DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **32**, Number **4**, April **2012**

doi:10.3934/dcds.2012.32.1309

pp. **1309–1353**

COMPUTATION OF WHISKERED INVARIANT TORI AND THEIR ASSOCIATED MANIFOLDS: NEW FAST ALGORITHMS

Gemma Huguet

Center for Neural Science, New York University
New York, NY 10003, USA
and
Centre de Recerca Matemàtica
Apartat 50, 08193 Bellaterra (Barcelona), Spain

Rafael de la Llave

Department of Mathematics, The University of Texas at Austin
Austin, TX, 78712-1082, USA

Yannick Sire

Université Paul Cézanne, Laboratoire LATP UMR 6632
Marseille, France

(Communicated by Amadeu Delshams)

Abstract. We present efficient (low storage requirement and low operation count) algorithms for the computation of several invariant objects for Hamiltonian dynamics, namely KAM tori (i.e diffeomorphic copies of tori such that the motion on them is conjugated to a rigid rotation) both Lagrangian tori(of maximal dimension) and whiskered tori (i.e. tori with hyperbolic directions which, together with the tangents to the torus and the symplectic conjugates span the whole tangent space). We also present algorithms to compute the invariant splitting and the invariant manifolds of whiskered tori. We present the algorithms for both discrete-time dynamical systems and differential equations.

The algorithms do not require that the system is presented in action-angle variables nor that it is close to integrable and are backed up by rigorous *a-posteriori* bounds. We will report on the implementation results elsewhere.

1. **Introduction.** The goal of this paper is to present efficient algorithms to compute accurately several objects of interest in Hamiltonian dynamical systems (both discrete-time dynamical systems and differential equations). More precisely, we present algorithms to compute:

- Lagrangian KAM tori.
- Whiskered KAM tori.
- The invariant bundles of the whiskered tori.
- The stable and unstable manifolds of the whiskered tori.

2000 *Mathematics Subject Classification.* Primary: 70K43; Secondary: 37J40 .

Key words and phrases. Quasi-periodic solutions, whiskered KAM tori, invariant manifolds, numerical computation.

The work of G. H. and R. L. has been partially supported by NSF grants. G. H. has also been supported by the Spanish Grants MTM2006-00478, MTM2009-06973 and CUR-DIUE grant 2009SGR859 and the Spanish Fellowships AP2003-3411 and "Proyectos Flechados". R. L. was supported by NHARP 0223.

The algorithms are very different. For example, the algorithms for tori require the use of small divisors and symplectic geometry and the algorithms for invariant bundles and invariant manifolds rely on the theory of normal hyperbolicity and dichotomies. The computation of whiskered tori has to combine both.

We recall that KAM tori are manifolds diffeomorphic to a torus which are invariant for a map or flow, on which the motion of the system is conjugate to a rotation. As we will see later, this is also equivalent to quasi-periodic solutions. The tori are called Lagrangian when they are Lagrangian manifolds, which in our case amounts to the tori having a dimension equal to the number of degrees of freedom of the system. The tori are called whiskered when the linearized equation has directions that decrease exponentially either in the future (stable) or in the past (unstable) and these directions together with the tangent to the torus and its symplectic conjugate span the whole tangent space. These invariant spaces for the linearization have non-linear analogues, namely invariant manifolds. It has been recognized since [1] that whiskered tori and their invariant manifolds are very interesting landmarks that organize the long-term behavior of many systems.

The algorithms we present are based on running an *efficient* Newton method to solve a functional equation, which expresses the dynamical properties above. What we mean by *efficient* is that if we discretize the problem using N Fourier coefficients and N points in a grid, we require $O(N)$ storage and only $O(N \log(N))$ operations for the Newton step. Since the functions we are considering are analytic, we see that the truncation error is $O(\exp(-CN^{1/d}))$ where d is the dimension of the object. Note that, in contrast, a straightforward implementation of a Newton method would require to use $O(N^2)$ storage – to store the linearization matrix and its inverse – and $O(N^3)$ operations to invert.

In practical applications, using the algorithms described in this paper, computing with several million coefficients becomes quite practical in a typical desktop computer of today. Implementation details and the results of several runs will be discussed in another companion paper [25]. Given the characteristics of today's computers, savings in storage space are more crucial than savings in operations for these problems.

The algorithms we present here are inspired by the rigorous results of [36] – for KAM tori – and [16, 15] –for whiskered tori. The algorithms to compute the stable and unstable manifolds had not been previously discussed. The rigorous results of the above papers are also based on a Newton method applied to the same functional equation that we consider here.

Of course, going from a mathematical treatment to a practical algorithm requires making non-trivial choices of algorithms and specifying significantly many more details. Sometimes, perfectly good theoretical arguments are quite unsuitable for practical algorithms. In our case, for example, the algorithms to compute the invariant splittings and the invariant manifolds are different from those in the above references. The papers above use that there is a metric in which the splittings are orthogonal, so that the matrix representing the symplectic matrix, has a block structure. Even if this is perfectly correct for a theoretical paper, the algorithms require to know the metric explicitly. Hence, in this paper we present algorithms that do not require this assumption.

This paper concentrates on explaning the algorithmic issues but we will postpone for a future paper the discussion of practical issues such as data structures as well as the results of the runs.

The results of the papers [36, 15] give a justification of the algorithms for invariant tori and their invariant splittings presented here. The theorems in [36, 15], have been formulated in an *a-posteriori* way, i.e. the theorems assert that if we have a function which solves the invariance equation up to a small error (e.g. the outcome of a successful run of the algorithms presented here) and which also satisfies some explicit non-degeneracy conditions, then, we can conclude that there is a true solution which is close to the computed solution. Hence, by supplementing our calculations with the (very simple) computations of the non-degeneracy conditions (they play a role very similar to the condition numbers common in numerical analysis), we can be sure that the computation that we are performing is meaningful. This allows to compute with confidence even close to the limit of validity of the KAM theorem (a rather delicate boundary since the smooth KAM tori do not disappear completely but rather morph into Cantor sets).

This boundary of parameters between the region with smooth tori and Cantor sets, called *analyticity breakdown* has received a great deal of attention in the applied literature [2, 3, 5, 22, 39, 40, 7, 8]). Having an a-posteriri method that can distinguish with confidence between rather similar objects and in regions when spurious solutions abound has proved invaluable.

Since the papers [36, 15] contain estimates, in the present paper, we will only discuss the algorithmic issues. For example, we will detail how solutions of equations (whose existence was shown in the above papers) can be computed with small requirements of storage and small operation count. Note that different algorithms of a same mathematical operation can have widely different operation counts and storage requirements. (See, for example, the discussion in [35] on the different algorithms to multiply matrices, polynomials, etc.) On the other hand, we will not include some implementation issues (methods of storage of arrays, ordering of loops, precision, etc.) needed to obtain actual results in a real computer. They will be given in another paper together with experimental results obtained by running the algorithms.

One remarkable feature of the algorithms presented here is that they do not require the system to be close to integrable. We only need a good initial guess for the Newton method. Typically, one uses a continuation method starting from an integrable case, where solutions can be computed analytically. However, in the case of secondary KAM tori, which do not exist in the integrable case, one can use, for instance, Lindstedt series, variational methods or approximation by periodic orbits to obtain an initial guess.

As for the algorithms to compute invariant splittings, we depart from the standard mathematical methods (most of the them based on graph transforms) and we have found more efficient to device an equation for the invariant projections. We also present some acceleration of convergence methods that give superexponential convergence. They are based on fast algorithms to solve the cohomology equations which could be of independent interest (see Appendix A).

The algorithms to compute invariant manifolds are based on the parameterization method [9, 11]. Compared to standard methods such as the graph transform the parameterization method has the advantage that to compute geometric objects of dimension ℓ, we only need to compute functions of dimension ℓ. In contrast with [9, 11], which was based on contractive iterations, our method is based on a Newton iteration which we also implement without requiring large matrices and requiring only $N \log(N)$ operations.

An overview of the method. The numerical method we use is based on the parameterization methods introduced in [9, 10]. In this section, we provide a sketch of the issues, postponing some important details. We make the presentation for maps first and the details for flows will be done later, specially in Section 2.2.

Invariant tori. We observe that if F is a map and we can find an embedding K in which the motion on the torus is a rotation ω, it should satisfy the equation

$$F(K(\theta)) - K(\theta + \omega) = 0. \tag{1}$$

Given an approximate solution of (1), i.e.

$$F(K(\theta)) - K(\theta + \omega) = E(\theta),$$

the Newton method aims to find Δ solving

$$DF(K(\theta))\Delta(\theta) - \Delta(\theta + \omega) = -E(\theta), \tag{2}$$

so that $K + \Delta$ will be a better approximate solution.

The fact that the approximate solution is whiskered means that there is a splitting as in (26).

The main idea of the Newton method is that, using the assummed decomposition of the phase space into subspaces invariant for $DF \circ K$, one can decompose (2) into three components

$$\begin{aligned} DF(K(\theta))\Delta^s(\theta) - \Delta^s(\theta + \omega) &= -E^s(\theta) \\ DF(K(\theta))\Delta^u(\theta) - \Delta^u(\theta + \omega) &= -E^u(\theta) \\ DF(K(\theta))\Delta^c(\theta) - \Delta^c(\theta + \omega) &= -E^c(\theta) \end{aligned} \tag{3}$$

where the s, u refer to the stable, unstable components and c refers to the component along the tangent to the torus and its symplectic conjugate. For Lagrangian tori, only the E^c part appears in the equations.

The algorithm requires:

- Efficient methods to evaluate the LHS of (1).
- Efficient methods to compute the splitting.
- Efficient methods to solve the equations (3).

As we will see in Section 3.6, to evaluate (1), it is efficient to use both a Fourier representation which makes easy to evaluate $K(\theta + \omega)$ and a real representation which makes easy to evaluate $F(K(\theta))$. Of course, both of them are linked through the Fast Fourier Transform (FFT from now on).

The methods to compute the splitting are discussed in Section 4.4. More precisely, we present a numerical procedure to compute the projections on the linear stable/unstables subspaces based on a Newton method. In [25], we present an alternative procedure for the computation of the projections based on the calculation of invariant bundles for cocycles which is also of interest to other problems. We note that these methods to compute invariant splittings do not use symplectic geometry and are applicable to any dynamical system.

The solution of the hyperbolic components in equation (3) is discussed in Section 4.2 and Appendix A. Indeed, equations of this form appear as well in the calculation of the invariant splitting discussed in Section 4.4. A first method is based on an acceleration of the fixed point iteration (Appendix A.1). We note that to obtain superexponential convergence for the solution of (3), we need to use both the Fourier representation and the real space representation.

In the case that the bundles are one-dimensional, there is yet another algorithm, which is even faster than the previous ones (see Appendix A.2). The algorithms are discussed for maps, and they do not have an easy analog for flows except by passing through the integration of differential equations. We think that this is one case where working with time-1 maps is advantageous.

The most challenging step is the solution of the center component of (3). This depends on cancelations which use the symplectic structure, involves small divisors and requires that certain obstructions vanish. Using several geometric identities that take advantage of the fact that the map is symplectic (see Section 4.3), the solution of (3) in the center direction is reduced to solving the following equation for φ given η,

$$\varphi(\theta) - \varphi(\theta + \omega) = \eta(\theta). \tag{4}$$

Equation (4) can be readily solved using Fourier coefficients provided that $\int \eta = 0$ (and that ω is sufficiently irrational). The solution is unique up to addition of a constant.

The existence of obstructions – which are finite dimensional – is one of the main complications of the problem. It is possible to show that, when the map F is exact symplectic, the obstructions for the solution are $O(||E||^2)$. An alternative method to deal with these obstructions is to add some new – finite dimensional – unknowns λ and solve, instead of (1), the equation

$$F(K(\theta)) - K(\theta + \omega) + G(\theta + \omega)\lambda = 0$$

where $G(\theta)$ is an explicit function. Even if λ is kept through the iteration involving approximate solutions, it can be shown that, if the map is exact symplectic, we have $\lambda = 0$. This counterterm approach also helps to weaken the non-degeneracy assumptions.

A minor issue is that the solutions of (1) are not unique. If K is a solution, $\tilde{K}$ defined by $\tilde{K}(\theta) = K(\theta + \sigma)$ is also a solution for any $\sigma \in \mathbb{R}^\ell$. This can be easily solved by taking an appropriate normalization that fixes the origin of coordinates in the torus. In [15] it is shown that this is the only non-uniqueness phenomenon of the equation. Furthermore, this local uniqueness property allows to deduce results for vector fields from the results for maps. For the algorithms, uniquess is not as crucial if we can prove that they converge. Nevertheless, if one tries to use a Newton method, one needs to make sure that one does not try to find an inverse when there is none. One of the advantage of the algorithms presented here is that, in contrast with the straightforward Newton method, they do not try to produce an inverse of the derivative, but rather produce a left inverse.

It is important to remark that the algorithms that we will present can compute in a unified way both primary and secondary tori. We recall here that *secondary tori* are invariant tori which are contractible to a torus of lower dimension, whereas this is not the case for primary tori. The tori which appear in integrable systems in action-angle variables are always primary. In quasi-integrable systems, the tori which appear through Lindstedt series or other perturbative expansions starting from those of the integrable system are always primary. Secondary tori, however, are generated by resonances. In numerical explorations, secondary tori are very prominent features that have been called "islands". In [28], one can find arguments showing that these solutions are very abundant in systems of coupled oscillators. As an example of the importance of secondary tori we will mention that in the recent paper [13] they constituted the essential object to overcome the "large gap

problem" and prove the existence of diffusion. In [12], one can find a detailed analysis of secondary tori.

In this paper, we will mainly discuss algorithms for systems with dynamics described by diffeomorphisms. For systems described through vector fields, we note that, taking surfaces of section in the energy surface, we can reduce the problem with vector fields to a problem with diffeomorphisms. However, in some practical applications, it may be convenient to have a direct treatment of the system described by vector fields. For this reason, present the invariance equations for flows in parallel with the discussion for maps, see Section 2.2. In Appendix B we present algorithms for flows.

Invariant manifolds attached to invariant tori. When the torus is a whiskered torus, it has invariant manifolds attached to it. For simplicity, in this presentation we will discuss the case of one dimensional directions – even if the torus can be of higher dimension.

We use again a parameterization method. Consider an embedding W in which the motion on the torus is a rotation ω and the motion on the stable (unstable) whisker consists of a contraction (expansion) at rate μ. The embedding W has to satisfy the invariance equation

$$F(W(\theta, s)) - W(\theta + \omega, \mu s) = 0. \tag{5}$$

Again, the key point is that taking advantage of the geometry of the problem we can devise algorithms which implement a Newton step to solve equation (5) without having to store—and much less invert—a large matrix. We first discuss the so-called order by order method, which serves as a comparison with more efficient methods based on the reducibility. Although the methods for whikers are based on the same idea as for the case of tori, they have not been introduced previously and constitute one of the main novelties of this paper. We present algorithms that given a torus and the associated linear spaces compute the invariant manifold tangent to it. It is clearly possible to extend the method to compute stable and unstable manifolds in general dimensions (or even non-resonant bundles). To avoid increasing the length of this paper and since higher dimensional examples are harder numerically, we postpone this to a future paper.

Some remarks on the literature. Invariant tori in Hamiltonian dynamics have been recognized as important landmarks in Hamiltonian dynamics. In the case of whiskered tori, their manifolds have also been crucial for the study of Arnold diffusion.

Since the mathematical literature is so vast, we cannot hope to summarize it here. We refer to the rather extensive references of [37] for Lagrangian tori and those of [15] for whiskered tori. We will just briefly mention that [21, 45] the earliest references on whiskered tori, as well as most of the later references, are based on *transformation theory*, that is making changes of variables that reduce the perturbed Hamiltonian to a simple form which obviously presents the invariant torus. From the point of view of numerics, this has the disadvantage that transformations are very hard to implement.

The numerical literature is not as broad as the rigorous one, but it is still quite extensive. The papers closest to our problems are [31, 30, 32], which consider tori of systems under quasi-periodic perturbation. These papers also contain a rather wide bibliography on papers devoted to numerical computation of invariant

circles. Among the papers not included in the references of the papers above because these appeared later, we mention [6], which presents other algorithms which apply to variational problems (even if they do not have a Hamiltonian interpretation). Another fast method is the *fractional iteration method* [44]. Note that the problems considerd in [31, 30, 32] do not involve center directions (and hence, do not deal with small divisors) and that the frequency and one of the coordinates of the torus is given by the external perturbation. The methods of [31, 30, 32] work even if the system is not symplectic (even if they can take advantage of the symplectic structure).

The papers [33, 34] present and implement calculations of *reducible* tori. This includes tori with normally elliptic directions. The use of reducibility indeed leads to very fast Newton steps, but it still requires the storage of a matrix for the changes of variables (it can still be $O(N)$ in the space discretization, even if it is $O(N^2)i$ in Fourier space). As seen in the examples in [32, 29], reducibility may fail in a codimension 1 set even for hyperbolic systems (a Cantor set of codimension 1 manifolds for elliptic tori in Hamiltonian systems) even if the geometric objetcs persist. For these reasons, we will not discuss methods based on reducibility in this paper (even if it is a useful and practical tool) and just refer to the references just mentioned. The solutions of the cohomology equations can be obtained using the hyperbolicity and not the reducibility. In this paper we also present algorithms to accelerate the convergence of the methods based on hyperbolicity so that they are as fast as the methods based on reducibility. See Appendix A.

The paper is organized as follows. In Section 2 we summarize the notions of mechanics and symplectic geometry we will use. In Section 3 we formulate the invariance equations for the objects of interest (invariant tori, invariant bundles and invariant manifolds) and we will present some generalities about the numerical algorithms.

Algorithms for whiskered tori are discussed in Section 4. In particular, we discuss how to compute the decomposition (3) of the linearized equation (2), and how to solve efficiently each equation in (3).

In Section 5 we discuss fast algorithms to compute rank-1 (un)stable manifolds of whiskered tori. More precisely, we present an efficient Newton method to solve equation (5).

In Appendix A one can find the fast algorithms to solve cohomology equations with non-constant coefficients that will be used in the computation of the splitting (3) as well as to solve the hyperbolic components of equations (3). In Appendices B-D, one can find the algorithms specially designed for flows, analogous to the ones for maps.

2. Setup and conventions.

We will be working with systems defined on an Euclidean phase space endowed with a symplectic structure. The phase space under consideration will be

$$\mathcal{M} \subset \mathbb{R}^{2d-\ell} \times \mathbb{T}^{\ell}.$$

We do not assume that the coordinates in the phase space are action-angle variables. Indeed, there are several systems (even quasi-integrable ones) which are very smooth in Cartesian coordinates but less smooth in action-angle variables (e.g., neighborhoods of elliptic fixed points [17, 20], hydrogen atoms in crossed electric and magnetic fields [41, 42] and several problems in celestial mechanics [4]).

We will assume that the Euclidean manifold $\mathcal{M}$ is endowed with an exact symplectic structure $\Omega = d\alpha$ (for some one-form α) and we have

$$\Omega_z(u, v) = \langle u, J(z)v\rangle, \tag{6}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product on the tangent space of $\mathcal{M}$ and $J(z)$ is a skew-symmetric matrix ($J(z) = -J(z)^\top$).

Remark 1. To associate a matrix to a symplectic form we assume a metric on $\mathcal{M}$. Some of the formulas in [16] assume that the metric is such that the center and the stable and unstable subspaces are orthogonal. Of course, such a metric indeed exists, but we found that it is algorithmically convenient to develop formulas that do not assume that. Hence, we will also have formulas that do not corespond to the formulas in [16].

An important particular case is when J induces an almost-complex structure, i.e.

$$J^2 = -\operatorname{Id}. \tag{7}$$

Most of our calculations do not need this assumption. One important case, where the identity (7) is not satisfied, is when J is a symplectic structure on surfaces of section chosen arbitrarily in the energy surface or when J is the symplectic form expressed in symplectic polar coordinates near an elliptic fixed point. When (7) holds, some calculations can be made faster.

As previously mentioned, we will be considering systems described either by diffeomorphisms or by vector-fields.

2.1. **Systems described by diffeomorphisms.** We will consider maps $F : \mathcal{U} \subset \mathcal{M} \mapsto \mathcal{M}$ which are not only symplectic (i.e. $F^*\Omega = \Omega$, where F^* denotes the pullback by F) but exact symplectic, that is

$$F^*\alpha = \alpha + dP,$$

for some smooth function P, called the *primitive function*.

We will also need Diophantine properties for the frequencies of the torus. For the case of maps, the useful notion of a Diophantine frequency is:

$$\mathcal{D}(\nu,\tau) = \left\{\omega \in \mathbb{R}^\ell \;\middle|\; |\omega\cdot k - n|^{-1} \le \nu|k|^\tau \;\forall\; k \in \mathbb{Z}^\ell - \{0\},\; n \in \mathbb{Z}\right\},\quad \nu > \ell.$$

2.2. **Systems described by vector fields.** We will assume that the system is described by a globally Hamiltonian vector-field X, that is

$$X = J\nabla H,$$

where H is a globally defined function on $T^*\mathcal{M}$.

In the case of flows, the appropriate notion of Diophantine numbers is:

$$\mathcal{D}^{\text{aff}}(\nu,\tau) = \left\{\omega \in \mathbb{R}^\ell \;\middle|\; |\omega\cdot k|^{-1} \le \nu|k|^\tau \;\forall\; k \in \mathbb{Z}^\ell - \{0\}\right\},\quad \nu \ge \ell - 1$$

Remark 2. It is well known that for non-Diophantine frequencies substantially complicated behavior can appear [27, 19]. Observing convincingly Liouvillian behaviors seems a very ambitious challenge for numerical exploration.

3. **Equations for invariance.** In this section, we discuss the functional equations for the objects of interest, that is, the invariant tori and the associated whiskers. These functional equations, which describe the invariance of the objects under consideration, are the cornerstone of the algorithms.

3.1. **Functional equations for whiskered invariant tori for diffeomorphisms.** At least at the formal level, it is natural to search quasi-periodic solutions with frequency ω (independent over the integers) under the form of Fourier series

$$x^{(n)} = \sum_{k \in \mathbb{Z}^\ell} \hat{x}_k e^{2\pi i k \cdot \omega n} , \tag{8}$$

where $\omega \in \mathbb{R}^\ell$ and $n \in \mathbb{Z}$.

We allow some components of x in (8) to be angles. In that case, it suffices to take some of the components of x modulo 1.

It is then natural to describe a quasi-periodic function using the so-called "hull" function $K : \mathbb{T}^\ell \to \mathcal{M}$ defined by

$$K(\theta) = \sum_{k \in \mathbb{Z}^\ell} \hat{x}_k e^{2\pi i k \cdot \theta},$$

so that we can write

$$x^{(n)} = K(n\omega).$$

The geometric interpretation of the hull function is that it gives an embedding from $\mathbb{T}^\ell$ into the phase space. In our applications, the embedding will actually be an immersion.

It is clear that quasi-periodic functions will be orbits for a map F if and only if the hull function K satisfies:

$$F \circ K - K \circ T_\omega = 0, \tag{9}$$

where T_ω denotes a rigid rotation

$$T_\omega(\theta) = \theta + \omega. \tag{10}$$

A modification of the invariance equations (9) which we will be important for our purposes is considering

$$F \circ K - K \circ T_\omega - G \circ T_\omega \lambda = 0, \tag{11}$$

where the unknowns are now $K : \mathbb{T}^\ell \to \mathcal{M}$ (as before) and $\lambda \in \mathbb{R}^\ell$. Here, G is a function of θ taking values in $2d \times \ell$ matrices, such that translations along the direction of G move the constant term in the center directions. In the case that we have chosen a metric on $\mathcal{M}$ (to associate a matrix to a 2-form) such that the invariant subspaces are orthogonal, we can choose a matrix G of the form

$$G(\theta) = (J \circ K_0)^{-1} DK_0,$$

where K_0 denotes a given approximate (in a suitable sense which will be given below) solution of the equation (9).

It has been shown in [16, 15] (see the *vanishing lemma*) that, for exact symplectic maps, if (K, λ) satisfy the equation (11) with K_0 close to K, then at the end of the iteration of the Newton method, we have $\lambda = 0$ and, therefore, K is a solution of the invariance equation (9). In other words, the formulations (11) and (9) are equivalent. Of course, for approximate solutions of the invariance equation (9), there is no reason why λ should vanish and it is numerically advantageous to solve the equation with the extra variable λ.

The advantage of equation (11) for numerical calculations is that, at the initial stages of the method, when the error in the invariance equation is large, it is not easy to ensure that certain compatibility conditions are satisfied approximately, so that the standard Newton method has problems proceeding. On the other hand,

we can always proceed by adjusting the λ (see the discussion on how to solve them in Section 4.3). This is particularly important for the case of secondary tori that we will discuss in Section 3.4. We also note that this procedure makes possible to deal with tori when the twist condition degenerates.

The equations (9) and (11) will be the centerpiece of our treatment. We will discretize them using Fourier series and study numerical methods to solve the discretized equations.

It is important to remark that there are *a posteriori* theorems (see [16, 15]) for equations (9), (11) (as well as their analogous for flows (12), (14)). These theorems ensure that given a function that satisfies (9), (11) up to a small error and that, at the same time, satisfies some non-degeneracy conditions (which are given quite explicitly), then there is a true solution close to the computed one. Hence, if we monitor the non-degeneracy conditions, we can be sure that the computed solutions correspond to some real effects and are not spurious solutions.

Remark 3. Notice that for whiskered tori the dimension of the torus ℓ is smaller than half the dimension of the phase space $2d$. Hence, the algorithms presented here have the advantage that they look for a function K of ℓ variables to compute invariant objects of dimension ℓ. This is important because the cost of handling functions grows exponentially fast with the number of variables. Indeed, to discretize a function of ℓ variables in a grid of side h into $\mathbb{R}^{2d}$, one needs to store $(1/h)^\ell \cdot 2d$ real values.

Remark 4. Equations (9) and (11) do not have unique solutions. Observe that if K is a solution, for any $\sigma \in \mathbb{R}^\ell$, $K \circ T_\sigma$ is also a solution. In [15], it is shown that, under the non-degeneracy assumptions of the main theorem, this is the only non uniqueness phenomenon in a sufficiently small neighborhood of K. Hence, it is easy to get rid of it by imposing some normalization. See Section 3.5.2 for a discussion on this issue.

3.2. **Functional equations for whiskered invariant tori for vector-fields.** In this case, one can write

$$x(t) = \sum_{k \in \mathbb{Z}^\ell} \hat{x}_k e^{2\pi i k \cdot \omega t}$$

where $\omega \in \mathbb{R}^\ell$, $t \in \mathbb{R}$ and then the hull function K is defined by

$$x(t) = K(\omega t).$$

The invariance equation for flows is written:

$$\partial_\omega K - X \circ K = 0, \tag{12}$$

where ∂_ω denotes the derivative in direction ω

$$\partial_\omega = \sum_{k=1}^{\ell} \omega_k \partial_{\theta_k}. \tag{13}$$

The modification of (12) incorporating a counterterm is:

$$\partial_\omega K - X \circ K - G\lambda = 0, \tag{14}$$

where, under the assumption that subspaces are orthogonal for the metric on $\mathcal{M}$, $G(\theta) = J(K_0)^{-1}(DX \circ K_0)$ with K_0 being a given embedding satisfying some non-degeneracy conditions.

Remark 5. Recall that, since autonomous Hamiltonian systems preserve energy, we can take a surface of section and deal with the return map. This reduces by 2 the dimension of the phase space and the parameterization of the torus requires 1 variable less. In practice, it is much more efficient to use a numerical integrator to compute the point of intersection with the surface of section than to deal with functions of one more variable and with two more components.

3.3. **Some global topological considerations.** In our context, both the domain $\mathbb{T}^\ell$ and the range of K have topology. As a consequence, there will be some topological considerations in the way that the torus $\mathbb{T}^\ell$ gets embedded in the phase space. More explicitly, the angle variables of $\mathbb{T}^\ell$ can get wrapped around in different ways in the phase space.

A concise way of characterizing the topology of the embedding is to consider the lift of K to the universal cover, i.e.

$$\widehat{K} : \mathbb{R}^\ell \to \mathbb{R}^{2d-\ell} \times \mathbb{R}^\ell,$$

in such a way that K is obtained from $\widehat{K}$ by identifying variables in the domain and in the range that differ by an integer.

It is therefore clear that $\forall\ e \in \mathbb{Z}^\ell$

$$\begin{aligned} \widehat{K}_p(\theta + e) &= \widehat{K}_p(\theta), \\ \widehat{K}_q(\theta + e) &= \widehat{K}_q(\theta) + I(e), \end{aligned} \tag{15}$$

where $\widehat{K}_p, \widehat{K}_q$ denote the projections of the lift on the p and q coordinates of $\mathbb{R}^{2d-\ell} \times \mathbb{R}^\ell$. It is easy to see that $I(e)$ is a linear function of e, namely

$$I(e)_{i=1,\ldots,\ell} = \left(\sum_{j=1}^{\ell} I_{ij} e_j \right)_{i=1,\ldots,\ell} \tag{16}$$

with $I_{ij} \in \mathbb{Z}$.

We note that if a function $\widehat{K}_q$ satisfies

$$\widehat{K}_q(\theta + e) = \widehat{K}_q(\theta) + I(e)\ ,$$

the function

$$\widetilde{K}_q(\theta) \equiv \widehat{K}_q(\theta) - I(\theta) \tag{17}$$

is $e-$periodic. The numerical methods will always be based on studying the periodic functions $\widetilde{K}_q$, but we will not emphasize this unless it can lead to confusion.

Of course, the integer valued matrix $I = \{I_{ij}\}_{ij}$ remains constant if we modify the embedding slightly. Hence, it remains constant under continuous deformation. For example, in the integrable case with $\ell = d$, invariant tori satisfy $\widehat{K}_q(\theta) = \theta$, so that we have $I = \mathrm{Id}$. Hence, all the invariant tori which can be continued from tori of the integrable system will also have $I = \mathrm{Id}$.

3.4. **Secondary tori.** One can produce other ℓ-dimensional tori for which the range of I is of dimension less than ℓ. These tori are known as *secondary tori*. It is easy to see that if $\mathrm{rank}(I) < \ell$ we can contract $K(\mathbb{T}^\ell)$ to a diffeomorphic copy of $\mathbb{T}^{\mathrm{rank}(I)}$. Even in the case of maximal tori $\ell = d$, one can have contractible directions. The most famous example of this phenomenon are the "islands" generated in twist maps around resonances.

Secondary tori do not exist in the integrable system and they cannot be even continuously deformed into some of the tori presented in the integrable system.

This is often described informally as saying that the secondary tori are generated by the resonances.

Perturbative proofs of existence of secondary tori are done in [38] and in [13] and in more detail in [12]. In [14] one can find rigorous results showing that these islands have to be rather abundant (in different precise meanings) in many classes of 2D-maps. In particular, for standard-like maps, secondary tori appear at arbitrarily large values of the parameter.

In [28], there are heuristic arguments and numerical simulations arguing that in systems of coupled oscillators, the tori with contractible directions are much more abundant than the invariant tori which can be continued from the integrable limit.

In view of these reasons, we will pay special attention to the computation of secondary tori.

We want to emphasize on some features of the method presented here, which are crucial for the computation of secondary tori:

- The method does not require either the system to be close to integrable nor to be written in action-angle variables.
- The modification of the invariance equations (9) and (12) allows one to adjust some global averages required to solve the Newton equations (see equations (45) and the accompanying discussion on how to solve them in Section 4.3).
- The periodicity of the function $\widetilde{K}$ can be adjusted by the matrix I introduced in (15). Hence, the rank of the matrix I has to be chosen according to the number of contractible directions.

3.5. **Equations for the invariant whiskers.** Invariant tori with $\ell < d$ may have associated invariant bundles and whiskers. We are interested in computing the invariant manifolds which contain the torus and are tangent to the invariant bundles of the linearization around the torus. This includes the stable and unstable manifolds but also invariant manifolds associated to other invariant bundles of the linearization, such as the slow manifolds, associated to the less contracting directions.

Using the parameterization method, it is natural to develop algorithms for invariant manifolds tangent to invariant sub-bundles that satisfy a non-resonance condition (see [9]). This includes as particular cases, the stable/unstable manifolds, the strong stable and strong unstable ones as well as some other slow manifolds satisfying some non-resonance conditions.

To avoid lengthening the paper, we restrict in this paper just to the one-dimensional manifolds (see Section 5), where we do not need to deal with resonances as it is the case in higher dimensions. We think that, considering this particular case, we can state in a more clear and simpler way the main idea behind the algorithms. We will come back to the study of higher dimensional manifolds in future work.

3.5.1. *Invariant manifolds of rank 1.* Again we use a parameterization to describe the whiskers. This amounts to finding a solution u of the equations of motion under the form

$$u^{(n)} = W(\omega n, \mu^n s)$$

in the discrete time case and

$$u(t) = W(\omega t, s e^{\mu t})$$

in the continuous time case, where $W : \mathbb{T}^\ell \times (V \subset \mathbb{R}) \to \mathcal{M}$ and $\mu \in \mathbb{R}$. The function W has then to satisfy the following invariance equations

$$\begin{aligned} F(W(\theta, s)) &= W(\theta + \omega, \mu s), \\ \partial_\omega W(\theta, s) + \mu s \frac{\partial}{\partial s} W(\theta, s) &= (X \circ W)(\theta, s), \end{aligned} \tag{18}$$

for the case of maps and flows, respectively.

Note that equations (18) imply that in variables (θ, s) the motion on the torus consists of a rigid rotation of frequency ω whereas the motion on the whiskers consists of a contraction (or an expansion) by a constant μ (e^μ in the case of flows). We call contractive the situation $|\mu| < 1$ for maps (or $\mu < 0$ for flows). We call expansive the case when $|\mu| > 1$ for maps (or $\mu > 0$ for flows). Note that if $W(\theta, s)$ satisfies (18) then $W(\theta, 0)$ is a solution of the invariance equations (9) or (12).

3.5.2. *Uniqueness of solutions of the invariance equation for whiskers.* The solutions of equations (18) are not unique. Indeed, if $W(\theta, s)$ is a solution, for any $\sigma \in \mathbb{T}^\ell$, $b \in \mathbb{R}$, we have that $\tilde{W}(\theta, s) = W(\theta + \sigma, sb)$ is also a solution. This non-uniqueness of the problem can be removed by supplementing the invariance equation with a normalization condition.

Some suitable normalization conditions (in the case of maps) that make the solutions unique are

$$\begin{aligned} &\int_{\mathbb{T}^\ell} (W(\theta, 0) - I(\theta)) \cdot \nu_i = 0, \\ &DF(W(\theta, 0)) \partial_s W(\theta, 0) = \mu \partial_s W(\theta, 0), \\ &||\partial_s W(\cdot, 0)|| = \rho, \end{aligned} \tag{19}$$

where $\{\nu_i\}_{i=1}^L$ is a basis for Range(I) (L is the dimension) and I is a linear function introduced in (16), $\partial_s W$ denotes the derivative with respect to the variable s, $\rho > 0$ is any arbitratrily chosen number and $\|.\|$ stands for a suitable norm.

The fact that the solutions of (9) supplemented by (19) are locally unique is proved in [15]. In this paper, we will see that these normalizations uniquely determine the Taylor expansions (in s) of the function W whenever the first term $W_1(\theta) \equiv \partial_s W(\theta, 0)$ is fixed, and we will present algorithms to perform these computations.

The first equation in (19) amounts to choosing the origin of coordinates in the parameterization of the torus and, therefore eliminates the ambiguity corresponding to σ.

The second equation in (19) indicates that $W_1(\theta)$ is chosen to be a vector in the hyperbolic direction. We furthermore require that we have chosen the coordinate so that it is an eigenvector of the expanding/contracting direction.

The third equation in (19) chooses the eigenvalue. Equivalently, it fixes the scale in the variables s. Observe that, setting b amounts to multiplying W_1 by b. Hence, setting the norm of $\partial_s W$ sets the scale in s.

From the mathematical point of view, all choices of ρ are equivalent. Nevertheless, from the numerical point of view, it is highly advantageous to choose $||W_1||$ so that the numerical coefficients of the expansion (in s) of W have norms that neither grow nor decrease fast. This makes the computation more immune to round off error since round-off becomes more important when we add/subtract numbers of very different sizes.

3.6. **Fourier-Taylor discretization.** One of the ingredients of algorithms to solve the functional equations is to consider discretizations of functions one searches for.

In this section, we introduce the discretizations we will use. Roughly, for periodic functions, we will use *both* a Fourier series discretization *and* a real discretization on a grid. We will show that the Newton step can be decomposed into substeps which require only $O(N)$ operations in either of the representations. Of course, one can switch between both representations using $O(N \log(N))$ operations using FFT algorithms. For the study of invariant manifolds, we will use Taylor series in the real variables.

3.6.1. *Fourier series discretization.* Since we are seeking functions K which are periodic in the angle variable θ, it is natural to discretize them retaining a finite number of their Fourier coefficients

$$K_N(\theta) = \sum_{k \in \mathcal{O}_N} c_k e^{2i\pi k \cdot \theta}, \tag{20}$$

where

$$\mathcal{O}_N = \left\{ k \in \mathbb{Z}^\ell \mid |k| \leq N \right\}.$$

Since we will deal with real-valued functions, we have $c_k = \bar{c}_{-k}$ and one can just consider the cosine and sine Fourier series,

$$K_N(\theta) = a_0 + \sum_{k \in \mathcal{O}_N} a_k \cos(2\pi k \cdot \theta) + b_k \sin(2\pi k \cdot \theta). \tag{21}$$

These Fourier discretizations have a very long history going back to classical astronomy, but have become much more widely used with computers and go under different name such as *"automatic differentiation"*. The manipulation of these polynomials are reviewed in [35]. A recent review of their applications in dynamics – including implementation issues and examples – is [24].

The main shortcoming of Fourier series discretization of a function is that they are not adaptive and that for discontinuous functions, they converge very slowly and not uniformly. These shortcomings are however not very serious for our applications. Since the tori are invariant under rigid rotations, they tend to be very homogeneous, so that adaptativity is not a great advantage. Also, it is known [15] that if tori are C^r for sufficiently large r, they are in fact analytic so that Fourier series converge fast.

The fact that the Fourier series converge slowly for functions with discontinuities is a slight problem if one wants to compute tori close to the breakdown of analyticity, when the tori transform into Aubry-Mather objects. Of course, when they are far from breakdown – as it happens in many interesting problems in celestial mechanics – the Fourier coefficients converge very fast. To perform calculations close to breakdown, the *a posteriori* theorems in [15] prove invaluable help to have confidence in the computed objects.

3.6.2. *Fourier vs grid representation.* Another representation of the function K is to store the values in a regularly spaced grid. For functions of ℓ variables, we see that if we want to use N variables, we can store either the Fourier coefficients of index up to $O(N^{1/\ell})$ or the values on a grid of step $O(N^{-1/\ell})$.

Some operations are very fast on the real space variables, for example multiplication of functions (it suffices to multiply values at the points of the grid). Also, the evaluation of $F \circ K$ is very fast if we discretize on a grid (we just need to evaluate the function F for each of the points on the grid). Other operations are fast in Fourier

representation. For example, in Fourier space representation it is fast to shift the functions, to take derivatives and, as we will see in Section 3.7, to solve cohomology equations. Hence, our iterative step will consist in the application of several operations, all of which are fast – $O(N)$ – either in Fourier mode representation or in a grid representation. Of course, using the Fast Fourier Transform, we can pass from a grid representation to Fourier coefficients or viceversa in $O(N \ln N)$ operations. There are extremely efficient implementations of the FFT algorithm that take into account not only operation counts but also several other characteristics (memory access, cache, etc.) of modern computers. (See for example [18] for a public domain implementation that takes advantage of all these issues.)

3.6.3. *Fourier-Taylor series.* For the computation of whiskers of invariant tori, we will use Fourier-Taylor expansions of the form

$$W(\theta, s) = \sum_{n=0}^{\infty} W_n(\theta) s^n, \tag{22}$$

where W_n are 1-periodic functions in θ which we will approximate using Fourier series (20).

To manipulate this type of series we will use the so called *automatic differentiation algorithms* (see [35],[24]). For the basic algebraic operations and the elementary transcendental functions (exp, sin, cos, log, power, etc.), they provide an expression for the Taylor coefficients of the result in terms of the coefficients of each of the terms.

3.7. **Cohomology equations and Fourier discretization.** In the Newton step to construct KAM tori, one faces solving cohomology equations, that is, given a periodic (on $\mathbb{T}^\ell$) function η, we want to find another periodic function φ solving

$$\begin{aligned} \varphi - \varphi \circ T_\omega &= \eta, \\ \partial_\omega \varphi &= \eta. \end{aligned} \tag{23}$$

(the first and second equations are the small divisors equations for maps and flows, respectively).

As it is well known, equations (23) have a solution provided that

$$\hat{\eta}_0 \equiv \int_{\mathbb{T}^\ell} \eta = 0, \tag{24}$$

and that ω is Diophantine in the appropriate sense. The Fourier coefficients $\hat{\varphi}_k$ of the solution φ of (23) are then given respectively by

$$\begin{aligned} \hat{\varphi}_k &= \frac{\hat{\eta}_k}{1 - e^{2\pi i k \cdot \omega}}, \\ \hat{\varphi}_k &= \frac{\hat{\eta}_k}{2\pi i \omega \cdot k}. \end{aligned} \tag{25}$$

where $\hat{\eta}_k$ are the Fourier coefficients of the function η.

Notice that the solution φ is unique up to the addition of a constant (the average $\hat{\varphi}_0$ of φ is arbitrary).

Equations (23) and their solutions (25) are very standard in KAM theory (see the exposition in [37]). Very detailed estimates can be found in [43], when ω is Diophantine (which is our case).

4. **Fast Newton methods for (possibly) whiskered tori.** In this section we develop an efficient Newton method to solve the invariance equations (9)-(12) and (11)-(14). We mainly focus on the case of maps (the case for vector fields being similar is described in Appendix B).

We emphasize that the algorithm applies both to whiskered tori and to Lagrangian tori. Indeed, the case of Lagrangian tori is simpler. The hyperbolic part of the Lagrangian tori is just empty so that we do not need to compute the splittings (see Remark 7).

We will assume that the motion on the torus is a rigid rotation with a Diophantine frequency $\omega \in \mathbb{R}^\ell$.

We will consider tori that have a hyperbolic splitting

$$T_{K(\theta)}\mathcal{M} = \mathcal{E}^c_{K(\theta)} \oplus \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)}, \tag{26}$$

such that there exist $0 < \mu_1, \mu_2 < 1$, $\mu_3 > 1$ satisfying $\mu_1\mu_3 < 1$, $\mu_2\mu_3 < 1$ and $C > 0$ such that for all $n \geq 1$ and $\theta \in \mathbb{T}^\ell$

$$\begin{aligned} v \in \mathcal{E}^s_{K(\theta)} &\iff |\mathcal{Z}(n,\theta)v| \leq C\mu_1^n|v| \qquad \forall n \geq 1 \\ v \in \mathcal{E}^u_{K(\theta)} &\iff |\mathcal{Z}(n,\theta)v| \leq C\mu_2^n|v| \qquad \forall n \leq 1 \\ v \in \mathcal{E}^c_{K(\theta)} &\iff |\mathcal{Z}(n,\theta)v| \leq C\mu_3^n|v| \qquad \forall n \in \mathbb{Z} \end{aligned} \tag{27}$$

where $\mathcal{Z}(n,\theta)$ is the cocycle with generator $Z(\theta) = DF(K(\theta))$ and frequency ω, i.e. $\mathcal{Z} : \mathbb{Z} \times \mathbb{T}^\ell \to GL(2d, \mathbb{R})$ is given by

$$\mathcal{Z}(n,\theta) = \begin{cases} Z(\theta + (n-1)\omega) \cdots Z(\theta) & n \geq 1, \\ \mathrm{Id} & n = 0, \\ Z^{-1}(\theta + (n+1)\omega) \cdots Z^{-1}(\theta) & n \leq 1. \end{cases} \tag{28}$$

We will also assume that

$$\dim \mathcal{E}^c_{K(\theta)} = 2\ell, \qquad \dim \mathcal{E}^s_{K(\theta)} = \dim \mathcal{E}^u_{K(\theta)} = d - \ell. \tag{29}$$

We associate to the splitting (26) the projections $\Pi^c_{K(\theta)}$, $\Pi^s_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ over the invariant spaces $\mathcal{E}^c_{K(\theta)}$, $\mathcal{E}^s_{K(\theta)}$ and $\mathcal{E}^u_{K(\theta)}$.

In [25], we provide a method to compute the rank-1 bundles by iterating the cocycle. Of course, once we have computed the vector spanning the rank-1 (un)stable bundle it is very easy to obtain the projections. In Section 4.4 we discuss an alternative to compute the projections by means of a Newton method for which we do not need to assume that the bundle is 1-dimensional.

4.1. **Some remarks on the symplectic geometry of the assumptions.** In this section, we collect some well known remarks about the symplectic properties of the tori. The most important one is that they are isotropic and that the center direction has to have a dimension at least twice that of the torus.

Since F is symplectic (i.e. $F^*\Omega = \Omega$, where Ω is the symplectic 2-form (6)), for all $n \geq 1$ and $n \leq -1$

$$\Omega(u, v) = \Omega(DF^n u, DF^n v),$$

so that, if u, v have rates of decrease, by taking limits in the appropriate direction we obtain that Ω is zero. That is, we get

$$\begin{aligned} &\Omega(\mathcal{E}^s, \mathcal{E}^s) = 0, \quad \Omega(\mathcal{E}^u, \mathcal{E}^u) = 0, \\ &\Omega(\mathcal{E}^c, \mathcal{E}^s) = 0, \quad \Omega(\mathcal{E}^c, \mathcal{E}^u) = 0. \end{aligned} \tag{30}$$

Notice that (30) implies that the symplectic form Ω is non-degenerate on $\mathcal{E}^c$, that is, given $u \in \mathcal{E}^c$, if $\Omega(u, v) = 0 \ \forall v \in \mathcal{E}^c \Rightarrow u = 0$. Indeed, using (30) we have that given $u \in \mathcal{E}^c$, $\Omega(u, v) = 0 \ \forall v \in T_{K(\theta)}\mathcal{M}$. Since Ω is non-degenerate on $\mathcal{M}$, then $u = 0$.

Therefore, $\mathcal{E}^c$ has its own symplectic form $\Omega_{|\mathcal{E}^c}$ and we can choose any metric in $\mathcal{E}^c$ and the associated matrix $J_{|\mathcal{E}^c}$ is a full-rank matrix.

It is also well known that the range of DK is isotropic (see [45]). This is because

$$K^*\Omega_{|\mathcal{E}^c} = K^*F^*\Omega_{|\mathcal{E}^c} = (F \circ K)^*\Omega_{|\mathcal{E}^c} = (K \circ T_\omega)^*\Omega_{|\mathcal{E}^c} = T_\omega^* K^*\Omega_{|\mathcal{E}^c},$$

where K^*, F^* and T_ω^* denote the pullback by K, F and T_ω, respectively and $\Omega_{|\mathcal{E}^c}$ is the symplectic 2-form on $\mathcal{E}^c$. Since the only forms invariant under irrational rotations are constant and $K^*\Omega$ is exact because $\Omega_{|\mathcal{E}^c}$ is exact we conclude that $K^*\Omega_{|\mathcal{E}^c} = 0$. In coordinates, we have that

$$DK^T (J \circ K)_{|\mathcal{E}^c} DK = 0.$$

This means that Range DK is orthogonal to Range $(J \circ K)_{|\mathcal{E}^c} DK$. Hence,

$$\begin{aligned} &\text{Range}\, DK \cap \text{Range}\, (J \circ K)_{|\mathcal{E}^c} DK \\ &= (J \circ K)_{|\mathcal{E}^c} (\text{Range}\, (J \circ K)^{-1}_{|\mathcal{E}^c} DK \cap \text{Range}\, DK) \\ &= \{0\}, \end{aligned}$$

where we have used that $J_{|\mathcal{E}^c}$ is a full rank matrix.

Therefore, the assumption (29) implies that the only non-hyperbolic directions are those spanned by the tangent to the torus and its symplectic conjugate, that is, there are no elliptic directions except those that are forced by the symplectic structure and the fact that the motion on the torus is a rotation.

4.2. **General strategy of the Newton method to solve the invariance equation.** In this section we will design a Newton method to solve the invariance equation (9) and the modified one (11), and discuss several algorithms to deal with the linearized equations.

We first define the following concept of approximate solution.

Definition 4.1. We say that K (resp. (K, λ)) is an approximate solution of equation (9) (resp. (11)) if

$$\begin{aligned} &F \circ K - K \circ T_\omega = E, \\ &(\text{resp.} \quad F \circ K - K \circ T_\omega - G \circ T_\omega \lambda = E), \end{aligned} \tag{31}$$

where E is small.

The Newton method consists in computing Δ in such a way that replacing K by $K + \Delta$ in (31) and expanding the LHS in Δ up to order $\|\Delta\|^2$, it cancels the error term E.

Remark 6. Throughout the paper, we are going to denote $\|.\|$ some norms in functional spaces without specifying what they are exactly. We refer the reader to [36, 15], where the whole theory is developed and the convergence of the algorithms

is proved. Recall that one of the key ideas of KAM theory is that the norms are modified at each step.

Performing a straightforward calculation, we obtain that the Newton procedure to solve equation (9) and (12), given an approximate solution K, consists in finding Δ satisfying

$$(DF \circ K)\Delta - \Delta \circ T_\omega = -E. \tag{32}$$

For the modified invariance equation (11), given an approximate solution (K, λ), the Newton method consists in looking for (Δ, δ) in such a way that $K + \Delta$ and $\lambda + \delta$ eliminate the error in first order. The linearized equation in this case is

$$(DF \circ K)\Delta - \Delta \circ T_\omega - G \circ T_\omega \delta = -E, \tag{33}$$

where one can take $K_0 = K$.

As it is well known, the Newton method converges quadratically in $\|E\|$ and the error $\widetilde{E}$ at step $K + \Delta$ is such that

$$\|\widetilde{E}\| \leq C\|E\|^2,$$

where E is the error at the previous step.

In order to solve the linearized equations (32) and (33), we will first project them on the invariant subspaces $\mathcal{E}^c, \mathcal{E}^u$ and $\mathcal{E}^s$, and then solve an equation for each subspace.

Thus, let us denote

$$\begin{aligned} \Delta^{s,c,u}(\theta) &= \Pi^{s,c,u}_{K(\theta)} \Delta(\theta), \\ E^{s,c,u}(\theta) &= \Pi^{s,c,u}_{K(\theta+\omega)} E(\theta), \end{aligned} \tag{34}$$

such that $\Delta(\theta) = \Delta^s(\theta) + \Delta^c(\theta) + \Delta^u(\theta)$. Then, by the invariant properties of the splitting, the linearized equations for the Newton method (32) and (33) split into:

$$\begin{aligned} DF(K(\theta))\Delta^c(\theta) - \Delta^c \circ T_\omega(\theta) &= -E^c(\theta), \\ DF(K(\theta))\Delta^s(\theta) - \Delta^s \circ T_\omega(\theta) &= -E^s(\theta), \\ DF(K(\theta))\Delta^u(\theta) - \Delta^u \circ T_\omega(\theta) &= -E^u(\theta), \end{aligned} \tag{35}$$

and

$$\begin{aligned} DF(K(\theta))\Delta^c(\theta) - \Delta^c \circ T_\omega(\theta) + \Pi^c_{K(\theta+\omega)} G(\theta+\omega)\delta &= -E^c(\theta), \\ DF(K(\theta))\Delta^s(\theta) - \Delta^s \circ T_\omega(\theta) + \Pi^s_{K(\theta+\omega)} G(\theta+\omega)\delta &= -E^s(\theta), \\ DF(K(\theta))\Delta^u(\theta) - \Delta^u \circ T_\omega(\theta) + \Pi^u_{K(\theta+\omega)} G(\theta+\omega)\delta &= -E^u(\theta). \end{aligned} \tag{36}$$

Notice that once δ is obtained, the equations (36) on the hyperbolic spaces reduce to equations of the form (35). More precisely,

$$DF(K(\theta))\Delta^{s,u}(\theta) - \Delta^{s,u} \circ T_\omega(\theta) = -\widetilde{E}^{s,u}(\theta) \tag{37}$$

where

$$\widetilde{E}^{s,u} = E^{s,u}(\theta) + \Pi^{s,u}_{K(\theta+\omega)} G(\theta+\omega)\delta.$$

Equations (35) and (36) for the stable and unstable spaces can be solved iteratively using the contraction properties of the cocycles on the hyperbolic spaces

given in (27). Indeed, a solution for equations (37) is given explicitly by

$$\Delta^s(\theta) = \widetilde{E}^s \circ T_{-\omega}(\theta) + \sum_{k=1}^{\infty} (DF \circ K \circ T_{-\omega}(\theta) \times \cdots \times DF \circ K \circ T_{-k\omega}(\theta))(\widetilde{E}^s \circ T_{-(k+1)\omega}(\theta)) \tag{38}$$

for the stable equation, and

$$\Delta^u(\theta) = -\sum_{k=0}^{\infty} (DF^{-1} \circ K(\theta) \times \cdots \times DF^{-1} \circ K \circ T_{k\omega}(\theta))(\widetilde{E}^u \circ T_{k\omega}(\theta)) \tag{39}$$

for the unstable direction. Of course, the contraction of the cocycles guarantees the uniform convergence of these series.

The algorithms presented in Appendix A allow us to compute the solutions $\Delta^{s,u}$ of equations (37) efficiently.

In Section 4.3 we discuss how to solve equations (35) and (36) for the center direction.

Hence, the Newton step of the algorithm for whiskered tori that we summarize here will be obtained by combining several algorithms.

Algorithm 4.2. Consider given F, ω, K_0 and an approximate solution K (resp. K, λ), perform the following operations:

A) Compute the invariant splittings and the projections associated to the cocycle $Z(\theta) = DF \circ K(\theta)$ and ω using the algorithms described in Section 4.4 (or in [25]).
B) Project the linearized equation to the hyperbolic space and use the algorithms described in Appendix A to obtain $\Delta^{s,u}$.
C) Project the linearized equation on the center subspace and use the Algorithm 4.3 in Section 4.3 to obtain Δ^c and δ.
D) Set $K + \Delta^s + \Delta^u + \Delta^c \to K$ and $\lambda + \delta \to \lambda$

Of course, since this is a Newton step, it will have to be iterated repeatedly until one reaches solutions up to a small tolerance error.

We will start by some remarks on the different steps of Algorithm 4.2 and, later, we will provide more details on them.

Remark 7. It is important to remark that the above Algorithm 4.2 also applies to the case of Lagrangian tori. In this case the center space is the whole phase space, so that there is no need to compute the splitting. Hence, for Lagrangian tori, the steps A) and B) of Algorithm 4.2 are trivial and do not need any work.

Remark 8. The main issue of the Newton method is that it needs a good initial guess to start the iteration. On the other hand, any reasonable approximation can be used as an input to the Newton method. The algorithm used to generate the approximation does not need to be justified theoretically. It only needs to work in practice. Indeed, our problems have enough structure so that one can use Lindstedt series, variational methods, approximation by periodic orbits, frequency methods, besides the customary continuation methods.

Remark 9. As we have mentioned in Remark 4, the solutions of (9) and (12) are not unique. Therefore, in order to avoid dealing with non-invertible matrices in the Newton procedure, we will impose the normalization condition

$$\int_{\mathbb{T}^\ell} (K(\theta) - I(\theta)) \cdot \nu_i = 0$$

where $\{\nu_i\}_{i=1}^{L}$ is a basis for Range(I) (L being the dimension) and I is the linear function introduced in (16).

4.3. **Fast Newton method for (whiskered) tori: the center directions.** We present here the Newton method to solve the equations in the center subspace for the case of maps.

For Lagrangian tori, the hyperbolic directions are empty and the study of the center direction is the only component which is needed. Hence, the algorithms discussed in this section allow to solve, in particular, equations (32) and (33) in the case of Lagrangian tori. For a discussion of the center equations for Hamiltonian flows, we refer the reader to Appendix B.

The key observation is that the linearized Newton equations (32) and (33) are closely related to the dynamics and therefore, we can use geometric identities to find a linear change of variables that reduces the Newton equations to upper diagonal difference equations with constant coefficients. This phenomenon is often called "automatic reducibility".

The idea is stated in the following proposition:

Proposition 1 (Automatic reducibility, see [16, 15])**.** *Given an approximation K of the invariance equation as in (31), denote*

$$\begin{aligned} \alpha(\theta) &= DK(\theta) \\ N(\theta) &= \left([\alpha(\theta)]^\top \alpha(\theta)\right)^{-1} \\ \beta(\theta) &= \alpha(\theta) N(\theta) \\ \gamma(\theta) &= (J \circ K(\theta))^{-1}_{|\mathcal{E}^c} \beta(\theta) \end{aligned} \tag{40}$$

where $(J \circ K(\theta))^{-1}_{|\mathcal{E}^c}$ denotes the matrix associated to the 2-form of the center subspace. Form the following matrix

$$M(\theta) = [\alpha(\theta)\,,\, \gamma(\theta)], \tag{41}$$

where by $[\cdot,\cdot]$ we denote the $2d \times 2\ell$ matrix obtained by juxtaposing the two $2d \times \ell$ matrices that are in the arguments.

Then, we have

$$(DF \circ K(\theta)) M(\theta) = M(\theta + \omega) \begin{pmatrix} \mathrm{Id} & A(\theta) \\ 0 & \mathrm{Id} \end{pmatrix} + \widehat{E}(\theta) \tag{42}$$

where

$$A(\theta) = \beta(\theta + \omega)^\top [(DF \circ K(\theta)) \gamma(\theta) - \gamma(\theta + \omega)], \tag{43}$$

and $\|\widehat{E}\| \leq \|DE\|$ in the case of (32) or $\|\widehat{E}\| \leq \|DE\| + |\lambda|$ in the case of (33).

Remark 10. If the symplectic structure is almost-complex (i.e. $J^2 = -\mathrm{Id}$), we have that

$$\beta(\theta + \omega)^\top \gamma(\theta + \omega) = 0,$$

since the torus is isotropic. Then $A(\theta)$ has a simpler expression given by

$$A(\theta) = \beta(\theta + \omega)^\top (DF \circ K)(\theta) \gamma(\theta).$$

Once again, we omit the definition of the norms used in the bounds for $\widehat{E}$. For these precisions, we refer to the paper [15], where the convergence of the algorithm is established.

It is interesting to pay attention to the geometric interpretation of the identity (42). Note that, taking derivatives with respect to θ in (31), we obtain that

$$(DF \circ K)DK - DK \circ T_\omega = DE,$$

which means that the vectors DK are invariant under $DF \circ K$ (up to a certain error). Moreover, $(J \circ K)^{-1}_{\mathcal{E}^c} DKN$ are the symplectic conjugate vectors of DK, so that the preservation of the symplectic form clearly implies (42). The geometric interpretation of the matrix $A(\theta)$ is a shear flow near the approximately invariant torus. See Figure 1.

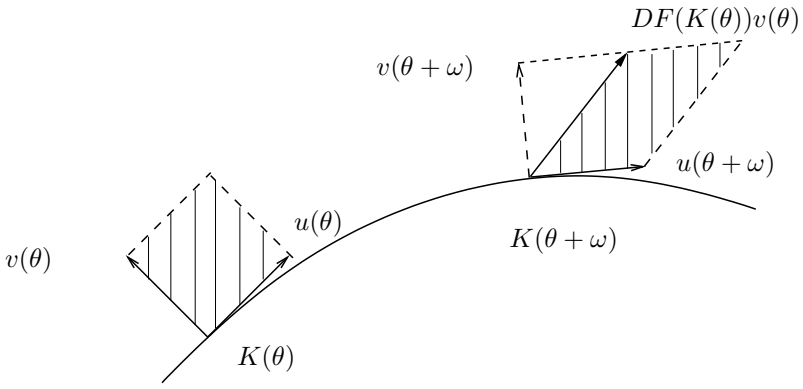

FIGURE 1. Geometric representation of the automatic reducibility where $u = DK$, $v = (J \circ K)^{-1}_{\mathcal{E}^c} DKN$

To be able to use the change of unknowns via the matrix M previously introduced on the center subspace, one has to ensure that one can identify the center space $\mathcal{E}^c_{K(\theta)}$ with the range of M. This is ensured by taking the matrix $J_{\mathcal{E}^c}$ associated to the symplectic form of the center manifold. More details of the proof are given in [15] to which we refer.

For our purposes it is important to compute not just the invariant spaces, but also the projections over invariant subspaces. Knowing one invariant subspace is not enough to compute the projection, since it also depends on the complementary space chosen.

Next, we will see that the result stated in Proposition 1 allows us to design a very efficient algorithm for the Newton step.

Notice first that if we change the unknowns $\Delta = MU$ in (32) and (33) and we use (42) we obtain

$$\begin{aligned} M(\theta+\omega)\begin{pmatrix} \mathrm{Id} & A(\theta) \\ 0 & \mathrm{Id} \end{pmatrix} U(\theta) - M(\theta+\omega)U(\theta+\omega) \\ - G(\theta+\omega)\delta = -E(\theta) \end{aligned} \tag{44}$$

Of course, the term involving δ has to be omitted when considering (32).

Multiplying (44) by $M(\theta+\omega)^\top J(K(\theta+\omega))_{|\mathcal{E}^c}$ and using the invertibility of the matrix $M(\theta+\omega)^\top J(K(\theta+\omega))_{|\mathcal{E}^c} M(\theta+\omega)$, we are left with the system of equations

$$\begin{aligned} U_1(\theta) + A(\theta)U_2(\theta) - B_1(\theta)\delta - U_1(\theta+\omega) = -\widetilde{E}_1(\theta) \\ U_2(\theta) - U_2(\theta+\omega) - B_2(\theta)\delta = -\widetilde{E}_2(\theta) \end{aligned} \tag{45}$$

where

$$\widetilde{E}(\theta) = (M(\theta+\omega)^\top J(K(\theta+\omega))_{|\mathcal{E}^c} M(\theta+\omega))^{-1} M(\theta+\omega)^\top J(K(\theta+\omega))_{|\mathcal{E}^c} E(\theta)$$

$$B(\theta) = [(M^\top J(K)_{|\mathcal{E}^c} M)^{-1} M^\top J(K)_{|\mathcal{E}^c} G] \circ T_\omega(\theta)$$

and the subindices $i = 1, 2$ indicate symplectic coordinates.

When K is close to K_0, we expect that B_2 is close to the ℓ-dimensional identity matrix and B_1 is small.

The next step is to solve equations (45) for U (and δ). Equations (45) are equations of the form considered in (23) and they can be solved very efficiently in Fourier space.

More precisely, the second equation of (45) is uncoupled from the first one and allows us to determine U_2 (up to a constant) and δ. The role of the parameter δ is now clear. It allows us to adjust some global averages that we need to be able to solve equations (45). Indeed, we choose δ so that the term $B_2(\theta)\delta - \widetilde{E}_2$ has zero average (which is a necessary condition to solve small divisor equations as described in Section 3.7). This allows us to solve equation (23) for U_2. We then denote

$$U_2(\theta) = \widetilde{U}_2(\theta) + \overline{U}_2$$

where $\widetilde{U}_2(\theta)$ has average zero and $\overline{U}_2 \in \mathbb{R}$.

Once we have $\widetilde{U}_2$, we can substitute U_2 in the first equation. We get $\overline{U}_2$ imposing that the average of

$$B_1(\theta)\delta - A(\theta)\widetilde{U}_2(\theta) - A(\theta)\overline{U}_2 - \widetilde{E}_1(\theta)$$

is zero and then we can find U_1 up to a constant according to (25).

We therefore have the following algorithm to solve (3) in the center direction,

Algorithm 4.3 (Newton step in the center direction)**.** Consider given F, ω, K_0 and an approximate solution K (resp. K, λ). Perform the following calculations

1. (1.1) Compute $F \circ K$
   (1.2) Compute $K \circ T_\omega$
   (1.3) Compute the invariant projections, Π^s, Π^u, Π^c.
2. Set $E^c = \Pi^c(F \circ K - K \circ T_\omega)$ (resp. set $E^c = \Pi^c(F \circ K - K \circ T_\omega - G \circ T_\omega \lambda)$)
3. Following (40)
   (3.1) Compute $\alpha(\theta) = DK(\theta)$
   (3.2) Compute $N(\theta) = \left([\alpha(\theta)]^\top \alpha(\theta)\right)^{-1}$
   (3.3) Compute $\beta(\theta) = \alpha(\theta) N(\theta)$
   (3.4) Compute $\gamma(\theta) = (J(K(\theta)))^{-1}_{|\mathcal{E}^c} \beta(\theta)$
   (3.5) Compute $M(\theta) = [\alpha(\theta), \gamma(\theta)]$
   (3.6) Compute $M(\theta+\omega)$
   (3.7) Compute $(M(\theta+\omega)^\top J(K(\theta+\omega))_{|\mathcal{E}^c} M(\theta+\omega))^{-1}$
   (3.8) Compute $\widetilde{E}(\theta) = (M(\theta+\omega)^\top J(K(\theta+\omega))_{|\mathcal{E}^c} M(\theta+\omega))^{-1} E^c(\theta)$.
   We denote $\widetilde{E}_1, \widetilde{E}_2$ the components of $\widetilde{E}$ along $J^{-1}_{\mathcal{E}^c} DK$ and DK, respectively.
   (3.9) Compute

$$A(\theta) = \beta(\theta+\omega)^\top [(DF \circ K(\theta))\gamma(\theta) - \gamma(\theta+\omega)]$$

   as indicated in (43)

4. (4.1) Solve for U_2 satisfying

$$U_2 - U_2 \circ T_\omega = -\widetilde{E}_2 + \int_{\mathbb{T}^\ell} \widetilde{E}_2$$

(resp.
(4.1′) Solve for δ such that

$$\int_{\mathbb{T}^\ell} \widetilde{E}_2 - \left[\int_{\mathbb{T}^\ell} B_2\right]\delta = 0$$

(4.2′) Solve for U_2 satisfying

$$U_2 - U_2 \circ T_\omega = -\widetilde{E}_2 + B_2\delta$$

Set U_2 such that the average is 0.)

5. (5.1) Compute $A(\theta)U_2(\theta)$
(5.2) Solve for $\overline{U}_2$ satisfying

$$0 = \int_{\mathbb{T}^\ell} \widetilde{E}_1(\theta) + \int_{\mathbb{T}^\ell} A(\theta)U_2(\theta) + \left[\int_{\mathbb{T}^\ell} A(\theta)\right]\overline{U}_2$$

(5.3) Find U_1 solving

$$U_1 - U_1 \circ T_\omega = -\widetilde{E}_1 - A(U_2 + \overline{U}_2)$$

Normalize it so that $\int_{\mathbb{T}^\ell} U_1 = 0$
(resp.
(5.1′) Compute $A(\theta)U_2(\theta)$
(5.2′) Solve for $\overline{U}_2$ satisfying

$$0 = \int_{\mathbb{T}^\ell} \widetilde{E}_1(\theta) - \int_{\mathbb{T}^\ell} B_1(\theta)\delta + \int_{\mathbb{T}^\ell} A(\theta)U_2(\theta) + \left[\int_{\mathbb{T}^\ell} A(\theta)\right]\overline{U}_2$$

(5.3′) Find U_1 solving

$$U_1 - U_1 \circ T_\omega = -\widetilde{E}_1 - A(U_2 + \overline{U}_2) + B_1\delta$$

Normalize it so that $\int_{\mathbb{T}^\ell} U_1 = 0$.)

6. The improved K is $K(\theta) + M(\theta)U(\theta)$
(resp. the improved λ is $\lambda + \delta$).

Notice that steps (1.2), (3.1), (3.6), (4.1) (resp. (4.2′)), (5.3) (resp. (5.3′)) in Algorithm 4.3 are diagonal in Fourier series, whereas the other steps are diagonal in the real space representation. Note also that the algorithm only stores vectors whose size is of order N.

Remark 11. Using the symplectic properties of the matrix M, step (3.7) can be sped up.

When the torus is exactly invariant we have that the invariant torus is co-isotropic, that is $DK^\top (J \circ K)_{|\mathcal{E}^c} DK = 0$. Hence, when the torus is invariant, we have that

$$M(\theta+\omega)^\top J(K(\theta+\omega))_{|\mathcal{E}^c} M(\theta+\omega) = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & P(\theta) \end{pmatrix}$$

where $P(\theta) = -N^T(\theta)DK^T(\theta)(J\circ K)^{-1}_{|\mathcal{E}^c}(\theta)DK(\theta)N(\theta)$. Therefore, the inverse can be computed just rearranging terms, i.e.

$$[M(\theta+\omega)^\top J(K(\theta+\omega))_{|\mathcal{E}^c}M(\theta+\omega)]^{-1} = \begin{pmatrix} P(\theta) & -\mathrm{Id} \\ \mathrm{Id} & 0 \end{pmatrix} \tag{46}$$

Notice that when J induces an almost-complex structure, i.e. $J^2 = -\mathrm{Id}$, $P(\theta) = 0$ and the matrix M is symplectic from the original form to the standard one (it transforms the original symplectic structure into the standard symplectic structure).

For the purposes of a Newton Method, we can use the expression (46) for the inverse in step (3.7) and still obtain a quadratically convergent algorithm.

4.4. **A Newton method to compute the projections over invariant subspaces.** In this section we will discuss a Newton method to compute the projections $\Pi^c_{K(\theta)}$, $\Pi^s_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ associated to the linear spaces $\mathcal{E}^c_{K(\theta)}$, $\mathcal{E}^s_{K(\theta)}$ and $\mathcal{E}^u_{K(\theta)}$ where K is an (approximate) invariant torus. More precisely, we will design a Newton method to compute $\Pi^s_{K(\theta)}$ and $\Pi^{cu}_{K(\theta)} = \Pi^c_{K(\theta)} + \Pi^u_{K(\theta)}$. Similar arguments allow us to design a Newton method to compute $\Pi^u_{K(\theta)}$ and $\Pi^{cs}_{K(\theta)} = \Pi^c_{K(\theta)} + \Pi^s_{K(\theta)}$. Then, of course, $\Pi^c_{K(\theta)}$ is given by

$$\Pi^c_{K(\theta)} = \Pi^{cs}_{K(\theta)}\Pi^{cu}_{K(\theta)} = \Pi^{cu}_{K(\theta)}\Pi^{cs}_{K(\theta)} \ .$$

We note that this method does not use the symplectic geometry of the problem and that, hence, it is applicable to any dynamical system.

Let us discuss first a Newton method to compute $\Pi^s_{K(\theta)}$ and $\Pi^{cu}_{K(\theta)}$. To simplify notation, from now on, we will omit the dependence in $K(\theta)$.

Given a cocycle $Z(\theta)$ (which in our case will be $Z(\theta) = DF(K(\theta))$), we will look for maps $\Pi^s : \mathbb{T}^\ell \to \mathcal{M}_{2d\times 2d}(\mathbb{R})$ and $\Pi^{cu} : \mathbb{T}^\ell \to \mathcal{M}_{2d\times 2d}(\mathbb{R})$ satisfying the following equations:

$$\Pi^{cu}(\theta+\omega)Z(\theta)\Pi^s(\theta) = 0, \tag{47}$$
$$\Pi^s(\theta+\omega)Z(\theta)\Pi^{cu}(\theta) = 0, \tag{48}$$
$$\Pi^s(\theta) + \Pi^{cu}(\theta) = \mathrm{Id}, \tag{49}$$
$$[\Pi^s(\theta)]^2 = \Pi^s(\theta), \tag{50}$$
$$[\Pi^{cu}(\theta)]^2 = \Pi^{cu}(\theta), \tag{51}$$
$$\Pi^s(\theta)\Pi^{cu}(\theta) = 0, \tag{52}$$
$$\Pi^{cu}(\theta)\Pi^s(\theta) = 0. \tag{53}$$

Notice that the system of equations (47)–(53) is redundant. It is easy to see that equations (51), (52) and (53) follow from equations (49) and (50). Therefore, the system of equations that needs to be solved is reduced to equations (47)–(50).

We are going to design a Newton method to solve equations (47)–(48) and use equations (49)–(50) as constraints. In this context, by approximate solution of equations (47)–(48), we mean a solution (Π^s, Π^{cu}) such that

$$\Pi^{cu}(\theta+\omega)Z(\theta)\Pi^s(\theta) = E^{cu}(\theta), \tag{54}$$
$$\Pi^s(\theta+\omega)Z(\theta)\Pi^{cu}(\theta) = E^s(\theta), \tag{55}$$
$$\Pi^s(\theta) + \Pi^{cu}(\theta) = \mathrm{Id}, \tag{56}$$
$$[\Pi^s(\theta)]^2 = \Pi^s(\theta). \tag{57}$$

where E^i denotes the error in a certain component. Notice that the error in equation (54) has components only on the center and unstable "approximated" subspaces and we denote it by E^{cu}. The same happens with the equation (55) but on the "approximated" stable subspace. We assume that E^{cu} and E^s are both small.

As standard in the Newton method, we will look for increments Δ^s and Δ^{cu} in such a way that setting $\Pi^s \leftarrow \Pi^s + \Delta^s$ and $\Pi^{cu} \leftarrow \Pi^{cu} + \Delta^{cu}$, the new projections solve equations (47) and (48) up to order $\|E\|^2$ where $\|E\| = \|E^s\| + \|E^{cu}\|$ for some norm $\|.\|$.

The functions Δ^s and Δ^{cu} solve the following equations

$$\begin{aligned} \Delta^{cu}(\theta+\omega)Z(\theta)\Pi^s(\theta) + \Pi^{cu}(\theta+\omega)Z(\theta)\Delta^s(\theta) &= -E^{cu}(\theta) \\ \Delta^s(\theta+\omega)Z(\theta)\Pi^{cu}(\theta) + \Pi^s(\theta+\omega)Z(\theta)\Delta^{cu}(\theta) &= -E^s(\theta) \end{aligned} \tag{58}$$

with the constraints

$$\Delta^s(\theta) + \Delta^{cu}(\theta) = 0 \tag{59}$$

$$\Pi^s(\theta)\Delta^s(\theta) + \Delta^s(\theta)\Pi^s(\theta) = \Delta^s(\theta)\ . \tag{60}$$

By equation (59) we only need to compute Δ^s since $\Delta^{cu} = -\Delta^s$. We now work out equations (58), (59) and (60) so that we can find Δ^s.

Denote

$$\begin{aligned} \Delta^s_s &= \Pi^s\Delta^s, \\ \Delta^s_{cu} &= \Pi^{cu}\Delta^s, \end{aligned} \tag{61}$$

so that

$$\Delta^s = \Delta^s_s + \Delta^s_{cu}. \tag{62}$$

Then equation (60) reads

$$\Delta^s_s(\theta) + \Delta^s(\theta)\Pi^s(\theta) = \Delta^s_s(\theta) + \Delta^s_{cu}(\theta), \tag{63}$$

or equivalently,

$$\Delta^s(\theta)\Pi^s(\theta) = \Delta^s_{cu}(\theta)\ . \tag{64}$$

By (56), (64) and (62) we have that

$$\Delta^s(\theta)\Pi^{cu}(\theta) = \Delta^s(\theta) - \Delta^s(\theta)\Pi^s(\theta) = \Delta^s(\theta) - \Delta^s_{cu}(\theta) = \Delta^s_s(\theta). \tag{65}$$

Now, using (59), equations (58) transform to

$$\begin{aligned} -\Delta^s(\theta+\omega)Z(\theta)\Pi^s(\theta) + \Pi^{cu}(\theta+\omega)Z(\theta)\Delta^s(\theta) &= -E^{cu}(\theta), \\ \Delta^s(\theta+\omega)Z(\theta)\Pi^{cu}(\theta) - \Pi^s(\theta+\omega)Z(\theta)\Delta^s(\theta) &= -E^s(\theta). \end{aligned} \tag{66}$$

Denoting

$$\begin{aligned} N_s(\theta) &= \Pi^s(\theta+\omega)Z(\theta)\Pi^s(\theta), \\ N_{cu}(\theta) &= \Pi^{cu}(\theta+\omega)Z(\theta)\Pi^{cu}(\theta), \end{aligned}$$

and using that $\Pi^{cu}(\theta+\omega)Z(\theta)\Pi^s(\theta)$ and $\Pi^s(\theta+\omega)Z(\theta)\Pi^{cu}(\theta)$ are small by (54)–(55) and $\Pi^s(\theta) + \Pi^{cu}(\theta) = \mathrm{Id}$ by (56), it is enough for the Newton method to solve for Δ^s satisfying the following equations

$$\begin{aligned} -\Delta^s(\theta+\omega)\Pi^s(\theta+\omega)N_s(\theta) + N_{cu}(\theta)\Pi^{cu}(\theta)\Delta^s(\theta) &= -E^{cu}(\theta), \\ \Delta^s(\theta+\omega)\Pi^{cu}(\theta+\omega)N_{cu}(\theta) - N_s(\theta)\Pi^s(\theta)\Delta^s(\theta) &= -E^s(\theta). \end{aligned} \tag{67}$$

Finally, by expressions (64) and (65) and taking into account the notations introduced in (61), equations (67) read

$$-\Delta^s_{cu}(\theta+\omega)N_s(\theta) + N_{cu}(\theta)\Delta^s_{cu}(\theta) = -E^{cu}(\theta), \tag{68}$$

$$\Delta^s_s(\theta+\omega)N_{cu}(\theta) - N_s(\theta)\Delta^s_s(\theta) = -E^s(\theta). \tag{69}$$

In Appendix A, we discussed how to solve efficiently equations of the form (68)-(69). Notice that they are of the form (94) for $A(\theta) = N_{cu}(\theta)$, $B(\theta) = N_s(\theta)$ and $\eta(\theta) = -E^{cu}(\theta)$ in the case of equation (68) and $A(\theta) = N_s(\theta)$, $B(\theta) = N_{cu}(\theta)$ and $\eta(\theta) = +E^s(\theta)$ in the case of equation (69). Furthermore, $\|N_s\| < 1$ and $\|N^{-1}_{cu}\| < 1$. Hence, they can be solved iteratively using the fast iterative algorithms described in Appendix A.

The explicit expressions for Δ^s_{cu} and Δ^s_s are

$$\begin{aligned}\Delta^s_{cu}(\theta) = -\Big[&N^{-1}_{cu}(\theta)E^{cu}(\theta) + \sum_{n=1}^{\infty} N^{-1}_{cu}(\theta)\times\cdots\times \\ &N^{-1}_{cu}(\theta+n\omega)E^{cu}(\theta+n\omega)N_s(\theta+(n-1)\omega)\times\cdots\times N_s(\theta)\Big]\end{aligned} \tag{70}$$

and

$$\begin{aligned}\Delta^s_s(\theta) =& E^s(\theta-\omega)N^{-1}_{cu}(\theta-\omega) + \sum_{n=1}^{\infty} N_s(\theta-\omega)\times\cdots\times \\ &N_s(\theta-(n+1)\omega)E^s(\theta-(n+1)\omega)N^{-1}_{cu}(\theta-(n+1)\omega)\times\cdots\times N^{-1}_{cu}(\theta-\omega).\end{aligned} \tag{71}$$

Remark 12. Notice that $N_{cu}(\theta)$ can only be inverted when restricted to the right spaces. Hence we assume that $N_{cu}(\theta)$ is a matrix defined on $\mathcal{E}^c \oplus \mathcal{E}^u$. Similarly in other instances. Most of the time, one does not need to distinguish between $\mathcal{E}^{s,u,c}$ as an space on its own or as a subspace of the ambient space. This is one of the few cases where such a distintion is needed.

Finally, let us check that $\Delta^s = \Delta^s_{cu} + \Delta^s_s$ also satisfies the constraints. In order to check that constraint (60), which is equivalent to (64), is satisfied we will use the expressions (70) and (71). Notice first that

$$N_s(\theta)\Pi^s(\theta) = N_s(\theta) \tag{72}$$

and

$$N^{-1}_{cu}(\theta-\omega)\Pi^s(\theta) = 0\ . \tag{73}$$

Moreover, from (54) and using (57) one can see that

$$E^{cu}(\theta)\Pi^s(\theta) = \Pi^{cu}(\theta+\omega)Z(\theta)[\Pi^s(\theta)]^2 = E^{cu}(\theta)\ . \tag{74}$$

Then, from expressions (70) and (71) and the above expression (72), (73) and (74), it is clear that

$$\Delta^s(\theta)\Pi^s(\theta) = \Delta^s_s(\theta)\Pi^s(\theta) + \Delta^s_{cu}(\theta)\Pi^s(\theta) = 0 + \Delta^s_{cu},$$

hence, constraint (64) is satisfied.

Now, using equation (59) we get $\Delta^{cu}(\theta) = -(\Delta^s_s(\theta) + \Delta^s_{cu}(\theta))$ and the improved projections are

$$\begin{aligned}\widetilde{\Pi}^s(\theta) &= \Pi^s(\theta) + \Delta^s_s(\theta) + \Delta^s_{cu}(\theta)\\ \widetilde{\Pi}^{cu}(\theta) &= \Pi^{cu}(\theta) + \Delta^{cu}(\theta).\end{aligned}$$

The new error for equations (47) and (48) is now $\|\widetilde{E}\| \leq C\|E\|^2$ where $\|E\| = \|E^{cu}\| + \|E^s\|$. Of course equation (49) is clearly satisfied but (50) is satisfied up to an error which is quadratic in $\|E\|$. However it is easy to get an exact solution for (50) and the correction is quadratic in Δ^s (and therefore in Δ^{cu}). To do so, we just take the the singular value decomposition (SVD) (see [23]) of $\widetilde{\Pi}^s$ and we set the values in the singular value decomposition to be either 1 or 0.

In this way we obtain new projections Π^s_{new} and $\Pi^{cu}_{\text{new}} = \text{Id} - \Pi^s_{\text{new}}$ satisfying

$$\|\Pi^s_{\text{new}} - \widetilde{\Pi}^s\| < \|\Delta^s\|^2$$
$$\|\Pi^{cu}_{\text{new}} - \widetilde{\Pi}^{cu}\| < \|\Delta^{cu}\|^2,$$

so that the error for equations (47) and (48) is still quadratic in $\|E\|$. Moreover, they satisfy equations (50) and, of course, (49) exactly.

Hence, setting $\Pi^s \leftarrow \Pi^s_{\text{new}}$ and $\Pi^{cu} \leftarrow \Pi^{cu}_{\text{new}}$ we can repeat the procedure described in this section and perform another Newton step.

Consequently, the algorithm of the Newton method to compute the projections is:

Algorithm 4.4 (Computation of the projections by a Newton method)**.** Consider given F, K, ω and an approximate solution (Π^s, Π^{cu}) of equations (47)-(48). Perform the following calculations:

1. Compute $Z(\theta) = DF \circ K(\theta)$
2. (2.1) Compute $E^{cu}(\theta) = \Pi^{cu}(\theta+\omega)Z(\theta)\Pi^s(\theta)$
   (2.2) Compute $E^s(\theta) = \Pi^s(\theta+\omega)Z(\theta)\Pi^{cu}(\theta)$
3. (3.1) Compute $N_s(\theta) = \Pi^s(\theta+\omega)Z(\theta)\Pi^s(\theta)$
   (3.2) Compute $N_{cu}(\theta) = \Pi^{cu}(\theta+\omega)Z(\theta)\Pi^{cu}(\theta)$
4. (4.1) Solve for Δ^s_s satisfying
$$N_s(\theta)\Delta^s_s(\theta) - \Delta^s_s(\theta+\omega)N_{cu}(\theta) = E^s(\theta)$$
   (4.2) Solve for Δ^s_{cu} satisfying
$$N_{cu}(\theta)\Delta^s_{cu}(\theta) - \Delta^s_{cu}(\theta+\omega)N_s(\theta) = -E^{cu}(\theta)$$
5. (5.1)Compute $\widetilde{\Pi}^s(\theta) = \Pi^s(\theta) + \Delta^s_s(\theta) + \Delta^s_{cu}(\theta)$.
   (5.2) Compute the SVD decomposition of $\widetilde{\Pi}^s(\theta)$: $\widetilde{\Pi}^s(\theta) = U(\theta)\Sigma(\theta)V^\top(\theta)$.
   (5.3) Set the values in $\Sigma(\theta)$ equal to the closer integer (which will be either 0 or 1).
   (5.4) Recompute $\bar{\Pi}^s(\theta) = U(\theta)\Sigma(\theta)V^\top(\theta)$.
6. Set $\bar{\Pi}^s \to \Pi^s$
   $\text{Id} - \bar{\Pi}^s \to \Pi^{cu}$
   and iterate the procedure.

Notice that the matrix multiplication is diagonal in real space representation, whereas the phase shift is diagonal in Fourier space. A discussion on how to perform step 4 efficiently is given in Appendix A.

Remark 13. There are several variations that can improve the efficiency. One variation is that in (5.2) we do not need to compute the SVD from scratch. If we save the SVD, obtained at one step, it will be an approximate SVD for the next step. So that we can use it as the basis of an iterative algorithm to find the new SVD. This iterative solution only requires to use the application of algebraic operations among functions of θ.

5. **Computation of rank-1 whiskers of an invariant torus.** In this section, we present algorithms to compute the whiskers associated to an invariant torus, that is the invariant manifolds that contain the torus and are tangent to the invariant bundles.

For the sake of simplicity we will only discuss the case when the invariant whiskers are one-dimensional (i.e. $d - \ell = 1$). The same idea can be extended to compute invariant manifolds of any rank. However, there are several new phenomena (resonances) that can appear and need to be discussed. We plan to come back to this issue in the future.

As we already mentioned in Section 3.5.1, we will look for the whiskers by finding a parameterization for them, so we will search for a function $W : \mathbb{T}^\ell \times (V \subset \mathbb{R}) \to \mathcal{M}$ and a scalar μ satisfying equation (18).

We will consider two different methods to solve equation (18). We will first discuss the order by order method. The other method is based on the philosophy of quasi-Newton methods. Assuming that the invariant tori and the tangent bundles are already known, we use "automatic reducibility" to design an efficient Newton method.

We detail only the case of maps because the same ideas work for the case of vector fields and we refer the reader to Appendices C-D for the case of flows.

Similar algorithms were developed and implemented in [30, 32] for the slightly simpler case of quasi-periodic maps.

5.1. **The order by order method.** In this section we adapt the parameterization method introduced in [11]. The convergence of the Fourier-Taylor series in this paper can be easily adapted to the present case. We focus on the case of maps and refer the reader to Appendix C for the case of flows. We emphasixe that the order by order method does not take advantage of the geometry of the problem and, hence, it can be applied to many other problems.

We will find a solution (W, μ) of the invariance equation (18) discretizing it in Fourier-Taylor series. Hence, we will look for W as a power series

$$W(\theta, s) = \sum_{n=0}^{\infty} W_n(\theta) s^n, \tag{75}$$

and match similar coefficients in s^n on both sides of equation (18).

For $n = 0$, we obtain

$$F(W_0(\theta)) = W_0(\theta + \omega), \tag{76}$$

which is equation (9) for the invariant torus. Therefore, we have $W_0(\theta) = K(\theta)$, where K is a parametrization of the invariant torus.

For $n = 1$, we obtain

$$DF \circ K(\theta) W_1(\theta) = W_1(\theta + \omega)\mu, \tag{77}$$

so that $W_1(\theta)$ is an eigenfunction with eigenvalue μ of the operator $\mathcal{Z}(1, \theta)$ defined in equation (28).

Equation (77) states that the bundle spanned by W_1 is invariant for the linearization of F, and the dynamics on it is reduced to a contraction/expansion by a constant μ. Therefore, one can compute W_1 and μ using the algorithms given in Section 4.4.

Remark 14. Notice that if $W_1(\theta)$ is a solution of equation (77), then $bW_1(\theta)$, for any $b \in \mathbb{R}$, is also a solution. See Section 3.5.2 for a discussion on how to choose b.

For $n \geq 2$, we obtain

$$DF \circ K(\theta) W_n(\theta) = W_n(\theta + \omega)\mu^n + R_n[W_0, \ldots, W_{n-1}](\theta), \tag{78}$$

where R_n is an explicit polynomial in $W_0, \ldots, W_{n-1}$ whose coefficients are derivatives of F evaluated at W_0.

Equation (78) can be solved provided that μ is such that μ^n is not in the spectrum of the operator $\mathcal{Z}(1, \theta)$. This condition is clearly satisfied in the case of (un)stable bundles which are one-dimensional but it can also be satisfied by other bundles.

Equation (78) can be solved using the large matrix method. It consists on considering a discretization of equation (78) using Fourier series and reducing the problem to solving a linear system in Fourier space, where the unknowns are the Fourier coefficients of the matrix W_n.

There are also efficient algorithms which are variants of the methods devoted in the previous sections. The equation (78) is equivalent to

$$W_n(\theta) = (DF \circ K(\theta))^{-1} \left[\mu^n W_n(\theta + \omega) R_n[W_0, \ldots, W_{n-1}](\theta)\right],$$

which, for large enough n is a contraction, so that we can apply the fast methods of Section A.1. In particular Algorithm A.1. In the case that the stable and unstable directions are one dimensional – which is the one we discuss in this paper – this is enough (remember that we always have $n \geq 2$.) When the bundles are higher dimensional, we may need to find a splitting corresponding to the cocycle generated by $Z(\theta) = (DF \circ K(\theta))^{-1}\mu^n$.

Remark 15. Notice that once $W_0(\theta)$ and $W_1(\theta)$ are fixed, the solution $W_n(\theta)$ for $n \geq 2$ of equation (78) is uniquely determined. It is then clear that any analytic solution is unique. The existence of analytic solutions is discussed in [11].

Remark 16. Notice that the equations to compute the new term W_n do not involve small divisors.

5.2. **A Newton method to compute whiskers.** In this section we will use the automatic reducibility properties to design an efficient Newton method to compute the whiskers.

We assume that we have computed an invariant torus $K(\theta)$ with the associated stable direction $V^s(\theta)$ (resp. unstable direction $V^u(\theta)$) and the rate of contraction μ (resp. expansion). The goal of this section is to compute the power expansion of the whiskers using a quadratically convergent algorithm whose step involves solving a system of linear equations with constant coefficients.

We concentrate on the case of maps, referring to Appendix D for the case of flows.

We consider the invariance equation (18), and we assume that we have an initial approximation W for the whiskers, expressed as a power series

$$W(\theta, s) = \sum_{n=0}^{\infty} W^n(\theta) s^n. \tag{79}$$

We will develop a quasi-Newton method and use a version of automatic reducibility to show that the Newton method only requires to solve constant coefficient equations.

The automatic reducibility will work in general, but the solution of the constant coefficient equations will require that we have

$$W^0(\theta) = K(\theta) \quad \text{and} \quad W^1(\theta) = V^s(\theta),$$

(the unstable case is analogous).

The computation of invariant tori and the linearization has been discussed before in Section 4.

The initial approximation W satisfies

$$F(W(\theta, s)) - W(\theta + \omega, \mu s) = E(\theta, s), \tag{80}$$

where E is the error term. From now on, we use the notation $R(\theta, s) = (\theta + \omega, \mu s)$.

The Newton method to solve equation (18) would require to find $\Delta(\theta, s)$ solving

$$DF \circ W\, \Delta - \Delta \circ R = -E \tag{81}$$

The equation (81) is difficult to solve because the matrix $DF \circ W$ is not constant. Nevertheless, we proceed to show that a version of automatic reducibility holds. That is, using that W is an approximate solution of (18) and the geometry of the problem, there is an explicit change of variables in equation (81) which turns it into a constant coefficient equation.

5.2.1. *Automatic reducibility in the computation of whiskers.* We first give the idea of the automatic reducibility when W is an exact solution, i.e., W is such that

$$(F \circ W)(\theta, s) = (W \circ R)(\theta, s). \tag{82}$$

We will show later that if W is an approximate solution, the reducibility will also hold up to an error term E_R that can be estimated from the error E of the invariance equation (80).

Taking derivatives in (82) with respect to θ and s, we have that

$$\begin{aligned} DF \circ W(\theta, s) D_\theta W(\theta, s) &= D_\theta W(\theta + \omega, \mu s), \\ DF \circ W(\theta, s) \partial_s W(\theta, s) &= \mu \partial_s W(\theta + \omega, \mu s). \end{aligned}$$

From the above equations, we read that the quantity $D_\theta W(\theta, s)$ remains invariant under $DF \circ W(\theta, s)$, whereas the vector $\partial_s W(\theta, s)$ is multiplied by a factor μ.

We can write the above invariance properties more concisely as

$$DF \circ W\, DW = DW \circ R\, DR \tag{83}$$

where $DW = [D_\theta W\,,\, \partial_s W]$ and $DR = \begin{pmatrix} \mathrm{Id}_\ell & 0 \\ 0 & \mu \end{pmatrix}$.

A well known result, which will be important for our algorithm, is that the manifold given by the whiskered tori is Lagrangian (see [45]). Indeed, from (82) we have that $W^* F^* \Omega = R^* W^* \Omega$, where W^*, F^* and R^* denote the pullback by W, F and R, respectively, and Ω is the symplectic 2-form (6). Since $F^*\Omega = \Omega$, then $W^*\Omega = R^*(W^*\Omega)$. It is not hard to see that, if $\Omega = d\alpha$ the only exact 2-form invariant under R is zero.

In our context, the Lagrangian character of the whiskers means that, for a solution W of (18), we have

$$DW^T\, J \circ W\, DW = 0, \tag{84}$$

where J is a skew-symmetric matrix associated to the symplectic form Ω.

Using this property, we construct ss in Section 4.3 a basis of vectors for the tangent space $T\mathcal{M}$. Indeed, the Lagrangian property (84) means that Range DW is orthogonal to Range $(J \circ K)DW$. Hence,

$$\begin{aligned} &\{0\} \\ &= (\mathrm{Range}\, DW \cap \mathrm{Range}\, (J \circ W) DW) \\ &= (J \circ W) \left(\mathrm{Range}\, (J \circ K)^{-1} DW \cap \mathrm{Range}\, DW\right) \end{aligned}$$

where we have used that $J \circ K$, is a full rank matrix. Therefore the the vectors DW and $V = (J^{-1} \circ W)DWN$, where N is a normalization matrix that we choose to be $N = (DW^T DW)^{-1}$, are a basis of $T\mathcal{M}$. Note that the range of the matrix V is independent of the matrix N (provided that N has full range).

Since the differential of F at $W(\theta, s)$, that we denote $DF(W(\theta, s))$, is a linear map from the tangent space of $T\mathcal{M}$ at $W(\theta, s)$ to the tangent space of $T\mathcal{M}$ at $F \circ W(\theta, s) = W \circ R(\theta, s)$, then the $DF \circ W(\theta, s)\, DW(\theta, s)$ is a vector based at $W \circ R(\theta, s)$. Therefore, it can be expressed as a combination of $V \circ R$ and $DW \circ R$, which constitute a basis of $T\mathcal{M}$ at that point. Hence, we can write

$$(DF \circ W)(J^{-1} \circ W)\, DWN = (DW \circ R)A + [((J^{-1} \circ W)\, DWN) \circ R]B, \tag{85}$$

where $A(\theta, s)$ and $B(\theta, s)$ are some matrices, which we will compute in similar way as we did in Proposition 1 (see Proposition 2).

We introduce the matrix M, defined as

$$M = [DW\,,\, (J \circ W)^{-1} DWN] \tag{86}$$

where by $[\cdot, \cdot]$ we denote the $2d \times 2$ matrix obtained by juxtaposing the two $d \times 1$ matrices that are in the arguments. Then, using (83) and (85), we have that

$$(DF \circ W)M = (M \circ R) \begin{pmatrix} DR & A \\ 0 & B \end{pmatrix}. \tag{87}$$

The block A can be computed straightforwardly just as $\Pi_{DW \circ R}(DF \circ W)V$, where $\Pi_{DW \circ R}$ means the projection over the span of $DW \circ R$. There do not seem to be interesting cancellations on the block A. It will, however be important that a subblock on it is invertible. We note for future reference that, if we write $A = \begin{pmatrix} A_{\theta\theta} & A_{\theta,s} \\ A_{s,\theta} & A_{s,s} \end{pmatrix}$, we will assume that $\det(A_{\theta,\theta}) \neq 0$.

The following Proposition 2 provides an explicit expression for B in the invariant case.

Proposition 2. *With the notations above, we have:*

$$B = (DR^T)^{-1} \equiv DR^{-T}. \tag{88}$$

Remark 17. The important point of Proposition (2) is that the matrix in the right hand side of (87) is an upper triangular matrix with constants in the diagonal (note that it is block-upper-triangular and that the blocks in the diagonal are DR and DR^{-T}, which are diagonal with constant terms).

Proof. Consider the expression (85), and multiply both sides on the left by $(DW^T \circ R)\, ((J \circ W) \circ R)$, we obtain

$$\begin{aligned} (DW^T \circ R)\, (J \circ W \circ R)(DF \circ W)(J^{-1} \circ W)\, DWN \\ = (DW^T \circ R)\, ((J \circ W) \circ R)\, ((J^{-1} \circ W)\, DWN) \circ R)B \end{aligned} \tag{89}$$

where we have used that, by the Lagrangian character of the whisker (84), we have $(DW^T \circ R)\, ((J \circ W) \circ R)\, (DW \circ R) = 0$.

Since DF is a symplectic matrix, we have

$$((J \circ W) \circ R)\, (DF \circ W)\, (J^{-1} \circ W) = (DF^{-1} \circ W)^T.$$

Hence, using the equation above as well as equation (83) and the expression for N, we can compute the LHS of (89) as:

$$\begin{aligned}(DW^T \circ R)\,((J \circ W) \circ R)\,(DF \circ W)\,(J^{-1} \circ W)\,DWN \\ &= (DW^T \circ R)\,(DF^{-1} \circ W)^T\,DWN \\ &= [(DF^{-1} \circ W)\,(DW \circ R)]^T\,DWN \\ &= [DWDR^{-1}]^T\,DWN \\ &= DR^{-T}\,DW^T\,DWN \\ &= DR^{-T}.\end{aligned} \tag{90}$$

The RHS of (89) is computed straightforwardly to be just B (recall the expression for N). □

Remark 18. In the approximately invariant case, we note that the Lagrangian character does not hold exactly. However, expression (84) holds up to an error with is controlled by the error E of the invariance equation (81). Hence, (87) holds with an error which is bounded by the error E of the invariance equation. Nevertheless, we still get a quadratically convergent algorithm.

We proceed as in Section 4.3 and we introduce a change of variables in the unknowns of equation (81), by writing $\Delta = MU$, where M is given explicitly in (86). Then, equation (81) is equivalent to the following equation for U

$$(DF \circ W)MU - (M \circ R)(U \circ R) = -E$$

Using that the equation (87) holds up to an error E_R we obtain that (81) is equivalent to

$$(M \circ R)\begin{pmatrix} DR & A \\ 0 & DR^{-\top}\end{pmatrix} U + E_R\,U - (M \circ R)\,(U \circ R) = -E. \tag{91}$$

Ignoring the term $E_R\,U$, we obtain a quasi-Newton equation

$$\begin{pmatrix} DR & A \\ 0 & DR^{-\top}\end{pmatrix} U - U \circ R = -(M^{-1} \circ R)\,E, \tag{92}$$

that is readily solvable provided that the terms W^0 and W^1 in the power expansion (79) solve the invariance equation (82). Indeed, if we write

$$U(\theta, s) = \sum_{k,n} \hat{U}_{k,n} e^{2\pi i k\theta} s^n,$$

equation (92) is equivalent to

$$\begin{pmatrix} DR & A \\ 0 & DR^{-\top}\end{pmatrix} \hat{U}_{k,n} - e^{2\pi i k\omega}\mu^n \hat{U}_{k,n} = \hat{\eta}_{k,n}, \tag{93}$$

where $\hat{\eta}_{k,n}$ are the Fourier-Taylor coefficients of the function $-(M^{-1} \circ R)\,E$.

Notice that the system of equations (93) is a system of linear equations which can always be solved for $n \geq 2$. The reason is that the matrix entering in the LHS of (93) has spectrum $1, \mu, \mu^{-1}$ so that $e^{2\pi i k\omega}\mu^n$ is uniformly far away from the spectrum when $n \geq 2$. Hence, the equations (93) can be solved uniquely and with uniform bounds (there are no small divisors involved). We use here the assumption that we start from a knowledge of the torus and the invariant bundle, so that we do not need to consider $n = 0, 1$.

Remark 19. Notice that since we start with the error being of order 2, after L iterations it will be of order 2^{2^L}, so that in L iterations we have computed exactly $2^{2^L} - 1$ coefficients in the power expansion (79).

Remark 20. Notice that substantial parts of the algorithm do not require that the bundle is 1-dimensional nor that we start from the solution of the first two orders. Hence, there are many variants which can be pursued. We will not detail them here but we hope to come back to them in future projects.

5.2.2. *Formulation of the algorithm for the Newton Method to compute whiskers.* The algorithm to compute the torus and the whiskers follows in a similar way as the algorithm 4.2 to compute the torus in the Lagrangian case. Notice that we do not need to compute the splitting explicitly, but, in exchange, we have to work with functions of two variables.

Algorithm 5.1 (Newton step to compute the whiskers)**.** Given F, ω, μ as well as W, an approximate solution which solves the invariance equation (82) up to order 2, i.e.

$$F(W(\theta, s)) - W(\theta + \omega, \mu s) = E^{[\geq L]}(\theta, s)$$

with $L \geq 2$.

Perform the following calculations:

1. Compute $E^{[\geq L]}(\theta, s) = F \circ W(\theta, s) - W(\theta + \omega, \mu s)$
2. Compute
   - (2.1) $\alpha(\theta, s) = [D_\theta W(\theta, s)\,,\, \partial_s W(\theta, s)]$
   - (2.2) $N(\theta, s) = (\alpha^\top(\theta, s)\, \alpha(\theta, s))^{-1}$.
   - (2.3) $\beta(\theta, s) = \alpha(\theta, s)\, N(\theta, s)$
   - (2.4) $\gamma(\theta, s) = (J^{-1} \circ W)(\theta, s)\, \beta(\theta, s)$
   - (2.5) $M(\theta, s) = [\alpha(\theta, s)\,,\, \gamma(\theta, s)]$
   - (2.6) $M^{-1}(\theta, s)$
   - (2,8) Π_{DW}, Π_V, the projections corresponding to the decomposition of the space into the span of DW and the span of $V = (J^{-1} \circ W)DWN$.
3. Compute
$$\widetilde{E}^{[\geq L]}(\theta, s) = -M^{-1}(\theta + \omega, \mu s)\, E^{[\geq L]}(\theta, s)$$
4. Compute
   - (4.1) $A(\theta, s) = \Pi_{DW \circ R} DF \circ W(\theta, s)\, \beta(\theta, s)$,
   - (4.2) $B(\theta, s) = DR^{-T}$,
5. We denote $U_{DW} = \Pi_{DW} U$ and $U_V = \Pi_V U$.
   - (5.1) Solve for U_{DW} satisfying
$$\begin{pmatrix} \mathrm{Id}_\ell & 0 \\ 0 & \mu \end{pmatrix} U_{DW} - U_{DW} \circ R = \Pi_{DW} \tilde{E}$$
   - (5.2) Solve for U_V satisfying
$$\begin{pmatrix} \mathrm{Id}_\ell & 0 \\ 0 & 1/\mu \end{pmatrix} U_V - U_V \circ R = \Pi_V \tilde{E} - A U_{DW}$$

   Note that both equations can be solved in Fourier-Taylor coefficients.
6. Set $\tilde{W} = W + MU$

Remark 21. We think that Algorithm 5.1 is advantageous with respect to the order by order method described in Section 5.1. In the order–by–order method, errors in first step (round off errors) propagate in the subsequent steps and never get

corrected. However, using the Newton method described in this section, numerical errors eventually get corrected.

Remark 22. There are two important steps in Algorithm 5.1: computing the residual E and computing the correction U.

The computation of the residual of the Newton method involves manipulating a polynomial up to order 2^N. Depending on how the computation of composition is done, it could be the bottleneck step for the speed of the algorithm. (for example, if it is done using the direct substitution).

Nevertheles, we note that there are fast algorithms to compose polynomials. The strategy is to use the FFT to compute the polynomials in a complex circle. Then, one can evaluate the composition. A subsequent inverse FFT recovers the coefficients of the composition. More details on this algorithm appear in [35].

Depending on the algorithms one uses for composition and multiplication of series, the operation count of Algorithm 5.1 and the step by step method may be very different.

Appendix A. **Fast algorithms to solve difference equations with non constant coefficients.** In this section we present fast algorithms to solve for $\Delta(\theta)$ the cohomology equation with non constant coefficients

$$A(\theta)\Delta(\theta) - \Delta(\theta+\omega)B(\theta) = \eta(\theta) \tag{94}$$

for given $A(\theta)$, $B(\theta)$ and $\eta(\theta)$ satisfying either $\|A\| < 1$, $\|B^{-1}\| < 1$ or $\|A^{-1}\| < 1$, $\|B\| < 1$.

Equations of this form appear in the Netwon step for whiskered tori (See the informal description in Section 1). Equations of this form also appear in the calcultion of the invariant splitting (see (68)-(69)).

We will present two algorithms. The first one is an iterative method with an accelerated convergence and the second one very fast (see Section A.1). The second one is only for the case of one-dimensional bundles and it is faster (computations are $O(N)$)(see Section A.2).

A.1. **Fast iterative algorithms for the cohomology equation.** In this section we will present a fast algorithm to solve equation (94) using iterative methods. We refer the reader to [25] where a similar idea is used to compute iteration of cocycles.

We consider first the case $\|A^{-1}\| < 1$ and $\|B\| < 1$ or, more generally, $\|A^{-1}(\theta)\| \cdot \|B(\theta)\| < 1$ Then, multiplying (94) by $A^{-1}(\theta)$ on the LHS, we obtain

$$\Delta(\theta) = A^{-1}(\theta)\Delta(\theta+\omega)B(\theta) + A^{-1}(\theta)\eta(\theta). \tag{95}$$

This is a contraction mapping and it is straightforward to iterate it and obtain an algorithm that converges faster than exponentially.

Next, we compute $\Delta(\theta+\omega)$ by shifting (95) and substituting again in (95), so that we get

$$\begin{aligned}\Delta(\theta) = {} & A^{-1}(\theta)\eta(\theta) \\ & + A^{-1}(\theta)A^{-1}(\theta+\omega)\eta(\theta+\omega)B(\theta) \\ & + A^{-1}(\theta)A^{-1}(\theta+\omega)\Delta(\theta+2\omega)B(\theta+\omega)B(\theta).\end{aligned}$$

Notice that if we define

$$\bar\eta(\theta) = A^{-1}(\theta)\eta(\theta)$$

and

$$\begin{aligned} A_1^{-1}(\theta) &= A^{-1}(\theta)A^{-1}(\theta+\omega), \\ B_1(\theta) &= B(\theta+\omega)B(\theta), \\ \eta_1(\theta) &= \bar{\eta}(\theta) + A^{-1}(\theta)\bar{\eta}(\theta+\omega)B(\theta), \end{aligned}$$

we have that

$$\Delta(\theta) = \eta_1(\theta) + A_1^{-1}(\theta)\Delta(\theta+2\omega)B_1(\theta)$$

which has the same structure as (95) and we can repeat the same scheme. This leads to an iterative procedure to compute $A(\theta)$, converging superexponentially in the number of iterations. Thus, define

$$\begin{aligned} A_{n+1}^{-1}(\theta) &= A_n^{-1}(\theta)A_n^{-1}(\theta+2^n\omega), \\ B_{n+1}(\theta) &= B_n(\theta+2^n\omega)B_n(\theta), \\ \eta_{n+1}(\theta) &= \eta_n(\theta) + A_n^{-1}(\theta)\eta_n(\theta+2^n\omega)B_n(\theta), \end{aligned}$$

for $n \geq 0$, with

$$\begin{aligned} A_0^{-1}(\theta) &= A^{-1}(\theta), \\ B_0(\theta) &= B(\theta), \\ \eta_0(\theta) &= \bar{\eta}(\theta). \end{aligned}$$

Then,

$$\Delta(\theta) = \eta_{n+1}(\theta) + A_{n+1}^{-1}(\theta)\Delta(\theta+2^{n+1}\omega)B_{n+1}(\theta), \quad \forall\ n \geq 0$$

and

$$\Delta(\theta) = \lim_{n\to+\infty} \eta_{n+1}(\theta).$$

The convergence of the algorithm is guaranteed by the contraction of A^{-1} and B. The cost of computing 2^N terms in the sum is proportional to N since it involves only N steps of the algorithm.

The iterative algorithm is the following:

Algorithm A.1 (Solution of difference equations with non constant coefficient)**.** Given $A(\theta)$, $B(\theta)$ such that $\|A^{-1}(\theta)\| \cdot \|B(\theta)\| \leq \kappa < 1$, and $\eta(\theta)$ perform the following operations:

1. Compute $\Delta(\theta) = A^{-1}(\theta)\eta(\theta)$
2. Compute
   (2.1) $\tilde{\Delta}(\theta) = A^{-1}(\theta)\Delta(\theta+\omega)B(\theta) + \Delta(\theta)$
   (2.2) $\tilde{A}^{-1}(\theta) = A^{-1}(\theta)A^{-1}(\theta+\omega)$
   (2.3) $\tilde{B}(\theta) = B(\theta+\omega)B(\theta)$
3. Set $\tilde{\Delta} \to \Delta$
   $\tilde{A} \to A$
   $\tilde{B} \to B$
   $2\omega \to \omega$
4. Iterate points $2-3$

The case when $\|A\| < 1$ and $\|B^{-1}\| < 1$ can be done similarly. In this case, we multiply (94) by $B^{-1}(\theta)$ on the LHS so that we obtain

$$\Delta(\theta+\omega) = A(\theta)\Delta(\theta)B^{-1}(\theta) - \eta(\theta)B^{-1}(\theta).$$

Before applying the iterative scheme we shift by $-\omega$. In this way, we have

$$\Delta(\theta) = A(\theta')\Delta(\theta')B^{-1}(\theta') - \eta(\theta')B^{-1}(\theta')$$

where $\theta' = T_{-\omega}\theta$.

The iterative algorithm in this case is

Algorithm A.2. Given $A(\theta)$, $B(\theta)$ $\|A(\theta)\|\|B^{-1}(\theta)\| \leq \kappa < 1$ and $\eta(\theta)$, perform the following operations:

1. Compute $\Delta(\theta) = -\eta(\theta)B^{-1}(\theta)$
2. Compute
   (2.1) $\tilde{\Delta}(\theta) = A(\theta)\Delta(\theta-\omega)B^{-1}(\theta) + \Delta(\theta)$
   (2.2) $\tilde{A}(\theta) = A(\theta)A(\theta-\omega)$
   (2.3) $\tilde{B}^{-1}(\theta) = B^{-1}(\theta-\omega)B^{-1}(\theta)$
3. Set
   $\tilde{\Delta} \to \Delta$
   $\tilde{A} \to A$
   $\tilde{B} \to B$
   $2\omega \to \omega$
4. Iterate parts 2–3

This algorithm gives $\Delta(\theta+\omega)$. Shifting it by $-\omega$ we get $\Delta(\theta)$.

A.2. **Fast algorithm for the 1-D cohomology equation with non-constant coefficients.** In this section we present an efficient algorithm for the one-dimensional version of equation (94). It is an adaptation of methods used in [26].

Consider the following equation,

$$\frac{A(\theta)}{B(\theta)}\Delta(\theta) - \Delta(\theta+\omega) = \frac{\eta(\theta)}{B(\theta)} \tag{96}$$

which is obtained from (94) multiplying by $B^{-1}(\theta)$ (recall that in this case $B(\theta)$ is just a number).

We will solve (96) in two steps:

1. Find $C(\theta)$ and $\nu \in \mathbb{R}$ such that
$$\frac{A(\theta)}{B(\theta)} = \nu\frac{C(\theta)}{C(\theta+\omega)} \tag{97}$$
2. Applying (97) in (96) and multiplying by $C(\theta+\omega)$ we obtain
$$\nu C(\theta)\Delta(\theta) - C(\theta+\omega)\Delta(\theta+\omega) = \tilde{\eta}(\theta) \tag{98}$$

where $\tilde{\eta}(\theta) = C(\theta+\omega)B^{-1}(\theta)\eta(\theta)$.

If we change the unknowns in (98) by $W(\theta) = C(\theta)\Delta(\theta)$, we are left with the equation

$$\nu W(\theta) - W(\theta+\omega) = \tilde{\eta}(\theta). \tag{99}$$

Of course, if $|\nu| \neq 1$, equation (99) can be solved in Fourier space. That is, we can obtain the Fourier coefficients of W as:

$$\widehat{W}_k = \frac{\widehat{\tilde{\eta}}_k}{\nu - e^{2\pi i k\omega}} \ ,$$

and the solution is unique. Notice that whenever $|\nu| = 1$, equation (99) involves small divisors, which is not the case for the linearized Newton equations restricted to the hyperbolic subspaces.

Finally, once we have $W(\theta)$ we get

$$\Delta(\theta) = C^{-1}(\theta)W(\theta).$$

Step 1 can be achieved by taking logarithms of (97). Assume that $A(\theta)/B(\theta) > 0$, otherwise we change the sign. Then, we get

$$\log A(\theta) - \log B(\theta) = \log \nu + \log C(\theta) - \log C(\theta + \omega).$$

Taking $\log \nu$ to be the average of $\log A(\theta) - \log B(\theta)$, the problem reduces to solve for $\log C(\theta)$ an equation of the form (23). Then $C(\theta)$ and ν can be obtained by exponentiation. Notice that $\log C(\theta)$ is determined up to a constant. We will fix it by assuming that it has average 0.

Hence, we have the following algorithm:

Algorithm A.3 (Solution of difference equations with non constant coefficient $(1D)$)**.** Given $A(\theta)$, $B(\theta)$ and $\eta(\theta)$. Perform the following operations:

1. (1.1) Compute $L(\theta) = \log(A(\theta)) - \log(B(\theta))$
   (1.2) Compute $\overline{L} = \int_{\mathbb{T}^\ell} L$
2. Solve for L_C satisfying
$$L_C(\theta) - L_C \circ T_\omega(\theta) = L(\theta) - \overline{L}$$
as well as having zero average.
3. (3.1) Compute $C(\theta) = \exp(L_C(\theta))$
   (3.2) Compute $\nu = \exp(\overline{L})$
4. Compute $\tilde{\eta}(\theta) = C(\theta + \omega)B^{-1}(\theta)\eta(\theta)$
5. Solve for W satisfying
$$\nu W(\theta) - W(\theta + \omega) = \tilde{\eta}(\theta)$$
6. The solution of (94) is $\Delta(\theta) = C^{-1}(\theta)W(\theta)$

Appendix B. **Fast Newton method for whiskered tori in Hamiltonian flows: The center directions.** In this section, we provide the numerical algorithm to solve the invariance equation (12) and the modified one (14) using a Newton method analogous to the one described in Section 4.3.

The automatic reducibility can also be proved in this context (see [15]) and we provide here the algorithm only.

Algorithm B.1 (Newton step for flows in the center direction)**.** Consider given $X = J(K)\nabla H$, ω, K_0 and an approximate solution K (resp. K, λ). Perform the following calculations

1. (1.1) Compute $\partial_\omega K$.
   (1.2) Compute $X \circ K$ (1.3) Compute the invariant projections Π^c, Π^u, Π^s
2. Set $E^c = \Pi^c(\partial_\omega K - X \circ K)$ (resp. set $E^c = \Pi^c(\partial_\omega K - X \circ K - G\lambda)$)
3. Following (40)
   (3.1) Compute $\alpha(\theta) = DK(\theta)$
   (3.2) Compute $N(\theta) = \left([\alpha(\theta)]^\top \alpha(\theta)\right)^{-1}$
   (3.3) Compute $\beta(\theta) = \alpha(\theta)N(\theta)$
   (3.4) Compute $\gamma(\theta) = J(K_0(\theta))_{\mathcal{E}^c}^{-1}\beta(\theta)$
   (3.5) Compute $M(\theta) = [\alpha(\theta) \mid \gamma(\theta)]$
   (3.6) Compute $M(\theta + \omega)$
   (3.7) Compute $(M(\theta + \omega)^\top J(K_0(\theta + \omega))_{\mathcal{E}^c}^{-1} M(\theta + \omega))^{-1}$
   (3.8) Compute $\widetilde{E}(\theta) = (M(\theta + \omega)^\top J(K_0(\theta + \omega))_{\mathcal{E}^c}^{-1} M(\theta + \omega))^{-1} E^c(\theta)$
   (3.9) Compute
$$S(\theta) = \beta^\top(\theta)(\mathrm{Id}_{2d} - \beta(\theta)\alpha(\theta)^\top)(DX(K(\theta)) + DX(K(\theta))^\top)\beta(\theta).$$

4. (4.1) Solve for U_2 satisfying

$$\partial_\omega U_2 = -\widetilde{E}_2 + \int_{\mathbb{T}^\ell} \widetilde{E}_2$$

(resp.

(4.1′) Solve for δ satisfying

$$\int_{\mathbb{T}^\ell} \widetilde{E}_2 - \left[\int_{\mathbb{T}^\ell} B_2\right]\delta = 0$$

(4.2′) Solve for U_2 satisfying

$$\partial_\omega U_2 = -\widetilde{E}_2 + B_2\delta$$

Set U_2 such that its average is 0.)

5. (5.1) Compute $S(\theta)U_2(\theta)$
(5.2) Solve for $\overline{U}_2$ satisfying

$$\int_{\mathbb{T}^\ell} \widetilde{E}_1(\theta) + \int_{\mathbb{T}^\ell} S(\theta)U_2(\theta) + \left[\int_{\mathbb{T}^\ell} S(\theta)\right]\overline{U}_2 = 0$$

(5.3) Find U_1 solving

$$\partial_\omega U_1 = -\widetilde{E}_1 - S(U_2 + \overline{U}_2)$$

Normalize it so that $\int_{\mathbb{T}^\ell} U_1 = 0$
(resp.
(5.1′) Compute $S(\theta)U_2(\theta)$
(5.2′) Solve for $\overline{U}_2$ satisfying

$$\int_{\mathbb{T}^\ell} \widetilde{E}_1(\theta) + \int_{\mathbb{T}^\ell} B_1(\theta)\delta - \int_{\mathbb{T}^\ell} S(\theta)U_2(\theta) - \left[\int_{\mathbb{T}^\ell} S(\theta)\right]\overline{U}_2 = 0$$

(5.3′) Find U_1 solving

$$\partial_\omega U_1 = -\widetilde{E}_1 - S(U_2 + \overline{U}_2) + B_1\delta$$

Normalize it so that $\int_{\mathbb{T}^\ell} U_1 = 0$).

6. The improved K is $K(\theta) + M(\theta)U(\theta)$
(resp. the improved λ is $\lambda + \delta$).

Notice that steps (1.1), (3.1), (4.1) (resp. (4.2′)), (5.3) (resp. (5.3′)) in Algorithm B.1 are diagonal in Fourier series, whereas the other steps are diagonal in the real space representation. The algorithm only stores vectors which are of length of order N.

Appendix C. The order by order method for whiskers in Hamiltonian flows. In this section we present the result analogous to the one described in Section 5.1 to solve the invariance equation (18) for the whiskers in the case of Hamiltonian flows.

As in Section 5.1 we look for W as a power series

$$W(\theta, s) = \sum_{n=0}^{\infty} W_n(\theta)s^n,$$

and match similar coefficients in s^n on both sides of equation (18).

For $n = 0$, one obtains

$$\partial_\omega W_0(\theta) = (X \circ W_0)(\theta) \tag{100}$$

which admits the solution $W_0(\theta) = K(\theta)$, where K is a parametrization of the invariant torus.

For $n = 1$, we obtain

$$\partial_\omega W_1(\theta) + W_1(\theta)\mu = (DX \circ K(\theta))W_1(\theta), \tag{101}$$

from where we read that $W_1(\theta)$ is an eigenfunction with eigenvalue $-\mu$ of the operator L_ω

$$L_\omega := \partial_\omega - DX \circ K(\theta).$$

Again, we note that, multiplying a solution of (101) by a scalar $b \in \mathbb{R}$, we also obtain a solution. See Remark 14.

For $n \geq 2$, we obtain

$$\partial_\omega W_n(\theta) + W_n(\theta)n\mu = (DX \circ K(\theta))W_n(\theta) + R_n(W_0, \ldots, W_{n-1}), \tag{102}$$

where R_n is an explicit polynomial in $W_0, \ldots, W_{n-1}$ whose coefficients are derivatives of X evaluated at $W_0 = K$.

Notice that, in this case, equation (102) can be solved provided that $n\mu$ is not in the spectrum of the operator L_ω (this is a non-resonance condition which is clearly satisfied since the stable spaces are 1-dimensional). As in the case of maps, the previous equation can be solved using the large matrix method.

Appendix D. **A Newton method to compute the whiskers for flows.** In this section we present the result analogous to the one described in Section 5.2 to solve the invariance equation (18) for the whiskers in the case of Hamiltonian flows.

As in Section 5.2, we start with an initial approximation W for the invariance equation (18), that solves it exactly up to order 2, i.e.

$$X(W(\theta, s)) - (\partial_\omega + \mu s \partial_s)\, W(\theta, s) = E^{[\geq L]}(\theta, s), \tag{103}$$

for $L \geq 2$. From now on, we will omit the superindex $[\geq L]$ in the error term.

The Newton method to solve (18) involves looking for an improved solution

$$W \to W + \Delta,$$

where Δ is a solution of the linearized equation

$$(DX \circ W)(\theta, s)\Delta(\theta, s) - (\partial_\omega + \mu s \partial_s)\, \Delta(\theta, s) = -E(\theta, s). \tag{104}$$

Here, we will show that we can use a version of automatic reducibility for flows to transform equation (104) into a constant coefficient equation by means of a change of coordinates.

Automatic reducibility for flows in the computation of whiskers. As we did for the case of maps is Section 5.2 we will discuss the automatic reducibility for the case when W is an exact solution, i.e.

$$X(W(\theta, s)) - (\partial_\omega + \mu s \partial_s)\, W(\theta, s) = 0. \tag{105}$$

Applying the operators D_ω and ∂_s to equation (105), we have that

$$\begin{aligned} DX(W(\theta, s))D_\theta W(\theta, s) - (\partial_\omega + \mu s \partial_s)\, D_\theta W(\theta, s) &= 0, \\ DX(W(\theta, s))\partial_s W(\theta, s) - (\partial_\omega + \mu s \partial_s)\, \partial_s W(\theta, s) &= \mu \partial_s W(\theta, s) \end{aligned}$$

Defining the vector field

$$\mathcal{X} := \begin{bmatrix} \omega \\ \mu s \end{bmatrix},$$

we can write the above invariance equations in a compact way as

$$(DX \circ W)DW - \mathcal{L}_{\mathcal{X}} DW = DW D\mathcal{X} \tag{106}$$

where $\mathcal{L}_{\mathcal{X}} := \omega\partial_\theta + \lambda\mu\partial_s$, $DW = [D_\theta W, \partial_s W]$ and $D\mathcal{X} = \begin{pmatrix} 0_\ell & 0 \\ 0 & \mu \end{pmatrix}$.

As in the case of maps, the whisker for the torus is a Lagrangian manifold, i.e. for a solution W of (105) we have that

$$DW^T\, J \circ W\, DW = 0, \tag{107}$$

where J is a skew-symmetric matrix associated to the symplectic form Ω.

Exactly as in the case of maps, the vectors DW and $V = (J \circ W)^{-1} DWN$, with $N = (DW^T DW)^{-1}$ constitute a basis of the tangent space $T\mathcal{M}$. Hence, we can write

$$(DX \circ W)(J^{-1} \circ W)DWN - \mathcal{L}_{\mathcal{X}}((J^{-1} \circ W)DWN) = DWA + (J^{-1} \circ W)DWNB. \tag{108}$$

Thereby, introducing the matrix M defined as

$$M = [DW, (J^{-1} \circ W)DWN], \tag{109}$$

and using equations (106) and (108) we have that

$$(DX \circ W)M - \mathcal{L}_{\mathcal{X}} M = M \begin{pmatrix} D\mathcal{X} & A \\ 0 & B \end{pmatrix}. \tag{110}$$

The following Proposition 3 provides explicit expressions for A and B in the invariant case.

Proposition 3. *With the notations above, we have:*

$$B = -D\mathcal{X}^T \tag{111}$$

and

$$A = -NDW^T(J^{-1} \circ W)[\mathrm{Id}_{2d} - DWNDW^T][DX + DX^T]DWN \tag{112}$$

Remark 23. As in the case of maps, the important point of Proposition (3) is that the matrix in the right hand side of (87) is an upper triangular matrix with constants in the diagonal (note that it is block-upper-triangular and that the blocks in the diagonal are $D\mathcal{X}$ and $-D\mathcal{X}^T$, which are diagonal with constant terms).

Remark 24. If the symplectic structure is almost-complex (i.e. $J^2 = -\mathrm{Id}$), we have that

$$DW^T(J^{-1} \circ W)DW = 0,$$

since the whisker is a Lagrangian manifold. Then A has a simpler expression given by

$$A = -NDW^T(J^{-1} \circ W)[DX + DX^T]DWN.$$

Proof. Consider the expression (108), and multiply both sides on the left by $DW^T(J \circ W)$, we obtain

$$DW^T(J \circ W)[(DX \circ W)(J^{-1} \circ W)DWN - \mathcal{L}_{\mathcal{X}}((J^{-1} \circ W)DWN)] = B \tag{113}$$

where we have used the Lagrangian character of the whisker (84) and the definition of N.

We first expand the term $\mathcal{L}_{\mathcal{X}}((J^{-1}\circ W)DWN)$:

$$\begin{aligned}\mathcal{L}_{\mathcal{X}}((J^{-1}\circ W)DWN) =&\mathcal{L}_{\mathcal{X}}(J^{-1}\circ W)DWN + (J^{-1}\circ W)\mathcal{L}_{\mathcal{X}}(DW)N\\ &+ (J^{-1}\circ W)DW\mathcal{L}_{\mathcal{X}}(N).\end{aligned} \tag{114}$$

By differentiation of $(J\circ W)(J^{-1}\circ W) = \mathrm{Id}$ and using the Hamiltonian character of the vector field

$$DX^T(J\circ W) + (J\circ W)DX + (DJ\circ K)(X\circ W) = 0,$$

we have that

$$\mathcal{L}_{\mathcal{X}}(J^{-1}\circ W) = (J^{-1}\circ W)[DX^T(J\circ W) + (J\circ W)DX](J^{-1}\circ W). \tag{115}$$

Similarly, differentiating $N\,N^{-1} = \mathrm{Id}$ and using the invariance equation (106), we have that

$$\mathcal{L}_{\mathcal{X}}(N) = -NDW^T[DX^T + DX]DWN + ND\mathcal{X}^T + D\mathcal{X}N. \tag{116}$$

Using the previous calculations (115) and (116) as well as the invariance equation (106), from expression (113) we have that

$$B = -D\mathcal{X}^T.$$

Now, multiplying equation (108) by NDW^T from the left, and using the definition of N we have

$$\begin{aligned}&NDW^T[(DX\circ W)(J^{-1}\circ W)DWN - \mathcal{L}_{\mathcal{X}}((J^{-1}\circ W)DWN)]\\ &\quad = A + N\,DW^T(J^{-1}\circ W)DWNB.\end{aligned} \tag{117}$$

Using the previous calculations (116) and (115), the invariance equation (106) and the expression for B, we have

$$A = -NDW^T(J^{-1}\circ W)[\mathrm{Id}_{2d} - DWNDW^T][DX + DX^T]DWN.$$

□

Remark 25. The expressions in Proposition 3 are true when W is an exact solution. For the case the W is an approximate solution, reducibility (110) holds up to an error which is controlled by the error E of the invariance equation (103). Nevertheless, we still obtain a quadratically convergent algorithm.

We proceed as in Section 5.2 and we introduce a change of variables in the unknowns of equation (104) by writing $\Delta = MU$, where M is given explicitly in (109). Then, equation (104) reads:

$$(DX\circ W)MU - \mathcal{L}_{\mathcal{X}}(MU) = -E.$$

and working out the expression for $\mathcal{L}_{\mathcal{X}}(MU)$ we have

$$[(DX\circ W)M - \mathcal{L}_{\mathcal{X}}(M)]U - M\mathcal{L}_{\mathcal{X}}(U) = -E. \tag{118}$$

Using the reducibility equation (110), which holds up to an error E_R we obtain that (104) is equivalent to

$$M\begin{pmatrix} D\mathcal{X} & A\\ 0 & -D\mathcal{X}^T\end{pmatrix}U + E_RU - M\mathcal{L}_{\mathcal{X}}(U) = -E. \tag{119}$$

Multiplying the above expression by M^{-1} and ignoring the term E_RU, we obtain the quasi-Newton equation

$$\begin{pmatrix} D\mathcal{X} & A\\ 0 & -D\mathcal{X}^T\end{pmatrix}U - \mathcal{L}_{\mathcal{X}}(U) = -M^{-1}E. \tag{120}$$

Equation (120) can be solved provided that the term W^0, W^1 in the power expansion (79) solve the invariance equation (80). Indeed, if we write

$$U(\theta, s) = \sum_{k,n} \hat{U}_{k,n} e^{2\pi i k\theta} s^n,$$

equation (120) is equivalent to

$$\begin{pmatrix} D\mathcal{X} & A \\ 0 & -D\mathcal{X}^T \end{pmatrix} \hat{U}_{k,n} - (2\pi i k \cdot \omega + \mu n)\hat{U}_{k,n} = \hat{\eta}_{k,n}, \tag{121}$$

where $\hat{\eta}_{k,n}$ are the Fourier-Taylor coefficients of the function $-M^{-1}\, E$.

Notice that the system of equations (93) is a system of linear equations which can always be solved for $n \geq 2$, provided that $2\pi i k \cdot \omega + n\mu$ is not in the spectrum of the matrix entering in the LHS of (121). This condition is true because the matrix has spectrum $0, \mu, -\mu$ so that $2\pi i k \cdot \omega + n\mu$ is uniformly far from the spectrum when $n \geq 2$. Hence, the equations (93) can be solved uniquely and with uniform bounds (there are no small divisors involved). We use here the the assumption that we start from a knowledge of the torus and the invariant bundle, so that we do not need to consider the cases $n = 0, 1$.

Formulation of the algorithm for the Newton Method to compute whiskers. The algorithm to compute the whiskers for flows follows in the following way:

Algorithm D.1 (Newton step to compute the whiskers)**.** Given X, ω, μ as well as W, an approximate solution which solves the invariance equation (82) up to order 2, i.e.

$$X \circ W(\theta, s) - (\partial_\omega + \mu s \partial_s) W(\theta, s) = E^{[\geq L]}(\theta, s)$$

with $L \geq 2$.

Perform the following calculations:

1. Compute $E^{[\geq L]}(\theta, s) = X \circ W(\theta, s) - (\partial_\omega + \mu s \partial_s) W(\theta, s)$
2. Compute
   - (2.1) $\alpha(\theta, s) = [D_\theta W(\theta, s)\,,\, \partial_s W(\theta, s)]$
   - (2.2) $N(\theta, s) = (\alpha^\top(\theta, s)\, \alpha(\theta, s))^{-1}$.
   - (2.3) $\beta(\theta, s) = \alpha(\theta, s)\, N(\theta, s)$
   - (2.4) $\gamma(\theta, s) = (J^{-1} \circ W)\beta(\theta, s)$
   - (2.5) $M(\theta, s) = [\alpha(\theta, s)\,,\, \gamma(\theta, s)]$
   - (2.6) $M^{-1}(\theta, s)$
   - (2,8) Π_{DW}, Π_V, the projections corresponding to the decomposition of the space into the span of DW and the span of $V = (J^{-1} \circ W)DWN$.
3. Compute
$$\widetilde{E}^{[\geq L]}(\theta, s) = -M^{-1}(\theta, s)\, E^{[\geq L]}(\theta, s)$$
4. Compute
   - (4.1) $A(\theta, s) = \gamma^T(\theta, s)[\mathrm{Id}_{2d} - \beta(\theta, s)\alpha^T(\theta, s)](DX \circ W + DX^T \circ W)(\theta, s)\beta(\theta, s)$,
   - (4.2) $B(\theta, s) = -D\mathcal{X}^T$,
5. We denote $U_{DW} = \Pi_{DW} U$ and $U_V = \Pi_V U$.
   - (5.1) Solve for U_{DW} satisfying
$$\begin{pmatrix} 0_\ell & 0 \\ 0 & \mu \end{pmatrix} U_{DW} - (\partial_\omega + \mu s \partial_s) U_{DW} = \Pi_{DW} \tilde{E}$$

(5.2) Solve for U_V satisfying

$$\begin{pmatrix} 0_\ell & 0 \\ 0 & -\mu \end{pmatrix} U_V - (\partial_\omega + \mu s \partial_s) U_V = \Pi_V \tilde{E} - A U_{DW}$$

Note that both equations can be solved in Fourier-Taylor coefficients.

6. Set $\tilde{W} = W + MU$

Acknowledgments. We are specially thankful to Á Haro for several important improvements on the algorithms and clarifications.

We also want to thank R. Calleja, A. Celletti, A. Luque, J. M. Mondelo and C. Simó for several discussions and for comments on the paper. The manuscript was written while we were visiting CRM during the Research Programme *Stability and Instability in Mechanical Systems (SIMS08)*, for whose hospitality we are very grateful. G. H. and Y. S. would like to thank the hospitality of the department of Mathematics of University of Texas at Austin, where part of this work was carried out. GH and RL also thank IMA where we taught a course on these methods, which clarified the presentation.

REFERENCES

[1] V. I. Arnold, *Instability of dynamical systems with several degrees of freedom*, Sov. Math. Doklady, **5** (1964), 581–585.

[2] S. Aubry, *The twist map, the extended Frenkel-Kontorova model and the devil's staircase. Order in chaos*, (Los Alamos, N.M., 1982), Phys. D, **7** (1983), 240–258.

[3] S. Aubry and P. Y. Le Daeron, *The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states*, Phys. D, **8** (1983), 381–422.

[4] A. Celletti and L. Chierchia, *KAM stability and celestial mechanics*, Mem. Amer. Math. Soc., **187** (2007), viii+134 pp.

[5] A. Celletti, C. Falcolini and U. Locatelli, *On the break-down threshold of invariant tori in four dimensional maps*, Regul. Chaotic Dyn., **9** (2004), 227–253.

[6] R. Calleja and R. de la Llave, *Fast numerical computation of quasi-periodic equilibrium states in 1D statistical mechanics, including twist maps*, Nonlinearity, **22** (2009), 1311–1336.

[7] R. Calleja and R. de la Llave, *A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification*, Nonlinearity, **23** (2010), 2029–2058.

[8] R. Calleja and R. de la Llave, *Computation of the breakdown of analyticity in statistical mechanics models: Numerical results and a renormalization group explanation*, J. Stat. Phys, **141** (2010), 940–951.

[9] X. Cabré, E. Fontich and R. de la Llave, *The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces*, Indiana Univ. Math. J., **52** (2003), 283–328.

[10] X. Cabré, E. Fontich and R. de la Llave, *The parameterization method for invariant manifolds. II. Regularity with respect to parameters*, Indiana Univ. Math. J., **52** (2003), 329–360.

[11] X. Cabré, E. Fontich and R. de la Llave, *The parameterization method for invariant manifolds. III. Overview and applications*, J. Differential Equations, **218** (2005), 444–515.

[12] A. Delshams and G. Huguet, *Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems*, Nonlinearity, **22** (2009), 1997–2077.

[13] A. Delshams, R. de la Llave and T. 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** (2006), viii+141 pp.

[14] P Duarte, *Plenty of elliptic islands for the standard family of area preserving maps*, Ann. Inst. H. Poincaré Anal. Non Linéaire, **11** (1994), 359–409.

[15] E. Fontich, R. de la Llave and Y. Sire, *Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions*, J. Differential Equations, **246** (2009), 3136–3213.

[16] E. Fontich, R. de la Llave and Y. Sire, *A method for the study of whiskered quasi-periodic and almost-periodic solutions in finite and infinite dimensional Hamiltonian systems*, Electron. Res. Announc. Math. Sci., **16** (2009), 9–22.

[17] F. Fassò, M. Guzzo and G. Benettin, *Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems*, Comm. Math. Phys., **197** (1998), 347–360.
[18] M. Frigo and S. G. Johnson, *The design and implementation of FFTW3*, Proceedings of the IEEE, **93** (2005), 216–231.
[19] B. Fayad, A. Katok and A. Windsor, *Mixed spectrum reparameterizations of linear flows on* $\mathbb{T}^2$, Mosc. Math. J., **1** (2001), 521–537, 644.
[20] M. Guzzo, F. Fassò and G. Benettin, *On the stability of elliptic equilibria*, Math. Phys. Electron. J., **4** (1998), Paper 1, 16 pp. (electronic).
[21] Samuel M. Graff, *On the conservation of hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations, **15** (1974), 1–69.
[22] J. M. Greene, *A method for determining a stochastic transition*, Jour. Math. Phys., **20** (1979), 1183–1201.
[23] G. H. Golub and C. F. Van Loan, "Matrix Computations," Third edition, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 1996.
[24] Alex Haro, *Automatic differentiation tools in computational dynamical systems*, Manuscript, 2008.
[25] G. Huguet, R. de la Llave and Y. Sire, *Fast iteration of cocyles over rotations and Computation of hyperbolic bundles*, preprint, `arXiv:1102.2461`.
[26] Michael-R. Herman, *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1*, With an appendix by Albert Fathi, With an English summary, in "Astérisque," 103–104, Société Mathématique de France, Paris, 1983.
[27] M.-R. Herman, *On the dynamics of Lagrangian tori invariant by symplectic diffeomorphisms*, in "Progress in Variational Methods in Hamiltonian Systems and Elliptic Equations" (L'Aquila, 1990), Pitman Res. Notes Math. Ser., **243**, Longman Sci. Tech., Harlow, (1992), 92–112.
[28] A. Haro and R. de la Llave, *New mechanisms for lack of equipartion of energy*, Phys. Rev. Lett., **89** (2000), 1859–1862.
[29] À. Haro and R. de la Llave, *Manifolds on the verge of a hyperbolicity breakdown*, Chaos, **16** (2006), 013120, 8 pp.
[30] À. Haro and 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** (2006), 1261–1300.
[31] À. Haro and 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** (2006), 530–579.
[32] À. Haro and R. de la Llave, *A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: Explorations and mechanisms for the breakdown of hyperbolicity*, SIAM Jour. Appl. Dyn. Syst., **6** (2007), 142–207.
[33] À. Jorba and E. Olmedo, *A parallel method to compute quasi-periodic solutions*, in "EQUADIFF 2003," 181–183, World Sci. Publ., Hackensack, NJ, 2005.
[34] À. Jorba and E. Olmedo, *On the computation of reducible invariant tori on a parallel computer*, SIAM J. Appl. Dyn. Syst., **8** (2009), 1382–1404.
[35] D. E. Knuth, "The Art of Computer Programming. Vol. 2: Seminumerical Algorithms," Third revised edition, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1997.
[36] R. de la Llave, A. González, À. Jorba and J. Villanueva, *KAM theory without action-angle variables*, Nonlinearity, **18** (2005), 855–895.
[37] R. de la Llave, *A tutorial on KAM theory*, in "Smooth Ergodic Theory and its Applications" (Seattle, WA, 1999), Proc. Sympos. Pure Math., **69**, Amer. Math. Soc., Providence, RI, (2001), 175–292.
[38] R. de la Llave and C. E. Wayne, *Whiskered and low dimensional tori in nearly integrable Hamiltonian systems*, Math. Phys. Electron. J., **10** (2004), Paper 5, 45 pp. (electronic).
[39] R. S. McKay, "Renormalisation in Area Preserving Maps," Ph.D thesis, Princeton University, 1982.
[40] A. Olvera and N. P. Petrov, *Regularity properties of critical invariant circles of twist maps and their universality*, SIAM J. Appl. Dyn. Syst., **7** (2008), 962–987.
[41] M. J. Raković and Shih-I Chu, *New integrable systems: Hydrogen atom in external fields*, Phys. D, **81** (1995), 271–279.

[42] M. J. Raković and Shih-I Chu, *Phase-space structure of a new integrable system related to hydrogen atoms in external fields*, J. Phys. A, **30** (1997), 733–753.
[43] 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" (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974), Lecture Notes in Phys., Vol. 38, Berlin, Springer, (1975), 598–624.
[44] C. Simó, private communication, 2000.
[45] E. Zehnder, *Generalized implicit function theorems with applications to some small divisor problems. II*, Comm. Pure Appl. Math., **29** (1976), 49–111.

Received September 2010; revised August 2011.

E-mail address: gh707@nyu.edu
E-mail address: llave@math.utexas.edu
E-mail address: sire@cmi.univ-mrs.fr