J. Differential Equations 246 (2009) 3136–3213

Contents lists available at ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

# Construction of invariant whiskered tori by a parameterization method. Part I: Maps and flows in finite dimensions

Ernest Fontich [a], Rafael de la Llave [b,*], Yannick Sire [c,d]

[a] Departament de Matemàtica Aplicada i Anàlisi, Universitat de Barcelona, Gran Via, 585, 08007 Barcelona, Spain
[b] Department of Mathematics, University of Texas at Austin, 1 University station, Austin, USA
[c] Université Aix Marseille 3 Paul Cézanne, Avenue Escadrille Normandie-Niemen, Laboratoire LATP, UMR, 6632 Marseille, France
[d] Laboratoire Poncelet UMI 2615, 11 Bolshoy Vlasyevskiy Pereulok, 119002 Moscow, Russia

A R T I C L E I N F O

*Article history:*
Received 23 January 2008
Revised 11 January 2009
Available online 28 February 2009

*MSC:*
37J40

*Keywords:*
Whiskered tori
Hamiltonian systems
Small divisors
KAM theory

A B S T R A C T

We present theorems which provide the existence of invariant whiskered tori in finite-dimensional exact symplectic maps and flows. The method is based on the study of a functional equation expressing that there is an invariant torus.
We show that, given an approximate solution of the invariance equation which satisfies some non-degeneracy conditions, there is a true solution nearby. We call this an *a posteriori* approach.
The proof of the main theorems is based on an iterative method to solve the functional equation.
The theorems do not assume that the system is close to integrable nor that it is written in action-angle variables (hence we can deal in a unified way with primary and secondary tori). It also does not assume that the hyperbolic bundles are trivial and much less that the hyperbolic motion can be reduced to constant linear map.
The *a posteriori* formulation allows us to justify approximate solutions produced by many non-rigorous methods (e.g. formal series expansions, numerical methods). The iterative method is not based on transformation theory, but rather on successive corrections. This makes it possible to adapt the method almost verbatim to several infinite-dimensional situations, which we will discuss in a forthcoming paper. We also note that the method leads to fast and efficient algorithms. We plan to develop these improvements in forthcoming papers.


* Corresponding author.
*E-mail addresses:* fontich@mat.ub.es (E. Fontich), llave@math.utexas.edu (R. de la Llave), sire@cmi.univ-mrs.fr (Y. Sire).

0022-0396/$ – see front matter 
doi:10.1016/j.jde.2009.01.037

## Contents

## 1. Introduction

The goal of this paper is to prove some results on persistence of invariant tori for symplectic and exact symplectic maps and flows.

We will assume that the motion on the torus is a Diophantine rotation and that the remaining directions are as hyperbolic as allowed by the symplectic structure (if the remaining directions are not void such tori are commonly called *whiskered tori*).

More precisely, as it is well known, the preservation of the symplectic structure, together with the fact that the motion on the torus is a rotation, implies that the symplectic conjugate direction to the tangent of the torus is not hyperbolic. We will assume that the remaining directions in the tangent bundle of the phase space at the torus are spanned by a basis of vectors which contract exponentially in the future or in the past.

To make the previous statements more precise, we discuss first the case of maps. As we will show, results for flows can be readily deduced from the ones for maps. Given an exact symplectic map $F$ from an exact symplectic manifold $(\mathcal{M}, \Omega = d\alpha)$ into itself (for the purposes of this preliminary exposition, we will take $\mathcal{M}$ to be an Euclidean manifold, even if we will indicate how to eliminate this restriction later), and a frequency vector $\omega \in \mathbb{R}^l$, we seek an embedding $K : \mathbb{T}^l \to \mathcal{M}$ satisfying

$$(F \circ K)(\theta) = K(\theta + \omega), \quad \theta \in \mathbb{T}^l = \mathbb{R}^l/\mathbb{Z}^l. \tag{1}$$

Eq. (1) implies that the range of $K$ is invariant under $F$. If $K$ is an embedding, we obtain that $K(\mathbb{T}^l)$ is a torus contained in $\mathcal{M}$, invariant by $F$ and that the dynamics on it is, up to a change of coordinates, just a rotation of rotation vector $\omega$.

The main result of this paper will show that if we can find a function $K$ which satisfies some non-degeneracy assumptions and which satisfies (1) up to a sufficiently small error, then there is a true solution nearby.

Differentiating the functional equation (1) with respect to $\theta$ one gets

$$DF\big(K(\theta)\big)DK(\theta) = DK(\theta + \omega).$$

Geometrically, this shows that the tangent vector-field $DK(\theta)$ is invariant and does not grow or contract under iteration of the action by the map.

As we will see in more detail in Section 4.2.1, if the map preserves the symplectic form $\Omega$ and $K$ is a solution of the invariance equation, there exists an analytic matrix valued function $A(\theta)$, such that

$$DF\big(K(\theta)\big)\big[J(K)^{-1}DKN\big](\theta) = DK(\theta + \omega)A(\theta) + \big[J(K)^{-1}DKN\big](\theta + \omega), \tag{2}$$

where $J$ is the matrix representation of the symplectic form and

$$N(\theta) = \big[DK(\theta)^\top DK(\theta)\big]^{-1}.$$

As a consequence, $[J(K)^{-1}DKN](\theta)$ cannot grow more than polynomially. Hence we obtain that the center subspace of $T_{K(\theta)}\mathcal{M}$ is at least a $2l$-dimensional space spanned by $DK(\theta)$ and $[J(K)^{-1}DKN](\theta)$ (we will show that range $DK(\theta) \cap$ range $[J(K)^{-1}DKN](\theta) = \{0\}$ because the image of the torus is a isotropic manifold).

For approximately invariant systems, the previous identities are just approximate and this implies that the center direction is at least $2l$-dimensional. We will assume that indeed the dimension of the center subspace is exactly $2l$. That is, we will assume that the tori we consider are as hyperbolic as allowed by the fact that the motion on them are rotations and that the system preserves the symplectic structure.

The main non-degeneracy assumptions on the approximate solution are (a) that the other directions in $T_{K(\theta)}\mathcal{M}$ are hyperbolic. That is, they are spanned by vectors which contract exponentially fast in the future or in the past. (b) That there is some twist condition, that is, that the matrix $A$ in (2) is invertible.

We will use a KAM iterative method to show that, if we are given a function $K$ which solves (1) up to an error which is sufficiently small with respect to the properties of the non-degeneracy conditions (a), (b) above, then there is a true solution close to this approximate solution.

These results based on validating an approximate solution—which we call *a posteriori*—imply the usual persistence results (one can take as approximate solution of the modified system the exact solution for the original one). Nevertheless, the *a posteriori* results can be used for other purposes. For example *a posteriori* results can be used to validate solutions obtained through any method such as numerical approximations or asymptotic methods. The validation of Lindstedt series leads to estimates on their domain of analyticity. The paper [47] for instance considers Lindstedt series of whiskered tori.

*A posteriori* results also lead automatically to Lipschitz dependence on parameters and, with a bit more of work, to differentiable dependence on parameters. The *a posteriori* approach to KAM theorem was emphasized in [50,51,73,74]. There, it was pointed out that this *a posteriori* approach automatically allows to deduce results for finitely differentiable systems. We refer the reader to [13] for a comparison of different KAM methods.

In the present paper we deal with finite-dimensional maps and flows. In the forthcoming second part of it we consider coupled map lattices [22]. The case of partial differential equations, which can be treated in a similar way but involves technical difficulties, is postponed to a forthcoming paper [16].

Results on whiskered tori similar to the finite-dimensional ones of this paper have been considered several times in the literature. The first ones are [27,74].

The approach in [74]—which also takes the *a posteriori* format—is based on [73] which consists of finding a change of variables which reduces the system to a normal form which obviously possesses an invariant torus. This change of variables is accomplished by applying a sequence of canonical transformations. The method of proof introduced here is not based on successive transformations but rather on successive corrections introduced additively. This makes the estimates easier to establish and it leads to efficient numerical implementations. In order to be able to solve the equations, we take advantage of some cancellations due to the preservation of the symplectic structure that were also pointed out in [13,14,42].

The method of [74] proves the result for periodic Hamiltonian flows. The result for diffeomorphisms is proved in [74] by interpolating diffeomorphisms by periodic flows and then applying the results for periodic flows. The proof we present here proceeds along the opposite route. We prove first the result for diffeomorphisms and, then, deduce the result for flows taking time-one maps. Giving a direct proof of the result for persistence of whiskered tori for maps has been suggested as somewhat desirable in J. Moser's Mathematical review for [74]. We also provide such a direct proof. Of course, if one uses normal forms—as in [74]—it is natural to consider flows since the normal forms require only the study of the Hamiltonian function, which transforms very well. In the method presented here, the geometric cancellations are much more transparent in the case of diffeomorphisms.

Among other results for finite-dimensional systems, we call attention to [71], which uses a method similar to that of [1]. The paper [71] has the advantage that it is a first order method (i.e. that each step of the Newton iteration requires to solve only one small divisors equation). As a consequence, the size of the gaps among tori in near integrable systems, the loss of regularity as a function of the Diophantine exponent and the required minimum regularity are smaller than these of the second order methods. A comparison between first and second order methods to prove KAM results can be found in [13]. The paper [66] (see also the sketch in [48]) uses a reduction to a normally hyperbolic manifold and then applies the standard KAM theorem for Lagrangian tori. Of course, since normally hyperbolic manifolds are in general only $C^r$, the above method cannot produce $C^\infty$ or analytic tori. On the other hand, we note that the method of [48,66] leads to very good regularity conclusions for finite differentiable systems and also to good estimates on the measure occupied by the tori. We also call attention to [19,21,37,45,46,64,65,75] which consider also tori with hyperbolic and elliptic directions and relax the twist conditions and the differentiability requirements. The paper [26] considers analytic perturbations which depend only on the angles of reducible tori satisfying a twist condition and uses a direct resummation method.

As compared with previous finite-dimensional results, the method presented here has the advantage that one does not need to assume that the hyperbolic bundles are trivial (and much less that the motion in the hyperbolic directions is reducible to a constant linear map). Tori with non-trivial invariant bundles appear naturally in one parameter families after crossing a resonance, see [35].

Also, we do not need to assume that the system is given in action-angle coordinates, something which is convenient if we are working in situations when the action-angle coordinates are singular. For instance, in the study of diffusion one is lead naturally to the study of whiskered tori near resonances (see [7,8]). In this case, the action-angle variables are singular and avoiding its use leads to better estimates.

For symplectic ODEs we will also prove a translated torus theorem. From this general version we will deduce the results for exact symplectic ODEs using a vanishing lemma. We note that the approach of proving a translated torus theorem was introduced in [61] in the one degree of freedom case.

The method presented here lends itself to a very efficient numerical implementation (see [36]). The only functions to be considered are functions with a number of variables equal to the dimension of the torus itself (independently of the number of variables of the ambient space). Of course, when studying infinite-dimensional systems—PDEs or coupled map lattices or chains of oscillators—studying functions with the number of variables of the phase space is prohibitive. When implementing our method, if we discretize the tori by $N$ Fourier coefficients, the algorithm presented here only requires storage of order of $N$ and a Newton step takes only order of $N\log(N)$ operations using the fast Fourier transform. This seems to be significantly faster than other algorithms. Actual implementations are now being pursued and will be the subject of a forthcoming paper (see [36]). We refer the reader to [32,33,35] for analysis and implementation of related algorithms.

## 2. Definitions and notations

Before presenting the basic ideas and the results of our method, we introduce some notations and definitions which are useful for our purposes. All definitions are rather standard and we collect them here mainly to set the notation.

### 2.1. Diophantine vectors

In the study of invariant tori one needs an arithmetic condition over the frequency vector. In the case of maps the notion of Diophantine vector is the following.

**Definition 2.1.** Given $\kappa > 0$ and $\nu \geqslant l$, we define $D(\kappa,\nu)$ as the set of frequency vectors $\omega \in \mathbb{R}^l$ satisfying the Diophantine condition:

$$|\omega \cdot k - n|^{-1} \leqslant \kappa |k|^{\nu}, \quad \text{for all } k \in \mathbb{Z}^l \setminus \{0\} \text{ and } n \in \mathbb{Z},$$

where $\cdot$ means scalar product, $|k| = |k_1| + \cdots + |k_l|$ and $k_i$ are the coordinates of $k$. We will say that $\omega \in D(\kappa,\nu)$ is Diophantine.

For vector-fields the corresponding notion is the following.

**Definition 2.2.** Given $\kappa > 0$ and $\nu \geqslant l-1$, we define $D_h(\kappa,\nu)$ as the set of frequency vectors $\omega \in \mathbb{R}^l$ satisfying the Diophantine condition:

$$|\omega \cdot k|^{-1} \leqslant \kappa |k|^{\nu}, \quad \text{for all } k \in \mathbb{Z}^l \setminus \{0\}$$

with the same notation as in Definition 2.1.

The two conditions are closely related since $\omega \in D_h(\kappa,\nu)$ with $\omega_1 \neq 0$ if and only if $(\omega_2/\omega_1, \ldots, \omega_l/\omega_1) \in D(\kappa',\nu)$ for some $\kappa'$. The geometric and measure properties of the sets of Diophantine vectors have been extensively studied. These results translate immediately into statements about the abundance of KAM tori.

**Definition 2.3.** Let $\mathbb{T}^l = \mathbb{R}^l/\mathbb{Z}^l$ and $f \in L^1(\mathbb{T}^l)$. We denote $\operatorname{avg}(f)$ its average on the $l$-dimensional torus, i.e.

$$\operatorname{avg}(f) = \int_{\mathbb{T}^l} f(\theta)\, d\theta.$$

**Definition 2.4.** Given $\omega \in \mathbb{R}^l$ we introduce the rotation over $\mathbb{T}^l$ of rotation vector $\omega$:

$$T_\omega(\theta) = \theta + \omega.$$

*2.2. Functional spaces, functions and operators*

We will denote $D_\rho$ the complex extension of the torus of width $\rho$, i.e.

$$D_\rho = \{ z \in \mathbb{C}^l/\mathbb{Z}^l \mid |\operatorname{Im} z_i| \leqslant \rho,\ i = 1, \ldots, l \}. \tag{3}$$

We denote by $|\cdot|$ the supremum norm on $\mathbb{R}^N$ or $\mathbb{C}^N$. The sup norm makes several estimates independent of the dimension of the manifold, which are useful when considering infinite-dimensional problems. However on $\mathbb{Z}^l$ we will use the norm $|k| = |k_1| + \cdots + |k_l|$. Furthermore, for finite differentiability purposes, we consider the following norms: given $g$ analytic, with bounded derivatives in a complex domain $\mathcal{B}$, and $m \in \mathbb{N}$ we introduce the following $C^m$-norm for $g$

$$|g|_{C^m(\mathcal{B})} = \sup_{0 \leqslant |k| \leqslant m} \sup_{z \in \mathcal{B}} \left| D^k g(z) \right|.$$

Let $\mathcal{A}_\rho$ be the set of continuous functions on $D_\rho$, analytic in the interior of $D_\rho$ with values on a manifold $\mathcal{M}$, which is assumed to be Euclidean. We endow the space $\mathcal{A}_\rho$ with the usual supremum norm

$$\|u\|_\rho = \sup_{z \in D_\rho} \left| u(z) \right|.$$

We have that $(\mathcal{A}_\rho, \|\cdot\|_\rho)$ is a Banach space. In particular, $\|u\|_0 = \|u\|_{L^\infty(\mathbb{T}^l)}$.

We also recall the following convexity property (see [59, Lemma 12.8]).

**Proposition 2.5.** *Let $0 \leqslant \rho_1 \leqslant \rho_2$ and assume that $f \in \mathcal{A}_{\rho_2}$. Then, for every $\theta \in [0,1]$ we have*

$$\|f\|_{\theta\rho_1 + (1-\theta)\rho_2} \leqslant \|f\|_{\rho_1}^{\theta} \|f\|_{\rho_2}^{1-\theta}. \tag{4}$$

*In particular, taking $\rho_1 = 0$ and $\theta = 1/2$,*

$$\|f\|_{\rho_2/2} \leqslant \|f\|_{L^\infty(\mathbb{T}^l)}^{1/2} \|f\|_{\rho_2}^{1/2}. \tag{5}$$

We will also consider spaces of continuous functions on $D_\rho$ analytic in its interior and taking values on finite-dimensional vector spaces, for instance in spaces of matrices. When endowed with the supremum norm, these function spaces are also Banach spaces.

In particular, we will also need some norms of linear maps on the tangent space $T_{K(\theta)}\mathcal{M}$ with $K(\theta) \in \mathcal{M}$, where $K$ is an embedding. More concretely, let $A(\theta)$ be a continuous linear operator from $T_{K(\theta)}\mathcal{M}$ into itself depending on the variable $\theta \in D_\rho$. Then we define $\|A\|_\rho$ by

$$\|A\|_\rho = \sup_{\theta \in D_\rho} \sup_{v \in T_{K(\theta)}\mathcal{M},\, |v|=1} \big\|A(\theta)v\big\|_\rho.$$

## 3. Setting of the problem and results

### 3.1. Geometric setup

We will consider the Euclidean manifolds $\mathcal{M} = \mathbb{R}^{2d}$ and $\mathcal{M} = \mathbb{R}^{2d-2l} \times \mathbb{R}^l \times \mathbb{T}^l$. In the second case we can consider the universal covering $\mathbb{R}^{2d}$ of $\mathcal{M}$ and lift the maps defined on $\mathcal{M}$ to maps $\bar{F}$, defined on $\mathbb{R}^{2d}$, such that $\pi \bar{F} = \bar{F}\pi$, where $\pi : \mathbb{R}^{2d} \to \mathcal{M}$ is the canonical projection. Even if we pass to the covering we will use the symbol $\mathcal{M}$ to refer to the manifold. These manifolds obviously admit complex extensions by considering $\mathbb{R} \subset \mathbb{C}$ and $\mathbb{T} \equiv \mathbb{R}/\mathbb{Z} \subset \mathbb{C}/\mathbb{Z}$. As we will see, these different possibilities are convenient when we consider tori whose embeddings are topologically different. For example, tori which are contractible to tori with different dimensions. We will use the same symbol $\mathcal{M}$ for the complex extension of the manifold or its covering.

For convenience of notation, we will endow these manifolds with the standard Riemannian metric, even if this may not be natural for the problem at hand. For us, the metric will only play the role to measure sizes and therefore any equivalent metric will give a similar result. The standard metric will have the advantage that it will allow us to use matrix notation for adjoints. In matrix notation, thinking of vectors as column vectors, we can write $a^\top b = \langle a, b\rangle$. On the other hand, we note that the length of vectors will always be the supremum norm and the norm of matrices will be the operator norm associated to the supremum norm on vectors. Of course, for finite-dimensional problems the supremum norm is equivalent to the Euclidean norm.

We will assume that the Euclidean manifold $\mathcal{M}$ has an analytic exact symplectic form $\Omega$ with primitive $\alpha$, i.e. $\Omega = d\alpha$. For each $z \in \mathcal{M}$, let $J(z) : T_z\mathcal{M} \to T_z\mathcal{M}$ be the isomorphism such that

$$\Omega(\xi, \eta) = \big\langle \xi, J(z)\eta \big\rangle,$$

where $\langle , \rangle$ is the Euclidean product on $T_z\mathcal{M}$.

We will not assume that $J(z)$ has the standard form. We do not assume either that $J$ induces an almost-complex structure on $T\mathcal{M}$. This generality is useful in some applications (celestial mechanics, numerics, ...) when we use some system of coordinates—e.g. polar coordinates—which lead to non-standard symplectic matrices.

**Remark 3.1.** As we will see in the proof, we are not using much the Euclidean structure of the manifolds. In Section 7.6, we will present the modifications needed to work on other manifolds.

More precisely, we will show that it is possible to work out the proof in a neighborhood $U$ of the zero section of a bundle $E^c \oplus E^s \oplus E^u$. The bundle $E^c$ will be shown to be trivial (as a consequence of the fact that the motion on the torus is a rotation, the preservation of the symplectic structure and the fact that the dimension of the center space is $2l$, see Section 4.2.2), but the other bundles—which correspond to the hyperbolic directions—need not be trivial.

We note that Eq. (1) is geometrically natural since it can be formulated in any manifold.

In the following write up the Euclidean structure enters in two ways: one is a purely notational one and can be eliminated at the price of a typographical nightmare. When we only have approximate solutions, we will denote the error just as $F \circ K(\theta) - K(\theta + \omega)$ rather than $\exp^{-1}_{K(\theta+\omega)}(F \circ K(\theta))$. We will also compare vectors in $T_{F\circ K(\theta)}\mathcal{M}$ with vectors in $T_{K(\theta+\omega)}\mathcal{M}$. This can be done by introducing *connectors* as in [39], so that what we denote $DF \circ K(\theta + \omega) DF \circ K(\theta)$ is really $S^{F\circ K(\theta+\omega)}_{K(\theta+2\omega)} DF \circ K(\theta + \omega) S^{F\circ K(\theta)}_{K(\theta+\omega)} DF \circ K(\theta)$. See Definition 3.9 and the discussion thereinafter (particularly Eq. (17)).

A second, and more serious way in that the Euclidean space enters is that, to implement the iterative step in KAM theory, we will use Fourier series. This certainly requires that the functions take values in a vector space. Fortunately, this happens only in the center directions. In the hyperbolic directions there are geometrically natural ways to solve the iterative equation. This is why we are requiring that the center bundle is trivial, but we do not need the triviality of the hyperbolic bundles.

Of course, the fact that we work in a set $U$ as above is no loss of generality because, if there is a whiskered torus, by the tubular neighborhood theorem, we can identify a neighborhood of the torus with a neighborhood of the zero section of the normal bundle.

### 3.2. *Setting of the problem and results for maps*

The main purpose of the theory we are going to develop is to construct invariant tori for exact symplectic maps. We recall the following.

**Definition 3.2.** Let $(\mathcal{M}, \Omega = d\alpha)$ be an exact symplectic manifold. A map $F$ from $\mathcal{M}$ into itself is exact symplectic if there exists a smooth function $W$ on $\mathcal{M}$ such that

$$F^*\alpha = \alpha + dW.$$

In particular, every exact symplectic map is symplectic, i.e. $F^*\Omega = \Omega$.

Heuristically, our problem is the following: let $F$ be an exact symplectic map and $\omega \in D(\kappa, \nu)$. We want to construct an invariant torus for $F$ such that the dynamics of $F$ on it is conjugated to $T_\omega$. To this end, we search for an embedding $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ in $\mathcal{A}_\rho$ such that for all $\theta \in D_\rho$, $K$ satisfies the functional equation

$$F\big(K(\theta)\big) = K\big(T_\omega(\theta)\big). \tag{6}$$

Notice that if (6) is satisfied, the image under $F$ of a point in the range of $K$ will be also in the same range. Hence, since $K$ is an embedding, the range of $K$ will be an invariant torus.

The assumptions of our results will be that we are given a mapping $K$ that satisfies (6) up to a very small error and which satisfies some non-degeneracy and hyperbolicity assumptions. We will prove that then, there is a true solution of (6) close to $K$. We will also prove that the solution of (6) is unique up to composition on the right with translations.

The exactness of the map $F$ is important for the existence of a solution to (6). It is easy to construct examples of symplectic non-exact symplectic maps without invariant tori. For instance, consider $\mathcal{M} = \mathbb{T} \times \mathbb{R}$ with the standard symplectic structure. The translation in the $\mathbb{R}$-direction is a symplectic non-exact symplectic map without any invariant torus.

To construct the desired invariant torus, we consider a parameter $\lambda \in \mathbb{R}^l$ and introduce a translation term in Eq. (6) depending on $\theta$.

We then consider the following functional equation, where $G$ is a suitably chosen function of $\theta$ taking values in $2d \times l$ matrices and whose unknowns are both $K$ and $\lambda$

$$F\big(K(\theta)\big) + G(\theta)\lambda = K\big(T_\omega(\theta)\big). \tag{7}$$

The introduction of this parameter $\lambda$ will allow us to sidestep several technical complications and then we will show that, since $F$ is exact symplectic, the geometry implies that $\lambda = 0$. The fact that the dimension of the parameter $\lambda$ is $l$ is important for our purpose. We also mention that it is possible to use the parameter $\lambda$ to weaken non-degeneracy conditions by taking $\lambda \in \mathbb{R}^{2l}$ instead of $\mathbb{R}^l$. In such a case, $G$ is a $2d \times 2l$-matrix.

**Remark 3.3.** The introduction of the parameter $\lambda$ is also motivated by numerical calculations (see [36]). It leads to more stable computations. More importantly, it is useful in the numerical computation of *secondary* tori (i.e. tori generated by resonances, which have some contractible directions).

We go through a KAM technique to prove the existence of such a pair $(\lambda, K)$. To this end, we introduce the operator $\mathcal{F}_\omega$

$$\mathcal{F}_\omega(\lambda, K) = F \circ K + G\lambda - K \circ T_\omega, \tag{8}$$

where

$$G = \left[J(K_0)^{-1} DK_0\right] \circ T_\omega \tag{9}$$

is a function defined on $\mathbb{T}^l$ and where $K_0$ stands for an approximate whiskered torus. We will write $G$ instead of its explicit form in many of the following results. As we will see later, the important property of $G$ is that translations along the direction of $G$ can change the cohomology of the push-forward in the center directions.

The method is based on a careful study of the linearization (around a given pair $(\lambda, K)$) of the operator $\mathcal{F}_\omega$. We will show that this linear operator is approximately invertible in a suitable sense.

For that, we have to introduce several non-degeneracy conditions.

**Definition 3.4.** Given $\lambda \in \mathbb{R}^l$ and an embedding $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ we say that the pair $(\lambda, K)$ is non-degenerate for the functional equation (7) (and we denote $(\lambda, K) \in ND(\rho)$) if it satisfies the following conditions:

- *Spectral condition:* the tangent space $T_{K(\theta)}\mathcal{M}$ has an invariant splitting for all $\theta \in \mathbb{T}^l$,

$$T_{K(\theta)}\mathcal{M} = \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^c_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)}, \tag{10}$$

  where $\mathcal{E}^s_{K(\theta)}$, $\mathcal{E}^c_{K(\theta)}$ and $\mathcal{E}^u_{K(\theta)}$ are the stable, center and unstable invariant spaces respectively, i.e.

$$DF\big(K(\theta)\big)\mathcal{E}^{s,c,u}_{K(\theta)} = \mathcal{E}^{s,c,u}_{K(\theta+\omega)}.$$

  This splitting is analytic in $\theta$. To this splitting we associate the projections $\Pi^s_{K(\theta)}$, $\Pi^c_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ respectively, which are analytic with respect to $\theta$.
  Moreover, the splitting (10) is characterized by asymptotic growth conditions (co-cycles over $T_\omega$): there exist $0 < \mu_1, \mu_2 < 1$, $\mu_3 > 1$ such that $\mu_1\mu_3 < 1$, $\mu_2\mu_3 < 1$ and $C_h > 0$ such that for all $n \geqslant 1$ and $\theta \in D_\rho$

$$\left|(DF) \circ K \circ T^{n-1}_\omega(\theta) \times \cdots \times (DF) \circ K(\theta)v\right| \leqslant C_h \mu_1^n |v| \quad \Leftrightarrow \quad v \in \mathcal{E}^s_{K(\theta)}, \tag{11}$$

$$\left|(DF)^{-1} \circ K \circ T^{-(n-1)}_\omega(\theta) \times \cdots \times (DF)^{-1} \circ K(\theta)v\right| \leqslant C_h \mu_2^n |v| \quad \Leftrightarrow \quad v \in \mathcal{E}^u_{K(\theta)} \tag{12}$$

  and

$$\left|(DF) \circ K \circ T^{n-1}_\omega(\theta) \times \cdots \times (DF) \circ K(\theta)v\right| \leqslant C_h \mu_3^n |v|,$$
$$\left|(DF)^{-1} \circ K \circ T^{-(n-1)}_\omega(\theta) \times \cdots \times (DF)^{-1} \circ K(\theta)v\right| \leqslant C_h \mu_3^n |v| \quad \Leftrightarrow \quad v \in \mathcal{E}^c_{K(\theta)}. \tag{13}$$

- Furthermore, we assume that the dimension of the center subspace is $2l$.
  That is, the torus is as hyperbolic as allowed by the symplectic structure and there are no elliptic directions in the normal direction.

- *Twist condition:* We introduce the notation

$$N(\theta) = \left[DK(\theta)^\top DK(\theta)\right]^{-1},$$
$$P(\theta) = DK(\theta)N(\theta). \tag{14}$$

Assume that the averages on $\mathbb{T}^l$ of the matrices

$$Q(\theta) = DK(\theta+\omega)^\top J\big(K(\theta+\omega)\big)G(\theta) \tag{15}$$

and

$$A(\theta) = P(\theta+\omega)^\top\big[\big[DF(K)J(K)^{-1}\ P\big](\theta) - \big[J(K)^{-1}P\big](\theta+\omega)\big] \tag{16}$$

are non-singular.

**Remark 3.5.** With a view to applications, we note that in Proposition 5.2, we will show that we can deduce the existence of an invariant splitting from the existence of an approximately invariant one which satisfies the hyperbolicity conditions (11)–(13). Consequently, Definition 3.4 can be verified with finite precision calculation on a given numerical approximation. We anticipate that the basic idea is that, if we can verify that for some operator $B$ we have $\|B^N\| \leqslant \mu^N < 1$ for some $N > 0$, it follows that $\|B^n\| \leqslant C\mu^n$ for all $n > 0$. This gives a way to obtain all inequalities from finite computations.

**Remark 3.6.** Note that since $K$ is an embedding—hence $DK(\theta)$ is one-to-one for all $\theta$—and $d \geqslant l$ we have that $DK(\theta)^\top DK(\theta)$ is invertible for all $\theta$.

**Remark 3.7.** If we take $G(\theta) = J(K_0(\theta+\omega))^{-1}DK_0(\theta+\omega)$, $Q$ becomes $DK(\theta+\omega)^\top J(K_0(\theta+\omega))J(K(\theta+\omega))^{-1}DK_0(\theta+\omega) \approx N(\theta+\omega)^{-1}$ and hence one of the twist conditions becomes automatic because, under the smallness assumptions, $\text{avg}(Q) := \int_{\mathbb{T}^l} Q(\theta)\,d\theta \approx \text{avg}(N^{-1})$ and $\text{avg}(N^{-1})$ is invertible. Indeed, assume that $v \in \text{Ker}(\text{avg}(N^{-1}))$. Then $v^\top \text{avg}(N^{-1})v = 0$. This last expression is approximately

$$0 = \int_{\mathbb{T}^l} v^\top \big(DK^\top DK\big)(\theta)v\,d\theta = \int_{\mathbb{T}^l} \big|DK(\theta)v\big|^2\,d\theta$$

which implies $DK(\theta)v = 0$ for all $\theta \in \mathbb{T}^l$. Since $DK(\theta)$ is one-to-one for all $\theta$ we obtain $v = 0$. Hence the condition on the invertibility of $\text{avg}(Q)$ is just a quantitative statement of the fact that $K$ is indeed an embedding. The condition on $A$ is a twist condition.

**Remark 3.8.** Note that if the torus $K$ was exactly invariant (i.e. $F \circ K = K \circ T_\omega$) then

$$DF \circ K \circ T_\omega^{n-1}(\theta) \times \cdots \times DF \circ K(\theta) = DF^n \circ K(\theta),$$

so that conditions (11)–(13) are the usual growth conditions in the theory of normally hyperbolic manifolds (see [24,40,56]). Of course, for our applications, we only assume that the tori are approximately invariant.

When the manifolds are Euclidean, the conditions (11)–(13) make perfect sense. Nevertheless, if the phase space is a general manifold $\mathcal{M}$, we have $DF(K(\theta)) : T_{K(\theta)}\mathcal{M} \to T_{F\circ K(\theta)}\mathcal{M}$. If $F \circ K(\theta) \neq K(\theta+\omega)$, then, we should write the conditions (11)–(13) using connectors (see [39]).

We recall that

**Definition 3.9.** A connector $S_x^y$ is an isomorphism from $T_y\mathcal{M}$ to $T_x\mathcal{M}$, defined when $d(x, y)$ is small enough, such that $S_x^x = \mathrm{Id}$ and $S_x^y S_y^z = S_x^z$, when both make sense.

A concrete way of implementing the connectors is to take parallel transport along the shortest geodesic joining $x, y$ (equivalently, the differential of the exponential map).

In the case that we formulate the result in a general manifold, (11) should be written

$$\left|(DF)\circ K\circ T_\omega^{n-1}(\theta)S_{K\circ T_\omega^{n-1}(\theta)}^{F\circ K\circ T_\omega^{n-2}(\theta)}\times\cdots\times S_{K(\theta+\omega)}^{F(K(\theta))}(DF)\circ K(\theta)v\right|\leqslant C_h\mu_1^n|v| \quad\Leftrightarrow\quad v\in\mathcal{E}_{K(\theta)}^s \tag{17}$$

and analogously the others.

**Remark 3.10.** The technical reason why we introduced the extra parameter $\lambda$ in (7) is the following: in the iteration of the KAM scheme, one has to prove that some equations are approximately solved up to a quadratic error. To this end, we have to show that some averages are quadratic in the error. To avoid these technicalities, we introduce this parameter $\lambda$ which allows us to cancel some terms in the equation so that we can reach the suitable approximate solution (see Propositions 4.18 and 4.19). Then we use the exact symplecticness of the map to keep the parameter $\lambda$ under control.

We can now state our main theorem, which provides the existence of a solution $K$ to the functional equation (6) with $F$ exact symplectic, provided we are given a sufficiently approximate one.

**Theorem 3.11.** *Let $\omega\in D(\kappa,\nu)$ for some $\kappa>0$, $\nu\geqslant l$. Assume that:*

(1) *$F:\mathcal{U}\subset\mathcal{M}\to\mathcal{M}$ is an exact symplectic map and $\mathcal{U}$ is an open connected set, which we will assume without loss of generality has a smooth boundary.*
(2) *$K_0\in ND(\rho_0)$ (the embedding $K_0$ is non-degenerate) in the sense that it satisfies the spectral condition in Definition 3.4 and the average on $\mathbb{T}^l$ of the matrices $Q_0(\theta)$ and $A_0(\theta)$ are non-singular, where $Q_0$ and $A_0$ are as $Q$ and $A$ in Definition 3.4 with $K=K_0$.*
(3) *The map $F$ is real analytic and it can be extended holomorphically to some complex neighborhood of the image under $K_0$ of $D_{\rho_0}$:*

$$B_r=\left\{z\in\mathbb{C}^{2d}\ \middle|\ \exists\theta\in\left\{|\operatorname{Im}\theta|<\rho_0\right\}\text{ s.t. }\left|z-K_0(\theta)\right|<r\right\},$$

*for some $r>0$ and such that $|F|_{C^2(B_r)}$ is finite.*

*Denote $E_0=F\circ K_0-K_0\circ T_\omega$ the initial error. Then there exists a constant $C>0$ depending on $l$, $\nu$, $|F|_{C^2(B_r)}$, $\|DK_0\|_{\rho_0}$, $\|N_0\|_{\rho_0}$, $\|A_0\|_{\rho_0}$, $|(\mathrm{avg}(A_0))^{-1}|$, $|(\mathrm{avg}(Q_0))^{-1}|$, $|J|_{C^1(B_r)}$ and the norms of the projections $\|\Pi_{K_0(\theta)}^{c,s,u}\|_{\rho_0}$ such that, if $E_0$ satisfies the estimates*

$$C\kappa^4\delta^{-4\nu}\|E_0\|_{\rho_0}<1 \tag{18}$$

*and*

$$C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0}<r,$$

*where $0<\delta\leqslant\min(1,\rho_0/12)$ is fixed, then there exists an embedding $K_\infty\in ND(\rho_\infty:=\rho_0-6\delta)$ such that*

$$F\circ K_\infty=K_\infty\circ T_\omega.$$

*Furthermore, we have the following estimate*

$$\|K_\infty-K_0\|_{\rho_\infty}\leqslant C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0}. \tag{19}$$

**Remark 3.12.** The previous theorem provides a construction of whiskered tori without assuming the existence of action-angle variables for the original system. Moreover, the method of proof does not involve the sequence of transformations by symplectomorphisms, which is often used to prove this kind of results, but hard to implement numerically.

**Remark 3.13.** It is important to remark that the non-degeneracy conditions we use in Theorem 3.11 depend only on the approximate solution under consideration. As one can see, Definition 3.4 only depends on averages of the approximately computed solutions. This latter fact is useful in the validation of numerical computations. Indeed, numerical computations provide an approximate solution and this is the only information that is available. The non-degeneracy conditions needed to apply Theorem 3.11 can be verified by straightforward computations on the numerical approximation.

This leads directly to the so-called *small twist theorems*. See Section 7.3 and in particular Proposition 7.1 and the subsequent comments for more details on the dependence of the constants on the non-degeneracy assumptions.

After introducing an additional term in the functional equation (6), namely

$$F \circ K + \big(J(K_0)^{-1} DK_0\big) \circ T_\omega \lambda = K \circ T_\omega$$

and performing a KAM iteration on $(K, \lambda)$, the final task consists of proving that $\lambda_\infty = 0$ using the geometry. This is done by using the exact symplecticness of $F$ and a suitable representation of the center subspace. Indeed, the center subspace in $T_{K(\theta)}\mathcal{M}$, which will be shown to be non-trivial, will be very close to the vector space spanned by $DK(\theta)$ and its symplectically conjugate $J(K(\theta))^{-1}DK(\theta)$.

### 3.3. Uniqueness

A natural question to ask is whether the embedding $K$ provided by Theorem 3.11 is unique. Notice that if $K$ is a solution of (6), for any $\sigma \in \mathbb{R}^l$, $K \circ T_\sigma$ is also a solution, hence one can only hope for uniqueness up to a composition with a translation on the right.

The following theorem provides a local uniqueness result. We will see in the next section that there is a simple general argument that shows that uniqueness results allow us to deduce results for flows from results for diffeomorphisms.

**Theorem 3.14.** *Let $F$ be exact symplectic and analytic in $B_r \subset \mathcal{M}$. Let $\omega \in D(\kappa, \nu)$ for some $\kappa > 0$, $\nu \geqslant l$. Assume $K_1, K_2 \in ND(\rho)$ with $\rho > 0$ are two solutions of Eq. (6) such that $K_1(D_\rho) \subset B_r$, $K_2(D_\rho) \subset B_r$. Then there exists a constant $C > 0$ depending on $l$, $\nu$, $|F|_{C^2(B_r)}$, $\|DK_1\|_\rho$, $\|N_1\|_\rho$, $|J|_{C^1(B_r)}$, $\|A_1\|_\rho$, $\|\Pi^{c,s,u}_{K(\theta)}\|_\rho$, $|(\mathrm{avg}(A_1))^{-1}|$ such that if for some $\tau \in \mathbb{R}^l$ the norm $\|K_1 \circ T_\tau - K_2\|_\rho$ satisfies*

$$C\kappa^2 \rho^{-2\nu} \|K_1 \circ T_\tau - K_2\|_\rho \leqslant 1 \tag{20}$$

*with $\delta = \rho/4$, there exists a phase $\tilde\tau \in \mathbb{R}^l$ such that $K_1 \circ T_{\tilde\tau} = K_2$ in $D_\rho$. Moreover $|\tilde\tau - \tau| \leqslant C\kappa^2\rho^{-2\nu}\|K_1 - K_2\|_\rho$.*

The proof of this theorem is postponed to Section 6.

### 3.4. Result for flows

As a by-product of the previous uniqueness theorem, we get a result on the existence of invariant whiskered tori for flows. This follows from a time-one map argument (see [18]). The argument we present here comes from [3,14].

**Theorem 3.15.** *Let $\omega \in D(\kappa, \nu)$ for some $\kappa > 0, \nu \geqslant l$. Let $(S_t)_{t\in\mathbb{R}}$ be the flow generated by a finite-dimensional analytic exact symplectic vector-field*

$$\frac{du}{dt} = f(u),$$

*where $u : I \subset \mathbb{R} \to \mathcal{M}$. Assume that there exist a time $t = 1$ and an embedding $K \in ND(\rho)$ for some $\rho > 0$ such that $S_1 \circ K(\theta) = K(\theta + \omega)$ for all $\theta \in \mathbb{T}^l$. Then for all time $t \in \mathbb{R}$, we have*

$$S_t \circ K(\theta) = K(\theta + \omega t).$$

**Proof.** If we have $S_1 \circ K(\theta) = K(\theta + \omega)$, then for all $t$ this yields

$$S_1 \circ S_t \circ K(\theta) = S_t \circ (S_1 \circ K)(\theta) = S_t \circ K(\theta + \omega).$$

By Theorem 3.14, if $\|S_t \circ K - K\|_\rho$ is sufficiently small, which is achieved if $t$ is sufficiently small, this implies that there exists a phase $\phi(t)$ such that $S_t \circ K(\theta) = K(\theta + \phi(t))$. From the flow property $S_{t+s} = S_t \circ S_s$ and the fact that $K$ is one-to-one, we have $\phi(t + s) = \phi(t) + \phi(s)$. We now prove that the function $\phi$ is continuous. The map $K$ from $\mathbb{T}^l$ into its image is one-to-one and continuous over a compact (for the topology of $\mathbb{T}^l$). Then its inverse is continuous. This leads to the continuity of the function $\phi$. Using this fact and the additivity condition we deduce, that for $t$ small enough, $\phi(t) = \beta t$ for some $\beta \in \mathbb{R}^l$. Then in this case we have

$$S_t \circ K(\theta) = K(\theta + \beta t). \tag{21}$$

Since both sides of (21) are analytic with respect to $t \in [0, 1]$ we obtain the result for all $t \in [0, 1]$. Putting $t = 1$ we get $\beta = \omega$. Expression (21) shows that the torus $K(\mathbb{T}^l)$ is invariant by the flow. Since the torus is compact, the flow on it is defined for all $t \in \mathbb{R}$ and hence (21) holds for all $t \in \mathbb{R}$. This ends the proof. □

In Section 9, we will give a more precise version of this result and a direct proof (i.e. a proof which does not pass through a reduction to a time-1 map). This is useful since the method of proof leads to numerical algorithms for differential equations. The direct proof can also be used as a model for results for some ill-posed partial differential equations (see [16]).

## 4. The linearized operator $D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)$

In this section, we describe the inductive step of the procedure. As most of the KAM proofs, it will be a modification of the classical Newton method.

Using the Taylor theorem, given an approximate solution, we write

$$\mathcal{F}_\omega(\lambda + \Lambda, K + \Delta) = \mathcal{F}_\omega(\lambda, K) + D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) + O\big(\big\|(\Lambda, \Delta)\big\|^2\big)$$

and, following the idea of Newton's method, we look for $(\Lambda, \Delta)$ such that $\mathcal{F}_\omega(\lambda + \Lambda, K + \Delta)$ is quadratically small so we are lead to consider the following equation

$$D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) = -E, \tag{22}$$

where $(\lambda, K)$ is a pair satisfying approximately Eq. (7) with an error $E(\theta) = \mathcal{F}_\omega(\lambda, K)(\theta)$ with $\theta \in \mathbb{T}^l$. Using the definition of the operator $\mathcal{F}_\omega$ in (8), we see that the derivative of the operator can be written more explicitly as

$$D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta)(\theta) = G(\theta)\Lambda + DF\big(K(\theta)\big)\Delta(\theta) - \Delta(\theta + \omega).$$

The study of the Newton equation (22) is mainly done in three steps:

- One projects Eq. (22) on the hyperbolic space and the center space, by using the invariant splitting (see Definition 3.4).
- One reduces the equation of the projection on the center subspace to two classical small divisors equations. Thanks to a suitable change of coordinates on the tangent space (which does not use action-angle variables) these equations are then solved approximately (i.e. up to quadratic error) by using the extra variable $\Lambda \in \mathbb{R}^l$.
- One solves (with “tame” estimates) the equations corresponding to the projections onto the stable and unstable invariant subspaces, by using the conditions on the co-cycles over $T_\omega$.

**Remark 4.1.** We note that the equation on the center subspace will not be solved exactly. We will just solve it up to quadratic errors. The reason is that the change of variables mentioned in the above discussion will be constructed taking advantage of approximate identities obtained by differentiating with respect to $\theta$ the equation for the initial error and applying geometric identities. The procedure of comparing the linearized Newton equation with the equations that appear taking derivatives is very common in KAM theory. It is certainly used systematically in [50,51,74]. See [73, Section 5] for some remarks on the relation of these identities with a group structure of conjugacy problems. We note that some of these remarks in the above references work also for some semi-conjugacy problems.

Of course, the above-mentioned strategy uses the non-degeneracy assumptions. In subsequent sections, we will show that these assumptions are changed only by a small amount, so that the procedure can be iterated.

The main goal of this section is to prove the following result.

**Lemma 4.2.** *Consider the linearized equation*

$$D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) = -E. \tag{23}$$

*Then there exists a constant $C$ that depends on $\nu$, $l$, $\|DK\|_\rho$, $\|N\|_\rho$, $\|\Pi^{s,c,u}_{K(\theta)}\|_\rho$, $\|G\|_\rho$, $|(\mathrm{avg}(A))^{-1}|$, $|(\mathrm{avg}(Q))^{-1}|$ and the hyperbolicity constants such that assuming that $\delta \in (0, \rho/2)$ satisfies*

$$C\kappa\delta^{-(\nu+1)}\big(\|E\|_\rho + \|G\|_\rho|\lambda|\big) < 1 \tag{24}$$

*we have*:

(1) *There exists an approximate solution $(\Lambda, \Delta)$ of* (23)*, in the following sense*: *there exits a function $\tilde{E}(\theta)$ such that $(\Lambda, \Delta)$ solves exactly*

$$D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) = -E + \tilde{E},$$

*with the following estimates*: *for all $\delta \in (0, \rho/2)$*

$$\|\Delta\|_{\rho-2\delta} \leqslant C\kappa^2\delta^{-2\nu}\|E\|_\rho, \tag{25}$$

$$\|D\Delta\|_{\rho-2\delta} \leqslant C\kappa^2\delta^{-2\nu-1}\|E\|_\rho, \tag{26}$$

$$|\Lambda| \leqslant C\|E\|_\rho, \tag{27}$$

$$\|\tilde{E}\|_{\rho-\delta} \leqslant C\kappa^2\delta^{-(2\nu+1)}\|E\|_\rho\big\|\mathcal{F}_\omega(\lambda, K)\big\|_\rho. \tag{28}$$

(2) *If $\Delta_1$ and $\Delta_2$ solve the linearized equation in the previous approximate sense, then there exists $\alpha \in \mathbb{R}^l$ such that for all $\delta \in (0, \rho)$*

$$\left\| \Delta_1 - \Delta_2 - DK(\theta)\alpha \right\|_{\rho-\delta} \leqslant C\kappa^2 \delta^{-(2\nu+1)} \|E\|_\rho \left\| \mathcal{F}_\omega(\lambda, K) \right\|_\rho. \tag{29}$$

**Remark 4.3.** The form of the previous inductive lemma corresponds very closely to Zehnder's implicit function theorem in [73]. Once Lemma 4.2 is proved, we then follow the strategy in [73]. The most crucial step is the verification of how the hypothesis of hyperbolicity are changed when the embedding changes in the iterative step.

More precise information on the dependence of the constants $C$ on the non-degeneracy conditions will be provided in Proposition 7.1. We anticipate that, roughly speaking, the constants $C$ can be bounded by universal powers of the non-degeneracy constants. We postpone the precise formulation since it will involve some notations that will be developed along the proof. This power dependence on the constants has some applications to the study of tori close to resonance and to small twist theorems.

We will need the following classical proposition (see [13,60–62]) which provides existence of a solution together with estimates for small divisors equations.

**Proposition 4.4.** *Let $\omega \in D(\kappa, \nu)$ and assume the mapping $h : \mathbb{T}^l \to \mathcal{M}$ is analytic on $D_\rho$ and has zero average. Then for any $0 < \sigma < \rho$ the difference equation*

$$v(\theta + \omega) - v(\theta) = h(\theta)$$

*has a unique zero average solution $v : \mathbb{T}^l \to \mathcal{M}$, real analytic on $D_{\rho-\sigma}$ for any $0 < \sigma < \rho$. Moreover, we have the estimate*

$$\|v\|_{\rho-\sigma} \leqslant C\kappa\sigma^{-\nu} \|h\|_\rho, \tag{30}$$

*where $C$ only depends on $\nu$ and the dimension of the torus $l$.*

**Remark 4.5.** It is important for our purposes to have estimates independent of the dimension of the manifold $\mathcal{M}$ since in a follow-up paper [22] we apply the procedure of this paper in an infinite-dimensional context.

The independence of the estimates on the number of dimensions comes from the fact that we consider the supremum norm and the equation is solved component-wise.

### *4.1. Geometric considerations*

#### *4.1.1. Isotropic character of the torus*

We start by recalling the definition of isotropy.

**Definition 4.6.** Let $(\mathcal{M}, \Omega)$ be a symplectic manifold. A submanifold $\mathcal{N}$ of $\mathcal{M}$ is isotropic if $\mathcal{N} \subset \mathcal{N}^\perp$, where $\mathcal{N}^\perp$ is the orthogonal space of $\mathcal{N}$ with respect to the 2-form $\Omega$.

We formulate in our framework the well-known fact that a torus supporting an irrational rotation is isotropic. The manifold $K(\mathbb{T}^l)$ is isotropic if the pull-back $K^*\Omega(\theta)$ vanishes for all $\theta \in \mathbb{T}^l$. In other words, noting

$$K^*\Omega(\theta)(\xi, \eta) = \langle \xi, L(\theta)\eta \rangle$$

for all $\xi, \eta \in \mathbb{R}^l$, the isotropic character is equivalent to

$$L(\theta) = DK(\theta)^\top J\big(K(\theta)\big)DK(\theta) = 0$$

for all $\theta \in \mathbb{T}^l$. We first deal with the case of an exact solution of (6) (see Lemma 4.7). The approximate case is the purpose of Lemma 4.8. We note that the fact that exactly invariant tori are isotropic manifolds remains true for all irrational rotations and is well known [74]. The fact that approximately invariant tori carrying an irrational rotation are approximately isotropic seems to require that the rotation is Diophantine, see [14]. For the sake of completeness, we present the simple proofs of both results.

**Lemma 4.7.** *Assume that $\mathcal{M}$ is exact symplectic, $K$ satisfies (6) and $\omega$ is rationally independent. Then $L(\theta)$ is identically zero.*

**Proof.** Since $F$ is symplectic we have

$$F^*\Omega = \Omega.$$

Consequently, this yields

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

Since $\omega$ is rationally independent, $T_\omega$ is ergodic and this implies that $K^*\Omega$ is constant and so is $L(\theta)$. Using that $\mathcal{M}$ is exact symplectic, we have that $K^*\Omega = dK^*\alpha$ and, the only constant form which is exact is zero.

Similarly, a computation shows that $L(\theta)$ has the form $DL_1(\theta)^\top - DL_1(\theta)$ for some matrix $L_1(\theta)$. Since the average on $\mathbb{T}^l$ of $DL_1(\theta)$ is zero, we get the result. $\square$

**Lemma 4.8.** *Assume that $\mathcal{M}$ is an exact symplectic manifold, $F : B_r \to \mathcal{M}$ is analytic and symplectic. Let $K$ be real analytic on the complex strip $D_\rho$ for some $\rho > 0$ and such that $K(D_\rho) \subset B_r$. Assume also that $\omega \in D(\kappa, \nu)$ and denote*

$$E = F \circ K + G\lambda - K \circ T_\omega.$$

*Then there exists a constant $C$ depending on $l$, $\nu$, $\|DK\|_\rho$, $|F|_{C^1(B_r)}$, $|J|_{C^1(B_r)}$ such that for all $\delta \in (0, \rho/2)$ we have*

$$\|L\|_{\rho-2\delta} \leqslant C\kappa\delta^{-(\nu+1)}\big(\|E\|_\rho + \|G\|_\rho|\lambda|\big). \tag{31}$$

**Proof.** We want to estimate the norm of the matrix $L$. Recalling $F^*\Omega = \Omega$, one gets

$$K^*\Omega - (K \circ T_\omega)^*\Omega = E^*\Omega - (G\lambda)^*\Omega.$$

Performing the same computations as in [14], this leads to the following equation

$$L - L \circ T_\omega = g,$$

where $g$ is a function on $\mathbb{T}^l$ such that (here we just use Cauchy estimates)

$$\|g\|_{\rho-\delta} \leqslant C\delta^{-1}\big(\|E\|_\rho + \|G\|_\rho|\lambda|\big).$$

We now make use of Proposition 4.4 to complete the proof. $\square$

Recall that we are assuming that $K$ is an embedding. Hence the range of $DK$ is $l$-dimensional.

*4.1.2. Vanishing lemma*

This section is devoted to an estimate which allows to control the extra parameter $\lambda$ through the iterative step. We consider the functional equation

$$F \circ K + G\lambda = K \circ T_\omega + E,$$

where $F$ is *exact* symplectic (see Definition 3.2) and $G = [J(K_0)^{-1} DK_0] \circ T_\omega$.

Recall that $\lambda \in \mathbb{R}^l$ and $K_0 \in ND(\rho_0)$. Note that the term $(J(K_0)^{-1} DK_0) \circ T_\omega$ is very close to $(J(K)^{-1} DK) \circ T_\omega$ and hence close to the center subspace associated to the torus $K(\mathbb{T}^l)$.

The following lemma provides the desired vanishing result.

**Lemma 4.9.** *Assume $F$ maps $\mathcal{M}$ into itself and $\omega \in D(\kappa, \nu)$. Let $K \in ND(\rho)$ be a solution of*

$$F \circ K + G\lambda = K \circ T_\omega + E, \tag{32}$$

*with $G = [J(K_0)^{-1} DK_0] \circ T_\omega$ and $\lambda$ is such that*

$$\begin{gathered} \|E\|_\rho + \|G\|_\rho |\lambda| \leqslant r, \\ \|K - K_0\|_\rho \leqslant r, \qquad \|DK - DK_0\|_\rho \leqslant r, \end{gathered} \tag{33}$$

*where $r > 0$ is sufficiently small (precise conditions will be given along the proof).*

*Assume furthermore that:*

(1) *$F$ is exact symplectic.*
(2) *$F$ extends analytically to a neighborhood of $K(\mathbb{T}^l)$.*

*Then, there exists a constant $C$ such that*

$$|\lambda| \leqslant C \|E\|_\rho.$$

**Proof.** We follow a method used in [42]. We refer the reader to Fig. 1 for an illustration of the method.

We denote by

$$\hat{\theta}_i = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_l) \in \mathbb{T}^{l-1} \tag{34}$$

and similarly $\hat{\omega}_i = (\omega_1, \dots, \omega_{i-1}, \omega_{i+1}, \dots, \omega_l) \in \mathbb{R}^{l-1}$.

We also denote $\sigma_{i,\hat{\theta}_i} : \mathbb{T} \to \mathbb{T}^l$ the path given by

$$\sigma_{i,\hat{\theta}_i}(\eta) = (\theta_1, \dots, \theta_{i-1}, \eta, \theta_{i+1}, \dots, \theta_l). \tag{35}$$

We will compute $\int_{\mathbb{T}^{l-1}} \int_{K \circ \sigma_{i,\hat{\theta}_i}} F^*\alpha$ in two different ways. On one hand, using the fact that $F$ is exact symplectic, we have

$$\int\limits_{K \circ \sigma_{i,\hat{\theta}_i + \hat{\omega}_i}} F^*\alpha = \int\limits_{K \circ \sigma_{i,\hat{\theta}_i + \hat{\omega}_i}} \alpha + dW = \int\limits_{K \circ \sigma_{i,\hat{\theta}_i + \hat{\omega}_i}} \alpha. \tag{36}$$

On the other hand, using (32)

$$\int\limits_{K \circ \sigma_{i,\hat{\theta}_i}} F^*\alpha = \int\limits_{F \circ K \circ \sigma_{i,\hat{\theta}_i}} \alpha = \int\limits_{(K \circ T_\omega - G\lambda) \circ \sigma_{i,\hat{\theta}_i}} \alpha + R_i, \tag{37}$$

where $|R_i| \leqslant C \|E\|_\rho$.

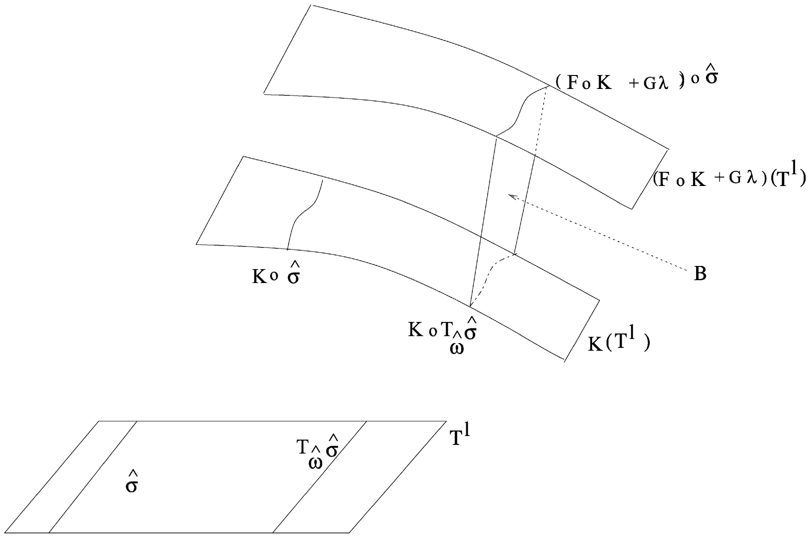


**Fig. 1.** Illustration of the vanishing lemma.

Since we want to compare the last integrals in (36) and (37) it is natural to introduce a two-cell whose boundary is the difference between the two paths $K \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}$ and $(K \circ T_\omega - G\lambda) \circ \sigma_{i,\hat{\theta}_i}$. We denote $B_{i,\hat{\theta}_i,\lambda}$ this two-cell, which we parametrize by $(\xi,\eta) \in (0,1) \times (0,1)$ as follows:

$$B_{i,\hat{\theta}_i,\lambda}(\xi,\eta) = K \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}(\eta) - G \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}(\eta)\lambda\xi.$$

By Stokes's theorem, since $d\alpha = \Omega$, we have

$$\int\limits_{(K\circ T_\omega - G\lambda)\circ\sigma_{i,\hat{\theta}_i}} \alpha = \int\limits_{K\circ\sigma_{i,\hat{\theta}_i+\hat{\omega}_i}} \alpha + \int\limits_{B_{i,\hat{\theta}_i,\lambda}} \Omega. \tag{38}$$

We have

$$\int\limits_{B_{i,\hat{\theta}_i,\lambda}} \Omega = \int_0^1\int_0^1 \Omega_{B_{i,\hat{\theta}_i,\lambda}(\xi,\eta)}\big(\partial_\xi B_{i,\hat{\theta}_i,\lambda}(\xi,\eta), \partial_\eta B_{i,\hat{\theta}_i,\lambda}(\xi,\eta)\big)\,d\xi\,d\eta.$$

Note that

$$\partial_\eta B_{i,\hat{\theta}_i,\lambda} = \partial_{\theta_i} K \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}(\eta) - \partial_{\theta_i} G \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}(\eta)\lambda\xi$$

and

$$\partial_\xi B_{i,\hat{\theta}_i,\lambda} = -G \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}(\eta)\lambda.$$

Using the previous expressions

$$\begin{aligned}&\Omega_{B_{i,\hat{\theta}_i,\lambda}(\xi,\eta)}\big(\partial_\xi B_{i,\hat{\theta}_i,\lambda}(\xi,\eta), \partial_\eta B_{i,\hat{\theta}_i,\lambda}(\xi,\eta)\big)\\ &\quad= -\lambda^\top G \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}(\eta)^\top J\big(B_{i,\hat{\theta}_i,\lambda}(\xi,\eta)\big)\big(\partial_{\theta_i} K \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}(\eta) - \partial_{\theta_i} G \circ \sigma_{i,\hat{\theta}_i+\hat{\omega}_i}(\eta)\lambda\xi\big).\end{aligned}$$

Using (33), we have

$$J\big(B_{i,\hat\theta_i,\lambda}(\xi,\eta)\big) = J\big(K\circ\sigma_{i,\hat\theta_i+\hat\omega_i}(\eta)\big) + O\big(|\lambda|\big)$$
$$= J\big(K_0\circ\sigma_{i,\hat\theta_i+\hat\omega_i}(\eta)\big) + O(r) + O\big(|\lambda|\big).$$

Therefore, we end up with (using the expression of $G(\theta) = J(K_0(\theta))^{-1}DK_0(\theta)$)

$$\int_{B_{i,\hat\theta_i,\lambda}} \Omega = -\int_0^1\int_0^1 \lambda^\top\big[DK_0^\top J(K_0)^{-\top}J(K_0)\partial_{\theta_i}K_0\big]\circ\sigma_{i,\hat\theta_i+\hat\omega_i}(\eta)\,d\xi\,d\eta$$
$$+ O\big(r|\lambda|\big) + O\big(|\lambda|^2\big) + O\big(\|E\|_\rho\big).$$

Joining these expressions for all values of $i$ and integrating over $\mathbb{T}^{l-1}$ we get

$$\int_{\mathbb{T}^{l-1}}\int_{B_{i,\hat\theta_i,\lambda}} \Omega = \lambda^\top\bigg[\int_{\mathbb{T}^l}\tilde Q\bigg] + O\big(r|\lambda|\big) + O\big(|\lambda|^2\big) + O\big(\|E\|_\rho\big), \tag{39}$$

where $\tilde Q = DK_0^\top DK_0$. Since $DK_0$ has rank $l$ then the matrix $\tilde Q$ has rank $l$. See Remark 3.6.

We now integrate with respect to $\hat\theta_i$ both (36) and (37). By a simple change of variables we have that the following integrals are equal

$$\int_{\mathbb{T}^{l-1}} d\hat\theta_i \int_{K\circ\sigma_{i,\hat\theta_i+\hat\omega_i}} F^*\alpha = \int_{\mathbb{T}^{l-1}} d\hat\theta_i \int_{K\circ\sigma_{i,\hat\theta_i}} F^*\alpha.$$

Therefore, from (38) we obtain

$$\int_{\mathbb{T}^{l-1}}\int_{B_{i,\hat\theta_i,\lambda}} \Omega = -\int_{\mathbb{T}^{l-1}} R_i.$$

Now, Eq. (39), the fact that $\tilde Q$ is invertible, the assumption

$$\|E\|_\rho + \|G\|_\rho|\lambda| \leqslant C,$$

and $r$ sufficiently small (this is the condition we imposed in (33)), lead to the desired result invoking the implicit function theorem. $\square$

**Remark 4.10.** The assumption in Lemma 4.9 that

$$\|E\|_\rho + \|G\|_\rho|\lambda| \leqslant C$$

will be an inductive assumption in the iteration of the KAM method that we will deal with later.

**Remark 4.11.** In the KAM iteration, we will generate a sequence $\{\lambda_n, K_n\}_{n\in\mathbb{N}}$ of approximations of the solution $(\lambda_\infty, K_\infty)$ of the equation

$$F\circ K + G\lambda = K\circ T_\omega.$$

As a corollary of Lemma 4.9, the sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ in the KAM iteration converges to 0 since $\|E_n\|_{\rho_n}$ converges to 0.

#### 4.1.3. Basis of $\mathcal{E}^c_{K(\theta)}$ when $K$ is an exact parameterization

To avoid the use of action-angle variables we are going to perform a change of variables, using the geometric structure on the tangent bundle.

For that we will first find a useful basis of the center space $\mathcal{E}^c_{K(\theta)}$ in the case that $K:\mathbb{T}^l\to\mathcal{M}$ is a solution of $F\circ K = K\circ T_\omega$.

Here we are assuming that the dimension of the center subspace is $2l$ and hence $\mathcal{E}^c_{K(\theta)}\sim\mathbb{R}^{2l}$. In [14], the authors studied the case when $d=l$, i.e. the dimension of the range of $K(\mathbb{T}^l)$ is half the dimension of the space and the tori are Lagrangian submanifolds.

In [14] it is shown that, in the Lagrangian case, the perturbative equations can be studied very conveniently applying the change of variables given by the following matrix

$$\big[DK(\theta),\, J\big(K(\theta)\big)^{-1}DK(\theta)N(\theta)\big]. \tag{40}$$

In the Lagrangian case, the range of (40) is the tangent space of the manifold at $K(\theta)$. In our case, however, the range of (40) is not the whole space, but it will be a very good approximation of the center space. Then, we can apply a method very similar to the method in [14] for the equations in the center directions. The hyperbolic directions will be solved by other methods.

By the symplecticness of $F$ and the dynamical properties we have that the matrix of the symplectic structure with respect to the splitting $\mathcal{E}^s_{K(\theta)}\oplus\mathcal{E}^c_{K(\theta)}\oplus\mathcal{E}^u_{K(\theta)}$ has the form

$$J\big(K(\theta)\big)=\begin{pmatrix}0 & 0 & J^{su}\\ 0 & J^{cc} & 0\\ J^{us} & 0 & 0\end{pmatrix}, \tag{41}$$

where $J^{cc}$ is an antisymmetric form and $J^{su}(e_s,e_u)=-J^{us}(e_u,e_s)$.

Indeed from

$$u^\top J\big(K(\theta)\big)v=\Omega_{K(\theta)}(u,v)=\Omega_{F^n(K(\theta))}\big(DF^n\big(K(\theta)\big)u,\,DF^n\big(K(\theta)\big)v\big),\quad n\in\mathbb{Z},$$

we deduce, sending $n\to+\infty$ and using the hyperbolic conditions (expansion/contraction properties), that $u^\top J(K(\theta))v=0$ in the following cases:

- $u,v\in\mathcal{E}^s_{K(\theta)}$,
- $u,v\in\mathcal{E}^u_{K(\theta)}$,
- $u\in\mathcal{E}^s_{K(\theta)}\cup\mathcal{E}^u_{K(\theta)}$ and $v\in\mathcal{E}^c_{K(\theta)}$,
- $v\in\mathcal{E}^c_{K(\theta)}$ and $v\in\mathcal{E}^s_{K(\theta)}\cup\mathcal{E}^u_{K(\theta)}$

which implies the form (41). See also [9]. The form (41) proves that $J(K(\theta))^{-1}$ sends the center subspace into itself.

Since range $DK(\theta)$ is the tangent space of the torus $K(\mathbb{T}^l)$ and the dynamics on the torus is conjugated to a rotation, $DK(\theta)\mathbb{R}^l$ is contained in $\mathcal{E}^c_{K(\theta)}$. Moreover the previous property of $J(K)^{-1}$ implies that $J(K(\theta))^{-1}DK(\theta)\mathbb{R}^l$ also is contained in $\mathcal{E}^c_{K(\theta)}$. Instead of $J(K(\theta))^{-1}DK(\theta)$ we will consider the matrix $J(K(\theta))^{-1}DK(\theta)N(\theta)$ where $N(\theta)$ is the normalization $l\times l$-matrix $N(\theta)=[DK(\theta)^\top DK(\theta)]^{-1}$ introduced in (14). Both have the same range because $N(\theta)$ is non-singular. The role of $N$ is to provide some normalization for the symplectic conjugate.

Now we check that the range of $[DK(\theta), J(K(\theta))^{-1}DK(\theta)N(\theta)]$ is $2l$-dimensional. Indeed, assume that there is a linear combination

$$f = \sum_{j=1}^{l} \alpha_j DK(\theta)e_j + \sum_{j=1}^{l} \beta_j J\big(K(\theta)\big)^{-1} DK(\theta)N(\theta)e_j = 0.$$

Then, for $1 \leqslant k \leqslant l$, using the isotropic character of $T_{K(\theta)}K(\mathbb{T}^l)$

$$\begin{aligned} 0 = \Omega\big(DK(\theta)e_k, f\big) &= \sum_{j=1}^{l} \beta_j e_k^\top DK(\theta)^\top J\big(K(\theta)\big) J\big(K(\theta)\big)^{-1} DK(\theta)N(\theta)e_j \\ &= \sum_{j=1}^{l} \beta_j \langle e_k, e_j\rangle = \beta_k. \end{aligned}$$

This calculation shows that $f$ reduces to $\sum_{j=1}^{l} \alpha_j DK(\theta)e_j$. Moreover, for $1 \leqslant k \leqslant l$

$$\begin{aligned} 0 &= \Omega\big(J\big(K(\theta)\big)^{-1} DK(\theta)N(\theta)e_k, f\big) \\ &= \sum_{j=1}^{l} \alpha_j e_k^\top N(\theta)^\top DK(\theta)^\top J\big(K(\theta)\big)^{-\top} J\big(K(\theta)\big) DK(\theta)e_j \\ &= -\sum_{j=1}^{l} \alpha_j \langle e_k, e_j\rangle = \alpha_k. \end{aligned}$$

Hence $\alpha_j = \beta_j = 0$ for all $j = 1, \ldots, l$. We conclude that

$$\operatorname{range}\big[DK(\theta), J\big(K(\theta)\big)^{-1} DK(\theta)N(\theta)\big] = \mathcal{E}^c_{K(\theta)}.$$

Finally we define

$$\tilde{M}(\theta) = \big[DK(\theta), J\big(K(\theta)\big)^{-1} DK(\theta)N(\theta)\big]. \tag{42}$$

*4.2. Solving the linearized equation on the center subspace*

This section is devoted to the study of Eq. (22) projected on the center subspace. We denote

$$\Delta^c(\theta) = \Pi^c_{K(\theta)} \Delta(\theta).$$

Projecting the linearized equation (23) into the center space, we end up with the following equation

$$\Pi^c_{K(\theta+\omega)} G(\theta)\Lambda + DF\big(K(\theta)\big)\Delta^c(\theta) - \Delta^c(\theta+\omega) = -E^c(\theta), \tag{43}$$

where $E^c(\theta) = \Pi^c_{K(\theta+\omega)} E(\theta)$.

In Section 4.2.1, we will develop several identities and approximate identities that have a geometric nature. These identities will be used to reduce the equation on the center to constant coefficients equations of the form considered in Proposition 4.4. One important step is accomplished in Section 4.2.2 where we use the geometric identities and the theory of hyperbolic systems to obtain an approximate representation of the center space. Once this material is developed, we can establish the main result of this section, Proposition 4.19.

*4.2.1. Normalization procedure*

In the following, we construct a suitable representation for the matrix $DF(K(\theta))\tilde{M}(\theta)$. Recall that the $2d \times 2l$-matrix $\tilde{M}$ is given by

$$\tilde{M} = \left[DK,\, J(K)^{-1}DKN\right]. \tag{44}$$

As a motivation, we first consider the case when $K$ is a solution of (6). We search for a matrix $\mathcal{S}(\theta)$ satisfying

$$DF\big(K(\theta)\big)\tilde{M}(\theta) = \tilde{M}(\theta+\omega)\mathcal{S}(\theta), \tag{45}$$

where $\mathcal{S}(\theta)$ is upper triangular with identity matrices on the diagonal. Explicit expressions for $\mathcal{S}$ will be given later.

Differentiating Eq. (6) with respect to $\theta$, we get

$$DF\big(K(\theta)\big)DK(\theta) = DK(\theta+\omega).$$

This shows that $\mathcal{S}(\theta)$ has the form

$$\begin{pmatrix} \mathrm{Id}_l & A(\theta) \\ 0_l & B(\theta) \end{pmatrix}, \tag{46}$$

where $A(\theta)$ and $B(\theta)$ are $l \times l$ matrices. We will see that the choice of the second column of $\tilde{M}$ and the symplectic structure forces that $B(\theta) = \mathrm{Id}_l$. Then, it will be easy to compute an expression for $A$.

Indeed, from (44)–(46) we should have

$$\left[DF(K)J(K)^{-1}DKN\right](\theta) = DK(\theta+\omega)A(\theta) + \left[J(K)^{-1}DKN\right](\theta+\omega)B(\theta). \tag{47}$$

By the isotropic character of $K(\mathbb{T}^l)$ we have $DK^\top J(K)DK = 0$. Hence

$$\left[DK^\top J(K)\right](\theta+\omega)\left[DF(K)J(K)^{-1}DKN\right](\theta) = \left[DK^\top DKN\right](\theta+\omega)B(\theta). \tag{48}$$

Also by the symplecticness of $F$

$$J\big(K(\theta+\omega)\big)DF\big(K(\theta)\big) = J\big(F\big(K(\theta)\big)\big)DF\big(K(\theta)\big) = \left[DF(K)^{-\top}J(K)\right](\theta).$$

Then the left-hand side of (48) becomes

$$DK^\top(\theta+\omega)\left[DF(K)^{-\top}DKN\right](\theta) = \left[DK^\top DKN\right](\theta) = \mathrm{Id}_l.$$

With this we conclude that $B(\theta) = \mathrm{Id}_l$.

To obtain the expression of $A(\theta)$ we multiply (47) by $(DKN)(\theta+\omega)^\top$. Using $N^\top = N$ and $NDK^\top DK = \mathrm{Id}_l$ we get

$$A(\theta) = P(\theta+\omega)^\top\left[\left[DF(K)J(K)^{-1}P\right](\theta) - \left[J(K)^{-1}P\right](\theta+\omega)\right]. \tag{49}$$

We sum up the previous computations in the following lemma.

**Lemma 4.12.** *Let $K$ be a solution of Eq.* (6)*. Then we can write*

$$DF\big(K(\theta)\big)\tilde{M}(\theta) = \tilde{M}(\theta+\omega)\mathcal{S}(\theta),$$

*with*

$$\mathcal{S}(\theta) = \begin{pmatrix} \mathrm{Id}_l & A(\theta) \\ 0_l & \mathrm{Id}_l \end{pmatrix} \tag{50}$$

*and*

$$A(\theta) = P(\theta+\omega)^\top \big[\big[DF(K)J(K)^{-1}P\big](\theta) - \big[J(K)^{-1}P\big](\theta+\omega)\big],$$

*where the notation* $P(\theta) = DK(\theta)N(\theta)$ *was introduced in* (14).

The matrix $\tilde{M}(\theta)$ is not invertible since it is not square. However we can derive a generalized inverse for $\tilde{M}(\theta)$. As a motivation for subsequent developments, we first present Lemma 4.14 which deals with the geometric cancellations in the case of an exactly invariant torus. The case of interest for a KAM algorithm—when the torus is only approximately invariant—will be studied in Lemma 4.15 as a perturbation of Lemma 4.14.

A straightforward calculation shows that

$$\tilde{M}^\top J(K)\tilde{M} = \begin{pmatrix} L & \mathrm{Id}_l \\ -\mathrm{Id}_l & N^\top DK^\top J(K)^{-\top} DKN \end{pmatrix}. \tag{51}$$

**Remark 4.13.** When $J^2 = -\mathrm{Id}_{2d}$ we have $J^{-\top} = -J^\top = J$ and then

$$\tilde{M}^\top J(K)\tilde{M} = \begin{pmatrix} L & \mathrm{Id}_l \\ -\mathrm{Id}_l & N^\top LN \end{pmatrix}. \tag{52}$$

If moreover $K$ is a solution of $F \circ K = K \circ T_\omega$ the right-hand side matrix of (52) reduces to the standard symplectic matrix $J_0 = \begin{pmatrix} 0 & \mathrm{Id}_l \\ -\mathrm{Id}_l & 0 \end{pmatrix}$.

**Lemma 4.14.** *Let $K$ be a solution of* (6)*. Then the matrix $\tilde{M}^\top J(K)\tilde{M}$ is invertible and*

$$\big(\tilde{M}^\top J(K)\tilde{M}\big)^{-1} = \begin{pmatrix} N^\top DK^\top J(K)^{-\top} DKN & -\mathrm{Id}_l \\ \mathrm{Id}_l & 0 \end{pmatrix}.$$

**Proof.** It follows from (51) and the isotropic character of the invariant torus, i.e. $L = 0$. □

We now establish a similar result for approximate solutions, i.e. solutions of (7) up to error $E(\theta) = \mathcal{F}_\omega(\lambda, K)(\theta)$. We can expect this type of normalization to be true if the error and $\lambda$ are small enough. Following the calculations in Lemma 4.12, we obtain

$$DF\big(K(\theta)\big)\tilde{M}(\theta) = \tilde{M}(\theta+\omega)\begin{pmatrix} \mathrm{Id}_l & A(\theta) \\ 0_l & \mathrm{Id}_l \end{pmatrix} + O\big(E^c, DE^c\big).$$

More precisely, we introduce

$$e(\theta) = DF\big(K(\theta)\big)\tilde{M}(\theta) - \tilde{M}(\theta+\omega)\mathcal{S}(\theta), \tag{53}$$

where $\mathcal{S}$ is given by (50). If we denote $e(\theta) = (e_1(\theta), e_2(\theta))$, a simple algebraic computation yields

$$e_1(\theta) = DE^c(\theta) - D_\theta G^c(\theta)\lambda,$$
$$e_2(\theta) = \big[\big(DFJ^{-1}\big)(K)DKN\big](\theta) - DK(\theta+\omega)A(\theta) - \big[J^{-1}DKN\big](\theta+\omega) = O(E, DE)$$

by the choice of $A$, where $G^c(\theta) = \Pi^c_{K(\theta+\omega)}G(\theta)$.

The next step is to ensure the invertibility of the $2l \times 2l$-matrix $\tilde{M}^\top J(K)\tilde{M}$. According to expression (51), we can write

$$\tilde{M}(\theta)^\top J\big(K(\theta)\big)\tilde{M}(\theta) = V(\theta) + R(\theta),$$

where

$$V = \begin{pmatrix} 0 & \mathrm{Id}_l \\ -\mathrm{Id}_l & N^\top DK^\top J(K)^{-\top} DKN \end{pmatrix}$$

and

$$R = \begin{pmatrix} L & 0 \\ 0 & 0 \end{pmatrix}.$$

We have the following lemma, providing the desired invertibility result under a smallness assumption on $E$, namely (54) in the next lemma. Note that (54) has the same form as (24), but the constants could be slightly different since (24) should also accommodate (33), which is implied by conditions of the same form.

**Lemma 4.15.** *There exists a constant $C > 0$ such that if*

$$C\kappa\delta^{-(\nu+1)}\|E\|_\rho \leqslant 1/2 \tag{54}$$

*for some $0 < \delta < \rho/2$ then the matrix $\tilde{M}^\top(\theta)J(K(\theta))\tilde{M}(\theta)$ is invertible for $\theta \in D_{\rho-2\delta}$ and there exists a matrix $\tilde{V}(\theta)$ such that*

$$\big(\tilde{M}(\theta)^\top J\big(K(\theta)\big)\tilde{M}(\theta)\big)^{-1} = V(\theta)^{-1} + \tilde{V}(\theta)$$

*with*

$$\tilde{V}(\theta) = \left(\sum_{k=1}^{\infty}\big(V(\theta)^{-1}R(\theta)\big)^k\right)V(\theta)^{-1},$$

*where the series is absolutely convergent. Furthermore, we have the estimate*

$$\|\tilde{V}\|_{\rho-2\delta} \leqslant C'\kappa\delta^{-(\nu+1)}\|E\|_\rho, \tag{55}$$

*where the constant $C' > 0$ depends on $l$, $\nu$, $|F|_{C^1(B_r)}$, $|J|_{C^1(B_r)}$, $\|DK\|_\rho$, $\|N\|_\rho$ and $\|\Pi^c_{K(\theta)}\|_\rho$.*

**Proof.** The matrix $V(\theta)$ is invertible with

$$V^{-1} = \begin{pmatrix} N^\top DK^\top J(K)^{-\top} DKN & -\mathrm{Id}_l \\ \mathrm{Id}_l & 0 \end{pmatrix}.$$

We can write

$$\tilde{M}(\theta)^\top J\big(K(\theta)\big)\tilde{M}(\theta) = V(\theta)\big(\mathrm{Id}_{2l} + V(\theta)^{-1}R(\theta)\big).$$

To apply the Neumann series (and consequently justify the existence of the inverse of $\mathrm{Id}_{2l} + V^{-1}R$ as well as the estimates for its size), we have to estimate the term $V^{-1}R$. According to Lemma 4.8, we have the estimate for $L$

$$\|L\|_{\rho-2\delta} \leqslant C\kappa\delta^{-(\nu+1)}\big(\|E\|_\rho + \|G\|_\rho|\lambda|\big)$$

for all $\delta \in (0, \rho/2)$. Using Lemma 4.9 this leads to the estimate

$$\big\|V^{-1}R\big\|_{\rho-2\delta} \leqslant C\kappa\delta^{-(\nu+1)}\|E\|_\rho$$

for $0 < \delta < \rho/2$, where $C > 0$ depends on $l$, $\nu$, $|F|_{C^1(B_r)}$, $|J|_{C^1(B_r)}$, $\|DK\|_\rho$, $\|N\|_\rho$ and $\|\Pi^c_{K(\theta)}\|_\rho$. Because of assumption (54), we have that the right-hand side of the last equation is less than $1/2$.

Then the matrix $\mathrm{Id}_{2l} + V(\theta)^{-1}R(\theta)$ is invertible with

$$\big\|\big(\mathrm{Id}_{2l} + V^{-1}R\big)^{-1}\big\|_{\rho-2\delta} \leqslant \frac{1}{1 - \|V^{-1}R\|_{\rho-2\delta}} \leqslant 2.$$

This ends the proof of Lemma 4.15. □

*4.2.2. Identification of the center space*

In this section, we identify the center space as being very close (up to terms that can be bounded by the error in the invariance equation) to the range of the matrix $\tilde{M}$ introduced in (42), see Proposition 4.16. This will allow us to use the range of $\tilde{M}$ in place of $\mathcal{E}^c_{K(\theta)}$ without changing the quadratic character of the method.

**Proposition 4.16.** *Denote by $\Gamma_{K(\theta)}$ the range of $\tilde{M}(\theta)$ and by $\Pi^\Gamma_{K(\theta)}$ the projection onto $\Gamma_{K(\theta)}$ according to the splitting $\mathcal{E}^s_{K(\theta)} \oplus \Gamma_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)}$.*

*Then there exists a constant $C > 0$ such that if*

$$\delta^{-1}\|E\|_\rho \leqslant C$$

*we have the estimates (here $\mathrm{dist}_\rho$ stands for the distance between subspaces at the Grassmanian level)*

$$\mathrm{dist}_{\rho-2\delta}\big(\Gamma_{K(\theta)}, \mathcal{E}^c_{K(\theta)}\big) \leqslant C\delta^{-1}\|E\|_\rho,$$
$$\big\|\Pi^c_{K(\theta)} - \Pi^\Gamma_{K(\theta)}\big\|_{\rho-2\delta} \leqslant C\delta^{-1}\|E\|_\rho \tag{56}$$

*for every $\delta \in (0, \rho/2)$ and where $C$, as usual, depends on the non-degeneracy constants of the problem.*

**Proof.** Of course, the two inequalities in (56) are equivalent.

From (53) and Cauchy estimates, we have

$$\mathrm{dist}_{\rho-\delta}\big((DF\circ K)\Gamma_{K(\theta)}, \Gamma_{K(\theta)}\circ T_\omega\big) \leqslant C\delta^{-1}\|E\|_\rho.$$

Using again Eq. (53) and iterating it, we obtain for $n \geqslant 1$

$$DF\big(K(\theta+n\omega)\big)\times\cdots\times DF\big(K(\theta)\big)\tilde{M}(\theta) = \tilde{M}(\theta+n\omega)\mathcal{S}\big(\theta+(n-1)\omega\big)\times\cdots\times\mathcal{S}(\theta) + R_n,$$

where

$$\|R_n\|_{\rho-\delta} \leqslant C_n\delta^{-1}\|E\|_\rho$$

and $C_n$ depends on $n$.

Since $\mathcal{S}(\theta)$ is upper triangular with $\mathrm{Id}_l$ on the diagonal, we have

$$\mathcal{S}\big(\theta+(n-1)\omega\big)\times\cdots\times\mathcal{S}(\theta)=\begin{pmatrix}\mathrm{Id}_l & A(\theta+(n-1)\omega)+\cdots+A(\theta)\\ 0 & \mathrm{Id}_l\end{pmatrix}.$$

Therefore, by induction, we have for every $n\in\mathbb{N}$

$$\big\|DF\big(K(\theta+n\omega)\big)\cdots DF\big(K(\theta)\big)\tilde{M}(\theta)\big\|_{\rho-\delta}\leqslant Cn+C_n\delta^{-1}\|E\|_\rho.$$

Identical calculations give that

$$\big\|DF^{-1}\big(K(\theta-n\omega)\big)\cdots DF^{-1}\big(K(\theta)\big)\tilde{M}(\theta)\big\|_{\rho-\delta}\leqslant Cn+C_n\delta^{-1}\|E\|_\rho.$$

Note that, given any $\mu_3>1$ (as in Definition 3.4), there exists an integer $n_{\mu_3}\geqslant 0$ such that for all $n\geqslant n_{\mu_3}$, we have $Cn<\mu_3^n$. Consequently, choosing such $n_{\mu_3}$ there exists a constant $C$ such that if the error satisfies

$$\delta^{-1}\|E\|_\rho\leqslant C,$$

we have $Cn+C_n\delta^{-1}\|E\|_\rho<\mu_3^n$. In other words, the above estimates hold for all sufficiently large $n$, provided that we impose a suitable smallness condition on $\delta^{-1}\|E\|_\rho$.

As a consequence, $\varGamma_{K(\theta)}$ is an approximately invariant bundle, and we also have bounds on the rate of growth of the co-cycle both in positive and negative times. Using standard tools in the theory of hyperbolic systems (see Proposition 5.2 below where we prove the result for all the bundles), this shows that indeed one can find a true invariant subspace $\tilde{\mathcal{E}}_{K(\theta)}$ close to $\varGamma_{K(\theta)}$. Since this invariant subspace should be of the same dimension of the center space $\mathcal{E}^c_{K(\theta)}$, we deduce that

$$\tilde{\mathcal{E}}_{K(\theta)}=\mathcal{E}^c_{K(\theta)}.$$

See also Remark 5.6 below. □

*4.2.3. Final estimates of the solution on the center subspace*

We can now finish the solution of Eq. (23) on the center subspace. We recall the linearized equation around $(\lambda,K)$ projected on the center subspace:

$$G^c(\theta)\Lambda+DF\big(K(\theta)\big)\Delta^c(\theta)-\Delta^c(\theta+\omega)=-E^c(\theta). \tag{57}$$

We make the change the unknowns in (57)

$$\Delta^c(\theta)=\tilde{M}(\theta)W(\theta)+\hat{e}(\theta)W(\theta), \tag{58}$$

where

$$\hat{e}=\Pi^c_{K(\theta+\omega)}-\Pi^\Gamma_{K(\theta+\omega)} \tag{59}$$

which was estimated in Proposition 4.16.

Substituting (58) into Eq. (57) we get

$$\begin{aligned}&DF\big(K(\theta)\big)\tilde{M}(\theta)W(\theta)-\tilde{M}(\theta+\omega)W(\theta+\omega)\\&\quad=-E^c(\theta)-G^c(\theta)\Lambda+\hat{e}(\theta+\omega)W(\theta+\omega)-DF\big(K(\theta)\big)\hat{e}(\theta)W(\theta).\end{aligned} \tag{60}$$

We anticipate that the term $\hat{e}W$ will be quadratic in the error. Similarly, writing

$$G^c = \Pi^{\Gamma}_{K(\theta+\omega)} G + \hat{e}G,$$

we also anticipate that the term $\hat{e}G\Lambda$ will be quadratic in the error. Since the function $G$ will be chosen to be $J(K_0)^{-1}DK_0 \circ T_\omega$, namely in $\Gamma_{K_0(\theta+\omega)}$, we drop the index from $G^c$, writing $G$ directly. As a consequence, we will ignore these two terms and consider instead the equation

$$DF\big(K(\theta)\big)\tilde{M}(\theta)W(\theta) - \tilde{M}(\theta+\omega)W(\theta+\omega) = -E^c(\theta) - G(\theta)\Lambda \tag{61}$$

which differs from the linearized equation in the term $(\hat{e}W)\circ T_\omega - DF(K)(\hat{e}W) - \hat{e}G\Lambda$. Note that, ignoring this term we obtain an equation where all the terms are in the range of $\tilde{M}$.

We multiply Eq. (61) by $\tilde{M}(\theta+\omega)^\top J(K(\theta+\omega))$

$$\begin{aligned}&\big[\tilde{M}^\top J(K)\big](\theta+\omega)DF\big(K(\theta)\big)\tilde{M}(\theta)W(\theta) - \big[\tilde{M}^\top J(K)\big](\theta+\omega)\tilde{M}(\theta+\omega)W(\theta+\omega)\\ &\quad= -\big[\tilde{M}^\top J(K)\big](\theta+\omega)\big[E^c(\theta)+G(\theta)\Lambda\big].\end{aligned}$$

Using Lemma 4.15 (invertibility of $\tilde{M}^\top J(K)\tilde{M}$) and Eq. (53), we can write

$$\begin{aligned}&\left[\begin{pmatrix}\mathrm{Id}_l & A(\theta)\\ 0_l & \mathrm{Id}_l\end{pmatrix} + B(\theta)\right]W(\theta) - W(\theta+\omega)\\ &\quad= p_1(\theta)+p_2(\theta) - \big[\tilde{M}^\top J(K)\tilde{M}\big](\theta+\omega)^{-1}\big[\tilde{M}^\top J(K)\big](\theta+\omega)G(\theta)\Lambda,\end{aligned} \tag{62}$$

where

$$B(\theta) = \big[\tilde{M}^\top J(K)\tilde{M}\big](\theta+\omega)^{-1}\big[\tilde{M}^\top J(K)\big](\theta+\omega)e(\theta), \tag{63}$$

$$p_1(\theta) = -V(\theta+\omega)^{-1}\big[\tilde{M}^\top J(K)\big](\theta+\omega)E^c(\theta) \tag{64}$$

and

$$p_2(\theta) = -\tilde{V}(\theta+\omega)\big[\tilde{M}^\top J(K)\big](\theta+\omega)E^c(\theta). \tag{65}$$

In the following lemma, we sum up the previous computations and estimate the terms in Eq. (62).

**Lemma 4.17.** *Assume $\omega \in D(\kappa,\nu)$ and $\delta$ and $\|E\|_\rho$ satisfy* (54)*. Eq.* (61) *can be written in the form*

$$\begin{aligned}&\left[\begin{pmatrix}\mathrm{Id}_l & A(\theta)\\ 0_l & \mathrm{Id}_l\end{pmatrix} + B(\theta)\right]W(\theta) - W(\theta+\omega)\\ &\quad= p_1(\theta)+p_2(\theta) - \big[\tilde{M}^\top J(K)\tilde{M}\big]^{-1}(\theta+\omega)\big[\tilde{M}^\top J(K)\big](\theta+\omega)G(\theta)\Lambda,\end{aligned} \tag{66}$$

*where the matrix $B$ and the vectors $p_1$ and $p_2$ are given by expressions* (63)–(65) *respectively.*

*The following estimates hold*:

$$\|p_1\|_\rho \leqslant C\|E\|_\rho, \tag{67}$$

*where $C$ only depends on $|J|_{C^1(B_r)}$, $\|N\|_\rho$, $\|DK\|_\rho$ and $\|\Pi^c_{K(\theta)}\|_\rho$. For $p_2$ and $B$ we have*

$$\|p_2\|_{\rho-2\delta} \leqslant C\kappa\delta^{-(\nu+1)}\|E\|^2_\rho \tag{68}$$

*and*

$$\|B\|_{\rho-2\delta} \leqslant C\delta^{-1}\big(\|E\|_\rho + |\lambda|\big), \tag{69}$$

*where C depends on l, $\nu$, $\|N\|_\rho$, $\|DK\|_\rho$, $|F|_{C^1(B_r)}$, $|J|_{C^1(B_r)}$ and $\|\Pi^c_{K(\theta)}\|_\rho$.*

**Proof.** Since the matrix $V^{-1}$ does not depend on $L$ the estimate (67) is obvious from the formula (64) for $p_1(\theta)$.

According to the proof of Lemma 4.15 the estimate (68) then comes from estimate (55). We turn to the estimate on $B$. We have

$$B(\theta) = \big(V(\theta+\omega)^{-1} + \tilde{V}(\theta+\omega)\big)\tilde{M}(\theta+\omega)^\top J\big(K(\theta+\omega)\big)e(\theta).$$

This leads to

$$\begin{aligned}\|B\|_{\rho-2\delta} &\leqslant \big\|V(\theta+\omega)^{-1}\big\|_{\rho-2\delta}\big\|\tilde{M}(\theta+\omega)^\top J\big(K(\theta+\omega)\big)e(\theta)\big\|_{\rho-2\delta}\\ &\quad + \big\|\tilde{V}(\theta+\omega)\tilde{M}(\theta+\omega)^\top J\big(K(\theta+\omega)\big)e(\theta)\big\|_{\rho-2\delta}.\end{aligned}$$

Therefore, using estimate (55) and Cauchy estimates, we end up with

$$\|B\|_{\rho-2\delta} \leqslant C\big(\delta^{-1}\|E\|_\rho + \delta^{-1}|\lambda| + \kappa\delta^{-(\nu+1)}\|E\|_\rho\big(\delta^{-1}\|E\|_\rho + \delta^{-1}|\lambda|\big)\big).$$

This leads to the desired result thanks to the smallness assumption on $\|E\|_\rho$. $\square$

*4.2.4. Approximate solvability of the equations on the center subspace*

This section is devoted to solving approximately (up to quadratic error) the linearized equation (66), as is usual in KAM theory.

To this end, we introduce the following operator

$$\mathcal{L}W(\theta) = \begin{pmatrix}\mathrm{Id}_l & A(\theta)\\ 0_l & \mathrm{Id}_l\end{pmatrix}W(\theta) - W(\theta+\omega).$$

Eq. (66) can be written as

$$\begin{aligned}&\mathcal{L}W(\theta) + B(\theta)W(\theta)\\ &\quad = p_1(\theta) + p_2(\theta) - \big[\tilde{M}^\top J(K)\tilde{M}\big](\theta+\omega)^{-1}\big[\tilde{M}^\top J(K)\big](\theta+\omega)G(\theta)\Lambda.\end{aligned} \tag{70}$$

According to estimates in Lemma 4.17, we have $p_2 = O(\|E\|^2_\rho)$, $p_1 = O(\|E\|_\rho)$ and $B = O(\|E\|_\rho + |\lambda|)$. Solving approximately Eq. (70) with an error "quadratic" in $E$ does not affect the convergence of the Newton scheme. See [73] for an abstract discussion and [74] for several concrete applications.

Eq. (70) does not fit into the framework of Proposition 4.4 since the average of the right-hand side is generically non-zero. However, by using the increment parameter $\Lambda$, we can make this average equal to zero. Furthermore, Eq. (70) has two unknowns (the two symplectic coordinates). Thanks to Lemma 4.15, one can write the term

$$\big[\tilde{M}^\top J(K)\tilde{M}\big]^{-1}(\theta+\omega)\big[\tilde{M}^\top J(K)\big](\theta+\omega)G(\theta)\Lambda = q_1(\theta)\Lambda + q_2(\theta)\Lambda,$$

where the matrix $q_1$ (which is $2l \times l$) is

$$q_1(\theta) = V(\theta+\omega)^{-1}\tilde{M}(\theta+\omega)^\top J\big(K(\theta+\omega)\big)G(\theta)$$

and $q_2$ satisfies for all $\delta \in (0, \rho/2)$

$$\|q_2\|_{\rho-2\delta} \leqslant C\kappa\delta^{-(\nu+1)}\|G\|_\rho\|E\|_\rho,$$

where the constant $C$ depends on $l$, $\nu$, $\|N\|_\rho$, $\|DK\|_\rho$, $|F|_{C^1(B_r)}$, $|J|_{C^1(B_r)}$ and $\|\Pi^c_{K(\theta)}\|_\rho$.

We define an approximate solution of (70) as a solution of the following equation (71), obtained by removing the terms containing $B$ and $q_2$ from Eq. (66), which was equivalent to (61). We recall that (61) was obtained from the Newton step by removing the terms that contained $\hat{e}$. As we will see, all these eliminations do not change the quadratic convergence of the method. Consider now

$$\mathcal{L}v(\theta) = p_1(\theta) - q_1(\theta)\Lambda. \tag{71}$$

Thanks to the non-degeneracy conditions (see Definition 3.4), we obtain the following result.

**Proposition 4.18.** *Assume $\omega \in D(\kappa, \nu)$ and $(\lambda, K)$ is a non-degenerate pair (i.e. $(\lambda, K) \in ND(\rho)$). If the error $\|E\|_\rho$ satisfies* (54) *and the smallness assumptions in Proposition* 4.16*, there exist a mapping $v$, analytic on $D_{\rho-2\delta}$ and a vector $\Lambda \in \mathbb{R}^l$ solving Eq.* (71)*.*

*Moreover there exists a constant $C > 0$ depending on $\nu$, $l$, $\|K\|_\rho$, $|(\mathrm{avg}(Q))^{-1}|$, $|(\mathrm{avg}(A))^{-1}|$, $\|N\|_\rho$ and $\|\Pi^c_{K(\theta)}\|_\rho$ such that*

$$\|v\|_{\rho-2\delta} < C\kappa^2\delta^{-2\nu}\|E\|_\rho$$

*and*

$$|\Lambda| < C\|E\|_\rho.$$

**Proof.** We denote $R(\theta)$ the right-hand side of Eq. (71), i.e. we solve

$$\mathcal{L}v(\theta) = R(\theta), \tag{72}$$

with

$$R = p_1 - q_1\Lambda.$$

We now decompose Eq. (72) into symplectically conjugate coordinates, i.e. $v = (v_1, v_2)^\top$, $R(\theta) = (R_1(\theta), R_2(\theta))^\top$. Therefore, Eq. (72) is equivalent to

$$v_1(\theta) + A(\theta)v_2(\theta) = v_1(\theta+\omega) + R_1(\theta),$$
$$v_2(\theta) = v_2(\theta+\omega) + R_2(\theta).$$

A simple computation shows that

$$R_2(\theta) = -\big[DK^\top J(K)\big] \circ T_\omega G(\theta).$$

We choose $\Lambda \in \mathbb{R}^l$ such that

$$\mathrm{avg}(R_2) = 0.$$

According to Proposition 4.4, if $\mathrm{avg}(R_2) = 0$ the equation in $v_2$ admits an analytic solution with arbitrary average on $D_{\rho-\delta}$ and we have the estimate

$$\|v_2\|_{\rho-\delta} \leqslant C\kappa\delta^{-\nu}\|R_2\|_\rho + \big|\mathrm{avg}(v_2)\big|. \tag{73}$$

Then we choose $\operatorname{avg}(v_2)$ such that $\operatorname{avg}(R_1 - Av_2) = 0$, which allows us to solve uniquely the equation in $v_1$, the function $v_1$ being of zero average. Furthermore, we have the estimate

$$\|v_1\|_{\rho-2\delta} \leqslant C\kappa\delta^{-\nu}\|R_1 - Av_2\|_{\rho-\delta}.$$

We now turn to the estimates. First we estimate $\Lambda$. The vector $\Lambda \in \mathbb{R}^l$ is such that

$$\operatorname{avg}\big(DK^\top(\omega+\theta)J\big(K(\omega+\theta)\big)\big(E^c(\theta) + G(\theta)\Lambda\big)\big) = 0.$$

This leads to

$$\begin{aligned}
&\operatorname{avg}\big(DK^\top(\omega+\theta)J\big(K(\omega+\theta)\big)G(\theta)\big)\Lambda \\
&\quad = -\operatorname{avg}\big(DK^\top(\omega+\theta)J\big(K(\omega+\theta)\big)E^c(\theta)\big).
\end{aligned}$$

Note that by the definition of $P$ and the fact that $N$ is symmetric, the matrix which applies to $\Lambda$ is the average of $Q$ which, by hypothesis, is invertible. This leads to the desired estimate for $\Lambda$.

We now estimate the solution $v$. From the expression of $R$ and the value of $\Lambda$ obtained above, we have that there exists a constant $C$ such that

$$\|R_i\|_\rho \leqslant C\|E\|_\rho,$$

for $i = 1, 2$. Furthermore, we choose $\operatorname{avg}(v_2)$ such that $\operatorname{avg}(R_1 - Av_2) = 0$, i.e.

$$\operatorname{avg}(v_2) = \operatorname{avg}(A)^{-1}\big(\operatorname{avg}(R_1) - \operatorname{avg}\big(Av_2^\perp\big)\big),$$

where $v_2 = v_2^\perp + \operatorname{avg}(v_2)$. This is possible since by the twist condition $\operatorname{avg}(A)$ is invertible. Thanks to estimate (73), this leads to the desired result. $\square$

We now come back to the solutions of (43). The above procedure allows us to prove the following proposition, providing an approximate solution of the projection of $D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\Lambda, \Delta) = -E$ on the center subspace.

**Proposition 4.19.** *Let $(\Lambda, W)$ be as in Proposition 4.18 and assume the hypotheses of that proposition hold. Define $\Delta^c(\theta) = \tilde{M}(\theta)W(\theta) + \hat{e}(\theta)W(\theta)$ and obtain $W$ and $\lambda$ as indicated above.*

*Then, $(\lambda, \Delta^c)$ is an approximate solution of (43) and we have the following estimates*

$$\begin{gathered}
\big\|\Delta^c\big\|_{\rho-2\delta} \leqslant C\kappa^2\delta^{-2\nu}\|E\|_\rho, \\
|\Lambda| \leqslant C\|E\|_\rho,
\end{gathered}$$

*where the constant $C$ depends on $\nu$, $l$, $|(\operatorname{avg}(Q))^{-1}|$, $|(\operatorname{avg}(A))^{-1}|$, $\|N\|_\rho$, $\|G\|_\rho$ and $\|\Pi^c_{K(\theta)}\|_\rho$. Moreover*

$$\begin{aligned}
\big\|D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)\big(\Lambda, \Delta^c\big) + E^c\big\|_{\rho-2\delta} &\leqslant C\kappa^3\delta^{-(3\nu+1)}\big(\|E\|_\rho^2 + \|E\|_\rho|\lambda|\big) + C\delta^{-1+\nu}\|E\|_\rho^2 \\
&\leqslant C\kappa^3\delta^{-(3\nu+1)}\|E\|_\rho^2,
\end{aligned} \tag{74}$$

*where the constant $C$ depends on $l$, $\kappa$, $\nu$, $|F|_{C^1(B_r)}$, $\|DK\|_\rho$, $\|N\|_\rho$, $|(\operatorname{avg}(A))^{-1}|$, $|(\operatorname{avg}(Q))^{-1}|$ and $\|G\|_\rho$.*

**Proof.** The first estimate comes from the previous Proposition 4.18.

For the second one (74), we recall that we have

$$\begin{aligned} &D_{\lambda,K}\mathcal{F}_\omega\big(\Lambda,\Delta^c\big)(\theta)+E^c(\theta)\\ &\quad=-\big[\tilde{M}^\top J(K)\tilde{M}\big](\theta+\omega)\big[B(\theta)v(\theta)-p_2(\theta)\big]-q_2(\theta)\Lambda\\ &\qquad+\hat{e}(\theta+\omega)W(\theta+\omega)-DF\big(K(\theta)\big)\hat{e}(\theta)W(\theta). \end{aligned} \tag{75}$$

The first term in the right-hand side of (75) is estimated in Proposition 4.18, see estimates (68)–(69). The second one comes from the vanishing Lemma 4.9. The third term is estimated in Proposition 4.16. □

### 4.3. Solving the linearized equations on the hyperbolic subspaces

According to the splitting (10), there exist projections on the linear spaces $\mathcal{E}^s_{K(\theta)}$ and $\mathcal{E}^u_{K(\theta)}$. The analytic regularity of the splitting implies that the dependence of these projections in $\theta$ is analytic in the same domain as the spaces. We denote $\Pi^s_{K(\theta+\omega)}$ (resp. $\Pi^u_{K(\theta+\omega)}$) the projections (of base $K(\theta+\omega)$) on the stable (resp. unstable) invariant subspace.

We project Eq. (22) on the stable and unstable spaces to obtain

$$\Pi^s_{K(\theta+\omega)}\big(G(\theta)\Lambda+DF\big(K(\theta)\big)\Delta(\theta)-\Delta(\theta+\omega)\big)=-\Pi^s_{K(\theta+\omega)}E(\theta), \tag{76}$$

$$\Pi^u_{K(\theta+\omega)}\big(G(\theta)\Lambda+DF\big(K(\theta)\big)\Delta(\theta)-\Delta(\theta+\omega)\big)=-\Pi^u_{K(\theta+\omega)}E(\theta). \tag{77}$$

Furthermore, thanks to the invariance of the splitting, we can write

$$\Pi^s_{K(\theta+\omega)}DF\big(K(\theta)\big)\Delta(\theta)=DF\big(K(\theta)\big)\Pi^s_{K(\theta)}\Delta(\theta)$$

for the stable part and

$$\Pi^u_{K(\theta+\omega)}DF\big(K(\theta)\big)\Delta(\theta)=DF\big(K(\theta)\big)\Pi^u_{K(\theta)}\Delta(\theta)$$

for the unstable one. Introducing the change of variables $\theta'=T_\omega(\theta)$ and the notation $\Delta^{s,u}(\theta')=\Pi^{s,u}_{K(\theta')}\Delta(\theta')$, Eqs. (76)–(77) can be written in the following form

$$DF(K)\circ T_{-\omega}(\theta')\Delta^s\big(T_{-\omega}(\theta')\big)-\Delta^s(\theta')=-\tilde{E}^s(\theta',\Lambda), \tag{78}$$

where

$$\tilde{E}^s(\theta',\Lambda)=\Pi^s_{K(\theta')}\big(G\big(T_{-\omega}(\theta')\big)\Lambda\big)+\Pi^s_{K(\theta')}E\circ T_{-\omega}(\theta')$$

and

$$DF(K)\circ T_{-\omega}(\theta')\Delta^u\big(T_{-\omega}(\theta')\big)-\Delta^u(\theta')=-\tilde{E}^u(\theta',\Lambda), \tag{79}$$

where

$$\tilde{E}^u(\theta',\Lambda)=\Pi^u_{K(\theta')}\big(G\big(T_{-\omega}(\theta')\big)\Lambda\big)+\Pi^u_{K(\theta')}E\circ T_{-\omega}(\theta').$$

The following proposition provides an existence result together with estimates for Eqs. (78)–(79).

**Proposition 4.20.** *Fix $\rho > 0$. Then Eq.* (78) (*resp.* (79)) *admits a unique analytic solution $\Delta^s : D_\rho \to \mathcal{E}^s_{K(\theta)}$ (resp. $\Delta^u : D_\rho \to \mathcal{E}^u_{K(\theta)}$). Furthermore there exists a constant $C$ such that*

$$\|\Delta^{s,u}\|_\rho \leqslant C\big(\|E\|_\rho + |\Lambda|\big), \tag{80}$$

*where the constant $C$ depends on the hyperbolicity constant $\mu_1$ (resp. $\mu_2$), the norm of the projection $\|\Pi^s_{K(\theta)}\|_\rho$ (resp. $\|\Pi^u_{K(\theta)}\|_\rho$), $\|G(\theta)\|_\rho$ and the constant $C_h$ involved in* (11) (*resp.* (12)).

**Proof.** We give the proof for the stable case, the unstable one being similar and left to the reader. Using Eq. (78) iteratively, we claim that its solution is given by

$$\Delta^s(\theta') = \sum_{k=0}^{\infty} \big(DF(K) \circ T_{-\omega}(\theta') \times \cdots \times DF(K) \circ T_{-\omega}(\theta')\big)\tilde{E}^s\big(T_{-(k+1)\omega}(\theta'), \Lambda\big). \tag{81}$$

Using the condition on the co-cycles over $T_{-\omega}$ (see Eq. (11)), the series converges uniformly on $D_\rho$ and one can estimate

$$\|\Delta^s\|_\rho \leqslant C_h \|\tilde{E}^s\|_\rho \sum_{k=0}^{\infty} \mu_1^k \leqslant C\big(\|E\|_\rho + |\Lambda|\big) \tag{82}$$

since $\mu_1 < 1$. Once we know that the series converges uniformly, we can rearrange the terms and get that (81) is indeed a solution. The proof in the case of the unstable space follows in the same way, multiplying Eq. (79) by $(DF(K) \circ T_{-\omega})^{-1}$ and using the condition (12) on the co-cycles. $\square$

## 5. Iteration of the Newton step and convergence

In the following we describe precisely the iteration of the Newton method. As it is standard in KAM theory, we show that if the initial error $\|E_0\|_{\rho_0}$ is small enough, one can choose the domain loss, so that the iterative scheme converges to a solution of (7) which moreover is close to the initial approximate solution. As a consequence of the vanishing lemma (i.e. Lemma 4.9) one gets $\lambda = 0$ and then a solution of

$$F \circ K = K \circ T_\omega.$$

In the rest of this section, we are under the assumptions of Theorem 3.11.

### 5.1. Estimates for one step of the Newton method

Recall that we have implemented a step showing that, given an approximate solution, $(\lambda_{m-1}, K_{m-1})$ of (7), which is non-degenerate in the sense of Definition 3.4 and satisfies the conditions (33) of Lemma 4.9 and (24) of Lemma 4.2, then we find an approximate solution $(\Lambda_{m-1}, \Delta_{m-1})$ of the Newton equation. That is, we can find

$$D_{\lambda,K}\mathcal{F}_\omega(\lambda_{m-1}, K_{m-1})(\Lambda_{m-1}, \Delta_{m-1}) = -E_{m-1} + R_m$$

with $E_{m-1}(\theta) = \mathcal{F}_\omega(\lambda_{m-1}, K_{m-1})(\theta)$ and $R_m$ "quadratically" small. If $E$ is defined in $D_{\rho_{m-1}}$, the Newton correction $\Delta_{m-1}$ is defined in a smaller domain $D_{\rho_m}$, $\rho_m = \rho_{m-1} - \delta_m$. The precise results on the step are collected in Lemma 4.2 and the description of the step is given along the proof.

The next result Proposition 5.1, makes precise the observation that, if we can define $F \circ K_m$ then it is possible to show that the new remainder is quadratic. Furthermore, we will show that the change in the non-degeneracy assumptions can be estimated by the size of the error.

The assumption that $F \circ K_m$ can be defined, requires only that the range of $K_m = K_{m-1} + \Delta_{m-1}$ does not get out the domain of $F$. This will be implied by smallness assumptions on $\Delta$ that, using the conclusions of Lemma 4.2, are implied by assumption (84). As it will turn out, the assumption (84) is stronger than (24) so that (84) is enough to ensure that we can carry out a Newton step as indicated.

In subsequent sections, we will show that if we choose the sequence of domain losses $\delta_m = \frac{1}{4}\delta_0 2^{-m}$, and the error is small enough, the process can be iterated infinitely often and converges to a solution of the equation. The argument also shows that the hyperbolic splitting converges.

**Proposition 5.1.** *Choose an initial approximation $\lambda_0 = 0$, $K_0$, where $K_0 \in ND(\rho_0)$. Assume that $K_0(D_{\rho_0})$, the range of $K_0$ is at a distance $r > 0$ from complement of the domain of definition of $F$.*

*Assume $(\lambda_{m-1}, K_{m-1}) \in ND(\rho_{m-1})$ is an approximate solution of Eq.* (7) *and that the following holds*

$$\|K_{m-1} - K_0\|_{\rho_{m-1}} < r/2, \tag{83}$$

*where $r$ is chosen sufficiently small so that we can apply Lemma* 4.9*, the constants in $ND(\rho_{m-1})$ are chosen uniformly and that the range of $K_{m-1}$ is inside the domain of definition of $F$. Assume furthermore that* (24) *holds so that we can apply Lemma* 4.2*.*

*Denote by $C$ expressions that depend only on $\nu$, $l$, $|F|_{C^1(B_r)}$, $\|DK_{m-1}\|_{\rho_{m-1}}$, $\|\Pi^{s,c,u}_{K_{m-1}(\theta)}\|_{\rho_{m-1}}$, $|(\mathrm{avg}(Q_{m-1}))^{-1}|$ and $|(\mathrm{avg}(A_{m-1}))^{-1}|$ and, hence, can be chosen uniformly if $K_{m-1}$ is in a sufficiently small neighborhood of $K_0$ as indicated in* (83)*.*

*Let $\Lambda_{m-1}$, $\Delta_{m-1}$ be the corrections produced in Lemma* 4.2*.*

*If $E_{m-1}$ is small enough such that*

$$C\kappa\delta_{m-1}^{-2\nu-1}\|E_{m-1}\|_{\rho_{m-1}} < r/2, \tag{84}$$

*then the set $(K_{m-1} + \Delta_{m-1})(D_{\rho_{m-1}-\delta_{m-1}})$ is well inside the domain of definition of $F$ and $E_m(\theta) = \mathcal{F}_\omega(\lambda_m, K_m)(\theta)$ satisfies (defining $\rho_m = \rho_{m-1} - 3\delta_{m-1}$)*

$$\|E_m\|_{\rho_m} \leqslant C\kappa^4\delta_{m-1}^{-4\nu}\|E_{m-1}\|^2_{\rho_{m-1}}. \tag{85}$$

**Proof.** We have $\Delta_{m-1}(\theta) = \Pi^h_{K_{m-1}(\theta)}\Delta_{m-1}(\theta) + \Pi^c_{K_{m-1}(\theta)}\Delta_{m-1}(\theta)$, where $\Pi^h_{K_{m-1}(\theta)}$ is the projection on the hyperbolic subspace. Propositions 4.19 and 4.20 respectively, particularized to $\delta_{m-1}$ give us that

$$\|\Delta_{m-1}\|_{\rho_m} \leqslant C\kappa^2\delta_{m-1}^{-2\nu}\|E_{m-1}\|_{\rho_{m-1}}$$

and using Cauchy inequalities

$$\|D\Delta_{m-1}\|_{\rho_m} \leqslant C\kappa^2\delta_{m-1}^{-2\nu-1}\|E_{m-1}\|_{\rho_{m-1}}.$$

Using (84), and the previous estimates on $\Delta_{m-1}$, we see that the range of $K_m \equiv K_{m-1} + \Delta_{m-1}$ is well inside the domain of definition of $F$ so that we can define $F \circ K_m$.

Define the remainder of the Taylor expansion

$$\begin{aligned}\mathcal{R}(\lambda, \lambda', K, K') = {} & \mathcal{F}_\omega(\lambda, K) - \mathcal{F}_\omega(\lambda', K') \\ & - D_{\lambda,K}\mathcal{F}_\omega(\lambda, K)(\lambda - \lambda', K - K').\end{aligned}$$

Then we have

$$E_m(\theta) = E_{m-1}(\theta) + D_{\lambda,K}\mathcal{F}_\omega\big(\lambda_{m-1}, K_{m-1}(\theta)\big)\big(\Lambda_{m-1}, \Delta_{m-1}(\theta)\big)$$
$$+ \mathcal{R}(\lambda_{m-1}, \lambda_m, K_{m-1}, K_m)(\theta).$$

Using estimate (74), for the error in solving the center equation and recalling that the equations on the hyperbolic subspace are *exactly* solved, we have

$$\big\| E_{m-1} + D_{\lambda,K}\mathcal{F}_\omega(\lambda_{m-1}, K_{m-1})(\Lambda_{m-1}, \Delta_{m-1}) \big\|_{\rho_m}$$
$$\leqslant c_{m-1}\kappa^3 \delta_{m-1}^{-(3\nu+1)} \|E_{m-1}\|^2_{\rho_{m-1}}.$$

Estimate (85) then follows from Taylor's remainder bound

$$\big| F\circ(K_{m-1}+\Delta_{m-1})(\theta) - F\circ K_{m-1}(\theta) - DF\circ K_{m-1}(\theta)\Delta_{m-1}(\theta)\big|$$
$$\leqslant C\big\|D^2F\big\|_{\mathcal{B}}\big|\Delta_{m-1}(\theta)\big|^2.$$

Note that, since $\delta_n$ go to zero, we can assume that the estimates from the Taylor remainder are larger than those from the error of the solution. □

### 5.2. Change of the hyperbolicity and the non-degeneracy conditions in the iterative step

The main goal of this section is to estimate the change of the non-degeneracy conditions in terms of the size of the error at the beginning of the iterative step.

We begin by estimating the change in the invariant splitting. Later, we will estimate the change in the twist conditions.

The first result Proposition 5.2 is a standard result in the theory of normally hyperbolic sets that allows us to conclude that if we are given an approximately invariant splitting, which has some hyperbolicity, then there is a truly invariant splitting nearby. The proof is a reformulation in an *a posteriori* format of standard arguments on the stability of hyperbolic splittings [23,38,56,58,67]. Since this will be part of an iterative procedure, we also need to obtain rather detailed estimates.

As a corollary, we will obtain that, when we change the embeddings $K$ in the iterative step, the change of the invariant subspaces will be controlled by the change in the embedding. Of course, since the twist conditions are just properties of the restriction of the derivative to an appropriate subspace, we will obtain that the size of the change in the twist conditions is controlled by the size of the change of the embedding.

Notice also that Proposition 5.2 provides a way to verify the hyperbolicity out of a finite calculation and in particular, out of the results of a numerical calculation. We have also used Proposition 5.2 to identify the center space in Section 4.2.2.

**Proposition 5.2.** *Assume that there is an analytic splitting*

$$T_{K(\theta)}\mathcal{M} = \tilde{\mathcal{E}}^s_{K(\theta)} \oplus \tilde{\mathcal{E}}^c_{K(\theta)} \oplus \tilde{\mathcal{E}}^u_{K(\theta)} \tag{86}$$

*which is approximately invariant under the co-cycle $DF\circ K$ over $T_\omega$. That is,*

$$\operatorname{dist}_\rho\big(DF\circ K(\theta)\tilde{\mathcal{E}}^{c,s,u}_{K(\theta)}, \tilde{\mathcal{E}}^{c,s,u}_{K(\theta+\omega)}\big) \leqslant \delta,$$

*where* $\operatorname{dist}_\rho$ *stands for the supremum of the distance when $\theta$ belongs to $D_\rho$, the complex extension of the torus defined in* (3)*. We denote by $\Pi^{s,c,u}$ the projections corresponding to the above splitting.*

*Assume, moreover that, for some $N \in \mathbb{N}$, $0 < \tilde{\mu}_1, \tilde{\mu}_2 < 1$, and some $1 \leqslant \tilde{\mu}_3$, such that* $\max(\tilde{\mu}_1, \tilde{\mu}_2)\cdot \tilde{\mu}_3 < 1$*, we have*

$$\left|DF\circ K\circ T_\omega^{N-1}(\theta)\times\cdots\times DF\circ K(\theta)v\right|\leqslant\tilde{\mu}_1^N|v|\quad\forall v\in\tilde{\mathcal{E}}^s_{K(\theta)},\tag{87}$$

$$\left|DF^{-1}\circ K\circ T_\omega^{-(N-1)}(\theta)\times\cdots\times DF^{-1}\circ K(\theta)v\right|\leqslant\tilde{\mu}_2^N|v|\quad\forall v\in\tilde{\mathcal{E}}^u_{K(\theta)}\tag{88}$$

*and*

$$\left|DF\circ K\circ T_\omega^{N-1}(\theta)\times\cdots\times DF\circ K(\theta)v\right|\leqslant\tilde{\mu}_3^N|v|,$$
$$\left|DF^{-1}\circ K\circ T_\omega^{-(N-1)}(\theta)\times\cdots\times DF^{-1}\circ K(\theta)v\right|\leqslant\tilde{\mu}_3^N|v|\quad\forall v\in\tilde{\mathcal{E}}^c_{(\theta)}.\tag{89}$$

*Assume that $\delta<\delta_0$, where $\delta_0$ is an expression depending on $N$, $\|DF\circ K\|_\rho$, $\|DF^{-1}\circ K\|_\rho$, $\|\Pi^{c,s,u}\|_\rho$.*

*Then, there is an analytic splitting*

$$T_{K(\theta)}\mathcal{M}=\mathcal{E}^s_{K(\theta)}\oplus\mathcal{E}^u_{K(\theta)}\oplus\mathcal{E}^c_{K(\theta)}$$

*invariant under the co-cycle $DF\circ K$ over $T_\omega$, which satisfies the characterization of hyperbolic splittings* (11)–(13)*.*

*The splitting above is unique among the splittings in a neighborhood of the original splitting of size $\delta_0$ measured in* $\mathrm{dist}_\rho$.

*Furthermore, we have that*

$$\mathrm{dist}_\rho\left(\mathcal{E}^{s,u,c}_{K(\theta)},\tilde{\mathcal{E}}^{s,u,c}_{K(\theta)}\right)\leqslant C\delta,$$
$$|\mu_{1,2,3}-\tilde{\mu}_{1,2,3}|\leqslant C\delta,\tag{90}$$

*where $C$ depends on the same quantities as $\delta_0$ does.*

The previous result is applicable to all co-cycles over $T_\omega$. It is important that the base is a rotation. As it is well known in the general theory of hyperbolic systems, if the base of the co-cycle had non-zero Lyapunov exponents, we expect that the invariant splittings are only finitely differentiable and not analytic even if the co-cycle and the base map are analytic. Some explicit examples are available in [12].

In the statement of Proposition 5.2, for typographical simplicity, we are assuming that the phase space is an Euclidean manifold so that we can compute the product $DF\circ K(\theta+\omega)DF\circ K(\theta)$ and consider $DF\circ K$ as a co-cycle over $T_\omega$. In case that the phase space is not an Euclidean manifold, the co-cycle is $S^{F\circ K(\theta)}_{K(\theta+\omega)}DF\circ K$, where $S$ is the connector introduced in Definition 3.9. This can be done provided that $\mathrm{dist}(F\circ K(\theta),K(\theta+\omega))$ is small enough so that the connectors can be defined. The proof of Proposition 5.2 does not require any change beyond that to work in non-Euclidean manifolds.

Note that Proposition 5.2 implies immediately the persistence of invariant bundles under perturbations of the co-cycle. Given a co-cycle, its invariant bundles are approximately invariant under the perturbed co-cycle. The approximately invariant co-cycles can be obtained in many different ways, for example through numerical computations or through formal expansions. For the numerical applications we refer to [33]. We also mention that [47] computes Lindstedt series expansions for quasi-periodic solutions in center manifolds for problems in celestial mechanics. These solutions are whiskered solutions in the full space and can be validated applying the results of this paper.

**Remark 5.3.** Notice that the statements of the hyperbolicity conditions in Proposition 5.2 do not involve any constant $C_h$ as in (11)–(13), but on the other hand, we include an $N$. From the point of view of mathematical theorems, both formulations are equivalent if we consider (11)–(13) for fixed $n=N$. Note that if $\tilde{\mu}>\mu$ and $N$ are such that $C_h(\mu/\tilde{\mu})^N<1$, the conditions in (11) imply those in Proposition 5.2. The converse is trivial.

We note that the constants $C_h$ depend on the norm used in the space. Indeed, in theoretical applications, it is convenient to choose a norm such that $C_h=1$. Equivalently, one can choose a norm such that $N = 1$. This indeed simplifies the notation. We have chosen not to take advantage of this simplification since the adapted norm is not commonly used in numerical applications.

**Remark 5.4.** Another application of Proposition 5.2 that we will not develop here, is a bootstrap of regularity. If an invariant splitting is continuous, smoothing it, we obtain an approximately invariant analytic one and, applying Proposition 5.2, we obtain an analytic invariant splitting which has to coincide with the original one. See [34,43].

**Remark 5.5.** With a view to the applications in [22], we note that the arguments in the proof of Proposition 5.2 are rather soft (contraction mapping principle and such). Hence, they go through without changes when the bundles are Banach bundles.

**Remark 5.6.** Notice that the proof of the existence of invariant subbundles given the approximately invariant ones is done one subbundle at a time. Hence, if we have two invariant subbundles (this is the situation considered in Proposition 4.16), the argument in the proof of Proposition 5.2 above leaves unchanged the invariant subspaces. Hence, the hyperbolicity constants $\mu_1, \mu_2, \mu_3$ and $C_h$ in these spaces are unaltered. On the other hand, the projections on the invariant subspaces are altered because the projections depend on the splitting. The change of one of the subbundles changes all the projections. Of course, the change of the projections can be estimated by the change of the spaces, which is in turn estimated by the error in the invariance equation.

The main application of Proposition 5.2 in this paper is the following result, Proposition 5.7, which estimates the change in the hyperbolicity hypotheses in an iterative step.

**Proposition 5.7.** *Assume that $\|K - \tilde{K}\|_\rho$ is small enough and hypotheses of Proposition* 5.1 *apply. Then there exists an analytic invariant splitting for $DF \circ \tilde{K}$.*

*Furthermore, there exists a constant $C > 0$ such that we have the estimates*

$$\left\| \Pi^{s,c,u}_{\tilde{K}(\theta)} - \Pi^{s,c,u}_{K(\theta)} \right\|_\rho \leqslant C \|\tilde{K} - K\|_\rho, \tag{91}$$

$$|\tilde{\mu}_i - \mu_i| \leqslant C \|\tilde{K} - K\|_\rho, \quad i = 1, 2, 3, \tag{92}$$

$$\tilde{C}_h = C_h. \tag{93}$$

**Proof of Proposition 5.2.** The proof we present is very similar to the proof in [33]. The ideas are very similar to the standard proof of the persistence of invariant splittings in [40,56,58] but we present them in an *a posteriori* format, obtaining very quantitative estimates and we take advantage of the fact that the motion in the base is a rotation. This requires only some minor rearrangements of the argument in the above references.

We will denote

$$\begin{aligned} \mathcal{E}^1_{K(\theta)} &= \tilde{\mathcal{E}}^s_{K(\theta)}, \\ \mathcal{E}^2_{K(\theta)} &= \tilde{\mathcal{E}}^c_{K(\theta)} \oplus \tilde{\mathcal{E}}^u_{K(\theta)}. \end{aligned} \tag{94}$$

We clearly have

$$T_{K(\theta)}\mathcal{M} = \mathcal{E}^1_{K(\theta)} \oplus \mathcal{E}^2_{K(\theta)} \tag{95}$$

and the splitting (95) is almost invariant under $DF \circ K$.

We consider the matrix of $DF(K(\theta))$ with respect to the splitting (95):

$$DF\big(K(\theta)\big) = \begin{pmatrix} a_{11}(\theta) & a_{12}(\theta) \\ a_{21}(\theta) & a_{22}(\theta) \end{pmatrix}.$$

The almost invariance of the splitting implies that $\|a_{12}\|_\rho, \|a_{21}\|_\rho \leqslant C\eta$.

We will construct the invariant subspaces corresponding to this splitting as graphs of linear functions $u^1(\theta) : \mathcal{E}^1_\theta \to \mathcal{E}^2_\theta$ and $u^2(\theta) : \mathcal{E}^2_\theta \to \mathcal{E}^1_\theta$.

Computing the image of the point $(x, u^1(\theta)x)$ (resp. $(u^2(\theta)y, y)$) and imposing that the images are in the graph of $u^1(\theta+\omega)$ (resp. $u^2(\theta+\omega)$), we obtain that the graphs of $u^1, u^2$ are invariant if and only if $u^1, u^2$ satisfy

$$u^1(\theta+\omega)\big(a_{11}(\theta) + a_{12}(\theta)u^1(\theta)\big) = a_{21}(\theta) + a_{22}(\theta)u^1(\theta), \tag{96}$$

$$a_{11}(\theta)u^2(\theta) + a_{12}(\theta) = u^2(\theta+\omega)\big(a_{21}(\theta)u^2(\theta) + a_{22}(\theta)\big). \tag{97}$$

As can be seen by elementary algebraic manipulations, Eqs. (96) and (97) are equivalent to

$$u^1(\theta) = a_{22}^{-1}(\theta)\big(u^1(\theta+\omega)\big(a_{11}(\theta) + a_{12}(\theta)u^1(\theta)\big) - a_{21}(\theta)\big), \tag{98}$$

$$u^2(\theta+\omega) = \big(a_{11}(\theta)u^2(\theta) + a_{12}(\theta)\big)\big(a_{22}(\theta) + a_{21}(\theta)u^2(\theta)\big)^{-1}. \tag{99}$$

We see that $u^1, u^2$ are fixed points of the operators $\mathcal{T}^1, \mathcal{T}^2$ which are defined as the right-hand side of Eq. (98) and the right-hand side of Eq. (99) shifted by $-\omega$, respectively:

$$\mathcal{T}^1\big[u^1\big](\theta) = a_{22}^{-1}(\theta)\big(u^1(\theta+\omega)\big(a_{11}(\theta) + a_{12}(\theta)u^1(\theta)\big) - a_{21}(\theta)\big),$$
$$\mathcal{T}^2\big[u^2\big](\theta) = \big(a_{11}(\theta-\omega)u^2(\theta-\omega) + a_{12}(\theta-\omega)\big)\big(a_{22}(\theta-\omega) + a_{21}(\theta-\omega)u^2(\theta-\omega)\big)^{-1}.$$

Now we concentrate on the operator $\mathcal{T}^1$. We introduce the space $\mathcal{S} = \mathcal{A}(D_\rho, \mathcal{L}_1)$ of analytic sections from $D_\rho$ to the unit bundle of linear operators from $\mathcal{E}^1_{K(\theta)}$ into $\mathcal{E}^2_{K(\theta)}$, i.e. the space of analytic maps $u$ such that $u(\theta) : \mathcal{E}^1_{K(\theta)} \to \mathcal{E}^2_{K(\theta)}$ is linear and $\|u(\theta)\| \leqslant 1$. Endowed with $\|u\|_{\mathcal{S}} = \sup_{\theta \in D_\rho} \|u(\theta)\|$, $\mathcal{S}$ is a Banach space. Moreover $\mathcal{S}$ satisfies Banach algebra properties under the natural multiplications.

We note that if $\eta$ is small enough and consequently $\|a_{12}\|$, $\|a_{21}\|$ are small, a reasonable linear approximation of $\mathcal{T}^1$ is (obtained by eliminating all the terms that contain $a_{12}, a_{21}$)

$$\mathcal{T}^1_0\big[u^1\big](\theta) := a_{22}^{-1}(\theta)u^1(\theta+\omega)a_{11}(\theta).$$

An elementary computation gives

$$\big(\mathcal{T}^1_0\big)^N\big[u^1\big](\theta) = a_{22}^{-1}(\theta)\cdots a_{22}^{-1}\big(\theta + (N-1)\omega\big)u^1(\theta + N\omega)a_{11}\big(\theta + (N-1)\omega\big)\cdots a_{11}(\theta).$$

Using the fact that $\mathcal{T}^1$ is a quadratic polynomial operator, by performing algebraic manipulations we obtain

$$\max_{\|u^1\|_\rho \leqslant \eta} \big\| \big(\mathcal{T}^1\big)^N\big[u^1\big] - \big(\mathcal{T}^1_0\big)^N\big[u^1\big] \big\|_\rho \leqslant C\eta \tag{100}$$

and

$$\mathrm{Lip}_{B_\eta}\big(\big(\mathcal{T}^1\big)^N - \big(\mathcal{T}^1_0\big)^N\big) \leqslant C\eta, \tag{101}$$

where $C$ depends on $N$.

Note that, by assumptions (87)–(89) we have that if $\eta$ is small enough, $(\mathcal{T}^1)^N$ maps $\mathcal{S}$ into $\mathcal{S}$.

Then using (87)–(89) together with the previous estimates we have that $(\mathcal{T}_0^1)^N$ is a contraction from $\mathcal{S}$ to $\mathcal{S}$.

This implies that also $(\mathcal{T}^1)^N$ is a contraction. It is well known that then $\mathcal{T}^1$ has a unique fixed point $u$ in $\mathcal{S}$.

Moreover the analyticity in $\theta \in D_\rho$ is inherited by the fixed point of the contraction. Hence $u$ depends analytically in $\theta$.

Furthermore, we have the standard fixed point estimate

$$\|u\|_{\mathcal{S}} \leqslant \frac{1}{1-\alpha}\left(\left(\mathcal{T}^1\right)^N(0) - 0\right) \leqslant C\eta, \tag{102}$$

where $\alpha = C'\delta$ and the constant $C'$ depends on $N, C_h, \mu_1, \mu_3$. This estimate gives that $d_\rho(\mathcal{E}^s_{K(\theta)}, \tilde{\mathcal{E}}^s_{K(\theta)}) \leqslant C\eta$. Then, since the spaces $\mathcal{E}^s$, $\tilde{\mathcal{E}}^s$ are $C\eta$-close we also have $|\tilde{\mu}_1 - \mu_1| \leqslant C\delta$.

The proof so far, gives us the existence of invariant spaces as in (94). This clearly gives us the existence of the invariant bundle $\mathcal{E}^s_{K(\theta)}$ and the invariant bundle $\mathcal{E}^{cu}_{K(\theta)}$.

We remark that exactly the same proof works if we take the splitting

$$\begin{aligned} \mathcal{E}^1_{K(\theta)} &= \tilde{\mathcal{E}}^s_{K(\theta)} \oplus \tilde{\mathcal{E}}^c_{K(\theta)}, \\ \mathcal{E}^2_{K(\theta)} &= \tilde{\mathcal{E}}^u_{K(\theta)}. \end{aligned} \tag{103}$$

Hence, we also obtain the existence of the bundles $\mathcal{E}^{sc}_{K(\theta)}$ and $\mathcal{E}^u_{K(\theta)}$. The invariant bundle $\mathcal{E}^c_{K(\theta)}$ is obtained as $\mathcal{E}^{cu}_{K(\theta)} \cap \mathcal{E}^{cs}_{K(\theta)}$.

This concludes the proof of Proposition 5.2. □

**Proof of Proposition 5.7.** We just observe that we can take the invariant splittings for $DF \circ K$ as approximately invariant for $DF \circ \tilde{K}$. Using Cauchy estimates, we see that we can take $\delta = C\|\tilde{K} - K\|_\rho$. Therefore, (91) follows from estimating the change of the spaces. The conclusions (92), (93) follow from the observations in Remark 5.3. □

The next Lemma 5.8 provides the perturbation for the remaining non-degenerate conditions. The idea is very simple. The twist condition is just the norm of a matrix obtained by restricting the derivative to the tangent and projecting it on the symplectic conjugate directions to the tangent. Cauchy estimates allows us to estimate easily the changes of these spaces. The estimate of the change of the derivative when we change the embedding is just the mean value theorem.

**Lemma 5.8.** *Assume that the hypotheses of Proposition* 5.1 *hold. If* $\|E_{m-1}\|_{\rho_{m-1}}$ *is small enough, then*:

- *If* $DK_{m-1}^\top DK_{m-1}$ *is invertible with inverse* $N_{m-1}$ *then* $DK_m^\top DK_m$ *is invertible with inverse* $N_m$ *and we have*

$$\|N_m\|_{\rho_m} \leqslant \|N_{m-1}\|_{\rho_{m-1}} + C_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|_{\rho_{m-1}}.$$

- *If* $\operatorname{avg}(A_{m-1})$ *is non-singular then also* $\operatorname{avg}(A_m)$ *is and we have the estimate*

$$\left|\left(\operatorname{avg}(A_m)\right)^{-1}\right| \leqslant \left|\left(\operatorname{avg}(A_{m-1})\right)^{-1}\right| + C'_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|_{\rho_{m-1}}.$$

- *If* $\operatorname{avg}(Q_{m-1})$ *is non-singular then also* $\operatorname{avg}(Q_m)$ *is and we have the estimate*

$$\left|\left(\operatorname{avg}(Q_m)\right)^{-1}\right| \leqslant \left|\left(\operatorname{avg}(Q_{m-1})\right)^{-1}\right| + C''_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|_{\rho_{m-1}}.$$

**Proof.** For the first, we refer the reader to [14, Section 5] since the proof is identical. We turn to the second and third points. The estimates just come from writing $K_m = K_{m-1} + \Delta_{m-1}$, using estimates (25)–(26) and neglecting quadratic error terms at the price of changing the constants. $\square$

*5.3. Convergence of the scheme*

It is by now classical that, under sufficiently strong smallness assumptions, the iterative scheme can be iterated indefinitely and that it converges. Similar arguments can be found in almost any paper in KAM theory, in particular [11,50,51,73,74]. The notation in this paper matches closely that in [14] so that the modifications, at this stage are rather minimal.

Recall that we have identified a set of embeddings in which we can obtain uniform constants in the Newton step, see Proposition 5.1.

In the following Lemma 5.9 we show that, with the choice of domain losses given in (104) if the initial error is small enough, the iterations do not leave the neighborhood where we have uniform estimates and converge to a solution of the problem, which also has hyperbolic splittings.

**Lemma 5.9.** *Using the previous notations, let $C_m$ be the sequence of positive numbers defined above. For a fixed $0 < \delta_0 \leqslant \min(1, \rho_0/12)$ define for $m \geqslant 0$,*

$$\delta_m = \delta_0 2^{-m}. \tag{104}$$

*Denote $\rho_m = \rho_{m-1} - 6\delta_{m-1}$ and $\epsilon_m = \|E_m\|_{\rho_m}$.*

*There exists a constant $C$ depending on $l$, $\nu$, $|F|_{C^2(B_r)}$, $|J|_{C^1(B_r)}$, $\|DK_0\|_{\rho_0}$, $\|N_0\|_{\rho_0}$, $|(\mathrm{avg}(Q_0))^{-1}|$, $|(\mathrm{avg}(A_0))^{-1}|$, $\|\Pi^{s,c,u}_{K_0(\theta)}\|_{\rho_0}$, $\|G\|_{\rho_0}$ such that if the error $\epsilon_0$ satisfies the following inequalities*

$$C 2^{4\nu} \kappa^4 \delta_0^{-4\nu} \epsilon_0 < 1/2$$

*and*

$$C\left(1 + \frac{2^{4\nu}}{2^{2\nu} - 1}\right) \kappa^2 \delta_0^{-2\nu} \epsilon_0 < r,$$

*then the modified Newton step can be iterated indefinitely and we obtain that $K_m$ converges to a map $K_\infty \in \mathcal{A}_{\rho_0 - 6\delta_0}$ which satisfies the non-degeneracy conditions, in particular, it is hyperbolic, and*

$$F \circ K_\infty = K_\infty \circ T_\omega.$$

*Moreover, there exists a constant $D > 0$ depending on $l$, $\nu$, $|F|_{C^2(B_r)}$, $|J|_{C^1(B_r)}$, $\|DK_0\|_{\rho_0}$, $\|N_0\|_{\rho_0}$, $|(\mathrm{avg}(Q_0))^{-1}|$, $|(\mathrm{avg}(A_0))^{-1}|$, $\|\Pi^{s,c,u}_{K_0(\theta)}\|_{\rho_0}$ and $\|G\|_{\rho_0}$ such that*

$$\|K_\infty - K_0\|_{\rho_0 - 6\delta_0} \leqslant D \kappa^2 \delta_0^{-2\nu} \|E_0\|_{\rho_0}.$$

**Remark 5.10.** We note again that by the vanishing Lemma 4.9, the sequence $\{\lambda_n\}_{n \geqslant 0}$ converges to 0 as $n$ goes to $+\infty$.

**Proof of Lemma 5.9.** As mentioned in the introduction of the section, the argument is quite standard.

To ensure that we can perform steps with the estimates in Proposition 5.1, we just need to verify that we do not leave the neighborhood of $K_0$ given by (83) and that we satisfy the bounds (84).

We note that, in a concise notation, Proposition 5.1 leads to the bounds

$$\epsilon_m \leqslant C \kappa^4 \delta_{m-1}^{-4\nu} \epsilon_{m-1}^2.$$

With the choice $\delta_m = \delta_0 2^{-m}$, we see that if we can perform $m$ steps, we have

$$\begin{aligned}\epsilon_m &\leqslant C\kappa^4\delta_0^{-4\nu}2^{4\nu(m-1)}\epsilon_{m-1}^2 \leqslant \left(C\kappa^4\delta_0^{-4\nu}\right)^{1+2}2^{4\nu[(m-1)+2(m-2)]}\epsilon_{m-2}^{2^2}\\ &\leqslant \left(C\kappa^4\delta_0^{-4\nu}\right)^{1+2+\cdots+2^{m-1}}2^{4\nu[(m-1)+2(m-2)+\cdots+2^{m-2}]}\epsilon_0^{2^m}\\ &\leqslant \left(C2^{4\nu}\kappa^4\delta_0^{-4\nu}\epsilon_0\right)^{2^m},\end{aligned}$$

for $m \geqslant 1$, where we have used that

$$(m-1)+2(m-2)+\cdots+2^{m-2} = 2^{m-2}\left[(m-1)2^{-(m-2)}+(m-2)2^{-(m-3)}+\cdots+1\right] \leqslant 2^m.$$

We see that if $\epsilon_0$ is small enough, then, $\epsilon_m\delta_m^{-4\nu}$ is so small than the conditions (84) are true for the next step. Indeed, we note that the smallness conditions that we need to impose in $\epsilon_0$ are independent of $m$.

Furthermore, we also observe that we also have $K_m - K_0 = \sum_{i=0}^{m-1}\Delta_i$. Hence

$$\|K_m - K_0\|_{\rho_m} \leqslant \sum_{i=0}^{m-1}\|\Delta_i\|_{\rho_i} \leqslant \sum_{i=0}^{m-1} C\kappa^2\left(C\kappa^4\delta_0^{-4\nu}\epsilon_0\right)^{2^i}\delta_0^{-2\nu}2^{2i\nu}.$$

We note that, by taking $\epsilon_0$ small enough we can make the right-hand side of the last formula as small as desired uniformly in $m$. In particular, by taking $\epsilon_0$ small enough we can ensure the assumption (83) for all $m$.

Therefore, if we assume that $\epsilon_0$ small enough, we can ensure that we can repeat the iterative step infinitely often and that the iteration never leaves the neighborhood identified in (83).

We also note that we have

$$\begin{aligned}\sum_{i=0}^{\infty}\|K_{i+1}-K_i\|_{\rho_\infty} = \sum_{i=0}^{\infty}\|\Delta_i\|_{\rho_\infty} &\leqslant \sum_{i=0}^{\infty}\|\Delta_i\|_{\rho_i}\\ &\leqslant C\kappa^2\delta_0^{-2\nu}\epsilon_0\left(1+\sum_{i=1}^{\infty}C\kappa^2\left(C2^{4\nu}\kappa^4\delta_0^{-4\nu}\epsilon_0\right)^{2^i}\delta_0^{-2\nu}2^{2i\nu}\epsilon_0^{-1}\right)\\ &\leqslant D\kappa^2\delta_0^{-2\nu}\epsilon_0.\end{aligned}$$

The absolute convergence of the above series shows that $K_m$ converge to a limit and the last bound establishes the conclusion (19). We note that since we had assumed (83), we have that $K_\infty$ admits a hyperbolic splitting. Since the change in the hyperbolic splittings is bounded by the change of the embedding (see Proposition 5.7), we see that the hyperbolic splittings also converge to the limiting one. □

## 6. Proof of the local uniqueness theorem

In this section, we prove Theorem 3.14. We closely follow the proof in [14]. Similar results are more or less implicit in the treatment of whiskered tori in [74]. For fully dimensional tori local uniqueness results appear in [51,63,69]. As we have argued before, local uniqueness results allow us to deduce results for flows from results for maps.

The proof of Theorem 3.14 is based on showing that the operator $D\mathcal{F}_\omega(K)$ has an approximate left inverse (as in [73]). Notice first that the composition on the right by every translation of a solution of (6) is also a solution. Therefore, one cannot expect a general uniqueness result. Moreover, the second statement in Lemma 4.2 and the calculation on the hyperbolic directions show that, roughly speaking, two solutions of the linearized equation differ by their average. Moreover this difference is in the direction of the tangent space of the torus.

The idea behind the local uniqueness result is to prove that one can transfer the difference of the averages of two solutions to a difference of phase between the two solutions.

Now we assume that the embeddings $K_1$ and $K_2$ satisfy the hypotheses in Theorem 3.14, in particular $K_1$ and $K_2$ are solutions of (6), or (7) with $\lambda = 0$. If $\tau \neq 0$ we write $K_1$ for $K_1 \circ T_\tau$ which is also a solution. Therefore $\mathcal{F}_\omega(0, K_1) = \mathcal{F}_\omega(0, K_2) = 0$. By Taylor's theorem we can write

$$0 = \mathcal{F}_\omega(0, K_1) - \mathcal{F}_\omega(0, K_2) = D_{\lambda,K}\mathcal{F}_\omega(0, K_2)(0, K_1 - K_2) + \mathcal{R}(0, 0, K_1, K_2), \tag{105}$$

where

$$\mathcal{R}(0, 0, K_1, K_2) = \frac{1}{2}\int_0^1 D^2 F\big(K_2 + t(K_1 - K_2)\big)(K_1 - K_2)^2\, dt.$$

Then, there exists $C > 0$ such that

$$\big\|\mathcal{R}(0, 0, K_1, K_2)\big\|_\rho \leqslant C\|K_1 - K_2\|_\rho^2.$$

Hence we end up with the following linearized equation

$$D_{\lambda,K}\mathcal{F}_\omega(0, K_2)(0, K_1 - K_2) = -\mathcal{R}(0, 0, K_1, K_2). \tag{106}$$

We denote $\Delta = K_1 - K_2$.

Projecting (106) on the center subspace with $\Pi^c_{K_2(\theta+\omega)}$, writing $\Delta^c(\theta) = \Pi^c_{K_2(\theta)}\Delta(\theta)$ and making the change of function $\Delta^c(\theta) = \tilde{M}(\theta)W(\theta)$, where $\tilde{M}$ is defined in (42) with $K = K_2$, we obtain

$$DF\big(K_2(\theta)\big)\tilde{M}(\theta)W(\theta) - \tilde{M}(\theta+\omega)W(\theta+\omega) = -\Pi^c_{K_2(\theta+\omega)}\mathcal{R}(0, 0, K_1, K_2)(\theta). \tag{107}$$

We note that since $K_2$ is an exact solution, range $\tilde{M}(\theta)$ coincides with $\mathcal{E}^c_{K(\theta)}$. See Section 4.1.3.

Applying the property $DF(K_2(\theta))\tilde{M}(\theta) = \tilde{M}(\theta+\omega)\mathcal{S}(\theta)$ for solutions of (6), multiplying both sides by $[\tilde{M}^\top J(K_2)](\theta+\omega)$ and using that $\tilde{M}^\top J(K_2)\tilde{M}$ is invertible we get

$$\begin{aligned}&\mathcal{S}(\theta)W(\theta) - W(\theta+\omega)\\ &\quad= -\big[\big(\tilde{M}^\top J(K_2)\tilde{M}\big)^{-1}\tilde{M}^\top J(K_2)\big](\theta+\omega)\Pi^c_{K_2(\theta+\omega)}\mathcal{R}(0, 0, K_1, K_2)(\theta).\end{aligned}$$

Since $W$ solves the previous equation, we get bounds for it using the methods in Section 4.2.4. We write $W = (W_1, W_2)$. Since $\mathcal{S}$ is triangular we begin by looking for $W_2$. We search it in the form $W_2 = W_2^\perp + \operatorname{avg}(W_2)$. We have $\|W_2^\perp\|_{\rho-\delta} \leqslant C\kappa\delta^{-\nu}\|K_1 - K_2\|_\rho^2$. For $W_1$ we have

$$\begin{aligned}&W_1(\theta) - W_1(\theta+\omega)\\ &\quad= T_2(\theta)\big(\Pi^c_{K_2(\theta+\omega)}\mathcal{R}(0, 0, K_1, K_2)\big)_1(\theta) - A(\theta)W_2^\perp(\theta) - A(\theta)\operatorname{avg}(W_2),\end{aligned} \tag{108}$$

where $T_2 = N_2^\top DK_2^\top J(K_2)^{-\top}[DK_2N_2DK_2^\top - \mathrm{Id}]J(K_2)$ and $N_2 = DK_2^\top DK_2$.

The condition the right-hand side of (108) to have zero average gives $|\operatorname{avg}(W_2)| \leqslant C\kappa\delta^{-\nu}\|K_1 - K_2\|_\rho^2$. Then

$$\big\|W_1 - \operatorname{avg}(W_1)\big\|_{\rho-2\delta} \leqslant C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_\rho^2$$

but $\mathrm{avg}(W_1)$ is free. Then

$$\left\| \Delta^c - \left(\mathrm{avg}\left(\Delta^c\right)_1, 0\right)^\top \right\|_{\rho-2\delta} \leqslant C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_\rho^2.$$

The next step is done in the same way as in [14]. We quote Lemma 14 of that reference using our notation. It is basically an application of the standard implicit function theorem.

**Lemma 6.1.** *There exists a constant $C$ such that if $C\|K_1 - K_2\|_\rho \leqslant 1$ then there exists an initial phase $\tau_1 \in \{\tau \in \mathbb{R}^l \mid |\tau| < \|K_1 - K_2\|_\rho\}$ such that*

$$\mathrm{avg}\left(T_2(\theta)\Pi^c_{K_2(\theta)}(K_1 \circ T_{\tau_1} - K_2)(\theta)\right) = 0.$$

The proof is based on an application of implicit function theorem in $\mathbb{R}^l$.

As a consequence of Lemma 6.1, if $\tau_1$ is as in the statement, then $K \circ T_{\tau_1}$ is a solution of (6) such that if

$$W = \left[\tilde{M}^\top J(K_2)\tilde{M}\right](\theta+\omega)^{-1}\left[\tilde{M}^\top J(K_2)\right](\theta+\omega)\Pi^c_{K_2(\theta)}(K_1 \circ T_{\tau_1} - K_2),$$

for all $\delta \in (0, \rho/2)$ and we have the estimate

$$\|W\|_{\rho-2\delta} < C\kappa^2\delta^{-2\nu}\|\mathcal{R}\|_\rho^2 \leqslant C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_\rho^2.$$

This leads to on the center subspace

$$\left\|\Pi^c_{K_2(\theta)}(K_1 \circ T_{\tau_1} - K_2)\right\|_{\rho-2\delta} \leqslant C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_\rho^2.$$

Furthermore, as in Section 4.3, taking projections on the hyperbolic subspace, we have that $\Delta^h = \Pi^h_{K_2(\theta)}(K_1 - K_2)$ satisfies the estimate

$$\left\|\Delta^h\right\|_{\rho-2\delta} < C\|\mathcal{R}\|_\rho.$$

All in all, we have proven the estimate for $K_1 \circ T_{\tau_1} - K_2$ (up to a change in the original constants)

$$\|K_1 \circ T_{\tau_1} - K_2\|_{\rho-2\delta} \leqslant C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_\rho^2.$$

We are now in position to carry out an argument very similar to the one used in Section 5.3. We can take a sequence $\{\tau_m\}_{m\geqslant 1}$ such that $|\tau_1| \leqslant \|K_1 - K_2\|_\rho$ and

$$|\tau_m - \tau_{m-1}| \leqslant \|K_1 \circ T_{\tau_{m-1}} - K_2\|_{\rho_{m-1}}, \quad m \geqslant 2,$$

and

$$\|K_1 \circ T_{\tau_m} - K_2\|_{\rho_m} \leqslant C\kappa^2\delta_m^{-2\nu}\|K_1 \circ T_{\tau_{m-1}} - K_2\|^2_{\rho_{m-1}},$$

where $\delta_1 = \rho/4$, $\delta_{m+1} = \delta_m/2$ for $m \geqslant 1$ and $\rho_0 = \rho$, $\rho_m = \rho_0 - \sum_{k=1}^m \delta_k$ for $m \geqslant 1$. By an induction argument we end up with

$$\|K_1 \circ T_{\tau_m} - K_2\|_{\rho_m} \leqslant \left(C\kappa^2\delta_1^{-2\nu}2^{2\nu}\|K_1 - K_2\|_{\rho_0}\right)^{2^m} 2^{-2\nu m}.$$

Therefore, under the smallness assumptions on $\|K_1 - K_2\|_{\rho_0}$, the sequence $\{\tau_m\}_{m\geqslant 1}$ converges and one gets

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

Since both $K_1 \circ T_{\tau_\infty}$ and $K_2$ are analytic in $D_\rho$ and coincide in $D_{\rho/2}$ we obtain the result.

## 7. Applications

In this section, we collect several consequences of our main theorem. We note that these consequences follow mainly from the fact that we have formulated the theorem in *a posteriori* style without reference to an integrable system.

### *7.1. Lipschitz dependence with respect to the frequency. Estimates of the measure occupied by the tori*

The basic idea is that if we have an embedding $K$ that solves the equation for one frequency, then it solves approximately the equation for a nearby frequency. Then, applying Theorem 3.11, there should be a solution for a new frequency which is close to the original one. Performing the argument with care, we see that this implies Lipschitz dependence of the solution on the frequency. Similar ideas were indicated in [73]. We remark that this Lipschitz dependence leads to estimates on the measure occupied by the tori in the perturbative case. We concentrate on the case of maps since, as we have shown, it implies the corresponding result for flows.

We assume that $F$ is defined and analytic in a sufficiently big complex domain of an Euclidean manifold $\mathcal{M}$. We consider $\omega \in D(\kappa, \nu)$ with $\kappa$ and $\nu$ fixed and we suppose that $K_\omega \in \mathcal{A}_\rho$ satisfies (6) and is non-degenerate in the sense of Definition 3.4. For all $\omega' \in D(\kappa, \nu)$ we have

$$F \circ K_\omega - K_\omega \circ T_{\omega'} = K_\omega \circ T_\omega - K_\omega \circ T_{\omega'}$$

and therefore, applying the mean value theorem and Cauchy estimates, we have

$$\|F \circ K_\omega - K_\omega \circ T_{\omega'}\|_{\rho-\delta} \leqslant C\delta^{-1}\|K_\omega\|_\rho |\omega - \omega'| \tag{109}$$

for all $\delta \in (0, \rho)$. If $|\omega - \omega'|$ is small enough, namely $C\kappa^4\delta^{-4\nu-1}\|K_\omega\|_\rho|\omega - \omega'| < 1$, applying Theorem 3.11 we obtain that there is an embedding $K_{\omega'}$ satisfying (6) with the frequency $\omega'$. Furthermore, taking the value of $\delta$ in Theorem 3.11 appropriately, one gets

$$\|K_\omega - K_{\omega'}\|_{(\rho-\delta)/2} \leqslant C\kappa^2(\rho - \delta)^{-2\nu}\delta^{-1}\|K_\omega\|_\rho|\omega - \omega'|.$$

We note that, in the domain of applicability of the previous argument, the Lipschitz constant is uniform since we are assuming that $\kappa$ and $\nu$ are fixed. Since the set of uniformly Diophantine vectors $D(\kappa, \nu)$ is locally compact, we can cover a bounded subset of $D(\kappa, \nu)$ with a finite number of balls in which the previous argument applies and therefore in this set we get Lipschitz dependence on $\omega$.

Moreover, the frequency of $K_\omega$ is given by the formula (up to some multiplicative constant)

$$\omega = \Pi_\varphi \int_{\mathbb{T}^l} \big(\tilde{F} \circ \tilde{K}_\omega(\theta) - \tilde{K}_\omega(\theta)\big)\, d\theta, \tag{110}$$

where $\tilde{F}$ and $\tilde{K}$ are the lifts of $F$ and $K$ to the universal cover of $\mathcal{M}$ and $\Pi_\varphi$ is the projection over the lift of the angle variables.

Thanks to formula (110), it is straightforward to see that the map $K_\omega \mapsto \omega$ is Lipschitz. Hence, we conclude that the map $\omega \mapsto K_\omega$ is bi-Lipschitz from the set of Diophantine vectors with fixed

Diophantine constants. Since the set $D(\kappa,\nu)$ has positive $l$-dimensional measure, we conclude that the set of tori also has $2l$-dimensional measure, i.e.

$$\mathcal{H}^{2l}\Bigl(\bigcup_{\omega\in D(\kappa,\nu)} K_\omega\bigl(\mathbb{T}^l\bigr)\Bigr) > 0,$$

where $\mathcal{H}^{2l}$ stands for the Hausdorff measure.

### 7.2. Analyticity with respect to parameters

The proof of the existence of tori associated to a fixed frequency $\omega$ presented here leads to analyticity in the dependence with respect to parameters. Later, we will see that this leads to analyticity properties of some series expansions, such as Lindstedt series. The argument is already contained in [52]. However the argument presented here is somewhat simpler than the one presented in that reference.

Given a family of functions $F_\eta$ and a family of approximate solutions $K_\eta$ both depending analytically on parameters $\eta \in U \subset \mathbb{C}^p$ $(p \geqslant 1)$ and continuous in $\overline{U}$, we see that the assumptions of Theorem 3.11 are satisfied uniformly in $\eta \in \overline{U}$. Consequently, there is a true solution nearby which also depends analytically on the parameters $\eta$.

The proof is very simple; we just observe that the iterative step is analytic for $\eta \in U$ (resp. continuous for $\eta \in \overline{U}$) if the family and the error are.

The procedure and the estimates for one step of the iterative procedure are stated in Lemma 4.2 and in a more detailed way in the statements and proofs of Propositions 4.19 and 4.20. We just call attention to the fact that the correction applied at each step of the application of Proposition 4.19 relies on some explicit algebraic formulas—involving derivatives with respect to $\theta$—and to use the solution of some small divisor equations. Note that the solution of the small divisor equations is obtained applying a linear operator which is independent of $\eta$. Also the method of obtaining the projection on the hyperbolic directions done in Section 4.3 and summarized in Proposition 4.20 preserves the analyticity on parameters since $\Delta^{s,u}$ are obtained as sums of uniformly convergent series.

Clearly, the analyticity properties with respect to $\eta$ are preserved by all these steps. Hence, the corrections applied in one step of the iterative Lemma 4.2 depend analytically on parameters when the error does. Of course, the error depends analytically on $\eta \in U$ (resp. continuously on $\eta \in \overline{U}$) if the approximate solution at the start of Lemma 4.17 depends analytically on the parameters, since to compute the error from the approximate solution, we just have to compose with the function $F_\eta$ and translate. Therefore, we conclude that the application of Proposition 4.19 preserves the analyticity properties with respect to parameters.

Hence, in all steps of the iterative process used in the proof of Theorem 3.11, the functions $\{K_m\}_{m\in\mathbb{N}}$ depend analytically on parameters ranging on the open set $U$ (and continuously on the boundary). Of course, in the iterative step, we decrease the analyticity domain in the variables $\theta$, but not the analyticity domain in $\eta$.

We also observe that in the proof of Proposition 4.19, the estimates on the correction applied at each step depend only on the sizes of the error and the non-degeneracy conditions. Also, we observe that the estimates on the change of the non-degeneracy conditions are uniform on the size of the corrections. In particular, if we assume that the smallness and non-degeneracy conditions hold uniformly for $\eta \in \overline{U}$, we can apply Lemma 5.9 to obtain that there is a sequence of analytic functions in $\theta, \eta$ converging uniformly for $\theta \in D_{\rho_\infty}$ and $\eta \in \overline{U}$.

In the paper [15] there is an alternative point of view for results with parameters. One can apply an abstract KAM implicit function theorem as in [73] to spaces of analytic functions in other Banach spaces. These kind of arguments were used to deal with rather degenerate problems. A more detailed study of KAM theorems with parameters appears in [72].

### 7.3. Small twist theorems and small hyperbolicity theorems

Small twist theorems have been introduced in [44,49] to deal with problems in celestial mechanics. The idea of small twist (and small hyperbolicity) theorems is to give conditions that ensure the convergence of the Newton-type method even if the twist is close to be degenerate. It goes through a more precise analysis of the constants involved in the Newton scheme.

We refer to [44,49,54,55] for applications of small twist theorems to celestial mechanics and to the stability of oscillators.

The goal of this section is to provide, as a corollary of the proof of our main Theorem 3.11, a small twist and small hyperbolicity result.

By examining carefully the proof involved in the iterative step in the KAM method (see Section 4.2.4), we get the following proposition.

**Proposition 7.1.**

- *There exist two positive numbers $\alpha$, $\beta$ such that the constant $C$ in Eq.* (74) (*which here will be denoted $C^c$*) *depending on $l$, $\kappa$, $\nu$, $|F|_{C^2(B_r)}$, $\|DK\|_\rho$, $\|N\|_\rho$, $|(\mathrm{avg}(A))^{-1}|$, $|(\mathrm{avg}(Q))^{-1}|$, $\|\Pi^c_{K(\theta)}\|_\rho$ and $\|G\|_\rho$ can be estimated by*

$$C^c \leqslant \left\|\Pi^c_{K_0(\theta)}\right\|_\rho \max\left(1, \|DK\|_\rho\right)^\alpha \max\left(1, \|N\|_\rho\right)^\beta \left(\left|\left(\mathrm{avg}(A)\right)^{-1}\right| + \left|\left(\mathrm{avg}(Q)\right)^{-1}\right|\right), \tag{111}$$

  *where $A$, $Q$ are defined in* (16), (15) *respectively.*
- *The constant $C$ in Eq.* (80) (*which here will be denoted $C^h$*) *depending on the hyperbolicity constant $\mu_1$* (*resp. $\mu_2$*)*, the norm of the projection $\|\Pi^s_{K(\theta)}\|_\rho$* (*resp. $\|\Pi^u_{K(\theta)}\|_\rho$*) *and $\|G\|_\rho$ and the constant $C_h$ involved in* (11) *and* (12) *can be estimated by*

$$C^h \leqslant C_h\left(1 + C^c\right) \max\left(\left\|\Pi^s_{K(\theta)}\right\|_\rho \frac{1}{1-\mu_1}, \left\|\Pi^u_{K(\theta)}\right\|_\rho \frac{1}{1-\mu_2}\right). \tag{112}$$

- *As a consequence of the two above items, the constant $C$ appearing in Theorem* 3.11 (*the one which enters in Eq.* (18)) *can be bounded by*

$$\begin{aligned} &C_h^2 \max\left(\left\|\Pi^s_{K(\theta)}\right\|_\rho \frac{1}{1-\mu_1}, \left\|\Pi^u_{K(\theta)}\right\|_\rho \frac{1}{1-\mu_2}\right)^2 \\ &\quad + \left\|\Pi^c_{K_0(\theta)}\right\|^2_\rho \max\left(1, \|DK\|_\rho\right)^\alpha \max\left(1, \|N\|_\rho\right)^\beta \left(\left|\left(\mathrm{avg}(A)\right)^{-1}\right| + \left|\left(\mathrm{avg}(Q)\right)^{-1}\right|\right)^2. \end{aligned} \tag{113}$$

The argument presented in this paper gives $\alpha = 4$, $\beta = 2$, but there are other variants of the argument which give better values. We have not optimized the bounds.

To prove Proposition 7.1 we note that to find $\Delta^{s,u}$ we just use formula (82) from which the claim follows by estimating the sum using the triangle inequality and using the sum of the geometric series.

The estimates claimed for the constants related to $\Delta^c$ follow by observing that the solution is obtained by applying the following operations: multiplying by the matrices $\tilde{M}$, $\tilde{M}^\top$, multiplying by the matrices $N$, modifying the constants $\Lambda$ and choosing the average of $W_2$. The latter steps are estimated by multiplying by $|(\mathrm{avg}(Q))^{-1}|$ and $|(\mathrm{avg}(A))^{-1}|$.

We also recall that the remainder of the Newton method is estimated by the remainder of the Taylor expansion. Hence, it is estimated by the square of $\|\Delta\|_{\rho-\delta}$.

Let $m \geqslant 0$ be an index for the Newton step and denote $\tilde{C}$ the constant involved in the third item of Proposition 7.1. We have for some $\upsilon > 0$

$$\|E_m\|_{\rho_m} \leqslant \tilde{C}\delta_m^{-\upsilon}\|E_{m-1}\|^2_{\rho_{m-1}},$$

where $\rho_m = \rho_{m-1} - \delta_{m-1}$.

It is standard in KAM theory (see Theorem 3.11) that if

$$\tilde{C}\delta_0^{-2\upsilon}\|E_0\|_{\rho_0} < C(\upsilon) \ll 1,$$

then the Newton method with $\delta_m = 2^{-m}\delta_0$ converges to a solution.

Therefore, even if the twist and the hyperbolicity are close to degenerate so that $\tilde{C}$ is large, if the initial error is small enough, one gets convergence of the scheme.

Small hyperbolicity arises naturally in perturbations of integrable systems. The integrable system, of course, has no hyperbolic behaviour, but an averaged system has some small hyperbolicity. Indeed, similar considerations for periodic orbits happen already in [57, Ch. 74, 79].

The papers [4,5,17,20,42,70] consider perturbations of integrable systems at resonances, where the hyperbolicity is small and the present result can be applied. These papers differ in several important aspects such as the dimension and the topology of the tori considered. The methods are also different.

All of the above papers consider perturbations of quasi-integrable systems $H_0 + \varepsilon H_1$.

The paper [42] shows that, given some appropriate non-degeneracy conditions on the perturbation, it is possible to construct formal series of approximate solutions in powers of $\varepsilon$. Truncating the series up to order $N$, it is shown that the error of the power series can be bounded by $CN^{aN}\varepsilon^N$. Similarly, for some of the solutions, $\Pi^{s,u,c}_{K(\theta)}$ are of order $(\operatorname{Re}\varepsilon)^{-1/2}$, the hyperbolicity constants $(1-\mu_1)^{-1}, (1-\mu_2)^{-1}$ are of order $(\operatorname{Re}\varepsilon)^{-1/2}$ but $(\operatorname{avg}(A))^{-1}$ and $(\operatorname{avg}(Q))^{-1}$ are of order 1. The Diophantine constants can be assumed to be fixed.

If we fix $r > 0$ sufficiently small and consider the set $r < |\varepsilon| < 2r$, we can choose the order of truncation so that the error is less that $C\exp(-b\varepsilon^{-c})$. Then, the small hyperbolicity result applies to show that there are invariant tori, which depend analytically in $\varepsilon$ for $\varepsilon$ such that $r < |\varepsilon| < 2r$ and $\operatorname{Im}\varepsilon \geqslant C\exp(-b(\operatorname{Re}\varepsilon)^c)$. Since $r$ is arbitrary, we obtain that there are hyperbolic invariant tori in a ball except, at most in a wedge around the positive real axis which is exponentially thin.

We refer to [42] for precise conditions on the series so that we can get the perturbation series as above. For the case of two degrees of freedom, the paper [42] considers the existence of elliptic tori. We note that the method of [42] is based on *reducibility*. This paper shows that we do not need to use reducibility for the hyperbolic directions. The study of elliptic directions has experienced very significant progress in the last years, but we will not mention it here.

The paper [17] considers also weakly hyperbolic tori around periodic orbits generated by resonances. Note that the tori considered in [17] are *secondary tori*, that is tori that cannot be deformed to tori with the same frequency in the unperturbed system. Indeed, the tori can be deformed into tori with less angles. Since [17] involves reduction to a center manifold it only concludes that the tori are finite differentiable even if the system is analytic. A result which is improved by using the results provided in this paper in Section 7.5.

### 7.4. *Secondary tori and whiskered tori close to rank-*1 *resonances*

The method described in this paper can accommodate to study secondary tori. Indeed, the development of algorithms which could deal with secondary tori was an important motivation to modify the invariance equation by adding a term containing $\lambda$.

We recall that secondary KAM tori are invariant tori, such that the motion on them are conjugate to an irrational rotation but in contrast to the usual KAM tori which are homotopic to $\mathbb{T}^l \times \{0 \in \mathbb{R}^{2d-l}\}$, the secondary tori are homotopic to $\mathbb{T}^{l-k} \times \{0 \in \mathbb{R}^{2d-l+k}\}$.

The existence of secondary KAM tori is very apparent in numerical explorations. For example, in two-dimensional maps, they are known as *islands*. In two-dimensional maps, islands are quite visible and they may occupy a large measure of the phase space.

Note that secondary tori are not present in the integrable system and their existence is not guaranteed by the standard perturbative KAM theory, which is concerned with the persistence of the invariant tori already present in the integrable system. In contrast, they are generated by the perturbation. The perturbation theory is somewhat unconventional since the unperturbed system does not present the phenomenon. Perturbative proofs of existence of secondary tori are done in [17] and in [8].

In the recent papers [7,8] it is shown that these secondary tori can be used as effective tools to generate diffusion and, in particular, to overcome the *large gap problem* in the study of diffusion. The paper [31] argues heuristically and verifies numerically that, in multiparticle systems that will be considered in a follow-up of this paper, in particular in the celebrated Fermi, Pasta and Ulam [25], the secondary tori occupy a much larger volume of phase space than the primary tori.

The method to construct whiskered tori in [17] was to show that, under explicit conditions on the perturbation, the rational frequencies give rise to periodic orbits, some of which admit center manifolds. Under appropriate non-degeneracy conditions, these center manifolds contain tori which are invariant under the restriction. These invariant tori, are whiskered tori for the full system. They are secondary since the directions corresponding to the center directions can be contracted to a point.

The paper [8] shows that secondary tori are generated by resonances in systems such that the unperturbed system has a two-dimensional normally hyperbolic manifold. The method of proof is to show that, near the resonances, one can approximate the system by a system which is pendulum like. This pendulum has orbits that are rotating. In [8], it is shown that one can consider the real system as a perturbation of the pendulum and, therefore that some of the tori present in the pendulum are also present in the real system. See also [10].

One of the difficulties of the method in [8] is that the action variables near the separatrix are singular. This difficulty is, of course, not a problem for the method developed in the present paper. The method used in [8] to overcome the singularity of the action variables was to perform more averaging steps, which required assuming more regularity of the perturbation. Applying the methods of this paper allows us to reduce the number of derivatives assumed in [8].

As we will discuss in more detail in Section 7.5, by using reduction to center or normally hyperbolic invariant manifolds, one can only prove that the obtained tori are finitely differentiable even if the mapping is analytic. Using the results of this paper, we will show that these tori are actually analytic if the map is.

### 7.5. Bootstrap of regularity of invariant tori

In this section, we show that if an analytic exact symplectic map $F$ admits an invariant torus, with the maximal number of hyperbolic directions permitted by the symplectic structure, of class $C^r$ with $r$ large enough, then the torus is actually analytic. Similar results for Lagrangian tori have been proved in [69].

**Proposition 7.2.** *Let $F : \mathcal{M} \to \mathcal{M}$ be an analytic exact symplectic map. Let $\omega \in D(\kappa, \nu)$ for some $\kappa > 0$ and $\nu \geqslant l$, and $K : \mathbb{T}^l \to \mathcal{M}$ satisfy*:

(1) *$K$ is a solution of the equation $F \circ K - K \circ T_\omega = 0$.*
(2) *$K$ is non-degenerate in the sense of Definition* 3.4.
(3) *$K$ is $C^r$ with*

$$r > 4\nu. \tag{114}$$

*Then $K$ is analytic.*

**Remark 7.3.** One case when Proposition 7.2 is useful is when the tori are produced by a reduction to a center manifold or a normally hyperbolic manifold. These invariant manifolds are only finitely differentiable. Applying the above result, Proposition 7.2, we can conclude that the tori are analytic.

The paper [17] constructs whiskered tori by applying the KAM theorem for Lagrangian tori to the restriction of the system to a center manifold. The papers [6–8] consider tori in a normally hyperbolic manifold. In particular, [8] constructs secondary tori. We conclude that, in the case that the considered system is analytic, Proposition 7.2 shows that the tori are analytic.

The idea of the proof is very simple. We approximate the function $K$ by an analytic one which will be an approximate solution of Eq. (6). Applying our main Theorem 3.11 we will obtain that there is an analytic invariant torus nearby. The smoothness in the assumption enters because the hypotheses of Theorem 3.11 involve the size of the error in a complex strip. The uniqueness result Theorem 3.14 will give that the analytic torus obtained this way coincides with the original one up to a translation in the "angles".

The construction of the analytic approximations could be done in many different ways. For example, truncating the Fourier series of $K$ would do, if one assumes a condition stronger than (114).

As it is well known since [51], a very efficient way of approximating smooth functions by analytic ones is performing a convolution with a suitable kernel.

Following [51,73], we introduce smoothing operators that provide natural ways of approximating smooth functions by analytic ones.

**Definition 7.4.** Let $u : \mathbb{R}^l \to \mathbb{R}$ be a $C^\infty$ even function identically 1 in a neighborhood of the origin and with support contained in the unit ball. Let $\hat{u} : \mathbb{R}^l \to \mathbb{C}$ be the Fourier transform of $u$ and denote by $v$ the holomorphic continuation of $\hat{u}$. For $f \in C^0(\mathbb{R}^l)$ and $t > 0$ we define

$$S_t[f](z) := t^l \int_{\mathbb{R}^l} v\bigl(t(y-z)\bigr) f(y)\, dy. \tag{115}$$

The map $S_t$ defines a linear operator from $C^0(\mathbb{T}^l)$ to the space of analytic maps from $D_\rho$ to $\mathbb{C}$, $\rho > 0$. Moreover, $S_t$ is an analytic smoothing operator in the sense of [51,73] since it satisfies the following proposition (see [73, Lemma 2.1] for a proof). We recall that if $g \in \mathcal{A}_\rho$, $\|g\|_\rho = \sup_{z \in D_\rho} |g(z)|$.

**Proposition 7.5.** *Let $r \in \mathbb{R}^+ \setminus \mathbb{N}$. There exists a constant $\kappa_1 = \kappa_1(l,r)$ such that for all $t \geqslant 1$ and all $f \in C^r(\mathbb{T}^l)$ we have*

(1) $|(S_t - \mathrm{Id})[f]|_{C^0} \leqslant \kappa_1 |f|_{C^r} t^{-r}$,
(2) $\|S_t[f]\|_{t^{-1}} \leqslant \kappa_1 |f|_{C^0}$,
(3) $\|(S_\tau - S_t)[f]\|_{\tau^{-1}} \leqslant \kappa_1 |f|_{C^r} t^{-r}$, *for all $\tau \geqslant t$.*

We note that, since the smoothing operator commutes with derivatives, we also have the following extensions of (1) and (2) for $s \leqslant r$

$$\bigl|(S_t - \mathrm{Id})[f]\bigr|_{C^s} \leqslant \kappa_1 |f|_{C^r} t^{-r+s}, \tag{116}$$

$$\bigl\|D^s S_t[f]\bigr\|_{t^{-1}} \leqslant \kappa_1 |f|_{C^s}. \tag{117}$$

**Proof of Proposition 7.2.** We consider $S_t[K]$, the smoothed version of $K$ with $t \geqslant 1$. Our first goal is to estimate the error in a domain of size $t^{-1}\xi$ with $\xi \in (0,1)$.

We note that, by (2) in Proposition 7.5 and (117), $\|S_t[K]\|_{t^{-1}} \leqslant \kappa_1 |K|_{C^0}$ and $\|DS_t[K]\|_{t^{-1}} \leqslant \kappa_1 |K|_{C^1}$ remain bounded uniformly in $t$.

**Lemma 7.6.** *For $t \geqslant 1$ and $f \in C^r(\mathbb{T}^l)$ we have*

$$\Bigl|\bigl\|D^s S_t[f]\bigr\|_{t^{-1}} - \bigl|D^s f\bigr|_{C^0}\Bigr| \leqslant 2\kappa_1 |f|_{C^r} t^{-1}, \quad 0 \leqslant s \leqslant r-1.$$

**Proof.** Since $f \in C^r(\mathbb{T}^l)$, $S_t[f]$ is analytic and $\mathbb{T}^l$ and $D_{t^{-1}}$ are compact, there exist $x_0 \in \mathbb{T}^l$ and $z_0 \in D_{t^{-1}}$ such that $|D^s f|_{C^0} = |D^s f(x_0)|$ and $\|D^s S_t[f]\|_{t^{-1}} = |D^s S_t[f](z_0)|$. Assume that $\|D^s S_t[f]\|_{t^{-1}} \geqslant |D^s f|_{C^0}$. Then applying (116) and (117) we have

$$
\begin{aligned}
0 &\leqslant \left|D^s S_t[f](z_0)\right| - \left|D^s f(x_0)\right| \\
&\leqslant \left|D^s S_t[f](z_0) - D^s S_t[f](\operatorname{Re} z_0)\right| + \left|D^s S_t[f](\operatorname{Re} z_0) - D^s f(\operatorname{Re} z_0)\right| \\
&\quad + \left|D^s f(\operatorname{Re} z_0)\right| - \left|D^s f(x_0)\right| \\
&\leqslant \kappa_1 |f|_{C^{s+1}} t^{-1} + \kappa_1 |f|_{C^r} t^{-r+s}.
\end{aligned}
$$

If $\|D^s S_t[f]\|_{t^{-1}} < |D^s f|_{C^0}$ we argue in a symmetric way and we obtain the result. $\square$

Since $K(\mathbb{T}^l)$ is real we have that for $t$ large enough, $S_t[K](D_{t^{-1}})$ is contained in the domain of $F$. Therefore, if $t$ is large enough, we have that $\|F \circ S_t[K]\|_{t^{-1}}$ and $\|F \circ S_t[K] - S_t[K] \circ T_\omega\|_{t^{-1}}$ remain uniformly bounded.

On the other hand, by (1) in Proposition 7.5 and the fact that $K$ satisfies the functional equation (6), we have that

$$
\begin{aligned}
\left|F \circ S_t[K] - S_t[K] \circ T_\omega\right|_{C^0} &\leqslant \left|F \circ S_t[K] - F \circ K\right|_{C^0} + \left|S_t[K] \circ T_\omega - K \circ T_\omega\right|_{C^0} \\
&\leqslant \kappa_1 |K|_{C^r} \left(|F|_{C^1} + 1\right) t^{-r}.
\end{aligned}
$$

Therefore, using the interpolation inequality (4) in Proposition 2.5 with $\rho_1 = t^{-1}$ and $\rho_2 = 0$, we obtain that

$$
\begin{aligned}
&\left\|F \circ S_t[K] - S_t[K] \circ T_\omega\right\|_{t^{-1}\xi} \\
&\quad \leqslant \left|F \circ S_t[K] - S_t[K] \circ T_\omega\right|_{C^0}^{1-\xi} \left\|F \circ S_t[K] - S_t[K] \circ T_\omega\right\|_{t^{-1}}^{\xi} \\
&\quad \leqslant C t^{-r(1-\xi)}.
\end{aligned}
\tag{118}
$$

Since all the non-degeneracy constants involve the first derivatives, by Lemma 7.6 we can perform the perturbative arguments in Section 5.

The constants in the non-degeneracy assumptions remain uniformly bounded for $S_t[K]$ in a neighborhood of size $t^{-1}$ and, *a fortiori*, in a neighborhood of size $t^{-1}\xi$.

Therefore, we can apply Theorem 3.11 with $\rho_0 = t^{-1}\xi$ and $\delta = t^{-1}\xi/12$ provided that we can find $t \geqslant 1$ such that

$$
C\left(t^{-1}\xi\right)^{-4\nu} t^{-r(1-\xi)} < 1
$$

for some constant $C > 0$, which depends on $l$, $\nu$, $\|DS_t[K]\|_{t^{-1}\xi}$, $\|N\|_{t^{-1}\xi}$, $\|A\|_{t^{-1}\xi}$, $|(\operatorname{avg}(A))^{-1}|$, $|(\operatorname{avg}(Q))^{-1}|$. By Lemma 7.6, if $t$ is big enough, the constant $C$ can be chosen independently on $t$.

The condition $r > 4\nu$ implies that there exist $\xi$ close to 0 and $t$ sufficiently large such that the previous inequality holds. Applying Theorem 3.11 with initial approximation $K_0 = S_t[K]$ we conclude that there exists an analytic solution $K_t^\infty$ of Eq. (6) defined on $D_{t^{-1}\xi/2}$ which satisfies

$$
\left\|K_t^\infty - S_t[K]\right\|_{t^{-1}\xi/2} \leqslant C_1 \left(t^{-1}\xi\right)^{-2\nu} t^{-r(1-\xi)},
$$

where $C_1$ depends on $l$, $\nu$, $\|DS_t[K]\|_{t^{-1}\xi}$, $\|N\|_{t^{-1}\xi}$, $\|A\|_{t^{-1}\xi}$, $|(\operatorname{avg}(A))^{-1}|, |(\operatorname{avg}(Q))^{-1}|$. As before $C_1$ can be taken independent on $t$. From (3) in Proposition 7.5 we have ($\tau \geqslant t$)

$$
\left\|S_\tau[K] - S_t[K]\right\|_{\tau^{-1}\xi/2} \leqslant \left\|S_\tau[K] - S_t[K]\right\|_{\tau^{-1}} \leqslant C_2 t^{-r}
$$

with $C_2$ independent on $t$.

We will apply Theorem 3.14 with $K_1$ and $K_2$ being $K_t^\infty$ and $K_\tau^\infty$ respectively, with $t,\tau \geqslant 1$. The application of this result requires condition (20) which in our case reads

$$\tilde{C}_3\kappa^2\left(\frac{\tau^{-1}\xi/2}{4}\right)^{-2\nu}\left\|K_t^\infty - K_\tau^\infty\right\|_{\tau^{-1}\xi/2} \leqslant 1. \tag{119}$$

The constant $\tilde{C}_3$ depends on $l$, $\nu$, $\|K_t^\infty\|_{\tau^{-1}\xi/2} \leqslant \|K_t^\infty\|_{t^{-1}\xi/2}$, $\|N_t\|_{t^{-1}\xi}$ $\|A_t\|_{t^{-1}\xi}$, $|(\mathrm{avg}(A_t))^{-1}|$, $|(\mathrm{avg}(Q_t))^{-1}|$, where $N_t$, $A_t$ and $Q_t$ are the expressions introduced in Definition 3.4 corresponding to $K_t^\infty$. As before $\tilde{C}_3$ can be chosen independently on $t,\tau \in [1,\infty)$, if $t$ is big enough. We write $C_3 = 8^{2\nu}\kappa^2\tilde{C}_3$.

**Lemma 7.7.** *There exists $t \geqslant 1$ such that if $\tau \geqslant t$ there exists $\varphi_{t,\tau} \in \mathbb{T}^l$ such that*

$$K_t^\infty \circ T_{\varphi_{t,\tau}} = K_\tau^\infty. \tag{120}$$

**Proof.** Using the previous notation we take $t$ big enough such that the constants $C_1$, $C_2$ and $C_3$ are independent on $t$ and such that

$$C_3 2^{2\nu}\xi^{-4\nu}\left(2C_1 t^{4\nu-r(1-\xi)} + C_2 t^{2\nu-r}\right) < 1. \tag{121}$$

We define $t_m = 2^m t$, $m \geqslant 0$, and we claim that for $t_m \leqslant \tau \leqslant 2t_m = t_{m+1}$ there exists $\varphi_{t_m,\tau} \in \mathbb{T}^l$ such that

$$K_{t_m}^\infty \circ T_{\varphi_{t_m,\tau}} = K_\tau^\infty.$$

Indeed, we apply Theorem 3.14 with $K_1 = K_{t_m}^\infty$ and $K_2 = K_\tau^\infty$. We have

$$\begin{aligned}\left\|K_{t_m}^\infty - K_\tau^\infty\right\|_{\tau^{-1}\xi/2} &\leqslant \left\|K_{t_m}^\infty - S_{t_m}[K]\right\|_{t_m^{-1}\xi/2} + \left\|S_{t_m}[K] - S_\tau[K]\right\|_{\tau^{-1}\xi/2} \\ &\quad + \left\|S_\tau[K] - K_\tau^\infty\right\|_{\tau^{-1}\xi/2} \\ &\leqslant C_1\left(t_m^{-1}\xi\right)^{-2\nu} t_m^{-r(1-\xi)} + C_2 t_m^{-r} + C_1\left(\tau^{-1}\xi\right)^{-2\nu}\tau^{-r(1-\xi)}.\end{aligned}$$

Using that $\tau \leqslant 2t_m$, condition (119) is implied by

$$C_3 2^{4\nu}\xi^{-4\nu}\left[2C_1 t_m^{4\nu-r(1-\xi)} + C_2 t_m^{2\nu-r}\right] < 1$$

which holds true by (121) since $t_m \geqslant t$.

If $\tau > t$ there exists $k \geqslant 0$ such that $t_k \leqslant \tau < t_{k+1}$. From the claim we can define $\varphi_{t,\tau} = \sum_{m=0}^{k-1}\varphi_{t_m,t_{m+1}} + \varphi_{t_k,\tau}$. Clearly $\varphi_{t,\tau}$ satisfies (120). □

Now consider $\tau_j \geqslant t$ going to $\infty$. Since $\varphi_{t,\tau_j} \in \mathbb{T}^l$ there exists a convergent subsequence, which we denote again $\varphi_{t,\tau_j}$, with limit $\varphi_\infty \in \mathbb{T}^l$.

Then

$$\begin{aligned}\left|K_{\tau_j}^\infty - K\right|_{C^0} &\leqslant \left|K_{\tau_j}^\infty - S_{\tau_j}[K]\right|_{C^0} + \left|S_{\tau_j}[K] - K\right|_{C^0} \\ &\leqslant C_1{\tau_j}^{2\nu-r} + \kappa_1|K|_{C^r}{\tau_j}^{-r}.\end{aligned}$$

Also, using that $K^\infty_{\tau_j} = K^\infty_t \circ T_{\varphi_{t,\tau_j}}$ we get

$$\left|K^\infty_t \circ T_{\varphi_\infty} - K\right|_{C^0} \leqslant \left|K^\infty_t \circ T_{\varphi_\infty} - K^\infty_t \circ T_{\varphi_{t,\tau_j}}\right|_{C^0} + \left|K^\infty_{\tau_j} - K\right|_{C^0}.$$

Finally, taking limit as $j$ goes to $\infty$ we get $K = K^\infty_t \circ T_{\varphi_\infty}$ and hence $K$ is analytic. □

*7.6. Non-trivial stable and unstable bundles*

*7.6.1. General comments and classification of bundles*

In this section we describe some examples of whiskered invariant tori with non-trivial stable/unstable bundles. Theorem 3.11 applies to these tori while other methods in the literature do not seem to apply. We note that, for some systems (see [35]), non-trivial bundles appear naturally when the systems experience resonances. We think that the study of bifurcations of the bundles of invariant tori deserves further exploration.

We are very grateful to Prof. R. Gompf for very enlightening discussions and, in particular, for constructing Example 7.6.2 and for providing us with a complete classification of rank 2 bundles over the torus, which we hope will be useful for future research.

We start from a non-trivial bundle $E \xrightarrow{\Pi} \mathbb{T}^l$ whose fibers are $\mathbb{R}^{d-l}$. Such examples are well known, but we detail a special one in Example 7.6.2.

As it is well known, when $l = 1$, the only obstruction to triviality is the orientation but when $l \geqslant 2$, there are other obstructions to triviality. We just mention the Euler characteristic or characteristic classes (Whitney–Stiefel or Pontryagin for real bundles or Chern classes for complex bundles). See [41,53,68]. The following construction is very similar to constructions in [29, Section 1.4].

We now consider a manifold $\mathcal{M}$ as a bundle given by

$$\mathcal{M} = \mathcal{E}^s \oplus \mathcal{E}^u \oplus T\mathbb{T}^l = \mathcal{E}^s \oplus \mathcal{E}^u \oplus \left(\mathbb{R}^l \times \mathbb{T}^l\right), \tag{122}$$

where $\mathcal{E}^s = E$, $\mathcal{E}^u = E^*$—the notation $E^*$ indicates the dual bundle of linear functions on the fibers—and $\oplus$ is the Whitney sum of bundles. We use the index $s, u$ to give an indication of future constructions.

We will also introduce the notation $T\mathbb{T}^l = \mathcal{E}^c$ so that we can write

$$\mathcal{M} = \mathcal{E}^s \oplus \mathcal{E}^u \oplus \mathcal{E}^c. \tag{123}$$

We denote the projections associated to each of the bundles $\mathcal{E}^s, \mathcal{E}^u, \mathcal{E}^c$ by $\Pi^s$, $\Pi^u$, $\Pi^c$ respectively.

The manifold $\mathcal{M}$ is a bundle over $\mathbb{T}^l$ whose fibers are $\mathbb{R}^{d-l} \times \mathbb{R}^{d-l} \times \mathbb{R}^l$. We can denote points in $\mathcal{M}$ as $(e^s, e^u, e^c, \theta)$, where $e^\sigma \in (\Pi^\sigma)^{-1}(\theta)$, $\sigma = s, u, c$.

We also recall that if $E$ is a linear bundle over a manifold $\mathcal{N}$, $TE$ can be canonically identified as a bundle over $T\mathcal{N}$ with fibers isomorphic to those of $E$. The basic idea is that the tangent directions along the fibers of $E$ can be identified with elements of the fibers since the space is linear.

Hence, we will write points in $T_{(e^s,e^u,e^c,\theta)}\mathcal{M}$ as $(v^s, v^u, v^c, v^t)$ where $v^\sigma \in \mathcal{E}^\sigma_\theta$, $\sigma = s, u, c$ and $v^t \in T_\theta\mathbb{T}^l$. Of course, we have the fact that the tangent bundle over the torus is trivial.

In a coordinate patch which trivializes the bundle, we can introduce the form $\alpha^{su} = \sum_{i=1}^{d-l} e^u_i de^s_i$. The key observation is that, even if the definition is in a coordinate patch, a change of coordinates in the patch leaves the form invariant. This is completely analogous to the coordinate construction of the canonical form in a cotangent bundle [2,28].

We also construct the canonical one-form in $\mathcal{E}^c$ by $\alpha^c = \sum_{i=1}^{l} e^c_i d\theta_i$ and consider the form $\alpha = \alpha^{su} + \alpha^c$.

The form $\Omega = d\alpha = d(\alpha^{su} + \alpha^c)$ is symplectic on $\mathcal{M}$. Indeed, it is clearly closed by definition. The fact that it is non-degenerate can be seen directly since, in the coordinate patch which trivializes the bundle, it has the standard form. As a consequence, $\mathcal{M}$ can be considered as an exact symplectic manifold.

We now relate the previous construction to our problem. We consider a linear bundle isomorphism on $\mathcal{E}^s$ over a rotation $T_\omega$, i.e. a family of invertible linear maps $A_\theta : \mathcal{E}^s_\theta \to \mathcal{E}^s_{\theta+\omega}$. We can then form a bundle isomorphism on $\mathcal{E}^s \oplus \mathcal{E}^u$ over the same rotation which preserves the form $\alpha^{su}$ by setting

$$A^{su}_\theta\big(e^s, e^u\big) = \big(A_\theta e^s, \big(A_\theta^{-1}\big)^\top e^u\big).$$

Then, the mapping

$$F\big(e^s, e^u, e^c, \theta\big) = \big(A^{su}_\theta\big(e^s, e^u\big), e^c, \theta + \omega\big)$$

is exact symplectic. The embedding $K : \mathbb{T}^l \to \mathcal{M}$ given by $K(\theta) = (0, 0, 0, \theta)$ clearly satisfies (6). If we compute the non-degeneracy conditions for this trivial solution, we obtain that $A(\theta) = \mathrm{Id}$ and $Q(\theta) = \mathrm{Id}$, which is the derivative of the frequency on the center direction.

The hyperbolicity condition is verified if

$$\big\|A^s\big\| < \mu_1 < 1$$

and

$$\big\|\big(A^u\big)^{-1}\big\| = \big\|\big(A^s\big)^\top\big\| < \mu_2 < 1.$$

This can be arranged by multiplying $A^s$ by a constant if necessary. Note that in this case, we can take $\mu_3$ to be as close to 1 as desired.

Furthermore, if $G$ is analytically close to $F$ (i.e. $\|F - G\|_{\mathcal{B}} < \varepsilon$, where $\mathcal{B}$ is a suitable complex subset of $\mathcal{M}$) and exact symplectic, then we have

$$\|G \circ K - K \circ T_\omega\|_{\rho_0} = \|F \circ K - G \circ K\|_{\rho_0} < \varepsilon$$

so that if $\varepsilon$ is small enough the hypotheses of Theorem 3.11 are met.

*7.6.2. An explicit example*

To make the whole construction more concrete, we just end with an explicit example of a non-trivial $\mathbb{R}^2$-bundle over $\mathbb{T}^2$ with positive Euler characteristic explained to us by Prof. Gompf. Many more examples can be found in [53]. Applying the construction in this section to these examples gives us symplectic manifolds and whiskered tori with non-trivial stable and unstable bundles. We construct a $\mathbb{R}^2$-bundle over $\mathbb{S}^2$ with non-zero Euler characteristic. If we identify $\mathbb{R}^2$ with $\mathbb{C}$ using the standard identification and $\mathbb{S}^2$ with the Riemann sphere, we can construct a non-trivial bundle in the semi-sphere, whose boundary is the circle $\mathbb{S}^1 \equiv \{|z| = 1\}$, by identifying the product bundle. We just give a gluing map on the unit sphere bundle, and extend it homogeneously. Hence, it suffices to give an identification mapping $i$ from $\mathbb{S}^1 \times \mathbb{S}^1$ to itself. The first factor is the boundary of the disk and the other factor is the unit bundle. We take $i(z, w) = (1/z, z^n w)$. Using partitions of identity, one can extend this bundle on a disk to a bundle of the torus.

## 8. Finite-dimensional flows

This section is devoted to the application of our method to find invariant tori for symplectic (locally Hamiltonian) vector-fields. Although we have already presented a result—in a rather abstract way—on existence of invariant tori for vector-fields in Theorem 3.15, we now present a direct proof of the results, following similar methods as in the case for maps. One motivation for writing this section is that the proof leads immediately to algorithms, which may be useful for applications involving vector-fields rather than maps. It may be of interest for practitioners to have algorithms for flows.

The proof for flows can also serve as a starting point for a proof for PDEs. We also note that the methods developed here apply to some ill-posed partial differential equations, which do not admit time-1 maps. Of course, the adaptation of the strategy of proof to PDEs involves several technical considerations (the generators of the evolutions are unbounded operators rather than differentiable ones). We postpone these considerations on PDEs to a forthcoming paper (see [16]).

We will study first the case of locally Hamiltonian flows. The case of globally Hamiltonian flows will be discussed in Section 9.

### 8.1. Some preliminaries on symplectic geometry

In this section we recall several well-known facts on symplectic geometry of vector-fields.

We will consider vector-fields on an exact symplectic manifold $\mathcal{M}$ with symplectic structure $\Omega = d\alpha$. We have the following definitions.

**Definition 8.1.** We say that a vector-field $X$ on $\mathcal{M}$ is symplectic when

$$\mathcal{L}_X \Omega = 0,$$

where $\mathcal{L}_X$ stands for the Lie derivative with respect to $X$.

**Definition 8.2.** We say that a vector-field $X$ is exact symplectic when there exists a smooth function $W$ on $\mathcal{M}$ such that

$$\mathcal{L}_X \alpha = dW.$$

An easy calculation checks that exact symplectic vector-fields are symplectic:

$$\mathcal{L}_X \Omega = \mathcal{L}_X \, d\alpha = d(\mathcal{L}_X \alpha) = d(dW) = 0.$$

However, the converse is not true. A well-known example is the following: consider the manifold $\mathcal{M} = \mathbb{T} \times \mathbb{R}$. We denote the corresponding coordinates $(q, p)$ and we set $\alpha = pdq$ and $\Omega = dp \wedge dq$. Consider now the vector-field $X = \partial_p$. It is symplectic but not exact symplectic.

Using Cartan's formula and the fact that $d\Omega = 0$, we obtain that $X$ is symplectic if and only if

$$0 = di_X \Omega + i_X d\Omega = di_X \Omega. \tag{124}$$

This means, by Poincaré lemma, that locally we can write

$$i_X \Omega = dH.$$

Of course, (124) does not imply that $H$ is a global function since in general it is only locally defined.

As a matter of fact, $H$ will be a global function if and only if the vector-field $X$ is exact symplectic. Indeed, since

$$dW = \mathcal{L}_X \alpha = d(i_X \alpha) + i_X \Omega,$$

we see that, if $X$ is exact symplectic, we can take $H = W - i_X \alpha$.

The above discussion shows that the only difference between symplectic and exact symplectic is the (de Rham) cohomology class of $i_X \Omega$. We introduce the following definition.

**Definition 8.3.** Let $K$ be an embedding from $\mathbb{T}^l$ into $\mathcal{M}$. We say that a family of vector-fields $X_\lambda$ with $\lambda \in \mathbb{R}^l$ spans the cohomology of $K(\mathbb{T}^l)$ at $\lambda = \bar{\lambda}$ if the map

$$\mathbb{R}^l \to H^1\big(\mathbb{T}^l\big)$$
$$v \mapsto \frac{d}{d\lambda}[K^* i_{X_\lambda}\Omega]_{|\lambda=\bar\lambda} v$$

is an isomorphism. Here we denote $H^1(\mathbb{T}^l)$ the first de Rham cohomology group of $\mathbb{T}^l$, which is well known to be $\mathbb{R}^l$ (see [30]).

In $\mathbb{T}^l \times \mathbb{R}^l$ with the standard symplectic form, we have that, denoting by $p_i$ the coordinates along $\mathbb{R}^l$, the family

$$X_\lambda = \sum_{i=1}^{l} \lambda_i \partial_{p_i}$$

spans the cohomology at every $\lambda$. Of course, in this case, the cohomology classes have a very simple characterization as the averages along each of the elementary cycles of $\mathbb{T}^l$.

### 8.2. Setting of the equations

The result for flows is based on the study of the equation

$$\partial_\omega K(\theta) = X\big(K(\theta)\big), \tag{125}$$

for $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$, where the operator $\partial_\omega$ (derivative in the direction $\omega$) is defined by

$$\partial_\omega K(\theta) = \sum_{i=1}^{l} \omega_i \frac{\partial K(\theta)}{\partial \theta_i}$$

and the vector-field $X : \mathcal{M} \to T\mathcal{M}$ is symplectic and real analytic.

Let $S_t$ be the flow of $X$. If $K : \mathbb{T}^l \to \mathcal{M}$ is a solution of (125) then

$$S_t\big(K(\theta)\big) = K(\theta + \omega t), \quad \theta \in \mathbb{T}^l,\ t \in \mathbb{R}, \tag{126}$$

and therefore the range of $K$ is invariant by $S_t$. Indeed, considering $\theta \in \mathbb{T}^l$ fixed, both sides of (126) satisfy the same Cauchy problem.

We first deal with a family of vector-fields $X_\lambda$ and we prove a version of the translated torus theorem. For an exact symplectic vector-field we will embed it into a family, then prove a vanishing lemma and finally prove the existence of an invariant torus. For families $X_\lambda$, the equation under consideration is

$$\partial_\omega K(\theta) = X_\lambda\big(K(\theta)\big), \tag{127}$$

where $\lambda \in \mathbb{R}^l$, the dependence of $X_\lambda$ in $\lambda$ is at least $C^1$ and we assume that the vector-field $X_\lambda$ spans the cohomology of $K_0(\mathbb{T}^l)$ in the sense of Definition 8.3, where $K_0$ is an approximate solution of (127).

A very important role will be played by the linearized equation

$$\frac{d\Delta}{dt} = A_\lambda(\theta + \omega t)\Delta, \tag{128}$$

where $A_\lambda(\theta) = DX_\lambda(K(\theta))$. Since $A_\lambda$ is a bounded operator, Eq. (128) admits an evolution operator, which is defined for all $t \in \mathbb{R}$, and we will denote it $U_\theta(t)$. It is characterized by

$$\frac{d}{dt}U_\theta(t) = A_\lambda(\theta + \omega t)U_\theta(t), \qquad U_\theta(0) = \mathrm{Id}. \tag{129}$$

*8.3. Non-degeneracy conditions*

To establish the existence of tori, we will require non-degeneracy conditions similar to the ones considered in the case of maps: namely, a spectral condition and a twist condition.

**Condition 8.4** *(Spectral non-degeneracy condition). Given $\lambda \in \mathbb{R}^l$ and an embedding $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ we say that the pair $(\lambda, K)$ is hyperbolic non-degenerate for the functional equation* (127) *if there is an analytic splitting*

$$T_{K(\theta)}\mathcal{M} = \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^c_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)}$$

*invariant under the linearized equation* (128) *in the sense that*

$$U_\theta(t)\mathcal{E}^{s,c,u}_{K(\theta)} = \mathcal{E}^{s,c,u}_{K(\theta+\omega t)}.$$

*Moreover the center subspace $\mathcal{E}^c_{K(\theta)}$ has dimension 2l. We denote $\Pi^s_{K(\theta)}$, $\Pi^c_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ the projections associated to this splitting and we denote*

$$U^{s,c,u}_\theta(t) = U_\theta(t)|_{\mathcal{E}^{s,c,u}_{K(\theta)}}.$$

*Furthermore, we assume that there exist $\beta_1, \beta_2, \beta_3 > 0$ and $C_h > 0$ independent of $\theta$ satisfying $\beta_3 < \beta_1$, $\beta_3 < \beta_2$ and such that the splitting is characterized by the following rate conditions*:

$$\begin{aligned}
\left\| U^s_\theta(t)U^s_\theta(\tau)^{-1} \right\|_\rho &\leqslant C_h e^{-\beta_1(t-\tau)}, \quad t \geqslant \tau \geqslant 0, \\
\left\| U^u_\theta(t)U^u_\theta(\tau)^{-1} \right\|_\rho &\leqslant C_h e^{\beta_2(t-\tau)}, \quad t \leqslant \tau \leqslant 0, \\
\left\| U^c_\theta(t)U^c_\theta(\tau)^{-1} \right\|_\rho &\leqslant C_h e^{\beta_3|t-\tau|}, \quad t, \tau \in \mathbb{R}.
\end{aligned} \tag{130}$$

**Remark 8.5.** As in the case of maps, if we have an approximately invariant splitting and

$$\begin{aligned}
\left\| U^s_\theta(t)U^s_\theta(\tau)^{-1} \right\|_\rho &\leqslant e^{-\tilde\beta_1(t-\tau)}, \quad T/2 \leqslant t-\tau \leqslant T, \\
\left\| U^u_\theta(t)U^u_\theta(\tau)^{-1} \right\|_\rho &\leqslant e^{\tilde\beta_2(t-\tau)}, \quad T/2 \leqslant \tau-t \leqslant T, \\
\left\| U^c_\theta(t)U^c_\theta(\tau)^{-1} \right\|_\rho &\leqslant e^{\tilde\beta_3|t-\tau|}, \quad T/2 \leqslant |t-\tau| \leqslant T,
\end{aligned}$$

for some $T$ large enough, then there exists a true invariant splitting, close to the approximately invariant one, and the bounds (130) with respect to this new splitting hold. This can be checked by using the time $T$ map.

**Remark 8.6.** The previous non-degeneracy condition just expresses that we can associate semi-groups in positive and negative times to the operator $A_\lambda(\theta+\omega t)$. More precisely, since the systems under consideration are non-autonomous, we should write

$$\begin{cases} \dfrac{dV}{dt} = A_\lambda(\tilde{\theta})V, \\ \dfrac{d\tilde{\theta}}{dt} = \omega, \quad \tilde{\theta}(0) = \theta. \end{cases}$$

Note that if the systems were autonomous, the exponential bounds would follow from the spectral properties of $A_\lambda$.

The linear operators $U_\theta^{s,c,u}(t)$ enjoy the following co-cycle property.

**Lemma 8.7.** *For all $\theta$ and $\omega$ and all times $t$, $\tau$ we have*

$$U_\theta^{s,c,u}(t+\tau) = U_{\theta+\omega t}^{s,c,u}(\tau)U_\theta^{s,c,u}(t), \quad t, \tau \in \mathbb{R}.$$

**Proof.** It follows from the classical argument of uniqueness for Cauchy ODE problems. Dropping the indexes $s$, $c$ and $u$, for $\theta$, $\omega$ and $t$ fixed, we define the functions

$$\psi_{1,t}(\tau) = U_\theta(t+\tau)\psi_0, \qquad \psi_{2,t}(\tau) = U_{\theta+\omega t}(\tau)U_\theta(t)\psi_0$$

for an arbitrary $\psi_0$. Since $U_\theta(0)$ is the identity operator, these two functions satisfy the same Cauchy problem and hence are equal. $\square$

**Condition 8.8** *(Twist condition). Let $A_\lambda(\theta) = DX_\lambda(K(\theta))$ and*

$$N(\theta) = \left[DK(\theta)^\top DK(\theta)\right]^{-1}.$$

*We say that the pair $(\lambda, K)$ satisfies the twist condition if the average on $\mathbb{T}^l$ of the matrix*

$$S_\lambda(\theta) = N(\theta)DK(\theta)^\top\left[\partial_\omega\left(J(K)^{-1}DKN\right) - A_\lambda J(K)^{-1}DKN\right](\theta)$$

*is non-singular.*

If a pair $(\lambda, K)$ with $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ satisfy both Conditions 8.4 and 8.8 we write $(\lambda, K) \in ND(\rho)$. If $X$ does not depend on $\lambda$ we simply write $K \in ND(\rho)$.

We note that Conditions 8.4 and 8.8 hold in open sets of $K$. The fact that Condition 8.8 holds for an open set (in the $C^1$ topology) is obvious since it is the non-degeneracy of a matrix that is just an explicit algebraic expression involving derivatives. The fact that Condition 8.4 is stable under perturbations will be the content of Section 8.7.

### *8.4. Statement of the results*

The first result below provides an existence result in the case of a family of symplectic vector-fields. From a sufficiently good approximate torus for a vector-field in the family it provides an invariant torus for a translated (with respect to the parameter) vector-field in the family.

**Theorem 8.9.** *Let $\omega \in D_h(\kappa, \nu)$ for some $\kappa > 0$ and $\nu \geqslant l-1$. Assume the following hypotheses*:

(1) *The vector-fields $X_\lambda$ are symplectic for every $\lambda \in \mathbb{R}^l$.*

(2) *The family $X_\lambda$ spans the cohomology of $K_0(\mathbb{T}^l)$ at $\lambda = \lambda_0$ in the sense of Definition* 8.3.
(3) *The pair $(\lambda_0, K_0)$ satisfies the non-degeneracy Conditions* 8.4 *and* 8.8.
(4) *The vector-fields $X_\lambda$ are real analytic and they can be extended holomorphically to a complex neighborhood of the image under $K_0$ of $D_{\rho_0}$:*

$$B_r = \left\{z \in \mathbb{C}^{2d} \;\middle|\; \exists \theta \in \left\{|\operatorname{Im}\theta| < \rho_0\right\} \text{ s.t. } \left|z - K_0(\theta)\right| < r\right\},$$

*for some $r > 0$, and are $C^1$ with respect to $\lambda$.*

*Define the error $E_0$ by*

$$E_0(\theta) = \partial_\omega K_0(\theta) - X_{\lambda_0}\big(K_0(\theta)\big).$$

*Then there exists a constant $C > 0$ depending on $l$, $\nu$, $|X_\lambda|_{C^2(B_r)}$, $\|DK_0\|_{\rho_0}$, $\|N_0\|_{\rho_0}$, $\|\frac{\partial X_\lambda(K)}{\partial\lambda}\|_{\rho_0}$, $\|S_0\|_{\rho_0}$, $|(\operatorname{avg}(S_0))^{-1}|$ (where $S_0$ and $N_0$ are as in Condition* 8.8 *replacing $\lambda$ by $\lambda_0$ and $K$ by $K_0$) and the norms of the projections $\|\Pi^{s,c,u}_{K_0(\theta)}\|_{\rho_0}$ such that, if $E_0$ satisfies the estimates*

$$C\kappa^4\delta^{-4\nu}\|E_0\|_{\rho_0} < 1$$

*and*

$$C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0} < r,$$

*where $0 < \delta \leqslant \min(1, \rho_0/12)$ is fixed, there exist an embedding $K_\infty$ and a vector $\lambda_\infty \in \mathbb{R}^l$ such that $(\lambda_\infty, K_\infty) \in ND(\rho_\infty := \rho_0 - 6\delta)$ and*

$$\partial_\omega K_\infty(\theta) = X_{\lambda_\infty}\big(K_\infty(\theta)\big). \tag{131}$$

*Furthermore, we have the estimates*

$$\|K_\infty - K_0\|_{\rho_\infty} \leqslant C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0}$$

*and*

$$|\lambda_\infty - \lambda_0| < C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0}.$$

The following theorem deals with the existence of invariant tori for exact symplectic vector-fields. It follows from the translated torus version Theorem 8.9 applied to a suitably chosen perturbation of the exact symplectic vector-field $X$ and a vanishing theorem whose proof is postponed to Section 8.8.

**Theorem 8.10.** *Let $\omega \in D_h(\kappa, \nu)$ for some $\kappa > 0$ and $\nu \geqslant l - 1$. Assume that:*

(1) *The vector-field $X$ is exact symplectic.*
(2) *$K_0$ satisfies the non-degeneracy Conditions* 8.4 *and* 8.8.
(3) *The vector-field $X$ is real analytic and it can be extended holomorphically to a complex neighborhood of the image under $K_0$ of $D_{\rho_0}$:*

$$B_r = \left\{z \in \mathbb{C}^{2d} \;\middle|\; \exists \theta \in \left\{|\operatorname{Im}\theta| < \rho_0\right\} \text{ s.t. } \left|z - K_0(\theta)\right| < r\right\},$$

*for some $r > 0$.*

*Denoting $E_0$ the initial error, there exists a constant $C > 0$ depending on $l$, $\nu$, $|X|_{C^2(B_r)}$, $\|DK_0\|_{\rho_0}$, $\|N_0\|_{\rho_0}$, $\|S_0\|_{\rho_0}$, $|(\operatorname{avg}(S_0))^{-1}|$ (where $S_0$ and $N_0$ are as in Condition 8.8 replacing $K$ by $K_0$) and the norms of the projections $\|\Pi^{s,c,u}_{K_0(\theta)}\|_{\rho_0}$ such that, if $E_0$ satisfies the estimates*

$$C\kappa^4\delta^{-4\nu}\|E_0\|_{\rho_0} < 1$$

*and*

$$C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0} < r,$$

*where $0 < \delta \leqslant \min(1, \rho_0/12)$ is fixed, then there exists an embedding $K_\infty \in ND(\rho_\infty := \rho_0 - 6\delta)$ such that*

$$\partial_\omega K_\infty(\theta) = X\big(K_\infty(\theta)\big). \tag{132}$$

*Furthermore, we have the estimate*

$$\|K_\infty - K_0\|_{\rho_\infty} \leqslant C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0}.$$

**Remark 8.11.** One could also formulate a local uniqueness result in the case of vector-fields. This can be done by a reduction to a time-one map (see [18]).

### *8.5. Linearized equation*

In this context we define the operator

$$\mathcal{G}_\omega(\lambda, K) = \partial_\omega K - X_\lambda \circ K$$

and we want to solve the equation $\mathcal{G}_\omega(\lambda, K) = 0$. As in the case of maps this will be done through a KAM iterative procedure, starting with $(\lambda_0, K_0)$ such that $E = \mathcal{G}_\omega(\lambda_0, K_0)$ is sufficiently small. Therefore we are lead to consider the linearized equation

$$\partial_\omega \Delta(\theta) - A_\lambda(\theta)\Delta(\theta) - \frac{\partial X_\lambda(K(\theta))}{\partial\lambda}\Lambda = -E(\theta), \tag{133}$$

where $A_\lambda(\theta) = DX_\lambda(K(\theta))$.

Let $\xi : \mathbb{T}^l \to \mathcal{M}$ be a function. From the spectral non-degeneracy condition we have

$$\Pi_{K(\theta+\omega t)} U_\theta(t)\xi(\theta) = U_\theta(t)\Pi_{K(\theta)}\xi(\theta), \tag{134}$$

where $\Pi$ stands for any of the projections $\Pi^s$, $\Pi^c$ and $\Pi^u$. Differentiating with respect to $t$ both sides of (134) and using (129) we obtain

$$\begin{aligned}&\frac{d}{d\theta}[\Pi_{K(\theta+\omega t)}]\omega U_\theta(t)\xi(\theta) + \Pi_{K(\theta+\omega t)} A_\lambda(\theta+\omega t) U_\theta(t)\xi(\theta)\\ &\quad = A_\lambda(\theta+\omega t)U_\theta(t)\Pi_{K(\theta)}\xi(\theta).\end{aligned}$$

Evaluating this expression at $t = 0$ and using the definition of $\partial_\omega$ we have

$$\partial_\omega\big[\Pi_{K(\theta)}\xi(\theta)\big] - \Pi_{K(\theta)}\partial_\omega\xi(\theta) + \Pi_{K(\theta)}A_\lambda(\theta)\xi(\theta) = A_\lambda(\theta)\Pi_{K(\theta)}\xi(\theta)$$

which implies

$$\Pi_{K(\theta)}\big[\partial_\omega - A_\lambda(\theta)\big]\xi(\theta) = \big[\partial_\omega - A_\lambda(\theta)\big]\Pi_{K(\theta)}\xi(\theta). \tag{135}$$

*8.5.1. Linearized equation on the center subspace*

We first project Eq. (133) on the center subspace. Using (135) we immediately obtain

$$\partial_\omega \Delta^c(\theta) - A_\lambda(\theta)\Delta^c(\theta) - \Pi^c_{K(\theta)} \frac{\partial X_\lambda(K(\theta))}{\partial \lambda} \Lambda = -E^c(\theta), \tag{136}$$

where $\Delta^c(\theta) = \Pi^c_{K(\theta)}\Delta(\theta)$ and $E^c(\theta) = \Pi^c_{K(\theta)}E(\theta)$.

*8.5.2. Small divisors equations and isotropic character of the torus*

The following result, which is completely analogous to Proposition 4.4, deals with the resolution of small divisors equations along characteristics (see [13,60–62]).

**Proposition 8.12.** *Assume that $\omega \in D_h(\kappa, \nu)$ with $\kappa > 0$ and $\nu \geqslant l - 1$. Let $h : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ be a real analytic function with zero average. Then, for any $0 < \delta < \rho$ there exists a unique analytic solution $v : D_{\rho-\delta} \supset \mathbb{T}^l \to \mathcal{M}$ of the linear equation*

$$\sum_{j=1}^{l} \omega_j \frac{\partial v}{\partial \theta_j} = h$$

*having zero average. Moreover, if $h \in \mathcal{A}_\rho$ then $v$ satisfies the following estimate*

$$\|v\|_{\rho-\delta} \leqslant C\kappa\delta^{-\nu}\|h\|_\rho, \quad 0 < \delta < \rho.$$

*The constant $C$ depends on $\nu$ and the dimension of the torus $l$.*

The following result provides the approximate isotropic character of the torus. This proposition is similar to the one in [14] and we do not reproduce its proof here. We note that it also follows by taking time-1 maps from the corresponding result for maps, which we have established in Section 4.1.1.

**Proposition 8.13.** *Let $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$, $\rho > 0$, be a real analytic mapping. Define the error*

$$E(\theta) := \partial_\omega K(\theta) - X_\lambda\big(K(\theta)\big).$$

*Let $L(\theta) = DK(\theta)^\top J(K(\theta))DK(\theta)$. There exists a constant $C$ depending on $l$, $\nu$ and $\|DK\|_\rho$ such that*

$$\|L\|_{\rho-2\delta} \leqslant C\kappa\delta^{-(\nu+1)}\|E\|_\rho, \quad 0 < \delta < \rho/2.$$

Once again, we use a normalization argument which allows us to write Eq. (136) in a suitable form. To do so, we need a result which allows to approximate the center subspace with the range of the $2d \times 2l$-matrix

$$\tilde{M}(\theta) = \big[DK(\theta), J\big(K(\theta)\big)^{-1}DK(\theta)N(\theta)\big], \tag{137}$$

where $N(\theta)$ is the normalization $l \times l$-matrix given by $N(\theta) = [DK(\theta)^\top DK(\theta)]^{-1}$, as in Proposition 4.16. One can prove the following result.

**Proposition 8.14.** *Denote by $\Gamma_{K(\theta)}$ the range of $\tilde{M}(\theta)$ and by $\Pi^\Gamma_{K(\theta)}$ the projection onto $\Gamma_{K(\theta)}$ according to the splitting $\mathcal{E}^s_{K(\theta)} \oplus \Gamma_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)}$.*

*Then there exists a constant $C > 0$ such that if*

$$\delta^{-1}\|E\|_\rho \leqslant C$$

*then we have the estimates (here* $\mathrm{dist}_\rho$ *has to be understood as the distance of subspaces in the Grassmanian sense)*

$$\operatorname{dist}_{\rho-2\delta}\big(\varGamma_{K(\theta)}, \mathcal{E}^c_{K(\theta)}\big) \leqslant C\delta^{-1}\|E\|_\rho,$$
$$\big\|\varPi^c_{K(\theta)} - \varPi^\varGamma_{K(\theta)}\big\|_{\rho-2\delta} \leqslant C\delta^{-1}\|E\|_\rho \tag{138}$$

*for every* $\delta \in (0, \rho/2)$ *and where* $C$*, as usual, depends on the non-degeneracy constants of the problem.*

The proof of the previous proposition follows the same lines as the one of Proposition 4.16. We refer the reader to Corollary 8.22 where we construct exact invariant splittings from approximate ones.

We introduce the change of function $\Delta^c = \tilde{M}\xi + \hat{e}\xi$, where $\xi : \mathbb{T}^l \to T\mathcal{M}$, with $\xi(\theta) \in T_{K(\theta)}\mathcal{M}$ and $\hat{e} = \varPi^c_{K(\theta)} - \varPi^\varGamma_{K(\theta)}$. We then get

$$\big[\partial_\omega \tilde{M}(\theta) - A_\lambda(\theta)\tilde{M}(\theta)\big]\xi(\theta) + \tilde{M}(\theta)\partial_\omega\xi(\theta) - \varPi^c_{K(\theta)}\frac{\partial X_\lambda(K(\theta))}{\partial\lambda}\varLambda = -E^c(\theta), \tag{139}$$

where we have dropped the terms depending on $\hat{e}\xi$, which are quadratic in the error. As in the case of maps, the matrix $\tilde{M}(\theta)$ is not invertible but the matrix $\tilde{M}(\theta)^\top J(K(\theta))\tilde{M}(\theta)$ is. Multiplying Eq. (139) by $\tilde{M}(\theta)^\top J(K(\theta))$ and then by $(\tilde{M}^\top J(K)\tilde{M})^{-1}$, we get the following equation

$$\begin{aligned}&\big(\tilde{M}(\theta)^\top J\big(K(\theta)\big)\tilde{M}(\theta)\big)^{-1}\tilde{M}(\theta)^\top J\big(K(\theta)\big)\big[\partial_\omega\tilde{M}(\theta) - A_\lambda(\theta)\tilde{M}(\theta)\big]\xi(\theta) + \partial_\omega\xi(\theta)\\&\quad= \big(\tilde{M}(\theta)^\top J\big(K(\theta)\big)\tilde{M}(\theta)\big)^{-1}\tilde{M}(\theta)^\top J\big(K(\theta)\big)\left[\varPi^c_{K(\theta)}\frac{\partial X_\lambda(K(\theta))}{\partial\lambda}\varLambda - E^c(\theta)\right].\end{aligned}$$

We are going to normalize the matrix $\partial_\omega\tilde{M}(\theta) - A_\lambda(\theta)\tilde{M}(\theta)$. To avoid some computational technicalities, we perform this normalization only when $K$ is a solution of (127). We refer the reader to the case of maps on how to handle the computations in the approximate case.

**Lemma 8.15.** *Let* $(\lambda, K)$ *be a solution of*

$$\partial_\omega K(\theta) = X_\lambda\big(K(\theta)\big) \tag{140}$$

*and* $\tilde{M}$ *be the matrix defined by* (137)*. Then there exists an* $l \times l$*-matrix* $S_\lambda(\theta)$ *such that*

$$\partial_\omega\tilde{M}(\theta) - A_\lambda(\theta)\tilde{M}(\theta) = \tilde{M}(\theta)\begin{pmatrix} 0_l & S_\lambda(\theta)\\ 0_l & 0_l\end{pmatrix}. \tag{141}$$

*The matrix* $S_\lambda(\theta)$ *has the form*

$$S_\lambda(\theta) = N(\theta)DK(\theta)^\top\big[\partial_\omega\big(J(K)^{-1}DKN\big) - A_\lambda J(K)^{-1}DKN\big](\theta).$$

**Proof.** Exactly in the same way as in the case of maps, if $K$ is a solution of (140) the columns of $\tilde{M}$ generate the center subspace. Since $\partial_\omega - A_\lambda(\theta)$ commute with $\varPi^c_{K(\theta)}$ we have that

$$\partial_\omega\tilde{M}(\theta) - A_\lambda(\theta)\tilde{M}(\theta) = \tilde{M}(\theta)C(\theta) \tag{142}$$

for some $2l \times 2l$-matrix $C(\theta)$. Differentiating Eq. (140) with respect to $\theta$ we obtain

$$\partial_\omega DK(\theta) = A_\lambda(\theta)DK(\theta). \tag{143}$$

This implies that

$$C(\theta) = \begin{pmatrix} 0_l & S_\lambda(\theta) \\ 0_l & R_\lambda(\theta) \end{pmatrix}.$$

Identifying blocks in (142) we end up with

$$\partial_\omega\big(J(K)^{-1}DKN\big) - A_\lambda J(K)^{-1}DKN = DKS_\lambda + J(K)^{-1}DKNR_\lambda. \tag{144}$$

Multiplying (144) by $DK^\top J(K)$ and using the isotropic character of the invariant torus, i.e.

$$L(\theta) = DK(\theta)^\top J\big(K(\theta)\big)DK(\theta) = 0,$$

it follows that

$$R_\lambda = DK^\top J(K)\big[\partial_\omega\big(J(K)^{-1}DKN\big) - A_\lambda J(K)^{-1}DKN\big]. \tag{145}$$

Expanding $\partial_\omega(J(K)^{-1}DKN)$ and using Eq. (143), we get

$$\partial_\omega\big(J(K)^{-1}DKN\big) = \partial_\omega\big(J(K)^{-1}\big)DKN + J(K)^{-1}A_\lambda DKN + J(K)^{-1}DK\partial_\omega N.$$

By differentiation of $NN^{-1} = \mathrm{Id}$, using (143) we easily obtain

$$\partial_\omega N = -NDK^\top\big[A_\lambda^\top + A_\lambda\big]DKN.$$

Also $\partial_\omega(J(K)^{-1}) = -J(K)^{-1}DJ(K)A_\lambda DKJ(K)^{-1}$.

Moreover the symplectic character of the vector-fields $X_\lambda$, i.e. $\mathcal{L}_{X_\lambda}\Omega = 0$ can be expressed by (recalling the definition of the Lie derivative)

$$\frac{d}{dt}\big[D\Phi_t^\top J(\Phi_t)D\Phi_t\big]_{|t=0} = 0, \tag{146}$$

where $\Phi_t$ is the flow solution of $X_\lambda$ and (146) implies

$$A_\lambda^\top J(K) + J(K)A_\lambda + DJ(K)X(K) = 0.$$

Using the previous calculations we obtain that the right-hand side of (145) vanishes, i.e. $R_\lambda = 0$.

Now multiplying (144) by $NDK^\top$ and using the definition of $N$ we have

$$S_\lambda(\theta) = N(\theta)DK(\theta)^\top\big[\partial_\omega\big(J(K)^{-1}DKN\big) - A_\lambda J(K)^{-1}DKN\big](\theta). \tag{147}$$

Using again the previous calculations we can express $S_\lambda$ as

$$S_\lambda = NDK^\top J(K)^{-1}\big[\mathrm{Id}_{2d} - DKNDK^\top\big]\big(A_\lambda + A_\lambda^\top\big)DKN.$$

We emphasize that this last formula coincides with (147) only when $K$ is an exact solution. If $K$ is only an approximate solution then both expressions are approximately equal. $\square$

We now turn to the case of *approximate* solutions. The procedure is similar to the one of the case of maps.

When $K$ is just an approximate solution, we define

$$(e_1, e_2) = \partial_\omega \tilde{M}(\theta) - A_\lambda(\theta)\tilde{M}(\theta) - \tilde{M}(\theta)\begin{pmatrix} 0_l & S_\lambda(\theta) \\ 0_l & 0_l \end{pmatrix}.$$

Some computations, using that $\partial_\omega DK(\theta) - A_\lambda(\theta)DK(\theta) = E(\theta)$ and the definition of $S_\lambda$ give $e_1 = DE$ and $e_2 = O(\|E\|_\rho, \|DE\|_\rho)$.

Next we just state the result without proof, but we indicate that it is quite analogous to the proof in the map case. We first identify—up to a small error—the center space with the span of the tangent and its symplectically conjugate and then compute the matrix of the derivative in these coordinates.

**Lemma 8.16.** *Assume $\omega \in D_h(\kappa, \nu)$ with $\kappa > 0$ and $\nu \geqslant l - 1$ and $\|E\|_\rho$ is small enough. Then there exist a matrix $B(\theta)$ and vectors $p_1$ and $p_2$ such that Eq.* (139) *can be written as*

$$\begin{aligned} &\left[\begin{pmatrix} 0_l & S(\theta) \\ 0_l & 0_l \end{pmatrix} + B(\theta)\right]\xi(\theta) + \partial_\omega \xi(\theta) \\ &\quad = p_1(\theta) + p_2(\theta) - \left(\tilde{M}(\theta)^\top J\big(K(\theta)\big)\tilde{M}(\theta)\right)^{-1} \tilde{M}(\theta)^\top J\big(K(\theta)\big)\Pi^c_{K(\theta)} \frac{\partial X_\lambda(K(\theta))}{\partial \lambda}\Lambda. \end{aligned} \tag{148}$$

*Moreover, the following estimates hold*:

$$\|p_1\|_\rho \leqslant C\|E\|_\rho, \tag{149}$$

*where $C$ just depends on $\|J(K)\|_\rho$, $\|N\|_\rho$, $\|DK\|_\rho$ and $\|\Pi^c_{K(\theta)}\|_\rho$. For $p_2$ and $B$ we have*

$$\|p_2\|_{\rho-2\delta} \leqslant C\kappa\delta^{-(\nu+1)}\|E\|^2_\rho \tag{150}$$

*and*

$$\|B\|_{\rho-2\delta} \leqslant C\kappa\delta^{-(\nu+1)}\|E\|_\rho \tag{151}$$

*for $\delta \in (0, \rho/2)$, where $C$ depends on $l$, $\nu$, $\|N\|_\rho$, $\|DK\|_\rho$, $|X_\lambda|_{C^2(B_r)}$, $|J|_{C^1(B_r)}$ and $\|\Pi^c_{K(\theta)}\|_\rho$.*

*8.5.3. Solution of the reduced equations*

The solution of the reduced equations works in the same way as in the case of maps. We sketch the procedure in this section and we emphasize on the cohomology obstructions on the equations.

We write $\xi = (\xi_1, \xi_2)$. We introduce the operator

$$\mathcal{L}\xi = \begin{pmatrix} 0_l & S(\theta) \\ 0_l & 0_l \end{pmatrix}\xi + \partial_\omega \xi = p_1(\theta) + Q(\theta)\Lambda, \tag{152}$$

where $p_1 = (p_{11}, p_{12})$ and $Q = (Q_1, Q_2)$. Using this decomposition of $\mathcal{E}^c_{K(\theta)}$ we can write Eq. (152) in the form

$$S(\theta)\xi_2(\theta) + \partial_\omega \xi_1(\theta) = p_{11}(\theta) + Q_1(\theta)\Lambda,$$

$$\partial_\omega \xi_2(\theta) = p_{12}(\theta) + Q_2(\theta)\Lambda.$$

We furthermore have

$$Q_1(\theta) = \left(N^\top DK^\top J(K)^{-\top}\right)(\theta)\left[\left(DKNDK^\top\right)(\theta) - \mathrm{Id}_{2d}\right] J\left(K(\theta)\right) \Pi^c_{K(\theta)} \frac{\partial X_\lambda(K(\theta))}{\partial \lambda} \Lambda,$$

$$Q_2(\theta) = DK(\theta)^\top J\left(K(\theta)\right) \Pi^c_{K(\theta)} \frac{\partial X_\lambda(K(\theta))}{\partial \lambda} \Lambda.$$

The assumption of spanning the cohomology of $K(\mathbb{T}^l)$ for the vector-field $X_\lambda$ ensures that we can choose $\Lambda$ such that the second equation is solvable in the sense of Proposition 8.12. Indeed, notice first that the cohomology along the hyperbolic bundle of the form $K^* i_{X_\lambda} \Omega$ is trivial and we have then

$$\left[\frac{d}{d\lambda} K^* i_{X_\lambda} \Omega\right] = \left[\frac{d}{d\lambda} K^* i_{\Pi^c X_\lambda} \Omega\right].$$

Identifying the cohomology class of a form in $H^1(\mathbb{T}^l)$ to its integral on the torus $\mathbb{T}^l$ and using the fact that the family $X_\lambda$ spans the cohomology of $K(\mathbb{T}^l)$ at $\lambda$ gives the result (since we can choose $\Lambda$ such that the average of $p_{12}(\theta) + Q_2(\theta)\Lambda$ vanishes).

The degree of freedom we get on the average of $\xi_2$ then allows us to solve the equation on $\xi_1$. Recall that we use the non-degeneracy conditions as stated in Condition 8.8. We obtain the following proposition.

**Proposition 8.17.** *Assume $\omega \in D_h(\kappa, \nu)$ with $\kappa > 0$ and $\nu \geqslant l - 1$, and $(\lambda, K)$ is a non-degenerate pair. If the error $\|E\|_\rho$ is small enough, there exist a mapping $\xi$, analytic on $D_{\rho-2\delta}$ and a vector $\Lambda \in \mathbb{R}^l$ solving Eq.* (152)*.*

*Moreover there exists a constant $C > 0$ depending on $\nu, l, \|K\|_\rho, |(\mathrm{avg}(A))^{-1}|, \|N\|_\rho$ and $\|\Pi^c_{K(\theta)}\|_\rho$ such that*

$$\|\xi\|_{\rho-2\delta} < C\kappa^2\delta^{-2\nu}\|E\|_\rho$$

*and*

$$|\Lambda| < C\|E\|_\rho.$$

*8.6. Linearized equation on the hyperbolic space*

We project the linearized equation (133)

$$\partial_\omega \Delta - A_\lambda(\theta)\Delta - \frac{\partial X_\lambda(K(\theta))}{\partial \lambda}\Lambda = -E(\theta)$$

on the stable and unstable subspaces by using the projections $\Pi^s_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ respectively. We denote $\Delta^s(\theta) = \Pi^s_{K(\theta)}\Delta(\theta)$, $\Delta^u(\theta) = \Pi^u_{K(\theta)}\Delta(\theta)$ and $\tilde{E}(\theta, \lambda, \Lambda) = \frac{\partial X_\lambda(K(\theta))}{\partial \lambda}\Lambda - E(\theta)$.

Using the previous notation and (135) we obtain

$$\partial_\omega \Delta^s(\theta) - A_\lambda(\theta)\Delta^s(\theta) = \Pi^s_{K(\theta)}\tilde{E}(\theta, \lambda, \Lambda) \tag{153}$$

for the stable part and

$$\partial_\omega \Delta^u(\theta) - A_\lambda(\theta)\Delta^u(\theta) = \Pi^u_{K(\theta)}\tilde{E}(\theta, \lambda, \Lambda) \tag{154}$$

for the unstable one.

The following result provides the solution of the previous equations.

**Proposition 8.18.** *Given $\rho > 0$, Eqs.* (153) *and* (154) *admit unique analytic solutions $\Delta^s : D_\rho \to \mathcal{E}^s$ and $\Delta^u : D_\rho \to \mathcal{E}^u$ respectively, such that $\Delta^{s,u}(\theta) \in \mathcal{E}^{s,u}_{K(\theta)}$. Furthermore there exist constants $C^{s,u}$ such that*

$$\left\| \Delta^{s,u} \right\|_\rho \leqslant C^{s,u}\left(\|E\|_\rho + |\Lambda|\right), \tag{155}$$

*where $C^{s,u}$ depend on $\beta_1$, $\|\Pi^s_{K(\theta)}\|_\rho$ (resp. $\beta_2$, $\|\Pi^u_{K(\theta)}\|_\rho$) and $C_h$, $\|\frac{\partial X_\lambda(K)}{\partial \lambda}\|_\rho$.*

**Proof.** The proof is based on the integration of the equation along the characteristics $\theta + \omega t$ and the use of the spectral non-degeneracy Condition 8.4. We give the proof for the stable case, the unstable case being symmetric (for negative times).

We introduce the function $\tilde{\Delta}(t) = \Delta^s(\theta + \omega t)$. If $\Delta^s$ has to satisfy (153) then $\tilde{\Delta}(t)$ has to satisfy the equation

$$\frac{d}{dt}\tilde{\Delta}(t) - A_\lambda(\theta + \omega t)\tilde{\Delta}(t) = \Pi^s_{K(\theta+\omega t)}\tilde{E}(\theta + \omega t, \lambda, \Lambda). \tag{156}$$

We first derive heuristically the formula (159) for $\Delta^s$. Then, examining the formula, it will be easy to justify the derivation.

Let $U_\theta(t)$ be the evolution operator characterized by

$$\frac{d}{dt}U_\theta(t) = A_\lambda(\theta + \omega t)U_\theta(t), \qquad U_\theta(0) = \mathrm{Id}. \tag{157}$$

Using the formula of the variation of parameters we have

$$\tilde{\Delta}(t) = U_\theta(t)\left[\tilde{\Delta}(0) + \int_0^t U_\theta^{-1}(s)\Pi^s_{K(\theta+\omega s)}\tilde{E}(\theta + \omega s, \lambda, \Lambda)\,ds\right]. \tag{158}$$

Using the co-cycle property given by Lemma 8.7 we have $U_\theta^{-1}(s) = U_{\theta+\omega s}(-s)$.

Since formula (158) is valid for all $\theta \in D_\rho \supset \mathbb{T}^l$ we can use it substituting $\theta$ by $\theta - \omega t$ and recovering the notation $\Delta^s$:

$$\Delta^s(\theta) = U_{\theta-\omega t}(t)\left[\Delta^s(\theta - \omega t) + \int_0^t U_{\theta-\omega(t-s)}(-s)\Pi^s_{K(\theta-\omega(t-s))}\tilde{E}\big(\theta - \omega(t-s), \lambda, \Lambda\big)\,ds\right].$$

We assume that $\Delta^s$, the solution we are looking for, stays in $\mathcal{E}^s$ and it is bounded; then $U_{\theta-\omega t}(t)\Delta^s(\theta - \omega t)$ goes to 0 when $t$ goes to $\infty$. Using again the co-cycle property we have

$$U_{\theta-\omega t}(t)U_{\theta-\omega(t-s)}(-s) = U_{\theta-\omega(t-s)}(t-s).$$

Then we write

$$\Delta^s(\theta) = U_{\theta-\omega t}(t)\Delta^s(\theta - \omega t) + \int_0^t U_{\theta-\omega(t-s)}(t-s)\Pi^s_{K(\theta-\omega(t-s))}\tilde{E}\big(\theta - \omega(t-s), \lambda, \Lambda\big)\,ds.$$

Performing the change of variable $\tau = t - s$ and letting $t$ go to $\infty$ we finally obtain

$$\Delta^s(\theta) = \int_0^\infty U_{\theta-\omega\tau}(\tau)\Pi^s_{K(\theta-\omega\tau)}\tilde{E}(\theta - \omega\tau, \lambda, \Lambda)\,d\tau. \tag{159}$$

Using the spectral non-degeneracy hypothesis the subintegral function is bounded by $C_h e^{-\beta_1 \tau} \|\Pi^s_{K(\theta-\omega\tau)}\|_\rho \|\tilde{E}\|_\rho$. The exponential bound assures the convergence of the integral and also permits to obtain the bound

$$\|\Delta^s\|_\rho = (C_h/\beta_1)\|\Pi^s_{K(\theta-\omega\tau)}\|_\rho \|\tilde{E}\|_\rho.$$

Once we have formula (159) we check directly that $\Delta^s$ is indeed a solution of (153). The absolute convergence of (159) justifies the exchange of limits and rearrangements used in the derivation.

The uniqueness follows from the fact that if we start with any solution $\Delta^s_1$ of (153), doing the previous manipulations we will end up with the same explicit formula (159). □

### *8.7. Change of non-degeneracy conditions in the iterative step*

The next result deals with the measure of the change of the splitting when we perturb a linear system in an Euclidean space $\mathcal{M}$.

Let $A_\lambda(\theta)$ be a family of linear maps from an Euclidean space $\mathcal{M}$ into itself, depending on $\theta \in D_\rho \supset \mathbb{T}^l$ and $\lambda \in \mathbb{R}^l$ and let $U_\theta$ be its evolution operator, i.e.

$$\frac{d}{dt} U_\theta(t) = A_\lambda(\theta + \omega t) U_\theta(t), \qquad U_\theta(0) = \mathrm{Id}.$$

Assume that $\mathcal{M}$ has an analytic family of splittings

$$\mathcal{M} = \mathcal{E}^s_\theta \oplus \mathcal{E}^c_\theta \oplus \mathcal{E}^u_\theta$$

invariant by $U_\theta$ in the sense that $U_\theta(t)\mathcal{E}^{s,c,u}_\theta = \mathcal{E}^{s,c,u}_{\theta+\omega t}$. Let $\Pi^{s,c,u}_\theta$ be the projections associated to this splitting and $U^{s,c,u}_\theta(t) = U_\theta(t)|_{\mathcal{E}^{s,c,u}_\theta}$. Assume furthermore there exist $\beta_1, \beta_2, \beta_3 > 0$ and $C_h > 0$ independent of $\theta$ satisfying $\beta_3 < \beta_1$, $\beta_3 < \beta_2$ and such that the splitting is characterized by the following rate conditions:

$$\begin{aligned}
\|U^s_\theta(t) U^s_\theta(\tau)^{-1}\|_\rho &\leqslant C_h e^{-\beta_1 (t-\tau)}, \quad t \geqslant \tau \geqslant 0,\\
\|U^u_\theta(t) U^u_\theta(\tau)^{-1}\|_\rho &\leqslant C_h e^{\beta_2 (t-\tau)}, \quad t \leqslant \tau \leqslant 0,\\
\|U^c_\theta(t) U^c_\theta(\tau)^{-1}\|_\rho &\leqslant C_h e^{\beta_3 |t-\tau|}, \quad t, \tau \in \mathbb{R}.
\end{aligned}$$

**Proposition 8.19.** *Assume that $A_\lambda(\theta)$ is a family of linear maps as before. Let $\tilde{A}_\lambda(\theta)$ be another family such that $\|\tilde{A}_\lambda - A_\lambda\|_\rho$ is small enough. Let $\tilde{U}_\theta(t)$ denote the evolution operator corresponding to $\tilde{A}_\lambda$, i.e.*

$$\frac{d}{dt} \tilde{U}_\theta(t) = \tilde{A}_\lambda(\theta + \omega t) \tilde{U}_\theta(t), \qquad \tilde{U}_\theta(0) = \mathrm{Id}.$$

*Then there exists a family of analytic splittings*

$$\mathcal{M} = \tilde{\mathcal{E}}^s_\theta \oplus \tilde{\mathcal{E}}^c_\theta \oplus \tilde{\mathcal{E}}^u_\theta$$

*which is invariant under the linearized equation*

$$\frac{d}{dt}\Delta = \tilde{A}_\lambda(\theta + \omega t)\Delta$$

*in the sense that*

$$\tilde{U}_\theta(t)\tilde{\mathcal{E}}_\theta^{s,c,u} = \tilde{\mathcal{E}}_{\theta+\omega t}^{s,c,u}.$$

*We denote* $\tilde{\Pi}_\theta^{s,c,u}$ *the projections associated to this splitting and denote*

$$\tilde{U}_\theta^{s,c,u}(t) = \tilde{U}_\theta(t)|_{\tilde{\mathcal{E}}_\theta^{s,c,u}}.$$

*Then there exist* $\tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3 > 0$ *and* $\tilde{C}_h > 0$ *independent of* $\theta$ *satisfying* $\tilde{\beta}_3 < \tilde{\beta}_1$, $\tilde{\beta}_3 < \tilde{\beta}_2$ *and such that the splitting is characterized by the following rate conditions*:

$$\left\| \tilde{U}_\theta^s(t)\tilde{U}_\theta^s(\tau)^{-1} \right\|_\rho \leqslant \tilde{C}_h e^{-\tilde{\beta}_1(t-\tau)}, \quad t \geqslant \tau \geqslant 0,$$

$$\left\| \tilde{U}_\theta^u(t)\tilde{U}_\theta^u(\tau)^{-1} \right\|_\rho \leqslant \tilde{C}_h e^{\tilde{\beta}_2(t-\tau)}, \quad t \leqslant \tau \leqslant 0,$$

$$\left\| \tilde{U}_\theta^c(t)\tilde{U}_\theta^c(\tau)^{-1} \right\|_\rho \leqslant \tilde{C}_h e^{\tilde{\beta}_3|t-\tau|}, \quad t, \tau \in \mathbb{R}.$$

*Furthermore the following estimates hold*

$$\left\| \tilde{\Pi}_\theta^{s,c,u} - \Pi_\theta^{s,c,u} \right\|_\rho \leqslant C\|\tilde{A}_\lambda - A_\lambda\|_\rho, \tag{160}$$

$$|\tilde{\beta}_i - \beta_i| \leqslant C\|\tilde{A}_\lambda - A_\lambda\|_\rho, \quad i = 1, 2, 3, \tag{161}$$

$$\tilde{C}_h = C_h. \tag{162}$$

**Proof.** We provide the proof of the statements concerning the stable subspace. We divide it into several steps. We use the notation of Condition 8.4.

*Step 1. Construction of the invariant splitting.* We look for the invariant splitting associated to the linearized equation

$$\frac{d}{dt}W(t) = \tilde{A}_\lambda(\theta + \omega t)W(t) \tag{163}$$

focusing on the stable bundle. We write (163) as

$$\frac{d}{dt}W(t) = A_\lambda(\theta + \omega t)W(t) + B_\lambda(\theta + \omega t)W(t) \tag{164}$$

with $B_\lambda = \tilde{A}_\lambda - A_\lambda$. Since we are interested in solutions decreasing exponentially at $\infty$, for $a > 0$ we introduce the space

$$C_a = \left\{ f : [0,\infty) \to \mathbb{C}^{2d} \;\middle|\; f \text{ continuous}, \sup_{t \geqslant 0} e^{at}\left|f(t)\right| < \infty \right\},$$

with the norm $|f|_a = \sup_{t\geqslant 0} e^{at}|f(t)|$.

Given $\alpha \in (\beta_3, \beta_1)$ we look for solutions of (164) in the space $C_\alpha$. We begin with the following auxiliary result.

**Lemma 8.20.** *Let $\theta \in D_\rho \supset \mathbb{T}^l$, $\xi \in \mathcal{E}^s_{K(\theta)}$ and $H \in C_\alpha$ with $\alpha \in (\beta_3, \beta_1)$. Consider the equation*

$$w' = A_\lambda(\theta + \omega t)w + H(t). \tag{165}$$

*Then there exists a unique function $\mathcal{K}(\xi, H) \in C_\alpha$ such that:*

(i) *$\mathcal{K}(\xi, H)$ is solution of (165).*
(ii) *$\Pi^s_\theta \mathcal{K}(\xi, H)(0) = \xi$.*

*Moreover $\mathcal{K}(\xi, H) = \mathcal{K}_1(\xi) + \mathcal{K}_2(H)$, where $\mathcal{K}_1 : \mathcal{E}^s_\theta \to C_\alpha$ and $\mathcal{K}_2 : C_\alpha \to C_\alpha$ are bounded linear operators and*

$$|\mathcal{K}_1| \leqslant C_h, \tag{166}$$

$$|\mathcal{K}_2| \leqslant C_h\left(\frac{|\Pi^s|}{\beta_1 - \alpha} + \frac{|\Pi^{cu}|}{\alpha - \beta_3}\right), \tag{167}$$

*where $|\Pi^{s,cu}| = \sup_{\theta \in D_\rho} |\Pi^{s,cu}_\theta|$.*

**Proof.** If $w \in C_\alpha$ is a solution of (165) in $[0, \infty)$ and $t, \tau \geqslant 0$ we have

$$w(t) = U_\theta(t)U_\theta(\tau)^{-1}w(\tau) + \int_\tau^t U_\theta(t)U_\theta(s)^{-1}H(s)\,ds. \tag{168}$$

Projecting (168) to the center-unstable subspace and using the invariance of the splitting $\mathcal{E}^s_\theta \oplus (\mathcal{E}^c_\theta \oplus \mathcal{E}^u_\theta)$ with respect to $U_\theta$ and writing $\Pi^{cu}_\theta$ the projection onto $\mathcal{E}^c_\theta \oplus \mathcal{E}^u_\theta$

$$\Pi^{cu}_{\theta+\omega t} w(t) = U_\theta(t)U_\theta(\tau)^{-1}\Pi^{cu}_{\theta+\omega\tau} w(\tau) + \int_\tau^t U_\theta(t)U_\theta(s)^{-1}\Pi^{cu}_{\theta+\omega s} H(s)\,ds. \tag{169}$$

If $\tau \geqslant t$ we have

$$\left|U_\theta(t)U_\theta(\tau)^{-1}\Pi^{cu}_{\theta+\omega\tau} w(\tau)\right| \leqslant C_h e^{\beta_3(\tau - t)}\left|\Pi^{cu}_{\theta+\omega\tau}\right| e^{-\alpha\tau}|w|_\alpha$$

which goes to zero as $\tau$ tends to $\infty$. Also, if $s > t$

$$\left|U_\theta(t)U_\theta(s)^{-1}\Pi^{cu}_{\theta+\omega s} H(s)\right| \leqslant C_h e^{\beta_3(s-t)}\left|\Pi^{cu}_{\theta+\omega s}\right| e^{-\alpha s}|H|_\alpha$$

guarantees that we can take limit $\tau \to \infty$ in the integral in (169). Then we have

$$\Pi^{cu}_{\theta+\omega t} w(t) = \int_\infty^t U_\theta(t)U_\theta(s)^{-1}\Pi^{cu}_{\theta+\omega s} H(s)\,ds.$$

Using the projection to the stable subspace, we obtain

$$\begin{aligned}
w(t) &= \Pi^s_{\theta+\omega t} w(t) + \Pi^{cu}_{\theta+\omega t} w(t) \\
&= U_\theta(t)\Pi^s_\theta w(0) + \int_0^t U_\theta(t)U_\theta(s)^{-1}\Pi^s_{\theta+\omega s} H(s)\,ds \\
&\quad + \int_\infty^t U_\theta(t)U_\theta(s)^{-1}\Pi^{cu}_{\theta+\omega s} H(s)\,ds. \qquad (170)
\end{aligned}$$

Once we have the explicit expression of $w$, we easily check that it actually belongs to $C_\alpha$. We define $\mathcal{K}_1(\xi)(t) = U_\theta(t)\xi$ and $\mathcal{K}_2(H)(t)$ to be the sum of the two integrals in (170). A simple calculation gives the bounds (166) and (167). □

By Lemma 8.20 the solutions of (164) belonging to $C_\alpha$ satisfy

$$w(t) = \mathcal{K}_1\big(\Pi^s_\theta w(0)\big) + \mathcal{K}_2\big(B_\lambda(\theta+\omega\cdot)w\big)(t).$$

Note that $B_\lambda(\theta+\omega t)$ is bounded in $t$ and moreover $|B_\lambda(\theta+\omega t)| \leqslant \gamma$, where $\gamma = \|\tilde{A}_\lambda - A_\lambda\|_\rho$. We introduce the linear map $\tilde{\mathcal{K}}_2 : C_\alpha \to C_\alpha$ defined by

$$\tilde{\mathcal{K}}_2(w) = \mathcal{K}_2\big(B_\lambda(\theta+\omega\cdot)w\big).$$

Clearly $|\tilde{\mathcal{K}}_2(w)| \leqslant |\mathcal{K}_2||B_\lambda(\theta+\omega\cdot)|$. With the above introduced notation, given $\xi \in \mathcal{E}^s_\theta$, there is a unique solution $w \in C_\alpha$ such that $\Pi^s_\theta w(0) = \xi$ which is given by

$$w = \mathcal{K}_1(\xi) + \tilde{\mathcal{K}}_2(w).$$

Since $|B_\lambda| \leqslant \gamma < 1$ we can write

$$w = (\mathrm{Id} - \tilde{\mathcal{K}}_2)^{-1}\mathcal{K}_1(\xi).$$

Therefore $\tilde{\mathcal{E}}^s_\theta$ is the graph of

$$\xi \mapsto \tilde{M}^s(\theta)\xi := \Pi^{cu}_\theta(\mathrm{Id} - \tilde{\mathcal{K}}_2)^{-1}\mathcal{K}_1(\xi)(0) = \Pi^{cu}_\theta \sum_{k=1}^\infty \tilde{\mathcal{K}}_2^k \mathcal{K}_1(\xi)(0),$$

where the sum starts with $k = 1$ because $\Pi^{cu}_\theta \mathcal{K}_1 = 0$. Note that the analyticity in $\theta$ is preserved in all the previous manipulations, hence $\tilde{M}_\theta$ depends analytically in $\theta$. Since $|\mathcal{K}_2| \leqslant C\gamma$ then $\|\tilde{M}^s(\theta)\|_\rho < C\gamma$. In a completely analogous way we find $\tilde{\mathcal{E}}^{cu}_\theta$, and integrating with negative times we get $\tilde{\mathcal{E}}^u_\theta$ and $\tilde{\mathcal{E}}^{sc}_\theta$. Finally $\tilde{\mathcal{E}}^c_\theta = \mathcal{E}^{sc}_\theta \cap \mathcal{E}^{cu}_\theta$.

*Step 2. Estimates on the projections.* To get the bounds for the projections we follow the same argument as in the case of maps. We only give the argument for the stable subspace. Let $\tilde{M}^{cu}(\theta)$ be the linear map whose graph gives $\tilde{\mathcal{E}}^{cu}_\theta$.

We write

$$\Pi^s_\theta\xi = \big(\xi^s, 0\big), \qquad \tilde{\Pi}^s_\theta\xi = \big(\tilde{\xi}^s, \tilde{M}^s(\theta)\tilde{\xi}^s\big),$$
$$\Pi^{cu}_\theta\xi = \big(0, \xi^{cu}\big), \qquad \tilde{\Pi}^{cu}_\theta\xi = \big(\tilde{M}^{cu}(\theta)\tilde{\xi}^{cu}, \tilde{\xi}^{cu}\big),$$

and then

$$\xi^s = \tilde{\xi}^s + \tilde{M}^{cu}(\theta)\tilde{\xi}^{cu},$$
$$\xi^{cu} = \tilde{M}^s(\theta)\tilde{\xi}^s + \tilde{\xi}^{cu}.$$

Since $\tilde{M}^s(\theta)$ and $\tilde{M}^{cu}(\theta)$ are $O(\gamma)$ we can write

$$\begin{pmatrix} \tilde{\xi}^s \\ \tilde{\xi}^{cu} \end{pmatrix} = \begin{pmatrix} \mathrm{Id} & \tilde{M}^{cu}(\theta) \\ \tilde{M}^s(\theta) & \mathrm{Id} \end{pmatrix}^{-1} \begin{pmatrix} \xi^s \\ \xi^{cu} \end{pmatrix}$$

and then deduce that

$$\left|\left(\tilde{\Pi}^s_\theta - \Pi^s_\theta\right)\xi\right| \leqslant \left|\left(\tilde{\xi}^s - \xi^s, \tilde{M}^s(\theta)\tilde{\xi}^s\right)\right| \leqslant C\gamma.$$

*Step 3. Estimates on the growth conditions.* To get the exponential bounds let

$$\psi(t) = \tilde{U}_\theta(t)\tilde{U}_\theta(\tau)^{-1}\psi(\tau)$$

with $\psi(\tau) = (\xi, \tilde{M}(\theta+\omega\tau)\xi) \in \tilde{\mathcal{E}}^s_{\theta+\omega\tau}$. The function $\psi$ satisfies Eq. (164) and hence

$$\left|\psi(t)\right| \leqslant \left|U_\theta(t)U_\theta(\tau)^{-1}\psi(\tau)\right| + \int_\tau^t \left|U_\theta(t)U_\theta(s)^{-1}(\tilde{A}_\lambda - A_\lambda)(\theta+\omega s)\psi(s)\right| ds,$$

for $t \geqslant \tau$. Let $\chi$ be the auxiliary function defined by $\chi(t) = e^{\beta_1 t}|\psi(t)|$. Using the bounds of Condition 8.4 we have

$$\chi(t) \leqslant C_h\chi(\tau) + C_h C\gamma \int_\tau^t \chi(s)\, ds, \quad t \geqslant \tau.$$

By Gronwall's lemma we have $\chi(t) \leqslant C_h\chi(\tau)e^{C_h C\gamma(t-\tau)}$ and hence

$$\psi(t) \leqslant e^{-\beta_1 t}C_h e^{\beta_1\tau}\psi(\tau)e^{C_h C\gamma(t-\tau)}.$$

We conclude that

$$\left|\tilde{U}_\theta(t)\tilde{U}_\theta(\tau)^{-1}\psi(\tau)\right| \leqslant C_h e^{-(\beta_1 - C_h C\gamma)(t-\tau)}\left|\psi(\tau)\right|, \quad t \geqslant \tau.$$

We take $\tilde{C}_h = C_h$ and $\tilde{\beta}_1 = \beta_1 - C_h C\gamma$, which proves (161). □

The first consequence of Proposition 8.19 is that in the iterative step the small change of $K$ produces a small change in the invariant splitting and in the hyperbolicity constants.

**Corollary 8.21.** *Assume that* $(\lambda, K)$ *satisfies the hyperbolic non-degeneracy Condition* 8.4 *and that* $\|K - \tilde{K}\|_\rho$ *is small enough. If we denote* $\tilde{A}_\lambda(\theta) = DX_\lambda(\tilde{K}(\theta))$*, we can define an evolution operator, denoted* $\tilde{U}_\theta(t)$ *such that*

$$\frac{d}{dt}\tilde{U}_\theta(t) = \tilde{A}_\lambda(\theta+\omega t)\tilde{U}_\theta(t), \qquad \tilde{U}_\theta(0) = \mathrm{Id}.$$

*Then there exists an analytic splitting for* $\tilde{K}$*, i.e.*

$$T_{\tilde{K}(\theta)}\mathcal{M} = \mathcal{E}^s_{\tilde{K}(\theta)} \oplus \mathcal{E}^c_{\tilde{K}(\theta)} \oplus \mathcal{E}^u_{\tilde{K}(\theta)}$$

*which is invariant under the linearized equation* (128) (*replacing* $K$ *by* $\tilde{K}$) *in the sense that*

$$\tilde{U}_\theta(t)\mathcal{E}^{s,c,u}_{\tilde{K}(\theta)} = \mathcal{E}^{s,c,u}_{\tilde{K}(\theta+\omega t)}.$$

*We denote* $\Pi^s_{\tilde{K}(\theta)}$, $\Pi^c_{\tilde{K}(\theta)}$ *and* $\Pi^u_{\tilde{K}(\theta)}$ *the projections associated to this splitting. Denoting*

$$\tilde{U}^{s,c,u}_\theta(t) = \tilde{U}_\theta(t)|_{\mathcal{E}^{s,c,u}_{\tilde{K}(\theta)}},$$

*there exist* $\tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3 > 0$ *and* $\tilde{C}_h > 0$ *independent of* $\theta$ *satisfying* $\tilde{\beta}_3 < \tilde{\beta}_1$, $\tilde{\beta}_3 < \tilde{\beta}_2$ *and such that the splitting is characterized by the following rate conditions*:

$$\left\| \tilde{U}^s_\theta(t)\tilde{U}^s_\theta(\tau)^{-1} \right\|_\rho \leqslant \tilde{C}_h e^{-\tilde{\beta}_1(t-\tau)}, \quad t \geqslant \tau \geqslant 0,$$

$$\left\| \tilde{U}^u_\theta(t)\tilde{U}^u_\theta(\tau)^{-1} \right\|_\rho \leqslant \tilde{C}_h e^{\tilde{\beta}_2(t-\tau)}, \quad t \leqslant \tau \leqslant 0,$$

$$\left\| \tilde{U}^c_\theta(t)\tilde{U}^c_\theta(\tau)^{-1} \right\|_\rho \leqslant \tilde{C}_h e^{\tilde{\beta}_3|t-\tau|}, \quad t, \tau \in \mathbb{R}.$$

*Furthermore the following estimates hold*

$$\left\| \Pi^{s,c,u}_{\tilde{K}(\theta)} - \Pi^{s,c,u}_{K(\theta)} \right\|_\rho \leqslant C\|\tilde{K} - K\|_\rho, \tag{171}$$

$$|\tilde{\beta}_i - \beta_i| \leqslant C\|\tilde{K} - K\|_\rho, \quad i = 1, 2, 3, \tag{172}$$

$$\tilde{C}_h = C_h. \tag{173}$$

**Proof.** We just take $A_\lambda(\theta) = DX_\lambda(K(\theta))$, $\tilde{A}_\lambda(\theta) = DX_\lambda(\tilde{K}(\theta))$, $\mathcal{E}^{s,c,u}_{K(\theta)} = \mathcal{E}^{s,c,u}_\theta$, $\mathcal{E}^{s,c,u}_{\tilde{K}(\theta)} = \tilde{\mathcal{E}}^{s,c,u}_\theta$, $\Pi^{s,c,u}_{K(\theta)} = \Pi^{s,c,u}_\theta$ and $\Pi^{s,c,u}_{\tilde{K}(\theta)} = \tilde{\Pi}^{s,c,u}_\theta$ in Proposition 8.19 and we use that $\|\tilde{A}_\lambda(\theta) - A_\lambda(\theta)\|_\rho \leqslant \|X\|_{C^2}\|\tilde{K}(\theta) - K(\theta)\|_\rho$. $\square$

The second consequence of Proposition 8.19 is that if we have a sufficiently good approximate splitting associated to Eq. (129) then there is a true invariant splitting nearby.

**Corollary 8.22.** *Assume that* $T_{K(\theta)}\mathcal{M} = \mathcal{E}^{*s}_{K(\theta)} \oplus \mathcal{E}^{*c}_{K(\theta)} \oplus \mathcal{E}^{*u}_{K(\theta)}$ *is a splitting approximately invariant under the linearized equation* (128) *with evolution operator* $U_\theta(t)$, *in the sense that* $A_\lambda(\theta) = DX_\lambda(K(\theta))$ *can be represented as*

$$A_\lambda(\theta) = \begin{pmatrix} A^{11}_\lambda(\theta) & A^{12}_\lambda(\theta) & A^{13}_\lambda(\theta) \\ A^{21}_\lambda(\theta) & A^{22}_\lambda(\theta) & A^{23}_\lambda(\theta) \\ A^{31}_\lambda(\theta) & A^{32}_\lambda(\theta) & A^{33}_\lambda(\theta) \end{pmatrix}$$

*with respect to this splitting with* $\|A^{ij}_\lambda(\theta)\|_\rho \leqslant C\delta^{-1}\|E\|_\rho$ *if* $i \neq j$. *Let* $\Pi^{*s,c,u}_{K(\theta)}$ *be the projections associated to this splitting.*

*Let* $\tilde{U}^{s,c,u}_\theta$ *be the evolution operators of* $\dot{\Delta} = A^{11}_\lambda(\theta + \omega t)\Delta$, $\dot{\Delta} = A^{22}_\lambda(\theta + \omega t)\Delta$ *and* $\dot{\Delta} = A^{33}_\lambda(\theta + \omega t)\Delta$ *respectively, and assume*

$$\left\| \tilde{U}^s_\theta(t)\tilde{U}^s_\theta(\tau)^{-1} \right\|_\rho \leqslant C^*_h e^{-\beta^*_1(t-\tau)}, \quad t \geqslant \tau \geqslant 0,$$

$$\left\| \tilde{U}^u_\theta(t)\tilde{U}^u_\theta(\tau)^{-1} \right\|_\rho \leqslant C^*_h e^{\beta^*_2(t-\tau)}, \quad t \leqslant \tau \leqslant 0,$$

$$\left\| \tilde{U}^c_\theta(t)\tilde{U}^c_\theta(\tau)^{-1} \right\|_\rho \leqslant C^*_h e^{\beta^*_3|t-\tau|}, \quad t, \tau \in \mathbb{R},$$

*for some $\beta^*_{1,2,3}$, $C^*_h > 0$ such that $\beta^*_3 < \beta^*_1$, $\beta^*_3 < \beta^*_2$. Then there exists an analytic splitting $T_{K(\theta)}\mathcal{M} = \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^c_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)}$ invariant under Eq. (128). Let $\Pi^{s,c,u}_{K(\theta)}$ be the projections associated to this splitting and $U^{s,c,u}_\theta(t) = U_\theta(t)|_{\mathcal{E}^{s,c,u}_{K(\theta)}}$. Moreover there exist $\beta_{1,2,3} > 0$ and $C_h > 0$ independent of $\theta$ satisfying $\beta_3 < \beta_1$, $\beta_3 < \beta_2$ and such that the splitting is characterized by the following rate conditions*:

$$\left\| U^s_\theta(t) U^s_\theta(\tau)^{-1} \right\|_\rho \leqslant C_h e^{-\beta_1(t-\tau)}, \quad t \geqslant \tau \geqslant 0,$$
$$\left\| U^u_\theta(t) U^u_\theta(\tau)^{-1} \right\|_\rho \leqslant C_h e^{\beta_2(t-\tau)}, \quad t \leqslant \tau \leqslant 0,$$
$$\left\| U^c_\theta(t) U^c_\theta(\tau)^{-1} \right\|_\rho \leqslant C_h e^{\beta_3|t-\tau|}, \quad t, \tau \in \mathbb{R},$$

*and*

$$\left\| \Pi^{*s,c,u}_{K(\theta)} - \Pi^{s,c,u}_{K(\theta)} \right\|_\rho \leqslant C\delta^{-1} \|E\|_\rho, \tag{174}$$
$$\left| \beta^*_i - \beta_i \right| \leqslant C\delta^{-1} \|E\|_\rho, \quad i = 1, 2, 3, \tag{175}$$
$$C^*_h = C_h. \tag{176}$$

**Proof.** We make the same identifications as in the proof of Corollary 8.21. Consider the auxiliary linear equation

$$\dot{\Delta}(t) = A^*_\lambda(\theta + \omega t)\Delta(t) \tag{177}$$

with

$$A^*_\lambda(\theta) = \begin{pmatrix} A^{11}_\lambda(\theta) & 0 & 0 \\ 0 & A^{22}_\lambda(\theta) & 0 \\ 0 & 0 & A^{33}_\lambda(\theta) \end{pmatrix}.$$

Clearly $U^*_\theta(t) = (\tilde{U}^s_\theta(t), \tilde{U}^c_\theta(t), \tilde{U}^u_\theta(t))$ is a solution of (177). By hypothesis $\|A^*_\lambda(\theta) - A_\lambda(\theta)\|_\rho$ is small. Then the application of Proposition 8.19 gives the results. □

**Remark 8.23.** We can give an alternative proof to Proposition 8.19, parallel to the one for maps. We just sketch it for the stable bundle in the following. Recall that we have the invariance condition for all times $t \geqslant 0$

$$U^s_\theta(t)\mathcal{E}^s_{K(\theta)} = \mathcal{E}^s_{K(\theta+\omega t)}.$$

The graph condition then writes for all times $t \geqslant 0$

$$U^s_\theta(t) \begin{pmatrix} \mathrm{Id} \\ M(\theta) \end{pmatrix} \in \mathrm{Graph}\big(M(\theta + \omega t)\big).$$

We now consider the time-one map $U_1 = U^s_\theta(1)$. The graph condition leads to a functional equation which is solved by a fixed point argument. To propagate the result to any time and get the estimates, one just has to use the co-cycle property as stated in Lemma 8.7.

The other non-degeneracy conditions can be checked in exactly the same way (as in the previous section) and we do not repeat the arguments.

**Lemma 8.24.** *If $\|E_{m-1}\|_{\rho_{m-1}}$ is small enough, then:*

- *If $DK_{m-1}^\top DK_{m-1}$ is invertible with inverse $N_{m-1}$ then $DK_m^\top DK_m$ is invertible with inverse $N_m$ and we have*

$$\|N_m\|_{\rho_m} \leqslant \|N_{m-1}\|_{\rho_{m-1}} + C_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|_{\rho_{m-1}}.$$

- *If* $\operatorname{avg}(S_{m-1})$ *is non-singular then also* $\operatorname{avg}(S_m)$ *is and we have the estimate*

$$\left|\left(\operatorname{avg}(S_m)\right)^{-1}\right| \leqslant \left|\left(\operatorname{avg}(S_{m-1})\right)^{-1}\right| + C'_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\|E_{m-1}\|_{\rho_{m-1}}.$$

The last lemma is devoted to the proof of the cohomology obstruction under the iterative step.

**Lemma 8.25.** *Assume $\|E_{m-1}\|_{\rho_{m-1}}$ is small enough. If $X_{\lambda_{m-1}}$ spans the cohomology of $K_{m-1}(\mathbb{T}^l)$ at $\lambda_{m-1}$, then $X_{\lambda_m}$ spans the cohomology of $K_m(\mathbb{T}^l)$ at $\lambda_m$.*

**Proof.** We have by assumption that the map

$$\frac{d}{d\lambda}\left[K_{m-1}^* i_{X_\lambda}\Omega\right]_{|\lambda=\lambda_{m-1}} : \mathbb{R}^l \to H^1\left(\mathbb{T}^l\right)$$

is an isomorphism. Thanks to the estimates on $\Delta_m$ and $|\lambda_m - \lambda_{m-1}|$ and the continuity of $X_\lambda$ and $DX_\lambda$ with respect to $\lambda$, we can write

$$\left\| \frac{d}{d\lambda}\left[K_m^* i_{X_\lambda}\Omega\right]_{|\lambda=\lambda_m} - \frac{d}{d\lambda}\left[K_{m-1}^* i_{X_\lambda}\Omega\right]_{|\lambda=\lambda_{m-1}} \right\|_{\rho-\delta} \leqslant C\kappa\delta^{-1}\|E_{m-1}\|_\rho.$$

The previous estimate comes from the identification of the cohomology with the integration over loops of $\mathbb{T}^l$ and the fact the quantity $\frac{d}{d\lambda}K^* i_{X_\lambda}\Omega$ is in matrix notation

$$DK(\theta)^\top J\left(K(\theta)\right)\left(\frac{\partial X_\lambda(K(\theta))}{\partial\lambda}\right).$$

This shows the invertibility of the map

$$\frac{d}{d\lambda}\left[K_m^* i_{X_\lambda}\Omega\right]_{|\lambda=\lambda_m}. \qquad \square$$

*8.8. Vanishing lemma*

This section is devoted to the proof of Theorem 8.10. First, recall that the Lie derivative of the 1-form $\alpha$ with respect to a vector-field $L$ is given by (Cartan formula)

$$\mathcal{L}_L\alpha = di_L\alpha + i_L d\alpha.$$

The following result is of general interest and is a vanishing lemma.

**Lemma 8.26.** *Assume $\omega \in D_h(\kappa, \nu)$ with $\kappa > 0$ and $\nu \geqslant l-1$, and $X_\lambda$ is a family of real analytic symplectic vector-fields. Let $K : D_\rho \supset \mathbb{T}^l \to \mathcal{M}$ be a solution of*

$$\partial_\omega K = X_\lambda \circ K + E, \tag{178}$$

*for $|\lambda - \lambda^*|$ small enough. Assume furthermore that:*

(1) *$X_{\lambda^*}$ is exact symplectic and for all $\lambda \in \mathbb{R}^l$ and $\lambda \neq \lambda^*$, the vector-fields $X_\lambda$ are symplectic but not exact symplectic.*
(2) *For all $\lambda \in \mathbb{R}^l$, $X_\lambda$ can be extended holomorphically to a complex neighborhood of $K(D_\rho)$.*
(3) *The family $X_\lambda$ spans the cohomology of $K(\mathbb{T}^l)$ at $\lambda = \lambda^*$, i.e. the map*

$$\begin{aligned} &\mathbb{R}^l \to H^1(\mathbb{T}^l) \\ &v \mapsto \frac{d}{d\lambda}[K^* i_{X_\lambda}\Omega]_{|\lambda=\lambda^*} v \end{aligned} \tag{179}$$

*is an isomorphism.*

*Then there exists a constant $C$ such that*

$$|\lambda - \lambda^*| \leqslant C\|E\|_\rho.$$

**Proof.** The proof is very similar to the proof of Lemma 4.9. Indeed, if we consider vector-fields as *"infinitesimal"* diffeomorphisms, the present proof can be considered as an infinitesimal version of the proof of Lemma 4.9. We define $\sigma_{i,\hat{\theta}_i}$ as in (34), (35).

The proof consists of computing

$$\int_{K\circ\sigma_{i,\hat{\theta}_i}} \mathcal{L}_{X_\lambda}\alpha \tag{180}$$

in two different ways. First, notice that by Cartan's formula, we have

$$\int_{K\circ\sigma_{i,\hat{\theta}_i}} \mathcal{L}_{X_\lambda}\alpha = \int_{K\circ\sigma_{i,\hat{\theta}_i}} i_{X_\lambda}\Omega.$$

Expanding this last expression in terms of $\lambda$ yields

$$\int_{K\circ\sigma_{i,\hat{\theta}_i}} i_{X_\lambda}\Omega = \int_{K\circ\sigma_{i,\hat{\theta}_i}} i_{X^*_\lambda}\Omega + \int_{K\circ\sigma_{i,\hat{\theta}_i}} \frac{d}{d\lambda}(i_{X_\lambda}\Omega)_{|\lambda=\lambda^*}(\lambda-\lambda^*) + O(|\lambda-\lambda^*|^2).$$

Furthermore, since the vector-field $X_{\lambda^*}$ is exact symplectic, we have $\mathcal{L}_{X^*_\lambda}\alpha = dW$ and then the first term in the right-hand side vanishes. We are led to

$$\int_{K\circ\sigma_{i,\hat{\theta}_i}} i_{X_\lambda}\Omega = \int_{K\circ\sigma_{i,\hat{\theta}_i}} \frac{d}{d\lambda}(i_{X_\lambda}\Omega)_{|\lambda=\lambda^*}(\lambda-\lambda^*) + O(|\lambda-\lambda^*|^2). \tag{181}$$

On the other hand, using the linearity of the Lie derivative w.r.t. the vector-field and Eq. (178), we have

$$\int_{K\circ\sigma_{i,\hat{\theta}_i}} \mathcal{L}_{X_\lambda}\alpha = \int_{K\circ\sigma_{i,\hat{\theta}_i}} i_{X_\lambda}\Omega = \int_{K\circ\sigma_{i,\hat{\theta}_i}} i_{\partial_\omega}\Omega + R,$$

where $R$ is such that $\|R\|_\rho \leqslant C\|E\|$.

Furthermore, by the change of variables formula and the exact symplecticness of the manifold we have

$$\int\limits_{K\circ\sigma_{i,\hat{\theta}_i}} i_{\partial_\omega}\Omega = \int\limits_{\sigma_{i,\hat{\theta}_i}} i_\omega K^*\Omega = \int\limits_{\sigma_{i,\hat{\theta}_i}} i_\omega \, dK^*\alpha.$$

Since $\omega$ is constant, the exterior differentiation commutes with the contraction operator and one gets for all $1 \leqslant i \leqslant l$

$$\int\limits_{K\circ\sigma_{i,\hat{\theta}_i}} i_{\partial_\omega}\Omega = \int\limits_{\sigma_{i,\hat{\theta}_i}} di_\omega K^*\alpha = 0,$$

yielding

$$\int\limits_{K\circ\sigma_{i,\hat{\theta}_i}} \mathcal{L}_{X_\lambda}\alpha = R. \tag{182}$$

We note that $i$-component of the map (179) is the integral of

$$\xi \mapsto \frac{d}{d\lambda}(K^* i_{X_\lambda}\Omega)_{|\lambda=\lambda^*}\xi$$

over the $i$th generator of the torus. Then, from (181)–(182) and the implicit function theorem, we get the desired result if $|\lambda - \lambda^*|$ is small. $\square$

We are now in position to prove Theorem 8.10.

**Proof of Theorem 8.10.** We have to introduce a family of vector-fields $X_\lambda$ satisfying the non-degeneracy condition (179) in Lemma 8.26. Since $\Omega$ is non-degenerate, given a family of closed 1-forms $\sigma_\lambda$ such that $\sigma_0 = 0$ and an exact symplectic vector-field $X$ there exists a family of symplectic vector-fields $X_\lambda$ such that

(1) $X_0 = X$.
(2) $i_{X_\lambda}\Omega = \sigma_\lambda$.

Condition (2) implies that $X_\lambda$ is indeed symplectic:

$$\mathcal{L}_{X_\lambda}\Omega = di_{X_\lambda}\Omega + i_{X_\lambda}d\Omega = d\sigma_\lambda = 0.$$

If we choose $\sigma_\lambda$ such that the cohomology class $[\sigma_\lambda] \neq 0$ for $\lambda \neq 0$ then $X_\lambda$ will not be exact symplectic for $\lambda \neq 0$. Indeed, this follows from the calculation

$$\mathcal{L}_{X_\lambda}\alpha = di_{X_\lambda}\alpha + i_{X_\lambda}d\alpha = di_{X_\lambda}\alpha + i_{X_\lambda}\Omega = dW_\lambda + \sigma_\lambda.$$

To choose $\sigma_\lambda$ consider the torus $K(\mathbb{T}^l)$. We take a tubular neighborhood $N^\varepsilon(K(\mathbb{T}^l))$ of $K(\mathbb{T}^l)$. Since it is contractible to $K(\mathbb{T}^l)$, then $H^1(K(\mathbb{T}^l)) \sim H^1(N^\varepsilon(K(\mathbb{T}^l)))$. Now we consider a basis $\{\delta_j\}_{1\leqslant j\leqslant l}$ of $H^1(K(\mathbb{T}^l))$. We define

$$\sigma_\lambda = \sum_{i=1}^{l} \lambda_i \delta_i.$$

Then we have

$$\frac{d}{d\lambda_j} K^* i_{X_\lambda} \Omega = \frac{d}{d\lambda_j} K^* \left( \sum_{i=1}^{l} \lambda_i \delta_i \right) = K^*(\delta_j).$$

Since $K$ is an embedding, $\{K^*\delta_j\}_{1 \leqslant j \leqslant l}$ is a basis of $H^1(\mathbb{T}^l)$ and then the map $v \mapsto D_\lambda(K^* i_{X_\lambda} \Omega)v$ is invertible. $\square$

## 9. Finite-dimensional Hamiltonian flows

This section is devoted to the application of our method in the Hamiltonian vector-field case. In the same spirit as the previous section, one of the motivations is the study of Hamiltonian PDEs.

The result for Hamiltonian flows is based on the study of the equation

$$\partial_\omega K(\theta) = J\big(K(\theta)\big) \nabla H\big(K(\theta)\big), \tag{183}$$

where the function $H : \mathcal{M} \to \mathbb{R}$ is the Hamiltonian which is supposed to be real analytic.

Eq. (183) expresses the invariance of the range of $K$ under the Hamiltonian vector-field $X_H = J\nabla H$. We assume that $\mathcal{M}$ is endowed with $\Omega = dx \wedge dy$ and $\alpha = -y\, dx$ and hence $J$ is constant. Note that the vector-field $J\nabla H$ is exact symplectic. Indeed, by definition, we have

$$i_{J\nabla H} \Omega = -dH.$$

Then taking $W = -H + i_{J\nabla H}\alpha$, we have $\mathcal{L}_X \alpha = dW$. More generally, if we consider an exact symplectic vector-field $X$ in the sense of Definition 8.1, then there exists a function $H$ such that $X = J\nabla H$.

Consequently, the Hamiltonian framework fits exactly in the exact case as described in the previous section (due to the lack of cohomology obstruction). However since equations of the type (183) occur in a lot of physical contexts, our motivation to write this section is to provide the formulas showing up for this type of system.

Again, the linearized equation

$$\frac{d\Delta}{dt} = JD\nabla H\big(K(\theta + \omega t)\big)\Delta \tag{184}$$

plays a crucial role. Since $A(\theta) \equiv JD\nabla H(K(\theta))$ is bounded, Eq. (184) admits an evolution operator, denoted $U_\theta(t)$. We have

$$\frac{d}{dt} U_\theta(t) = A(\theta + \omega t) U_\theta(t),$$

and $U_\theta(0) = \mathrm{Id}$. We now have the following definitions.

**Condition 9.1.**

- Spectral conditions: *The evolution operator $U_\theta(t)$ satisfies the non-degeneracy Condition* 8.4.
- Twist condition: *Let $N(\theta) = [DK(\theta)^\top DK(\theta)]^{-1}$ and $P(\theta) = DK(\theta)N(\theta)$. The average on $\mathbb{T}^l$ of the matrix*

$$S(\theta) = N(\theta) DK(\theta)^\top \big[A(\theta)J - JA(\theta)\big] DK(\theta) N(\theta)$$

*is non-singular. Here $A(\theta) = JD\nabla H(K(\theta))$.*

For the sake of completeness, we state a theorem for Eq. (183). It provides the existence of invariant tori.

**Remark 9.2.** To obtain the expression for $S$, we used the fact that $DK^\top JDK = 0$ and $J^{-1} = -J$.

**Theorem 9.3.** *Let $\omega$ satisfy the Diophantine condition given by Definition* 2.2. *Assume the following hypotheses*:

- *The embedding $K_0$ satisfies the non-degeneracy Condition* 9.1.
- *The map $H$ is real analytic and it can be extended holomorphically to some complex neighborhood of the image under $K_0$ of $D_{\rho_0}$*:

$$B_r = \{ z \in \mathbb{C}^{2d} \mid \exists \theta \in \{ |\operatorname{Im}\theta| < \rho_0 \} \text{ s.t. } |z - K_0(\theta)g| < r \},$$

*for some $r > 0$.*

*Define the error $E_0$ by*

$$E_0 = \partial_\omega K_0(\theta) - J\nabla H(K_0(\theta)).$$

*There exists a constant $C > 0$ depending on $l$, $\kappa$, $\nu$, $|H|_{C^3(B_r)}$, $\|DK_0\|_{\rho_0}$, $\|N_0\|_{\rho_0}$, $\|S_0\|_{\rho_0}$, $|(\operatorname{avg}(S_0))^{-1}|$ (where $S_0$ and $N_0$ are as in Condition* 9.1 *replacing $K$ by $K_0$) and the norms of the projections $\|\Pi^{c,s,u}_{K_0(\theta)}\|_{\rho_0}$ such that, if $E_0$ satisfies the estimates*

$$C\kappa^4\delta^{-4\nu}\|E_0\|_{\rho_0} < 1$$

*and*

$$C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0} < r,$$

*where $0 < \delta \leqslant \min(1, \rho_0/12)$ is fixed, then there exists an embedding $K_\infty$ such that $K_\infty \in ND(\rho_\infty := \rho_0 - 6\delta)$ and*

$$\partial_\omega K_\infty(\theta) = J\nabla H(K_\infty(\theta)). \tag{185}$$

*Furthermore, we have the estimate*

$$\|K_\infty - K_0\|_{\rho_\infty} \leqslant C\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0}.$$

**Remark 9.4.** One could also formulate a local uniqueness result in the case of vector-fields.

## Acknowledgments

E.F. acknowledges the support of the Spanish Grant MEC-FEDER MTM2006-05849/Consolider and the Catalan grant CIRIT 2005 SGR01028 and the hospitality of the University of Texas at Austin. R.L. has been supported by NSF grants. Several visits of R.L. to Barcelona supported by MCyT-FEDER grant MTN2006-00478 were useful for this project. Y.S. would like to thank the hospitality of the Department of Mathematics of the University of Texas at Austin and the Departament de Matemàtica Aplicada i Anàlisi of the Universitat de Barcelona.

## References


[1] V.I. Arnol'd, Proof of a theorem of A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations, Russian Math. Surveys 18 (5) (1963) 9–36.
[2] V.I. Arnol'd, Mathematical Methods of Classical Mechanics, second ed., Springer-Verlag, New York, 1989.
[3] Xavier Cabré, Ernest Fontich, Rafael de la Llave, The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces, Indiana Univ. Math. J. 52 (2) (2003) 283–328.
[4] Chong Qing Cheng, Lower-dimensional invariant tori in the regions of instability for nearly integrable Hamiltonian systems, Comm. Math. Phys. 203 (2) (1999) 385–419.
[5] Chong Qing Cheng, Shaoli Wang, The surviving of lower-dimensional tori from a resonant torus of Hamiltonian systems, J. Differential Equations 155 (2) (1999) 311–326.
[6] Amadeu Delshams, Rafael de la Llave, Tere M. Seara, A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by a potential of generic geodesic flows of $\mathbf{T}^2$, Comm. Math. Phys. 209 (2) (2000) 353–392.
[7] Amadeu Delshams, Rafael de la Llave, Tere M. Seara, A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: Announcement of results, Electron. Res. Announc. Amer. Math. Soc. 9 (2003) 125–134 (electronic).
[8] Amadeu Delshams, Rafael de la Llave, Tere M. Seara, A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: Heuristics and rigorous verification on a model, Mem. Amer. Math. Soc. 179 (844) (2006) 1–141.
[9] Amadeu Delshams, Rafael de la Llave, Tere M. Seara, Geometric properties of the scattering map of a normally hyperbolic invariant manifold, Adv. Math. 217 (3) (2008) 1096–1153.
[10] Amadeu Delshams, Gemma Huguet, Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems, MP_ARC 08-227, 2008.
[11] R. de la Llave, Persistence of normally hyperbolic manifolds: II, preprint, 2001.
[12] Rafael de la Llave, Remarks on Sobolev regularity in Anosov systems, Ergodic Theory Dynam. Systems 21 (4) (2001) 1139–1180.
[13] Rafael de la Llave, A tutorial on KAM theory, in: Smooth Ergodic Theory and Its Applications, Seattle, WA, 1999, Amer. Math. Soc., Providence, RI, 2001, pp. 175–292.
[14] R. de la Llave, A. González, À. Jorba, J. Villanueva, KAM theory without action-angle variables, Nonlinearity 18 (2) (2005) 855–895.
[15] R. de la Llave, R. Obaya, Decomposition theorems for groups of diffeomorphisms in the sphere, Trans. Amer. Math. Soc. 352 (3) (2000) 1005–1020.
[16] R. de la Llave, Y. Sire, Quasi-periodic solutions for PDEs, 2007, manuscript.
[17] R. de la Llave, C.E. Wayne, Whiskered and low dimensional tori in nearly integrable Hamiltonian systems, Math. Phys. Electron. J. 10 (Paper 5) (2004), 45 pp. (electronic).
[18] R. Douady, Une démonstration directe de l'équivalence des théorèmes de tores invariants pour difféomorphismes et champs de vecteurs, C. R. Acad. Sci. Paris Sér. I Math. 295 (2) (1982) 201–204.
[19] L.H. Eliasson, Hamiltonian systems with linear normal form near an invariant torus, in: Nonlinear Dynamics, Bologna, 1988, World Scientific, Teaneck, NJ, 1989, pp. 11–29.
[20] L.H. Eliasson, Biasymptotic solutions of perturbed integrable Hamiltonian systems, Bull. Braz. Math. Soc. (N.S.) 25 (1) (1994) 57–76.
[21] L.H. Eliasson, Almost reducibility of linear quasi-periodic systems, in: Smooth Ergodic Theory and Its Applications, Seattle, WA, 1999, in: Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 679–705.
[22] E. Fontich, R. de la Llave, Y. Sire, Construction of invariant whiskered tori by a parameterization method, part II: Quasi-periodic breathers in coupled map lattices, 2008, manuscript.
[23] Neil Fenichel, Persistence and smoothness of invariant manifolds for flows, Indiana Univ. Math. J. 21 (1971/1972) 193–226.
[24] N. Fenichel, Asymptotic stability with rate conditions, Indiana Univ. Math. J. 23 (1973/74) 1109–1137.
[25] E. Fermi, J. Pasta, S. Ulam, Studies on nonlinear problems, Document LA 1940, 1955.
[26] G. Gallavotti, G. Gentile, Hyperbolic low-dimensional invariant tori and summations of divergent series, Comm. Math. Phys. 227 (3) (2002) 421–460.
[27] Samuel M. Graff, On the conservation of hyperbolic invariant tori for Hamiltonian systems, J. Differential Equations 15 (1974) 1–69.
[28] Victor Guillemin, Shlomo Sternberg, Geometric Asymptotics, Math. Surveys, vol. 14, Amer. Math. Soc., Providence, RI, 1977.
[29] Robert E. Gompf, András I. Stipsicz, 4-Manifolds and Kirby Calculus, Grad. Stud. Math., vol. 20, Amer. Math. Soc., Providence, RI, 1999.
[30] Allen Hatcher, Algebraic Topology, Cambridge Univ. Press, Cambridge, 2002.
[31] A. Haro, R. de la Llave, New mechanisms for lack of equipartition of energy, Phys. Rev. Lett. 85 (9) (2000) 1859–1862.
[32] À. Haro, R. de la Llave, A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: Numerical algorithms, Discrete Contin. Dyn. Syst. Ser. B 6 (6) (2006) 1261–1300 (electronic).
[33] A. Haro, R. de la Llave, A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: Rigorous results, J. Differential Equations 228 (2) (2006) 530–579.
[34] A. Haro, R. de la Llave, Spectral theory and dynamical systems, ftp://ftp.ma.utexas.edu/pub/papers/llave/spectrum-dynamics.pdf, 2006.
[35] A. Haro, R. de la Llave, A parameterization method for the computation of whiskers in quasi periodic maps: Numerical implementation and examples, SIAM J. Appl. Dyn. Syst. 6 (1) (2007) 142–207.

[36] G. Huguet, R. de la Llave, Y. Sire, Fast numerical methods for the computation of quasi-periodic solutions of Hamiltonian systems, MP_ARC 09-02, 2009.
[37] Yuecai Han, Yong Li, Yingfei Yi, Degenerate lower-dimensional tori in Hamiltonian systems, J. Differential Equations 227 (2) (2006) 670–691.
[38] M.W. Hirsch, C.C. Pugh, Stable manifolds and hyperbolic sets, in: Global Analysis, Proc. Sympos. Pure Math., vol. XIV, Berkeley, CA, 1968, Amer. Math. Soc., Providence, RI, 1970, pp. 133–163.
[39] M. Hirsch, J. Palis, C. Pugh, M. Shub, Neighborhoods of hyperbolic sets, Invent. Math. 9 (1969) 121–134.
[40] M.W. Hirsch, C.C. Pugh, M. Shub, Invariant Manifolds, Lecture Notes in Math., vol. 583, Springer-Verlag, Berlin, 1977.
[41] Dale Husemoller, Fibre Bundles, third ed., Grad. Texts in Math., vol. 20, Springer-Verlag, New York, 1994.
[42] À. Jorba, R. de la Llave, M. Zou, Lindstedt series for lower-dimensional tori, in: Hamiltonian Systems with Three or More Degrees of Freedom, S'Agaró, 1995, Kluwer Acad. Publ., Dordrecht, 1999, pp. 151–167.
[43] Russell A. Johnson, Analyticity of spectral subbundles, J. Differential Equations 35 (3) (1980) 366–387.
[44] W.T. Kyner, Rigorous and Formal Stability of Orbits about an Oblate Planet, Mem. Amer. Math. Soc., vol. 81, Amer. Math. Soc., Providence, RI, 1968.
[45] Yong Li, Yingfei Yi, Persistence of hyperbolic tori in Hamiltonian systems, J. Differential Equations 208 (2) (2005) 344–387.
[46] Yong Li, Yingfei Yi, Persistence of lower dimensional tori of general types in Hamiltonian systems, Trans. Amer. Math. Soc. 357 (4) (2005) 1565–1600 (electronic).
[47] J.J. Masdemont, High-order expansions of invariant manifolds of libration point orbits with applications to mission design, Dyn. Syst. 20 (1) (2005) 59–113.
[48] Richard Moeckel, Transition tori in the five-body problem, J. Differential Equations 129 (2) (1996) 290–314.
[49] J. Moser, On invariant curves of area-preserving mappings of an annulus, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II 1962 (1962) 1–20.
[50] J. Moser, A rapidly convergent iteration method and non-linear differential equations. II, Ann. Sc. Norm. Super. Pisa (3) 20 (1966) 499–535.
[51] J. Moser, A rapidly convergent iteration method and non-linear partial differential equations. I, Ann. Sc. Norm. Super. Pisa (3) 20 (1966) 265–315.
[52] J. Moser, Convergent series expansions for quasi-periodic motions, Math. Ann. 169 (1967) 136–176.
[53] John W. Milnor, James D. Stasheff, Characteristic Classes, Ann. of Math. Stud., vol. 76, Princeton Univ. Press, Princeton, NJ, 1974.
[54] Rafael Ortega, Nonexistence of invariant curves of mappings with small twist, Nonlinearity 10 (1) (1997) 195–197.
[55] Rafael Ortega, Boundedness in a piecewise linear oscillator and a variant of the small twist theorem, Proc. London Math. Soc. (3) 79 (2) (1999) 381–413.
[56] Ya.B. Pesin, Lectures on partial hyperbolicity and stable ergodicity, in: Zürich Lectures in Advanced Mathematics, European Math. Soc. (EMS), Zürich, 2004.
[57] H. Poincaré, Les méthodes nouvelles de la mécanique céleste, vols. 1, 2, 3–Gauthier–Villars, Paris, 1892–1899.
[58] Victor A. Pliss, George R. Sell, Robustness of exponential dichotomies in infinite-dimensional dynamical systems, J. Dynam. Differential Equations 11 (3) (1999) 471–513.
[59] Walter Rudin, Real and Complex Analysis, second ed., McGraw–Hill Ser. in Higher Math., McGraw–Hill, New York, 1974.
[60] H. Rüssmann, On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus, in: Dynamical Systems, Theory and Applications, Battelle Rencontres, Seattle, WA, 1974, in: Lecture Notes in Phys., vol. 38, Springer-Verlag, Berlin, 1975, pp. 598–624.
[61] H. Rüssmann, On a new proof of Moser's twist mapping theorem, in: Proceedings of the Fifth Conference on Mathematical Methods in Celestial Mechanics, Part I, Oberwolfach, 1975, Celestial Mech. 14 (1976) 19–31.
[62] H. Rüssmann, On optimal estimates for the solutions of linear difference equations on the circle, Celestial Mech. 14 (1) (1976) 33–37.
[63] Dietmar A. Salamon, The Kolmogorov–Arnold–Moser theorem, Math. Phys. Electron. J. 10 (Paper 3) (2004), 37 pp. (electronic).
[64] M.B. Sevryuk, The lack-of-parameters problem in the KAM theory revisited, in: Hamiltonian Systems with Three or More Degrees of Freedom, S'Agaró, 1995, Kluwer Acad. Publ., Dordrecht, 1999, pp. 568–572.
[65] Mikhail B. Sevryuk, Partial preservation of frequencies in KAM theory, Nonlinearity 19 (5) (2006) 1099–1140.
[66] A. Sorrentino, Sulle soluzioni quasi-periodiche di sistemi Hamiltoniani differenziabili, PhD thesis, Univ. di Roma Tre, 2002.
[67] R.J. Sacker, G.R. Sell, Existence of dichotomies and invariant splittings for linear differential systems. I, J. Differential Equations 15 (1974) 429–458.
[68] Norman Steenrod, The Topology of Fibre Bundles, Princeton Math. Ser., vol. 14, Princeton Univ. Press, Princeton, NJ, 1951.
[69] D. Salamon, E. Zehnder, KAM theory in configuration space, Comment. Math. Helv. 64 (1) (1989) 84–132.
[70] D.V. Treschev, Hyperbolic tori and asymptotic surfaces in Hamiltonian systems, Russ. J. Math. Phys. 2 (1) (1994) 93–110.
[71] Enrico Valdinoci, Families of whiskered tori for a priori stable/unstable Hamiltonian systems and construction of unstable orbits, Math. Phys. Electron. J. 6 (Paper 2) (2000), 31 pp. (electronic).
[72] John A. Vano, A Nash–Moser implicit function theorem with Whitney regularity and applications, PhD thesis, Univ. of Texas at Austin, 2002, MP_ARC #02-276.
[73] E. Zehnder, Generalized implicit function theorems with applications to some small divisor problems. I, Comm. Pure Appl. Math. 28 (1975) 91–140.
[74] E. Zehnder, Generalized implicit function theorems with applications to some small divisor problems. II, Comm. Pure Appl. Math. 29 (1) (1976) 49–111.
[75] Wenzhuang Zhu, Baifeng Liu, Zhenxin Liu, The hyperbolic invariant tori of symplectic mappings, Nonlinear Anal. 68 (1) (2008) 109–126.