EXISTENCE OF WHISKERED KAM TORI OF CONFORMALLY SYMPLECTIC SYSTEMS

RENATO C. CALLEJA, ALESSANDRA CELLETTI, AND RAFAEL DE LA LLAVE

2010 *Mathematics Subject Classification.* 70K43, 70K20, 34D35.

Key words and phrases. Whiskered tori, Conformally symplectic systems, KAM theory.

R.C. was partially supported by UNAM-PAPIIT IA 102818. A.C. was partially supported by GNFM-INdAM and acknowledges MIUR Excellence Department Project awarded to the Department of Mathematics of the University of Rome "Tor Vergata" (CUP E83C18000100006). R.L. was partially supported by NSF grant DMS-1800241.

This material is based upon work supported by the National Science Foundation under Grant No. DMS-1440140 while the authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Fall 2018 semester.

Abstract. Many physical problems are described by conformally symplectic systems (i.e., systems whose evolution in time transforms a symplectic form into a multiple of itself). We study the existence of whiskered tori in a family f_μ of conformally symplectic maps depending on parameters μ (often called *drifts*). We recall that whiskered tori are tori on which the motion is a rotation, but they have as many expanding/contracting directions as allowed by the preservation of the geometric structure.

Our main result is formulated in an *a-posteriori* format. We fix ω satisfying Diophantine conditions. We assume that we are given 1) a value of the parameter μ_0, 2) an embedding of the torus K_0 into the phase space, approximately invariant under f_{μ_0} in the sense that $f_{\mu_0} \circ K_0 - K_0 \circ T_\omega$ (where T_ω is the shift by ω) is small (in some norm), 3) a splitting of the tangent space at the range of K_0, into three bundles which are approximately invariant under Df_{μ_0} and such that the derivative satisfies *"rate conditions"* on each of the components.

Then, if some non-degeneracy conditions (verifiable by a finite calculation on the approximate solution and which do not require any global property of the map) are satisfied, we show that there is another parameter μ_∞, an embedding K_∞ and splittings close to the original ones which are invariant under f_{μ_∞}. We also bound $|\mu_\infty - \mu_0|$, $\|K_\infty - K_0\|$ and the distance of the initial and final splittings in terms of the initial error.

We allow that the stable/unstable bundles are nontrivial (i.e., not homeomorphic to a product bundle). On the other hand, we show that the geometric set up has the global consequence that the center bundle is necessarily trivial (i.e., homeomorphic to a product bundle).

The proof of the main theorem consists in describing an iterative process that takes advantage of cancellations coming from the geometry. Then, we show that the process converges to a true solution when started from an approximate enough solution. The iterative process leads to an efficient algorithm that is quite practical to implement.

The a-posteriori format of the theorem implies the usual formulation of persistence under perturbations (the solutions for the original systems are approximate solutions for the perturbation), but it also allows to justify approximate solutions produced by any method (for example numerical solutions or asymptotic formal expansions). As an application, we study the singular problem of effects of small dissipation on whiskered tori. We develop formal (presumably not convergent) expansions in the perturbative parameter (which generates dissipation) and use them as input for the a-posteriori theorem. This allows to obtain lower bounds for the domain of analyticity of the tori as function of the perturbative parameter.

Even if we state only the theory for maps, our results apply also to flows.

1. Introduction

Several physical problems are modeled by Hamiltonian systems affected by a dissipation which enjoys a remarkably geometric property, namely that the symplectic structure (preserved by the Hamiltonian evolution) is transformed into a multiple of itself. Such systems are called *conformally symplectic systems*.

Examples of physical problems that are described by conformally symplectic systems are:

- Hamiltonian systems with a dissipative effect proportional to the velocity; a concrete example is given by the spin–orbit problem in Celestial Mechanics with a tidal torque – see [Cel10, CL04, CL09];
- Euler-Lagrange equations of exponentially discounted systems; these models are often found in finance, when inflation is present and one needs to minimize the cost in present money – see [Ben88, LR16, ISM11, DFIZ16a, DFIZ16b]. The exponential discount is also common in control theory *finite horizon models* - see [MHER95];
- Gaussian thermostats, which are used in computations of non-equilibrium molecular dynamics - see [DM96b, WL98];
- Nosé-Hoover dynamics in Statistical Mechanics - see [DM96a, Hoo91][1]

Besides the interest in applications, (locally) conformally symplectic systems were also studied as natural problems by differential geometers (see [Ban02, Agr10, Vai85]). For a detailed comparison between the conformally symplectic systems and the slightly more general locally conformally symplectic systems see [CCdlL19a].

Remark 1. *Conformally symplectic systems are a very special kind of dissipative systems (for example, mechanical systems with a friction proportional to a power different from* 1 *of the velocity are not conformally symplectic). Therefore, conformally symplectic systems enjoy remarkable properties which are not present in more general dissipative systems.*

For example, in [CCdlL13] *it is shown that for conformally symplectic systems in a neighborhood of a Lagrangian invariant torus, the motion is smoothly conjugate to a rotation along the torus and a constant contraction in the normal direction. This is clearly false in general dissipative systems. In* [CCFdlL14] *it is shown how this rigid behaviour of conformally symplectic systems leads to quantitative properties of phase locking regions (*[CCFdlL14]*). The existence of a time-dependent variational principle also leads to very dramatic qualitative properties not shared by arbitrary dissipative systems (*[MS17a]*). The conformal symplectic properties of a system also affect the dimension of the parameters one needs to adjust to get a quasiperiodic solution (*[CCdlL13]*).*

For us, the most remarkable property of conformally symplectic systems is the so-called "automatic reducibility" *that shows that the "center" bundle around a rotational torus*

[1]There are several formulations of the Nosé-Hoover dynamics in the literature. Some of them are not conformally symplectic.

allows a canonically defined system of coordinates in which the derivative is particularly simple. We will also show in Section 4 that the conformal symplectic geometry implies that the "center" bundles around a rotational torus are trivial.

The goal of this paper is to study the existence of whiskered tori in conformally symplectic systems. Whiskered tori were introduced in [Arn64, Arn63] in Hamiltonian (a.k.a. symplectic) systems and conjectured to be the key geometric structures leading to instability for nearly integrable systems.

We will present more precise definitions of whiskered tori later (see Definition 19), but we indicate that these are tori in which the motion is conjugate to a rotation and which have many hyperbolic directions (exponentially contracting in the future or in the past under the linearized evolution). Definition 19 includes that the exponential rates characterizing the center bundle straddle the conformally symplectic constant and that the dimension of the hyperbolic directions is as large as possible given that the map is conformally symplectic.

The existence of quasi-periodic motions in dissipative systems is very different from the Hamiltonian case. The dissipation forces many orbits to have the same asymptotic behavior and, hence, the set of asymptotic behaviours is smaller in the dissipative case than in the Hamiltonian one. Therefore, adding a small dissipation to a Hamiltonian system is a very singular perturbation. In the Hamiltonian case, under mild non-degeneracy conditions, to find tori of a certain frequency, it suffices to choose the initial conditions. In the dissipative case, since many solutions have the same asymptotic behavior one cannot choose the initial conditions to obtain the desired long term behavior in a fixed map. One needs to consider families and adjust parameters to obtain a torus with a fixed frequency. The results here apply also to symplectic systems (the case considered in [Arn64, Arn63]), even if the parameter count is different. Indeed, we find a unified formalism in which one can continue from the symplectic to the weakly symplectic.

Our first main result is Theorem 20, which is expressed in the format of a-posteriori theorems of numerical analysis. Given a family f_μ of conformally symplectic mappings, we formulate an equation, called *invariance equation* for the parameterization of a torus, say K, for the parameter μ in the family and for the splittings of the space. The invariance equation expresses that the parameterization and the splittings are invariant by the map given by the parameter.

Theorem 20 allows us to validate approximate solutions of the invariance equation obtained by non-rigorous methods (one just needs to verify the accuracy of the solution and the condition numbers). Notably, one can justify numerical methods or asymptotic expansions. Note that Theorem 20 does not assume that the system is close to integrable and it is, therefore, very suitable to study phenomena that happen near the breakdown of the tori. A result on the local uniqueness of the solution is given in Theorem 34. The a-posteriori theorems can also be used to justify computations in concrete systems of interest. For example in Celestial Mechanics and Astrodynamics, one is interested in finding objects in N-body problems with very concrete values of the parameters.

It is also important to remark that the condition numbers we introduce just involve averages of algebraic expressions formed by the approximate solution and its derivatives, hence they are straightforwardly computable just knowing the approximate solution. They do not involve any global assumption on the map such as the global twist.

Our second main result is Theorem 38. We consider dissipative perturbations of a Hamiltonian system and study the domain of analyticity in terms of the perturbation parameter of the resulting whiskered tori (and the drift and the bundles).

The limit of small dissipation occurs naturally in many problems in Celestial Mechanics and many of the models for dissipation in common use are conformally symplectic ([Cel10], [CL04], [CL09]). Examples of dissipations going to zero might appear when a celestial body subject to tidal torque reaches a synchronous rotation state, or when a satellite is launched from Earth and reaches an altitude where the atmosphere is absent, thus moving to a region where the atmospheric drag dissipation is zero ([CG18], [Cha05]). The limit of small dissipation is also of interest in finance, where it corresponds to small inflation ([Ben88]). In control theory it is common to consider the limit when the finite horizon is taken to infinity ([MHER95]). In the discounted action model, in the limit of small dissipation the minimizing sets of the action and the solutions of the Hamilton-Jacobi equation have been considered in [DFIZ16a, DFIZ16b].

If we denote the perturbative parameter by ε (which we consider complex), we can study the analyticity domains for μ_ε, K_ε (see Theorem 38). In contrast with the Hamiltonian case, in the dissipative case, unless one adjusts parameters, one does not expect that there are quasi-periodic solutions. Hence, we do not expect that perturbative expansions converge and that there is no open neighborhood of $\varepsilon = 0$ contained in the analyticity domain of $\mu_\varepsilon, K_\varepsilon$. We conjecture that the domain we obtain in Theorem 38 (which indeed does not contain any ball centered in the origin) is essentially optimal, see Section 7.4.

1.0.1. *Application to flows.* Finally, we note that all the results we will present for maps apply also to flows in continuous time by taking sections and considering the return map (see the construction in [CCdlL13]). Of course, producing a direct proof for flows by adapting the steps in the proof for maps presented here is straightforward.

1.1. **Comparison with other results in the literature.** The literature on the existence of whiskered tori in symplectic systems is very extensive, see [Gra74, Zeh76, JV97, LY05b, LY05a].

One should also mention that there are persistence theorems for whiskered tori that apply to general systems (not necessarily symplectic). Indeed, the pioneering work [Mos67] already considered general perturbations (see also [BHTB90, BHS96, CH17b]). Of course, the conformally symplectic systems studied in this paper are a particular case of general perturbations and the results of [Mos67, BHTB90, BHS96, CH17b] apply a-fortiori. The extra assumption of conformally symplectic allows us to obtain some results that are false for more general perturbations. Notably, the results here for weakly dissipative systems could be false if the dissipation is not conformally symplectic (see Remark 1). We also note that, for conformally symplectic systems there are constraints on the possible exponents on the normal direction. We discuss these constraints in Section 4. We also note that the dynamical and geometric assumptions have the surprising global consequence that the center bundle is trivial in the sense of bundle theory (diffeomorphic to a product bundle).

From the point of view of techniques, we note that the papers [Mos67, BHTB90, BHS96] are based on transformation theory (i.e., making changes of variables till the system is in a form which manifestly does have an invariant torus), they also assume that the stable and unstable bundles are trivial in the sense of bundle theory (the bundle is a product bundle). This triviality of the bundle allows us to use a global system of coordinates in the center manifold in which the linearized equation can be reduced to constant coefficients.

This paper is based on making additive corrections to an embedding and by solving the linearized equations using geometrically natural operations.

From the point of view of computation, transformation theory is hard to implement as algorithms, since it involves operating on functions with many variables (as many variables as the dimension of the phase space). The algorithms in this paper only require to deal with functions with as many variables as the dimension of the objects considered.

On the other hand, the transformation theory gives not only the existence of the whiskered tori, but also substantial information on the behavior of orbits in a neighborhood of the torus [SL12]. A comparison of the transformation theory and methods similar to ours for Lagrangian tori in conformally symplectic systems appears in [LS15].

An approach for the study of whiskered tori close to ours was described in [FdlLS09b]. In particular, it produces an a-posteriori method, which was implemented in [FdlLS09a] for finite dimensional Hamiltonian systems. Algorithms based on the method of [FdlLS09a] and a more efficient reformulation appear in [HdlLS11]. Generalizations to Hamiltonian lattice systems appear in [FdlLS15] and in [dlLS18] for PDE's, while presymplectic systems are investigated in [dlLX19]. The paper [CH17b] uses similar techniques to give a result on the persistence of whiskered tori for general systems, while [CH17a] presents implementations of the algorithms and studies of the phenomena that happen at breakdown.

1.2. **The a-posteriori format.** The method in this paper is not based on transformation theory and on the reduction to normal forms, but it is rather based on formulating an equation for the parameterization which expresses the invariance and which is solved by a Newton-like method.

The solution of the Newton equations uses the contraction on the stable and unstable bundles as well as geometric identities obtained from the fact that the map is conformally symplectic. These identities have the geometric meaning that there is a special system of coordinates in which the Newton equations are particularly simple. They were used in [CCdlL13] for Lagrangian tori of symplectic systems and in [CC10] for conformally symplectic systems. In the symplectic case they were introduced in [dlL01, dlLGJV05] for Lagrangian tori. In the case of whiskered tori for symplectic systems, they were used in [FdlLS09b, FdlLS09a] (a more efficient variation was introduced in [HdlLS11, FdlLS15]).

The method in this paper allows that the hyperbolic bundles are non-trivial (a situation that happens near resonances, [HdlL07]). Some other examples were presented in [FdlLS09b]. Most methods based on normal forms assume that the bundles are trivial.

In contrast, in this paper, we will show that for conformally symplectic systems (including symplectic systems) the center bundles of whiskered tori are trivial (see Lemma 25): this is a rather surprising interaction between global geometry and dynamics.

The main result of this paper is stated in an a-posteriori format: assuming the existence of sufficiently approximate solutions with respect to *condition numbers*, then we conclude that there are true solutions.

Notice that the a-posteriori format implies the usual results of persistence under a small change of the system. If there is a system which has an invariant whiskered torus, the torus and its bundles will be approximately invariant for all the approximate systems. Then, applying the a-posteriori theorem we will conclude that the perturbed system also has an invariant torus. As an immediate consequence of the a-posteriori format, we automatically obtain Lipschitz dependence on the map or on the frequency (just observe that the solution corresponding to a frequency ω satisfies the equation for a frequency ω' up to an error bounded by $C|\omega - \omega'|$). Validating the Lindstedt series, we obtain sharper differentiability properties. As mentioned in [CCdlL13], the a-posteriori format has several consequences: smooth dependence on parameters, Whitney regularity in the frequency, bootstrap of regularity, etc.

The a-posteriori format is very well suited for numerical analysis, since it allows one to validate the approximate solutions produced by a numerical calculation. We also note that the method of proof leads to a very efficient algorithm. The iterative method is quadratically convergent, it only requires to discretize functions with a number of variables equal to the dimension of the torus searched, rather than the dimension of the phase space, the storage space requirements are small (order $O(N)$, where N is the number of discretization modes) and the operation count per step is small ($O(N \ln(N))$). A continuation method based on the algorithm presented here is guaranteed to converge till the boundary of existence of the torus or till some of the non-degeneracy conditions fail. Hence, a careful continuation algorithm gives a practical numerical algorithm to detect the breakdown of analyticity (these algorithms were implemented for Lagrangian rotational tori and models in statistical mechanics in [CdlL10], [CC10]). For conformally symplectic systems, the paper [LS15] includes a comparison of the method presented here with the transformation method from the point of view of applications.

1.3. **Main results of the paper.** The proof of our main result on the existence of the whiskered tori, Theorem 20, consists in describing an iterative Newton method to solve the invariance equation when started on an approximate solution. The Newton step takes advantage of the geometry of the system and of the existence of hyperbolic directions.

A step of the Newton method involves solving a linearized invariance equation. To solve this equation, the linearized invariance equation is projected on the hyperbolic and center subspaces. The equations projected on the hyperbolic equations can be solved taking advantage of the contractions (in the future or in the past). As for the equations projected to the center subspace, we use the so–called *automatic reducibility* which, near an approximate invariant torus, constructs a system of coordinates in which the linearized equation along the center directions takes a particularly simple form, which allows the use of Fourier methods.

Our second result, Theorem 38, is concerned with the limit of small dissipation. The low dissipation limit is very natural in Celestial Mechanics and Astrodynamics since many celestial bodies experience weak frictions, due to tidal forces or the atmospheric drag ([CC08], [Cha05]). In this case, the friction is indeed proportional to the velocity, which makes the system conformally symplectic.

We consider systems that depend on a small parameter[2] ε and the conformal factor is $\lambda(\varepsilon) = 1 + \alpha\varepsilon^a + O(|\varepsilon|^{a+1})$ for $a \in \mathbb{Z}_+$, $\alpha \in \mathbb{C} \setminus \{0\}$. When $\varepsilon = 0$ we recover the symplectic case and small ε means small dissipation. We study the analyticity domains in ε near $\varepsilon = 0$ of the whiskered tori (as well as the domains of the drift and of the stable/unstable bundles). We note that the domains we obtain do not contain any ball centered at $\varepsilon = 0$, and we conjecture that this is optimal, hence we conjecture that indeed the formal power series are divergent. For full dimensional tori, this has been studied numerically by [BC19].

Inspired by [CCdlL17], we start by constructing Lindstedt series for the problem. Even if the Lindstedt series, in this case, probably do not converge, a finite order truncation provides an approximate solution, which can be used as the approximate solution taken as input in the a-posteriori theorem, Theorem 20, for some (complex) values of ε for which we can verify the quantitative conditions of Theorem 20.

In this way, we obtain that the parameterization and the drift are analytic when ε ranges in a domain which we describe very explicitly. The domain of ε where we show that $K_\varepsilon, \mu_\varepsilon$ are analytic is obtained by removing from a ball centered at the origin a sequence of smaller balls with centers on a curve. The radii of the balls decrease exponentially fast with the distance of the centers to the origin.

It is interesting to remark that the Lindstedt expansions produce approximate solutions for all sufficiently small values of ε. What determines the shape of the analyticity domain is the values of ε for which we can verify the non-degeneracy conditions (notably the

[2]Since we are considering analyticity properties, it is natural to take ε complex and all the objects considered to be complex as well. This does not make any difference in the proof, since the Newton step involves just algebraic manipulations, derivatives and solutions of cohomology equations, which are the same for complex maps. Of course, if $f_{\mu,\varepsilon}$ is such that for real values of the arguments it gives real values, the parameterizations K_ε and μ_ε will also be real for real values of the arguments.

Diophantine conditions) of Theorem 20. This emphasizes the importance of the non-degeneracy conditions rather than just the smallness of the error.

The study of Lindstedt series followed by a validation step for low–dimensional tori in Hamiltonian systems was developed in [JdlLZ99]. The paper [JdlLZ99] considers also the case of weak hyperbolicity: systems which are perturbations of an integrable one, which does not have any hyperbolicity, so that the hyperbolicity is generated by the perturbation. We also mention [Mas05], where Lindstedt series expansions for whiskered tori have been constructed in a problem of Celestial Mechanics. More recently, [BC19] has computed Lindstedt series for Lagrangian tori for the dissipative case and determined the domain of analyticity of the tori, as well as several other properties of the series (monodromy, Gevrey properties). The numerical results of [BC19] present several interesting conjectures for the singular problem.

1.3.1. *Organization of the paper.* This paper is organized as follows. In Section 2 we provide basic notions, such as conformally symplectic systems, Diophantine frequency vectors, function spaces, cocycles, invariant bundles, dichotomies. In Section 3, we state the first main result, Theorem 20. In Section 4 we present some interactions of the hyperbolicity assumptions and the geometry. In particular, in Section 4.7 we present the automatic reducibility, a key ingredient of the proof of Theorem 20, and in Section 4.6 we prove that the center bundle has to be trivial. In Section 5 we provide the proof of Theorem 20, while the proof of the local uniqueness of the solution is given in Section 6. The study of the analyticity domains in the symplectic limit is presented in Section 7. We also call attention to [CCdlL19b], which can serve as a reading guide for this paper.

2. Some preliminary definitions and standard results

In this Section, we collect some definitions and some elementary lemmas that will be used in the formulation (and proof) of the results. Of course, this Section may be considered mainly as reference and could be skipped in a first reading.

2.1. **Conformally symplectic systems.** We start by introducing the definition of conformally symplectic mappings and flows (see, e.g., [Ban02, CCdlL13, DM96b, WL98]).

Let $\mathcal{M} = \mathbb{T}^n \times B$ be a symplectic manifold of dimension $2n$ with $B \subseteq \mathbb{R}^n$ an open, simply connected domain with smooth boundary. We assume that $\mathcal{M}$ is endowed with the standard scalar product and a symplectic form Ω. We do not assume that Ω has the standard form. In the study of the small dissipation limit we will assume that Ω is exact, but Theorem 20 does not need the assumption of exactness.

Definition 2. *We say that $f : \mathcal{M} \to \mathcal{M}$ is a conformally symplectic diffeomorphism, when there exists a constant λ such that*

$$f^*\Omega = \lambda\,\Omega\ . \tag{2.1}$$

Remark 3. *When $n = 1$, any orientable manifold is symplectic for any non-degenerate 2-form. If we allow that λ is a function, any diffeomorphism is conformally symplectic.*

When $n \geq 2$, any function λ satisfying (2.1) *has to be a constant for a connected manifold $\mathcal{M}$ (see* [Ban02]*). It suffices to observe that*

$$0 = f^*d\Omega = d(f^*\Omega) = d(\lambda\Omega) = d\lambda \wedge \Omega\ .$$

It is shown in [Ban02] *that, when the dimension of the space is* 4 *or higher, the above implies that $d\lambda = 0$ (a simple argument is just to use locally the Darboux form of Ω).*

Throughout this paper we always consider λ constant since the whiskered tori we are concerned with only appear when $n \geq 2$.

Remark 4. *The constant λ will be real when we consider real maps. It will be a complex constant in Section 7 devoted to the analyticity properties in ε, where it is natural to consider complex maps. In the physical applications, the complex maps are such that they take real values for real arguments.*

2.2. **Expressions in coordinates.** In some computations later, we will find it convenient to use matrix notation for the computations (recall that we are assuming that the phase space we are considering is Euclidean).

We will consider the tangent of the phase space endowed with the standard inner product, which does not depend on the base point (later in Section 4.6, we will find it useful to use metrics depending on the point).

Given a linear operator $A : \mathcal{T}_x\mathcal{M} \to \mathcal{T}_y\mathcal{M}$ (where $\mathcal{T}_x\mathcal{M}$ denotes the tangent space of $\mathcal{M}$ at x), we denote the adjoint A^T as the linear operator from $\mathcal{T}_y\mathcal{M}$ to $\mathcal{T}_x\mathcal{M}$ that satisfies

$$< u, A^T v >=< Au, v > \quad \forall u \in \mathcal{T}_x\mathcal{M}, v \in \mathcal{T}_y\mathcal{M}\ .$$

Once we fix the inner product, we can identify the symplectic form Ω with an operator:

$$\Omega_x(u, v) = \langle u, J_x v\rangle \quad \forall u, v \in \mathcal{T}_x\mathcal{M}\ . \tag{2.2}$$

The asymmetry of the symplectic form means that

$$J_x^T = -J_x \ .$$

Using the identification (2.2) between the symplectic form and the operator J_x, we see that a map f is conformally symplectic if and only if

$$Df^T(x) J_{f(x)} Df(x) = \lambda J_x \ . \tag{2.3}$$

In this paper we will not assume that J_x is constant or that it has the standard form $J_x = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}$. Non-constant symplectic forms appear naturally in the study of neighborhoods of fixed points, in the study of PDE's, etc.

Remark 5. *The relation between the operator J, the symplectic form and the metric is a very useful tool in modern symplectic geometry. We mention the books (*[Ber01, CdS01, MS17b]*). A much deeper study of the applications of J as a complex structure is in* [MS12].

In this paper, we will indeed use the relation between the operator J and the metric to obtain a result on the global structure of the center bundle, namely, that it is diffeomorphic to a product bundle, see Lemma 25.

2.2.1. *Exactness.* We say that the symplectic form is exact when there exists α such that

$$\Omega = d\alpha \ .$$

A map f is exact when there exists a single-valued function G, such that

$$f^*\alpha - \lambda\alpha = dG \ .$$

Note that the fact that the mapping is conformally symplectic can be written as $d(f^*\alpha) = f^*\Omega = \lambda d\alpha$, which gives $d(f^*\alpha - \lambda\alpha) = 0$. Hence, exact maps are conformally symplectic.

When $\lambda \neq 1$, the paper [CCdlL13] does not need exactness of the map to produce invariant tori. Nevertheless, in the symplectic case ($\lambda = 1$) it is well known that the exactness is a necessary condition to have Lagrangian invariant tori which are homotopically non-trivial (the case of maximal homotopically trivial tori is discussed in [FdlL15]).

Lower dimensional tori can exist in non-exact symplectic systems, but if the tori have some non-trivial homotopy, one needs that some cohomology of $f^*\alpha - \alpha$ vanishes. The considerations of exactness come into play only when we consider the symplectic limit in Section 7.

For simplicity, in this paper, we will assume that f is exact, even if parts of this assumption can be weakened to the vanishing of some cohomology class of $f^*\alpha - \alpha$ depending on the topology of the embedding of the torus.

2.3. **Diophantine properties.** We will assume that frequency vectors of the whiskered tori satisfy the following Diophantine inequality.

Definition 6. *Let $\omega \in \mathbb{R}^d$, $d \leq n$, $\tau \in \mathbb{R}_+$. Let the quantity $\nu(\omega;\tau)$ be defined as*

$$\nu(\omega;\tau) \equiv \sup_{k\in\mathbb{Z}^d\setminus\{0\}} \left(|e^{2\pi i k\cdot\omega} - 1|^{-1}\ |k|^{-\tau} \right) , \tag{2.4}$$

where $\cdot$ denotes the scalar product and $|k| \equiv |k_1| + ... + |k_d|$ with k_1, ..., k_d the coordinates of k. We say that ω is Diophantine of class τ and constant $\nu(\omega;\tau)$, if

$$\nu(\omega;\tau) < \infty \ .$$

We denote by $\mathcal{D}_d(\nu,\tau)$ the set of Diophantine vectors of class τ and constant ν.

For $\lambda \in \mathbb{C}$, we define the quantity $\nu(\lambda;\omega,\tau)$ as

$$\nu(\lambda;\omega,\tau) \equiv \sup_{k\in\mathbb{Z}^d\setminus\{0\}} \left(|e^{2\pi i k\cdot\omega} - \lambda|^{-1}\ |k|^{-\tau} \right) . \tag{2.5}$$

We say that λ is Diophantine with respect to ω of class τ and constant $\nu(\lambda;\omega,\tau)$ if

$$\nu(\lambda;\omega,\tau) < \infty \ .$$

Note that the quantity (2.5) makes sense for any complex number λ. In Theorem 20 we will need only to consider $\lambda \in \mathbb{R}$, but in Theorem 38, we will use complex λ. Note that when $|\lambda| \neq 1$, we have that $\nu(\lambda;\omega,\tau) < \infty$. But, as λ approaches the unit circle, depending on the limit on the unit circle, the ν could either become unbounded or remain bounded.

Note that the definition of λ Diophantine with respect to ω makes sense even if ω itself is not Diophantine (in particular if $|\lambda| \neq 1$, λ is Diophantine with respect to ω for all ω, even rational ones). Of course, the Diophantine exponents of ω and of λ with respect to ω are independent.

In our applications, however, it is natural to assume both that ω is Diophantine and that λ is Diophantine with respect to ω. For simplicity of the notation we will only use a common exponent that works both for the Diophantine exponent of ω and for the exponent of λ with respect to ω.

Remark 7. *In the study of analyticity domains in Section 7, we will consider ω fixed, but λ will change. Hence, we prefer to think of $\nu(\lambda; \omega, \tau)$ mainly as a function of λ and think of ω as a parameter that remains fixed.*

2.4. **Invariant rotational tori.** Let $\Upsilon \subset \mathcal{M}$ be diffeomorphic to a torus. We say that Υ is a rotational torus for a map f, when $f(\Upsilon) = \Upsilon$ and the dynamics of f restricted to Υ is conjugate to a rotation. Precisely, we start with the following definition of rotational invariant tori, non necessarily of maximal dimension.

Definition 8. *Let f be a differentiable diffeomorphism of $\mathcal{M}$. For $0 < d \leq n$, let $K : \mathbb{T}^d \to \mathcal{M}$ be a differentiable embedding. Let $\omega \in \mathbb{R}^d$, denote by $T_\omega : \mathbb{T}^d \to \mathbb{T}^d$ the rotation of vector ω, namely $T_\omega(\theta) = \theta + \omega$.*

We say that K parameterizes a rotational invariant torus, if the following invariance equation holds:

$$f \circ K = K \circ T_\omega \ . \tag{2.6}$$

The relation (2.6) appears frequently in ergodic theory and it is described as *"The rotation T_ω is a factor of f"* or *"The rotation T_ω is semiconjugate to f"*.

As it is well known ([Mos67]), to find an invariant torus with prescribed frequency of a conformally symplectic system, we will need to introduce a *drift* parameter ([CCdlL13]) and, precisely, we will consider a family f_μ of conformally symplectic mappings. Then, we will try to find a parameter vector $\mu \in \mathbb{R}^d$ and an embedding of the torus K in such a way that the following invariance equation is satisfied:

$$f_\mu \circ K = K \circ T_\omega \ . \tag{2.7}$$

Note that the equation (2.7) is an equation for both μ and K. The equation (2.7) will be the centerpiece of our analysis. We will develop a quasi-Newton method for it, using the geometric properties of the map to analyze the linearization of (2.7).

2.4.1. *Normalization.* The equation (2.7) is underdetermined. Note that if (μ, K) is a solution of (2.7), then, so is $(\mu, K \circ T_\sigma)$ for any $\sigma \in \mathbb{R}^d$. Note that this lack of uniqueness can be interpreted as choosing the origin in the space of the parametrization.

We will show that, in many cases, the choice of the origin of the parameterization is the only source of local non-uniqueness. By choosing some normalization that fixes the origin, we will show that one obtains local uniqueness (i.e. the normalized parameterizations of tori which are close enough coincide) and, hence, one can discuss smooth dependence on parameters, etc. Of course, global uniqueness is false unless one makes global assumptions.

In Section 6 we give a precise statement of the local uniqueness of the solution of the invariance equation (2.7) under some normalization that fixes the origin of coordinates.

Before starting with the main results let us prove the easy result given by Proposition 9.

If ω is nonresonant (i.e., for any $k \in \mathbb{Z}^d \setminus \{0\}$, we have $|\omega \cdot k| \notin \mathbb{N}$), then $\{T_\omega^n(\theta)\}$ is dense in the torus for any θ and the embedding K is almost uniquely determined as stated in the following Proposition.

Proposition 9. *Let f be a differentiable diffeomorphism of $\mathcal{M}$. Assume that $\omega \in \mathbb{R}^d$ is nonresonant. Let K_1, K_2 be continuous mappings satisfying* (2.6) *and their ranges have a non-empty intersection. Then, there exists $\sigma \in \mathbb{R}^d$ such that*

$$K_1 = K_2 \circ T_\sigma \ .$$

Proof. We can find $\sigma_1, \sigma_2 \in \mathbb{T}^d$ such that $K_2(\sigma_1) = K_1(\sigma_2)$. Applying f to both sides and using (2.6), we obtain $K_2(\sigma_1 + \omega) = K_1(\sigma_2 + \omega)$. Repeating the argument, we obtain that for all $j \in \mathbb{N}$, $K_2(\sigma_1 + j\omega) = K_1(\sigma_2 + j\omega)$, which gives

$$K_2(\theta) = K_1(\theta + \sigma_2 - \sigma_1) \ , \qquad \theta \in \{j\omega + \sigma_1\}_{j \in \mathbb{N}} \ . \tag{2.8}$$

Since $\{\sigma_1 + j\omega\}_{j \in \mathbb{N}}$ is dense in the torus and K_1, K_2 are continuous, we have the equality (2.8) for all θ in the torus. $\square$

2.5. **Definition of function spaces.** To make precise estimates on the quantities involved in the proof, most notably the error associated to an approximate solution of (2.7), we need to fix a function space and a norm, precisely the space and norm of analytic functions as in the definition given below.

Definition 10. *For $\rho > 0$ we denote by $\mathbb{T}^d_\rho$ the set*

$$\mathbb{T}^d_\rho = \{z \in \mathbb{C}^d/\mathbb{Z}^d \ : \ \mathrm{Re}(z_j) \in \mathbb{T} \ , \quad |\,\mathrm{Im}(z_j)| \le \rho \ , \quad j = 1, ..., d\} \ .$$

Given $\rho > 0$ and a Banach space X, we denote by $\mathcal{A}_\rho(X)$ the set of functions from $\mathbb{T}^d_\rho$ to X which are analytic in the interior of $\mathbb{T}^d_\rho$ and that extend continuously to the boundary of $\mathbb{T}^d_\rho$.

We endow $\mathcal{A}_\rho$ with the norm

$$\|f\|_\rho = \sup_{z \in \mathbb{T}^d_\rho} |f(z)| \ ,$$

which makes it into a Banach space.

For later use, we introduce the norm of a vector valued function $f = (f_1, \dots, f_n)$ as $\|f\|_\rho = \sqrt{\|f_1\|_\rho^2 + \ldots + \|f_n\|_\rho^2}$ and the norm of an $n_1 \times n_2$ matrix valued function F as $\|F\|_\rho = \sup_{\chi \in \mathbb{R}_+^{n_2}, |\chi|=1} \sqrt{\sum_{i=1}^{n_1} (\sum_{j=1}^{n_2} \|F_{ij}\|_\rho \, \chi_j)^2}$. In this way, we have the customary inequalities for the norm of the product of a matrix and a vector.

2.6. **Cocycles and invariant bundles.** In this Section, we recall the standard definitions on growth properties of the products (2.9) below. These properties are quite standard ([SS74, Cop78, MSK03]). The hyperbolicity of cocycles – over more general systems than rotation – is treated specially in [CL95]. For a pedagogical treatment, see also the expository chapters of [DV17].

Later on, in Section 4, we will see that there is a deep relation between the growth properties of the cocycle and the conformal symplectic properties of the map. These properties are somewhat surprising since they show an interplay between the dynamics and the geometric properties of the tori. Quite notably, we will establish that the center bundle of a rotational invariant torus is trivial in the sense of bundle theory. That is, it can be written as a product bundle.

For the sake of notation, we will assume that $\mathcal{M}$, the phase space, is an Euclidean manifold. In that way, we can identify the different tangent spaces.

Since we are identifying the spaces, we can think of all the factors $Df_\mu \circ K \circ T_{j\omega}$ as analytic functions from $\mathbb{T}^d_\rho$ to $n \times n$ matrices.

We will see that in our study of corrections to the invariance equation, to reduce the error we are led to the study of products of the form

$$\Gamma^j \equiv Df_\mu \circ K \circ T_{(j-1)\omega} \times Df_\mu \circ K \circ T_{(j-2)\omega} \times \cdots \times Df_\mu \circ K \ . \tag{2.9}$$

This is a particular case of products of the form

$$\Gamma^j = \gamma \circ T_{(j-1)\omega} \times \gamma \circ T_{(j-2)\omega} \times \cdots \times \gamma \ , \tag{2.10}$$

when we take:

$$\gamma(\theta) = Df_\mu \circ K(\theta) \ . \tag{2.11}$$

Note that if the torus was invariant, then (2.9) would become

$$\Gamma^j = Df_\mu^j \circ K \ ,$$

which makes it clear that the cocycle Γ^j has a dynamical interpretation for invariant tori.

An important property of products of the form (2.10) is:

$$\Gamma^{j+m} = \Gamma^j \circ T_{m\omega} \, \Gamma^m. \tag{2.12}$$

Products of the form (2.10) have been studied extensively in the mathematical literature under the name of *"quasiperiodic cocycles"* or *"cocycles over a rotation"*. The property (2.12) is called the *cocycle* property. In this Section, we will collect some properties of the cocycles, specially the properties of asymptotic growth and exponential trichotomies ([SS74, Cop78]).

2.6.1. *Exponential Trichotomies.* The asymptotic growths of the products (2.10) are important for the study of the invariance equation.

Definition 11. *We say that the cocycle* (2.9) *admits an exponential trichotomy when we can find a decomposition*

$$\mathbb{R}^n = E^s_\theta \oplus E^c_\theta \oplus E^u_\theta \ , \qquad \theta \in \mathbb{T}^d \tag{2.13}$$

and rates of decay

$$\begin{gathered} \lambda_- < \lambda_c^- \le \lambda_c^+ < \lambda_+ \ , \\ \lambda_- < 1 < \lambda_+ \end{gathered}$$

and a constant $C_0 > 0$ that characterize the decomposition:

$$\begin{aligned} v \in E^s_\theta &\iff |\Gamma^j(\theta)v| \le C_0 \lambda_-^j |v|, \quad j \ge 0 \\ v \in E^u_\theta &\iff |\Gamma^j(\theta)v| \le C_0 \lambda_+^j |v|, \quad j \le 0 \\ v \in E^c_\theta &\iff \begin{array}{l} |\Gamma^j(\theta)v| \le C_0 (\lambda_c^-)^j |v|, \quad j \ge 0 \\ |\Gamma^j(\theta)v| \le C_0 (\lambda_c^+)^j |v|, \quad j \le 0 \ . \end{array} \end{aligned} \tag{2.14}$$

We will refer to C_0 and the rates $\lambda_-, \lambda_c^-, \lambda_c^+, \lambda_+$ as the constants characterizing the splitting. We will not introduce the dependence of these quantities on the splitting in the notation.

Even if we assume that the space is Euclidean, we allow that the invariant sub-bundles are not trivial in the sense of bundle theory. That is, they are not isomorphic to a product bundle.

It will be convenient for notation to consider trichotomies as just a pair of dichotomies

$$\begin{aligned} \mathbb{R}^n &= E^s \oplus E^{\hat{s}} \\ \mathbb{R}^n &= E^{\hat{u}} \oplus E^u \ , \end{aligned} \tag{2.15}$$

where we denote by $E^{\hat{s}} = E^c \oplus E^u$ ($\hat{s}$ stands for the symbols that are not s). We use similar notations for other symbols.

The dichotomies in (2.15) are also characterized by rates of growth. If we have both dichotomies, we can reconstruct the spaces in the trichotomy taking $E^c_\theta = E^{\hat{s}}_\theta \cap E^{\hat{u}}_\theta$.

2.6.2. *General properties of splittings and their distances.* Given a splitting of the Euclidean space as in (2.13), we denote by $\Pi^s_\theta, \Pi^c_\theta, \Pi^u_\theta$ the projections corresponding to the spaces in the splitting. Note that the projections $\Pi^{s/c/u}_\theta$ depend on the whole splitting, but we do not include this in the notation.

We can think of Π^σ_θ both as mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ or as mappings from $\mathbb{R}^n$ to E^σ_θ. The advantage of the former is that we can compose mappings corresponding to different splittings. We will not always emphasize the interpretation that is taken.

The space of possible splittings is related to the Grassmanians ([MS74]). In this paper, in an iterative step, we will need to refine the splittings. Hence, it will be important for us to give an efficient description of the splittings close to another one and also give precise definitions of how close splittings are and define the convergence. Of course, for our applications, some smoothness considerations will be needed and the norms involved in the measure of distances will be smooth norms.

2.6.3. *Explicit description of a neighborghood of splittings.* Given a splitting E and another splitting $\widetilde{E}$ close to it, we can write uniquely each of the spaces in $\widetilde{E}$ as the graph of a function from the corresponding space in E to the complementary spaces in E. That is, we can find linear functions $A^\sigma_\theta : E^\sigma_\theta \to E^{\hat{\sigma}}_\theta$ (recall that $E^{\hat{\sigma}}$ denotes the sum of the spaces in the splitting which are not indexed by σ) in such a way that

$$\widetilde{E}^\sigma_\theta = \{v \in \mathbb{R}^n, v = x + A^\sigma_\theta x \,|\, x \in E^\sigma_\theta\} \, . \tag{2.16}$$

In the sequel we fix once and for all an analytic splitting that we will call the *"reference splitting"*, and we will describe all the splittings we use in terms of the A^σ as in (2.16).

In our application, we will see that we can take as the reference splitting the initial approximate splitting in the hypothesis of Theorem 20. Under the assumption that the initial error is small enough, we will see that all the splittings that appear during the iterative procedure are inside the neighborhood of maps that can be described as in (2.16).

It is useful to think of (2.16) as providing a system of coordinates of a neighborhood of the reference splitting in the (highly nonlinear) space of all splittings.

2.6.4. *Several notions of distance among bundles and splittings.* In Grassmanian geometry ([MS74]) it is customary to define the distance between subbundles of an ambient bundle, by fixing a metric in the ambient bundle (in our case, the ambient bundle will be the tangent space to the phase space) and introducing the orthogonal projection $P^{\perp}_{E_\theta}$, where E_θ is the fiber of a bundle E over θ.

For $\rho > 0$, the distance between two subbundles $E, \widetilde{E}$ of the ambient bundle is defined as:

$$dist_\rho(E, \widetilde{E}) = \|P^{\perp}_{E_\theta} - P^{\perp}_{\widetilde{E}_\theta}\|_\rho \ . \tag{2.17}$$

Note that the intrinsic distance between two bundles depends just on the bundles themselves and they do not depend on whether they are part of a splitting.

The distance between two splittings of the tangent space into bundles can be measured by the maximum of the distances between the corresponding bundles that give the splitting. Later we will discuss other notions of distance among splittings.

In our case, we will consider bundles based on a parameterization as $\|\cdot\|_\rho$, analytic norms in a complex extension of the torus as in Definition 10. This will induce an analytic distance between bundles based on a parameterization.

Other equivalent ways of measuring the distance among splittings

It is clear that (2.17) satisfies the triangle inequality and it is indeed a distance. As we will see later, for our application there will be other quantities (not distances) which bound (2.17) from above and below, but which are easier to compute. We will also need to study upper bounds for the distance between splittings and how do they change when the problem changes. We will try to work with the projections corresponding to the splitting Π^σ and not with the orthogonal projections.

If we have two splittings E^σ and $\widetilde{E}^\sigma$ with projections Π^σ and $\widetilde{\Pi}^\sigma$, we can measure the distance between the splittings by:

$$\left(\max_{\substack{\sigma,\sigma' \in \{s,c,u\} \\ \sigma \neq \sigma'}} \max(\|\widetilde{\Pi}^\sigma \Pi^{\sigma'}\|, \|\Pi^\sigma \widetilde{\Pi}^{\sigma'}\|) \right)^{-1} ,$$

where $\|\cdot\|$ can stand for any norm. As before, we will take analytic norms in a neighborhood of the torus.

We can also use as a measure of the distance between the reference splitting and a splitting given by (2.16), the quantity

$$\max_\sigma \|A^\sigma_\theta\| \ , \tag{2.18}$$

where again $\|\cdot\|$ stands for a smooth norm. Note that we are thinking of the linear maps A^σ_θ as mappings from E^σ_θ to $\mathbb{R}^n$.

More generally, if we fix a reference splitting, given two splittings E_1^σ, E_2^σ that can be parameterized by giving the mappings $A^\sigma_{\theta,1}$, $A^\sigma_{\theta,2}$ as above, the quantity

$$\max_\sigma \|A^\sigma_{\theta,1} - A^\sigma_{\theta,2}\| \tag{2.19}$$

gives a measure of the distance of the splittings.

Note that (2.18) and (2.19) are not distances. Nevertheless, they can be used in place of a distance in the sense that, in a small neighborhood in the space of splittings, there are constants that bound one in terms of any of the others. The reason for this equivalence is that there are algebraic expressions giving each of $P^\perp_{E^\sigma}$, Π^σ, A^σ_θ in terms of the others, see [DV17].

2.6.5. *Standard properties of splittings satisfying* (2.14). One of the consequences of (2.14) (see [Cop78]) is that the splittings depend continuously on θ (actually in Hölder fashion) and, hence that the projections Π^σ, $\sigma = s, u, c$, are uniformly bounded, see [SS74]. Using the fact that the dynamics on the base is a rotation, we will later bootstrap the regularity of the splittings to analytic (see [HdlL19]). The Hölder continuity remains valid (and is optimal) if the dynamics on the base is more complicated than a rotation.

Another (slightly non-trivial[3], see [SS74]) consequence of (2.14) is that the bundles characterized by (2.14) are invariant in the sense that:

$$\gamma(\theta) E^\sigma_\theta = E^\sigma_{\theta+\omega} \ .$$

2.6.6. *Approximately invariant splittings.* Given a splitting $E^s_\theta \oplus E^c_\theta \oplus E^u_\theta$ and a cocycle, $\gamma(\theta)$, we define

$$\gamma_\theta^{\sigma,\sigma'} = \Pi^\sigma_{\theta+\omega} \gamma(\theta) \Pi^{\sigma'}_\theta \ . \tag{2.20}$$

The splitting is invariant under the cocycle if and only if

$$\gamma_\theta^{\sigma,\sigma'} \equiv 0 \ , \qquad \sigma \neq \sigma' \ .$$

Again, we note that we can think of the $\gamma^{\sigma,\sigma'}$ either as linear maps in $\mathbb{R}^n$ or, more geometrically, as maps from E^σ to $E^{\sigma'}$. Thinking of them as maps in $\mathbb{R}^n$ allows us to compare maps in different spaces and will be useful in perturbative calculations.

[3]The result would be trivial if the characterization was valid for any C_0 rather than a different number. The key of the proof is to show that one can redefine the norms in such a way that $C_0 = 1$ and that the norm of the operators is also small.

Hence, it is natural to measure the lack of invariance of the splitting under the cocyle γ by

$$\mathcal{I}_\rho(\gamma, E) \equiv \max_{\substack{\sigma,\sigma' \in \{s,c,u\} \\ \sigma \neq \sigma'}} \|\gamma_\theta^{\sigma,\sigma'}\|_\rho \ , \tag{2.21}$$

where $\|\cdot\|_\rho$ is the supremum on $\mathbb{T}^d_\rho$. The choice of a smooth norm is consistent with what we will do in the paper, but of course, other smooth norms could also be considered.

We note that given a splitting specified with some finite error and some γ also specified with a finite error, it is possible to obtain error bounds for (2.21) with finite calculations. In particular, they can be verified with computer assisted proofs.

The following definition taken from [DV17], formulates a notion of approximately invariant splittings and approximately hyperbolic cocycles.

Definition 12. *Given a cocycle γ and a 3-splitting $E = E^s_\theta \oplus E^c_\theta \oplus E^u_\theta$, we can write the cocycle in blocks as in* (2.20)*.*

We say that the splitting is η approximately invariant if

$$\max_{\substack{\sigma,\sigma' \in \{s,c,u\} \\ \sigma \neq \sigma'}} \|\gamma_\theta^{\sigma,\sigma'}\|_\rho \leq \eta \ . \tag{2.22}$$

We say that the cocycle is approximately hyperbolic with respect to the splitting E, if the cocycle

$$\tilde{\gamma}_\theta = \begin{pmatrix} \gamma_\theta^{s,s} & 0 & 0 \\ 0 & \gamma_\theta^{c,c} & 0 \\ 0 & 0 & \gamma_\theta^{u,u} \end{pmatrix}$$

satisfies the trichotomy properties in Definition 11, with $\gamma_\theta^{\sigma,\sigma}$ defined in (2.20)*.*

We again note that the notion of measurements of approximate invariance involves the use of a smooth norm to measure the distance.

2.7. **The closing lemma for approximately invariant splittings.** In this Section, we present and prove a result generalizing slightly and making more precise Proposition 5.2 of [FdlLS09a]. The main result in this Section is Lemma 14 that roughly says that if a splitting is approximately invariant (with a sufficiently small error) for a cocycle and the cocycle is approximately hyperbolic for this splitting in the sense of Definition 12, then there is a true invariant splitting. The proof of Lemma 14 will be given in Appendix A, where we present very detailed estimates.

For the applications we have in mind, it will be important that the size of the corrections of the bundles needed to make them invariant can be bounded by the error in the invariance and that the constants involved in the lemma can be chosen uniformly in a neighborhood of splittings and cocycles. We have included these precisions in the statement of Lemma 14.

The ideas of the proof are standard among the specialists in hyperbolic dynamical systems. The basic idea is that we use the parameterization of splittings by the mappings A^{σ}_{θ} as in (2.16), formulate an invariance equation and manipulate it in a form that can be shown to be a contraction in appropriate spaces. After that, we will need to estimate the changes in the hyperbolicity characteristics.

If we fix a reference splitting $E = E^s \oplus E^c \oplus E^u$, we can parameterize all the splittings close to it by a triple of linear functions as in (2.16). Note that we are not assuming that the reference splitting is invariant, but we will assume it is approximately invariant. The results described below will hold for a sufficiently small neighborhood of the splittings in the Grasmannian space. More precisely, we will consider splittings that can be expressed as in (2.16) with sufficiently small $\|A^{\sigma}_{\theta}\|$.

Remark 13. *In our application to KAM theorem, we can take as the reference splitting the one in the first iterative step. As we will show, if the error in the first approximation is small enough, all the splittings remain in this neighborhood. This smallness condition in the first approximation will be one of the conditions that appear in the final smallness conditions of our main result.*

Lemma 14. *Assume that we have fixed an analytic reference 3-splitting E_0 defined on $\mathbb{T}^d_\rho$. Assume that this splitting is approximately hyperbolic and η_0 approximately invariant with η_0 sufficiently small.*

We denote by $\mathcal{U}$ a sufficiently small neighborhood of this splitting, so that all the splittings can be parameterized as graphs of linear maps A^{σ}_{θ} as in (2.16) with $\|A^{\sigma}_{\theta}\|_\rho < M_1$ for some $M_1 > 0$.

Let γ be an analytic cocycle over a rotation defined on $\mathbb{T}^d_\rho$. Assume that $\|\gamma\|_\rho < M_2$ for some $M_2 > 0$.

Let E be an analytic 3-splitting in the neighborhood $\mathcal{U}$.

Assume that E is η approximately invariant under γ and that γ is approximately hyperbolic for the reference splitting in the sense of Definition 12.

Assume that η is sufficiently small (depending only on the neighborhood $\mathcal{U}$ and M_2).

Then, there is a locally unique splitting $\widetilde{E}$ invariant under γ close to E in the sense that

$$\mathrm{dist}_\rho(E, \widetilde{E}) \le C\eta$$

for some constant $C > 0$.

The splitting $\widetilde{E}$ satisfies a trichotomy in the sense of Definition 11.

The constant C can be chosen uniformly depending only on M_1, M_2.

We stress that Lemma 14 does not require any non-resonance condition on the frequency ω, but it is quite important that the dynamics in the base is a rotation.

Note that there is no domain loss. The changes required are in same spaces where we assumed the error. This is because the proof of Lemma 14 will be just an application of the contraction mapping theorem, so it does not incur in any loss of domain. If the dynamics on the base of the cocycle was not a rotation, we could not be using analytic regularity.

The proof of Lemma 14 is given in Appendix A. It is based on formulating a functional equation for the quantities A^σ given in (2.16). For future use, it is convenient to mention now that after some manipulation one is led to solve the following equations for the dichotomy between $s, \hat{s}$ spaces:

$$\begin{aligned} &(\gamma_\theta^{\hat{s},\hat{s}})^{-1}\left[A^s_{\theta+\omega}\big(\gamma_\theta^{s,s}+\gamma_\theta^{s,\hat{s}}A^s_\theta\big)-\gamma_\theta^{\hat{s},s}\right]=A^s_\theta\ ,\\ &\left[\gamma^{s,\hat{s}}_{\theta-\omega}+\gamma^{s,s}_{\theta-\omega}A^{\hat{s}}_{\theta-\omega}-A^{\hat{s}}_\theta\gamma^{\hat{s},s}_{\theta-\omega}A^{\hat{s}}_{\theta-\omega}\right]\left(\gamma^{\hat{s},\hat{s}}_{\theta-\omega}\right)^{-1}=A^{\hat{s}}_\theta \end{aligned} \tag{2.23}$$

and the following equations for the dichotomy between $u, \hat{u}$ spaces:

$$\begin{aligned} &\left[-A^u_\theta\gamma^{u,\hat{u}}_{\theta-\omega}A^u_{\theta-\omega}+\gamma^{\hat{u},u}_{\theta-\omega}+\gamma^{\hat{u},\hat{u}}_{\theta-\omega}A^u_{\theta-\omega}\right](\gamma^{u,u}_{\theta-\omega})^{-1}=A^u_\theta\\ &(\gamma_\theta^{u,u})^{-1}\left[A^{\hat{u}}_{\theta+\omega}\big(\gamma_\theta^{\hat{u},\hat{u}}+\gamma_\theta^{\hat{u},u}A^{\hat{u}}_\theta\big)-\gamma_\theta^{u,\hat{u}}\right]=A^{\hat{u}}_\theta\ . \end{aligned} \tag{2.24}$$

2.7.1. *Estimating the change of hyperbolicity properties in terms of the change in the cocycles.* In Lemma 14 we showed that given an approximately invariant splitting satisfying the hyperbolicity conditions in Definition 12 there is a truly invariant splitting; we estimated the distance between the approximately invariant splitting and the truly invariant one and the diagonal blocks of the cocycle.

The goal of this Section is to obtain estimates on the hyperbolicity properties of the new cocycles in terms of the size of the changes. Estimates are, of course, not unique and we find it interesting to provide several versions which may be useful in different circumstances. In our case, hyperbolicity properties involve rates and constants and there are trade–offs among them. Note that in contrast to [SS74] we do not study optimal rates, but only bounds on the rates.

In Lemma 15 we provide simple, but generally applicable, estimates and in Lemma 16 we provide more quantitative estimates. We will use Lemma 15 to justify the repeated application of our Newton step and to obtain estimates of the change of the splitting and the diagonal blocks of the whole process. Then, we can apply Lemma 16 to obtain information on the change experienced in the whole Newton process.

We now need to introduce the following notation. Given γ as in (2.11), we denote $\gamma_j(\theta) = \gamma(\theta + j\omega)$ so that

$$\Gamma^k_m \Gamma^m_j(\theta) = \Gamma^k_j(\theta) \ , \tag{2.25}$$

where $\Gamma^k_0 = \Gamma^k$ and $\Gamma^m_j(\theta) = \Gamma^{m-j}_0 \circ T_{j\omega}(\theta)$.

We will assume that for some $C_0, \xi \in \mathbb{R}_+$

$$\|\Gamma^m\|_\rho \le C_0 \xi^m \ .$$

Note that multiplying the generator of the cocycle by a constant number (hence the Γ is multiplied by an exponential in n), we can arrange that the quantity ξ which measures the growth of the cocycle satisfies $\xi < 1$. The multiplication by a constant does not change the invariant spaces and only changes the smallness conditions by a constant.

We want to investigate conditions on $\|\gamma - \tilde\gamma\|_\rho$ so that we can ensure that $\widetilde\Gamma$ obtained by iterating as in (2.25) satisfies (5.24) for other constants $\tilde C_0$, $\tilde\xi$.

A general estimate on changes of hyperbolicity properties is given by the following result, whose proof is given in Appendix A.

Lemma 15. *Assume that Γ obtained from a cocycle γ satisfies* (5.24)*. Let $\varepsilon^* > 0$ and let $\tilde\gamma$ be a cocycle such that*

$$\|\gamma - \tilde\gamma\|_\rho \le \varepsilon^* \ ;$$

then, there exist $\tilde C_0$ and $\tilde\xi$ such that $\|\widetilde\Gamma^m\|_\rho \le \tilde C_0 \tilde\xi^m$.

A more quantitative estimate on changes of hyperbolicity properties is given by the following result, whose proof is given in Appendix A.

Lemma 16. *With the notations of Lemma 15, let $a \equiv \|\gamma - \tilde\gamma\|_\rho$ be small enough, say $a \le \frac{1}{4C_0}$. Then, we can take $\tilde\xi = \xi + Ca$ with an explicit constant C.*

Furthermore, the quantity $\tilde C_0 = \tilde C_0(C_0, \mu, a)$ in Lemma 15 can be bounded as

$$\tilde C_0 \le 4 C_0^2 a \xi^{-1} \frac{\tilde\xi}{\tilde\xi - \xi} \ . \tag{2.26}$$

An important corollary of Lemma 14 is the following.

Corollary 17. *If the reference splitting is approximately hyperbolic in the sense of Definition 12 for some cocycle γ^0_θ, then for all the γ's in a neighborhood of γ^0_θ, there is an* *(locally unique) invariant splitting which is hyperbolic and the hyperbolicity constants can be chosen uniformly.*

Remark 18. *The estimates on the hyperbolicity constants in* (2.26) *of Lemma 16 involve choices. One can make C_0 change or λ's change.*

Much of the theory (e.g. [SS74]*) is concerned with the optimal λ's. Note that, even for constant 2×2 cocycles, the optimal λ can change with a fractional power of the perturbation. Once we choose a slightly less optimal λ's we can make the change to be linear in the perturbation. Of course, the range of validity, may be smaller if the chosen upper bound is close to the optimal one.*

Note that, the C_0 depends on the choice of metrics – but the rates λ_* do not. Indeed, it is customary in the theory that deals with perturbations to observe that we can choose an "adapted metric" *so that $C_0 = 1$. Of course, the size of the perturbations allowed is measured in this metric and, when C_0 increases, the adapted metric becomes more inequivalent to the original one and the perturbations allowed may decrease.*

Numerical explorations ([HdlL07, HdlL06a]*) suggest that if one fixes a metric and studies the optimal rates λ's and the optimal C_0, there is a very interesting scenario for the loss of hyperbolicity called* "bundle collapse"*. In this scenario, the rates λ's remain bounded and the C_0 explodes. This scenario empirically presents remarkable scaling properties. The bundle collapse scenario is particularly important in the breakdown of KAM tori in conformally symplectic systems (see* [CF12]*). The papers* [HdlL07, HdlL06a] *presented numerical conjectures of the blow-up of the optimal values of the rates and of the geometric properties of the bundles. These conjectures were recently proved in several cases in* [BS08, OT17, FT18, Tim18].

We think it would be interesting to study the breakdown of hyperbolicty in conformally symplectic systems. It seems possible that the limit of zero dissipation will have some interactions with the previously studied phenomena. For these numerical implementations, the a-posteriori results and the fast algorithms developed here are likely to be useful.

2.8. **Whiskered tori.** The main result of this paper concerns whiskered tori, which are defined as follows.

Definition 19. *Let f_μ be a conformally symplectic system with conformal factor λ of a symplectic manifold $\mathcal{M}$.*

We say that $K : \mathbb{T}^d \to \mathcal{M}$ is a whiskered torus when:

(1) *K is the embedding of a rotational torus, that is, $f_\mu \circ K = K \circ T_\omega$.*

(2) *The cocycle* $Df_\mu \circ K$ *over the rotation* T_ω *admits a trichotomy as in Definition 11 with the rates* $\lambda_-, \lambda_c^-, \lambda_c^+, \lambda_+$.
(3) *The rates satisfy* $\lambda_c^- \leq \lambda \leq \lambda_c^+$.
(4) *The spaces* E^c *have dimension* $2d$.

Somewhat surprisingly there are relations between the conformal symplectic properties, the rates and the properties of the bundles. They will be explored in Section 4. Notably we will show that if an embedding satisfies Definition 19, then it is isotropic, there are relations between the rates of growth and, more surprisingly, the E^c bundle is trivial.

3. Statement of the main result, Theorem 20

The main result of this paper is an a-posteriori result about solutions of a parameterized version of (2.7).

As motivation for the hypothesis of Theorem 20, assume that K, μ satisfy approximately (2.7) with a small error term e, i.e.

$$f_\mu \circ K - K \circ T_\omega = e \ .$$

If we want that $K + \Delta$, $\mu + \beta$ for some corrections Δ, β is a better solution, the Newton-Kantorovich method would prescribe to choose Δ, β satisfying

$$Df_\mu \circ K\Delta - \Delta \circ T_\omega + (D_\mu f_\mu) \circ K\beta = -e \ . \tag{3.1}$$

If one tries to solve (3.1) by iterating, one is quickly led to the cocycles that were discussed in Section 2.6. Hence, it is clear that the asymptotic growth of the cocycles plays a role.

As it turns out, the geometry of the problem plays also a very important role and one of the most surprising facts is that the conformal symplectic geometry leads to constraints on the rates of growth. These interactions of the geometry with the dynamics will be explored in Section 4. We anticipate that the most important results will be a surprising triviality result for the bundle of vectors with intermediate slow decay and the *"automatic reducibility"* that constructs a natural system of coordinates in which the linearized equations are very simple. In this paper we go beyond the results in previous papers and show in Lemma 25 that the center bundle is trivial.

Using the geometry, we will show that the equations (3.1) can be solved. As mentioned in Section 1, the result is a very efficient algorithm. Of course, the a-posteriori format of the theorem gives an analytical support to the results.

Given the important role played by the geometry, it is clear that the limit when the geometry changes from conformally symplectic to symplectic is very singular. In Section 7 we will study this singular limit in which the dissipation becomes weak.

The following Theorem 20 on the persistence of whiskered tori is the main result of this paper. Later, we will use it to obtain information on the analyticity properties of the tori under dissipative perturbations (see Theorem 20).

We will consider specially the case $0 < |\lambda| < 1$, but Theorem 20 can be stated as well for $|\lambda| > 1$, just taking the inverse of the mapping. In the discussion of analyticity properties with respect to perturbations, we will need to consider even complex values. We will also consider the case $\lambda = 1$, but, as pointed out in [CCdlL17], the case of complex λ with $|\lambda| = 1$, $\lambda \neq 1$ requires special considerations. Indeed, when λ is a root of the identity, we do not expect that the solutions persist in general. Indeed, for generic perturbations, it is impossible to find even formal asymptotic expansions.

Theorem 20. *Let $\omega \in \mathcal{D}_d(\nu, \tau)$, $d \leq n$, as in (2.4), let $\mathcal{M}$ be as in Section 2.1 and let $f_\mu : \mathcal{M} \to \mathcal{M}$, $\mu \in \mathbb{R}^d$, be a family of real analytic, conformally symplectic mappings as in (2.1) with $0 < \lambda < 1$. We make the following assumptions.*

$(H1)$ Approximate solution:

Let (K_0, μ_0) with $K_0 : \mathbb{T}^d \to \mathcal{M}$, $K_0 \in \mathcal{A}_\rho$, and $\mu_0 \in \mathbb{R}^d$ define an approximate whiskered torus with frequency ω for f_{μ_0}, so that

$$\|f_{\mu_0} \circ K_0 - K_0 \circ T_\omega\|_\rho \leq \mathcal{E} \tag{3.2}$$

for some $\mathcal{E} > 0$.

To ensure that the composition of f_μ and K can be defined, we will assume that the range of K_0 is well inside the domain of f_μ for all μ sufficiently close to μ_0.

We will assume that there is a domain $\mathcal{U} \subset \mathbb{C}^n/\mathbb{Z}^n \times \mathbb{C}^n$ such that for all μ_0 such that $|\mu - \mu_0| \leq \eta$, f_μ has domain $\mathcal{U}$. Moreover, we assume that the range of K_0 is inside the domain $\mathcal{U}$:

$$\mathrm{dist}(K_0(\mathbb{T}^d_\rho), \mathbb{C}^n/\mathbb{Z}^n \times \mathbb{C}^n \setminus \mathcal{U}) \geq \eta\ . \tag{3.3}$$

$(H2)$ Approximate splitting:

For all the points in the torus, there exists a splitting of the tangent space of the phase space, depending analytically on the angle $\theta \in \mathbb{T}^d_\rho$.

These bundles are approximately invariant under the cocycle $\gamma(\theta) = Df_{\mu_0} \circ K_0(\theta)$, namely the quantity in (2.21) is smaller than $\mathcal{E}_h$, for some $\mathcal{E}_h > 0$.

$(H3)$ Spectral condition for the bundles (exponential trichotomy):

For all $\theta \in \mathbb{T}^d_\rho$ the spaces in $(H2)$ are approximately hyperbolic for the cocycle $\gamma(\theta)$ (see Definition 12). We recall that this just entails that the diagonal cocycles have different rates of growth and hyperbolicity constant that satisfy (2.14).

$(H3')$ *Since we are dealing with conformally symplectic systems and are interested in the almost symplectic limit, we will also assume*[45]:

$$\lambda_- < \lambda\lambda_+ < \lambda_c^- \ , \qquad \lambda_c^- \le \lambda \le \lambda_c^+ \ .$$

$(H4)$ *We assume that the dimension of the center subspace*[6] *is* $2d$.

$(H5)$ *Non–degeneracy:* [7]

Denote by J_c the operator J restricted to the center space (we will show in Lemma 22 that J_c is a non-degenerate matrix).

Let

$$N(\theta) = (DK(\theta)^T DK(\theta))^{-1} \ , \tag{3.4}$$

$$P(\theta) = DK(\theta)N(\theta) \ ,$$

$$\chi(\theta) = DK(\theta)^T (J^c)^{-1} \circ K(\theta) DK(\theta) \ .$$

Let M, S be auxiliary quantities defined as

$$M(\theta) = [DK(\theta) \mid (J^c)^{-1} \circ K(\theta)\ DK(\theta)N(\theta)] \tag{3.5}$$

and

$$S(\theta) \equiv P(\theta+\omega)^T Df_\mu \circ K(\theta)(J^c)^{-1} \circ K(\theta)P(\theta) - N(\theta+\omega)^T \chi(\theta+\omega)N(\theta+\omega)\ \lambda\ \mathrm{Id}_d \ . \tag{3.6}$$

We assume that the following non–degeneracy condition is satisfied, precisely that the matrix $\mathcal{S}$ defined below is invertible:

$$\mathcal{S} \equiv \begin{pmatrix} \overline{S} & \overline{S(W_b^c)^0} + \overline{\widetilde{A}_1^c} \\ (\lambda - 1)\,\mathrm{Id}_d & \overline{\widetilde{A}_2^c} \end{pmatrix} \ , \qquad \det \mathcal{S} \neq 0 \ , \tag{3.7}$$

---

[4]As we will show in Section 4, the interaction between conformally symplectic systems and the exponential trichotomy implies further restrictions which follow from the present assumptions.

[5]Note that we have used λ for the conformal factor and λ_σ for the different bounds on rates. Even if these are conceptually very different things, we will show that they are related. This justifies using similar letters.

[6]The content of this assumption is that the dimension of the center bundle is exactly twice the dimension of the invariant torus. As we will show later, the dimension has to be at least twice.

[7]The idea of the condition is that a very explicit $2d \times 2d$ matrix is invertible. We will formulate it here in detail, but the main point is that the condition can be verified with a finite computation on the approximate solution and the approximate bundles given in $(H1)$, $(H2)$.

where the bar denotes the average, $\widetilde{A}^c_1$, $\widetilde{A}^c_2$ denote the first d and the last d rows of the $2d\times d$ matrix $\widetilde{A}^c \equiv [\widetilde{A}^c_1|\widetilde{A}^c_2] = M^{-1}\circ T_\omega D_\mu f_\mu \circ K$, $(W^c_b)^0$ is the solution of $\lambda(W^c_b)^0 - (W^c_b)^0\circ T_\omega = -(\widetilde{A}^c_2)^0$, where $(\widetilde{A}^c_2)^0 = \widetilde{A}^c_2 - \overline{\widetilde{A}^c_2}$.

Let $\alpha(\tau)$ be an explicit number (see the discussion later for the values that come from the proof). Assume that for some δ, $0<\delta<\rho$, we have

$$\mathcal{E}_h \le \mathcal{E}^*_h\ , \qquad \mathcal{E} \le \delta^{2\alpha}\mathcal{E}^*\ ,$$

where $\mathcal{E}^*_h$, $\mathcal{E}^*$ are explicit functions given along the proof and depending on the following quantities:

$$\begin{aligned}&\nu, \tau, C_0, \lambda_+, \lambda_-, \lambda_c^+, \lambda_c^-, \|\Pi_\theta^{s/u/c}\|_\rho\\ &\|DK_0\|_\rho, \|(DK_0^T DK_0)^{-1}\|_\rho, |\mathcal{S}^{-1}|, \max_{j=0,1,2}\ \sup_{|\mu-\mu_0|\le\eta_0} \|D^j f_\mu\|_{\mathcal{U}}\ .\end{aligned} \tag{3.8}$$

Then, there exists an exact solution (K_e,μ_e), such that

$$f_{\mu_e}\circ K_e - K_e\circ T_\omega = 0$$

with

$$\|K_e - K_0\|_{\rho-2\delta} \le C\mathcal{E}\delta^{-\tau}, \qquad |\mu_e-\mu_0| \le C\mathcal{E}\ , \tag{3.9}$$

where C is a constant whose explicit expression can be obtained from the proof and which depends on the same variables as $\mathcal{E}^*_h, \mathcal{E}^*$.

Furthermore, the invariant torus K_e is hyperbolic in the sense that there exists an invariant splitting

$$\mathcal{T}_{K_e(\theta)}\mathcal{M} = E^s_\theta \oplus E^c_\theta \oplus E^u_\theta\ ,$$

that satisfies Definition 11.

The splitting of the invariant torus is close to the original one in the sense that, for some constant $C>0$, one has

$$\|\Pi_0^{s/u/c} - \Pi_f^{s/u/c}\|_{\rho-2\delta} \le C(\mathcal{E}\delta^{-\tau} + \mathcal{E}_h) \tag{3.10}$$

(as remarked above, this is equivalent to the analytic Grassmanian distance).

Moreover, the hyperbolicity constants corresponding to the invariant splitting of the invariant torus (which we denote by a tilde) can be taken to be close to those of the approximately invariant splitting of the approximate invariant torus assumed to exist in (H1), (H2):

$$\begin{aligned}|\lambda_\pm - \widetilde{\lambda}_\pm| &\le C(\mathcal{E}\delta^{-\tau} + \mathcal{E}_h),\\ |\lambda_c^\pm - \widetilde{\lambda}_c^\pm| &\le C(\mathcal{E}\delta^{-\tau} + \mathcal{E}_h)\ .\end{aligned} \tag{3.11}$$

The proof of Theorem 20 is postponed to Section 5, since we devote Section 4 to discuss some properties stemming from the geometry of conformally symplectic systems. Later in this paper, we will present other results. Notably, we will study the domain of analyticity of the tori for the small dissipation regime (see Theorem 38).

We also obtain explicit estimates on the new hyperbolicity constants C_0, but they are too cumbersome to state now (see Lemma 16 for more detailed estimates on the new hyperbolicity constants).

3.1. **Some remarks and comments on the statement of Theorem 20.** We collect in this Section some useful comments on the content of Theorem 20 and comparisons with other results in the literature.

• Note that the non-degeneracy quantities in (3.8) are quantities that can be estimated just on the approximate solution. The only global property of the function needed is an estimate on $\sup_{|\mu-\mu_0|\leq\eta_0} \|D^j f_\mu\|_{\mathcal{U}}$ and we do not need delicate global properties of the map such as a global twist condition.

• Note that Theorem 20 is stated without any reference to an integrable system. We just need an approximate solution of the invariance equation.

• We are not assuming that any of the invariant bundles are trivial (but we will show that the center bundle is trivial as a consequence of the other hypotheses).

• The twist non-degeneracy condition $(H5)$ is just that a very explicit $2d \times 2d$ matrix is non-degenerate. This matrix is formed by the derivatives of the approximate solution, performing algebraic operations and averages. It can be computed with a finite number of computations from the approximate solution.

• The above formulation gives a very transparent proof of several *"small twist results"*. One can construct perturbative expansions that satisfy the invariance equation to arbitrarily high powers of the perturbation parameter. At the same time (performing calculations) one can prove that the twist, hyperbolicity, etc., start to grow like a finite power of the perturbation. Then, the theorem will imply the existence of a solution.

Another application included in this paper is that we will prove small hyperbolicity assumptions, see Theorem 38.

Another non-degeneracy assumption we will need is that the matrix M introduced in (3.5) is invertible if the initial error is small enough. We will also show that the iterative procedure maintains the uniform bounds in M^{-1}. In computer assisted proofs, and more explicit treatments, it is advantageous to obtain precise estimates for M^{-1} at the initial step.

• The hypothesis $(H5)$ is analogue to the Kolmogorov non-degeneracy condition.We note that if $\lambda = 1$ – the symplectic case – then, the condition just becomes $\overline{S}$ being invertible. For an integrable system, this is the Kolmogorov non-degeneracy condition.

• The condition $(H5)$ is not a global property of the map. It is only a numerical condition evaluated on the approximate solution. It can be readily computed by taking derivatives, performing algebraic operations and taking averages.

• It is possible to use the method of [Mos67] or the method of [Yoc92, Sev99] to obtain the result under much weaker non–degeneracy conditions than $(H5)$ such as Rüssmann non-degeneracy conditions.

The proof of this result is particularly transparent taking advantage of the a-posteriori format which gives very easily the dependence on parameters.

• We note that, thanks to Lemma 14, instead of the approximate invariance of the splitting included in $(H2)$, we could have assumed that the approximately invariant torus has an invariant splitting.

We have chosen the present formulation to emphasize that all the hypotheses of Theorem 20 can be verified from an approximate solution with just a finite precision computation.

• We have treated separately the smallness conditions in the invariance of the torus $\mathcal{E}$ and the smallness condition in the invariance of the hyperbolic splitting $\mathcal{E}_h$. As we will see, the error in the invariance of the hyperbolic splitting can be eliminated with a contraction point argument. Eliminating the error $\mathcal{E}$ requires a Nash-Moser iteration to beat the small divisors that appear.

• The proof of Theorem 20 will be based on describing an iterative process which leads to a very efficient algorithm. To obtain an algorithm from the proof of Theorem 20, one needs to present also descriptions of the discretizations of the bundles and finite calculations that allow to verify the hypotheses. These algorithmic details are presented in [HdlLS11] for the symplectic case and they do not need to be modified in our case.

• The error in the hyperbolicity plays a very different role than the error in the invariance in the iterative process. We could think of the hyperbolicity as a preconditioner for the Newton method for the invariance equation. As we will see, the iterative step has an upper triangular structure. The error in the invariant splittings can be eliminated without affecting the embeddings. On the other hand, if we modify the embedding K and the drift μ, we modify $Df_\mu \circ K$ and have to correct for the invariant embedding. This elementary remark will be important for the study of Lindstedt series in Section 7.

• The conformal symplectic properties of the map imposes many relations between the properties of the invariant splitting. These will be discussed in Section 4.

• Notice that when $\lambda \neq 1$, the invariant torus is normally hyperbolic since the center direction, as remarked above, has the conformal factor λ as the multiplier.

This observation allows one to obtain several results, slightly weaker than Theorem 20.

(i) Using a-posteriori formulations of the theory of normally hyperbolic invariant manifolds ([BLZ08, CZ11]), we obtain from the hypotheses on the approximate invariance and the approximate invariant splitting that there are smooth invariant tori for all perturbations (no need to adjust the drift!). Of course, we do not know that the motion in the manifold will be conjugate to a rotation.

If we change the drift, using the theory of [Mos66] we obtain, under some non-degeneracy conditions that, for appropriate choices of the drift, the motion on the torus is conjugate to a rotation. We refer to [CCFdlL14] for more details on the argument and for an application of this strategy to discuss phase locking and other situations when the motion is not conjugate to a rotation. Notice that this method produces only finitely differentiable objects and not analytic ones as the present method. Also, the algorithms they give rise are very different.

(ii) An alternative approach is in [CH17b], which deals with normally hyperbolic tori using the fact that in the stable and unstable directions we can use an iterative method to solve the linearized invariance equation. For the tangent directions one needs to adjust parameters to solve the conjugacy equation. Notice that this method is different from the normally hyperbolic method, since it produces analytic manifolds but needs to adjust parameters. This technique leads to very efficient numerical methods that have been implemented in [CH17a].

(iii) Note that the above mentioned approaches do not require (and do not take advantage of) the conformally symplectic geometry. This generality is useful for some models of friction in which the friction does not lead to a conformally symplectic system. The limit of weak dissipation in such cases seems a challenging problem.

(iv) On the other hand, we note that these methods have estimates that blow up as λ goes to 1, whereas the method of this paper leads to a comfortable study of the small dissipation limit. Indeed, one of the main results of this paper is the study of the analyticity domain in the zero dissipation limit, see Section 7. One of our motivations was precisely the studies in celestial mechanics where the dissipation is indeed small and the zero dissipation limit is very relevant.

• It is important to remark that λ_+, λ_-, λ_c^+, λ_c^- appearing in $(H3)$ are only upper bounds. Hence, they are not uniquely defined. When we consider such values optimal, we obtain the Sacker-Sell spectrum [SS74]. The pairing rule provides relations between the bounds, but they become equalities for the optimal values.

4. Some consequences of the geometry

4.1. **Introduction.** In this Section, we present consequences of the conformally symplectic systems and the trichotomy assumptions for an (approximately) invariant torus with an (approximately) invariant splitting.

The geometrically natural arguments (leading to the sharper results) happen when the torus is invariant and the bundle is invariant. The main reason is that we need to compare vectors and forms in $f(K(\theta))$ and in $K(\theta+\omega)$. An alternative to invariance, is that Ω is constant. Of course, our iterative process to improve approximate solutions, needs to take advantage of the geometry for the approximate solutions, which are slightly weaker than the geometrically natural ones.

Hence we will introduce provisionally the hypothesis (HI) below to be able to carry out geometrically natural arguments. In Section 4.9, we show how to remove this assumption.

(HI) Assume either:

(HI.A)**:** K is an embedding of the torus satisfying (2.6) and E is a splitting of the tangent bundle to the phase space invariant under the cocycle $Df\circ K(\theta)$ under the rotation ω.

(HI.B)**:** The phase space is Euclidean and the symplectic form Ω is constant.

In many practical applications the case B) in the alternative above holds.

Of course, in the iterative step of the KAM theorem, we cannot assume (HI.A), that the torus is invariant. The assumption (HI.A) is natural in the development of Lindstedt series that will take place in Section 7.

The removal of (HI.A) in Section 4.9 will be obtained just by examining carefully the naturally geometric argument and adding some extra terms that are controlled by the invariance error and its derivatives.

4.2. **The results in this Section.** The first result we will present is the well known pairing rule ([DM96b, WL98]), which relates the stable/unstable exponential rates (see Section 4.3).

In Section 4.4 we show the isotropic property of invariant tori, namely that the symplectic form restricted to the torus is zero.

We will also show that, because of the conformally symplectic structure, we have that the symplectic form restricted to the center is non-degenerate, see Lemma 22 in Section 4.5.

A rather remarkable result obtained here is that the center bundle has to be trivial, see Section 4.6. This solves a question raised in [FdlLS09a], which constructed examples where the stable and unstable bundles were non-trivial. The automatic reducibility is discussed in Section 4.7 and some consequences in Section 4.8. Geometric identities for approximately invariant tori are presented in Section 4.9.

4.3. **Geometry and rates.** The conformal symplectic properties of the maps imply constraints on the rates assumed in $(H3)$. In this Section, we develop two of them: the pairing rule and the rigidity of rates in the center. These properties will not be used in the proof of Theorem 20 and, hence, can be omitted, but we include them since the method of proof is useful in other parts of the paper.

4.3.1. *The Pairing rule.* The paper [DM96b] studies the effect on the geometry of eigenvalues of periodic orbits; the paper [WL98] studies the effect on the Lyapunov exponents of cocyles. In our case, we want to study the relation with the optimal rates appearing in $(H3)$, which are known also as Sacker-Sell spectrum ([SS74]). We note that the paper [WL98], since it worked for general cocycles, did not take advantage of the fact that, for diffeomorphisms, the factor has to be a constant when $n \geq 2$ (see the argument after Definition 2). Therefore, some of the formulas in [WL98] can be simplified for the applications of this paper. We will revisit a more detailed comparison with these papers in [CCdlL19b].

The key observation is that since

$$\Omega(Df^n(x)u, Df^n(x)v) = \lambda^n \Omega(u, v) \ ,$$

then, if $|Df^n(x)v| \leq C\lambda_-^n|v|$ for any $n \geq 0$, we should have

$$|Df^n(x)u| \geq \tilde{C}\ \lambda^n\ \lambda_-^{-n}\ \Omega(u,v)\ |v|^{-1}$$

for some positive constant $\tilde{C}$. That is, if there is a vector that decreases exponentially fast, there should be others which grow faster than the rates.

Hence, if there is a vector with Lyapunov multiplier smaller than λ, there should be another one with Lyapunov multiplier bigger than ${\lambda_-}^{-1}\lambda$. By reversing the argument we obtain that the set of Lyapunov multipliers $\{\lambda_i\}_{i=1}^{2d}$ should satisfy the pairing rule

$$\lambda_i\ \lambda_{i+d} = \lambda \tag{4.1}$$

(compare with the corresponding formula in [WL98], which involves an integral of $\log\lambda$). We remark that our desired result is different than that of [WL98], since we want to obtain uniform bounds rather than Lyapunov exponents.

In our case, we can obtain uniform bounds on the growth, using Corollary 24 below and other elementary arguments.

Remark 21. *There are more general arguments in Sacker-Sell theory, relating the edge of the Sacker-Sell spectrum and the supremum of Lyapunov exponents of all measures (see* [SS74]*). We will not emphasize those arguments, since we will not use them.*

We note that for every x and any $u \in E_x^u$, $|u| = 1$, there exists $v \in E_x^s$, $|v| = 1$, such that $\Omega(u,v) \neq 0$ (see Corollary 24 below). Using the continuity and compactness of $\mathbb{T}^d$ and the spheres in the unit bundle, we obtain

$$\inf_{x\in K(\mathbb{T}^d_\rho)}\ \inf_{u\in E^u_{K(x)}, |u|=1}\ \inf_{v\in E^s_{K(x)}, |v|=1}\ |\Omega_x(u,v)| \geq \zeta$$

for some positive constant ζ. Therefore, given $u \in E^u_{K(x)}$, we can choose $v \in E^s_{K(x)}$ so that

$$|Df^n(K(x))u| \geq \tilde{C}\zeta(\lambda\lambda_-^{-1})^n|u|\ . \tag{4.2}$$

Using that the bounds (4.2) are uniform, we see that given $u \in E^u_{K(x)}$, we can apply them to $Df^{-n}(K(x))u$ and obtain

$$|Df^{-n}(K(x))u| \leq \tilde{C}^{-1}\zeta^{-1}(\lambda^{-1}\lambda_-)^n|u|\ .$$

Therefore we obtain that we should have

$$\lambda^{-1}\lambda_- \leq \lambda_+\ .$$

By applying a similar argument to the bounds along the stable direction, we obtain the other inequality, $\lambda^{-1}\lambda_- \geq \lambda_+$, which leads to (4.1), that is $\lambda_-\lambda_+^{-1} = \lambda$ for the optimal values.

4.4. **Isotropic properties of rotational invariant tori.** The isotropic property means that the symplectic form restricted to the invariant torus is zero.

To establish that rotational tori are indeed isotropic, we note that by the invariance equation (2.7) we have

$$K^* f_\mu^* \Omega = T_\omega^* K^* \Omega \ .$$

Since f_μ is conformally symplectic, according to (2.1) we have

$$\lambda \ K^*\Omega = T_\omega^* K^* \Omega \tag{4.3}$$

and if $|\lambda| \neq 1$, we obtain by iterating the relation (4.3) (either in the future or in the past) that

$$K^*\Omega = 0 \ ,$$

thus proving that the tori are isotropic.

In the case that $\lambda = 1$ (this is a case that has been discussed in [Zeh75]), it is required that ω is non-resonant and that the map is exact. We note that, under the non-resonant hypothesis, we obtain that $K^*\Omega$ is given by a constant matrix. Moreover, if $\Omega = d\alpha$, we have $K^*\Omega = K^* d\alpha = dK^*\alpha$. The only exact form with a constant matrix is 0. Note that in the case that $|\lambda| \neq 1$ we do not need that the symplectic form is exact, nor that ω is nonresonant to conclude that the rotational tori are isotropic.

In Section 4.5, we will see that approximately invariant tori are also approximately isotropic. In the $\lambda = 1$ case it requires that ω is Diophantine.

In coordinates, the isotropic property of the invariant torus, using the matrix J defined in Section 2.2, is written as

$$DK^T(\theta) J_{K(\theta)} DK(\theta) = 0 \ . \tag{4.4}$$

The equation (4.4) can be interpreted geometrically as saying that any vector in the range of $DK(\theta)$ is orthogonal to any vector in the range of $J_{K(\theta)} DK(\theta)$.

4.5. **Non degeneracy of the symplectic form restricted to the center bundle of a rotational invariant torus.** The following result shows that there are many cases where the symplectic form Ω has to vanish. As a corollary, we will deduce that the symplectic form is non-degenerate when restricted to the center bundle E_θ^c.

Lemma 22. *Let E^s_θ, E^c_θ, E^u_θ be an invariant splitting around a rotational torus with growth rates as in* $(H3)$*. Then,*

$$\begin{aligned}
&\Omega(s,c)=0 \quad \forall\, s\in E^s_\theta, c\in E^c_\theta\\
&\Omega(u,c)=0 \quad \forall\, u\in E^u_\theta, c\in E^c_\theta\\
&\Omega(s_1,s_2)=0 \quad \forall\, s_1,s_2\in E^s_\theta\\
&\Omega(u_1,u_2)=0 \quad \forall\, u_1,u_2\in E^u_\theta\ .
\end{aligned} \tag{4.5}$$

Proof. Let $s\in E^s_\theta$, $c\in E^c_\theta$; then, one finds that

$$\Omega(s,c)=0\ ,$$

since the following bounds hold for a suitable constant C and for $j\geq 1$:

$$\begin{aligned}
|\Omega(s,c)| &= \left|\frac{1}{\lambda}\,\Omega(Df_\mu\ s, Df_\mu\ c)\right|\\
&= \frac{1}{\lambda^j}\ \left|\Omega(Df^j_\mu\ s, Df^j_\mu\ c)\right|\\
&\leq C\ \left(\frac{\lambda_-\,\lambda_c^-}{\lambda}\right)^j\ ,
\end{aligned}$$

whose limit tends to zero as j goes to infinity for λ_- and λ_c^- as in $(H3)$ and $(H3')$, i.e. using that $\lambda_-\lambda_c^- < (\lambda_c^-)^2\leq\lambda^2$.

A similar argument holds for the unstable bundle,

$$\begin{aligned}
|\Omega(u,c)| &= \left|\lambda\Omega(Df^{-1}_\mu\ u, Df^{-1}_\mu\ c)\right|\\
&= \lambda^j\ \left|\Omega(Df^{-j}_\mu\ u, Df^{-j}_\mu\ c)\right|\\
&\leq C\ (\lambda\,\lambda_+\,\lambda_c^+)^j\ ,
\end{aligned}$$

whose limit goes to zero as j tends to infinity under the condition $(H3')$, i.e. using that $\lambda\lambda_+\lambda_c^+ < \lambda\lambda_+\lambda_+ < \lambda_c^-\lambda_+ < \lambda\lambda_+ < \lambda_c^- \leq \lambda$.

Next we prove the third of (4.5); for any s_1, $s_2\in E^s_\theta$, we have:

$$\begin{aligned}
|\Omega(s_1,s_2)| &= \left|\frac{1}{\lambda}\,\Omega(Df_\mu\ s_1, Df_\mu\ s_2)\right|\\
&= \frac{1}{\lambda^j}\ \left|\Omega(Df^j_\mu\ s_1, Df^j_\mu\ s_2)\right|\\
&\leq C\ \left(\frac{\lambda_-^2}{\lambda}\right)^j\ ,
\end{aligned}$$

which goes to zero for $j\to\infty$ due to $(H3)$ and $(H3')$, since $\lambda_-^2/\lambda < \lambda_-\lambda_c^-/\lambda < 1$. The fourth equation in (4.5) holds under the assumption $\lambda\lambda_+^2 < 1$, which is guaranteed by

$(H3)$ and recalling that $\lambda < 1$. In fact,we have:

$$\begin{aligned}|\Omega(u_1, u_2)| &= \left|\lambda\, \Omega(Df_\mu^{-1}\, u_1, Df_\mu^{-1}\, u_2)\right| \\ &= \lambda^j\, \left|\Omega(Df_\mu^{-j}\, u_1, Df_\mu^{-j}\, u_2)\right| \\ &\leq C\, (\lambda\,\lambda_+^2)^j\ .\end{aligned}$$

□

As we will see, the above results lead to some automatic non-degeneracy conclusions which will be important to develop structures on the theorem.

Corollary 23. *In the hypotheses of Lemma 22 we have that Ω restricted to E_θ^c is non-degenerate.*

Proof. To conclude that Ω restricted to E_θ^c is non–degenerate, we observe that if for some $w \in E_\theta^c$, we have that

$$\Omega(c, w) = 0 \qquad \forall c \in E_\theta^c\ ;$$

then, using that

$$\Omega(s, w) = 0\ , \quad \Omega(u, w) = 0\ , \quad \forall s \in E_\theta^s\ , \quad \forall u \in E_\theta^u\ ,$$

we obtain that $\Omega(v, w) = 0$ for any $v \in \mathcal{T}_\theta\mathcal{M}$. Therefore, since Ω is non-degenerate in the whole space, we conclude that $w = 0$. □

Corollary 24. *If $v \in E_x^s$ and for any $u \in E_x^u$ we have that $\Omega(v, u) = 0$, then $v = 0$. If $\tilde{u} \in E_x^u$ and for any $\tilde{v} \in E_x^s$ we have that $\Omega(\tilde{u}, \tilde{v}) = 0$, then $\tilde{u} = 0$.*

Proof. The proof of Corollary 24 is identical with that of Corollary 23. We note that the hypotheses of Corollary 24 and the results of Lemma 22 imply that $\Omega(v, u) = 0$ for any $u \in \mathcal{T}_x M$ which, by the non-degeneracy of Ω, implies the conclusion of Corollary 24. □

Corollary 24 can be interpreted as saying that some of the matrix elements giving Ω are not degenerate. This will be useful later when we discuss pairing rules for exponents.

4.6. **Triviality of the center bundle.** The main goal of this Section is to show that the bundle E_θ^c based on a rotational invariant torus satisfying our hypotheses (notably that the dimension of the fibers of the bundle is $2d$) is trivial in the sense of bundle theory. That is, we will show that E_θ^c is isomorphic to a product bundle (namely, a trivial bundle in the language of bundle theory).

Furthermore, we show that there is a natural system of coordinates on E_θ^c, see Lemma 25. In this system of coordinates, the linearization of the invariance equation (2.7) restricted to the center space becomes a constant coefficient equation and, hence, can be solved by using Fourier methods, see Lemma 26 in Section 4.9.

Note that the triviality of E^c_θ is in contrast with the stable and unstable bundles, which can be nontrivial (see examples in [FdlLS09a]). Note also that the proof works when the phase space is a manifold and it applies a-fortiori for symplectic systems.

Lemma 25. *Assume that K is an approximate solution of* (2.7). *Then, we can find a linear operator*

$$B_\theta : \mathrm{Range}(DK(\theta)) \to E^c_\theta \ ,$$

such that the center bundle is given by

$$E^c_\theta = \{v + B_\theta v : \ v \in \mathrm{Range}(DK(\theta))\} \ . \tag{4.6}$$

Notice that (4.6) shows that E^c_θ is the range under $\mathrm{Id} + B_\theta$ of the tangent bundle of the torus. This shows that E^c_θ is a trivial bundle.

Proof. We start by remarking that, as it is standard, if we fix a Riemannian metric g, we can identify the two form Ω_x with a linear operator $J(x) : \mathcal{T}_x\mathcal{M} \to \mathcal{T}_x\mathcal{M}$ by requiring

$$g_x(u, J(x)v) = \Omega_x(u, v) \quad \forall u, v \in \mathcal{T}_x\mathcal{M} \ . \tag{4.7}$$

Of course, the operator J depends on the metric chosen (we omit the dependence on the metric from the notation, unless it can lead to error). It will be advantageous for us to choose the metric so that the operator J has extra properties.

We will choose a metric g_x such that the spaces E^c_x, E^s_x, E^u_x are orthogonal under g_x. A possibly degenerate (i.e., assigning zero length to non-zero vectors) metric can be easily constructed in coordinate patches. By adding constructions in different coordinate patches, we can ensure that the resulting metric is not degenerate.

We denote by $\widetilde{J}_x$ the operator corresponding via (4.7) with Ω_x using the metric constructed above, which makes the splitting orthogonal. The properties established in (4.5) imply that if we decompose the operator $\widetilde{J}_x$ in blocks corresponding to the decomposition $\mathcal{T}_x\mathcal{M} = E^c_x \oplus E^s_x \oplus E^u_x$, then we have the block structure:

$$\widetilde{J}_x = \begin{pmatrix} \widetilde{J}^{cc}_x & 0 & 0 \\ 0 & 0 & \widetilde{J}^{su}_x \\ 0 & \widetilde{J}^{us}_x & 0 \end{pmatrix} \ . \tag{4.8}$$

The inverse of the operator $\widetilde{J}_x$ also has the same structure as (4.8).

The key of the construction is that the metric g is globally defined in a neighborhood of the approximately invariant torus and, therefore, so are the operators J and J^{-1}.

We also note that we established in Section 4.5 that the form restricted to the tangent space vanishes for invariant tori (we will see that it is small for approximately invariant tori in many cases). Thus, we obtain that

$$\Omega_{K(\theta)}\left(\text{Range}(DK(\theta)), \text{Range}(DK(\theta))\right) = \left(\widetilde{J}^{cc}_{K(\theta)}\, \text{Range}(DK(\theta)), \text{Range}(DK(\theta))\right)$$

is very small (identically zero for exactly invariant tori). In particular, we obtain that the operator $\widetilde{J}^{cc}_{K(\theta)}$ maps $\text{Range}(DK(\theta))$ into a linearly independent space.

Using that the dimension of the center manifold is $2d$ as in assumption $(H4)$, we obtain that:

$$\text{Range}(DK(\theta)) \oplus \widetilde{J}^{cc}_{K(\theta)}\, \text{Range}(DK(\theta)) = E^c_\theta\ .$$

Since $\widetilde{J}^{cc}_{K(\theta)}$ is a linear operator, we obtain that the center bundle can be expressed as in (4.6). □

4.7. **Automatic reducibility.** A key ingredient in the proof of our main result on whiskered tori is the so–called *automatic reducibility*: in a neighborhood of an invariant torus, one can construct a change of coordinates such that the linearization of the invariance equation (2.7) becomes a constant coefficient equation.

This technique is presented in full detail in [CCdlL13] for conformally symplectic systems (see [dlLGJV05] for symplectic systems), but we will present the details again. An important reason is that, by examing the proof carefully we will discover the surprising global result that the center bundle has to be trivial in the sense of bundle theory.

It will be important to note that there is also a version of approximate reducibility when the torus is only approximately invariant, see Section 4.9. The proof of the results in Section 4.9 will be based on walking through the arguments in this Section and checking how they are affected by the error in the invariance equation.

We will assume that the tangent space of $\mathcal{M}$ at $K(\theta)$, say $\mathcal{T}_{K(\theta)}\mathcal{M}$ with $\theta \in \mathbb{T}^d$, admits an invariant splitting as

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

Taking the derivative of (2.7) we obtain

$$Df_\mu \circ K(\theta)\, DK(\theta) - DK \circ T_\omega(\theta) = 0\ . \tag{4.9}$$

This implies that the range of $DK(\theta)$ is contained in E^c_θ.

Let $\Omega^c_{K(\theta)}$ denote the symplectic form Ω restricted to E^c_θ with

$$\Omega^c_{K(\theta)}(u,v) = (u, J^c_0 \circ K(\theta) v)\ , \quad \forall\, u, v \in E^c_\theta\ ,$$

where J_0^c is the $2d \times 2d$ matrix representing $\Omega^c_{K(\theta)}$ on the center space. Let J^c be the $2n \times 2n$ matrix of the embeddings of the center space into the ambient space.

As indicated above, we have that $\mathrm{Range}(DK(\theta)) \subset E^c_\theta$. Hence, we can write (4.4) as

$$DK^T(\theta) J^c \circ K(\theta) DK(\theta) = 0 \ . \tag{4.10}$$

Let us introduce the $2d \times 2d$ matrix $M(\theta)$ on E^c_θ, obtained juxtaposing the two matrices $DK(\theta)$, $(J^c)^{-1} \circ K(\theta)\ DK(\theta) N(\theta)$:

$$M(\theta) = [DK(\theta) \mid (J^c)^{-1} \circ K(\theta)\ DK(\theta) N(\theta)] \ , \tag{4.11}$$

where we have introduced the normalization factor N as in (3.4). For typographic reasons, we will write

$$v(\theta) = (J^c)^{-1} \circ K(\theta)\ DK(\theta) N(\theta) \ .$$

Note that, because of (4.10) we have that the range of M has dimension $2d$ and, due to our assumption on the dimension of the center, we obtain that

$$\mathrm{Range}(M(\theta)) = E^c_\theta \ . \tag{4.12}$$

Because of (4.12), we know that there exists a matrix $\mathcal{B}(\theta)$ such that

$$Df_\mu \circ K(\theta) M(\theta) = M(\theta + \omega)\ \mathcal{B}(\theta) \ , \tag{4.13}$$

where $\mathcal{B}(\theta)$ is required to be upper triangular with constant matrices on the diagonal.

The goal now is to identify the matrix $\mathcal{B}$. We observe that (4.9) identifies the first column of $\mathcal{B}$ to be $\begin{pmatrix} \mathrm{Id}_d \\ 0 \end{pmatrix}$.

To identify the second column of $\mathcal{B}(\theta)$, by (4.12), we know that

$$Df_\mu \circ K(\theta)\, v(\theta) = DK(\theta + \omega) S(\theta) + v(\theta + \omega) U(\theta) \ , \tag{4.14}$$

for some function $U = U(\theta)$ that we compute as follows. According to [CCdlL13], we multiply (4.14) on the right by $DK^T(\theta + \omega) J^c \circ K(\theta + \omega)$. Using (4.10), we obtain

$$\begin{aligned} DK^T(\theta+\omega) J^c \circ K(\theta+\omega)\, Df_\mu \circ K(\theta)\, v(\theta) &= \\ &= DK^T(\theta+\omega) J^c \circ K(\theta+\omega) (J^c(\theta+\omega))^{-1} DK(\theta+\omega) N(\theta+\omega) U(\theta) \\ &= DK^T(\theta+\omega) DK(\theta+\omega) N(\theta+\omega) U(\theta) \\ &= U(\theta) \ . \end{aligned} \tag{4.15}$$

Working on the other side, using the conformally symplectic property (2.3) and the invariance property of the center foliation, we obtain:

$$Df_\mu^T(x) J^c_{f(x)} Df_\mu(x) = \lambda J^c_{f(x)} \ .$$

Therefore, $J^c_{f(x)} Df_\mu(x)(J^c_{f(x)})^{-1} = \lambda Df_\mu^{-T}(x)$.

Hence, we see that the left hand side of (4.15) can be computed as

$$\begin{aligned} DK^T(\theta+\omega)J^c \circ K(\theta+\omega)\, Df_\mu \circ K(\theta)\, (J^c)^{-1} \circ K(\theta)DK(\theta)N(\theta) \\ = \lambda DK^T(\theta+\omega)Df_\mu^{-T} \circ K(\theta)DK(\theta)N(\theta) \\ = \lambda DK^T(\theta)DK(\theta)N(\theta)\ , \end{aligned}$$

where we have used (4.9).

Therefore, we conclude that $U(\theta) = \lambda\,\mathrm{Id}_d$.

The matrix S can be computed in similar way (it just suffices to multiply in the right to compute the projections): it does not require any change from the calculations in [CCdlL13]. The result is given by (3.6).

In conclusion, we can write (4.13) as

$$Df_\mu \circ K(\theta)M(\theta) = M(\theta+\omega)\begin{pmatrix} \mathrm{Id}_d & S(\theta) \\ 0 & \lambda\,\mathrm{Id}_d \end{pmatrix}\ . \tag{4.16}$$

We note that the average of the matrix $S(\theta)$ computed here is the matrix S appearing in $(H2)$ in Theorem 20. Hypothesis $(H2)$ is just that the average of the matrix S – which is a $d\times d$ matrix – is invertible. Again, we emphasize that this is a condition that is computed out of the approximate solution taking derivatives, performing algebraic operations and taking averages.

4.8. **Consequences of automatic reducibility.** In Section 4.7 we showed that the preservation of the geometric structure yields that we can find a matrix $M(\theta)$ in such a way that

$$M^{-1}(\theta+\omega)\Pi^u_{\theta+\omega}Df_\mu \circ K(\theta)\ \Pi^u_\theta\ M(\theta) = \begin{pmatrix} \mathrm{Id}_d & S(\theta) \\ 0 & \lambda\,\mathrm{Id}_d \end{pmatrix}\ .$$

This shows that we can choose λ_c^-, λ_c^+ as close as desired to $|\lambda|$ (at the price of choosing an appropriate proportionality constant).

For some λ's it is possible to do a further linear change of variables $A(x) = \begin{pmatrix} \mathrm{Id}_d & B(\theta) \\ 0 & \mathrm{Id}_d \end{pmatrix}$ in such a way that the matrix is even simpler. Computing

$$\begin{pmatrix} \mathrm{Id}_d & -B(\theta+\omega) \\ 0 & \mathrm{Id}_d \end{pmatrix}\begin{pmatrix} \mathrm{Id}_d & S(\theta) \\ 0 & \lambda\,\mathrm{Id}_d \end{pmatrix}\begin{pmatrix} \mathrm{Id}_d & B(\theta) \\ 0 & \mathrm{Id}_d \end{pmatrix} = \begin{pmatrix} \mathrm{Id}_d & S(\theta) - \lambda B(\theta+\omega) + B(\theta) \\ 0 & \lambda\,\mathrm{Id}_d \end{pmatrix}\ ,$$

one is led to solve the following equation for B given S:

$$S(\theta) - \lambda B(\theta+\omega) + B(\theta) = 0\ . \tag{4.17}$$

The equation (4.17) can be solved when $|\lambda| \neq 1$ or, for $\lambda \in \mathbb{C}$, when λ is Diophantine with respect to ω. In such a case, we can reduce the cocycle derivative to $\begin{pmatrix} \mathrm{Id}_d & 0 \\ 0 & \lambda\, \mathrm{Id}_d \end{pmatrix}$ and, hence, we can take $\lambda_c^+ = \lambda_c^- = |\lambda|$.

Nevertheless, when λ is close to one (or a root of unity), the B appearing in the last change of variables may be very large.

This means that if we take $\lambda_c^+ = \lambda_c^- = |\lambda|$, in $(H3)$ we can take a very large constant. Geometrically, this means that the center direction (which is a weak center direction) and the tangent span the symplectically complement to the tangent bundle. We can obtain the stable direction by taking the simplectic conjugate and add to it vectors on the stable direction.

The limit of λ close to one appears naturally in many physical problems and will be considered in great detail in Section 7 (see also [CCdlL17]). We note that this is a singular limit because some part of the normal hyperbolicity is lost. It can be controlled precisely because the geometry forces that this loss of hyperbolicity is done in a very specific way.

4.9. **Geometric properties for approximately invariant tori.** Of course, in the iterative procedure, we will not be dealing with invariant tori but with approximately invariant tori. Hence, it will be important for us to show that the geometric identities we developed for invariant tori – notably the automatic reducibility – hold approximately.

The main result of this Section is to show that indeed, this is the case, see Lemma 27. The reason is that, to obtain the main equation (4.13), we just took derivatives of the invariance equation and applied algebraic transformations. Hence, if the invariance equation holds up to an error, we obtain that (4.13) will hold up to errors which can be estimated by derivatives of the error in the invariance equation. A subtle point in the derivation is the use of the isotropic properties of the torus. We will also show that if the torus is approximately invariant, then it has to be approximately isotropic (with quantitative bounds).

As a preliminary result, we recall the following classical lemma which gives the solution of a cohomological equation and which will be needed in the proof of Lemma 27.

Lemma 26. *Let $\lambda \in [A_0, A_0^{-1}]$ for some $0 < A_0 < 1$ and let $\omega \in \mathcal{D}_d(\nu, \tau)$.*

Consider a cohomological equation of the form

$$w(\varphi + \omega) - \lambda w(\varphi) = \eta(\varphi) \tag{4.18}$$

for some functions w and η with $\eta \in \mathcal{A}_\rho$, $\rho > 0$, and with zero average:

$$\int_{\mathbb{T}^d} \eta(\theta)\, d\theta = 0 \ .$$

Then, there is one and only one solution of (4.18) with zero average. Moreover, if $\varphi \in \mathcal{A}_{\rho-\delta}$ for some $0 < \delta < \rho$, then we have

$$\|\varphi\|_{\rho-\delta} \leq C\, \nu\, \delta^{-\tau} \|\eta\|_\rho \ ,$$

where C is a constant that depends on A_0 and the dimension of the space, but it is uniform in λ and it is independent of the Diophantine constant ν.

The proof of Lemma 26 can be found, e.g., in [CCdlL13], see also [Rüs75, Rüs76a, Rüs76b].

Lemma 27. *Consider an approximately invariant torus, satisfying (3.2) where f_{μ_0} is a family of conformally symplectic maps.*

Assume that the cocycle over T_ω given by $\gamma(\theta) = Df_{\mu_0} \circ K(\theta)$ admits an invariant splitting which is hyperbolic and whose center dimension is $2d$ dimensional.

Assume furthermore that for some constant $C > 0$:

$$\begin{aligned} &\|DK\|_\rho, \|N\|_\rho, \|(J^c \circ K)^{-1}\|_\rho \leq C \ , \\ &C\|e\|_\rho \delta^{-1} < 1 \ . \end{aligned} \tag{4.19}$$

Then, defining the matrix M as in (3.5), we have in the center direction

$$Df_{\mu_0} \circ K(\theta) M(\theta) = M(\theta + \omega) \begin{pmatrix} \mathrm{Id}_d & S(\theta) \\ 0 & \lambda\, \mathrm{Id}_d \end{pmatrix} + \mathcal{E}_R \ , \tag{4.20}$$

where

$$||\mathcal{E}_R\|_{\rho-\delta} \leq C\delta^{-1} \|e\|_\rho \ . \tag{4.21}$$

Proof. The proof is basically walking though the proof of the invariant case.

The first step is to study how the approximate invariance modifies the approximately invariance properties.

We note that (2.3) gives that, defining $a(\theta) = DK^T(\theta) J_{K(\theta)} DK(\theta)$, we have:

$$
\begin{aligned}
a(\theta+\omega) - \lambda a(\theta) &= DK^T(\theta+\omega) J_{K(\theta+\omega)} DK(\theta+\omega) - \lambda DK^T(\theta) J_{K(\theta)} DK(\theta) \\
&= \left(DK^T(\theta) Df_{\mu_0}^T \circ K(\theta) - De^T(\theta)\right) J_{K(\theta+\omega)} \left(Df_{\mu_0} \circ K(\theta)\, DK(\theta) - De(\theta)\right) \\
&\quad - \lambda DK^T(\theta) J_{K(\theta)} DK(\theta) \\
&= \left(DK^T(\theta) Df_{\mu_0}^T \circ K(\theta) - De^T(\theta)\right) \left[J_{f_{\mu_0}(K(\theta))} + (J_{K(\theta+\omega)} - J_{f_{\mu_0}(K(\theta))})\right] \\
&\quad \left(Df_{\mu_0} \circ K(\theta)\, DK(\theta) - De(\theta)\right) - \lambda DK^T(\theta) J_{K(\theta)} DK(\theta) \\
&= DK^T(\theta) Df_{\mu_0}^T \circ K(\theta) J_{f_{\mu_0}(K(\theta)} Df_{\mu_0} \circ K(\theta) DK(\theta) \\
&\quad - \lambda DK^T(\theta) J_{K(\theta)} DK(\theta) + e_I(\theta) \\
&= e_I(\theta) \ ,
\end{aligned}
\tag{4.22}
$$

where the expression for e_I is just products of derivatives of the invariance equation (with other terms).

Hence, we can bound e_I by using the Cauchy estimates for De and the smallness assumptions and obtain

$$\|e_I\|_{\rho-\delta} \le C\delta^{-1} \|e\|_\rho \ .$$

We note that (4.22) shows that a satisfies a cohomology equation of the form considered in Lemma 26. Therefore, we can obtain estimates on a as $\|a\|_{\rho-2\delta} \le C\delta^{-\tau}\ \nu\ \|e\|_\rho$.

Again, we can interpret the estimates on a as approximate orthogonality relations.

We now walk through the calculations used in the computation leading to (4.16).

The first column in (4.16) is just the derivative of the invariance equation and we can use Cauchy estimates to get the estimates (4.21) for the first column.

To study the second column, we see that (4.14) is still true, since it only depends on the property that $\mathrm{Range}(DK(\theta+\omega)), \mathrm{Range}(v(\theta+\omega))$ span $E^c_{\theta+\omega}$, which is an easy consequence of the approximate orthogonality.

As before, we multiply (4.14) by $DK^T(\theta+\omega) J^c \circ K(\theta+\omega)$. We can follow the calculations used in (4.15) adding and subtracting the terms that we have estimated. □

5. Proof of Theorem 20

We now proceed to the proof of Theorem 20. As standard in KAM theorem in an a-posteriori format, the proof can be divided into two parts. The first one is an algorithm that, given an approximate solution, produces a much more approximate one (in a slightly weaker sense of approximation) and with slightly worse quality properties. In a second part, we show that, if we start with a sufficiently small error, we can repeat the procedure indefinitely and that the solution indeed converges in some appropriate sense, and that the limit inherits properties such as the hyperbolicity and twist. See the following Proposition 28 in Section 5.1.

The iterative algorithm will be discussed in Section 5.1. At the end of the step, the error will be roughly the square of the original error, but measured in a norm corresponding to functions in a smaller domain than that of the original approximation (there is also a factor depending on the loss of domain). This iterative algorithm will be affected by condition numbers (hyperbolicity properties, twist, etc.) and we need to estimate how do they deteriorate.

We note that the step we will discuss in this paper will be numerically very efficient. It does not require that the system is close to integrable, it only requires to handle functions of the dimension of the torus, the storage requirement is small and the operation count is small.

We anticipate that to be able to carry the step and obtain estimates, we will need to introduce *inductive assumptions*. One – very standard in KAM theory – requires that Δ, the correction to K, is small enough in its domain so that the range of $(K+\Delta)$ is well inside the domain of f (so that we can define $f \circ (K+\Delta)$ and study Taylor expansions in Δ). In our case, we will also need another inductive assumption that guarantees that the hyperbolicity constants are still bounded. As it is well known, once we fix the domain loss, the first inductive assumption can be guaranteed by the requirement that the error is small enough (so that the correction Δ is small enough). As for the assumption on the uniformity of the hyperbolicity, using Lemma 15, it will amount to the block diagonal cocycles remaining in a neighborhood.

The iterative process is discussed in Section 5.1.8. The main, well known, idea is that if we fix a sequence of domain losses that goes to zero not too fast (e.g. exponentially fast), if the original error of the invariance is small enough, the error of the invariance decreases very fast in the iterative step so that the ranges of the K do not get close to the boundary of the domain of f and the block diagonal cocycles do not move out of the neighborhood specified in Lemma 15.

5.1. **Results on the iterative step.** The main result of this Section is Proposition 28 that specifies how, given an approximate solution with some non-degeneracy properties (if some quantitative assumptions are satisfied), we produce a more approximate one with only slightly worse non-degeneracy assumptions. The quantitative assumptions, that allow to perform the step, are standardly called *inductive assumptions*, since we will use an inductive argument to show that they can be satisfied for all steps of the iteration.

Proposition 28. *Let $\omega \in \mathcal{D}_d(\nu,\tau)$, $d \le n$, and let $f_\mu : \mathcal{M} \to \mathcal{M}$, $\mu \in \mathbb{R}^d$, be a family of real-analytic, conformally symplectic maps as in Theorem 20 with $0 < \lambda < 1$.*

Let (K,μ), $K : \mathbb{T}^d \to \mathcal{M}$, $K \in \mathcal{A}_\rho$, be an approximate solution of the invariance equation (2.6):

$$f_\mu \circ K(\theta) - K \circ T_\omega(\theta) = e(\theta) \tag{5.1}$$

for some function $e = e(\theta)$. Denote $\mathcal{E} = \|e\|_\rho$.

Let $E^s_\theta \oplus E^c_\theta \oplus E^u_\theta$ be an approximately invariant, hyperbolic splitting based on K. Denote by $\mathcal{E}_h$ the quantity appearing on (2.21).

Assume that $(K, \mu, E^s_\theta, E^c_\theta, E^s_\theta)$ satisfy assumptions $(H2)$-$(H3)$-$(H3')$-$(H4)$-$(H5)$ of Theorem 20.

We will assume that $\mathcal{E}, \mathcal{E}_h$ are sufficiently small depending on the quantities in (3.8) and on η with η as in $(H1)$. The constants C denote expressions depending only the quantities in (3.8), and the formulas for C will be made explicit along the proof.

Then, we have the following results.

1) There exists an exact invariant splitting based on K, say $\widetilde{E}^{s/c/u}_\theta$. Denote by $\tilde{\gamma}^{\sigma,\sigma}_\theta$ the projections of $Df_\mu \circ K(\theta)$ corresponding to the invariant splitting. Then, we have:

$$\begin{aligned} \mathrm{dist}_\rho(E^{s/c/u}_\theta, \widetilde{E}^{s/c/u}_\theta) &\le C\mathcal{E}_h \ , \\ \|\gamma^{\sigma,\sigma}_\theta - \tilde{\gamma}^{\sigma,\sigma}_\theta\|_\rho &\le C\mathcal{E}_h \ . \end{aligned} \tag{5.2}$$

2) Assume that δ is such that

$$C\delta^{-\tau}\|e\|_\rho + \mathrm{dist}(K(\mathbb{T}^d_\rho)), \mathbb{C}^{2n} \backslash \mathcal{D}) > \eta/2 \ . \tag{5.3}$$

Then, we have that $K' = K + \Delta$, $\mu' = \mu + \beta$ for suitable corrections Δ, β, satisfy

$$f_{\mu'} \circ K'(\theta) - K' \circ T_\omega(\theta) = e'(\theta)$$

with

$$\|e'\|_{\rho-\delta} \le C\ \delta^{-2\tau}\ \|e\|^2_\rho \ .$$

Moreover, the corrections can be bounded as

$$\begin{aligned} &\|\Delta\|_{\rho-\delta} \le C\delta^{-\tau}\mathcal{E} \\ &\|D\Delta\|_{\rho-\delta} \le C\delta^{-1-\tau}\mathcal{E} \\ &|\beta| \le C\mathcal{E} \ . \end{aligned} \tag{5.4}$$

3) Furthermore, the splitting $\widetilde{E}^{s/c/u}_\theta$ is approximately invariant for $Df_{\mu+\beta} \circ (K + \Delta)$.

3.1) The error in the change of the invariance is smaller than $C\delta^{-\tau}\mathcal{E}$.

3.2) The block diagonal cocycles corresponding to $Df_{\mu+\beta}\circ(K+\Delta)$ (which we denote by $\hat{\gamma}_\theta^{\sigma,\sigma}$) satisfy

$$\|\hat{\gamma}_\theta^{\sigma,\sigma}-\tilde{\gamma}_\theta^{\sigma,\sigma}\|_{\rho-\delta}\leq C\delta^{-\tau}\|e\|_\rho+\mathcal{E}_h\ .$$

5.1.1. *Overview of the argument.* We look for a correction (Δ,β), such that $K'=K+\Delta$, $\mu'=\mu+\beta$ satisfy (5.1) with an error quadratically smaller.

Expanding the composition to first order in Δ,β we obtain:

$$\begin{aligned}f_{\mu'}\circ K'(\theta)-K'(\theta+\omega)\ &=\ f_\mu\circ K(\theta)+Df_\mu\circ K(\theta)\,\Delta(\theta)\\&\quad+D_\mu f_\mu\circ K(\theta)\,\beta\\&\quad-K(\theta+\omega)-\Delta(\theta+\omega)\\&\quad+O(\|\Delta\|^2)+O(|\beta|^2)\ .\end{aligned}$$

Taking into account (5.1), the new error is quadratically smaller if (Δ,β) satisfies

$$Df_\mu\circ K(\theta)\ \Delta(\theta)+D_\mu f_\mu\circ K(\theta)\ \beta-\Delta(\theta+\omega)=-e(\theta) \tag{5.5}$$

or, more generally, if

$$Df_\mu\circ K(\theta)\ \Delta(\theta)+D_\mu f_\mu\circ K(\theta)\ \beta-\Delta(\theta+\omega)+e(\theta)=O(\|e\|_\rho^2)\ . \tag{5.6}$$

Finding corrections by solving (5.5) can be thought of as an infinite dimensional version of Newton method. The small modification (5.6) is called a quasi-Newton method, since as we will see, the errors are also reduced quadratically.

To establish part 1) in Proposition 28, we just invoke the Lemma 14 to change the assumed approximately invariant splitting into an exactly invariant splitting.

We project (5.5) on the hyperbolic and center spaces, using the invariant splitting (2.13).

Due to the invariance of the splitting, we have that

$$\Pi_{\theta+\omega}^{s/c/u}\Big(Df_\mu\circ K(\theta)\ \Delta(\theta)\Big)=Df_\mu\circ K(\theta)\ \Delta^{s/c/u}(\theta)\ ,$$

where $\Delta^{s/c/u}(\theta)\equiv\Pi_{\theta+\omega}^{s/c/u}\Delta(\theta)$. Therefore, (5.5) is equivalent to the three equations:

$$Df_\mu\circ K(\theta)\ \Delta^{s/c/u}(\theta)+\Pi_{\theta+\omega}^{s/c/u}D_\mu f_\mu\circ K(\theta)\beta-\Delta^{s/c/u}(\theta+\omega)=-e^{s/c/u}(\theta)\ , \tag{5.7}$$

where we have defined $e^{s/c/u}(\theta)\equiv\Pi_{\theta+\omega}^{s/c/u}e(\theta)$.

Note that in the three linear equations (5.7), we have four unknowns given by Δ^s, Δ^c, Δ^u and β, which we solve using a substitution method. We will first solve (approximately) the equation for Δ^c, which will determine both Δ^c and β.

Then, we will solve the equations for Δ^s, Δ^u.

It is convenient for us to start by solving the equation in the center space since it is the equation that allows to determine the β, which also enters in the equations in the stable/unstable directions. The alternative of solving the stable/unstable equations with a floating parameter seems more cumbersome.

The equation in the center can be solved approximately, using the automatic reducibility established in Section 2.4. Note that the automatic reducibility depends on the geometry. We will also use a non–degeneracy condition (as in $(H5)$); from the analysis point of view, it is the most delicate equation since it involves small divisors (here we use the assumption that the frequency is Diophantine) and it entails a loss of domain. In contrast, the equations along the stable/unstable directions can be solved by soft methods (iteration and contraction) and do not involve any loss of domain.

Along the argument, we will make some side remarks about an efficient numerical implementation, which just need to implement the correction step. Of course, to be convincing one needs to monitor also the condition numbers (which we make explicit) to ensure that the numerical solutions produced correspond to the true ones. We anticipate that all the operations required are algebraic operations on the approximate solutions, taking derivatives, shifting and solving cohomology equations. If we discretize the parameterization with N terms, a Newton step requires $O(N \ln(N))$ operations, either in a grid discretization or in a Fourier discretization. Of course, one can generate a grid from a Fourier series by using FFT. The procedure gets quadratic convergence, but does not need to store (much less solve) an $N \times N$ matrix. Indeed, the storage required only $O(N)$.

Once we have the estimates on the correction, the nonlinear estimates for the error can be obtained by elementary methods such as adding and subtracting terms and applying Taylor's theorem to first order (the corrections have been chosen precisely to cancel out the first order approximation).

5.1.2. *Approximate solution of the linearized invariance equation in the center space.* As for the center subspace, we will construct an approximate solution of the Newton equation in the center direction. That is, we will construct a function that solves the projection of the linearized equation up to a quadratically small error (which does not affect the quadratic convergence of the method). The construction of this approximate solution, which will take the rest of the Section, is somewhat subtle, since it requires taking advantage of some geometric properties and of the Diophantine properties of the frequency. This solution will also involve a loss of domain and we will obtain estimates for the correction only in a slightly smaller domain than the domain of the error.

The linearized equation (5.5) on the center subspace is

$$Df_\mu \circ K(\theta)\ \Delta^c(\theta) + \Pi^c_{\theta+\omega} D_\mu f_\mu \circ K(\theta)\beta - \Delta^c(\theta+\omega) = -e^c(\theta)\ . \tag{5.8}$$

We will take advantage of the geometry, see Section 4.9, to find an explicit linear change of variables that approximately reduces the equation in (5.8) to constant coefficients difference equations. These equations can be solved using Fourier coefficients (but they need the Diophantine conditions on the frequency). Let us introduce W^c such that we can write Δ^c as

$$\Delta^c = M\ W^c$$

with M as in (3.5) and satisfying (4.20).

Using (5.8) and (4.20), we obtain that the Newton equation projected in the center is equivalent to:

$$\begin{aligned} M(\theta+\omega) \begin{pmatrix} \mathrm{Id}_d & S(\theta) \\ 0 & \lambda\,\mathrm{Id}_d \end{pmatrix} W^c(\theta) - M(\theta+\omega)\ W^c(\theta+\omega) \\ + \mathcal{E}_R(\theta) W^c(\theta) + \Pi^c_{\theta+\omega} D_\mu f_\mu \circ K(\theta)\beta = -e^c(\theta)\ . \end{aligned}$$

Since the above equation is hard to solve, we argue heuristically that the term $\mathcal{E}_R$ is comparable to $\mathcal{E}$ because of (4.21), hence the term $\mathcal{E}_R W^c$ is second order. Hence, we omit it and consider the following equation (5.9). As we will see, the equation (5.9) is readily solvable, admits tame estimates and, indeed, omitting the term $\mathcal{E}_R W^c$ does not change the fact that the error remaining after the iterative step is quadratic in the original error (we will obtain estimates for the W^c and we have estimates for $\mathcal{E}_R$ in (4.21)):

$$M(\theta+\omega) \begin{pmatrix} \mathrm{Id}_d & S(\theta) \\ 0 & \lambda\,\mathrm{Id}_d \end{pmatrix} W^c(\theta) - M(\theta+\omega)\ W^c(\theta+\omega) + \Pi^c_{\theta+\omega} D_\mu f_\mu \circ K(\theta)\beta = -e^c(\theta)\ . \tag{5.9}$$

Multiplying (5.9) on the left by $M^{-1}(\theta+\omega)$, we obtain:

$$\begin{pmatrix} \mathrm{Id}_d & S(\theta) \\ 0 & \lambda\,\mathrm{Id}_d \end{pmatrix} W^c(\theta) - W^c \circ T_\omega(\theta) = -\widetilde{e}^c(\theta) - \widetilde{A}^c(\theta)\beta\ , \tag{5.10}$$

where $\widetilde{e}^c(\theta) \equiv M^{-1} \circ T_\omega(\theta) e^c(\theta)$, $\widetilde{A}^c(\theta) \equiv M^{-1} \circ T_\omega(\theta)\ \Pi^c_{\theta+\omega} D_\mu f_\mu \circ K(\theta)$.

Writing (5.10) in components we obtain

$$\begin{aligned} W_1^c(\theta) - W_1^c \circ T_\omega(\theta) &= -S(\theta)\, W_2^c(\theta) - \widetilde{e}_1^c(\theta) - \widetilde{A}_1^c(\theta)\beta \\ \lambda W_2^c(\theta) - W_2^c \circ T_\omega(\theta) &= -\widetilde{e}_2^c(\theta) - \widetilde{A}_2^c(\theta)\beta\ , \end{aligned} \tag{5.11}$$

where $\widetilde{A}^c \equiv [\widetilde{A}^c_1|\widetilde{A}^c_2]$ with $\widetilde{A}^c_1$, $\widetilde{A}^c_2$ denoting the first d and the last d rows of the $2d \times d$ matrix $\widetilde{A}^c$.

Denoting by $\overline{W^c}$ the average of W^c and setting $(W^c)^0 \equiv W^c - \overline{W^c}$, we obtain

$$\begin{aligned}
(W^c_1)^0(\theta) - (W^c_1)^0 \circ T_\omega(\theta) &= -(SW^c_2)^0(\theta) - (\widetilde{e}^c_1)^0(\theta) - (\widetilde{A}^c_1)^0(\theta)\beta \\
\lambda(W^c_2)^0(\theta) - (W^c_2)^0 \circ T_\omega(\theta) &= -(\widetilde{e}^c_2)^0(\theta) - (\widetilde{A}^c_2)^0(\theta)\beta \ .
\end{aligned} \tag{5.12}$$

Since $(W^c_2)^0$ is an affine function of β, we can write $(W^c_2)^0 = (W^c_a)^0 + \beta(W^c_b)^0$ for some functions W^c_a, W^c_b. Therefore, the second equation in (5.12) can be split as

$$\begin{aligned}
\lambda(W^c_a)^0(\theta) - (W^c_a)^0 \circ T_\omega(\theta) &= -(\widetilde{e}^c_2)^0(\theta) \\
\lambda(W^c_b)^0(\theta) - (W^c_b)^0 \circ T_\omega(\theta) &= -(\widetilde{A}^c_2)^0(\theta) \ .
\end{aligned} \tag{5.13}$$

On the other hand, taking the average of (5.11) we obtain

$$\begin{aligned}
&\overline{S}\,\overline{W^c_2} + (\overline{S(W^c_b)^0} + \overline{\widetilde{A}^c_1})\beta = -\overline{S(W^c_a)^0} - \overline{\widetilde{e}^c_1} \\
&(\lambda - 1)\overline{W^c_2} + \overline{\widetilde{A}^c_2}\beta = -\overline{\widetilde{e}^c_2} \ .
\end{aligned} \tag{5.14}$$

Provided that the non–degeneracy condition (3.7) in $(H5)$ is satisfied, equations (5.14) yield $\overline{W^c_2}$ and β as the solution of the finite dimensional system:

$$\begin{pmatrix} \overline{S} & \overline{S(W^c_b)^0} + \overline{\widetilde{A}^c_1} \\ (\lambda-1)\,\mathrm{Id}_d & \overline{\widetilde{A}^c_2} \end{pmatrix} \begin{pmatrix} \overline{W^c_2} \\ \beta \end{pmatrix} = \begin{pmatrix} -\overline{S(W^c_a)^0} - \overline{\widetilde{e}^c_1} \\ -\overline{\widetilde{e}^c_2} \end{pmatrix} \ .$$

Using Lemma 26, we can solve (5.13) to get $(W^c_a)^0$, $(W^c_b)^0$, which provide $(W^c_2)^0$ and finally we solve the first of (5.12) in order to compute $(W^c_1)^0$. This yields the solution of (5.10), which allows to find the correction (Δ^c, β) on the center subspace.

Note that this correction (Δ^c, β) does not eliminate completely the error in the center direction, but reduces it to $W^c\mathcal{E}_R$.

Note that we obtain W^c_2 solving a small divisor equation and then we obtain β, W^c_1 by performing algebraic equations and applying a contraction argument, leading to the following estimates:

$$\|W^c\|_{\rho-\delta} \le C\delta^{-\tau}\|e\|_\rho \ , \tag{5.15}$$

$$|\beta| \le C\|e\|_\rho \ . \tag{5.16}$$

Then, Δ^c is obtained multiplying M and W^c and, hence, under the assumption that $\|M\|_{\rho-\delta}$ (which we will prove holds inductively) is bounded by a constant, we obtain that:

$$\|\Delta^c\|_{\rho-\delta} \le C\delta^{-\tau}\|e\|_\rho \ . \tag{5.17}$$

5.1.3. *Uniqueness properties of the approximate solutions in the center direction.* It will be important for future studies (e.g., for the local uniqueness in Theorem 34) to note that the W^c, out of which we construct Δ^c, is obtained applying Lemma 26.

We observe that, by following the procedure indicated in the previous paragraph, we obtain that the W^c solving the equation (5.12) is unique up to adding a constant to W_1^c. The estimates claimed in (5.15) correspond to taking W_1^c with zero average.

This lack of uniqueness of the corrections has a geometric interpretation related to the underdetermination of (2.7) remarked at the beginning of Section 2.4.1. Since $\Delta^c = MW^c$, with M introduced in (4.11), we have that adding a constant σ to W^c is tantamount to adding to Δ^c the quantity $DK\sigma$. That is, we are changing the corrections by adding to them a movement along the space of solutions.

Note that the above statement of uniqueness refers only to the solutions of (5.18). If our goal was to improve the accuracy of the solutions, we have a flexibility of changing the Δ^c by other terms that are much smaller than the error.

5.1.4. *Solutions of the linearized equation in the hyperbolic directions.* Let us now consider the stable subspace. Let $\theta' = T_\omega(\theta)$, so that we can write (5.7) as

$$Df_\mu(K \circ T_{-\omega}(\theta'))\ \Delta^s(T_{-\omega}(\theta')) + \Pi^s_{\theta'} D_\mu f_\mu(K \circ T_{-\omega}(\theta'))\beta - \Delta^s(\theta') = -\widetilde{e}^s(\theta')\ , \tag{5.18}$$

where

$$\widetilde{e}^s(\theta') \equiv \Pi^s_{\theta'} e \circ T_{-\omega}(\theta')\ .$$

We proceed now to solve (5.18) for the stable subspace. For any β, we can write

$$\Delta^s(\theta') = \widetilde{e}^s(\theta') + Df_\mu(K \circ T_{-\omega}(\theta'))\ \Delta^s(T_{-\omega}(\theta')) + \Pi^s_{\theta'} D_\mu f_\mu(K \circ T_{-\omega}(\theta'))\beta\ , \tag{5.19}$$

which leads to the solution for Δ^s in the form

$$\begin{aligned}
\Delta^s(\theta') \;&=\; \widetilde{e}^s(\theta') + \sum_{k=1}^{\infty} \Big(Df_\mu(K \circ T_{-\omega}(\theta')) \times \cdots \times Df_\mu(K \circ T_{-k\omega}(\theta')) \Big)\ \widetilde{e}^s(T_{-k\omega}(\theta')) \\
&+\; \Pi^s_{\theta'} D_\mu f_\mu(K \circ T_{-\omega}(\theta'))\beta \\
&+\; \sum_{k=1}^{\infty} \Big(Df_\mu(K \circ T_{-\omega}(\theta')) \times \ldots \times Df_\mu(K \circ T_{-k\omega}(\theta'))\ \Pi^s_{\theta'} D_\mu f_\mu(K \circ T_{-(k+1)\omega}(\theta')) \Big)\ \beta\ .
\end{aligned} \tag{5.20}$$

By the variations of parameters formula, we guess a solution of the form (5.20). Then, we observe that the series in (5.20) converges uniformly in $\mathcal{A}_\rho$, because of the bounds assumed in (2.14). Hence, we can substitute (5.20) in (5.19) and rearrange the terms, so that we can verify that (5.19) is satisfied by (5.20).

Due to the growth conditions (2.14) and due to the fact that $D_\mu f_\mu$ is a bounded operator, we obtain that

$$\|\Delta^s\|_\rho \le C \left(\|\widetilde{e}^s\|_\rho + |\beta| \ \|\Pi^s_{\theta+\omega} D_\mu f_\mu \circ K\|_\rho\right) \cdot \sum_{k=0}^{\infty} \lambda_-^k \tag{5.21}$$

for some constant $C > 0$.

Concerning the unstable subspace, from (5.7) we can write in a similar way for any β:

$$\Delta^u(\theta) = (Df_\mu)^{-1}(K(\theta)) \ [-e^u(\theta) - \Pi^u_{\theta+\omega} D_\mu f_\mu(K(\theta))\beta + \Delta^u(T_\omega(\theta))] \ ,$$

which leads to the expression

$$\begin{aligned}\Delta^u(\theta) \ &= \ -\sum_{k=0}^{\infty} \Big((Df_\mu)^{-1}(K(\theta)) \times ... \times (Df_\mu)^{-1}(K \circ T_{k\omega}(\theta))\Big) \ e^u(T_{k\omega}(\theta)) \\ &- \ \sum_{k=0}^{\infty} \Big((Df_\mu)^{-1}(K(\theta)) \times ... \times (Df_\mu)^{-1}(K \circ T_{k\omega}(\theta)) \ \Pi^u_{\theta+\omega} D_\mu f_\mu(K \circ T_{k\omega}(\theta)\Big)\beta \ .\end{aligned} \tag{5.22}$$

From (5.22) and (2.14) we obtain:

$$\|\Delta^u\|_\rho \le C \left(\|\widetilde{e}^u\|_\rho + |\beta| \ \|\Pi^u_{\theta+\omega} D_\mu f_\mu \circ K\|_\rho\right) \cdot \sum_{k=0}^{\infty} \lambda_+^k \tag{5.23}$$

for some constant $C > 0$. Expressions (5.21) and (5.23) lead to

$$\|\Delta^s\|_\rho \le C\|e\|_\rho \ , \qquad \|\Delta^u\|_\rho \le C\|e\|_\rho \ .$$

Remark 29. *In numerical implementations, it may be advantagous to solve* (5.18) *by adding an equation that makes the cocycle diagonal. The effort required is not that much, since the linearizing transformation can be computed iteratively. In this case, we can use the Fourier methods to compute the solutions (see* [HdlL06b]*).*

5.1.5. *Nonlinear estimates for the step.* Notice that, in all the cases above, we have obtained estimates for the corrections in terms of the error. The most delicate ones are those in the center direction since they involve the small divisors.

We have that for all $0 < \delta < \rho$, the corrections satisfy (5.4), namely:

$$\begin{aligned}&\|\Delta\|_{\rho-\delta} \le C\delta^{-\tau}\|e\|_\rho \\ &\|D\Delta\|_{\rho-\delta} \le C\delta^{-\tau-1}\|e\|_\rho \\ &|\beta| \le C\|e\|_\rho \ .\end{aligned} \tag{5.24}$$

The first line of (5.24) is wasteful in the stable and unstable directions. The second line can be obtained from the first line estimates for $\|\Delta\|_{\rho-\frac{1}{2}\delta}$ and then, using Cauchy estimates for the derivative.

5.1.6. *Nonlinear estimates.* Let us now conclude the proof of part 2) of Proposition 28.

We see that, under the assumption (5.3), we have that we can define the composition $f_{\mu+\beta}(K+\Delta)$ and that $(K+\Delta)(\mathbb{T}^d_{\rho-\delta})$ is away from the boundary of the domain of the function f.

We can write the error of $K' = K + \Delta$, $\mu' = \mu + \beta$ as

$$\begin{aligned} f_{\mu+\beta} \circ (K+\Delta) - (K+\Delta) \circ T_\omega &= f_\mu \circ K + Df_\mu \circ K\Delta + D_\mu f_\mu \circ K\beta + \mathcal{E}_T \\ &\quad - K \circ T_\omega - \Delta \circ T_\omega \\ &= \mathcal{E}_T + \mathcal{E}_R W \ , \end{aligned}$$

where $\mathcal{E}_T$ is the Taylor reminder of the expansion of the composition and we note that the correction Δ has been chosen precisely to cancel the error. The Taylor estimates are elementary

$$\|\mathcal{E}_T\|_{\rho-\delta} \le C(\|\Delta\|_\rho^2 + |\beta|^2) \le C\delta^{-2\tau}\|e\|_\rho^2$$

and the estimates for $\mathcal{E}_R$ – the error in the approximate reducibility – are in (4.21), while the estimates for W can be obtained from (5.24).

Remark 30. *The algorithm we have implemented uses an invariant splitting to solve the linearized equation. Of course, after the step, we need to apply the closing Lemma 14 to obtain a new invariant splitting. The algorithm we presented is very efficient from the theoretical point of view. Nevertheless in a numerical implementation, it is wasteful to obtain an exact splitting at each step. In numerical implementations it suffices to obtain invariant splittings up to an error comparable to the error in the invariance equation. Refining the splittings to more accuracy does not lead to significant improvements of the step.*

Hence, in numerical implementations, it would be advantageous to develop a Newton method for the invariant splittings and for the invariance equation, simultaneously.

A rigorous presentation and estimates for the Newton method for the invariance of the torus, and the splittings (which also allows some elliptic directions) can be found in [DV17].

5.1.7. *Change in the non-degeneracy conditions after an iterative step.* Now, we estimate the changes on the non-degeneracy conditions induced by the change of the μ, K in the iterative step. Since we start with an invariant splitting, we have that $\gamma_\theta^{\sigma,\sigma'} = 0$. Hence, denoting by $\hat{\gamma}_\theta^{\sigma,\sigma'}$ the cocycle associated to $K' = K + \Delta$, $\mu' = \mu + \beta$, then

$$\|\hat{\gamma}_\theta^{\sigma,\sigma'} - \gamma_\theta^{\sigma,\sigma'}\|_\rho < C(\|D^2 f_\mu \circ K\|_\rho \, \|\Delta\|_\rho + \|D_\mu D f_\mu \circ K\|_\rho \, |\beta|) < C\delta^{-\tau}\mathcal{E} \ ,$$

thus yielding 3.1 of Theorem 28.

The results of Proposition 31 below give part 3.2 of Proposition 28, beside implying the estimate (3.10) of Theorem 20.

Proposition 31. *With the notations and assumptions of Proposition 28, we have the following bounds.*

(i) *Let $\mathcal{S}'$ be the quantity $\mathcal{S}$ in* (3.7) *after one iterative step, namely with K, μ replaced by $K' = K + \Delta$, $\mu' = \mu + \beta$ as in Proposition 28. Let $0 < \delta < \rho$; then:*

$$\|\mathcal{S}'\|_{\rho-\delta} \leq \|\mathcal{S}\|_\rho + C\delta^{-\tau}\|e\|_\rho \tag{5.25}$$

for a suitable constant $C > 0$.

(ii) *The changes in the projections are bounded as*

$$\|\widetilde{\Pi}_\theta^{s/c/u} - \Pi_\theta^{s/c/u}\|_{\rho-\delta} \leq C\|K' - K\|_\rho \leq C\delta^{-\tau}\|e\|_\rho \ . \tag{5.26}$$

(iii) *The changes in the diagonal cocycles are bounded by*

$$\|\hat{\gamma}_\theta^{\sigma,\sigma} - \tilde{\gamma}_\theta^{\sigma,\sigma}\|_{\rho-\delta} \leq C(\delta^{-\tau}\mathcal{E} + \mathcal{E}_h) \ . \tag{5.27}$$

Proof. We note that the matrix $\mathcal{S}$ is an algebraic expression of the derivatives of K and the derivatives of f_μ evaluated at K. The projections of the above quantities are taken on the center directions. We can estimate all the changes by adding and subtracting, so that we get only terms in which one changes. The derivatives of K change by the derivatives of Δ, which can be bounded using Cauchy estimates from the estimates for Δ. Hence, the change of these terms can be bounded by $C\delta^{-\tau}\|e\|_\rho$, thus leading to (5.25).

The change in the projections is also estimated already and come from the application of Lemma 14, thus leading to (5.26).

The changes of the term $Df_{\mu+\beta}$ can be bounded by the size of β and the size of Δ multiplied by the second derivatives of f with respect to its arguments. The terms change by quantities that are smaller than the previous one. The estimate in (5.27) is bounded by the sum of the norms of $D^2 f_\mu \circ K \, \Delta$ and $D_\mu D f_\mu \circ K \, \beta$ and using (5.2). $\square$

5.1.8. *Iteration of the iterative step.* It is a classical argument in KAM theory that, if the initial error is small enough, the inductive step can be iterated indefinitely and that it converges to the true solution. We can also estimate the difference between the limit and the initial approximate solutions.

Remark 32. *It is important to note that assumptions of smallness in Lemma 15 are independent of the ρ considered. In the applications, we will be considering a sequence of corrections in a sequence of decreasing domains. We will just need that the norm of the corrections (in the smallest domain) are sufficiently small.*

Remark 33. *If we assume that*

$$\|DK_0 - DK_j\|_{\rho_j} \leq \alpha$$

for some $\alpha > 0$ we obtain that (3.7) *and that the $\|(DK_j^{\mathcal{T}} DK_j)^{-1}\|_{\rho_j}$ are bounded by a value slightly bigger than the original one, say twice.*

Similarly, if $\|K_0 - K_j\|_{\rho_j}$ are sufficiently small, we can ensure that the range of K_j is inside the domain of f_μ and at a distance bigger than $\eta/2$ from the boundary of the domain.

Provided that during the iteration the DK does not leave the neighborhood more than the distance α, we can use the bounds corresponding to the chosen values of the condition numbers.

Recalling the definition of M in (4.11)*, we notice that $M_j - M_0$ is an algebraic expression of DK_j and DK_0. By the mean value theorem, $M_j - M_0$ is bounded by $DK_j - DK_0$ and hence it is small, if DK_j is close to DK_0.*

We will just repeat the standard argument. We start by fixing the sequences of domains where we will be doing estimates. Let δ_j, ρ_{j+1} be defined as

$$\delta_j = \frac{\delta_0}{2^{j+2}} \ , \qquad \rho_{j+1} = \rho_j - \delta_j \ .$$

We take δ_0 to be $1/2$ of the total loss of domain that we allow in the conclusions of Theorem 20. The ρ_j will be the domains where we will carry estimates in the j-th step.

We recall that to perform the iterative step, we need to make two inductive assumptions that ensure that we can define the composition. We will also need to assume that we are in the region when all the hyperbolicity constants are uniform (slightly worse than those in the original problem) and we will also assume that other non-degeneracy conditions are satisfied.

The important thing to observe is that, if we can carry out j steps, the conditions for being able to carry out the next step and remaining in the neighborhood are implied by

$$\begin{aligned} \|K_0 - K_j\|_{\rho_j} &\leq \alpha \ , \\ \|DK_0 - DK_j\|_{\rho_j} &\leq \alpha \end{aligned} \tag{5.28}$$

for some $\alpha > 0$, which is independent of j, see Remark 32 and Remark 33.

We will assume that for j steps, we have been able to carry out the step and remain in the set of analytic functions where we have (5.28) and, hence, we can perform the iteration. We will show that, if the initial error is small enough, the error has decreased so much after the j-th step that we can apply again the result. Note that this is very similar to the estimates that one carries out in the elementary Newton Method.

Denoting by $\varepsilon_h \equiv \|e_h\|_{\rho_h}$, we have:

$$\begin{aligned} \varepsilon_h &\leq C\delta_h^{-\tau}\varepsilon_{h-1}^2 \\ &= (C\delta_0^{-\tau}2^{3\tau})\ 2^{(h-1)\tau}\ \varepsilon_{h-1}^2 \\ &\leq (C\delta_0^{-\tau}2^{3\tau})^{1+2+\ldots+2^{h-1}}\ 2^{\tau\ [(h-1)+2(h-2)+\ldots+2^{h-2}]}\ \varepsilon_0^{2^h} \\ &\leq (C\delta_0^{-\tau}2^{3\tau}\,2^{\tau}\varepsilon_0)^{2^h-1}\ \varepsilon_0 \\ &= (\kappa_0\varepsilon_0)^{2^h-1}\ \varepsilon_0\ , \end{aligned}$$

where

$$\kappa_0 = C\delta_0^{-\tau}\,2^{4\tau}\ .$$

Starting from $K_h = K_{h-1} + \Delta_{h-1}$, a bound on $K_h - K_0$ and its derivative is obtained as follows:

$$\begin{aligned} \|K_h - K_0\|_{\rho_h} &= \|\sum_{j=0}^{h}\Delta_j\|_{\rho_h} \leq \sum_{j=0}^{h}\|\Delta_j\|_{\rho_h} \\ &\leq C\ \sum_{j=0}^{h}\delta_0^{-\tau}2^{\tau(j+2)}\ (\kappa_0\varepsilon_0)^{2^j-1}\varepsilon_0\ , \end{aligned}$$

$$\begin{aligned} \|DK_h - DK_0\|_{\rho_h} &= \|\sum_{j=0}^{h}D\Delta_j\|_{\rho_h} \leq \sum_{j=0}^{h}\|D\Delta_j\|_{\rho_h} \\ &\leq C\ \sum_{j=0}^{h}\delta_0^{-\tau-1}2^{\tau(j+2)+1}\ (\kappa_0\varepsilon_0)^{2^j-1}\varepsilon_0\ , \end{aligned}$$

$$
\begin{aligned}
|\mu_h - \mu_0| &= |\sum_{j=0}^{h} \beta_j| \le \sum_{j=0}^{h} |\beta_j| \\
&\le C\ \sum_{j=0}^{h} \varepsilon_j \le C\ \sum_{j=0}^{h} (\kappa_0\varepsilon_0)^{2^j-1} \le 2C\ \varepsilon_0\ ,
\end{aligned}
\tag{5.29}
$$

provided that $\kappa_0\varepsilon_0$ is smaller than $1/2$.

The important points of the above estimates are:

(1) The bound for ε_h can be made as small as we want by making ε_0 sufficiently small. Hence, by imposing some smallness assumption in ε_0, we can ensure that K_h does not leave a neighborhood of K_0.
(2) In particular, under suitable smallness assumptions on ε_0 (independent of h), the range of $K_h(\mathbb{T}^d_{\rho_h})$ is $\eta/2$ away from the boundary of the domain of f_μ.
(3) Under suitable smallness assumption on ε_0 (independent of h), $\|DK_h - DK_0\|_{\rho_h}$ is so small that we have the estimates on the hyperbolic splitting assumed in Lemma 15. The K_h are also in the region where we have uniform bounds on the non-degeneracy constants.
(4) Since $\|\Delta_h\|_{\rho_j} \le \varepsilon_h\delta_h$ decreases very fast, much faster than an exponential, we obtain that by taking ε_0 sufficiently small, we recover the assumption (4.19).
(5) Putting together the above two remarks, we obtain that under a finite number of smallness assumptions on ε_0, we obtain that we can iterate the procedure indefinitely and remain in the neighborhood where we have uniform estimates for all the non-degeneracy assumptions.
(6) Since $\sum_h \|\Delta_h\|_{\rho_j} \le \varepsilon_h\delta_h \le C\varepsilon_0$, we obtain that the K_h converge in $\mathcal{A}_{\rho_\infty}$ and they satisfy the conclusions of Theorem 20, including the estimates (3.9).

We conclude by noticing that (3.11) come from the fact that, due to Lemma 16, the estimates $|\lambda_\pm - \tilde\lambda_\pm|$, $|\lambda_c^\pm - \tilde\lambda_c^\pm|$ are bounded by a constant time $\|\gamma - \tilde\gamma\|_\rho$; with the same argument as in (5.27), we obtain (3.11).

6. Local uniqueness of the solution

In this Section we give a result on the local uniqueness of the solution of the invariance equation, see Theorem 34.

Theorem 34. *Let $\omega \in \mathcal{D}_d(\nu,\tau)$, $d \le n$; let $\mathcal{M}$ be as in Section 2.1 and let $f_\mu : \mathcal{M} \to \mathcal{M}$, $\mu \in \mathbb{R}^d$, be a family of real analytic, conformally symplectic maps with* $0 < \lambda < 1$.

Let $(K^{(1)}, \mu^{(1)})$, $(K^{(2)}, \mu^{(2)})$ be solutions of (2.7) *and assume that*

$$\int_{\mathbb{T}^d} M^{-1}(\theta)\ [K^{(2)}(\theta) - K^{(1)}(\theta)]_1\ d\theta = 0\ , \tag{6.1}$$

Recall that we defined M on the central bundle using J^c, where M is as in (3.5) *and $[\cdot]_1$ means that we take the first component.*

Assume that the non-degeneracy condition $(H5)$ is satisfied at $(K^{(1)}, \mu^{(1)})$. Let W^c, W^s, W^u be the projections of $M^{-1}(\theta)(K^{(2)}(\theta) - K^{(1)}(\theta))$ on the center, stable, unstable subspaces. Let $0 < \delta < \rho$ and assume that the following inequalities are satisfied:

$$\begin{aligned} C\ \nu\ \delta^{-\tau}\ \max(\|W^c\|_{\rho+2\delta}, |\mu^{(2)} - \mu^{(1)}|)\ &<\ 1 \\ C\ (\|W^s\|_{\rho+\delta} + \|W^u\|_{\rho+\delta})\ &<\ 1\ . \end{aligned} \tag{6.2}$$

Then, we have:

$$K^{(1)} = K^{(2)}\ , \qquad \mu^{(1)} = \mu^{(2)}\ .$$

Remark 35. *The normalization condition* (6.1) *has a very transparent geometric interpretation. Remember that the M is a linear change of variables in the torus that selects the tangent and the symplectic conjugate. The normalization chosen roughly imposes that the average of the increase in phase of $K^{(2)}$ is the same as that of $K^{(1)}$, since the average over the angles of the difference of the two solutions in the parametric coordinates is zero (compare with Section 2.4.1).*

Remark 36. *Note that if we take $K^{(2)}(\theta) = K^{(1)}(\theta + \sigma)$, then*

$$\frac{d}{d\sigma} \int_{\mathbb{T}^d} M^{-1}(\theta)\ [K^{(1)}(\theta + \sigma) - K^{(1)}(\theta)]_1\ d\theta = 0\ \big|_{\sigma=0} = \mathrm{Id}\ .$$

Therefore, the finite dimensional implicit function theorem shows that given any $K^{(2)}$ close to $K^{(1)}$, there is a unique σ such that $K^{(2)} \circ T_\sigma$ satisfies the normalization (6.1)*. The estimates on σ show that if there existed a solution $K^{(2)}$ close to $K^{(1)}$, then, we could, without loss of generality, get a normalized solution by composing it with a translation. This normalized translation satisfies similar hypothesis of proximity to $K^{(1)}$ as the original $K^{(2)}$.*

Hence, the normalization can be interpreted as fixing the element in the family of solutions mentioned at the beginning of Section 2.4.1.

Proof. Assume that $(K^{(1)}, \mu^{(1)})$, $(K^{(2)}, \mu^{(2)})$ satisfy the invariance equations

$$f_{\mu^{(1)}} \circ K^{(1)} = K^{(1)} \circ T_\omega\ , \qquad f_{\mu^{(2)}} \circ K^{(2)} = K^{(2)} \circ T_\omega\ .$$

Let us define the quantity $\widetilde{R}$ as

$$\widetilde{\mathcal{R}} = f_{\mu^{(2)}}\circ K^{(2)} - f_{\mu^{(1)}}\circ K^{(1)} - Df_{\mu^{(1)}}\circ K^{(1)}\ (K^{(2)}-K^{(1)}) - D_\mu f_{\mu^{(1)}}\circ K^{(1)}\ (\mu^{(2)}-\mu^{(1)})\ . \quad (6.3)$$

By Taylor's theorem, setting the constant $\widetilde{C}$ as

$$\widetilde{C} = \frac{1}{2}\ \max(\|D^2 f_{\mu^{(1)}}\circ K^{(1)}\|_{\rho-\delta}, \|D^2_\mu f_{\mu^{(1)}}\circ K^{(1)}\|_{\rho-\delta})\ ,$$

one has the quadratic estimates

$$\|\widetilde{\mathcal{R}}\|_{\rho-\delta} \le \widetilde{C}\ (\|K^{(2)} - K^{(1)}\|^2_\rho + |\mu^{(2)} - \mu^{(1)}|^2)\ . \quad (6.4)$$

Projecting (6.3) on the center subspace, one obtains

$$\begin{aligned} &\Pi^c_{K^{(2)}(\theta)}\{K^{(2)}\circ T_\omega - K^{(1)}\circ T_\omega - Df_{\mu^{(1)}}\circ K^{(1)}\ (K^{(2)}-K^{(1)}) - D_\mu f_{\mu^{(1)}}\circ K^{(1)}\ (\mu^{(2)}-\mu^{(1)})\}\\ -\ \ &\Pi^c_{K^{(2)}(\theta)}\widetilde{\mathcal{R}} = 0\ . \end{aligned} \quad (6.5)$$

Let $W = W(\theta)$ be defined[8] by:

$$K^{(2)}(\theta) - K^{(1)}(\theta) = M(\theta)W(\theta)\ ,$$

where we intend that $M(\theta)$ is computed for $K^{(1)}$, i.e.

$$M(\theta) = [DK^{(1)}(\theta)\ |\ (J^c)^{-1}\circ K^{(1)}(\theta)\ DK^{(1)}(\theta)\ (DK^{(1)}(\theta)^T DK^{(1)}(\theta))^{-1}]\ .$$

Let $W^c(\theta) = \Pi^c_{K^{(2)}(\theta)}W(\theta)$; similarly to (6.4), one has that

$$\|\Pi^c_{K^{(2)}(\theta)}\widetilde{\mathcal{R}}\|_{\rho-\delta} \le \widetilde{C}\ (\|W^c\|^2_\rho + |\mu^{(2)} - \mu^{(1)}|^2)\ . \quad (6.6)$$

Using the automatic reducibility[9] (4.13) on the center subspace, we obtain:

$$\begin{aligned} &\mathcal{B}(\theta)W^c(\theta) - W^c(\theta+\omega) + M^{-1}(\theta+\omega)\ \Pi^c_{K^{(2)}(\theta)}(D_\mu f_{\mu^{(1)}}\circ K^{(1)}(\theta))(\mu^{(2)}-\mu^{(1)})\\ +\ \ &M^{-1}(\theta+\omega)\Pi^c_{K^{(2)}(\theta)}\widetilde{\mathcal{R}}(\theta) = 0\ . \end{aligned} \quad (6.7)$$

Hence, the W^c is a solution of the equation (6.7) By the remarks in Section 5.1.3 we obtain that the solutions are unique up to adding a constant vector to W^c_1. The estimates for the zero average solution are obtained from (6.7). We note that using (6.5), we obtain that the average of W^c_1 should be bounded by the size of R. Hence, we obtain that the W^c used here satisfies:

$$\|W^c\|_{\rho-\delta} \le C\nu\delta^{-\tau}\ \|\Pi^c_{K^{(2)}(\theta)}\widetilde{\mathcal{R}}\|_\rho\ , \qquad |\mu^{(2)} - \mu^{(1)}| \le C\ \|\Pi^c_{K^{(2)}(\theta)}\widetilde{\mathcal{R}}\|_\rho\ . \quad (6.8)$$

[8]Notice that the formulas are the same as the iterative step. Therefore, even if we use the same letters as in the study of the iterative step the interpretation of the letters is slightly different. In the iterative step, W was an unknown to be found. Here, it is a known quantity and our goal is to show that it satisfies some equations and that we have estimates for it.

[9]Note that because $(K^{(1)}, \mu^{(1)})$ is an exact solution of (2.7), the approximate reducibility is exact.

Next, we recall Hadamard's three circle theorem ([Ahl76]), which gives

$$\|W^c\|_\rho \leq C\ \|W^c\|_{\rho-2\delta}^{\frac{1}{2}}\ \|W^c\|_{\rho+2\delta}^{\frac{1}{2}}\ . \tag{6.9}$$

Hence, from (6.6), (6.8), (6.9), we obtain

$$\begin{aligned}
\max(\|W^c\|_{\rho-2\delta}, |\mu^{(2)}-\mu^{(1)}|) &\leq C\nu\delta^{-\tau}\ \|\Pi^c_{K^{(2)}(\theta)}\widetilde{\mathcal{R}}\|_{\rho-\delta} \\
&\leq C\nu\delta^{-\tau}\ \max(\|W^c\|_\rho^2, |\mu^{(2)}-\mu^{(1)}|^2) \\
&\leq C\nu\delta^{-\tau}\ \max(\|W^c\|_{\rho+2\delta}, |\mu^{(2)}-\mu^{(1)}|)\ \max(\|W^c\|_{\rho-2\delta}, |\mu^{(2)}-\mu^{(1)}|) \\
&< \max(\|W^c\|_{\rho-2\delta}, |\mu^{(2)}-\mu^{(1)}|)\ ,
\end{aligned}$$

if the first in (6.2) holds. This allows to conclude that $W^c = 0$ and $\mu^{(1)} = \mu^{(2)}$.

We now project on the hyperbolic subspaces; let

$$W^h = \Pi^h_{K^{(2)}(\theta)} M(K^{(2)} - K^{(1)})\ ,$$

where $h = s$ or $h = u$. From Section 5.1.4, we have that

$$\|W^h\|_{\rho-\delta} \leq C\ \|\widetilde{\mathcal{R}}\|_\rho\ .$$

From (6.4), knowing that $W^c = 0$ and $\mu_1 = \mu_2$, we have

$$\|\widetilde{\mathcal{R}}\|_{\rho-\delta} \leq \widetilde{C}\ (\|W^s\|_\rho^2 + \|W^u\|_\rho^2)\ .$$

From Hadamard's three circle theorem, we have:

$$\begin{aligned}
\|W^s\|_\rho^2 + \|W^u\|_\rho^2 &\leq C(\|W^s\|_{\rho+\delta}\|W^s\|_{\rho-\delta} + \|W^u\|_{\rho+\delta}\|W^u\|_{\rho-\delta}) \\
&\leq C(\|W^s\|_{\rho+\delta} + \|W^u\|_{\rho+\delta})\ \|\widetilde{\mathcal{R}}\|_{\rho-\delta} \\
&\leq C(\|W^s\|_{\rho+\delta} + \|W^u\|_{\rho+\delta})\ (\|W^s\|_\rho^2 + \|W^u\|_\rho^2)\ .
\end{aligned}$$

This relation, together with the second inequality in (6.2), leads to $W^s = W^u = 0$. □

7. Domains of analyticity of Lindstedt expansions of whiskered tori

In this Section we investigate the domains of analyticity of whiskered tori in conformally symplectic systems in the limit of small dissipation. The discussion below proceeds along the lines of [CCdlL17]. We develop an asymptotic expansion (Lindstedt series) and use this as a starting point for the application of Theorem 20.

The perturbative expansions for the parameterization of the torus are not very different from the treatment in [CCdlL17] – the main difference is that the equation that needs to be studied in the iterative step requires to consider the hyperbolic directions, but this will be very similar to the treatment in Section 5.1.4.

The construction of expansions of invariant bundles will be based on a perturbative treatment of the equations (2.23), (2.24).

It is interesting to note that the Lindstedt series have a triangular structure. The series of the parameterization of the torus solve the invariance equation by themselves; on the other hand the series for the parameterization of the invariant spaces require the series of parameterization of the torus.

The formal expansions of K, μ, A^σ constructed to order N by the Lindstedt method produce objects that satisfy the invariance equations up to an error which has norm less than $C_N|\varepsilon|^{N+1}$. We apply Theorem 20 for ε in a domain such that the Diophantine properties of λ are good enough. We conclude that there exist exactly invariant K, μ solutions and that the distance between K, μ and the formal expansions to order N are bounded in the domain, namely the formal series expansions obtained by Lindstedt series are an asymptotic expansion of the true solution.

7.1. **Description of the set up.** Consider a family of mappings $f_{\mu_\varepsilon,\varepsilon} : \mathcal{M} \to \mathcal{M}$, such that

$$f^*_{\mu_\varepsilon,\varepsilon}\Omega = \lambda(\varepsilon)\Omega\ ,$$

where in this Section the conformal factor λ is assumed to be an analytic function of a small parameter ε and it is such that $\lambda(0) = 1$, thus allowing to recover the symplectic case for $\varepsilon = 0$. In particular, we assume that $\lambda = \lambda(\varepsilon)$ takes the form

$$\lambda(\varepsilon) = 1 + \alpha\varepsilon^a + O(|\varepsilon|^{a+1})$$

for some $a > 0$ integer and $\alpha \in \mathbb{C} \setminus \{0\}$.

7.2. **Some standard definitions.** In our results, we will show that some functions depending on a parameter ε are analytic in ε. The stronger (and most natural) way to study the dependence is to consider the functions for fixed ε taking values in a Banach space of functions. The analyticity is taken to mean that the function can be expressed as a Taylor series in the neighborhood of each point. Other seemingly weaker definitions turn out to be equivalent ([HP74, Chapter III]).

As standard, given a sequence B_j of elements in a Banach space, a formal power series in ε is an expression of the form $B^\infty_\varepsilon = \sum_{j=0}^\infty \varepsilon^j B_j$ (the sum is not meant to converge). We denote by $B^{[\leq N]}_\varepsilon = \sum_{j=0}^N \varepsilon^j B_j$.

In our application, we will consider series in which the Banach spaces are just $\mathbb{C}^d$ (in the case of K_ε, the maps take values in the phase space and in the case of A^σ_ε, it is a space of bundle maps).

7.3. **Description of the domains of analyticity.** Recalling Definition 6, we introduce the following sets, where the Diophantine constants behave in such a way that one can start an iterative convergent procedure (see [CCdlL17]).

Definition 37. *For $A > 0$, $N \in \mathbb{Z}_+$, $\omega \in \mathbb{R}^d$, let the sets $\mathcal{G} = \mathcal{G}(A;\omega,\tau,N)$ and $\Lambda = \Lambda(A;\omega,\tau,N)$ be defined as*

$$\begin{aligned}\mathcal{G}(A;\omega,\tau,N) &\equiv \{\varepsilon \in \mathbb{C} :\ \nu(\lambda(\varepsilon);\omega,\tau)\ |\lambda(\varepsilon)-1|^{N+1} \leq A\}\ ,\\ \Lambda(A;\omega,\tau,N) &\equiv \{\lambda \in \mathbb{C} :\ \nu(\lambda;\omega,\tau)\ |\lambda-1|^{N+1} \leq A\}\ .\end{aligned}$$

For $r_0 \in \mathbb{R}$, let

$$\mathcal{G}_{r_0}(A;\omega,\tau,N) = \mathcal{G} \cap \{\varepsilon \in \mathbb{C} :\ |\varepsilon| \leq r_0\}\ . \tag{7.1}$$

7.4. **Statement of the main result on domains of analyticity, Theorem 38.** We will prove that the parametrization and the drift are analytic in a domain $\mathcal{G}_{r_0}$ as in (7.1) for a sufficiently small r_0. This domain is obtained by removing from a ball centered at zero a sequence of smaller balls whose center lies along smooth lines going through the origin. The radii of the balls which have been removed decrease faster than any power of the distance of their center from the origin. As in [CCdlL17], we conjecture that this domain is essentially optimal.

Theorem 38. *Let $\omega \in \mathcal{D}_d(\nu,\tau)$, $d \leq n$, as in (2.4), let $\mathcal{M}$ be as in Section 2.1, and let $f_{\mu,\varepsilon}$ with $\mu \in \Gamma$ with $\Gamma \subseteq \mathbb{C}^d$ open, $\varepsilon \in \mathbb{C}$, be a family of conformally symplectic maps.*

Assume that the family of maps $f_{\mu,\varepsilon}$ is conformally symplectic, with a conformal symplectic factor that depends analytically on ε, $\lambda_\varepsilon = 1 + \alpha\varepsilon^a + O(\varepsilon^{a+1})$ for $\alpha \in \mathbb{R}$, $\alpha \neq 0$, $a \in \mathbb{N}$.

(A1) Assume that for $\mu = \mu_0, \varepsilon = 0$ the map $f_{\mu_0,0}$ admits a whiskered invariant torus. This assumption amounts to the following requirements:

(A1.1) There exists an embedding $K_0 : \mathbb{T}^d \to \mathcal{M}$, $K_0 \in \mathcal{A}_\rho$ for some $\rho > 0$, such that

$$f_{\mu_0,0} \circ K_0 = K_0 \circ T_\omega\ . \tag{7.2}$$

(A1.2) There exists a splitting $T_K(\theta)\mathcal{M} = E^s_\theta \oplus E^c_\theta \oplus E^u_\theta$. This splitting is invariant under the cocycle $\gamma^0_\theta = Df_{\mu_0,0} \circ K_0(\theta)$ and satisfies Definition 11. The ratings of the splitting satisfy the assumptions (H3), (H3') and (H4) of Theorem 20.

(A.2) We assume that the function $f_{\mu,\varepsilon}(x)$ is analytic in all of its arguments and that the analyticity domains are large enough. That is:

(A2.1) Both the embedding $K_0(\theta)$ and the splittings $E^{s,c,u}_\theta$ considered as a function of θ are in $\mathcal{A}_{\rho_0}$ for some $\rho_0 > 0$.

(A2.2) Assume that there is a domain $\mathcal{U} \subset \mathbb{C}^n/\mathbb{Z}^n \times \mathbb{C}^n$ *such that for* $|\varepsilon| \le \varepsilon^*$ *and all* μ *with* $|\mu - \mu_0| \le \mu^*$, *we have that* $f_{\mu,\varepsilon}$ *is defined in* $\mathcal{U}$ *and we also have* (3.3).

(A3) The invariant torus satisfies the twist condition (H5) of Theorem 20.
Then, we have:

B.1) We can compute formal power series expansions

$$K_\varepsilon^{[\infty]} = \sum_{j=0}^{\infty} \varepsilon^j K_j \qquad \mu_\varepsilon^{[\infty]} = \sum_{j=0}^{\infty} \varepsilon^j \mu_j$$

satisfying (2.7) *in the sense of formal power series, which means that for any* $0 < \rho' \le \rho$ *and* $N \in \mathbb{N}$, *we have*

$$||f_{\mu_\varepsilon^{[\le N]},\varepsilon} \circ K_\varepsilon^{[\le N]} - K_\varepsilon^{[\le N]} \circ T_\omega||_{\rho'} \le C_N |\varepsilon|^{N+1} \ .$$

B.2) We can compute four formal power series expansions

$$A_\varepsilon^{\sigma,\infty} = \sum_{j=0}^{\infty} \varepsilon^j A_j^\sigma \ , \qquad \sigma = s, \hat{s}, u, \hat{u} \ ,$$

$$A_j^\sigma(\theta) : E_0^\sigma(\theta) \to E_0^{\hat{\sigma}}(\theta)$$

and the $A_j^\sigma \in \mathcal{A}_\rho$ *in such a way that the operators satisfy the equation* (2.23), (2.24) *for invariant dichotomies in the sense of power series.*

B.3) For the set $\mathcal{G}_{r_0}$ *as in* (7.1) *with* r_0 *sufficiently small and for* $0 < \rho' < \rho$, *there is* $K_\varepsilon : \mathcal{G}_{r_0} \to \mathcal{A}_{\rho'}$, $\mu_\varepsilon : \mathcal{G}_{r_0} \to \mathbb{C}^d$, *analytic in the interior of* $\mathcal{G}_{r_0}$ *taking values in* $\mathcal{A}_{\rho'}$ *which extends continuously to the boundary of* $\mathcal{G}_{r_0}$, *such that for* $\varepsilon \in \mathcal{G}_{r_0}$ *the invariance equation is satisfied exactly:*

$$f_{\mu_\varepsilon,\varepsilon} \circ K_\varepsilon - K_\varepsilon \circ T_\omega = 0 \ .$$

Moreover, the above solutions admit the formal series in A) as an asymptotic expansion, namely for $0 < \rho' < \rho$, $N \in \mathbb{N}$, *one has:*

$$||K_\varepsilon^{[\le N]} - K_\varepsilon||_{\rho'} \le C_N |\varepsilon|^{N+1} \ , \qquad |\mu_\varepsilon^{[\le N]} - \mu_\varepsilon| \le C_N |\varepsilon|^{N+1} \ .$$

The proof of Theorem 38 is very similar to the main theorem in [CCdlL17] and we sketch in the following Sections just the main ingredients of the proof.

7.5. **Proof of Theorem 38.** By hypothesis we can find an embedding $K_0 : \mathbb{T}^d \to \mathcal{M}$ satisfying (7.2). Since $f_{\mu_0,0}$ is symplectic, one has

$$f_{\mu_0,0}^* \Omega = \Omega \ .$$

Substituting $K_\varepsilon^{[\leq N]}$, $\mu_\varepsilon^{[\leq N]}$ in (7.2), and equating the coefficients with the same power of ε, we obtain recursive relations for the terms K_j, μ_j.

Indeed, the first order in ε is

$$(Df_{\mu_0,0} \circ K_0)K_1 - K_1 \circ T_\omega + (D_\mu f_{\mu_0,0} \circ K_0)\mu_1 + D_\varepsilon f_{\mu_0,0} \circ K_0 = 0 \ .$$

More generally, for any order j, we obtain that

$$(Df_{\mu_0,0} \circ K_0)K_j - K_j \circ T_\omega + (D_\mu f_{\mu_0,0} \circ K_0)\mu_j = R_j \ , \tag{7.3}$$

where R_j is a polynomial in $K_0, \dots, K_{j-1}, \mu_1, \dots, \mu_{j-1}, D_\varepsilon f_{\mu_0,\varepsilon=0} \circ K_0, \dots, D_\varepsilon^j f_{\mu_0,\varepsilon=0} \circ K_0$.

It is important to note that the coefficients multiplying the unknowns K_j, μ_j in (7.3) are $(Df_{\mu_0,0} \circ K_0)$ and $(D_\mu f_{\mu_0,0} \circ K_0)$, respectively. These coefficients do not depend on j and they can be evaluated on the zero order approximation.

Using that the zero order corresponds to a whiskered torus in the Hamiltonian case which is exactly invariant, we obtain that the coefficient $(Df_{\mu_0,0} \circ K_0)$ is exactly reducible (see Lemma 27 and the discussion around it).

That is, defining the matrix M_0 as in (3.5), we have in the center direction,

$$Df_{\mu_0,0} \circ K_0 M_0(\theta) = M_0(\theta + \omega) \begin{pmatrix} \mathrm{Id}_d & S_0(\theta) \\ 0 & \lambda \, \mathrm{Id}_d \end{pmatrix} \ . \tag{7.4}$$

Note that, in our case, using that K_0 is exactly invariant, there is no $\mathcal{E}_R$ term (see (4.21)).

The equation (7.3) for K_j, μ_j can be solved taking advantage of the exact reducibility. It is the same procedure that we had in Section 5.1.2, but now the reducibility equation holds exactly.

Let us write $K_j(\theta) = M_0(\theta) W_j(\theta)$. Using (7.3), we have:

$$(Df_{\mu_0,0} \circ K_0) \, M_0 W_j - M_0 \circ T_\omega \ W_j \circ T_\omega + (D_\mu f_{\mu_0,0} \circ K_0)\mu_j = R_j \ ;$$

using (7.4) it follows that

$$M_0 \circ T_\omega \begin{pmatrix} \mathrm{Id}_d & S_0(\theta) \\ 0 & \lambda \, \mathrm{Id}_d \end{pmatrix} W_j - M_0 \circ T_\omega \ W_j \circ T_\omega + (D_\mu f_{\mu_0,0} \circ K_0)\mu_j = R_j \ ,$$

namely

$$\begin{pmatrix} \mathrm{Id}_d & S_0(\theta) \\ 0 & \lambda \, \mathrm{Id}_d \end{pmatrix} W_j - W_j \circ T_\omega + (M_0 \circ T_\omega)^{-1} (D_\mu f_{\mu_0,0} \circ K_0)\mu_j = (M_0 \circ T_\omega)^{-1} R_j \ .$$

In components, namely taking the first d rows and the last d rows, say $W = (W_{j1}|W_{j2})$, denoting by $[\cdot]_k$ the k-th component, we have:

$$\begin{aligned}\lambda W_{j2} - W_{j2}\circ T_\omega + [(M_0\circ T_\omega)^{-1}(D_\mu f_{\mu_0,0}\circ K_0)]_2\mu_j &= [(M_0\circ T_\omega)^{-1}R_j]_2\\ W_{j1} - W_{j1}\circ T_\omega + S_0 W_{j2} + [(M_0\circ T_\omega)^{-1}(D_\mu f_{\mu_0,0}\circ K_0)]_1\mu_j &= [(M_0\circ T_\omega)^{-1}R_j]_1\ .\end{aligned}$$

Define $\widetilde{E}_j = (M_0\circ T_\omega)^{-1}R_j$, $\widetilde{A}_0 = (M_0\circ T_\omega)^{-1}(D_\mu f_{\mu_0,0}\circ K_0)$, we obtain:

$$\begin{aligned}\lambda W_{j2} - W_{j2}\circ T_\omega + \widetilde{A}_{20}\mu_j &= \widetilde{E}_{2j}\\ W_{j1} - W_{j1}\circ T_\omega + \widetilde{A}_{10}\mu_j &= \widetilde{E}_{1j} - S_0 W_{j2}\ .\end{aligned}\tag{7.5}$$

Taking the average of the first and second equation in (7.5), one has

$$\begin{aligned}(\lambda-1)\overline{W_{j2}} &= \overline{\widetilde{E}_{2j}} - \overline{\widetilde{A}_{20}}\ \mu_j\\ \overline{\widetilde{A}_{10}}\ \mu_j &= \overline{\widetilde{E}_{1j}} - \overline{S_0 W_{j2}}\ .\end{aligned}\tag{7.6}$$

Defining $W_{j2} = \overline{W_{j2}} + (W_{j2})^o$, one has $\overline{S_0 W_{j2}} = \overline{S_0}\,\overline{W_{j2}} + \overline{S_0 (W_{j2})^o}$. Since W_{j2} is an affine function of μ_j, let $(W_{j2})^o = (B_{a0})^o + (B_{b0})^o\mu_j$, where $(B_{a0})^o$, $(B_{b0})^o$ are solutions of

$$\begin{aligned}\lambda(B_{a0})^o - (B_{a0})^o\circ T_\omega &= (\widetilde{E}_{2j})^o\\ \lambda(B_{b0})^o - (B_{b0})^o\circ T_\omega &= -(\widetilde{A}_{20})^o\ .\end{aligned}$$

Using the second of (7.6) we have

$$\overline{\widetilde{A}_{10}}\ \mu_j + \overline{S_0}\ \overline{W_{j2}} + \overline{S_0(B_{b0})^o}\mu_j = -\overline{S_0(B_{a0})^o} + \overline{\widetilde{E}_{1j}}\ ,$$

so that we have:

$$\begin{pmatrix}\overline{S_0} & \overline{\widetilde{A}_{10}} + \overline{S_0(B_{b0})^o}\\ (\lambda-1)\,\mathrm{Id}_d & \overline{\widetilde{A}_{20}}\end{pmatrix}\begin{pmatrix}\overline{W_{j2}}\\ \mu_j\end{pmatrix} = \begin{pmatrix}-\overline{S_0(B_{a0})^o} + \overline{\widetilde{E}_{1j}}\\ \overline{\widetilde{E}_{2j}}\end{pmatrix}\ ,$$

which can be solved for $\overline{W_{j2}}$, μ_j under the non-degeneracy condition (3.7). This concludes the proof of B.1).

To establish B.2), we just observe that, once we have the expansions in powers of ε for the K_j, μ_j, we can obtain the power series expansion in ε for $Df_{\mu_0,0}\circ K_j$, hence for the γ's, which are obtained from $Df_{\mu_0,0}\circ K_0$ by taking projections on fixed spaces. We also note that, if we take projections over the original splittings, we have that $A^{\sigma,(0)} = 0$. Also, the approximate invariance of the initial splitting tells that $\gamma^{s,\hat{s},(0)}$, $\gamma^{\hat{s},s,(0)}$, $\gamma^{u,\hat{u},(0)}$, $\gamma^{\hat{u},u,(0)}$ are small.

If we substitute the expansions for A^σ and equate terms of order ε^j in (2.23), (2.24), we obtain that these equations are satisfied in the sense of power series if and only if for all j we have

$$
\begin{aligned}
&(\gamma^{\hat{s},\hat{s}}_{\theta,0})^{-1}\left[A^{s}_{\theta+\omega,j}\gamma^{s,s}_{\theta,0}-\gamma^{\hat{s},s}_{\theta,0}\right]=A^{s}_{\theta,j}+R^{I}_{s,\theta,j}\\
&[\gamma^{s,\hat{s}}_{\theta-\omega,0}+\gamma^{s,s}_{\theta-\omega,0}A^{\hat{s}}_{\theta-\omega,j}](\gamma^{\hat{s},\hat{s}}_{\theta-\omega,0})^{-1}=A^{\hat{s}}_{\theta,j}+R^{II}_{s,\theta,j}\\
&[\gamma^{\hat{u},u}_{\theta-\omega,0}+\gamma^{\hat{u},\hat{u}}_{\theta-\omega,0}A^{u}_{\theta-\omega,j}](\gamma^{u,u}_{\theta-\omega,0})^{-1}=A^{u}_{\theta,j}+R^{I}_{u,\theta,j}\\
&(\gamma^{u,u}_{\theta,0})^{-1}\left[A^{\hat{u}}_{\theta+\omega,j}\gamma^{\hat{u},\hat{u}}_{\theta,0}-\gamma^{u,\hat{u}}_{\theta,0}\right]=A^{\hat{u}}_{\theta,j}+R^{II}_{u,\theta,j}\ ,
\end{aligned}
\tag{7.7}
$$

where the $R^{I}_{\sigma,\theta,j}$, $R^{II}_{\sigma,\theta,j}$ are explicit polynomial expressions involving only $A^{\sigma}_{\theta,l}$ for $l\le j-1$. Notice that only the γ coefficients enter.

Remark 39. *We note that all the equations in* (7.7) *have the form of fixed points of an operator, some of whose power is a contraction. Note that finding the $A^{\sigma,(j)}$ does not entail loosing any domain of analyticity. Because the structure of the $R^{(k)}_{\sigma,\theta,j}$, we can solve the equations recursively and proceed to find solutions which have the same domain as the $R^{(k)}_{\sigma,\theta,j}$, γ. The domain of the γ can be taken to be as close to the domain of K_0 as desired.*

We fix N sufficiently large (say 5). Then, by choosing ε sufficiently small, all assumptions of Theorem 38 are satisfied. We note that if we take the approximate solution as $(K,\mu)=(K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]})$ and as the approximate splitting the results of the expansion, we have that $\mathcal{E}_h,\mathcal{E}\le C_{N,\rho}|\varepsilon|^{N+1}$ for an analytic norm in a fixed radius ρ slightly smaller than the analytic domain of the original radius.

Since we assume that the torus in the Hamiltonian case is non degenerate, we get that the non-degeneracy conditions are uniform for $|\varepsilon|$ small. If we choose a $\delta<\rho/2$, we can obtain the smallness conditions of Theorem 20 for small enough $|\varepsilon|$. Also the assumption of the range of K being inside the domain of $f_{\mu,\varepsilon}$ are uniform for $|\varepsilon|$ small enough.

We also note that the non-degeneracy conditions for Theorem 20 are uniform in the sets

$$\{a_-<|\varepsilon|<a_+\}\cap\mathcal{G} \tag{7.8}$$

for sufficiently small $0<a_-<a_+$. As we argued before, the condition that ε is small enough ensures that the non-degeneracy conditions are uniform. The intersection with $\mathcal{G}$ ensures that the Diophantine properties are uniform.

Therefore, the iterative procedure in the proof of Theorem 20 is uniform for all the ε in the sets (7.8). We also recall that the iterative step to prove Theorem 20 consists just in performing algebraic operations, shifting functions and solving cohomology equations.

In all these operations, it is clear that if the data depend analytically on ε, so does the correction.

Putting together the two remarks above, we obtain that the procedure leads to a sequence of functions all of which are analytic in ε and which converge uniformly in sets of the form (7.8). Therefore the solution will be analytic in sets of the form (7.8). Due to the local uniqueness of the solution, we obtain that the suitably normalized solutions in different patches that overlap have to agree.

Remark 40. *Note that the conditions of smallness in ε, so that we can apply Theorem 20 to the truncated series, depend on the size of the coefficients and the domain loss. It would be interesting to try to optimize the choices of the orders of truncation and the domain losses depending on ε – similar calculations are often done in the study of Birkhoff Normal forms.*

We conclude by writing the Lindstedt series in B.2) associated to (2.23), (2.24). We start from the first of (2.23). We expand A^s as

$$A^s_\theta = \sum_{j=0}^\infty \varepsilon^j A^s_{\theta,j} \ ,$$

and we expand $\gamma^{\sigma\eta}_\theta$ as

$$\begin{aligned} \gamma^{\sigma\eta}_\theta &= \sum_{j=0}^\infty \varepsilon^j \gamma^{\sigma\eta}_{\theta,j} \qquad if\ \sigma=\eta\ , \\ \gamma^{\sigma\eta}_\theta &= \sum_{j=1}^\infty \varepsilon^j \gamma^{\sigma\eta}_{\theta,j} \qquad if\ \sigma\neq\eta\ . \end{aligned}$$

Inserting the above series expansions in the first of (7.7) and equating same orders of ε^j, $j\geq 1$, one obtains:

$$A^s_{\theta,j} = (\gamma^{\hat s,\hat s}_{\theta,0})^{-1} A^s_{\theta+\omega,j}\gamma^{s,s}_{\theta,0} - (\gamma^{\hat s,\hat s}_{\theta,0})^{-1}\gamma^{\hat s,s}_{\theta,j} - R^I_{s,\theta,j}(A^s_{\theta,1},\dots,A^s_{\theta,j-1})\ . \tag{7.9}$$

Equation (7.9) can be solved by iteration to obtain:

$$\begin{aligned} A^s_{\theta,j} &= (\gamma^{\hat s,\hat s}_{\theta,0})^{-1}\ \{(\gamma^{\hat s,\hat s}_{\theta+\omega,0})^{-1} A^s_{\theta+2\omega,j}\gamma^{s,s}_{\theta+\omega,0} - (\gamma^{\hat s,\hat s}_{\theta+\omega,0})^{-1}\gamma^{\hat s,s}_{\theta+\omega,j} - R^I_{s,\theta+\omega,j}\}\ \gamma^{s,s}_{\theta,0} \\ &- (\gamma^{\hat s,\hat s}_{\theta,0})^{-1}\ \gamma^{\hat s,s}_{\theta,j} + R^I_{s,\theta,j} \\ &= \sum_{k=0}^\infty 1\times\dots\times(\gamma^{\hat s,\hat s}_{\theta+(k-1)\omega,0})^{-1}\ R^I_{s,\theta+k\omega,j}\ (\gamma^{s,s}_{\theta+(k-1)\omega,0})\times\dots\times 1 \\ &- \sum_{k=0}^\infty (\gamma^{\hat s,\hat s}_{\theta,0})^{-1}\times\dots\times(\gamma^{\hat s,\hat s}_{\theta+(k-1)\omega,0})^{-1}(\gamma^{\hat s,\hat s}_{\theta+k\omega,0})^{-1}\ \gamma^{\hat s,s}_{\theta+k\omega,j}\ (\gamma^{s,s}_{\theta+(k-1)\omega,0})\times\dots\times 1\ , \end{aligned}$$

where $1 \times \ldots \times (\gamma^{\hat{s},\hat{s}}_{\theta+(k-1)\omega,0})^{-1} = 1$ and $(\gamma^{s,s}_{\theta+(k-1)\omega,0}) \times \ldots \times 1 = 1$ for $k = 0$. We remark that the products of $1 \times \ldots \times (\gamma^{\hat{s},\hat{s}}_{\theta+(k-1)\omega,0})^{-1}$ by $(\gamma^{s,s}_{\theta+(k-1)\omega,0}) \times \ldots \times 1$ are contractions. The other equations in (2.23), (2.24) are treated in the same way; we omit the details and provide just the final results. Analogously, the second equation in (2.23) gives the following solution:

$$\begin{aligned} A^{\hat{s}}_{\theta,j} &= -\sum_{k=0}^{\infty} 1 \times \ldots \times \gamma^{s,s}_{\theta-k\omega,0} \; R^{II}_{s,\theta-k\omega,j} \; (\gamma^{\hat{s},\hat{s}}_{\theta-k\omega,0})^{-1} \times \ldots \times (\gamma^{\hat{s},\hat{s}}_{\theta-\omega,0})^{-1} \\ &+ \sum_{k=0}^{\infty} 1 \times \ldots \times \gamma^{s,s}_{\theta-k\omega,0} \; \gamma^{s,\hat{s}}_{\theta-(k+1)\omega,j} \; (\gamma^{\hat{s},\hat{s}}_{\theta-(k+1)\omega,0})^{-1} \times \ldots \times (\gamma^{\hat{s},\hat{s}}_{\theta-\omega,0})^{-1} \,, \end{aligned}$$

where $1 \times \ldots \times \gamma^{s,s}_{\theta-k\omega,0} = 1$ and $(\gamma^{\hat{s},\hat{s}}_{\theta-k\omega,0})^{-1} \times \ldots \times (\gamma^{\hat{s},\hat{s}}_{\theta-\omega,0})^{-1} = 1$ for $k = 0$. As for the first equation in (2.24), we obtain:

$$\begin{aligned} A^{u}_{\theta,j} &= -\sum_{k=0}^{\infty} 1 \times \ldots \times \gamma^{\hat{u},\hat{u}}_{\theta-k\omega,0} \; R^{I}_{u,\theta-k\omega,j} \; (\gamma^{u,u}_{\theta-k\omega,0})^{-1} \times \ldots \times (\gamma^{u,u}_{\theta-\omega,0})^{-1} \\ &+ \sum_{k=0}^{\infty} 1 \times \ldots \times \gamma^{\hat{u},\hat{u}}_{\theta-k\omega,0} \; \gamma^{\hat{u},u}_{\theta-(k+1)\omega,j} \; (\gamma^{u,u}_{\theta-(k+1)\omega,j})^{-1} \times \ldots \times (\gamma^{u,u}_{\theta-\omega,j})^{-1} \,, \end{aligned}$$

where $1 \times \ldots \times \gamma^{\hat{u},\hat{u}}_{\theta-k\omega,0} = 1$ and $(\gamma^{u,u}_{\theta-k\omega,0})^{-1} \times \ldots \times (\gamma^{u,u}_{\theta-\omega,0})^{-1} = 1$ for $k = 0$. The second equation in (2.24) is solved as follows:

$$\begin{aligned} A^{\hat{u}}_{\theta,j} &= -\sum_{k=0}^{\infty} (\gamma^{u,u}_{\theta,0})^{-1} \times \ldots \times (\gamma^{u,u}_{\theta+(k-1)\omega,0})^{-1} R^{II}_{u,\theta+k\omega,j} \; (\gamma^{\hat{u},\hat{u}}_{\theta+(k-1)\omega,j}) \times \ldots \times 1 \\ &- \sum_{k=0}^{\infty} (\gamma^{u,u}_{\theta,0})^{-1} \times \ldots \times (\gamma^{u,u}_{\theta+k\omega,0})^{-1} \gamma^{u,\hat{u}}_{\theta+k\omega,j} \; \gamma^{\hat{u},\hat{u}}_{\theta+(k-1)\omega,0} \times \ldots \times 1 \,, \end{aligned}$$

where $(\gamma^{u,u}_{\theta,0})^{-1} \times \ldots \times (\gamma^{u,u}_{\theta+(k-1)\omega,0})^{-1} = 1$ and $\gamma^{\hat{u},\hat{u}}_{\theta+(k-1)\omega,0} \times \ldots \times 1 = 1$ for $k = 0$.

Appendix A. Proof of the closing lemma and its consequences

In this Appendix we provide the proof of Lemma 14 and of some of its consequences, precisely Lemma 15 and Lemma 16.

A.1. **Proof of Lemma 14.** Assume that we have a splitting in the neighborhood of the reference splitting, so that we can describe the splitting by the functions A^{σ}_{θ} as in (2.16). Let γ be a cocycle over a rotation.

Our first task is to formulate a functional equation for the A^{σ}_{θ} that is equivalent to their graphs being invariant. Then, we will transform this equation into a contraction mapping theorem. These constructions are very standard in the theory of hyperbolic systems ([Ano69, HPS77]).

We see that for a vector in the graph of A^σ_θ (which we write as $x + A^\sigma_\theta x$ with $x \in E^\sigma_\theta$), we have that its image under $\gamma_\theta = \gamma(\theta)$, expressed in components, is:

$$\gamma_\theta(x + A^\sigma_\theta x) = \left(\gamma^{\sigma,\sigma}_\theta x + \gamma^{\sigma,\hat\sigma}_\theta A^\sigma_\theta x\right) + \left(\gamma^{\hat\sigma,\sigma}_\theta x + \gamma^{\hat\sigma,\hat\sigma}_\theta A^\sigma_\theta x\right) \ . \tag{A.1}$$

The point (A.1) is in the graph of $A^\sigma_{\theta+\omega}$ for all $x \in E^\sigma$, if and only if the matrices A^σ_θ satisfy:

$$A^\sigma_{\theta+\omega}\big(\gamma^{\sigma,\sigma}_\theta + \gamma^{\sigma,\hat\sigma}_\theta A^\sigma_\theta\big) = \gamma^{\hat\sigma,\sigma}_\theta + \gamma^{\hat\sigma,\hat\sigma}_\theta A^\sigma_\theta \ . \tag{A.2}$$

Conversely, since the derivation of (A.2) is just algebra, we see that if (A.2) holds, all the points in the graph of A^σ_θ will be transformed into maps in the graph of $A^\sigma_{\theta+\omega}$.

Hence, our treatment will be based on discussing (A.2), manipulating it algebraically till it becomes a contraction. Note that (A.2) is a very general calculation and that it applies to any dichotomy.

To guess the algebraic transformations that make (A.2) into a contraction in our cases, it is useful to remark that $\gamma^{\sigma,\hat\sigma}$, $\gamma^{\hat\sigma,\sigma}$ will be assumed to be sufficiently small and that the cocycles generated by $\gamma^{\sigma,\sigma}$ and $\left(\gamma^{\hat\sigma,\hat\sigma}\right)^{-1}$ have different contraction/growth rates, see (2.14).

Hence, (A.2) is heuristically a small perturbation of

$$A^\sigma_{\theta+\omega}\gamma^{\sigma,\sigma}_\theta = \gamma^{\hat\sigma,\hat\sigma}_\theta A^\sigma_\theta \ . \tag{A.3}$$

The rearrangements of the equation (A.2) that are useful to reformulate it as a contraction are different depending on the cases we consider. Note that we need to study two dichotomies: $\sigma = s, \hat\sigma = \hat s$ and $\sigma = u, \hat\sigma = \hat u$. Hence, we will need two equations for each of the two dichotomies.

The manipulations needed can be understood by looking at (A.3). We want to isolate one of the A^σ appearing in (A.3) in such a way that the RHS is a contraction. Once we get that the main part is a contraction, it will follow that the (arbitrarily) small terms omitted from (A.3) do not affect the contraction properties.

For the dichotomy between $s, \hat s$ spaces we use:

$$\begin{aligned}
&(\gamma^{\hat s,\hat s}_\theta)^{-1}\left[A^s_{\theta+\omega}\big(\gamma^{s,s}_\theta + \gamma^{s,\hat s}_\theta A^s_\theta\big) - \gamma^{\hat s,s}_\theta\right] = A^s_\theta \ , \\
&\left[\gamma^{s,\hat s}_{\theta-\omega} + \gamma^{s,s}_{\theta-\omega}A^{\hat s}_{\theta-\omega} - A^{\hat s}_\theta \gamma^{\hat s,s}_{\theta-\omega} A^{\hat s}_{\theta-\omega}\right]\left(\gamma^{\hat s,\hat s}_{\theta-\omega}\right)^{-1} = A^{\hat s}_\theta \ .
\end{aligned} \tag{A.4}$$

For the dichotomy corresponding to $u, \hat u$, we obtain the pair of equations:

$$
\begin{aligned}
&\Big[-A^u_\theta \gamma^{u,\hat u}_{\theta-\omega} A^u_{\theta-\omega} + \gamma^{\hat u,u}_{\theta-\omega} + \gamma^{\hat u,\hat u}_{\theta-\omega} A^u_{\theta-\omega}\Big] (\gamma^{u,u}_{\theta-\omega})^{-1} = A^u_\theta \\
&(\gamma^{u,u}_\theta)^{-1} \Big[A^{\hat u}_{\theta+\omega}\big(\gamma^{\hat u,\hat u}_\theta + \gamma^{\hat u,u}_\theta A^{\hat u}_\theta\big) - \gamma^{u,\hat u}_\theta\Big] = A^{\hat u}_\theta\ .
\end{aligned}
\tag{A.5}
$$

The two systems (A.4) and (A.5) can be dealt with by the same methods. So, we will only discuss (A.5). It will be important to note that the estimates that we obtain for the solutions depend only on the constants C_0 and the rates entering into (2.14).

We realize that, if we eliminate from (A.5) the blocks of γ that can be made small by assuming that the splitting is almost invariant, then we are led to consider the fixed point of the operator $\mathcal{N}_0$ defined as

$$
\mathcal{N}_0(A^u, A^{\hat u})_\theta = \begin{pmatrix} \gamma^{\hat u,\hat u}_{\theta-\omega} A^u_{\theta-\omega} (\gamma^{u,u}_{\theta-\omega})^{-1} \\ (\gamma^{u,u}_\theta)^{-1} A^{\hat u}_{\theta+\omega} \gamma^{\hat u,\hat u}_\theta \end{pmatrix}\ .
$$

In the following we present some (rather arbitrary choices) that work.

Note that $\mathcal{N}_0$ is a linear operator and that powers of it are obtained by multiplying the arguments by cocycles in the right and in the left (and by shifting the arguments).

Due to the rate conditions, there exists an $L > 0$ such that, by iterating $\mathcal{N}_0$, L times, we can make the Lipschitz constant of the iterate (in the analytic norm) smaller than $1/2$, $\mathrm{Lip}(\mathcal{N}_0^L) < 1/2$ – where the Lipschitz constant is in the space $\mathcal{A}_\rho$.

If we consider the ball $\|A^s_\theta\|_\rho, \|A^u_\theta\|_\rho < M_1$ for some $M_1 > 0$, we can find smallness conditions on $\|\gamma^{u,\hat u}\|_\rho, \|\gamma^{\hat u,u}\|_\rho$, so that the contraction of $\mathcal{N}_0$ in this ball is smaller than $3/4$.

We recall that the splitting E is η-approximately invariant and that the distance between the splittings can be measured by (2.18). Using that $(\mathcal{N}_0)^L$ is a contraction, it follows that $\mathcal{N}_0$ has a unique fixed point. We conclude with the standard fixed point estimates which, together with (2.22), lead to

$$
\max_\sigma \|A^\sigma_\theta\| \le C\eta\ .
$$

A.2. **Proof of Lemma 15.** There exists $N \in \mathbb{N}$ such that $\|\Gamma^N\|_\rho \le \frac14$ (indeed just take $N = \ln(\frac{1}{4C_0})/\ln(\xi)$). Now, there exists an $\varepsilon^* > 0$ so that $\|\gamma - \tilde\gamma\|_\rho \le \varepsilon^*$ implies that

$$
\|\widetilde\Gamma^N\|_\rho \le \frac12\ .
$$

We recall that $\widetilde\Gamma^\ell_0 = \widetilde\Gamma^\ell$, we write $\widetilde\Gamma^{\ell+Nk}_{Nk} = \widetilde\Gamma^{\ell+Nk-Nk}_0 \circ T_{Nk\omega}$, so that $\widetilde\Gamma^{\ell+Nk}_0 = \widetilde\Gamma^{\ell+Nk}_{Nk}\widetilde\Gamma^{Nk}_0 = \widetilde\Gamma^\ell_0 \circ T_{Nk\omega}\widetilde\Gamma^{Nk}_0$. Then, we have:

$$
\|\widetilde\Gamma^{Nk+\ell}\|_\rho \le \left(\frac12\right)^k \sup_{0<\ell\le N} \|\widetilde\Gamma^\ell\|_\rho = \left(\frac{1}{2^{1/N}}\right)^{Nk+\ell} \frac{1}{2^{-\ell/N}} \sup_{0<\ell\le N} \|\widetilde\Gamma^\ell\|_\rho\ .
$$

Hence we obtain the desired result with $\tilde{\xi} = (1/2)^{1/N}$, $\tilde{C}_0 = 2\sup_{0<\ell\le N} \|\widetilde{\Gamma}^\ell\|_\rho$.

A.3. **Proof of Lemma 16.** The method of proof is very similar to perturbation arguments of semigroups ([HP74]).

The first observation is that, using the operators A_θ^σ as in (2.16), we can identify the approximately invariant spaces E_θ^σ with the invariant ones $\widetilde{E}_\theta^\sigma$. Let the invariant cocycle be

$$\tilde{\gamma}_\theta^{\sigma,\sigma} : \widetilde{E}_\theta^\sigma \to \widetilde{E}_{\theta+\omega}^\sigma \ .$$

We will only consider the case of the forward cocycles. The case of the inverse requires only to change ω to $-\omega$ and to consider the inverse cocycle.

Adding and subtracting appropriate terms, using the notation in (2.9)-(2.10), we have

$$\begin{aligned}\widetilde{\Gamma}_\theta^n &\equiv \tilde{\gamma}_{\theta+(k-1)\omega}^{\sigma,\sigma} \cdots \tilde{\gamma}_\theta^{\sigma,\sigma} \\ &= \gamma_{\theta+(k-1)\omega}^{\sigma,\sigma} \cdots \gamma_\theta^{\sigma,\sigma} + \sum_{j=0}^{k-1} \gamma_{\theta+(k-1)\omega}^{\sigma,\sigma} \cdots \gamma_{\theta+(j+1)\omega}^{\sigma,\sigma} \left[\tilde{\gamma}_{\theta+j\omega}^{\sigma,\sigma} - \gamma_{\theta+j\omega}^{\sigma,\sigma}\right] \tilde{\gamma}_{\theta+(j-1)\omega}^{\sigma,\sigma} \cdots \tilde{\gamma}_\theta^{\sigma,\sigma} \\ &= \Gamma_\theta^k + \sum_{j=0}^{k-1} \Gamma_{j+1}^k \left[\tilde{\gamma}_{\theta+j\omega}^{\sigma,\sigma} - \gamma_{\theta+j\omega}^{\sigma,\sigma}\right] \widetilde{\Gamma}_\theta^j \ ,\end{aligned} \tag{A.6}$$

where we define $\widetilde{\Gamma}_\theta^0$ and Γ_θ^0 to be the identity, and we intend that $\tilde{\gamma}_{\theta+(j-1)\omega}^{\sigma,\sigma} \cdots \tilde{\gamma}_\theta^{\sigma,\sigma} = 0$ for $j = 0$ and $\gamma_{\theta+(k-1)\omega}^{\sigma,\sigma} \cdots \gamma_{\theta+(j+1)\omega}^{\sigma,\sigma} = 0$ for $j = k-1$. Note that (A.6) is a discrete version of Duhamel formula, so that the rest of the argument is very similar to the arguments in perturbation theory of semigroups.

We will consider (A.6) as a fixed point equation for $\widetilde{\Gamma}^k$ lying in an appropriate space of sequences with an appropriate norm. We therefore write (A.6) as

$$\widetilde{\Gamma} = \Gamma + \mathcal{L}\widetilde{\Gamma}. \tag{A.7}$$

We will show that the operator $\mathcal{L}$ given by

$$(\mathcal{L}\widetilde{\Gamma})^k \equiv \sum_{j=0}^{k-1} \Gamma_{j+1}^k \left[\tilde{\gamma}_{\theta+j\omega}^{\sigma,\sigma} - \gamma_{\theta+j\omega}^{\sigma,\sigma}\right] \widetilde{\Gamma}_\theta^j$$

is a contraction in a space of sequences endowed with a norm that captures the rate.

Iterating (A.7) we obtain

$$\widetilde{\Gamma} = \Gamma + \mathcal{L}\Gamma + \mathcal{L}^2\Gamma + ... + \mathcal{L}^k\Gamma + ... \tag{A.8}$$

The above treatment is very similar to perturbation theory of semigroups in the Physics literature; equation (A.8) is known as the Dyson expansion.

To study (A.7) we introduce appropriate norms in spaces of sequences of operators, precisely:

$$\|\widetilde{\Gamma}\|_{\tilde{\xi}} = \sup_{k\in\mathbb{N}} \left(\tilde{\xi}^{-k}\|\widetilde{\Gamma}^k\|_\rho\right) \ ;$$

we suppress the $\mathcal{A}_\rho$ from the notation for $\|\Gamma\|$ since it will be fixed in this argument. On the other hand, the ξ will be important for us.

We fix $\xi < \tilde{\xi}$ and estimate $\mathcal{L}$ in the norm. We introduce the quantity a as

$$a \equiv \|\tilde{\gamma}^{\sigma,\sigma} - \gamma^{\sigma,\sigma}\|_\rho \ .$$

We note that for any choice of $\tilde{\xi}$ for which we can show that $\mathcal{L}$ is a contraction, we show that there is a solution of (A.7) in the space of functions with rate $\tilde{\xi}$ (and of course, that the solution is unique). Obviously the $\widetilde{\Gamma}$ produced by the recursion (A.8) is a solution. Therefore, we show that the (A.8) has growth with exponent $\tilde{\xi}$:

$$\begin{aligned}
\|(\mathcal{L}\widetilde{\Gamma})^k\|_\rho\tilde{\xi}^{-k} &\leq \tilde{\xi}^{-k}\sum_{j=0}^{k-1}\|\Gamma^k_{j+1}\|_\rho a\|\widetilde{\Gamma}^j\|_\rho \\
&\leq \tilde{\xi}^{-k}\sum_{j=0}^{k-1} C_0\xi^{k-(j+1)}a\tilde{\xi}^j\|\widetilde{\Gamma}\|_{\tilde{\xi}} \\
&= C_0 a\sum_{j=0}^{k-1}\xi^{k-j}\tilde{\xi}^{-k+j}\|\widetilde{\Gamma}\|_{\tilde{\xi}}(\xi^{-1}) \\
&\leq C_0 a\tilde{\xi}^{-1}\frac{1}{1-\xi/\tilde{\xi}}\|\widetilde{\Gamma}\|_{\tilde{\xi}} \ .
\end{aligned}$$

Hence, if we take $\xi/\tilde{\xi} < \frac{1}{2}$ and

$$C_0 \ a \ \tilde{\xi}^{-1} \leq 1/4,$$

we ensure that $\|\mathcal{L}\|$ is a contraction.

The estimate of the constant $\tilde{C}_0$ follows from the fact that $\|\mathcal{L}\| = C_0 a\tilde{\xi}^{-1}\frac{1}{1-\xi/\tilde{\xi}} < 1/2$.

Appendix B. Non-Euclidean manifolds

In this Section, we discuss how one can adapt the results for non-Euclidean manifolds. For lower dimensional tori, this is interesting because there are examples with lower dimensional tori with a non-trivial topology of the neighborhoods (in Lagrangian tori, this cannot happen).

In non-Euclidean manifolds, we run into two problems.

One is that for approximately invariant tori, $\mathcal{T}_{f_\mu\circ K(\theta))} \neq \mathcal{T}_{K(\theta+\omega)}$ and, hence for tori in a non-Euclidean manifold we cannot write $Df_\mu \circ K(\theta) = DK(\theta+\omega)$, which is a very suggestive notation for effects of iterations. This problem has appeared very frequently in dynamics. A standard way of fixing the problem is to construct connectors S_x^y ([HPPS70]) which identify the tangent spaces of close enough points[10]. Hence, the cocycles one should consider are

$$Df_\mu \circ K(\theta + n\omega) S^{K(\theta+n\omega)}_{f\circ K(\theta+(n-1)\omega)} \times \cdots \times Df_\mu \circ K(\theta+\omega) S^{K(\theta+\omega)}_{f_\mu \circ K(\theta)} .$$

Note, however, that if the bundles are not trivial, these cannot be identified with matrix cocycles.

A second problem is that $f_\mu \circ K(\theta)$ transforms $\Omega_{K(\theta)}$ into a multiple of $\Omega_{f_\mu(K(\theta))}$. Even if one identifies the tangent spaces, it is not clear what are the geometric properties of the product (2.9).

There are standard ways of correcting this. For example [GEdlL08] uses the Global Darboux theorem from [Mos65, Wei73] to change slightly the map in such a way that approximate cocycles are exactly conformally symplectic. As shown in [GEdlL08] these changes do not alter the quadratic convergence of the algorithm because the size of the required changes can be bounded by the error in the invariance equation.

In our case, this problem appears only in Section 4.

References

[Agr10] A. A. Agrachëv. Invariant Lagrangian submanifolds of dissipative systems. *Uspekhi Mat. Nauk*, 65(5(395)):185–186, 2010.

[Ahl76] L. Ahlfors. *Complex analysis.* McGraw-Hill, 1976.

[Ano69] D. V. Anosov. *Geodesic flows on closed Riemann manifolds with negative curvature.* American Mathematical Society, Providence, R.I., 1969.

[Arn63] V. I. Arnol'd. Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations. *Russian Math. Surveys*, 18(5):9–36, 1963.

[Arn64] V.I. Arnold. Instability of dynamical systems with several degrees of freedom. *Sov. Math. Doklady*, 5:581–585, 1964.

[Ban02] A. Banyaga. Some properties of locally conformal symplectic structures. *Comment. Math. Helv.*, 77(2):383–398, 2002.

[BC19] A. P. Bustamante and R. Calleja. Computation of domains of analyticity for the dissipative standard map in the limit of small dissipation. *Preprint*, 2019.

[Ben88] Alain Bensoussan. *Perturbation methods in optimal control.* Wiley/Gauthier-Villars Series in Modern Applied Mathematics. John Wiley & Sons Ltd., Chichester, 1988. Translated from the French by C. Tomson.

[Ber01] Rolf Berndt. *An introduction to symplectic geometry*, volume 26 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2001. Translated from the 1998 German original by Michael Klucznik.

[10] A very natural definition of connectors is $y = \exp_x(v)$ with v sufficiently small; we fix x and the derivative of the exponential mapping at v identifies the tangent space at x with the tangent space at y. The chain rule gives that $S_y^z S_x^y = S_x^z$, when x is sufficiently close to y, z and y, z are sufficiently close.

[BHS96] H. W. Broer, G. B. Huitema, and M. B. Sevryuk. *Quasi-Periodic Motions in Families of Dynamical Systems.* Order Amidst Chaos. Springer-Verlag, Berlin, 1996.

[BHTB90] H. W. Broer, G. B. Huitema, F. Takens, and B. L. J. Braaksma. Unfoldings and bifurcations of quasi-periodic tori. *Mem. Amer. Math. Soc.*, 83(421):viii+175, 1990.

[BLZ08] Peter W. Bates, Kening Lu, and Chongchun Zeng. Approximately invariant manifolds and global dynamics of spike states. *Invent. Math.*, 174(2):355–433, 2008.

[BS08] Kristian Bjerklöv and Maria Saprykina. Universal asymptotics in hyperbolicity breakdown. *Nonlinearity*, 21(3):557–586, 2008.

[CC08] A. Celletti and L. Chierchia. Measures of basins of attraction in spin-orbit dynamics. *Celestial Mechanics and Dynamical Astronomy*, 101:159–170, May 2008.

[CC10] R. Calleja and A. Celletti. Breakdown of invariant attractors for the dissipative standard map. *CHAOS*, 20(1):013121, 2010.

[CCdlL13] Renato C. Calleja, Alessandra Celletti, and Rafael de la Llave. A KAM theory for conformally symplectic systems: efficient algorithms and their validation. *J. Differential Equations*, 255(5):978–1049, 2013.

[CCdlL17] Renato C. Calleja, Alessandra Celletti, and Rafael de la Llave. Domains of analyticity and Lindstedt expansions of KAM tori in some dissipative perturbations of Hamiltonian systems. *Nonlinearity*, 30(8):3151, 2017.

[CCdlL19a] Renato C. Calleja, Alessandra Celletti, and Rafael de la Llave. Locally conformally symplectic and conformally symplectic systems: geometry and dynamics. *Manuscript*, 2019.

[CCdlL19b] Renato C. Calleja, Alessandra Celletti, and Rafael de la Llave. Whiskered tori in conformally symplectic systems. *Preprint*, 2019.

[CCFdlL14] Renato C. Calleja, Alessandra Celletti, Corrado Falcolini, and Rafael de la Llave. An extension of Greene's criterion for conformally symplectic systems and a partial justification. *SIAM J. Math. Anal.*, 46(4):2350–2384, 2014.

[CdlL10] R. Calleja and R. de la Llave. A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification. *Nonlinearity*, 23(9):2029–2058, 2010.

[CdS01] Ana Cannas da Silva. *Lectures on symplectic geometry*, volume 1764 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.

[Cel10] Alessandra Celletti. *Stability and Chaos in Celestial Mechanics*. Springer-Verlag, Berlin; published in association with Praxis Publishing Ltd., Chichester, 2010.

[CF12] R. Calleja and J-Ll. Figueras. Collision of invariant bundles of quasi-periodic attractors in the dissipative standard map. *Chaos*, 23(021203), 2012.

[CG18] Alessandra Celletti and Catalin Gales. Dynamics of resonances and equilibria of low earth objects. *SIAM J. Appl. Dyn. Syst.*, 17:203–235, 2018.

[CH17a] M. Canadell and A. Haro. Computation of quasi-periodic normally hyperbolic invariant tori: algorithms, numerical explorations and mechanisms of breakdown. *Jour. Nonlinear Sci.*, 2017.

[CH17b] M. Canadell and A. Haro. A KAM-like theorem for quasi-periodic normally hyperbolic invariant tori. *Jour. Nonlinear. Sci.*, 2017.

[Cha05] Chia-Chun George Chao. *Applied Orbit Perturbation and Maintenance*, volume 1764 of *Aerospace Press Series*. AIAA, 2005.

[CL95] S. N. Chow and H. Leiva. Dynamical spectrum for skew product flow in Banach spaces. In *Boundary value problems for functional-differential equations*, pages 85–105. World Sci. Publ., River Edge, NJ, 1995.

[CL04] A. C. M. Correia and J. Laskar. Mercury's capture into the 3/2 spin-orbit resonance as a result of its chaotic dynamics. *Nature*, 429:848–850, June 2004.

[CL09] A. C. M. Correia and J. Laskar. Mercury's capture into the 3/2 spin-orbit resonance including the effect of core-mantle friction. *Icarus*, 201:1–11, May 2009.

[Cop78] W. A. Coppel. *Dichotomies in stability theory.* Lecture Notes in Mathematics, Vol. 629. Springer-Verlag, Berlin-New York, 1978.

[CZ11] Maciej J. Capiński and Piotr Zgliczyński. Cone conditions and covering relations for topologically normally hyperbolic invariant manifolds. *Discrete Contin. Dyn. Syst.*, 30(3):641–670, 2011.

[DFIZ16a] Andrea Davini, Albert Fathi, Renato Iturriaga, and Maxime Zavidovique. Convergence of the solutions of the discounted equation: the discrete case. *Math. Z.*, 284(3-4):1021–1034, 2016.

[DFIZ16b] Andrea Davini, Albert Fathi, Renato Iturriaga, and Maxime Zavidovique. Convergence of the solutions of the discounted Hamilton-Jacobi equation: convergence of the discounted solutions. *Invent. Math.*, 206(1):29–55, 2016.

[dlL01] R. de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001.

[dlLGJV05] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.

[dlLS18] R. de la Llave and Y. Sire. An a posteriori KAM theorem for whiskered tori in Hamiltonian partial differential equations with applications to some ill-posed equations. *Arch. Rational. Mech. Anal., https://doi.org/10.1007/s00205-018-1293-6*, 2018.

[dlLX19] Rafael de la Llave and Lu Xu. Whiskered tori for exact presymplectic dynamical systems. *Preprint*, 2019.

[DM96a] C. P. Dettmann and G. P. Morriss. Hamiltonian formulation of the Gaussian isokinetic thermostat. *Phys. Rev. E*, 54:2495–2500, Sep 1996.

[DM96b] C. P. Dettmann and G. P. Morriss. Proof of Lyapunov exponent pairing for systems at constant kinetic energy. *Phys. Rev. E*, 53(6):R5545–R5548, 1996.

[DV17] Mikel J. De Viana. *Results on invariant whiskered tori for fibered holomorphic maps and on compensated domains, Dissertation.* PhD thesis, Georgia Inst. of Technology, 2017.

[FdlL15] Adam M. Fox and Rafael de la Llave. Barriers to transport and mixing in volume-preserving maps with nonzero flux. *Phys. D*, 295/296:1–10, 2015.

[FdlLS09a] Ernest Fontich, Rafael de la Llave, and Yannick Sire. Construction of invariant whiskered tori by a parameterization method. Part I: maps and flows in finite dimensions. *J. Differential Equations,*, 246:3136–3213, 2009.

[FdlLS09b] Ernest Fontich, Rafael de la Llave, and Yannick 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.*, 16:9–22, 2009.

[FdlLS15] Ernest Fontich, Rafael de la Llave, and Yannick Sire. Construction of invariant whiskered tori by a parameterization method. Part II: Quasi-periodic and almost periodic breathers in coupled map lattices. *J. Differential Equations*, 259(6):2180–2279, 2015.

[FT18] Jordi-Lluís Figueras and Thomas Ohlson Timoudas. Sharp $\frac{1}{2}$-Hölder continuity of the lyapunov exponent at the bottom of the spectrum for a class of Schrödinger cocycles. *Preprint, https://arxiv.org/abs/1809.05420*, 2018.

[GEdlL08] A. González Enriquez and de la Llave. Analytic smoothing of geometric maps with applications to KAM theory. *J. Diff. Eq.*, 245:1243–1298, 2008.

[Gra74] Samuel M. Graff. On the conservation of hyperbolic invariant tori for Hamiltonian systems. *J. Differential Equations*, 15:1–69, 1974.

[HdlL06a] À. Haro and R. de la Llave. Manifolds on the verge of a hyperbolicity breakdown. *Chaos*, 16(1):013120, 8, 2006.

[HdlL06b] 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(6):1261–1300 (electronic), 2006.

[HdlL07] A. 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 J. Appl. Dyn. Syst.*, 6(1):142–207, 2007.

[HdlL19] Alex Haro and Rafael de la Llave. Spectral theory of transfer operators over irrational rotations. *Preprint*, 2019.

[HdlLS11] Gemma Huguet, Rafael de la Llave, and Yannick Sire. Fast numerical algorithms for the computation of invariant tori in Hamiltonian systems. *Disc. Cont. Dynam. Syst. A*, 17(4):1309–1353, 2011. MP_ARC #09-02.

[Hoo91] W.G. Hoover. *Computational Statistical Mechanics*, volume 11 of *Studies in modern thermodynamics*. Elsevier, Amsterdam, 1991.

[HP74] Einar Hille and Ralph S. Phillips. *Functional analysis and semi-groups.* American Mathematical Society, Providence, R. I., 1974. Third printing of the revised edition of 1957, American Mathematical Society Colloquium Publications, Vol. XXXI.

[HPPS70] M. Hirsch, J. Palis, C. Pugh, and M. Shub. Neighborhoods of hyperbolic sets. *Invent. Math.*, 9:121–134, 1969/1970.

[HPS77] M.W. Hirsch, C.C. Pugh, and M. Shub. *Invariant manifolds.* Springer-Verlag, Berlin, 1977. Lecture Notes in Mathematics, Vol. 583.

[ISM11] Renato Iturriaga and Héctor Sánchez-Morgado. Limit of the infinite horizon discounted Hamilton-Jacobi equation. *Discrete Contin. Dyn. Syst. Ser. B*, 15(3):623–635, 2011.

[JdlLZ99] Àngel Jorba, Rafael de la Llave, and Maorong Zou. Lindstedt series for lower-dimensional tori. In *Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995)*, volume 533 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 151–167. Kluwer Acad. Publ., Dordrecht, 1999.

[JV97] À. Jorba and J. Villanueva. On the persistence of lower-dimensional invariant tori under quasi-periodic perturbations. *J. Nonlinear Sci.*, 7(5):427–473, 1997.

[LR16] Héctor Lomelíand Beatriz Rumbos. *Métodos dinámicos en economia Otra búsqueda del tiempo perdido.* Jit Press, S.A. de CV, 3 edition, 2016.

[LS15] Ugo Locatelli and Letizia Stefanelli. Quasi-periodic motions in a special class of dynamical equations with dissipative effects: a pair of detection methods. *Discrete Contin. Dyn. Syst. Ser. B*, 20(4):1155–1187, 2015.

[LY05a] Yong Li and Yingfei Yi. Persistence of hyperbolic tori in Hamiltonian systems. *J. Differential Equations*, 208(2):344–387, 2005.

[LY05b] Yong Li and Yingfei Yi. Persistence of lower dimensional tori of general types in Hamiltonian systems. *Trans. Amer. Math. Soc.*, 357(4):1565–1600 (electronic), 2005.

[Mas05] J. J. Masdemont. High-order expansions of invariant manifolds of libration point orbits with applications to mission design. *Dyn. Syst.*, 20(1):59–113, 2005.

[MHER95] Edward S. Meadows, Michael A. Henson, John W. Eaton, and James B. Rawlings. Receding horizon control and discontinuous state feedback stabilization. *Internat. J. Control*, 62(5):1217–1229, 1995.

[Mos65] Jürgen Moser. On the volume elements on a manifold. *Trans. Amer. Math. Soc.*, 120:286–294, 1965.

[Mos66] J. Moser. A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:265–315, 1966.

[Mos67] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169:136–176, 1967.

[MS74] John W. Milnor and James D. Stasheff. *Characteristic classes.* Princeton University Press, Princeton, N. J.; University of Tokyo Press, Tokyo, 1974. Annals of Mathematics Studies, No. 76.

[MS12] Dusa McDuff and Dietmar Salamon. *J-holomorphic curves and symplectic topology*, volume 52 of *American Mathematical Society Colloquium Publications.* American Mathematical Society, Providence, RI, second edition, 2012.

[MS17a] Stefano Marò and Alfonso Sorrentino. Aubry–Mather Theory for Conformally Symplectic Systems. *Comm. Math. Phys.*, 354(2):775–808, 2017.

[MS17b] Dusa McDuff and Dietmar Salamon. *Introduction to symplectic topology.* Oxford Graduate Texts in Mathematics. Oxford University Press, Oxford, third edition, 2017.

[MSK03] Yu. A. Mitropolsky, A. M. Samoilenko, and V. L. Kulik. *Dichotomies and stability in nonautonomous linear systems*, volume 14 of *Stability and Control: Theory, Methods and Applications.* Taylor & Francis, London, 2003.

[OT17] Thomas Ohlson Timoudas. Power law asymptotics in the creation of strange attractors in the quasi-periodically forced quadratic family. *Nonlinearity*, 30(12):4483–4522, 2017.

[Rüs75] H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical Systems, Theory and Applications (Battelle Rencontres, Seattle, Wash., 1974)*, pages 598–624. Lecture Notes in Phys., Vol. 38, Berlin, 1975. Springer.

[Rüs76a] H. Rüssmann. Note on sums containing small divisors. *Comm. Pure Appl. Math.*, 29(6):755–758, 1976.

[Rüs76b] H. Rüssmann. On optimal estimates for the solutions of linear difference equations on the circle. *Celestial Mech.*, 14(1):33–37, 1976.

[Sev99] M. B. Sevryuk. The lack-of-parameters problem in the KAM theory revisited. In *Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995)*, pages 568–572. Kluwer Acad. Publ., Dordrecht, 1999.

[SL12] Letizia Stefanelli and Ugo Locatelli. Kolmogorov's normal form for equations of motion with dissipative effects. *Discrete Contin. Dynam. Systems*, 17(7):2561–2593, 2012.

[SS74] Robert J. Sacker and George R. Sell. Existence of dichotomies and invariant splittings for linear differential systems. I. *J. Differential Equations*, 15:429–458, 1974.

[Tim18] Thomas Ohlson Timoudas. Asymptotic laws for a class of quasi-periodic Schrödinger cocycles at the lowest energy of the spectrum. *Preprint, https://arxiv.org/abs/1809.05418*, 2018.

[Vai85] Izu Vaisman. Locally conformal symplectic manifolds. *Internat. J. Math. Math. Sci.*, 8(3):521–536, 1985.

[Wei73] A. Weinstein. Lagrangian submanifolds and Hamiltonian systems. *Ann. of Math.*, 98:377–410, 1973.

[WL98] M. P. Wojtkowski and C. Liverani. Conformally symplectic dynamics and symmetry of the Lyapunov spectrum. *Comm. Math. Phys.*, 194(1):47–60, 1998.

[Yoc92] J.-C. Yoccoz. Travaux de Herman sur les tores invariants. *Astérisque*, 206:Exp. No. 754, 4, 311–344, 1992. Séminaire Bourbaki, Vol. 1991/92.

[Zeh75] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.

[Zeh76] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. II. *Comm. Pure Appl. Math.*, 29(1):49–111, 1976.

Department of Mathematics and Mechanics, IIMAS, National Autonomous University of Mexico (UNAM), Apdo. Postal 20-726, C.P. 01000, Mexico D.F., Mexico
E-mail address: calleja@mym.iimas.unam.mx

Department of Mathematics, University of Rome Tor Vergata, Via della Ricerca Scientifica 1, 00133 Rome (Italy)
E-mail address: celletti@mat.uniroma2.it

School of Mathematics, Georgia Institute of Technology, 686 Cherry St.. Atlanta GA. 30332-1160
E-mail address: rll6@math.gatech.edu