Journal of Dynamics and Differential Equations (2021) 33:1–34
https://doi.org/10.1007/s10884-020-09823-w

Whiskered Tori for Presymplectic Dynamical Systems

Rafael de la Llave[1] **· Lu Xu**[2]

Received: 12 November 2018 / Revised: 4 January 2020 / Published online: 24 January 2020

Abstract

We prove persistence result of whiskered tori for the dynamical system which preserves an exact presymplectic form. The results are given in an a-posteriori format. Given an approximate solution of an invariance equation which satisfies some non-degeneracy assumptions, we conclude that there is a true solution close by. The proof is based on certain iterative procedure by which the accuracy of the approximate solutions of the invariance equation can be improved. The iterative procedure is not based on transformation theory, which is cumbersome for presymplectic systems, but on finding corrections to the solutions of the invariance equation. This iterative procedure takes advantage of identities that come from the preservation of the geometric structure and leads to a very efficient numerical method which has low storage requirements, low operator count per step and it is quadratically convergent. We note that a particular case of presymplectic systems is symplectic perturbed by quasi-periodic systems.

Keywords Whiskered tori · KAM theory · Small divisors · Presymplectic system

Mathematics Subject Classification Primary 37K55 · 70K43 · 58F05 · 70H15

R.L. has been partially supported by NSF Grant DMS-1800241. L.X. has been supported by NSFC Grant 11401251. The collaboration of the authors benefited from the hospitality of the JLU-GT Joint institute for Theoretical Science and the author L.X visited Georgia Institute of Technology during January to July, 2017 supported by the CSC and Jilin University. This material is also based upon work supported by the National Science Foundation under Grant No. DMS-1440140 while both authors were in residence at the Mathematical Science Research Institute in Berkeley, California, during the Fall 2018 semester.

✉ Lu Xu
xulu@jlu.edu.cn

Rafael de la Llave
rafael.delallave@math.gatech.edu

1 School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, GA 30332, USA

2 School of Mathematics, Jilin University, No. 2699 Qianjin St., Changchun 130012, Jilin, People's Republic of China

1 Introduction and Motivation

In this paper, we prove persistence of whiskered tori in presymplectic systems. See later for more precise definitions. We anticipate informally that whiskered tori are tori in which the motion is a rotation, such that there are many infinitesimal perturbations that grow exponentially. Presymplectic systems are dynamical systems whose evolution leaves it invariant under a presymplectic form and a presymplectic form is just a closed form which may be degenerate. See Sect. 2.1. As it happens often in the KAM theory of persistence of quasi-periodic motions, there is a deep interaction between the dynamics and the geometric structures preserved by the systems [17,26,27]. For example, the preservation of a form requires that the tori and their manifolds have special properties. In this paper, we do not deal with the most general case, we only consider some cases when the geometric structures have some extra properties with respect to the map. These assumptions happen in applications and are stable under perturbation. We hope that the methods developed here can cover several other cases.

The presymplectic systems appear in a variety of applications, including time dependent perturbations of Hamiltonian systems [25]. The original motivation for the study of presymplectic manifolds was the Bergmann–Dirac theory of constraints [4,14,15], [16, Section V] [12,21]. The presymplectic geometry provided a framework for quantization [8,22]. The theory for presymplectic systems has been found useful in many situations that involve constraints such as gauge theory [24] or even financial models [9]. Several mechanics textbooks that contain general presentations of presymplectic systems in mechanics are [13,25,30]. Similarly, we mention that whiskered tori were shown to be important structures in the instability of quasi-integrable systems. Indeed, till recently whiskered tori were considered as the only structures causing instability in nearly integrable systems [2].

Our main result is a persistence theorem for whiskered tori in presymplectic systems formulated in the *a-posteriori* format standard in numerical analysis. We formulate an invariance equation for an embedding and a splitting of the tangent space at the range of the embedding in such a way that the zeros of this functional equation are whiskered tori (with the corresponding stable and unstable splittings).

The proof of the persistence result involves *adjusting parameters*, as it is common in KAM theory [26–28]. Since it will be hard to find an invariant torus for a fixed presymplectic map, we consider a family of presymplectic maps indexed by parameters λ. Our first assumption is the existence of a torus which is approximately invariant under a map indexed by some parameter λ_0 and the approximately invariant torus has an approximately invariant hyperbolic splitting. Assuming that the errors are small enough compared to some explicit condition numbers computed on the approximate torus considered (no global conditions such as twist to verify), then, by modifying the approximately invariant torus and the parameter in the family, we obtain a truly invariant torus and the corresponding presymplectic map indexed by some parameter λ which is slightly different from the parameter λ_0. As to the dimension of the parameter λ, we refer to the Remark 3.5. We also prove local uniqueness of the invariant tori and the parameter λ.

Note that this a-posteriori format implies the usual formulation of persistence under small perturbations or the existence of tori in nearly integrable systems. An exactly invariant torus is approximately invariant for a slight perturbation of the map and the whiskered invariant tori of an integrable system are approximately invariant for a nearly integrable system.

With a view to numerical applications, it is relevant to note that the method of our proof is to prove convergence of an iterative process. This iterative process can be implemented as a very efficient algorithm: we only need to consider functions of a number of variables equal to the dimension of the torus considered (we do not need to deal with functions of the dimension of the phase space). A step of the process is quadratically convergent. If we discretize our functions using N numbers, we need a storage $O(N)$ and a step requires $O(N \ln(N))$ operations. Note that thanks to the a-posteriori theorem, by computing the condition numbers we can be assured that the computations are reliable even in regions when other more conventional methods of diagnostic become equivocal. Applying the a-posteriori result to validate Lindstedt series leads to smooth dependence on parameters, even parameters in ranging over a Cantor set, for example, frequencies of the torus which range over the set of Diophantine vectors.

The method of our proof is based on formulating an equation for parameterizations of tori that captures the fact that the torus is invariant and that the motion in it is a rotation as well as another equation that captures the invariance of the bundles. One should think of these equations as equations for the embedding and for the parameters in the family. Then, we describe an iterative procedure that given an approximate solution (embedding of the torus, splitting and parameter value in the family) produces a more approximate solution which satisfies the equation much more accurately. Furthermore, we show that if the initial error is small enough, the procedure can be iterated infinitely many times and it converges to a true solution. To implement this procedure, it is important to take advantage of identities that follow from the fact that the system preserves an exact presymplectic form.

Similar methods were described in [19,20] (a more efficient modification, which we will follow, and numerical algorithms are in [23]) for symplectic systems and in [6,7] for conformally symplectic systems. It could be instructive to compare the differences between these different contexts.

Note that the method of proof used here does not use transformation theory. This is advantageous for numerical implementations since transformations require to discretize functions of as many variables as the dimension of the phase space whereas the present method only requires to discretize functions of as many variables as the dimension of the torus. We also note that, in contrast to Hamiltonian theory where a small transformation can be parameterized by a function (the Hamiltonian—using e.g the Weinstein chart [3]—), a good parameterization of the infinitesimal presymplectic transformations is not easy [7,10,11].

2 Preliminaries and Notations

In this section, we present some preliminary set up and some standard notations. This section can be used as reference.

2.1 Presymplectic Form

We recall that a 2-form Ω is called presymplectic when $d\Omega = 0$. In contrast with symplectic forms, which are assumed to be non-degenerate, presymplectic forms are allowed to have non-trivial kernel, that is:

$$\begin{aligned} \ker \Omega_x &= \{v \in T_x M \,|\, i(v)\Omega = 0\} \\ &= \{v \in T_x M \,|\, \Omega(v, w) = 0, \quad \forall\, w \in T_x M\}. \end{aligned}$$

We will assume that the dimension of $\ker(\Omega_x)$ is independent of x. We warn the readers that some references allow the dimension of the kernel to change over the phase space but we assume it is constant. On the other hand, some authors include the assumption that the dimension of the kernel is 1, but we will allow any dimension of the kernel. See the discussion in [25].

In [1], we can find the result that the distribution $\ker(\Omega_x)$ satisfies the Frobenius condition, therefore, there is a foliation whose leaves have $\ker(\Omega_x)$ as tangent space. This is usually described as *"there is a foliation integrating the distribution* $\ker(\Omega_x)$*"*. Note that this uses the fact that the kernel of Ω has several derivatives so that the proof uses the Frobenius integrability theorem. In our case, we will assume that the kernel of the presymplectic form Ω is analytic.

An important particular case of presymplectic forms are exact forms which satisfy $\Omega = d\alpha$ for certain 1-form α. The presymplectic forms that appear in the Dirac–Bergmann theory of mechanical systems subject to constraints are exact. In our paper, we will assume that presymplecitc form is exact.

2.2 The Phase Space We Will Consider

Since we are looking for invariant tori, without too much loss of generality, we consider the manifold

$$M = \mathbb{R}^{2d} \times \left(\mathbb{T}^m \times \mathbb{R}^m\right) \times \mathbb{T}^l \tag{2.1}$$

endowed with a presymplectic 2-form $\Omega = \mathrm{d}\alpha$ and denote $n = 2d + 2m + l$. We also recall that, a manifold endowed with a presymplectic form is called presymplectic manifold.

The factor $\mathbb{T}^l$ in (2.1) in our applications will correspond to the kernel of the presymplectic form.

Note that the integrability of the kernel established in [1] does not imply that the leaves integrating the kernel are compact or that the phase space is a product manifold. The compactness of the leaves integrating the kernel and the product manifold structure of the phase space are extra assumptions in our set up. See Sect. 8 for an example of a presymplectic system which is not a product. These assumptions hold in examples, which, as we will show, are stable.

We hope that the method presented here can incorporate other situations at the price of a more complicated notation and adding parameters. This seems very interesting question and we hope to come back to them.

In this paper we obtain quasi-periodic motions on the kernel, which does not make sense if the foliation tangent to the kernel does not have leaves with compact closure. One notable case where the phase space indeed factorizes is the case of symplectic systems subject to quasi-periodic perturbations. In the case of quasi-periodic perturbations, the system is a skew system and the presymplectic form is constant. We do not make the assumption that the presymplectic form is the standard one.

In the models we will consider $\mathbb{R}^{2d}$ corresponds to the hyperbolic directions. As in [19,20], we note that we do not need that the stable/unstable bundles are trivial. See examples in [19] of systems with non-trivial stable and unstable bundles.

We note that the manifold M is Euclidean and we can compare vectors at different points. This is mainly for convenience (in general manifolds one can use connectors) and without a big loss of generality, since we are seeking tori (See Remark 3.4). Note that all our arguments happen in a vicinity of a whiskered torus, we can assume that the manifolds are as in M. The metric in M is only used in the analysis and we can just as well use the Euclidean metric.

2.3 Norms and Spaces of Embeddings

In this subsection, we collect several definitions of norms that we will use. They are norms in spaces of analytic functions and they are very standard.

If $x = (x_1, \ldots, x_{m+l}) \in \mathbb{R}^{m+l}$, we set

$$|x| := \max_{j=1,\ldots,l+m} |x_j|.$$

Then, denote U_ρ as the complex strip of width $\rho > 0$, that is:

$$U_\rho = \{z \in \mathbb{C}^{m+l}/\mathbb{Z}^{m+l} : |\mathrm{Im}(z)| < \rho\}.$$

Definition 2.1 The space $(\mathcal{P}_\rho, \|\cdot\|_\rho)$ consists of functions $K : U_\rho \to M$ which are one periodic in all their arguments, real analytic on the interior of U_ρ and continuous on the closure of U_ρ, with the norm

$$\|K\|_{C^K, U_\rho} := \sup_{0 \le |k| \le K, z \in U_\rho} |\mathrm{D}^k K(z)|.$$

It is well known that $(\mathcal{P}_\rho, \|\cdot\|_\rho)$ is a Banach space.

2.4 Diophantine Properties

To deal with the small divisors as the previous KAM type results, we introduce the following definition, which is standard.

Definition 2.2 Given $\gamma > 0$ and $\sigma \ge l + m$, we will denote by $D(\gamma, \sigma)$ the set of frequency vectors $\omega \in \mathbb{R}^{l+m}$ satisfying the *Diophantine condition*:

$$|\langle k, \omega\rangle - p| \ge \gamma |k|^{-\sigma}, \quad \forall k \in \mathbb{Z}^{l+m} \setminus \{0\}, \quad p \in \mathbb{Z}, \tag{2.2}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean product.

We will introduce the classical result (see [29]) which provides existence and estimate of a solution of cohomology equations for Diophantine rotations.

Lemma 2.1 *Let $\omega \in D(\gamma, \sigma)$ and assume that $h : \mathbb{T}^{m+l} \to M$ is analytic on U_ρ and has zero average,* $\mathrm{avg}(h) = 0$. *Then for all $0 < \delta < \rho$, the difference equation*

$$v(\theta) - v(\theta + \omega) = h(\theta) \tag{2.3}$$

has a unique zero average solution $v : \mathbb{T}^{m+l} \to \mathbb{R}^{2m+l}$ which is analytic in $U_{\rho-\delta}$. Moreover, this solution satisfies the following estimate:

$$\|v\|_{\rho-\delta} \le c_0 \gamma^{-1} \delta^{-\sigma} \|h\|_\rho, \tag{2.4}$$

where c_0 is a constant depending only on the dimension of the torus $m + l$ and on σ.

2.5 Families of (Exact) Presymplectic Maps

Let M be a manifold endowed with a presymplectic 2-form Ω. We introduce that:

Definition 2.3 A diffeomorphism $f : M \to M$ is called presymplectic diffeomorphism if

$$f^*\Omega = \Omega. \tag{2.5}$$

Moreover, if the presymplectic 2-form Ω is exact, that is, $\Omega = d\alpha$, we say that f is an exact presymplectic diffeomorphism if there exists a function $\mathcal{S}$ such that

$$f^*\alpha = \alpha + \mathrm{d}\mathcal{S}, \tag{2.6}$$

which also equals that if at the level of de Rham cohomology one has :

$$[f^*\alpha - \alpha] = 0. \tag{2.7}$$

Note that, since $\mathrm{d}(f^*\alpha - \alpha) = f^*\Omega - \Omega$, an exact presymplectic diffeomorphism is a presymplectic diffeomorphism.

Definition 2.4 An $(m + l)$-parametric family of presymplectic (resp. exact presymplectic) diffeomorphisms f_λ is a function

$$f : M \times B \to M, \quad B \subset \mathbb{R}^{m+l},$$

such that for each $x \in M$, the map $f(x, \lambda)$ is C^2 with respect to the variable λ varying the B and for each $\lambda \in B$, the map $f_\lambda := f(\cdot, \lambda)$ is a real analytic presymplectic (resp. exact presymplectic) diffeomorphism. We also assume that f and its derivatives with respect to λ extend to a complex neighborhood.

In our paper, we assume that the presymplectic form Ω is exact. In many other papers, it is also assumed the diffeomorphism is exact. Since we will dealing with a family of parameterized presymplectic diffeomorphisms f_λ, we do not assume they are exact, but we will consider a larger dimension of parameters, which is $(m+l)$-dimension. See Remark 3.5.

2.6 Invariant Tori

We say that $K : \mathbb{T}^{m+l} \to M$ is an invariant torus with frequency $\omega \in \mathbb{R}^{m+l}$ for a map $f : M \to M$, if

$$f(K(\theta)) - K(\theta + \omega) = 0, \quad \forall\, \theta \in \mathbb{T}^{m+l}, \tag{2.8}$$

and we say that f has an approximate invariant torus K with frequency ω if

$$f(K(\theta)) - K(\theta + \omega) = e(\theta), \quad \forall\, \theta \in \mathbb{T}^{m+l}. \tag{2.9}$$

As we will see, it will be hard to find invariant tori of frequency ω for a fixed map. A natural problem [26–28] is to consider families and to search at the same time for the embedding of the torus and the parameter of the family.

Given a family $f_\lambda : M \to M$, we will seek a parameter λ and an embedding K in such a way that

$$f_\lambda(K(\theta)) - K(\theta + \omega) = 0, \quad \forall\, \theta \in \mathbb{T}^{m+l}. \tag{2.10}$$

The Eq. (2.10) for (λ, K) will be the centerpiece of our analysis. We will develop an iterative procedure that starting from an approximate solution, (i.e a parameter λ and a embedding K so that (2.10) has a small right hand side) if applied repeatedly, produces a true solution. This improvement algorithm will require to include assumptions about the linearized behavior that we will discuss in the following sections. The notions of norms and spaces are discussed in Sect. 2.3.

2.7 Approximately Whiskered Invariant Tori

Definition 2.5 Let f_λ be a presymplectic diffeomorphism on M and K be an embedding which satisfies (2.9). We say that K is an approximately invariant whiskered torus if the followings hold.

(H1) Spectral condition: the tangent space $T_{K(\theta)}M$ has an invariant splitting for all $\theta \in \mathbb{T}^{m+l}$,

$$T_{K(\theta)}M = \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)} \oplus \mathcal{E}^c_{K(\theta)}, \tag{2.11}$$

where $\mathcal{E}^s_{K(\theta)}$, $\mathcal{E}^u_{K(\theta)}$, $\mathcal{E}^c_{K(\theta)}$ are the stable, unstable, center invariant spaces, respectively.
Moreover, the splitting (2.11) is characterized by asymptotic growth conditions, that is, there exist constants $0 < \mu_1, \mu_2 < 1,\ \mu_3 > 1$ such that $\mu_1\mu_3 < 1,\ \mu_2\mu_3 < 1$ and $C_h > 0$ such that

$$\begin{aligned} &\nu \in \mathcal{E}^s_{K(\theta)} \iff \forall n \in \mathbb{N}, \\ &\quad \|(\mathrm{D}f)(K) \circ T_\omega^{n-1}(\theta) \times \cdots \times (\mathrm{D}f)(K(\theta))\nu\| \le C_h\mu_1^n\|\nu\|, \end{aligned} \tag{2.12}$$

$$\begin{aligned} &\nu \in \mathcal{E}^u_{K(\theta)} \iff \forall n \in \mathbb{N}, \\ &\quad \|(\mathrm{D}f)^{-1}(K) \circ T_\omega^{-(n-1)}(\theta) \times \cdots \times (\mathrm{D}f)^{-1}(K(\theta))\nu\| \le C_h\mu_2^n\|\nu\|, \end{aligned} \tag{2.13}$$

and

$$\begin{aligned} &\nu \in \mathcal{E}^c_{K(\theta)} \iff \forall n \in \mathbb{N}, \\ &\|(\mathrm{D}f)(K) \circ T_\omega^{n-1}(\theta) \times \cdots \times (\mathrm{D}f)(K(\theta))\nu\| \le C_h\mu_3^n\|\nu\|, \\ &\|(\mathrm{D}f)^{-1}(K) \circ T_\omega^{-(n-1)}(\theta) \times \cdots \times (\mathrm{D}f)^{-1}(K(\theta))\nu\| \le C_h\mu_3^n\|\nu\|. \end{aligned} \tag{2.14}$$

(H2) We assume that the dimension of the center subspace is $2m + l$, that is, the torus is as hyperbolic as allowed by the presymplectic structure and there are no elliptic directions in the normal direction.

Remark 2.1 Note that, if K satisfies (2.10), that is, K is an invariant torus for some λ, the factor $\mathrm{D}K \circ T_\omega^m$ is just $\mathrm{D}f_\lambda^m \circ K$. The definition as it is written makes sense even when K is approximately invariant as (2.9) in an Euclidean manifold. Moreover, if $(\lambda, K(\theta))$ satisfies (2.10), the following holds,

$$Df_\lambda(K(\theta))\mathcal{E}^{s,u,c}_{K(\theta)} = \mathcal{E}^{s,u,c}_{K(\theta+\omega)}. \tag{2.15}$$

Remark 2.2 A crucial result later will be the fact that assuming $\mathcal{E}^{s,u,c}_{K(\theta)}$ are analytic splitting as in (H1) and

$$dist_\rho(Df_\lambda(K(\theta))\mathcal{E}^{s,u,c}_{K(\theta)}, \mathcal{E}^{s,u,c}_{K(\theta+\omega)}) \le \tilde{\delta}, \quad \theta \in U_\rho$$

for some sufficiently small $\tilde{\delta}$, where $dist_\rho$ stands for the infimum of the distances when θ varies in U_ρ, then there is an analytic invariant splitting satisfying (2.15) close to the approximately invariant splitting. Results of this type appear as Proposition 5.2 in [20]. Much more detailed and quantitative versions appear in [7]. We note that the proof does not rely on geometry, only on hyperbolicity. See more details in Sect. 5.5.

Remark 2.3 Defining

$$C^n(\theta) = \mathrm{D}f(K(\theta + (n-1)\omega)) \times \cdots \times \mathrm{D}f(K(\theta)),$$

then $C^n(\theta)$ satisfies

$$C^{n+m}(\theta) = C^n(\theta + m\omega)C^m(\theta). \tag{2.16}$$

The identity (2.16) is usually referred as saying C is a "cocycle" over the rotation by ω. This property will be used to solve the Eq. (2.10) on the hyperbolic subspace in Sect. 5.4.

Remark 2.4 Given a splitting as in (2.11), we find that it is convenient to introduce the projections $\Pi^s_{K(\theta)}$, $\Pi^u_{K(\theta)}$, $\Pi^c_{K(\theta)}$ which take a vector in $T_{K(\theta)}M$ and assign its components into each of the sub-bundles. Note that each projection depends on the whole splitting, not just on its range.

Remark 2.5 The invariance of the bundles expressed in terms of the projections means

$$\Pi^{s,u,c}_{K\circ T_\omega(\theta)} \mathrm{D}f(K(\theta)) = \mathrm{D}f(K(\theta))\Pi^{s,u,c}_{K(\theta)}, \tag{2.17}$$

which is not a geometrically natural equation except when K is invariant, that is $K \circ T_\omega(\theta) = f(K(\theta))$. The geometrically natural equation would be

$$\Pi^{s,u,c}_{f\circ K(\theta)} \mathrm{D}f(K(\theta)) = \mathrm{D}f(K(\theta))\Pi^{s,u,c}_{K(\theta)}. \tag{2.18}$$

One can find that $\mathrm{D}f(K)$ sends $T_{K(\theta)}M$ to $T_{f(K(\theta))}M$. The condition (2.18) agrees with (2.17) only when K is an invariant torus. For our purposes, the Eq. (2.17) is relevant. This is related to the fact that for us, the relevant cocycle (2.16) is not geometrically natural either. In our formulation we use the fact that the phase space is Euclidean, so that we can identify tangent spaces at different points. Formulating the results in general manifolds will require slight modifications of the treatment such as using connectors. The modifications needed are mathematically straightforward, but typographically awkward.

We anticipate that, in the main result we will impose that the dynamical spaces we have included, have certain relation with the kernel of the form Ω.

2.8 Lagrangian Properties of Invariant Tori

Due to the Diophantine condition on ω and the exactness of the presymplectic form Ω, we have the following lemmas:

Lemma 2.2 *Assume $K(\theta) \in \mathcal{P}_\rho$ and $K(\theta)$ satisfies (2.8), ω is rationally independent, then $K^*\Omega$ is identically zero.*

Proof Using the identity for pullback and the fact that $f^*\Omega = \Omega$ since f is a presymplectic diffeomorphism, we have

$$(f \circ K)^*\Omega = K^* \circ f^*\Omega = K^*\Omega,$$
$$(K \circ T_\omega)^*\Omega = T^*_\omega \circ K^*\Omega.$$

Together with $K(\theta)$ satisfying (2.8), then

$$K^*\Omega = T^*_\omega \circ K^*\Omega.$$

Since ω is rationally independent, $K^*\Omega$ is constant. Moreover, if we write $K^*\Omega$ in matrix form, we have

$$K^*\Omega(\xi, \eta) = \langle \xi, L(\theta)\eta \rangle, \tag{2.19}$$

where $L(\theta)$ is constant matrix. Then we will prove $K^*\Omega$ is identically zero.

Since the 2-form Ω is exact, that is, $\Omega = \mathrm{d}\alpha$, where $\alpha(u) = a(u)\mathrm{d}u$. Then we have the following expression

$$(K^*\alpha) = \sum_{i=1}^{m+l} C_i(\theta)\mathrm{d}\theta_i$$

with components $C_i(\theta) = [\mathrm{D}(K(\theta))a(K(\theta))]_i$, $i = 1, 2, \ldots, m + l$. It follows that

$$L(\theta) = \mathrm{D}C^\top(\theta) - \mathrm{D}C(\theta).$$

Moreover, we obtain that $\mathrm{avg}(L(\theta)) = 0$. Together with the fact $L(\theta)$ is constant matrix, we proved that $L(\theta) = 0$, that is, $K^*\Omega$ is identically zero. □

The following lemma is an analogous of a result found in [1].

Lemma 2.3 *Let $f_\lambda : M \to M$ be a presymplectic analytic diffeomorphism with fixed $\lambda \in B$ and $K \in \mathcal{P}_\rho$ be an approximate invariant torus with frequency $\omega \in D(\gamma, \sigma)$, that is, the pair $(f_\lambda,\ K(\theta))$ satisfies (2.9). Assume f_λ extends holomorphically to some complex neighborhood of the image of U_ρ under K:*

$$\mathcal{B}_r = \{z \in \mathbb{C}^{2d+2m+l} : \sup_{\theta \in U_\rho} |z - K(\theta)| < r\}.$$

Then, there exists a constant $c_0 > 0$, depending on $l, d, m, \sigma, \rho, \|DK\|_\rho, |f_\lambda|_{C^1,\mathfrak{B}_r}$, such that for $0 < \delta < \frac{\rho}{2}$

$$\|L\|_{\rho-2\delta} \leq c_0\gamma^{-1}\delta^{-\sigma-1}\|e\|_\rho, \tag{2.20}$$

where L is the matrix representing the pullback form $K^*\Omega$.

The reason for Lemma 2.3 is that if $f_\lambda \circ K = K \circ T_\omega + e$ with e being a small error, then denote by

$$\tilde{e} := K^*\Omega - T_\omega^* \circ K^*\Omega = K^* f_\lambda^* - T_\omega^* \circ K^*\Omega,$$

where $\tilde{e}$ can be estimated by e and its derivative. Using the result in Lemma [29], we conclude that $K^*\Omega$ is a constant plus a term whose size can be estimated by $\tilde{e}$. Besides the fact that, if Ω is exact and ω is rationally independent, the average of $K^*\Omega$ over the torus is zero. Consequently, we conclude that $K^*\Omega$ is bounded by a term whose size can be estimated by e. The formal proof can be found in [1]. Similar results happen in the symplectic or conformally symplectic contexts as well. See [18].

3 Statement of the Results

In this section, we will formulate our main result after introducing some notations and definitions of non-degenerate torus.

For the sake of convenience, we will introduce matrix notations for linear operations in the following discussion and we will take advantage of the fact that our phase space is an Euclidean manifold. We denote by

$$V := \{(u, (x, y, 0)) \in TM \ : \ u \in M, \ x \in \mathbb{R}^{2d}, \ y \in \mathbb{R}^{2m}\},$$
$$N := \text{Ker } (\Omega_{K(\theta)}) = \{(u, (0, 0, z)) \in TM \ : \ u \in M, \ z \in \mathbb{R}^{l}\}. \tag{3.1}$$

We will assume that $V \oplus N$ is the center space, it follows that

$$T_{K(\theta)}M = \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)} \oplus V \oplus N. \tag{3.2}$$

As it will be specified in the assumption (A1) in Theorem 1 below, we assume that we have chosen a metric in which the subspaces V and N are orthogonal and the center subspace $\mathcal{E}^c_{K(\theta)} = V \oplus N$ is trivial. As it was mentioned in Definition 2.5, the dimension of $\mathcal{E}^c_{K(\theta)}$ is $2m + l$, hence $\mathcal{E}^c_{K(\theta)} \sim \mathbb{R}^{2m+l}$.

Moreover, for each $\xi, \eta \in \mathcal{E}^c_{K(\theta)}$, we have the linear map $J^{cc}(K(\theta)) : \ \mathcal{E}^c_{K(\theta)} \to \mathcal{E}^c_{K(\theta)}$ defined by

$$\Omega^c_{K(\theta)}(\xi, \eta) := \langle \xi, J^{cc}(K(\theta))\eta \rangle, \tag{3.3}$$

where $\Omega^c_{K(\theta)}$ is the restriction of the presymplectic form on the center subspace,

$$J^{cc}(K(\theta)) = \begin{pmatrix} J(K(\theta)) & 0 \\ 0 & 0 \end{pmatrix},$$

and $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^{2m+l}$. The skew-symmetry of Ω^c yields that $J^\top = -J$. It is important to notice that the matrix J^{cc} also depends on the metric in the center subspace, as indicated in (3.2). But we do not need to keep track of the basis and metric for the hyperbolic subspace. Because of our assumption that the rank of the presymplectic form is constant, we have that the rank of J^{cc} is constant.

For $K(\theta) \in \mathcal{P}_\rho$ satisfying (2.8) or (2.10) for some fixed λ, we denote by

$$DK(\theta) = (X(\theta), Z(\theta)), \tag{3.4}$$

where $X(\theta), \ Z(\theta)$ are the first m and last l columns of $DK(\theta)$, $X^c_V(\theta), \ Z^c_V(\theta)$ are projections of $X(\theta), \ Z(\theta)$ on subspace V and $X^c_N(\theta), \ Z^c_N(\theta)$ are projections on subspace N, respectively. Then we have the following definitions.

Definition 3.1 We say that $K(\theta) \in \mathcal{P}_\rho$ is a non-degenerate torus if

(1) There exists an $m \times m$ matrix valued function $A(\theta)$ that satisfies the following relation:

$$A(\theta)(X^c_V(\theta))^\top X^c_V(\theta) = I_m,$$

where $X^c_V(\theta)$ are the projections of $X(\theta)$ on the subspace V.

(2) The matrix $B(\theta)$, which will be specifically defined in (4.12), is invertible.

To formulate the non-degeneracy condition for a pair $(f_\lambda, K(\theta))$, we denote the $(2m + l) \times (m + l)$ matrix $\Lambda(\theta)$ as follows:

$$\Lambda(\theta) := B^{-1}(\theta)Q(\theta)\left[\Pi^c_{K(\theta+\omega)}\frac{\partial f_\lambda}{\partial \lambda}(K(\theta))\right] := \begin{pmatrix} \Lambda_{11}(\theta) & \Lambda_{12}(\theta) \\ \Lambda_{21}(\theta) & \Lambda_{22}(\theta) \\ \Lambda_{31}(\theta) & \Lambda_{32}(\theta) \end{pmatrix}, \tag{3.5}$$

where the sub-matrices $\Lambda_{11}, \ \Lambda_{12}, \ \Lambda_{21}, \ \Lambda_{22}, \ \Lambda_{31}, \ \Lambda_{32}$ are of the sizes $m \times m$, $m \times l$, $m \times m$, $m \times l$, $l \times m$, $l \times l$, respectively. The matrices B^{-1} and Q are defined as (4.20), (4.15).

Definition 3.2 We say a pair $(f_\lambda, K(\theta))$ is non-degenerate at λ_0, if f_λ is an $(m+l)$-parametric family of presymplectic diffeomorphisms, $K(\theta) \in \mathcal{P}_\rho$ is a non-degenerate torus, and the $(2m+l) \times (2m+l)$ matrix

$$\Theta := \text{avg}\begin{pmatrix} \Lambda_{11}(\theta) & \Lambda_{12}(\theta) & S(\theta) \\ \Lambda_{21}(\theta) & \Lambda_{22}(\theta) & 0 \\ \Lambda_{31}(\theta) & \Lambda_{32}(\theta) & \bar{A}(\theta) \end{pmatrix} \tag{3.6}$$

is invertible, where $\bar{A}(\theta)$ will be defined in (4.18), $S(\theta)$ is the last expression in (5.11).

Remark 3.1 Note that the matrix $A(\theta)$ and $\bar{A}(\theta)$ have no connection with each other.

Remark 3.2 The role of Definition 3.2 is to reduce the Newton equation to a constant coefficient equation up to a small error. Even if we will not give the formula for the matrices $B,\ Q,\ \bar{A},\ S$ now and we postpone till they are motivated, we note that they are explicit matrices obtained from the approximate solution (K, λ), by taking derivatives, performing algebraic operations and averaging. Therefore, the assumptions on matrices $B,\ Q,\ \bar{A},\ S$ are explicit assumptions on the approximate solution and do not involve any global hypothesis about the map.

Now, we formulate our main theorem of the present paper:

Theorem 1 *Let $\omega \in D(\gamma, \sigma)$, let f_λ be the (m+l)-parametric family of analytic presymplectic diffeomorphisms, defined as Definition 2.4 and let $K_0 \in \mathcal{P}_{\rho_0}$.*

Assume that:

(A1) We have chosen a metric in which the subspaces V and N are orthogonal and the dimension of the center subspace $\mathcal{E}^c_{K(\theta)}$ is 2m+l.
(A2) The pair (f_λ, K_0) is non-degenerate at the point $\lambda = \lambda_0$ in the sense of Definition 3.2.
(A3) The family f_λ can be holomorphically extended to some complex neighborhood of the image of U_ρ under $K(\theta)$:

$$\mathcal{B}_r = \{z \in \mathbb{C} : \ \sup |z - K(\theta)| < r\}$$

such that $\|f_\lambda\|_{C^2,\ \mathcal{B}_r}$ is finite.

Denote the error term $e_0(\theta)$ as

$$e_0(\theta) := f_{\lambda_0}(K_0(\theta)) - K_0(\theta + \omega). \tag{3.7}$$

Then, there exists a positive constant $C_T > 0$, depending on $\sigma, m, l, d, \rho_0, r, \tilde{\delta}$, $\|f_\lambda\|_{C^2,\ \mathcal{B}_r}$, $\|\mathrm{D}K_0\|_{\rho_0}$, $\|A_0(\theta)\|_{\rho_0}$, $\|\frac{\partial f_\lambda}{\partial \lambda}|_{\lambda=\lambda_0}(K_0(\theta))\|_{\rho_0}$, $|\Theta^{-1}|$, $\|(\mathrm{D}K_0^\top \mathrm{D}K_0)^{-1}\|_{\rho_0}$, $\|\Pi^{s,u,c}_{K_0(\theta+\omega)}\|_{\rho_0}$, such that if $0 < \delta_0 \le \min\{1, \ \frac{\rho_0}{12}\}$ and

$$\|e_0\|_{\rho_0} < \min\left[\frac{1}{2C_T}\gamma^4\delta_0^{4\sigma},\ \frac{r}{C_T}\gamma^2\delta_0^{2\sigma}\right], \tag{3.8}$$

then there exists a mapping $K_\infty \in \mathcal{P}_{\rho_0 - 6\delta_0}$ and a vector $\lambda_\infty \in \mathbb{R}^{m+l}$ satisfying

$$f_{\lambda_\infty} \circ K_\infty = K_\infty \circ T_\omega. \tag{3.9}$$

Moreover, the following estimates hold:

$$\|K_\infty - K_0\|_{\rho_0 - 6\delta_0} < C_T\gamma^{-2}\delta_0^{-2\sigma}\|e_0\|_{\rho_0}, \tag{3.10}$$

$$|\lambda_\infty - \lambda_0| < C_T\gamma^{-2}\delta_0^{-2\sigma}\|e_0\|_{\rho_0}. \tag{3.11}$$

Remark 3.3 The assumption (A1) has the geometric meaning that the kernel of Ω—which is an invariant space under a presymplectic diffeomorphism by Lemma 4.2—is contained in the center. Equivalently, we assume that the vectors in the kernel satisfy the growth conditions (2.14). Due to the stability properties of the hyperbolic splittings, if the kernel vectors satisfy (2.14), they will satisfy similar hypothesis if we change slightly the map or the splitting. In this paper, we will consider the case when the kernel is a factor and the dynamics on it is a rotation, which is stronger than (A1).

Once we have that the kernel is contained in the center, it is standard in bundle theory that we can define a metric which makes the two bundles orthogonal.

We note that these assumptions, even if non-trivial,[1] remain valid under perturbations and hold in systems that appear in applications. The triviality of bundles assumed here enters into the study of cocycle equations using Fourier series.

Remark 3.4 One of the features of the method in the present paper is that it applies even when the stable and unstable bundles are non-trivial. We note that to solve the linearized invariance equations, we just project on the stable (unstable) and center bundles. Then, the components along the stable and unstable directions are solved by very simple iterative formulae (5.23) and (5.24) which make sense even for non-trivial bundles.

In contrast, the treatment of the center projection of the linearized equations relies on Fourier methods and on identifications which only make sense for trivial bundles. The paper [19] includes examples of whiskered tori in symplectic systems with non-trivial stable (unstable) bundles. One important result of [7] is that the center bundle of conformally symplectic whiskered tori is trivial in the sense of bundle theory, i.e., it is a product bundle.

Remark 3.5 We remark that the m components of parameters λ that we use have a geometric meaning to adjust the averages of the presymplectic conjugate of the tangent to the torus.

We can eliminate the use of these m parameters if we add to the non-degeneracy Definition 3.2 the assumption that m components of the cohomology of the presymplectic form vanish, corresponding to the directions of the embedding of the torus.

Of course, the above assumption is implied by the more natural assumption that the presymplectic diffeomorphism f_λ is exact, which is defined as in Definition 2.3. It means that all components of the cohomology of the presymplectic form vanish, which is satisfied in the mechanically constrained systems.

The proof of the result relies on a vanishing lemma as in [1] and reformulating the non-degeneracy assumptions. We will not do this in this paper.

3.1 Local Uniqueness

In this subsection, we claim the local uniqueness of the embedding K_∞ and the parameter λ_∞ provided in Theorem 1. Note that, if K_∞ is the solution of (3.9) for some λ_∞, then for every $\psi \in \mathbb{T}^m \times \mathbb{T}^l$, the solution $\hat{K}_\infty(\theta) := K_\infty(\theta + \psi)$ is also a solution for (3.9). Hence, we consider that the solution K_∞ and $\hat{K}_\infty$ are equivalent. As a result, we mean uniqueness up to this equivalence relation.

[1] A easy example when the assumption A1) does not hold is a product, $M = A \times \mathbb{T}^l$, A is a symplectic manifold. When $f(a, b) = (g(a), h(b))$, $g(0) = 0$ and g preserves the symplectic form in A. In this case, M could be a manifold with presymplectic form. Then $0 \times \mathbb{T}^l$ is an invariant torus of the presymplectic form in M. Depending on the dynamics of h, the kernel may contain (un)stable directions.

Theorem 2 *Let $\omega \in D(\gamma, \sigma)$ and assume that (f_{λ_1}, K_1), (f_{λ_2}, K_2) represent non-degenerate pairs satisfying the hypotheses of Theorem 1 and*

$$f_{\lambda_1}(K_1(\theta)) - K_1(\theta + \omega) = 0,$$
$$f_{\lambda_2}(K_2(\theta)) - K_2(\theta + \omega) = 0,$$

such that $K_1(U_\rho) \subset \mathcal{B}_r$ and $K_2(U_\rho) \subset \mathcal{B}_r$. Then, there exists a positive constant $\hat{C}_T$ depending on σ,m,l,d,ρ,r, $\|f_\lambda\|_{C^2,\ \mathcal{B}_r}$, $\|DK_1\|_\rho$, $\|A(\theta)\|_\rho$,$\|\Pi^{s,u,c}_{K_1(\theta)}\|_\rho$, $|\Theta^{-1}|$, such that

$$\|K_1 - K_2\|_\rho < \hat{C}_T \gamma^2 \delta^{2\sigma}, \quad |\lambda_1 - \lambda_2| \leq \hat{C}_T \gamma^2 \delta^{2\sigma} \tag{3.12}$$

with $\delta = \frac{\rho}{8}$. Furthermore, there exists an initial phase $\tau_\infty \in \mathbb{T}^m \times \mathbb{T}^l$ such that in $U_{\rho/2}$

$$K_1 \circ T_{\tau_\infty} = K_2,$$
$$\lambda_1 = \lambda_2.$$

The proof will be given in Sect. 7. We anticipate the idea is that, since the proof of Theorem 1 is based on a quasi-Newton method, it suffices to study the iterative step and see that its only neutral directions are in the directions of change of phase. A related argument appears in [18].

4 Some Results in Presymplectic Geometry

In this section, we will establish some results in presymplectic geometry, which we will use later.

Lemma 4.1 *Given $K(\theta)$ that satisfies (2.8), we deduce that $\Omega_{K(\theta)}(u, v) = 0$ in any of the following case:*

(1) $u, v \in \mathcal{E}^s_{K(\theta)}$,
(2) $u, v \in \mathcal{E}^u_{K(\theta)}$,
(3) $u \in \mathcal{E}^c_{K(\theta)}$, $\ v \in \mathcal{E}^s_{K(\theta)} \bigcup \mathcal{E}^u_{K(\theta)}$,
(4) $v \in \mathcal{E}^c_{K(\theta)}$, $\ u \in \mathcal{E}^s_{K(\theta)} \bigcup \mathcal{E}^u_{K(\theta)}$.

Proof From

$$\Omega_{K(\theta)}(u, v) = \Omega_{f^n(K(\theta))}(Df^n(K(\theta))u, Df^n(K(\theta))v), \forall n \in \mathbb{Z}, \tag{4.1}$$

we see that, for the case (1), using the contraction property (2.12), we can make the term $\Omega_{K(\theta)}(u, v)$ as small as desired. Then, sending $n \to \infty$, we prove that

$$\Omega_{K(\theta)}(u, v) = 0, \quad \forall u, v \in \mathcal{E}^s_{K(\theta)}.$$

All the other cases can be proved in a similar way. It suffices to use (4.1) and make n tend to $\pm\infty$. The differences in the rates of growth, shows that there is always a choice of sign so that the right hand side of (4.1) goes to zero. □

Lemma 4.2 *Given the 2-form Ω defined as above, $K(\theta)$ satisfying (2.10) and the spectral condition (H1), (H2), we have*

$$\text{Ker}\ \left(\Omega_{K(\theta)}\right) \bigcap \mathcal{E}^c_{K(\theta)} = \text{Ker}\ \left(\Omega^c_{K(\theta)}\right),$$

where $\Omega^c_{K(\theta)}$ is defined as in (3.3) and

$$\text{Ker}\ \left(\Omega^c_{K(\theta)}\right) = \{w \in \mathcal{E}^c_{K(\theta)} |\ \Omega^c_{K(\theta)}(w, \mu) = 0,\ \forall \mu \in \mathcal{E}^c_{K(\theta)}\}. \tag{4.2}$$

Moreover, we have

$$f_*(\text{Ker}\ \Omega_{K(\theta)}) = \text{Ker}\ \Omega_{(f(K(\theta)))}. \tag{4.3}$$

Proof Given any $v \in \text{Ker}\ \left(\Omega^c_{K(\theta)}\right)$, then $v \in \mathcal{E}^c_{K(\theta)}$ and for any $w \in \mathbb{R}^n$, we denote $w = w^s + w^u + w^c$, where $w^s \in \mathcal{E}^s_{K(\theta)},\ w^u \in \mathcal{E}^u_{K(\theta)},\ w^c \in \mathcal{E}^c_{K(\theta)}$. By Lemma 4.1 and (4.2), we have

$$\begin{aligned}\Omega_{K(\theta)}(v, w) &= \Omega_{K(\theta)}(v, w^s) + \Omega_{K(\theta)}(v, w^u) + \Omega^c_{K(\theta)}(v, w^c)\\ &= 0,\end{aligned}$$

that is, $v \in \text{Ker}\left(\Omega_{K(\theta)}\right) \cap \mathcal{E}^c_{K(\theta)}$. The opposite is obvious.

As for (4.3), given any $v \in \text{Ker}\ \Omega_{K(\theta)}$, then, for any $w \in T_{K(\theta)}M,\ \ \Omega(v, w) = 0$. It also implies that

$$\Omega(f_*v, f_*w) = 0,\ \ \forall w \in T_{K(\theta)}M \Leftrightarrow \Omega(f_*v, \tilde{u}) = 0,\ \ \forall\ \tilde{u} \in T_{f(K(\theta))}M,$$

that is,

$$f_*v \in \text{Ker}\ \Omega_{(f(K(\theta)))}.$$

□

4.1 A Useful Basis for the Center Subspace

In this subsection, we will first find a useful basis for the center subspace $\mathcal{E}^c_{K(\theta)}$ in the case that $K(\theta)$ is a solution for (2.10). A similar basis will be constructed later for an approximately invariant solution, i.e., $K(\theta)$ satisfies (2.9) for some f_λ. For the Newton method, the latter is the one that will be useful, but the basis in the invariant case is geometrically natural.

Remember that, for $K(\theta) \in \mathcal{P}_\rho$, we decompose the Jacobian matrix in the form

$$DK(\theta) = (X(\theta), Z(\theta)), \tag{4.4}$$

where $X(\theta),\ Z(\theta)$ are the first m and last l columns of $DK(\theta)$. Denote by $X^c_V(\theta),\ Z^c_V(\theta)$ the projections of $X(\theta),\ Z(\theta)$ on the subspace V and by $X^c_N(\theta),\ Z^c_N(\theta)$ the projections on subspace N.

Assume that $K(\theta)$ satisfies (2.10) for some f_λ and there exists an $m \times m$ matrix $A(\theta)$ such that

$$A(\theta)(X^{c\top}_V(\theta) \cdot X^c_V(\theta)) = I_m. \tag{4.5}$$

The formula (4.5) is an important non-degeneracy assumption to solve the Eq. (5.3) on the center subspace. Roughly, it says that the matrix X^c_V has maximal rank. Also, denote

$$Y^c_V(\theta) := X^c_V(\theta)A(\theta)\ \ \text{and}\ \ Y^c(\theta) = (J^{-1}(K(\theta))Y^c_V(\theta),\ 0)^\top. \tag{4.6}$$

Note that the span of the columns $X^c_V(\theta)$ is the same as the span of $Y^c_V(\theta)$, but we use $A(\theta)$ as a convenient normalization. Then, the following matrix denoted by $\mathcal{M}$, will provide us with a linear change of variables which simplifies the Eq. (5.3) projected on the center space. That is,

$$\mathcal{M}(\theta) = \begin{pmatrix} X_V^c(\theta) & J^{-1}(K(\theta))Y_V^c(\theta) & Z_V^c(\theta) \\ X_N^c(\theta) & 0 & Z_N^c(\theta) \end{pmatrix}. \tag{4.7}$$

By the non-degeneracy assumption (4.5) and Lemma 2.2, we obtain that

$$\Omega^c_{K(\theta)}(X^c(\theta), Y^c(\theta)) = I_m, \tag{4.8}$$

$$\Omega^c_{K(\theta)}(X^c(\theta), X^c(\theta)) = 0, \tag{4.9}$$

$$\Omega^c_{K(\theta)}(X^c(\theta), Z^c(\theta)) = 0, \tag{4.10}$$

where $X^c(\theta),\ Y^c(\theta),\ Z^c(\theta)$ denote the three columns of $\mathcal{M}(\theta)$ respectively.

We will now check that the first $2m$-columns of $\mathcal{M}(\theta)$ are linearly independent. We will procede by contradiction. Assume that there exist $\alpha_j,\ \beta_j,\ j = 1, \ldots, m$, such that the linear combination L is equal to zero,

$$L := \sum_{j=1}^m \alpha_j X^c(\theta)e_j + \sum_{j=1}^m \beta_j Y^c(\theta)e_j = 0. \tag{4.11}$$

We will prove that $\alpha_j = \beta_j = 0$ for all $j = 1, \ldots, m$. This establishes that the first $2m$-columns of $\mathcal{M}(\theta)$ are linearly independent.

First, for any $1 \le k \le m$, we have due to (4.8), (4.9), (4.10)

$$\begin{aligned} 0 = \Omega^c(X^c(\theta)e_k, L) &= \sum_{j=1}^l \beta_j e_k^\top (X^c(\theta))^\top J(K(\theta))Y^c(\theta)e_j \\ &= \sum_{j=1}^l \beta_j \langle e_k, e_j\rangle = \beta_k. \end{aligned}$$

Once we have that $\beta_j = 0$, the linear combination (4.11) reduces to

$$L := \sum_{j=1}^m \alpha_j X^c(\theta)e_j = 0.$$

Again for any $1 \le k \le m$, we have

$$0 = \Omega^c(Y^c(\theta)e_k, L) = \alpha_k,$$

that is, $\alpha_k = 0$, for all $1 \le k \le m$. It shows that first $2m$-columns of $\mathcal{M}(\theta)$ are linearly independent.

Lemma 4.3 *The columns of $\mathcal{M}(\theta)$ form a basis for $\mathcal{E}^c_{K(\theta)}$ provided the matrix*

$$B(\theta) = \begin{pmatrix} 0 & I_m & 0 \\ -I_m & (J^{-1}(K)Y_V^c)^\top Y_V^c(\theta) & (J^{-1}(K)Y_V^c)^\top J(K)Z_V^c(\theta) \\ X_N^c(\theta) & 0 & Z_N^c(\theta) \end{pmatrix} \tag{4.12}$$

is invertible. In this case, we have

$$\begin{aligned} \mathrm{D}f_\lambda^c(K(\theta))(X^c(\theta), Z^c(\theta)) &= (X^c(\theta+\omega),\ Z^c(\theta+\omega)), \\ \mathrm{D}f_\lambda^c(K(\theta))(Y^c(\theta)) &= X^c(\theta+\omega)S_1(\theta) + Y^c(\theta+\omega)I_m \\ &\quad + Z^c(\theta+\omega)\bar{A}(\theta), \end{aligned} \tag{4.13}$$

where $\bar{A}(\theta)$ and $S_1(\theta)$ are matrices satisfying:

$$\mathrm{D}f_\lambda^c(K(\theta))\mathcal{M}(\theta) = \mathcal{M}(\theta+\omega)\begin{pmatrix} I_m & S_1(\theta) & 0 \\ 0 & I_m & 0 \\ 0 & \bar{A}(\theta) & I_l \end{pmatrix} \tag{4.14}$$

and $\mathrm{D}f_\lambda^c(K(\theta))$ denotes the last $(2m+l)\times(2m+l)$ block of matrix $\mathrm{D}f_\lambda(K(\theta))$.

Proof Let

$$Q(\theta) := \begin{pmatrix} X_V^c(\theta)^\top J(K(\theta)) & 0 \\ (J^{-1}(K(\theta))Y_V^c(\theta))^\top J(K(\theta)) & 0 \\ 0 & I_l \end{pmatrix}_{(2m+l)\times(2m+l)}. \tag{4.15}$$

By calculation, we have

$$Q(\theta)\mathcal{M}(\theta) = B(\theta). \tag{4.16}$$

Since we assume that the matrix B is invertible, we conclude that the matrices $Q,\ \mathcal{M}$ are both invertible and the columns of $\mathcal{M}(\theta)$ provide a basis for $\mathcal{E}^c_{K(\theta)}$. □

Since f_λ is presymplectic as (2.5), we have

$$\mathrm{D}f_\lambda^c(K(\theta))(X^c(\theta), Z^c(\theta)) = (X^c(\theta+\omega), Z^c(\theta+\omega)).$$

Then, we will find matrices $S_1(\theta),\ \bar{A}(\theta)$ satisfying the Eq. (4.14). Moving the term $Y^c(\theta+\omega)I_m$ to the left side of the second equation of (4.13), we have

$$\mathrm{D}f_\lambda^c(K(\theta))Y^c(\theta) - Y^c(\theta+\omega)I_m = X^c(\theta+\omega)S_1(\theta) + Z^c(\theta+\omega)\bar{A}(\theta). \tag{4.17}$$

Since $\mathcal{M}$ is invertible, there exists the inverse $T = (T_1, T_2, T_3)^\top$ such that $T_3X^c = 0_{l\times m},\ T_3Z^c = I_{l\times l}$. Multiplying T_3 on both sides of (4.17), we have

$$\bar{A}(\theta) = T_3(\theta+\omega)[\mathrm{D}f^c(K(\theta))Y^c(\theta) - Y^c(\theta+\omega)I_m]. \tag{4.18}$$

Then moving $Z^c(\theta+\omega)\bar{A}(\theta)$ to the left side and multiplying $X^c(\theta+\omega)^\top J(K(\theta+\omega))$ on both sides of (4.17), we have

$$\begin{aligned} &S_1(\theta) \\ &= X^c(\theta+\omega)^\top J(K(\theta+\omega))[\mathrm{D}f_\lambda^c(K(\theta))Y^c(\theta) - Y^c(\theta+\omega)I_m \\ &\quad -Z^c(\theta+\omega)\bar{A}(\theta)]. \end{aligned} \tag{4.19}$$

Remark 4.1 By a simple calculation, we have that the inverse of B has the following form:

$$B^{-1} = \begin{pmatrix} B_{11}^{-1} & B_{12}^{-1} & B_{13}^{-1} \\ I_m & 0 & 0 \\ B_{31}^{-1} & B_{32}^{-1} & B_{33}^{-1} \end{pmatrix}, \tag{4.20}$$

where B_{ij}^{-1} are matrices of appropriate dimensions and defined by (4.20). The content of the formula (4.20) is precisely the special form of the middle row. We note that the B_{ij}^{-1} are explicit formulas which involve only the projections on the bundles and algebraic operations with the derivatives of the function.

5 The Linearized Equation and the Quasi-Newton Method

In this section, we find an approximate solution of the linearized equation appearing in the Newton method. This approximate solution is the solution of a simplified equation obtained by dropping a term in the equation suggested by the Newton method. Hence, we call it the quasi-Newton method. As usual in KAM theory, the quadratic estimates for the new error hold in a slightly smaller domain, and have a factor that grows in a controlled way when the domain loss is small.

The analysis is similar to that in [19], but there is a new obstruction due to the fact that in the present equation we have to adjust parameters to take care of the directions in the kernel of Ω, which leads to a coupling among different equations. The difficulty of the correction of the parameters appears also in [5] in the study of Lagrangian tori in conformally symplectic systems and in the study of whiskered tori in conformally symplectic systems [6,7].

Now we give a more detailed sketch of the proof that can serve as a reading guide. We begin with a pair $(f_\lambda, K(\theta))$ satisfying (A1)–(A3) at some $\lambda = \bar{\lambda}$ as in Theorem 1 and define

$$e(\theta) := G(K,\lambda) := f_\lambda(K(\theta)) - K(\theta+\omega)$$

as the error term.

The prescription of the standard Newton method is to find a correction $(\Delta(\theta),\ \varepsilon)$ at each iteration step, which satisfies the infinitesimal equation up to a quadratic error term with respect to $\|e\|$. The infinitesimal equation prescribed by Newton method is as follows:

$$\begin{aligned} \mathrm{D}G(K,\lambda)(\Delta,\varepsilon) &:= \mathrm{D}f_\lambda(K(\theta))\Delta(\theta) - \Delta(\theta+\omega) + \frac{\partial f_\lambda(K(\theta))}{\partial\lambda}\varepsilon \\ &= -e(\theta). \end{aligned} \tag{5.1}$$

Then we denote $K_+ = K + \Delta,\ \lambda_+ = \bar{\lambda} + \varepsilon$ and

$$e_+(\theta) := G(K_+,\lambda_+) := f_{\lambda_+}(K_+(\theta)) - K_+(\theta+\omega)$$

as the new error term. In Sect. 5.6, we will prove that the estimate of $\|e_+\|$ is a quadratic term with respect to $\|e\|$ in a slightly smaller domain, which completes an iterative step.

We take the projections of Eq. (5.1) on the stable, unstable and center subspaces, respectively. Then Eq. (5.1) is equivalent to the system of three equations as follows:

$$\begin{aligned} &\Pi^{s,u,c}_{K(\theta+\omega)}\mathrm{D}f_\lambda(K(\theta))\Delta(\theta) - \Pi^{s,u,c}_{K(\theta+\omega)}\Delta(\theta+\omega) + \Pi^{s,u,c}_{K(\theta+\omega)}\frac{\partial f_\lambda(K(\theta))}{\partial\lambda}\varepsilon \\ &= -\Pi^{s,u,c}_{K(\theta+\omega)}e(\theta). \end{aligned} \tag{5.2}$$

Using the invariance of vector bundles (2.17), we obtain that (5.2) can be rewritten as:

$$\begin{aligned} &\mathrm{D}f_\lambda(K(\theta))\Pi^{s,u,c}_{K(\theta)}\Delta(\theta) - \Pi^{s,u,c}_{K(\theta+\omega)}\Delta(\theta+\omega) + \Pi^{s,u,c}_{K(\theta+\omega)}\frac{\partial f_\lambda(K(\theta))}{\partial\lambda}\varepsilon \\ &= -\Pi^{s,u,c}_{K(\theta+\omega)}e(\theta). \end{aligned}$$

We introduce the notation

$$\begin{aligned} \Delta^{s,u,c}(\theta) &:= \Pi^{s,u,c}_{K(\theta)}\Delta(\theta), \\ e^{s,u,c}(\theta) &:= \Pi^{s,u,c}_{K(\theta+\omega)}e(\theta). \end{aligned}$$

Taking into account that, by the invariance of the bundles, we can consider $\mathrm{D}f_\lambda(K(\theta))$ in diagonal blocks, so we rewrite that

$$\mathrm{D}f_\lambda(K(\theta))\Pi^{s,u,c}_{K(\theta)}\Delta(\theta) := Df^{s,u,c}_\lambda(K(\theta))\Delta^{s,u,c}(\theta).$$

Using the notation above, we see that (5.1) is equivalent to the system of three equations as follows:

$$\begin{aligned} &\mathrm{D} f_\lambda^{s,u,c}(K(\theta))\Delta^{s,u,c}(\theta) - \Delta^{s,u,c}(\theta+\omega) + \Pi^{s,u,c}_{K(\theta+\omega)} \frac{\partial f_\lambda(K(\theta))}{\partial \lambda}\varepsilon \\ &= -e^{s,u,c}(\theta). \end{aligned} \tag{5.3}$$

Note that (5.3) consists of three equations but four unknowns, namely, $(\Delta^{s,u,c}(\theta), \varepsilon)$. The study of the system (5.3) will be done in three steps:

(1) In Sect. 5.1–5.2, we will use the geometric structure on the tangent bundle to find a basis for $\mathcal{E}^c_{K(\theta)}$ for the case that $K(\theta)$ is an approximate invariant torus. The basis provides us a change of variable under which we can transform the Eq. (5.3) on the center space into a constant coefficients equation up to a small error.
(2) In Sect. 5.3, we will solve the Eq. (5.3) on the center subspace. The parameter ε will be chosen so that the compatibility conditions for Δ^c are satisfied and then, we will solve the equation for Δ^c. Hence, in this step, we use one equation to find two unknowns.
(3) In Sect. 5.4, we will be able to solve the equations for $\Delta^s,\ \Delta^u$ by using the conditions on the co-cycles over Ω. (See Remark 2.3).

Remark 5.1 Note that, as in [1], we will not solve the equation for Δ^c exactly, but solve it only up to a quadratic error. The equation for Δ^c is an elaborated equation that involves small divisors and entails a loss of domain for the estimates. We also remark that all the constants appearing in the iteration step are all positive constants.

After we have solved the linearized equation with detailed estimates, we show that, indeed, the error has decreased. This requires to verify that the composition involved in the function can be performed. Note that the estimates produced for the correction depend on the properties of the splitting and on the twist condition, so we have to estimate how do they change. Since the correction of the splittings is obtained by a contraction argument and the twist is an algebraic expression of derivatives, it is quite straightforward to show that the change in the constants can be estimated by the error times a factor depending on the loss of domain.

We call attention to the fact that the solutions of the linearized equation are not unique but they have some arbitrary parameters. As it is well known, the Newton method does not need an inverse of the linearization, but just a right inverse. The proof of the local uniqueness (Theorem 2) is based on reexamining the procedure and identifying the lack of uniqueness of the right inverse as the directions of change of origin in the parameterization.

5.1 Basis for the Center Subspace When $K(\theta)$ is an Approximate Solution

We consider the linearized equation projected on the center subspace, that is,

$$\begin{aligned} &\Pi^c_{K(\theta+\omega)} \mathrm{D}G(K,\lambda)(\Delta^c,\varepsilon) \\ &:= \left[\Pi^c_{K(\theta+\omega)} \frac{\partial f_\lambda(K(\theta))}{\partial \lambda}\right]\varepsilon + \mathrm{D} f^c_\lambda(K(\theta))\Delta^c(\theta) - \Delta^c(\theta+\omega) = -e^c(\theta). \end{aligned} \tag{5.4}$$

We will find an approximate solution $(\Delta^c,\ \varepsilon)$ such that

$$\|\Pi^c_{K(\theta+\omega)} DG(K,\lambda)(\Delta^c,\varepsilon) + e^c(\theta)\|_{\rho-\delta} \le C\gamma^{-3}\delta^{-(3\sigma+1)}\|e\|^2_\rho,$$

where C is a constant to be made explicit later.

Remark 5.2 We note that the process we use to go from e to (Δ^c, ε) is linear. We just take projections (to obtain e^c, multiply by explicit expressions obtained from derivatives of the approximate solution by algebraic operations). Hence we can formulate the construction of (Δ^c, ε) as creating a linear operator. A similar construction will happen for the hyperbolic directions. Hence, the process of constructing (Δ, ε) can be described as creating an approximate right inverse on the derivative (which involves losses of domain). This is the basis of abstract implicit function theorems. One abstract implicit function theorem particularly well adapted to the situation described here is in [5].

To find (Δ^c, ε), we first prove that the columns of $\mathcal{M}(\theta)$ still consist of a basis for $\mathcal{E}^c_{K(\theta)}$ if the error term $e(\theta)$ is small enough.

Since $(\lambda, K(\theta))$ satisfies (2.9), it follows that $\Omega^c(X^c_V, X^c_V),\ \Omega^c(X^c_V, Z^c_V) \neq 0$, then the Eq. (4.16) becomes:

$$Q(\theta)\mathcal{M}(\theta) = B(\theta) + R(\theta), \tag{5.5}$$

where

$$R(\theta) = \begin{pmatrix} X^c_V(\theta)^\top J(K(\theta))X^c_{V(\theta)} & 0 & X^c_V(\theta)^\top J(K(\theta))Z^c_{V(\theta)} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Applying Lemma 2.3, we estimate the error term $R(\theta)$:

Lemma 5.1 *Assume that all the hypotheses of Lemma 2.3 hold. Then there exists a constant C_1 depending on $m,\ l,\ d,\ \rho,\ \|f_\lambda\|_{C^1,\mathcal{B}_r},\ \|J\|_{C^1,\mathcal{B}_r},\ \|DK\|_\rho$ such that for every $0 < \delta < \frac{\rho}{2}$ we have*

$$\|B^{-1}(\theta)R(\theta)\|_{\rho-2\delta} \le C_1\gamma^{-1}\delta^{-(\sigma+1)}\|e\|_\rho.$$

As an easy corollary, we have the following lemma, which will be useful to show that the hypothesis of invertibility applies.

Lemma 5.2 *Assume that all the hypotheses of Lemma 2.3 hold and the error term $e(\theta)$ satisfies*

$$C_1\gamma^{-1}\delta^{-(\sigma+1)}\|e\|_\rho \le \frac{1}{2}. \tag{5.6}$$

Then, the matrix $\mathcal{M}$ is invertible and

$$\mathcal{M}^{-1}(\theta) = B^{-1}(\theta)Q(\theta) + \tilde{B}(\theta),$$

where

$$\tilde{B}(\theta) = -[(I_{2m+l} + B^{-1}R)^{-1}B^{-1}RB^{-1}Q](\theta). \tag{5.7}$$

Moreover, we have the estimate

$$\|\tilde{B}(\theta)\|_{\rho-2\delta} \le C_2\gamma^{-1}\delta^{-(\sigma+1)}\|e\|_\rho, \tag{5.8}$$

where C_2 is a constant depending on $m,\ l,\ d,\ \rho,\ r,\ \|f_\lambda\|_{C^1,\mathcal{B}_r},\ \|J\|_{C^1,\mathcal{B}_r},\ \|A\|_\rho$ and $\|DK\|_\rho$.

Proof Rewrite (5.5) as

$$Q(\theta)\mathcal{M}(\theta) = B(\theta)(I_{2m+l} + B^{-1}(\theta)R(\theta)).$$

We now use the Neumann series and the estimate of $B^{-1}R$, we conclude that $I_{2m+l} + B^{-1}(\theta)R(\theta)$ is invertible. □

5.2 Change of Variables in the Linearized Equation

In this subsection, we will apply a change of variable to solve the Eq. (5.4) up to a quadratic error.

Since $f_\lambda(K(\theta))$ is presymplectic, together with Lemma 4.2 and the assumptions of Theorem 1, we claim that

$$\mathrm{D}f_\lambda^c(K(\theta)) = \begin{pmatrix} F_1(\theta) & 0 \\ F_2(\theta) & F_4(\theta) \end{pmatrix}. \tag{5.9}$$

Lemma 5.3 *Let $K(\theta)$ be an approximate invariant torus with error $e(\theta)$. Assume that the pair $(f_\lambda, K(\theta))$ satisfies the non-degeneracy condition in Definition 3.2 and $e(\theta)$ satisfies (5.6). Then, we take the change of variable*

$$\Delta^c(\theta) = \mathcal{M}(\theta)\xi(\theta) \tag{5.10}$$

such that the Eq. (5.4) is transformed as follows:

$$\left[\begin{pmatrix} I_m & S(\theta) & 0 \\ 0 & I_m & 0 \\ 0 & \bar{A}(\theta) & I_l \end{pmatrix} + B_1(\theta)\right]\xi(\theta) - \xi(\theta+\omega)$$
$$= -B^{-1}(\theta)Q(\theta)e^c\theta - \Lambda(\theta)\varepsilon - \tilde{B}(\theta)e^c(\theta) - \tilde{B}(\theta)\Pi^c_{K(\theta+\omega)}\frac{\partial f_\lambda(K(\theta))}{\partial\lambda}\varepsilon, \tag{5.11}$$

where

$$\begin{aligned} E(\theta) &:= \mathrm{D}f_\lambda^c(K(\theta))\mathcal{M}(\theta) - \mathcal{M}(\theta+\omega)\begin{pmatrix} I_m & S_1(\theta) & 0 \\ 0 & I_m & 0 \\ 0 & \bar{A}(\theta) & I_l \end{pmatrix} \\ &:= (D_1 e^c(\theta), E_1(\theta), D_2 e^c(\theta)), \end{aligned}$$

where $D_1e^c(\theta)$ and $D_2e^c(\theta)$ are the first m columns and the last l columns of $E(\theta)$, respectively. We also denote by

$$\begin{aligned} E_1(\theta) &:= \mathrm{D}f_\lambda^c(K(\theta))Y^c(\theta) - X^c(\theta+\omega)S_1(\theta) - Y^c(\theta+\omega)I_m - Z^c(\theta+\omega)\bar{A}(\theta), \\ B_1(\theta) &:= \mathcal{M}^{-1}(\theta+\omega)E(\theta) - \begin{pmatrix} 0 & S_2(\theta) & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ S_2(\theta) &:= B_{13}^{-1}F_2(\theta)J^{-1}(K(\theta))Y_V^c(\theta) - X_N^c(\theta+\omega)S_1(\theta) - Z_N^c(\theta+\omega)\bar{A}(\theta), \\ S(\theta) &:= S_1(\theta) + S_2(\theta), \end{aligned}$$

and $\Lambda(\theta)$, $\tilde{B}(\theta)$, $S_1(\theta)$ and $\bar{A}(\theta)$ are defined as (3.5), (5.7), (4.19), (4.18), respectively. $F_2(\theta)$ is the sub-matrix in (5.9).

Moreover, we have the following estimates:

$$\begin{aligned} \|\tilde{B}(\theta)e^c(\theta)\|_{\rho-2\delta} &\le C_2\gamma^{-1}\delta^{-(\sigma+1)}\|e\|_\rho^2, \\ \left\|\tilde{B}(\theta)\Pi^c_{K(\theta+\omega)}\frac{\partial f_\lambda(K(\theta))}{\partial\lambda}\varepsilon\right\| &\le C_2\gamma^{-1}\delta^{-(\sigma+1)}\left\|\Pi^c_{K(\theta+\omega)}\frac{\partial f_\lambda(K(\theta))}{\partial\lambda}\right\|\|\varepsilon\|\|e\|_\rho, \\ \|B_1\|_{\rho-2\delta} &\le C_3\gamma^{-1}\delta^{-(\sigma+1)}\|e\|_\rho, \end{aligned} \tag{5.12}$$

where C_2 is the same constant as in (5.8) and C_3 is another positive constant depends on the same parameters with C_2.

Proof Note that when $K(\theta)$ satisfies (2.9) for some f_λ, we have

$$\begin{aligned} &\mathrm{D}f_\lambda^c(K(\theta))\mathcal{M}(\theta) \\ &= \mathcal{M}(\theta+\omega)\begin{pmatrix} I_m & S_1(\theta) & 0 \\ 0 & I_m & 0 \\ 0 & \bar{A}(\theta) & I_l \end{pmatrix} + (D_1e^c(\theta), E_1(\theta), D_2e^c(\theta)). \end{aligned} \tag{5.13}$$

Substituting (5.13) and (5.10) into Eq. (5.4), one can verify that ξ satisfies (5.11). To prove the estimates, note that e^c is the projection of the error e under Π^c and Π^c is bounded, we have bounds for e^c similar to the bounds of e in (5.6). Then, (5.12) directly follows from Lemma 5.2 and Cauchy estimates, so that

$$\|\mathcal{M}^{-1}(\theta+\omega)D_ie^c(\theta)\|_{\rho-2\delta} \le C_4\gamma^{-1}\delta^{-(\sigma+1)}\|e\|_\rho^2,$$

where $i=1,2$, C_4 is a constant depending on the same parameters as C_2. Then, we will estimate the term $\mathcal{M}^{-1}(\theta)E_1(\theta)$ to complete the third estimate of (5.12).

Recalling that we write the inverse of $\mathcal{M}$ as $\mathcal{M}^{-1}=(T_1,T_2,T_3)^\top$ and that $T_3(\theta+\omega)X^c(\theta+\omega)=0_{l\times m}$, $T_3(\theta+\omega)Z^c(\theta+\omega)=I_{l\times l}$, it follows that $T_3(\theta+\omega)E_1(\theta)=0_{l\times m}$.

Hence, we have

$$\mathcal{M}^{-1}(\theta+\omega)E_1(\theta) = \begin{pmatrix} T_1(\theta+\omega) \\ T_2(\theta+\omega) \\ 0 \end{pmatrix} E_1(\theta). \tag{5.14}$$

Moreover, it follows from Lemma 5.2 that $\mathcal{M}^{-1}(\theta) = B^{-1}(\theta)Q(\theta)+\tilde{B}(\theta)$, where $B^{-1}(\theta)$, $Q(\theta)$ are defined as in (4.20), (4.15), respectively. In the following, we will split the right side of (5.14) into three parts, that is,

$$\begin{aligned} &\begin{pmatrix} T_1(\theta+\omega) \\ T_2(\theta+\omega) \end{pmatrix} E_1(\theta) \\ &= \hat{B}^{-1}(\theta+\omega)\hat{Q}(\theta)E_1(\theta) + \begin{pmatrix} 0 & 0 & B_{13}^{-1} \\ 0 & 0 & 0 \end{pmatrix} E_1(\theta) + \bar{B}(\theta)E_1(\theta), \end{aligned} \tag{5.15}$$

where $\bar{B}$ is the first $2m$-lines of $\tilde{B}$ and $\hat{B}^{-1}$, $\hat{Q}$ are blocks from B^{-1}, Q as follows:

$$\hat{B}^{-1} := \begin{pmatrix} B_{11}^{-1} & B_{12}^{-1} \\ I_m & 0 \end{pmatrix}, \quad \hat{Q}(\theta) := \begin{pmatrix} X_V^{c\,\top}(\theta)J(K(\theta)) & 0 \\ (J^{-1}(K(\theta))Y_V^c(\theta))^\top & 0 \end{pmatrix}.$$

It is easy to verify that

$$\|\bar{B}(\theta)E_1(\theta)\|_{\rho-2\delta} \le C_5\gamma^{-1}\delta^{-(\sigma+1)}\|e\|_\rho,$$

where C_5 depends on C_4 and $\|E_1(\theta)\|$. Note that since $\|E_1(\theta)\|$ is not bounded by the error in the invariant equation, then $\|\bar{B}(\theta)E_1(\theta)\|_{\rho-2\delta}$ is not quadratically bounded by the error. This will not affect the argument since we will use only E_1 multiplied by other quantities, which are bounded by the error.

Consider (5.9) and rewrite $E_1(\theta)$ as follows:

$$\begin{aligned} E_1(\theta) &:= \begin{pmatrix} E_1^u \\ E_1^l \end{pmatrix} \\ &:= \begin{pmatrix} F_1(\theta)J^{-1}(K(\theta))Y_V^c(\theta) - X_V^c(\theta+\omega)S_1(\theta) - J^{-1}(K(\theta+\omega))Y_V^c(\theta+\omega) - Z_V^c(\theta+\omega)\bar{A}(\theta) \\ F_2(\theta)J^{-1}(K(\theta))Y_V^c(\theta) - X_N^c(\theta+\omega)S_1(\theta) - Z_N^c(\theta+\omega)\bar{A}(\theta) \end{pmatrix}, \end{aligned}$$

where $Y^c(\theta+\omega)$ and $Y_V^c(\theta+\omega)$ are defined as in (4.6) for the variable $(\theta+\omega)$ in place of (θ). Then, the first term on the right side of (5.15) becomes:

$$\hat{B}^{-1}(\theta+\omega)\hat{Q}(\theta)E_1(\theta) = \begin{pmatrix} B_{11}^{-1}X_V^{c\top}J(K)E_1^u + B_{12}^{-1}(J^{-1}(K)Y_V^c)^\top J(K)E_1^u \\ X_V^{c\top}J(K)E_1^u \end{pmatrix}.$$

One can verify that the lower term $X_V^{c\top}J(K)E_1^u$ is identically zero and the upper term is controlled by the error by Lemma 2.1.

Consider the second term on the right side of (5.15), that is,

$$\begin{pmatrix} 0 & 0 & B_{13}^{-1} \\ 0 & 0 & 0 \end{pmatrix} E_1(\theta) = \begin{pmatrix} S_2(\theta) \\ 0 \end{pmatrix}.$$

Since this term is not controlled by the error, we subtract this term from $\mathcal{M}^{-1}(\theta+\omega)E_1(\theta)$. That is the reason why we add $S_2(\theta)$ to the term $S(\theta)$. Putting together the discussion above, we proved Lemma 5.3. □

5.3 Solving the Linearized Equation on the Center Subspace

If the term $\xi(\theta)$, ε are controlled by the error, we can omit the quadratically small terms $B_1(\theta)\xi(\theta)$, $\tilde{B}(\theta)e^c(\theta)$, $\tilde{B}(\theta)\Pi^c_{K(\theta+\omega)}\frac{\partial f_\lambda(K(\theta))}{\partial\lambda}\varepsilon$ from Eq. (5.11). Then, we have the following linearized equation:

$$\begin{pmatrix} I_m & S(\theta) & 0 \\ 0 & I_m & 0 \\ 0 & \bar{A}(\theta) & I_l \end{pmatrix}\xi(\theta) - \xi(\theta+\omega) = R(\theta), \tag{5.16}$$

where $R(\theta) := -B^{-1}(\theta)Q(\theta)e^c\theta - \Lambda(\theta)\varepsilon$. Note that $R(\theta)$ has a term that depends on ε. This linear system (5.16) can be solved by Lemma 2.1 and we have the following Proposition:

Proposition 5.1 *Suppose that all the assumptions in Lemma 5.3 hold. Then, there exists a mapping $\xi(\theta)$ analytic on domain $U_{\rho-2\delta}$ and a vector $\varepsilon \in \mathbb{R}^{m+l}$ such that (5.16) holds. Moreover, there are constants C_6, C_7 depending on m, l, d, ρ, r, $\|f_\lambda\|_{C^1,\mathcal{B}_r}$, $\|J\|_{C^1,\mathcal{B}_r}$, $\|A\|_\rho$ and $\|\mathrm{D}K\|_\rho$ and $|\Theta^{-1}|$ such that*

$$\begin{aligned} \|\xi(\theta)\|_{\rho-2\delta} &\le C_6\gamma^{-2}\delta^{-2\sigma}\|e\|_\rho, \\ |\varepsilon| &\le C_7|\Theta^{-1}|\|e\|_\rho. \end{aligned} \tag{5.17}$$

Proof Let

$$R(\theta) = \begin{pmatrix} R_x(\theta) \\ R_y(\theta) \\ R_z(\theta) \end{pmatrix} := \begin{pmatrix} -\left(B^{-1}(\theta)Q(\theta)e^c(\theta)\right)_x - \Lambda_{11}\varepsilon_1 - \Lambda_{12}\varepsilon_2 \\ -\left(B^{-1}(\theta)Q(\theta)e^c(\theta)\right)_y - \Lambda_{21}\varepsilon_1 - \Lambda_{22}\varepsilon_2 \\ -\left(B^{-1}(\theta)Q(\theta)e^c(\theta)\right)_z - \Lambda_{31}\varepsilon_1 - \Lambda_{32}\varepsilon_2 \end{pmatrix}, \quad \xi(\theta) = \begin{pmatrix} \xi_x(\theta) \\ \xi_y(\theta) \\ \xi_z(\theta) \end{pmatrix},$$

where the subscripts x, y, z denote the directions of the columns of $\mathcal{M}$ and ε_1, ε_2 denote the first m components and last l components of ε, respectively. Write (5.16) as

$$\begin{aligned} \xi_x(\theta) - \xi_x(\theta+\omega) &= R_x(\theta) - S(\theta)\xi_y(\theta), \\ \xi_y(\theta) - \xi_y(\theta+\omega) &= R_y(\theta), \\ \xi_z(\theta) - \xi_z(\theta+\omega) &= R_z(\theta) - \bar{A}(\theta)\xi_y(\theta). \end{aligned} \tag{5.18}$$

We solve the second equation first by Lemma (2.1), that is, we have to choose parameter ε to guarantee that $\mathrm{avg}(R_y) = 0$. Based on the Definition 3.2, the average of matrix Θ is invertible implies that the sub-matrix avg $(\Lambda_{21}\Lambda_{22})$ is of the rank m. It follows that the following linear equation

$$\mathrm{avg}\,(\Lambda_{21}\Lambda_{22})\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\end{pmatrix} = -\mathrm{avg}\left(B^{-1}(\theta)Q(\theta)e^c(\theta)\right)_y \tag{5.19}$$

is solvable. Using Lemma 2.1, we find a unique solution $\tilde{\xi}_y$ with zero average and the estimate that

$$\|\tilde{\xi}_y\|_{\rho-\delta} \le C\gamma^{-1}\delta^{-\sigma}\|R_y\|_\rho.$$

Denote $\xi_y = \tilde{\xi}_y + \bar{\xi}_y$, where $\bar{\xi}_y$ is a constant vector. On one hand, ξ_y is still a solution of the second equation, on the other hand, the values of $\bar{\xi}_y$ affect the right hand side of the first and third equations. By choosing $\bar{\xi}_y$, we can make the first and third equations solvable, that is, we solve the following equation for unique $\bar{\xi}_y$:

$$\mathrm{avg}\begin{pmatrix}S(\theta)\\ \bar{A}(\theta)\end{pmatrix}\bar{\xi}_y = \begin{pmatrix}-\mathrm{avg}\left(B^{-1}(\theta)Q(\theta)e^c(\theta)\right)_x - \mathrm{avg}(S(\theta)\tilde{\xi}_y(\theta))\\ -\mathrm{avg}\left(B^{-1}(\theta)Q(\theta)e^c(\theta)\right)_z - \mathrm{avg}(\bar{A}(\theta)\tilde{\xi}_y(\theta))\end{pmatrix}.$$

Based on Definition 3.2, the matrix Θ is invertible, it follows that the matrix avg $\begin{pmatrix}S(\theta)\\ \bar{A}(\theta)\end{pmatrix}$ has rank m. Consequently, the equation above is uniquely solvable for $\bar{\xi}_y$. Then, we have

$$\|\xi_y\|_{\rho-\delta} \le \|\tilde{\xi}_y\|_{\rho-\delta} + |\overline{\xi_y}| \le C_8\gamma^{-1}\delta^{-\sigma}\|R_y\|_\rho.$$

Using Lemma 2.1 again, we have that there exist unique zero average solutions $\xi_x,\ \xi_z$ satisfying

$$\|\xi_x\|_{\rho-2\delta} \le C_9\gamma^{-1}\delta^{-\sigma}\|R_x(\theta) - S(\theta)\xi_y(\theta)\|_{\rho-\delta},$$
$$\|\xi_z\|_{\rho-2\delta} \le C_{10}\gamma^{-1}\delta^{-\sigma}\|R_z(\theta) - \bar{A}(\theta)\xi_y(\theta)\|_{\rho-\delta}.$$

The details of the proof are similar to proofs in [18]. □

Note that the solutions $\xi_x,\ \xi_z$ we have found would be modified by adding them an average and that is the reason for non-uniqueness of the solutions.

Corollary 5.1 *Assume that all the assumptions of Proposition 5.1 hold. Then*

$$\|\Delta^c\|_{\rho-2\delta} \le C_{11}\gamma^{-2}\delta^{-2\sigma}\|e\|_\rho,$$
$$\|\mathrm{D}\Delta^c\|_{\rho-3\delta} \le C_{11}\gamma^{-2}\delta^{-(2\sigma+1)}\|e\|_\rho, \tag{5.20}$$

and

$$\|\Pi^c_{K(\theta+\omega)}\mathrm{D}G(K,\lambda)(\Delta^c,\varepsilon) + e^c\|_{\rho-2\delta} \le C_{12}\gamma^{-3}\delta^{-(3\sigma+1)}\|e\|^2_\rho, \tag{5.21}$$

where $C_{11},\ C_{12}$ *depend on* $m,\ l,\ d,\ \rho,\ r,\ \|f_\lambda\|_{C^2,\ \mathcal{B}_r}\ \|J\|_\rho,\ \|\mathrm{D}K_0\|_\rho,\ \|A\|_\rho,$ $\|\frac{\partial f_\lambda}{\partial\lambda}(K(\theta))\|_\rho,\ |\Theta^{-1}|$ *and the norm of the projection* $\|\Pi^c_{K(\theta+\omega)}\|$.

Proof Remind the change of variable (5.10), we have

$$\|\Delta^c(\theta)\|_{\rho-2\delta} \le \|\mathcal{M}(\theta)\xi(\theta)\|_{\rho-\delta} \le C_{11}\gamma^{-2}\delta^{-2\delta}\|e\|_\rho.$$

Also,

$$\begin{aligned}
&\|\Pi^c_{K(\theta+\omega)}\mathrm{D}G(K,\lambda)(\Delta^c,\varepsilon)+e^c\|_{\rho-2\delta}\\
&\le \|B_1(\theta)\|_{\rho-2\delta}\|\xi(\theta)\|_{\rho-2\delta}+\left\|\tilde{B}(\theta)\Pi^c_{K(\theta+\omega)}\frac{\partial f_\lambda}{\partial\lambda}\right\|_{\rho-2\delta}|\varepsilon|\|\tilde{B}(\theta)e^c(\theta)\|\\
&\le C_{12}\gamma^{-3}\delta^{-(3\sigma+1)}\|e\|^2_\rho.
\end{aligned}$$

□

5.4 Solving the Linearized Equations on the Hyperbolic Subspaces

We project Eq. (5.1) on the stable and unstable subspaces, that is,

$$\begin{aligned}
&\Pi^{s,u}_{K(\theta+\omega)}\mathrm{D}f_\lambda(K(\theta))\Delta(\theta)-\Pi^{s,u}_{K(\theta+\omega)}\Delta(\theta+\omega)\\
&=-\Pi^{s,u}_{K(\theta+\omega)}\frac{\partial f_\lambda(K(\theta))}{\partial\lambda}\varepsilon-\Pi^{s,u}_{K(\theta+\omega)}e(\theta).
\end{aligned}\tag{5.22}$$

By the invariance of the splitting (2.11), we can write, for the stable one,

$$\Pi^s_{K(\theta+\omega)}\mathrm{D}f_\lambda(\theta)\Delta(\theta)=\mathrm{D}f^s_\lambda(K(\theta))\Delta^s(\theta),$$

and for the unstable one:

$$\Pi^u_{K(\theta+\omega)}\mathrm{D}f_\lambda(\theta)\Delta(\theta)=\mathrm{D}f^u_\lambda(K(\theta))\Delta^u(\theta).$$

Introducing the change of variable $\theta'=T_\omega(\theta)=\theta+\omega$ and $\Delta^{s,u}(\theta')=\Pi^{s,u}_{K(\theta')}\Delta(\theta')$, we can rewrite the Eq. (5.22) as follows.

For the stable part, we rewrite as:

$$\mathrm{D}f^s_\lambda(K)\circ T_{-\omega}(\theta')\Delta^s(T_{-\omega}(\theta'))-\Delta^s(\theta')=-E^s(\theta',\varepsilon),\tag{5.23}$$

where

$$E^s(\theta',\varepsilon)=\Pi^s_{K(\theta')}e\circ T_{-\omega}(\theta')+\Pi^s_{K(\theta')}\left(\frac{\partial f_\lambda(K)\circ T_{-\omega}(\theta')}{\partial\lambda}\varepsilon\right).$$

For the unstable part, we write

$$\mathrm{D}f^u_\lambda(K)\circ T_{-\omega}(\theta')\Delta^u(T_{-\omega}(\theta'))-\Delta^u(\theta')=-E^u(\theta',\varepsilon),\tag{5.24}$$

where

$$E^u(\theta',\varepsilon)=\Pi^u_{K(\theta')}e\circ T_{-\omega}(\theta')+\Pi^u_{K(\theta')}\left(\frac{\partial f_\lambda(K)\circ T_{-\omega}(\theta')}{\partial\lambda}\varepsilon\right).$$

The following proposition provides the existence and the estimate of the solutions for Eqs. (5.23) and (5.24).

Proposition 5.2 *For any fixed $\rho>0$, there exists a unique analytic solution $\Delta^s:U_\rho\to\mathbb{E}^s_{K(\theta)}$ (resp. $\Delta^u:U_\rho\to\mathbb{E}^u_{K(\theta)}$) for Eq. (5.23) (resp.for Eq. (5.24)) and a constant C_{13} depending on the constant C_h in Definition 2.5, the norm of $\|\frac{\partial f_\lambda(K)\circ T_{-\omega}(\theta')}{\partial\lambda}\|_\rho$, the norm of the projection $\|\Pi^s_{K(\theta)}\|_\rho$ (resp. $\|\Pi^u_{K(\theta)}\|_\rho$), the hyperbolic constant μ_1 (resp.μ_2), such that*

$$\|\Delta^{s,u}\|_\rho\le C_{13}(\|e^{s,u}\|_\rho+|\varepsilon|).\tag{5.25}$$

Proof Using Eq. (5.23) iteratively, we have that

$$\Delta^s(\theta') = \sum_{k=0}^{\infty} \left(\mathrm{D} f_\lambda^s(K) \circ T_{-\omega}(\theta') \times \cdots \times \mathrm{D} f_\lambda^s(K) \circ T_{-k\omega}(\theta')\right) E^s(T_{-k\omega}(\theta'), \varepsilon). \tag{5.26}$$

Combining with the condition on the cocycles over $T_{-\omega}$ (see (2.12)), $\mu_1 < 1$, the series converges uniformly on U_ρ, that is,

$$\|\Delta^s\|_\rho \leq C_h \|e^s\|_\rho \sum_{k=0}^{\infty} \mu_1^k.$$

As to the unstable part, we multiply (5.24) by $(\mathrm{D} f_\lambda^u(K) \circ T_{-\omega})^{-1}$, then using the condition (2.13), the unstable equation can be solved in the same way. □

5.5 Change of the Hyperbolicity and the Non-degeneracy Conditions in the Iterative Step

In order to complete one iterative step, we introduce the following proposition to prove the existence of the invariant splitting at each step. As a corollary, we estimate that the change of invariant splitting will be controlled by the change of the embedding torus and hence by the error. It is important that the change will be controlled by the size of the error, so that, as the error decreasing, we can assume that non-degeneracy conditions remain hold under the iteration.

Proposition 5.3 *Assume that there is an analytic splitting*

$$T_{K(\theta)} M = \tilde{\mathcal{E}}^s_{K(\theta)} \oplus \tilde{\mathcal{E}}^c_{K(\theta)} \oplus \tilde{\mathcal{E}}^u_{K(\theta)}, \tag{5.27}$$

which is approximately invariant under the cocycle $\mathrm{D} f_\lambda(K)$ *over* T_ω. *That is, for* $(f_\lambda, K(\theta))$ *satisfying* (2.9), *we have*

$$dist_\rho(\mathrm{D} f_\lambda(K(\theta))\tilde{\mathcal{E}}^{s,u,c}_{K(\theta)}, \tilde{\mathcal{E}}^{s,u,c}_{K(\theta+\omega)}) \leq \tilde{\delta}, \quad \theta \in U_\rho,$$

where $dist_\rho$ *is defined as in Remark* 2.2, $\Pi^{s,u,c}$ *are defined as in Definition* 2.5. *Moreover, assume that for some* $N \in \mathbb{N}$, $0 < \tilde{\mu}_1, \ \tilde{\mu}_2 < 1$, *and some* $\tilde{\mu}_3 \geq 1$, *such that* $\max\{\tilde{\mu}_1, \ \tilde{\mu}_2\} \cdot \tilde{\mu}_3 < 1$, *we have*

$$\|(\mathrm{D} f)(K) \circ T_\omega^{N-1}(\theta) \times \cdots \times (\mathrm{D} f)(K(\theta))\nu\| \leq \tilde{\mu}_1^N \|\nu\|, \ \forall \nu \in \tilde{\mathcal{E}}^s_{K(\theta)}, \tag{5.28}$$

$$\|(\mathrm{D} f)^{-1}(K) \circ T_\omega^{-(N-1)}(\theta) \times \cdots \times (\mathrm{D} f)^{-1} \circ (K(\theta))\nu\| \leq \tilde{\mu}_2^N \|\nu\|, \ \forall \nu \in \tilde{\mathcal{E}}^u_{K(\theta)}, \tag{5.29}$$

and

$$\begin{aligned} &\|(\mathrm{D} f)(K) \circ T_\omega^{N-1}(\theta) \times \cdots \times (\mathrm{D} f)(K(\theta))\nu\| \leq \tilde{\mu}_3^N \|\nu\|, \\ &\|(\mathrm{D} f)^{-1}(K) \circ T_\omega^{-(N-1)}(\theta) \times \cdots \times (\mathrm{D} f)^{-1}(K(\theta))\nu\| \leq \tilde{\mu}_3^N \|\nu\|, \qquad \forall \nu \in \tilde{\mathcal{E}}^c_{K(\theta)}. \end{aligned} \tag{5.30}$$

Assume that $\tilde{\delta} \leq \tilde{\delta}_0$, *where* $\tilde{\delta}_0$ *is a constant depending on* N, $\|\mathrm{D} f_\lambda(K(\theta))\|_\rho$, $\|\mathrm{D} f_\lambda^{-1}(K(\theta))\|_\rho$, $\|\Pi^{s,u,c}\|_\rho$. *Then, there is an analytic splitting*

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

invariant under the cocycle $\mathrm{D}f_\lambda(K)$ *over* T_ω*, which satisfies the characterization of hyperbolic splittings* (2.12)–(2.14).

The splitting above is unique among the splittings in a neighborhood of the original splitting of size δ_0 *measured in* $dist_\rho$*. Furthermore, we have that*

$$\begin{aligned} dist_\rho\left(\mathcal{E}^{s,u,c}_{K(\theta)}, \tilde{\mathcal{E}}^{s,u,c}_{K(\theta)}\right) &\le C_{15}\tilde{\delta}, \\ |\mu_{1,2,3} - \tilde{\mu}_{1,2,3}| &\le C_{15}\tilde{\delta}, \end{aligned} \tag{5.31}$$

where C_{15} *depends on the same parameters as* $\tilde{\delta}_0$ *does.*

Proof See more details in section 5.2 of [19] and in the appendix B of [7]. □

Remark 5.3 Note that the statements on the hyperbolicity condition in Proposition 5.3 do not involve C_h as the parameter in the asymptotic condition. But we involve N. It is easy to verify that both formulations are equivalent if we take (2.12)–(2.14) for some fixed N. For instance, if $\tilde{\mu}_1 > \mu_1$ and we take N as $C_h(\mu_1/\tilde{\mu}_1)^N < 1$, then the condition (5.28) directly follows from (2.12). The opposite is easy to verify.

Remark 5.4 The first result of Proposition 5.3 is a standard result in the theory of normally hyperbolic sets that allows us to conclude that if we are given an approximately invariant splitting, which has some hyperbolic characteristics, then we can find a truly invariant splitting nearby. The main use of Proposition 5.3 is to estimate the change in the hyperbolicity hypotheses at each iterative step, as the following corollary.

Corollary 5.2 *Assume that* $\|K - \tilde{K}\|_\rho$ *is small enough and the hypotheses of Proposition 5.1 and Proposition 5.2 hold for* $\tilde{K}$*. Then, there exists an analytic invariant splitting for* $\mathrm{D}f_\lambda \circ \tilde{K}$*. Furthermore, there exists a constant* C_{16} *which depends on the same parameters as* C_{15}*, such that the following estimates hold:*

$$\begin{aligned} \|\Pi^{s,u,c}_{K(\theta)} - \Pi^{s,u,c}_{\tilde{K}(\theta)}\|_\rho &\le C_{16}\|K - \tilde{K}\|_\rho, \\ |\mu_{1,2,3} - \tilde{\mu}_{1,2,3}| &\le C_{16}\|K - \tilde{K}\|_\rho, \\ C_h &= \tilde{C}_h, \end{aligned} \tag{5.32}$$

where $\tilde{C}_h$ *is the parameter in the asymptotic growth condition for* $\tilde{K}$*.*

Proof We take the invariant splitting for $\mathrm{D}f_\lambda \circ K$ as approximately invariant for $\mathrm{D}f_\lambda \circ \tilde{K}$. Moreover, we take the $\tilde{\delta} = C_{16}\|K - \tilde{K}\|_\rho$. Then the first estimate in (5.32) follows from the Proposition 5.3. The other two estimates follow from the discussion in Remark 5.3. See more details in section 5.2 of [19] and in [7]. □

Once we have estimates for the change in the spaces, we can also obtain estimates in the non-degeneracy conditions in Definition 3.1. Note that they are just matrices obtained by taking projections on the invariant spaces. See Lemma 6.2.

5.6 Estimate for One Step

In the previous subsections, we found the approximate solutions $(\Delta(\theta), \varepsilon)$ for Eq. (5.1) in smaller analyticity domains. The estimates of the solutions depend on $\|e\|_\rho$ and the loss of domain δ. Denote by

$$\lambda_+ = \bar{\lambda} + \varepsilon, \quad K_+(\theta) = K(\theta) + \Delta(\theta), \quad e_+(\theta) := f_{\lambda_+}(K_+(\theta)) - K_+(\theta + \omega),$$

we will show that, if the estimates of the solutions $\|\Delta\|_{\rho-\delta}$, $|\varepsilon|$ are sufficiently small, then the new torus K_+ is still a non-degenerate approximate torus for f_{λ_+} and the new error term e_+ is quadratically small with respect to the original error e.

Lemma 5.4 *Suppose that all the hypotheses for Propositions 5.1 and 5.2 hold and that*

$$(K+\Delta)(U_{\rho-\delta}, \bar{\lambda}+\varepsilon) \subset \mathrm{Domain}(f).$$

Then, there exists a constant $\bar{C}$ depending on σ, m, l, d, ρ, r, $\|f_\lambda\|_{C^2,\mathcal{B}_r}$, $\|DK\|_\rho$, $\|A(\theta)\|_\rho$, $\|\Pi^{s,u,c}_{K(\theta+\omega)}\|$, $\|\frac{\partial f_\lambda}{\partial \lambda}|_{\bar{\lambda}}\|_\rho$ and $|\mathrm{avg}(3)^{-1}|$ such that

$$\|f_{\lambda_+}(K_+(\theta)) - K_+(\theta+\omega)\|_{\rho-3\delta} \le \bar{C}\gamma^{-4}\delta^{-4\sigma}\|e\|^2_\rho. \tag{5.33}$$

Furthermore, the pair (f_λ, K_+) is non-degenerate at $\lambda = \lambda_+$ as defined in Definition 3.2 and the non-degeneracy constraints change by an amount bounded by $\bar{C}\delta^{-\sigma}\|e\|_\rho$.

Proof Adding and subtracting terms, we can rewrite the new error term as:

$$\begin{aligned} &f_{\lambda_+}(K_+(\theta)) - K_+(\theta+\omega) \\ &\quad= f_{\lambda_+}(K+\Delta) - f_\lambda(K) - \frac{\partial f_\lambda}{\partial \lambda}|_{\lambda=\bar{\lambda}}(K)\varepsilon - Df_\lambda(K)\Delta \\ &\qquad+ f_\lambda(K) - K\circ T_\omega + Df_\lambda(K)\Delta - \Delta\circ T_\omega + \frac{\partial f_\lambda}{\partial \lambda}|_{\lambda=\bar{\lambda}}(K)\varepsilon. \end{aligned} \tag{5.34}$$

Now we can estimate the norm of the error by the sum of the norms of the lines in the identity above. The second line can be estimated by Taylor's theorem applied to the function f. Recall that we made assumptions that ensured the range of K_+ was inside the domain of f. See more details in Remark 5.5. The third line exactly equals (5.1), that is,

$$DG(K,\lambda)(\Delta,\varepsilon) + e = f_\lambda(K) - K\circ T_\omega + Df_\lambda(K)\Delta - \Delta\circ T_\omega + \frac{\partial f_\lambda}{\partial \lambda}|_{\lambda=\bar{\lambda}}(K)\varepsilon$$

and the pair (Δ, ε) were chosen precisely in such a way that the norm of the third line is bounded by the square of the norm of the error.

It follows from Proposition 5.1, Corollary 5.1 and Proposition 5.2 that

$$\|f_{\lambda_+}(K_+(\theta)) - K_+(\theta+\omega)\|_{\rho-3\delta} \le \bar{C}\gamma^{-4}\delta^{-4\sigma}\|e\|^2_\rho.$$

□

Remark 5.5 Note that the solution of the linearized equation Δ is defined in domain $U_{\rho-\delta}$ for any $0 < \delta < \rho$ and the estimate of Δ depends on the norm of the error term $e(\theta)$ and the loss of domain δ. If δ is too small compared with $\|e\|_\rho$, the estimates of Δ will blow up. So the estimates on each step require some restrictions on δ. For an instance, given the estimates on $(\Delta(\theta), \varepsilon)$, we can see that the assumption we made

$$\bar{C}\gamma^{-2}\delta^{-2\sigma-1}\|e\|_\rho \le \eta, \tag{5.35}$$

where η is smaller than the distance of $K(U_\rho)$ to the complement of the domain f implies that the range of K_+ is contained in the domain of f. This is what allows us to use the Taylor's theorem with reminder to estimate the new error.

6 Iteration of Newton's Method and Convergence

In this section, we will show that if the initial error is small enough, we can repeat infinitely the iterative step and it converges to a solution of the original problem. Furthermore, we can estimate the distance between the true solution and the initial position.

Recall that we define G in the very begin of Sect. 5, now we start with $f_{\lambda_0},\ K_0,\ \omega,\ \rho_0$ satisfying the assumptions (A1)–(A3) as in Theorem 1, that is,

$$e_0(\theta) := G(K_0, \lambda_0) := f_{\lambda_0}(K_0(\theta)) - K_0(\omega + \theta). \tag{6.1}$$

By Taylor theorem with remainder, we have

$$\begin{aligned} &G(K_0 + \Delta_0, \lambda_0 + \varepsilon_0) \\ &\quad = G(K_0, \lambda_0) + \left[\frac{\partial f_\lambda(K_0(\theta))}{\partial \lambda}|_{\lambda=\lambda_0} \right] \varepsilon_0 + \mathrm{D} f_{\lambda_0}(K_0(\theta))\Delta_0(\theta) - \Delta_0(\theta + \omega) \\ &\qquad + O(|\varepsilon_0|^2, \|\Delta_0\|^2). \end{aligned} \tag{6.2}$$

If the error $\|e_0\|_{\rho_0}$ is small enough, applying the process in Sect. 5, we can get a pair of corrections $(\Delta_0,\ \varepsilon_0)$ such that the first line on the right hand side of Eq. (6.2) cancel almost exactly. Hence

$$f_\lambda(K(\theta)) = K(\theta + \omega)$$

has a new pair of approximate solution $(f_{\lambda_1}, K_1(\theta))$ defined on domain U_{ρ_1}, where $\lambda_1 = \lambda_0 + \varepsilon_0,\ K_1 = K_0 + \Delta_0,\ \rho_1 = \rho_0 - 3\delta_0$. By Lemma 5.4, we have that

$$\|e_1\|_{\rho_1} := \|f_{\lambda_1}(K_1(\theta)) - K_1(\theta + \omega)\| < C_0\gamma^{-4}\delta_0^{-4\sigma}\|e_0\|_{\rho_0}^2.$$

Assume that we have already found the (i-1)th approximate solutions $(f_{\lambda_{i-1}}, K_{i-1}(\theta))$ on domain $U_{\rho_{i-1}}$ which satisfies the assumptions in Theorem 1 and that $\|e_{i-1}\|_{\rho_{i-1}}$ is sufficiently small, for any $i = 1, 2, 3, \ldots$. Consider the (i − 1)th linearized equation as follow:

$$\mathrm{D}G(K_{i-1}, \lambda_{i-1})(\Delta_{i-1}, \varepsilon_{i-1}) = -e_{i-1}(\theta). \tag{6.3}$$

By Propositions 5.1 and 5.2, we can find approximate solutions $(\Delta_{i-1}^c, \varepsilon_{i-1})$ on the center subspace and $\Delta_{i-1}^{s,u}$ on the hyperbolic subspaces. Then we consider the corrections

$$\lambda_i = \lambda_{i-1} + \varepsilon_{i-1},\quad K_i(\theta) = K_{i-1}(\theta) + \Delta_{i-1}(\theta), \tag{6.4}$$

defined on U_{ρ_i}, where $\rho_i = \rho_{i-1} - 3\delta_{i-1}$ and $\Delta_{i-1} = (\Delta_{i-1}^s,\ \Delta_{i-1}^u,\ \Delta_{i-1}^c)^\top$ such that the new error term

$$e_i(\theta) := G(K_i,\ \lambda_i) := f_{\lambda_i}(K_i(\theta)) - K_i(\theta + \omega)$$

is quadratically small with respect to $\|e_{i-1}\|_{\rho_{i-1}}$.

Under this iterative process, we can find a sequence of approximate solutions

$$(\lambda_0, K_0(\theta)),\ (\lambda_1, K_1(\theta)),\ \cdots,\ (\lambda_i, K_i(\theta)),\ \cdots$$

defined on domains

$$U_{\rho_0} \supset U_{\rho_1} \supset \cdots \supset U_{\rho_i} \supset \cdots$$

for the equation

$$f_\lambda(K(\theta)) = K(\theta + \omega).$$

We will verify that $K_i \to K_\infty,\ \lambda_i \to \lambda_\infty,\ \rho_i \to \rho_\infty > 0,\ e_i \to 0$ as $i \to \infty$.

The following lemma follows from Propositions 5.1 and 5.2. It generalizes Lemma 5.4 for a general step.

Lemma 6.1 *Assume that* $(f_{\lambda_{i-1}},\ K_{i-1})$ *is a pair of non-degenerate approximate solutions for Eq.* (6.3) *such that*

$$r_{i-1} := \|K_{i-1} - K_0\|_{\rho_{i-1}} < r. \tag{6.5}$$

If $\|e_{i-1}\|_{\rho_{i-1}}$ *is small enough, so that the assumptions of Lemma* 5.2 *apply, then for any* $0 < \delta_{i-1} < \frac{\rho_{i-1}}{3}$, *there exists a function* $\Delta_{i-1}(\theta) \in U_{\rho_{i-1}-3\delta_{i-1}}$ *and* $\varepsilon_{i-1} \in \mathbb{R}^{m+l}$, *such that*

$$\|\Delta_{i-1}\|_{\rho_{i-1}-2\delta_{i-1}} \le \bar{C}_{i-1}\gamma^{-2}\delta_{i-1}^{-2\sigma}\|e_{i-1}\|_{\rho_{i-1}}, \tag{6.6}$$

$$\|\mathrm{D}\Delta_{i-1}\|_{\rho_{i-1}-3\delta_{i-1}} \le \bar{C}_{i-1}\gamma^{-2}\delta_{i-1}^{-(2\sigma+1)}\|e_{i-1}\|_{\rho_{i-1}},$$

$$|\varepsilon_{i-1}| \le \bar{C}_{i-1}|\Theta^{-1}|\|e_{i-1}\|_{\rho_{i-1}}, \tag{6.7}$$

where $\bar{C}_{i-1}$ *is a positive constant depending on* $m,\ l,\ d,\sigma,\ r,\ \rho_{i-1},\ \|f_{\lambda_{i-1}}\|_{\mathcal{B}_r},\ \|DK_{i-1}\|_{\rho_{i-1}}$, $\|A_{i-1}\|_{\rho_{i-1}},\ \|\Pi^{s,u,c}_{K_{i-1}(\theta+\omega)}\|_{\rho_{i-1}},\ \left\|\frac{\partial f_\lambda}{\partial\lambda}|_{\lambda=\lambda_{i-1}}\right\|_{\rho_{i-1}}$. *Moreover if*

$$\bar{C}_{i-1}\gamma^{-2}\delta_{i-1}^{-2\sigma-1}\|e_{i-1}\|_{\rho_{i-1}} < r, \tag{6.8}$$

then we denote $K_i,\ \lambda_i$ *as* (6.4), *which are the new non-degenerate approximate torus and the new parameter. The new error term* $e_i(\theta)$ *satisfies the following estimate:*

$$\|e_i\|_{\rho_i} \le \bar{C}_{i-1}\gamma^{-4}\delta_i^{-4\sigma}\|e_{i-1}\|^2_{\rho_{i-1}}.$$

Lemma 6.2 *Assume that the hypotheses in Lemma* 6.1 *hold and that*

$$\bar{C}_{i-1}\gamma^{-1}\delta_{i-1}^{-\sigma-1}\|e_{i-1}\|_{\rho_{i-1}} \le \frac{1}{2}. \tag{6.9}$$

Then, the following statements hold.

(1) *Denote by*

$$\mathrm{D}K_i(\theta) = (X_i(\theta), Z_i(\theta)), \quad i = 1, 2, \ldots$$

where $X_i,\ Z_i$ *are the first* m *and the last* l *columns and* $X^c_{i,V}$ *denote the projections on the subspace* V, *defined as in* (3.2). *If the matrix* $\left[(X^c_{i-1,V})^\top X^c_{i-1,V}\right]$ *is invertible, then the matrix* $\left[(X^c_{i,V})^\top X^c_{i,V}\right]$ *is also invertible and the inverse* A_i *satisfies the following estimate:*

$$\|A_i\|_{\rho_i} \le \|A_{i-1}\|_{\rho_{i-1}} + \bar{C}_{i-1}\gamma^{-2}\delta_{i-1}^{-(2\sigma+1)}\|e_{i-1}\|_{\rho_{i-1}}.$$

(2) *If the matrix* B_{i-1} *is invertible, then* B_i *is invertible and the inverse satisfies the estimate*

$$\|B_i^{-1}\|_{\rho_i} \le \|B_{i-1}^{-1}\|_{\rho_{i-1}} + \bar{C}_{i-1}\gamma^{-2}\delta_{i-1}^{-(2\sigma+1)}\|e_{i-1}\|_{\rho_{i-1}}.$$

(3) *If the matrix* Θ_{i-1} *is invertible, then* Θ_i *is invertible and the inverse satisfies the estimate*

$$|\Theta_i^{-1}|_{\rho_i} \le |\Theta_{i-1}^{-1}|_{\rho_{i-1}} + \bar{C}_{i-1}\gamma^{-2}\delta_{i-1}^{-(2\sigma+1)}\|e_{i-1}\|_{\rho_{i-1}}.$$

(4) *The assumption* (6.5) *guarantees that the range of* $(K_{i-1} + \Delta_{i-1},\ \lambda_{i-1} + \varepsilon_{i-1})$ *is inside of the domain of* f_{λ_i}.

The iterative lemmas above are very common arguments in the proof of KAM type theorem. The most important reason for the lemmas to hold is that the constants $\bar{C}_i$, together with the constants appearing in the discussion of Sect. 5, depending only on $m,\ l,\ d,\ r$ which do not change during the iterative process and the following quantities

$$\|f_{\lambda_i}\|_{\mathcal{B}_r},\ \|DK_i\|_{\rho_i},\ \|A_i\|_{\rho_i},\ \|\Pi^{s,u,c}_{K_i(\theta+\omega)}\|_{\rho_i},\ \left\|\frac{\partial f_\lambda}{\partial\lambda}\,|_{\lambda=\lambda_i}\,(K_i)\right\|_{\rho_i},$$

which can be controlled when K_i only changes in a neighborhood. By similar calculations, one can claim that there exists a positive constant $\bar{C}^*$ such that $0<\bar{C}_i<\bar{C}^*$, for $i=1,2,\ldots$. See the discussion in [19] and in [7].

To ensure that we can perform the iterative process in Lemmas 6.1 and 6.2 to prove Theorem 1, we only need to verify (6.5) and (6.9). Take the choice of the analyticity loss as

$$\delta_i=\frac{\delta_0}{2^i},\ \rho_i=\rho_{i-1}-3\delta_{i-1}=\rho_0-6(1-\frac{1}{2^i})\delta_0.$$

Denote $\epsilon_i=\|e_i\|_{\rho_i}$ and together with the choice of δ_i, we have that, for $i\geq 1$,

$$\begin{aligned}\epsilon_i&\leq\bar{C}\gamma^{-4}\delta_i^{-4\sigma}\epsilon_{i-1}^2\leq\bar{C}\gamma^{-4}\delta_0^{-4\sigma}2^{4\sigma(i-1)}\epsilon_{i-1}^2\\&\leq(\bar{C}\gamma^{-4}\delta_0^{-4\sigma})^{1+2+\cdots+2^{i-1}}2^{4\sigma[(i-1)+2(i-2)+\cdots+2^{i-2}]}\epsilon^{2^i}\\&\leq\left(\bar{C}2^{4\sigma}\gamma^{-4}\delta_0^{-4\sigma}\epsilon_0\right)^{2^i},\end{aligned}$$

where we used

$$\begin{aligned}&(i-1)+2(i-2)+\cdots+2^{i-2}\\&\quad=2^{i-1}\left[(i-1)2^{-(i-1)}+(i-2)2^{-(i-2)}+\cdots+2^{-1}\right]\leq 2^i-i.\end{aligned}$$

We see that if ϵ_0 is small enough, then we can ensure that $\delta_{i-1}^{-(\sigma+1)}\epsilon_{i-1}$ is small enough to satisfy (6.9) for any $i=1,2,\ldots$. We note that the smallness assumption of ϵ_0 is independent of the iterative step.

Moreover, we have

$$r_i=\|K_i-K_0\|_{\rho_i}\leq\sum_{j=0}^{i-1}\|\Delta_j\|_{\rho_j}\leq\bar{C}\gamma^{-2}\delta_0^{-2\sigma}\epsilon_0\left(1+\kappa\frac{2^{4\sigma}}{2^{2\sigma}-1}\right),$$

where $\kappa=\bar{C}\gamma^{-4}\delta_0^{-4\sigma}2^{4\sigma}\epsilon_0$. Note that, taking ϵ_0 sufficiently small ensures that κ is small, so that the assumption (6.5) is satisfied. Therefore, the new torus K_i never leaves the neighborhood of K_0 and we can repeat the iterative process of any $i\in\mathbb{N}$. The estimate above also establishes the estimate of (3.10) in Theorem 1. Together with (6.7) in Lemma 6.1, we have that $\lambda_i\to\lambda_0$ as $i\to\infty$.

This completes the proof of Theorem 1.

7 Proof of Local Uniqueness

In this section, we will prove Theorem 2 which is similar to Theorem 2 in [1]. We assume that the embeddings K_1 and K_2 satisfy the assumptions in Theorem 1. More specifically, we assume that $K_1,\ K_2$ are non-degenerate solutions for (2.10) for some vector $\lambda_1,\ \lambda_2 \in \mathbb{R}^{m+l}$. Denote by

$$\mathcal{G}(\lambda_1,\ \lambda_2,\ K_1,\ K_2) = f_{\lambda_2}(K_2(\theta)) - K_2(\theta+\omega) - f_{\lambda_1}(K_1(\theta)) + K_1(\theta+\omega).$$

Note that $\mathcal{G}(\lambda_1,\ \lambda_2,\ K_1,\ K_2) = 0$. By the Taylor's expression at $(\lambda_1,\ K_1)$, we have that

$$0 = \mathcal{G} = \mathrm{D}f_{\lambda_1}(K_1)\Delta(\theta) - \Delta(\theta+\omega) + \left[\frac{\partial f_\lambda(K_1)}{\partial\lambda}|_{\lambda=\lambda_1}\right]\varepsilon,$$

where $\Delta(\theta) = K_2(\theta) - K_1(\theta),\ \varepsilon = \lambda_2 - \lambda_1$. Moreover, we obtain the following linearized equation

$$\mathrm{D}f_\lambda(K_1(\theta))\Delta(\theta) - \Delta(\theta+\omega) + \left[\frac{\partial f_\lambda(K_1)}{\partial\lambda}\ |_{\lambda=\lambda_1}\right]\varepsilon = -R. \tag{7.1}$$

where $R = O(\|\Delta\|^2,\ |\varepsilon|^2)$ is the Taylor reminder. We solve this equation by repeating the process in Sect. 5. We first project Eq. (7.1) on the center subspace, that is,

$$\mathrm{D}f_\lambda^c(K_1(\theta))\Delta^c(\theta) - \Delta^c(\theta+\omega) + \left[\Pi^c_{K_1(\theta+\omega)}\frac{\partial f_\lambda(K_1)}{\partial\lambda}\ |_{\lambda=\lambda_1}\right]\varepsilon = -R^c, \tag{7.2}$$

where $\Delta^c(\theta) = \Pi^c_{K_1(\theta)}\Delta,\ R^c = \Pi^c_{K_1(\theta+\omega)}R$. Applying the change of variable $\Delta^c(\theta) = \mathcal{M}(\theta)\xi(\theta)$, where $\mathcal{M}(\theta)$ is defined in (4.7) by replacing K with K_1, and we omit the quadratically bounded terms as in Sect. 5.3, the Eq. (7.2) can be rewritten as

$$\xi_x(\theta) - \xi_x(\theta+\omega) = \tilde{R}^c_x - S(\theta)\xi_y(\theta), \tag{7.3}$$
$$\xi_y(\theta) - \xi_y(\theta+\omega) = \tilde{R}^c_y, \tag{7.4}$$
$$\xi_z(\theta) - \xi_z(\theta+\omega) = \tilde{R}^c_z - \bar{A}(\theta)\xi_y(\theta), \tag{7.5}$$

where $\tilde{R}^c_{x,y,z} = -\left[B^{-1}(\theta)Q(\theta)R^c(\theta) + \Lambda(\theta)\varepsilon\right]_{x,y,z}$, $B,\ Q,\ \Lambda$ are defined as in (4.20), (4.15), (3.5), respectively, by replacing $K(\theta)$ by $K_1(\theta)$.

Together with the proof of Proposition 5.1 and the following lemma, we can estimate the norm of $\Delta^c(\theta)$.

Lemma 7.1 *There exists $\tilde{C}$ depending on $m,\ l,\ \rho,\ |J|_\rho,\ \|K_1\|_\rho,\ \|\Pi^c_{K_1(\theta)}\|,\ \|\mathrm{D}K_1\|_\rho$ such that if $\tilde{C}\|K_1 - K_2\|_\rho \le 1$, then one can find $\tau_1 \in \{\tau \in \mathbb{R}^{m+l},\ |\tau| < \|K_1 - K_2\|_\rho\}$, such that*

$$\mathrm{avg}\left[\begin{pmatrix} T_1 \\ T_3 \end{pmatrix}\left(\Pi^c_{K_1(\theta)}(K_1\circ T_{\tau_1} - K_2)(\theta)\right)\right] = 0.$$

As a consequence, for any $0 < \delta < \frac{\rho}{2}$, there exists a constant $\hat{C}$ depending on $\tilde{C}$ and $|\Theta^{-1}|$, such that the following estimates hold:

$$\|\Pi^c_{K_1(\theta)}(K_1\circ T_{\tau_1} - K_2)(\theta)\|_{\rho-2\delta} \le \hat{C}\gamma^{-2}\delta^{-2\sigma}\|K_1 - K_2\|^2_\rho,$$
$$|\lambda_1 - \lambda_2| \le \hat{C}\|K_1 - K_2\|^2_\rho. \tag{7.6}$$

Moreover, we can project (7.1) on the hyperbolic subspace. By the discussion in Sect. 5.4, we have that

$$\|\Pi^{s,u}_{K_1(\theta)}(K_1\circ T_{\tau_1} - K_2)(\theta)\|_{\rho-2\delta} \le \bar{C}\|K_1 - K_2\|^2_\rho.$$

Combining with the estimates (7.6), we have proved that the following estimate holds:

$$\|K_1\circ T_{\tau_1} - K_2\|_{\rho-2\delta} \le \hat{C}\gamma^{-2}\delta^{-2\sigma}\|K_1 - K_2\|^2_\rho. \tag{7.7}$$

We end up this proof with a discussion similar to Sect. 6, by taking a sequence $\{\tau_m\}_{m\geq 0}$ such that

$$|\tau_m - \tau_{m-1}| \leq \|K_1 \circ T_{\tau_{m-1}} - K_2\|_{\rho_{m-1}},$$

and

$$\|K_1 \circ T_{\tau_m} - K_2\|_{\rho_m} \leq \hat{C}\gamma^{-2}\delta_m^{-2\sigma}\|K_1 \circ T_{\tau_{m-1}} - K_2\|^2_{\rho_{m-1}},$$

where $\tau_0 = 0$, $\rho_0 = \rho$, $\delta_1 = \frac{\rho}{8}$ and $\delta_{m+1} = \frac{\delta_m}{2}$, $\rho_m = \rho - \sum_{k=1}^{m} \delta_k$ for $m \geq 1$. By a simple calculation, we have the following estimates

$$\|K_1 \circ T_{\tau_m} - K_2\|_{\rho_m} \leq \left(\hat{C}\gamma^{-2}\delta_1^{-2\sigma}2^{2\sigma}\|K_1 - K_2\|_\rho\right)^{2^m} 2^{-2\sigma(m+1)}.$$

Therefore, by the assumption in Theorem 2 that $\|K_1 - K_2\|_\rho$ is sufficiently small, we proved that the sequence $\{\tau_m\}_{m\geq 0}$ converges to τ_∞ and

$$\|K_1 \circ T_{\tau_\infty} - K_2\|_{\frac{\rho}{2}} = 0, \quad \lambda_1 = \lambda_2,$$

that is, $K_1 \circ T_{\tau_\infty}$ and K_2 are analytic and coincide in $U_{\frac{\rho}{2}}$. This completes the proof of Theorem 2.

8 An Example of Presymplectic Systems

To end this paper, we will give a simple example in which the kernel does not integrate to a product. We hope that the methods of this paper can be extended to deal with the examples considered in this section, but it seems that the role of parameters will be very different.

Consider $M = \mathbb{T}^3$, which is a 3-dimensional torus, endowed with the form:

$$\Omega_{\alpha,\beta}(u, v) = \langle \alpha, u\rangle \cdot \langle \beta, v\rangle - \langle \alpha, v\rangle \cdot \langle \beta, u\rangle,$$

where $\alpha, \beta \in \mathbb{R}^3$. The kernel of this form is the span of $\alpha \times \beta$, where $\times$ is the standard cross product. Of course the kernel in this case is integrable, since it is a 1-dimensional vector field.

If $\alpha \times \beta$ is irrational, we see that all the leaves of the kernel can be dense in $\mathbb{T}^3$. On the other hand, if $\alpha \times \beta$ is resonant, it could be that each of the leaves is $\mathbb{T}^2$ or $\mathbb{T}^1$ (depending on the multiplicity of the resonance.)

Similar example can be constructed in higher dimensional tori $\mathbb{T}^d$. For α_i, $\beta_i \in \mathbb{R}^d$, $i = 1, \ldots, N$, we denote that

$$\Omega(u, v) = \sum_{i=1}^{N} \langle \alpha_i, u\rangle\langle \beta_i, v\rangle - \langle \alpha_i, v\rangle\langle \beta_i, u\rangle.$$

In this case the kernel is the $d - 2N$ dimensional space orthogonal to α_i, β_i. This is always integrable and it can be arranged easily that all the leaves are dense or that they are dense in lower dimensional tori.

References

1. Alishah, H.N., de la Llave, R.: Tracing KAM tori in presymplectic dynamical systems. J. Dyn. Differ. Equ. **24**(4), 685–711 (2012)
2. Arnol'd, V.I.: Small denominators and problems of stability of motion in classical and celestial mechanics. Russ. Math. Surv. **18**(6), 85–191 (1963)
3. Banyaga, A.: The Structure of Classical Diffeomorphism Groups, Volume 400 of Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht (1997)
4. Bergmann, P.G.: Introduction of "true observables" into the quantum field equations. Nuovo Cimento **10**(3), 1177–1185 (1956)
5. Calleja, R., Celletti, A., de la Llave, R.: A KAM theory for conformally symplectic systems: efficient algorithms and their validation. J. Differ. Equ. **255**(5), 978–1049 (2013)
6. Calleja, R., Celletti, A., de la Llave, R.: Whiskered KAM tori of conformally symplectic systems (2019). arXiv:1901.06059
7. Calleja, R., Celletti, A., de la Llave, R.: Existence of whiskered KAM tori of conformally symplectic systems. Nonlinearity **33**(1), 538–597 (2020)
8. da Silva, A.C., Karshon, Y., Tolman, S.: Quantization of presymplectic manifolds and circle actions. Trans. Am. Math. Soc. **352**(2), 525–552 (2000)
9. Contreras, M., Hojman, S.A.: Option pricing, stochastic volatility, singular dynamics and constrained path integrals. Phys. A **393**, 391–403 (2014)
10. Cariñena, J.F., Gomis, J., Ibort, L.A., Román, N.: Canonical transformations theory for presymplectic systems. J. Math. Phys. **26**(8), 1961–1969 (1985)
11. Cariñena, J.F., Ibort, L.A., Gomis, J., Román-Roy, N.: Applications of the canonical-transformation theory for presymplectic systems. Nuovo Cimento B (11) **98**(2), 172–196 (1987)
12. Carinena, J.F., Ranada, M.F.: Comments on the presymplectic formalism and the theory of regular Lagrangians with constraints. J. Phys. A **28**(3), L91–L97 (1995)
13. Deriglazov, A.: Classical Mechanics. Springer, Heidelberg (2010). Hamiltonian and Lagrangian formalism
14. Dirac, P.A.M.: Generalized hamiltonian dynamics. Can. J. Math. **2**, 129–148 (1950)
15. Dirac, P.A.M.: Generalized hamiltonian dynamics. Proc. R. Soc. Lond. Ser. A **246**, 326–332 (1958)
16. Dirac, P.A.M.: Lectures on Quantum Mechanics, Volume 2 of Belfer Graduate School of Science Monographs Series. Belfer Graduate School of Science, New York; produced and distributed by Academic Press, Inc., New York, (1967). Second printing of the 1964 original
17. de la Llave, R.: A tutorial on KAM theory. In: Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), volume 69 of Proceedings of Symposium. Pure Mathemtics, pp. 175–292. American Mathematical Society, Providence (2001)
18. de la Llave, R., González, A., Jorba, Á., Villanueva, J.: KAM theory without action-angle variables. Nonlinearity **18**(2), 855–895 (2005)
19. Fontich, E., de la Llave, R., Sire, Y.: Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions. J. Differ. Equ. **246**(8), 3136–3213 (2009)
20. Fontich, E., de la Llave, R., Sire, Y.: 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)
21. Gotay, M.J., Nester, J.M., Hinds, G.: Presymplectic manifolds and the Dirac–Bergmann theory of constraints. J. Math. Phys. **19**(11), 2388–2399 (1978)
22. Gotay, M.J., Śniatycki, J.: On the quantization of presymplectic dynamical systems via co-isotropic embeddings. Commun. Math. Phys. **82**(3), 377–389 (1981/1982)
23. Huguet, G., de la Llave, R., Sire, Y.: Computation of whiskered invariant tori and their associated manifolds: new fast algorithms. Discrete Contin. Dyn. Syst. **32**(4), 1309–1353 (2012)
24. Henneaux, M., Teitelboim, C.: Quantization of Gauge Systems. Princeton University Press, Princeton (1992)
25. Libermann, P., Marle, C.-M.: Symplectic Geometry and Analytical Mechanics, Volume 35 of Mathematics and its Applications. D. Reidel Publishing Co., Dordrecht (1987). Translated from the French by Bertram Eugene Schwarzbach
26. Moser, J.: A rapidly convergent iteration method and non-linear partial differential equations-I. Ann. Scuola Norm. Sup. Pisa **3**(20), 265–315 (1966)
27. Moser, J.: A rapidly convergent iteration method and non-linear differential equations-II. Ann. Scuola Norm. Sup. Pisa **3**(20), 499–535 (1966)
28. Moser, J.: Convergent series expansions for quasi-periodic motions. Math. Ann. **169**, 136–176 (1967)

29. Rußmann, H.: On optimal Estimates for the Solutions of Linear Partial Differential Equations of First Order with Constant Coefficients on the Torus. Lecture Notes in Physics, vol. 38, pp. 598–624 (1975)
30. Souriau, J.-M.: Structure of Dynamical Systems, volume 149 of Progress in Mathematics. Birkhäuser Boston, Inc., Boston, MA (1997). A Symplectic View of Physics, Translated from the French by C. H. Cushman-de Vries, Translation edited and with a preface by R. H. Cushman and G. M. Tuynman