Physica D 462 (2024) 134127

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

A modified parameterization method for invariant Lagrangian tori for partially integrable Hamiltonian systems

Jordi-Lluís Figueras [a,*], Alex Haro [b,c]

[a] Department of Mathematics, Uppsala University, Box 480, 751 06 Uppsala, Sweden
[b] Departament de Matemàtiques i Informàtica, Universitat de Barcelona, Gran Via 585, 08007 Barcelona, Spain
[c] Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, 08193 Bellaterra, Spain

ARTICLE INFO

Communicated by A. Celletti

Keywords:
KAM theory
Small divisors
First integrals

ABSTRACT

In this paper we present an a-posteriori KAM theorem for the existence of an $(n-d)$-parameter family of d-dimensional isotropic invariant tori with Diophantine frequency vector $\omega \in \mathbb{R}^d$, of type (γ, τ), for n degrees of freedom Hamiltonian systems with $(n-d)$ independent first integrals in involution. If the first integrals induce a Hamiltonian action of the $(n-d)$-dimensional torus, then we can produce n-dimensional Lagrangian tori with frequency vector of the form (ω, ω_p), with $\omega_p \in \mathbb{R}^{n-d}$. In the light of the parameterization method, we design a (modified) quasi-Newton method for the invariance equation of the parameterization of the torus, whose proof of convergence from an initial approximation, and under appropriate non-degeneracy conditions, is the object of this paper. We present the results in the analytic category, so the initial torus is real-analytic in a certain complex strip of width ρ, and the corresponding error in the functional equation is ε. We heavily use geometric properties and the so called automatic reducibility to deal directly with the functional equation and get convergence if $\gamma^{-2}\rho^{-2\tau-1}\varepsilon$ is small enough, in contrast with most of KAM results based on the parameterization method, that get convergence if $\gamma^{-4}\rho^{-4\tau}\varepsilon$ is small enough. The approach is suitable to perform computer assisted proofs.

1. Introduction

Persistence under perturbations of regular (quasi-periodic) motion is one of the most important problems in Mechanics and Mathematical Physics, and has deep implications in Celestial and Statistical Mechanics. The seminal works of Kolmogorov [1], Arnold [2] and Moser [3] put the name to the KAM theory, that has become a full body of knowledge that connects fundamental mathematical ideas in many different contexts around the so-called small divisors. See e.g. the popular book [4], and the surveys [5,6].

Although KAM theory held for general dynamical systems under very mild technical assumptions, its application to concrete systems became a challenging problem. With the advent of computers and new methodologies, the distance between theory and practice was shortened (see e.g. [7–9]). One direction that have experienced a lot of progress is the a-posteriori approach based on the parameterization method [6,10–13], that in this context was originally known as KAM theory without angle-action coordinates. This approach lead to the design of a general methodology to perform computer assisted proofs of existence of Lagrangian invariant tori [14] for symplectic maps such as the Chirikov standard map. But new impulses have to be made in order to make KAM theory fully applicable to realistic physical systems, which often have extra first integrals, or degeneracies.

The goal of this paper is to present a KAM theorem in a-posteriori format for the existence of invariant Lagrangian cylinders for Hamiltonian systems with first integrals in involution, see Theorem 2.18. If the first integrals induce a Hamiltonian action of a torus, then the theorem produces invariant Lagrangian tori. The presence of first integrals in involution is usually treated with symplectic reduction techniques [15,16], and then applying KAM theorems to the reduced systems. For the sake of versatility, we do not pursue such changes of variables, and we get n-dimensional Lagrangian invariant cylinders from a single d-dimensional isotropic torus through a Reduction lemma, see Lemma 2.5, that avoids the use of symplectic reduction techniques. Of course, our results work for systems without additional first integrals in involution, that could be obtained after reduction techniques.

Although the setting of the paper is close to that in [17], we incorporate some constructs in [18] to improve crucial estimates. We present a (modified) quasi-Newton method for these systems. The method solves exactly the same equations as in the quasi-Newton method in [17],

---

[*] Corresponding author.
E-mail addresses: figueras@math.uu.se (J.-L. Figueras), alex@maia.ub.es (A. Haro).

https://doi.org/10.1016/j.physd.2024.134127
Received 22 September 2023; Received in revised form 31 January 2024; Accepted 3 March 2024

but differ in the last step with [17] in the way the updates of the parameterization of the torus are made. In particular, normal corrections to the torus are made to improve the invariance of the torus, while tangent corrections to the torus are made to conjugate the internal dynamics of the torus to a linear flow. As a result, while in [17] we got convergence of the quasi-Newton method if $\gamma^{-4}\rho^{-4\tau}\varepsilon$ is small enough, as in most of KAM papers based on the parameterization method, in this paper we get convergence if $\gamma^{-2}\rho^{-2\tau-1}\varepsilon$ is small enough, as in [18,19] (see also the prequel [20]), and in the KAM Theorem [2,21,22]. But, while the inspiring paper [18] substitutes the invariance equation of the parameterization method by three different conditions which are altogether equivalent to invariance, here we suitably project the error of invariance in tangent and normal components, as in [17]. Also, our proof of the result heavily relies on geometrical and reducibility properties, complementing the approach in [18], that does not use (at least explicitly) these properties.

In order to emphasize the geometric properties, and for the sake of versatility, we have considered symplectic structures other than the canonical. This was one main point in the seminal paper [10]. Here, however, we have only considered symplectic structures that have a compatible almost-complex structure as in [23] (thus inducing a Riemannian metric). It is known that any symplectic manifold admits a compatible almost-complex structure (see e.g. [15]).

This paper complements [17] also in the sense that, from the algorithm derived in that paper (whose convergence is proved), one can produce approximations of the solutions of the invariance equations, and then apply the a posteriori KAM theorem derived in this paper, in combination with techniques introduced in [14]. There are situations in which the a posteriori KAM theorem in [17] could fail to be applied in practical situations, in which computer resources time and memory are finite. For instance, when the tori are high dimensional and/or about to break, thus needing high order Fourier approximations and/or being the width of the analyticity strip ρ very small. The improved estimates here could mitigate such problems, thus pushing further the domain of existence of KAM tori.

As usual in KAM theory, and in particular in the parameterization method, the proof of existence of invariant tori is pursued by means of a quasi-Newton method in a scale of Banach spaces, here analytic functions. At each step the analyticity strip in which the objects are defined is reduced, and one has to control the whole sequence to get convergence to an object defined in a final analyticity strip. The way the width reductions are produced on the analyticity strip seems to influence the results in practical situations. We have included a digression on the fact that the best choice seems to be close to consider geometric series of width reductions with ratio $\frac{1}{2}$.

The paper is organized as follows. Section 2 introduces the background and the geometric and analytical constructions, and the main result, Theorem 2.18, is presented at the end. The proof of Theorem 2.18 is detailed in Section 3. An auxiliary lemma to control inverses of matrices is included in Appendix A. In order to collect the long list of expressions leading to the explicit estimates and conditions of the theorems, we include separate tables in Appendix B. We pay special attention in providing explicit and rather optimal bounds, with an eye in the application of the theorems and in computer assisted proofs.

2. The setting and the KAM theorem

In this section we set the geometrical and analytical background of this paper, and present the main result. In Section 2.1 we establish the basic notation. In Section 2.2 we introduce the geometrical setting of this paper, i.e. symplectic structures on open sets of $\mathbb{R}^{2n}$ that admit compatible almost-complex structures. In Section 2.3 we review standard definitions for Hamiltonian systems and first integrals, with an eye in Hamiltonian actions of tori. In Section 2.4 we set the main equations of this paper, to find invariant tori carrying quasi-periodic motion, and some implications of the presence of first integrals in involution, that lead to a reduction of the dimensionality of the problem of finding Lagrangian invariant cylinders and tori, without the need of reducing the Hamiltonian itself, see Lemma 2.5. In Section 2.5 we review some geometrical constructions and reducibility properties of invariant tori, and present some implications of the presence of compatible triples, see Proposition 2.9. In Section 2.6 we review the quasi-Newton method introduced in [17], and present a new modified quasi-Newton method to solve the invariance equations. In Section 2.7 we introduce the spaces in which the invariance equations are considered, i.e. spaces of analytic functions, and review some key results regarding small divisors equations, see Lemma 2.14 and Corollary 2.15. Finally, in Section 2.8 we present the main result of this paper, Theorem 2.18, an a posteriori theorem on the existence of isotropic invariant tori for Hamiltonian systems with first integrals in involution. The proof of this theorems constitutes the bulk of this paper and is found in Section 3.

2.1. Basic notation

We denote by $\mathbb{R}^m$ and $\mathbb{C}^m$ the vector spaces of m-dimensional vectors with components in $\mathbb{R}$ and $\mathbb{C}$, respectively, endowed with the norm

$$|v| = \max_{i=1,\ldots,m} |v_i|.$$

Given $a, b \in \mathbb{C}$, we also often use the notation $|a, b| = \max\{|a|, |b|\}$. We consider the real and imaginary projections $\mathrm{Re}, \mathrm{Im} : \mathbb{C}^m \to \mathbb{R}^m$, and identify $\mathbb{R}^m \simeq \mathrm{Im}^{-1}\{0\} \subset \mathbb{C}^m$. Given $U \subset \mathbb{R}^m$ and $\rho > 0$, the complex strip of size ρ is $U_\rho = \{\theta \in \mathbb{C}^m \ : \ \mathrm{Re}\,\theta \in U\,,\ |\mathrm{Im}\,\theta| < \rho\}$. Given two sets $X, Y \subset \mathbb{C}^m$, $\mathrm{dist}(X, Y)$ is defined as $\inf\{|x-y| \ : \ x \in X\,,\ y \in Y\}$.

We denote $\mathbb{R}^{n_1\times n_2}$ and $\mathbb{C}^{n_1\times n_2}$ the spaces of $n_1 \times n_2$ matrices with components in $\mathbb{R}$ and $\mathbb{C}$, respectively, identifying $\mathbb{R}^m \simeq \mathbb{R}^{m\times 1}$ and $\mathbb{C}^m \simeq \mathbb{C}^{m\times 1}$. We denote I_n and O_n the $n \times n$ identity and zero matrices, respectively. The $n_1 \times n_2$ zero matrix is represented by $O_{n_1\times n_2}$. Finally, we use the notation 0_n to represent the column vector $O_{n\times 1}$, although we mostly write 0 when the dimension is known from the context. Matrix norms in both $\mathbb{R}^{n_1\times n_2}$ and $\mathbb{C}^{n_1\times n_2}$ are the ones induced from the corresponding vector norms. That is to say, for an $n_1 \times n_2$ matrix M, we have

$$|M| = \max_{i=1,\ldots,n_1} \sum_{j=1,\ldots,n_2} |M_{i,j}|.$$

In particular, if $v \in \mathbb{C}^{n_2}$, $|Mv| \leq |M||v|$. Moreover, $M^\top$ denotes the transpose of the matrix M, so that

$$|M^\top| = \max_{j=1,\ldots,n_2} \sum_{i=1,\ldots,n_1} |M_{i,j}|.$$

Notice that $|M^\top| \leq n_1|M|$ and $|(Mv)^\top| \leq |M^\top|\,|v^\top|$, but also $|(Mv)^\top| \leq n_1|M|\,|v|$ and $|(Mv)^\top| \leq n_2|M^\top|\,|v|$.

Given an analytic function $f : \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}$, defined on an open set $\mathcal{U}$, the action of the r-order derivative of f at a point $x \in \mathcal{U}$ on a collection of vectors $v_1, \ldots, v_r \in \mathbb{C}^m$, with $v_k = (v_{1k}, \ldots, v_{mk})$, is

$$\mathrm{D}^r f(x)[v_1, \ldots, v_r] = \sum_{\ell_1,\ldots,\ell_r} \frac{\partial^r f}{\partial x_{\ell_1} \ldots \partial x_{\ell_r}}(x)\ v_{\ell_1 1} \cdots v_{\ell_r r},$$

where the indices $\ell_1, \ldots, \ell_r$ run from 1 to m. This construction is extended to vector and matrix-valued maps as follows: given a matrix-valued map $M \ : \ \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}^{n_1\times n_2}$ (whose components $M_{i,j}$ are analytic functions), a point $x \in \mathcal{U}$, and a collection of (column) vectors $v_1, \ldots, v_r \in \mathbb{C}^m$, we obtain an $n_1 \times n_2$ matrix $\mathrm{D}^r M(x)[v_1, \ldots, v_r]$ such that

$$\left(\mathrm{D}^r M(x)[v_1, \ldots, v_r]\right)_{i,j} = \mathrm{D}^r M_{i,j}(x)[v_1, \ldots, v_r].$$

Notice that, if we split M in its columns $M_{\cdot,j}$ for $j = 1, \ldots, n_2$, so that $(M_{\cdot,j})_i = M_{i,j}$ for $i = 1, \ldots, n_1$, we have

$$\mathrm{D}^r M(x)[v_1, \ldots, v_r] = \begin{pmatrix} \mathrm{D}^r M_{\cdot,1}(x)[v_1, \ldots, v_r] & \ldots & \mathrm{D}^r M_{\cdot,n_2}(x)[v_1, \ldots, v_r] \end{pmatrix}.$$

For $r = 1$, we will often write $\mathrm{D}M(x)[v] = \mathrm{D}^1 M(x)[v]$ for $v \in \mathbb{C}^m$.

Given a function $f \ : \ \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}^n \simeq \mathbb{C}^{n\times 1}$, we can think of $\mathrm{D}f$ as a matrix function $\mathrm{D}f \ : \ \mathcal{U} \to \mathbb{C}^{n\times m}$, hence, $\mathrm{D}^1 f(x)[v] = \mathrm{D}f(x)v$

for $v \in \mathbb{C}^m$. Therefore, we can apply the transpose to obtain a matrix function $(\mathrm{D}f)^\top$, which acts on n-dimensional vectors, while $\mathrm{D}f^\top = \mathrm{D}(f^\top)$ acts on m-dimensional vectors. Hence, according to the above notation, the operators D and $(\cdot)^\top$ do not commute. Therefore, in order to avoid confusion, we must pay attention to the use of parenthesis.

A function $u : \mathbb{R}^d \to \mathbb{R}$ is 1-periodic if $u(\theta + e) = u(\theta)$ for all $\theta \in \mathbb{R}^d$ and $e \in \mathbb{Z}^d$. Abusing notation, we write $u : \mathbb{T}^d \to \mathbb{R}$, where $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ is the d-dimensional standard torus. Analogously, for $\rho > 0$, a function $u : \mathbb{R}^d_\rho \to \mathbb{C}$ is 1-periodic if $u(\theta + e) = u(\theta)$ for all $\theta \in \mathbb{R}^d_\rho$ and $e \in \mathbb{Z}^d$. We also abuse notation and write $u : \mathbb{T}^d_\rho \to \mathbb{C}$, where $\mathbb{T}^d_\rho = \{\theta \in \mathbb{C}^d/\mathbb{Z}^d \, : \, |\mathrm{Im}\, \theta| < \rho\}$ is the complex strip of $\mathbb{T}^d$ of width $\rho > 0$. We write the Fourier expansion of a periodic function as

$$u(\theta) = \sum_{k\in\mathbb{Z}^d} \hat{u}_k \mathrm{e}^{2\pi \mathrm{i} k\cdot\theta}, \qquad \hat{u}_k = \int_{\mathbb{T}^d} u(\theta)\mathrm{e}^{-2\pi \mathrm{i} k\cdot\theta}\mathrm{d}\theta.$$

and introduce the notation $\langle u\rangle := \hat{u}_0$ for the average. Given $\omega \in \mathbb{R}^d$, we define the operator $\mathfrak{L}_\omega$ acting on u as the Lie derivative of u in the direction of the constant vector field $\dot{\theta} = -\omega$ on the torus:

$$\mathfrak{L}_\omega u = -\mathrm{D}u\, \omega = -\sum_{i=1}^{d} \omega_i \frac{\partial u}{\partial \theta_i}. \tag{2.1}$$

The corresponding Fourier series of $v = \mathfrak{L}_\omega u$ is

$$v(\theta) = \sum_{k\in\mathbb{Z}^d} \hat{v}_k \mathrm{e}^{2\pi \mathrm{i} k\cdot\theta}, \qquad \hat{v}_k = -2\pi \mathrm{i}(k\cdot\omega)\hat{u}_k.$$

The notation in this paragraph is extended to $n_1 \times n_2$ matrix-valued periodic maps $M : \mathbb{T}^d_\rho \to \mathbb{C}^{n_1\times n_2}$, for which $\hat{M}_k \in \mathbb{C}^{n_1\times n_2}$ denotes the Fourier coefficient of index $k \in \mathbb{Z}^d$.

2.2. *Symplectic setting*

In this paper we consider the phase space is an open set U of $\mathbb{R}^{2n}$, whose points are denoted by $z = (z_1, \dots, z_{2n})$, endowed with an *exact symplectic form* $\boldsymbol{\omega} = \mathrm{d}\boldsymbol{\alpha}$, where the 1-form $\boldsymbol{\alpha}$ is called *action form*. The setting and the results of this paper can be easily adapted to other settings such as $U \subset \mathbb{T}^k\times\mathbb{R}^{2n-k}$ with $k \le n$. See e.g. [11] for a discussion.

In order to simplify some of the geometrical constructs of this paper, we also assume that U is endowed with a Riemannian metric $\boldsymbol{g}$ and an anti-involutive linear isomorphism $\boldsymbol{J} : \mathrm{T}U \to \mathrm{T}U$, i.e. $\boldsymbol{J}^2 = -I$, such that $\forall z \in U, \forall u, v \in T_zU$, $\boldsymbol{\omega}_z(\boldsymbol{J}_z u, v) = \boldsymbol{g}_z(u, v)$. It is said that $(\boldsymbol{\omega}, \boldsymbol{g}, \boldsymbol{J})$ is a compatible triple and that $\boldsymbol{J}$ endows U with an almost-complex structure. The anti-involution preserves both 2-forms $\boldsymbol{\omega}$ and $\boldsymbol{g}$.

We rather use the matrix representations of the previous objects, given by the matrix-valued maps $a : U \longrightarrow \mathbb{R}^{2n}$, representing the 1-form $\boldsymbol{\alpha}$, and $\Omega, G, J : U \longrightarrow \mathbb{R}^{2n\times 2n}$, representing the 2-forms $\boldsymbol{\omega}$ and $\boldsymbol{g}$, and the anti-involution $\boldsymbol{J}$, respectively. The fact that $\boldsymbol{\omega}$ is closed reads

$$\frac{\partial \Omega_{r,s}}{\partial z_t} + \frac{\partial \Omega_{s,t}}{\partial z_r} + \frac{\partial \Omega_{t,r}}{\partial z_s} = 0,$$

for any triplet (r, s, t), and the fact that $\boldsymbol{\omega}$ is exact, with $\boldsymbol{\omega} = \mathrm{d}\boldsymbol{\alpha}$ reads

$$\Omega = (\mathrm{D}a)^\top - \mathrm{D}a.$$

Moreover, $\Omega^\top = -\Omega$, and Ω is pointwise invertible. Moreover, the metric condition of $\boldsymbol{g}$ reads $G^\top = G$ and it is positive definite, and the compatibility conditions read

$$J^\top \Omega = -\Omega J = G, \qquad J^2 = -I_{2n}.$$

Notice also that $\Omega = GJ$. These properties also imply the relations

$$\Omega = J^\top \Omega J, \qquad G = J^\top G J.$$

Remark 2.1. The prototype example of compatible triple is $(\boldsymbol{\omega}_0, \boldsymbol{g}_0, \boldsymbol{J}_0)$ in $\mathbb{R}^{2n}$, where $\boldsymbol{\omega}_0$ is the *standard symplectic structure*, $\boldsymbol{\omega}_0 = \sum_{i=1}^n \mathrm{d}z_{n+i} \wedge \mathrm{d}z_i$, $\boldsymbol{g}_0$ is the Euclidian metric, and $\boldsymbol{J}_0$ is the linear complex structure in $\mathbb{R}^{2n}$ (as a real vector space), coming from the complex structure in $\mathbb{C}^n$. The matrix representations of these objects are

$$\Omega_0 = \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix}, \quad G_0 = \begin{pmatrix} I_n & O_n \\ O_n & I_n \end{pmatrix}, \qquad J_0 = \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix}.$$

Moreover, an action form for $\boldsymbol{\omega}_0$ is $\boldsymbol{\alpha}_0 = \frac{1}{2}\sum_{i=1}^n (z_{n+i}\, \mathrm{d}z_i - z_i\, \mathrm{d}z_{n+i})$, which is represented as

$$a_0(z) = \frac{1}{2}\begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix} z\,.$$

Remark 2.2. Even though usually KAM theory is presented in the standard case, a notable counterexample is the seminal paper [10] that was followed by other papers such as [14,23,24]. Even more general constructs were presented in chapter 4 of [11].

2.3. *Hamiltonian systems and first integrals in involution*

Given a function $h : U \to \mathbb{R}$, the corresponding Hamiltonian vector field $X_h : U \to \mathbb{R}^{2n}$ is the one such that $i_{X_h}\boldsymbol{\omega} = -\mathrm{d}h$. In coordinates, the Hamiltonian vector field X_h satisfies

$$X_h(z)^\top \Omega(z) = -\mathrm{D}h(z)\,, \qquad \text{i.e.,} \qquad X_h(z) = \Omega(z)^{-1}(\mathrm{D}h(z))^\top.$$

Using Cartan's magic formula, the Lie derivative of $\boldsymbol{\omega}$ in the direction of X_h vanishes,

$$\mathrm{D}\Omega[X_h] + (\mathrm{D}X_h)^\top \Omega + \Omega\, \mathrm{D}X_h = O_{2n}. \tag{2.2}$$

The Poisson bracket of two functions f, g is given by $\{f, g\} = -\boldsymbol{\omega}(X_f, X_g)$. It is related to the Lie bracket through the well-known formula $[X_f, X_g] = -X_{\{f,g\}}$, that in coordinates is written as

$$\{f, g\} = -(X_f)^\top \Omega\, X_g = \mathrm{D}f\, X_g.$$

In particular, f is a first integral or preserved quantity of X_g if and only $\{f, g\} = 0$, and it is said that f, g are in involution. As a consequence,

$$\mathrm{D}X_f\, X_g = \mathrm{D}X_g\, X_f$$

and the corresponding flows commute.

We assume there is a *moment map* $p : U \to \mathbb{R}^{n-d}$, meaning that its components $p_1, \dots, p_{n-d}$, jointly with h, are pairwise in involution functionally independent functions. We encode the corresponding Hamiltonian vector fields as the columns of $X_p : U \to \mathbb{R}^{2n\times(n-d)}$:

$$X_p = \Omega^{-1}(\mathrm{D}p)^\top.$$

The properties mentioned above are summarized as follows: $\mathrm{D}p\, X_p = 0$, $\mathrm{D}p\, X_h = 0$ (and, then, $\mathrm{D}X_p[X_h] = \mathrm{D}X_h\, X_p$), and the matrix $\begin{pmatrix} X_h(z) & X_p(z) \end{pmatrix}$ has (maximal) rank $n - d + 1$, for any $z \in U$.

For $j = 1, \dots, n-d$, the vector fields X_{p_j} generate (local) flows $\varphi^j : D_j \subset \mathbb{R}\times U \to U$, for which we write $\varphi^j_{s_j}(z) = \varphi^j(s_j, z)$ for $(s_j, z) \in D_j$. The flows commute (and also commute with the flow of X_h). We then define $\boldsymbol{\Phi} : D \subset \mathbb{R}^{n-d} \times U \to U$ as

$$\boldsymbol{\Phi}(s, z) = \varphi^1_{s_1} \circ \cdots \circ \varphi^{n-d}_{s_{n-d}}(z),$$

where

$$D = \{(s, z) \in \mathbb{R}^{n-d}\times U \mid (s_{n-d}, z) \in D_{n-d}, \dots, (s_1, \varphi_{s_2} \circ \dots \circ \varphi^{n-d}_{s_{n-d}}(z)) \in D_1\}.$$

Since

$$\mathrm{D}_s\boldsymbol{\Phi} = X_p \circ \boldsymbol{\Phi},$$

we say that $\boldsymbol{\Phi}$ is the (local) *moment flow* associated to X_p, and write $\boldsymbol{\Phi}_s(z) = \boldsymbol{\Phi}(s, z)$. Notice also that

$$\mathrm{D}_z\boldsymbol{\Phi}\, X_p = X_p \circ \boldsymbol{\Phi},$$

and

$$\mathrm{D}_z\boldsymbol{\Phi}\, X_h = X_h \circ \boldsymbol{\Phi}.$$

The case $d = 1$ corresponds to the integrable case, and from now on we will assume $d > 1$. The case $d = n$ corresponds to not assuming the existence of first integrals in involution, and all the results of this paper hold for such a case.

Remark 2.3. An important special case is when the moment map p induces a Hamiltonian torus action, that is $D = \mathbb{R}^{n-d} \times U$ and the generated flow $\Phi : \mathbb{R}^{n-d} \times U \to U$ is periodic in all components of s. By scaling *times* we can get all periods equal to one, and then, with a slight abuse of notation, consider $\Phi : \mathbb{T}^{n-d} \times U \to U$.

2.4. Invariant tori

In this paper, we refer to an embedding $K : \mathbb{T}^d \to U$ as a parameterization of the torus $\mathcal{K} = K(\mathbb{T}^d)$.

Given $\omega \in \mathbb{R}^d$ with $1 \leq d \leq n$, we say that K is invariant for X_h with frequency vector ω if

$$X_h \circ K + \mathfrak{L}_\omega K = 0. \tag{2.3}$$

This means that the d-dimensional torus $\mathcal{K} = K(\mathbb{T}^d)$ is invariant and the internal dynamics is given by the constant vector field $\dot\theta = \omega$. Eq. (2.3) is called *invariance equation for K and frequency ω*. Case $d = 1$ corresponds to $\mathcal{K}$ being a periodic orbit. We then assume $d \geq 2$ and ω to be *ergodic*: for $k \in \mathbb{Z}^d \backslash \{0\}, k \cdot \omega \neq 0$. Hence, the flow on the torus is quasi-periodic and non-resonant.

Remark 2.4. Other topologies can be considered for both the ambient manifold U and the torus $\mathcal{K}$. See e.g. [11] for a discussion.

Ergodicity of ω implies additional geometric and dynamical properties of the torus $\mathcal{K}$. In particular, it is *isotropic*, i.e. the pullback $K^*\boldsymbol{\omega}$ is zero (see [25,26]), and it is contained in an energy level of the Hamiltonian and of the additional first integrals (if any):

$$h \circ K = \langle h \circ K \rangle, \quad p \circ K = \langle p \circ K \rangle,$$

since $\mathfrak{L}_\omega(h \circ K) = 0$ and $\mathfrak{L}_\omega(p \circ K) = 0$.

Another consequence of the presence of extra first integrals in involution is that an invariant torus $\mathcal{K}$, with frequency vector ω, induces a family of invariant tori $\mathcal{K}_s = \Phi_s(\mathcal{K})$, with the same frequency vector (with s defined, a priori, in an open neighborhood B of $0 \in \mathbb{R}^{n-d}$). If $K_s = \Phi_s \circ K$ is the parameterization of $\mathcal{K}_s$:

$$X_h \circ K_s + \mathfrak{L}_\omega K_s = (\mathrm{D}\Phi_s) \circ K \; X_h \circ K + (\mathrm{D}\Phi_s) \circ K \; \mathfrak{L}_\omega K = 0.$$

We can think the family foliating an n-dimensional invariant object $\hat{\mathcal{K}}$ parameterized by $\hat K : \mathbb{T}^d \times B \to U$ defined as

$$\hat K(\theta, s) = \Phi_s(K(\theta)).$$

Notice that we can rephrase the previous argumentation by writing

$$X_h \circ \hat K + \mathfrak{L}_{(\omega, 0)} \hat K = 0.$$

This argument of getting n-dimensional invariant objects (including n-dimensional invariant tori) from d-dimensional invariant tori is extended in the following lemma.

Lemma 2.5 (*Reduction Lemma*). *Let $h : U \to \mathbb{R}$ be a Hamiltonian for which there exists a moment map $p : U \to \mathbb{R}^{n-d}$, being $\Phi : D \subset \mathbb{R}^{n-d} \times U \to U$ the corresponding moment flow. Let $f : p(U) \subset \mathbb{R}^{n-d} \to \mathbb{R}$ be a function defined on the set of possible momenta, and let $\hat h : U \to \mathbb{R}$ be the discounted Hamiltonian, defined as $\hat h = h - f \circ p$. Let $K : \mathbb{T}^d \to U$ be a parameterization of a torus $\mathcal{K} = K(\mathbb{T}^d)$, invariant for $X_{\hat h}$ with ergodic frequency $\omega \in \mathbb{R}^d$, i.e.*

$$X_{\hat h} \circ K + \mathfrak{L}_\omega K = 0. \tag{2.4}$$

Let B be an open neighborhood of $0 \in \mathbb{R}^{n-d}$ such that $B \times \mathcal{K} \subset D$. Define $\hat K : \mathbb{T}^d \times B \to U$ as

$$\hat K(\theta, s) = \Phi(s, K(\theta)),$$

and $p_0 = \langle p \circ K \rangle$, $\omega_{p_0} = (Df(p_0))^\top$. Then,

$$X_h \circ \hat K + \mathfrak{L}_{(\omega, \omega_{p_0})} \hat K = 0. \tag{2.5}$$

That is $\hat{\mathcal{K}} = \hat K(\mathbb{T}^d \times B)$ is an invariant object for X_h.

Moreover, if for a certain $\theta_0 \in \mathbb{T}^d$ there exists $S \in B$ (with positive coordinates S_i) such that $\hat K(\theta_0, S) = K(\theta_0)$, and $X_p \circ K$ and K are transversal, then for all $\theta \in \mathbb{T}^d$ we have $\hat K(\theta, S) = K(\theta)$, we can take $B = \mathbb{R}^{n-d}$ and $\hat K$ is S-periodic in the s variables (i.e., $\hat K(\theta, s + S) = \hat K(\theta)$). That is $\hat{\mathcal{K}} = \hat K(\mathbb{T}^d \times \mathbb{R}^{n-d})$ is an n-dimensional invariant torus for X_h.

Proof. The hypotheses (2.4) of invariance of K for the vector field $X_{\hat h}$ reads

$$0 = X_h \circ K - X_p \circ K \; (\mathrm{D} f \circ p \circ K)^\top + \mathfrak{L}_\omega K.$$

Then, since

$$\mathfrak{L}_\omega (p \circ K) = \mathrm{D}p \circ K \; \mathfrak{L}_\omega K = -\mathrm{D}p \circ K \; (X_h \circ K - X_p \circ K \; (\mathrm{D} f \circ p \circ K)^\top) = 0,$$

and ω is ergodic, then $p \circ K = \langle p \circ K \rangle = p_0$ is constant.

Then, since $\mathrm{D}_z \Phi_s(z) X_h(z) = X_h(\Phi_s(z))$ and $\mathrm{D}_z \Phi_s(z) X_p(z) = X_p(\Phi_s(z)) = \mathrm{D}_s \Phi_s(z)$, we get

$$\begin{aligned} X_h \circ \hat K + \mathfrak{L}_{(\omega, \omega_{p_0})} \hat K &= X_h \circ \Phi_s \circ K + \mathrm{D}_z \Phi_s \circ K \; \mathfrak{L}_\omega K - X_p \circ \Phi_s \circ K \; \omega_{p_0} \\ &= \mathrm{D}_z \Phi_s \circ K \; (X_h \circ K - X_p \circ K \; (\mathrm{D} f \circ p \circ K)^\top + \mathfrak{L}_\omega K) = 0, \end{aligned}$$

proving (2.5).

Let us assume now that $K(\theta_0) = \hat K(\theta_0, S) = \Phi_S(K(\theta_0))$. Denote $\hat\varphi_t$ the (local) flow of $X_{\hat h}$. Then, for all $t \in \mathbb{R}$,

$$K(\theta_0 + \omega t) = \hat\varphi_t(K(\theta_0)) = \hat\varphi_t(\Phi_S(K(\theta_0))) = \Phi_S(\hat\varphi_t(K(\theta_0))) = \Phi_S(K(\theta_0 + \omega t)),$$

and the result follows from ergodicity of ω. ☐

Remark 2.6. We can think of K as the generator of the n-dimensional object $\hat{\mathcal{K}}$. The generator is defined up to a change of phase, say $\alpha \in \mathbb{R}^d$, and flying times, say $\beta \in B$, since the parameterization $K_{\alpha,\beta}$ defined as

$$K_{\alpha,\beta}(\theta) = \Phi_\beta(K(\theta + \alpha)),$$

is also a generator of $\hat{\mathcal{K}}$. This indeterminacy of the generator can be fixed by imposing extra conditions apart from the invariance equation. Some strategies are described in [17].

Remark 2.7. In case of existence of $\theta_0 \in \mathbb{T}^d$ such that $\hat K(\theta_0, S) = K(\theta_0)$ for a certain $S \in \mathbb{R}^{n-d}$, Lemma 2.5 establishes the existence of an n-dimensional invariant torus $\hat{\mathcal{K}}$, with frequency vector $\tilde\omega = (\omega, \omega_{p_0}/S)$ (we understand here that the division of the two vectors is made componentwise). In other words, by scaling periods we get a parameterization $\tilde K : \mathbb{T}^d \times \mathbb{T}^{n-d} \to U$ defined as

$$\tilde K(\theta, \vartheta) = \hat K(\theta, S\vartheta)$$

(again considering the product $S\vartheta$ componentwise), that satisfies

$$X_h \circ \tilde K + \mathfrak{L}_{\tilde\omega} \tilde K = 0.$$

Remark 2.8. In case the moment map p induces a Hamiltonian torus action, see Remark 2.3, we can consider the moment flow as a map $\Phi : \mathbb{T}^{n-d} \times U \to U$ (this is in fact the Hamiltonian action of $\mathbb{T}^{n-d}$ on U). In particular, for a fixed complementary frequency $\omega_p \in \mathbb{R}^{n-d}$, to which we will refer to as a *moment frequency*, we consider the discounted Hamiltonian $\hat h_{\omega_p} = h - p^\top \omega_p$. Then, a parameterization $K : \mathbb{T}^d \to U$ invariant for $X_{\hat h_{\omega_p}}$ with frequency ω induces a parameterization $\tilde K : \mathbb{T}^n \to U$ invariant for X_h with frequency $\tilde\omega = (\omega, \omega_p)$. We can also think of the moment frequency ω_p as a parameter, and the discounted Hamiltonian $\hat h_{\omega_p}$ as a family of Hamiltonians. From the results of this paper, one can obtain $n - d$ parametric families of n-dimensional tori with fixed *internal frequency* ω, parameterized by ω_p.

2.5. Linearized dynamics and reducibility

In this section we describe the geometric construction of a suitable symplectic frame attached to a torus $\mathcal{K}$ with respect to a Hamiltonian system X_h, possibly with first integrals p, and an ergodic frequency $\omega \in \mathbb{R}^d$.

In the following, for a matrix-valued map $V : \mathbb{T}^d \to \mathbb{R}^{2n\times m}$, with $1 \leq m \leq 2n$, we introduce the matrix-valued map $\mathcal{X}_V : \mathbb{T}^d \to \mathbb{R}^{2n\times m}$ defined as

$$\mathcal{X}_V := \mathrm{D}X_h \circ K\, V + \mathfrak{L}_\omega V. \tag{2.6}$$

If we think of V as a parameterization of a frame of an m-dimensional vector bundle $\mathcal{V}$, then $\mathcal{X}_V$ corresponds to its infinitesimal displacement by the flow of X_h around $\mathcal{K}$. We say that V is invariant under (the linearized equations of) X_h if $\mathcal{X}_V = O_{2n\times m}$. We also define the pullback of Ω and G of V on K to be the matrix-valued maps $\Omega_V, G_V : \mathbb{T}^d \to \mathbb{R}^{m\times m}$ defined as

$$\Omega_V = V^\top \Omega \circ K\, V,$$

and

$$G_V = V^\top G \circ K\, V,$$

respectively. We say that V is isotropic if $\Omega_V = O_m$ and Lagrangian if, moreover, $m = n$. We also say that V is orthonormal if $G_V = I_m$.

We consider the matrix-valued map $L : \mathbb{T}^d \to \mathbb{R}^{2n\times n}$ given by

$$L = \begin{pmatrix} \mathrm{D}K & X_p \circ K \end{pmatrix}, \tag{2.7}$$

and we assume that $\operatorname{rank} L(\theta) = n$ for every $\theta \in \mathbb{T}^d$. Notice that, while $\mathrm{D}K$ parameterizes the tangent bundle of $\mathcal{K}$, $T\mathcal{K}$, L parameterizes a subbundle $\mathcal{L}$ of rank n of the bundle $\mathrm{T}_\mathcal{K} U$. We refer to L as the tangent frame (attached to K).

One can use the geometric structure of the problem to complement the above frame. From the several choices one could do (see e.g. [11] for a discussion), we consider here the one that is specially tailored for a compatible triple. In particular, we define $N : \mathbb{T}^d \to \mathbb{R}^{2n\times n}$ as

$$N = J \circ K\, L\, B$$

where

$$B = G_L^{-1}.$$

Notice that N parameterizes another subbundle $\mathcal{N}$ of rank n of $\mathrm{T}_\mathcal{K} U$. We refer to N as the normal frame (attached to K).

Finally, we define the matrix-valued map $P : \mathbb{T}^d \to \mathbb{R}^{2n\times 2n}$ as the juxtaposition of matrix-valued maps $L, N : \mathbb{T}^d \to \mathbb{R}^{2n\times n}$:

$$P = \begin{pmatrix} L & N \end{pmatrix}. \tag{2.8}$$

We refer to P as an (adapted) frame (attached to K).

We define then the torsion of the parameterization K (with respect to the Hamiltonian h) to be the matrix-valued map $T : \mathbb{T}^d \to \mathbb{R}^{n\times n}$ defined as

$$T = N^\top\, T_h \circ K\, N, \tag{2.9}$$

where $T_h : U \to \mathbb{R}^{2n\times 2n}$ is the torsion of the Hamiltonian h, defined as

$$\begin{aligned} T_h &= \Omega \left(\mathrm{D}X_h + \mathrm{D}J[X_h]\, J + J\mathrm{D}X_h J \right) \\ &= \Omega\, \mathrm{D}X_h - J^\top \Omega\, \mathrm{D}X_h J + J^\top \Omega\, \mathrm{D}J[X_h], \end{aligned}$$

where we use that $J^2 = -I_{2n}$, so $\mathrm{D}J[X_h]J = -J\mathrm{D}J[X_h]$, and $G = -\Omega J = J^\top \Omega$. Notice that T_h is symmetric:

$$\begin{aligned} T_h - T_h^\top &= \Omega\, \mathrm{D}X_h - J^\top \Omega\, \mathrm{D}X_h J + J^\top \Omega\, \mathrm{D}J[X_h] \\ &\quad + (\mathrm{D}X_h)^\top \Omega - J^\top (\mathrm{D}X_h)^\top \Omega J + \mathrm{D}J^\top [X_h]\, \Omega J \\ &= -\mathrm{D}\Omega[X_h] + J^\top \mathrm{D}\Omega[X_h] J + J^\top \Omega\, \mathrm{D}J[X_h] + \mathrm{D}J^\top[X_h]\, \Omega J \\ &= -\mathrm{D}\Omega[X_h] + \mathrm{D}(J^\top \Omega J)[X_h] = O_{2n}. \end{aligned}$$

As a result, the torsion T of the parameterization K is symmetric.

The use of the frame P has several advantages. Among them, it produces a natural and geometrically meaningful non-degeneracy condition (twist condition, that is the invertibility of the average of the torsion T) in the KAM theorem, and, most importantly, when the torus is invariant, it reduces the linearized dynamics to a block-triangular form.

Proposition 2.9. *If K is invariant for X_h with ergodic frequency ω, then:*

(1) P is symplectic:

$$P^\top \Omega \circ K\, P = \Omega_0,$$

and, in particular, L and N parameterize complementary Lagrangian bundles ($\mathrm{T}_\mathcal{K} U = \mathcal{L} \oplus \mathcal{N}$);

(2) P reduces $\mathrm{D}X_h \circ K$ to the block-triangular form $\Lambda : \mathbb{T}^d \to \mathbb{R}^{2n\times 2n}$ given by

$$\Lambda = \begin{pmatrix} O_n & T \\ O_n & O_n \end{pmatrix}, \tag{2.10}$$

where T is defined in (2.9), so

$$\mathrm{D}X_h \circ K\, P + \mathfrak{L}_\omega P = P\, \Lambda. \tag{2.11}$$

Proof. The facts that L is invariant, i.e. $\mathcal{X}_L = O_{2n\times n}$, and Lagrangian, i.e. $\Omega_L = O_n$, follow directly from the invariance of K and the fact that the frequency ω is ergodic. The construction leads to the Lagrangianity of N, i.e. $\Omega_N = O_n$, and the fact that P is symplectic. The construction also leads to the reducibility property (2.11) with

$$T = N^\top \Omega \circ K\, \mathcal{X}_N.$$

See e.g. [17]. Finally, $T = N^\top T_h N$ follows from the computation:

$$\begin{aligned} \mathfrak{L}_\omega N &= (\mathrm{D}J)\circ K[\mathfrak{L}_\omega K]\; LB + J\circ K\; \mathfrak{L}_\omega L\; B + J\circ K\; L\; \mathfrak{L}_\omega B \\ &= (\mathrm{D}J)\circ K[-X_h\circ K]\; LB - J\circ K\; \mathrm{D}X_h\circ K L\; B + J\circ K\; L\; \mathfrak{L}_\omega B \\ &= (\mathrm{D}J)\circ K[X_h\circ K]\; J\circ K\; N + J\circ K\; \mathrm{D}X_h\circ K\; J\circ K N + J\circ K\; L\; \mathfrak{L}_\omega B. \quad \square \end{aligned}$$

Remark 2.10. The torsion measures the symplectic area determined by the normal bundle and its infinitesimal displacement. Notice that, in the present paper, the torsion involves geometrical and dynamical properties of both the torus and the first integrals.

2.6. A (modified) quasi-Newton method

In this section we outline a (modified) quasi-Newton method to obtain a solution of the invariance Eq. (2.3) from an initial approximation. Sufficient conditions of the convergence of the method are provided in Theorem 2.18, whose proof is detailed in Section 3. We focus here on the geometry of the method.

Assume then we are given a parameterization $K : \mathbb{T}^d \to U$. The error of the invariance is $E : \mathbb{T}^d \to \mathbb{R}^{2n}$, defined as

$$E = X_h \circ K + \mathfrak{L}_\omega K.$$

We may obtain a new parameterization

$$\bar{K} = K + \Delta K$$

by considering the linearized equation

$$\mathrm{D}X_h \circ K\, \Delta K + \mathfrak{L}_\omega \Delta K = -E. \tag{2.12}$$

If the error E is sufficiently small and we obtain a good enough approximation of the solution ΔK of (2.12), then $\bar{K}$ provides a new parameterization with a new error $\bar{E}$ which is quadratically small in terms of E. To do so, following the nowadays standard practice [10,11] we may resort to a frame P (here the one defined in (2.8)), so by writing

$$\Delta K = P\, \xi = L \xi^L + N \xi^N, \tag{2.13}$$

$\xi = (\xi^L, \xi^N) : \mathbb{T}^d \to \mathbb{R}^n \times \mathbb{R}^n$ is the new unknown. We think of the components of ξ as the tangent and normal components of the correction. The new approximation is

$$\bar K = K + L\xi^L + N\xi^N. \tag{2.14}$$

Following e.g. [17,23], using (2.13), multiplying both sides of (2.12) by $\Omega_0^{-1} P^\top \Omega\circ K$ (an approximation of P^{-1}), and skipping second order small terms (in the error E and correction ξ) one reaches the block-triangular system

$$\Lambda\,\xi + \mathfrak{L}_\omega \xi = \eta, \tag{2.15}$$

(see (2.9) and (2.10)), where

$$\eta = -\Omega_0^{-1} P^\top \Omega\circ K\; E. \tag{2.16}$$

We think of the components of $\eta = (\eta^L, \eta^N) : \mathbb{T}^d \to \mathbb{R}^n \times \mathbb{R}^n$ as the tangent and normal components (of the negative) of the error E.

Inspired by [18], we may also consider the new approximation as

$$\bar{\bar K} = \Phi_{\xi^L_{X_p}} \circ (K + N\xi^N)\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) \tag{2.17}$$

where the unknowns are $\xi^L = (\xi^L_{\mathrm{D}K}, \xi^L_{X_p}) : \mathbb{T}^d \to \mathbb{R}^d \times \mathbb{R}^{n-d}$ and $\xi^N : \mathbb{T}^d \to \mathbb{R}^n$ and id is the identity map on $\mathbb{T}^d$. By Taylor expanding the previous expression, we get

$$\begin{aligned}\bar{\bar K} &= K + N\xi^N + \mathrm{D}K\;\xi^L_{\mathrm{D}K} + \mathrm{D}_s\Phi_0\circ K\;\xi^L_{X_p} + hot\\ &= K + N\xi^N + \mathrm{D}K\;\xi^L_{\mathrm{D}K} + X_p\circ K\;\xi^L_{X_p} + hot\\ &= \bar K + hot,\end{aligned}$$

where, as usual, *hot* stands for *higher order terms*. Hence, the approximations (2.14) and (2.17) differ in quadratically small terms, and the correction terms $\xi = (\xi^L, \xi^N)$ may be computed by solving the triangular system (2.15).

It turns out that the triangular system (2.15), requires to solve two cohomological equations consecutively. More specifically, the system is

$$T\xi^N + \mathfrak{L}_\omega \xi^L = \eta^L, \tag{2.18}$$

$$\mathfrak{L}_\omega \xi^N = \eta^N, \tag{2.19}$$

where

$$\begin{aligned}\eta^L &= -N^\top \Omega\circ K\; E,\\ \eta^N &= L^\top \Omega\circ K\; E,\end{aligned}$$

and the torsion T is given by (2.9).

In the solution of the triangular system it is crucial the fact that the average of the normal component of the error, η^N, is zero. Notice that

$$\eta^N = \begin{pmatrix} (\mathrm{D}K)^\top \Omega\circ K\; E \\ (X_p\circ K)^\top \Omega\circ K\; E \end{pmatrix} = \begin{pmatrix} (\mathrm{D}(h\circ K))^\top - \Omega_{\mathrm{D}K}\,\omega \\ \mathrm{D}(p\circ K)\,\omega \end{pmatrix}.$$

The fact that $\langle \Omega_{\mathrm{D}K} \rangle = O_d$ follows directly from the exact symplectic structure, since $K^*\boldsymbol{\omega} = \mathrm{d}(K^*\boldsymbol{\alpha})$, from where $\langle \eta^N \rangle = 0$ follows immediately. See [17].

Quantitative estimates for the solutions of such equations (under Diophantine conditions of ω) are obtained by applying Rüssmann estimates, see Lemma 2.14. Here we just mention that if we denote by $\mathfrak{R}_\omega v$ the only zero-average solution u of equation $\mathfrak{L}_\omega u = v - \langle v \rangle$ (see (2.21) for an explicit expression of $\mathfrak{R}_\omega v$), then, since $\langle \eta^N \rangle = 0$ (see the *compatibility condition* above) and $\langle T \rangle$ is invertible (the so-called *twist condition*), $\xi = (\xi^L, \xi^N)$ is given by

$$\begin{aligned}\xi^N &= \langle T\rangle^{-1}\langle \eta^L - T\mathfrak{R}_\omega \eta^N\rangle + \mathfrak{R}_\omega \eta^N,\\ \xi^L &= \mathfrak{R}_\omega(\eta^L - T\xi^N),\end{aligned}$$

is the solution of the system (2.18),(2.19) with $\langle \xi^L \rangle = 0_n$. While the average of the normal correction is selected to solve the equation for ξ^L, and it is $\hat\xi^N_0 = \langle \xi^N \rangle = \langle T\rangle^{-1}\langle \eta^L - T\mathfrak{R}_\omega \eta^N\rangle$, we have the freedom of choosing any value for the average of the tangent correction, $\langle \xi^L \rangle = \hat\xi^L_0 \in \mathbb{R}^n$. This is related with the freedom to select a particular generator of the n-dimensional torus, which is a d-dimensional torus, and a particular phase, see Remark 2.23. For the sake of simplicity, we select the solution with $\hat\xi^L_0 = \langle \xi^L \rangle = 0_n$.

Recapitulating, for future reference we describe one step of the (two) quasi-Newton methods described above as follows:

(1) Compute the error of invariance:

$$E = X_h\circ K + \mathfrak{L}_\omega K.$$

(2) Compute the tangent and normal frames to the torus:

$$\begin{aligned}L &= \begin{pmatrix}\mathrm{D}K & X_p\circ K\end{pmatrix},\\ N &= J\circ K\; L\, B,\end{aligned}$$

where

$$B = (L^\top G\circ K\; L)^{-1}.$$

(3) Compute the tangent and normal components of the error of invariance:

$$\begin{aligned}\eta^L &= -N^\top \Omega\circ K\; E,\\ \eta^N &= L^\top \Omega\circ K\; E.\end{aligned}$$

(4) Compute the torsion:

$$T = N^\top\; T_h\circ K\;\, N,$$

where

$$T_h = \Omega\left(\mathrm{D}X_h + \mathrm{D}J[X_h]\; J + J\mathrm{D}X_h J\right).$$

(5) Compute the tangent and normal components of the correction, provided that $\langle T \rangle$ is invertible:

$$\begin{aligned}\hat\xi^N_0 &= \langle T\rangle^{-1}\langle \eta^L - T\mathfrak{R}_\omega \eta^N\rangle,\\ \xi^N &= \hat\xi^N_0 + \mathfrak{R}_\omega \eta^N,\\ \xi^L &= \mathfrak{R}_\omega(\eta^L - T\xi^N),\end{aligned}$$

where $\mathfrak{R}_\omega$ is the Rüssmann operator (see Lemma 2.14).

(6) Compute the new approximation as:

- (quasi-Newton method)

  $$\bar K = K + L\xi^L + N\xi^N;$$

- (modified quasi-Newton method)

  $$\bar{\bar K} = \Phi_{\xi^L_{X_p}} \circ (K + N\xi^N)\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}).$$

As it is usual in the a-posteriori approach to KAM theory, the argument consists in refining an initial approximation K by means of the iterative method and proving the convergence to a solution of the invariance equation. The proof of the corresponding theorem to the quasi-Newton method is in [17]. The goal of this paper is proving such a result for the modified version.

Remark 2.11. We have seen that the difference of the two approaches (2.14) and (2.17) are quadratically small. As we will see, the way the correction ξ^L is included in (2.17) results in a better behavior of the analyticity properties. This is the main idea in [18], in which the invariance equation is replaced by three conditions which are altogether equivalent to invariance. Instead, we will keep the invariance equation formulation.

Remark 2.12. An important feature of the quasi-Newton method is that it can be implemented in a computer. See e.g. [11] for some details of implementations in similar contexts, using FFT. In this respect, it seems that the approach (2.14) is better than (2.17), since the second involves compositions of periodic functions that, in general, are approximated by truncated Fourier series. Composition of periodic functions is much harder computationally than multiplications of periodic functions, that can be done using Fast Fourier Transform.

Remark 2.13. Approaches as (2.14) have been implemented as computer assisted proofs [14], for invariant KAM tori for exact symplectic maps. We think that the new approach (2.17) could result in better posed analytical bounds, that could improve the efficiency of the computer assisted proofs.

2.7. Analytic setting

The proof of the convergence of the algorithm is presented in the analytic category. Hence, we work with real analytic functions defined in complex neighborhoods of real domains. We consider the sup-norms of (matrix-valued) analytic maps and their derivatives (see the notation in Section 2.1). That is, for $f : \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}$, we consider

$$\|f\|_{\mathcal{U}} = \sup_{x\in\mathcal{U}} |f(x)|,$$

and

$$\|\mathrm{D}^r f\|_{\mathcal{U}} = \sum_{\ell_1,\dots,\ell_r} \left\| \frac{\partial^r f}{\partial x_{\ell_1} \dots \partial x_{\ell_r}} \right\|_{\mathcal{U}},$$

that could be infinite. For $M : \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}^{n_1\times n_2}$, we consider the norms

$$\|M\|_{\mathcal{U}} = \max_{i=1,\dots,n_1} \sum_{j=1,\dots,n_2} \left\|M_{i,j}\right\|_{\mathcal{U}},$$

$$\|\mathrm{D}^r M\|_{\mathcal{U}} = \max_{i=1,\dots,n_1} \sum_{j=1,\dots,n_2} \left\|\mathrm{D}^r M_{i,j}\right\|_{\mathcal{U}},$$

and notice, of course, that the norms $\|M^\top\|_{\mathcal{U}}$ and $\|\mathrm{D}^r M^\top\|_{\mathcal{U}}$ are obtained simply by interchanging the role of the indices i and j.

The above norms present Banach algebra-like properties. For example, given r analytic functions $v_1, \dots, v_r : \mathcal{U} \to \mathbb{C}^m \simeq \mathbb{C}^{m\times 1}$, the function $\mathrm{D}^r M[v_1, \dots, v_r] : \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}^{n_1\times n_2}$ defined as

$$\mathrm{D}^r M[v_1, \dots, v_r](x) = \mathrm{D}^r M(x)[v_1(x), \dots, v_r(x)]$$

is also analytic and satisfies

$$\|\mathrm{D}^r M[v_1, \dots, v_r]\|_{\mathcal{U}} \le \|\mathrm{D}^r M\|_{\mathcal{U}}\ \|v_1\|_{\mathcal{U}} \cdots \|v_r\|_{\mathcal{U}}.$$

There is also a similar bound for the action of the transpose:

$$\left\|(\mathrm{D}^r M[v_1, \dots, v_r])^\top\right\|_{\mathcal{U}} \le \left\|\mathrm{D}^r M^\top\right\|_{\mathcal{U}}\ \|v_1\|_{\mathcal{U}} \cdots \|v_r\|_{\mathcal{U}}.$$

In addition, given $M_1 : \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}^{n_1\times n_3}$ and $M_2 : \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}^{n_3\times n_2}$, we have

$$\|M_1 M_2\|_{\mathcal{U}} \le \|M_1\|_{\mathcal{U}}\ \|M_2\|_{\mathcal{U}},$$

and

$$\|\mathrm{D}(M_1 M_2)\|_{\mathcal{U}} \le \|\mathrm{D}M_1\|_{\mathcal{U}}\ \|M_2\|_{\mathcal{U}} + \|M_1\|_{\mathcal{U}}\ \|\mathrm{D}M_2\|_{\mathcal{U}}.$$

In particular, if $f : \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}^{n_1}$, we may take $n_2 = 1$ and consider the matrix constructions just made. One recover bounds such as

$$\|\mathrm{D}f\ V\|_{\mathcal{U}} \le \|\mathrm{D}f\|_{\mathcal{U}}\ \|V\|_{\mathcal{U}},$$

or

$$\left\|(\mathrm{D}f\ V)^\top\right\|_{\mathcal{U}} \le \left\|(\mathrm{D}f)^\top\right\|_{\mathcal{U}} \left\|V^\top\right\|_{\mathcal{U}}, \qquad \left\|(\mathrm{D}f\ V)^\top\right\|_{\mathcal{U}} \le \left\|\mathrm{D}f^\top\right\|_{\mathcal{U}} \|V\|_{\mathcal{U}},$$

where $V : \mathcal{U} \subset \mathbb{C}^m \to \mathbb{C}^{n_1\times n_3}$. Notice we obtain two possible upper bounds for $\|(\mathrm{D}f\ V)^\top\|_{\mathcal{U}}$.

Spaces of periodic real-analytic functions

The particular case of real-analytic periodic functions deserves some additional definitions and comments. We denote by $\mathcal{A}(\mathbb{T}^d_\rho)$ the Banach space of holomorphic functions $u : \mathbb{T}^d_\rho \to \mathbb{C}$, that can be continuously extended to $\bar{\mathbb{T}}^d_\rho$ (the closure), and such that $u(\mathbb{T}^d) \subset \mathbb{R}$ (real-analytic), endowed with the norm

$$\|u\|_\rho = \|u\|_{\mathbb{T}^d_\rho} = \max_{|\mathrm{Im}\,\theta|\le\rho} |u(\theta)|.$$

We also denote by $\mathcal{A}_{C^r}(\mathbb{T}^d_\rho)$ the Banach space of holomorphic functions $u : \mathbb{T}^d_\rho \to \mathbb{C}$ whose partial derivatives up to order r can be continuously extended to $\bar{\mathbb{T}}^d_\rho$, and such that $u(\mathbb{T}^d) \subset \mathbb{R}$, endowed with the norm

$$\|u\|_{\rho,C^r} = \max_{k=0,\dots,r} \left\|\mathrm{D}^k u\right\|_{\mathbb{T}^d_\rho}.$$

As usual in the analytic setting, we use *Cauchy estimates* to control the derivatives of a function. Given $u \in \mathcal{A}(\mathbb{T}^d_\rho)$, with $\rho > 0$, then for any $0 < \delta < \rho$ the partial derivative $\partial u/\partial\theta_\ell$ belongs to $\mathcal{A}(\mathbb{T}^d_{\rho-\delta})$ and we have the estimates

$$\left\|\frac{\partial u}{\partial\theta_\ell}\right\|_{\rho-\delta} \le \frac{1}{\delta}\,\|u\|_\rho, \qquad \|\mathrm{D}u\|_{\rho-\delta} \le \frac{d}{\delta}\,\|u\|_\rho, \qquad \left\|(\mathrm{D}u)^\top\right\|_{\rho-\delta} \le \frac{1}{\delta}\,\|u\|_\rho.$$

The above definitions and estimates extend naturally to matrix-valued maps, that is, given $M : \mathbb{T}^d_\rho \to \mathbb{C}^{n_1\times n_2}$, with components in $\mathcal{A}(\mathbb{T}^d_\rho)$, we have

$$\|\mathrm{D}M\|_{\rho-\delta} = \max_{i=1,\dots,n_1} \sum_{j=1,\dots,n_2} \left\|\mathrm{D}M_{i,j}\right\|_{\rho-\delta} \le \frac{d}{\delta}\,\|M\|_\rho.$$

A direct consequence is that $\|\mathrm{D}M^\top\|_{\rho-\delta} \le \frac{d}{\delta}\,\|M^\top\|_\rho \le \frac{d\,n_1}{\delta}\,\|M\|_\rho$. In particular, given a real analytic vector function $w : \mathbb{T}^d_\rho \to \mathbb{C}^n \simeq \mathbb{C}^{n\times 1}$, we have:

$$\|\mathrm{D}w\|_{\rho-\delta} \le \frac{d}{\delta}\,\|w\|_\rho, \quad \left\|\mathrm{D}w^\top\right\|_{\rho-\delta} \le \frac{d}{\delta}\left\|w^\top\right\|_\rho \le \frac{nd}{\delta}\,\|w\|_\rho,$$
$$\left\|(\mathrm{D}w)^\top\right\|_{\rho-\delta} \le \frac{n}{\delta}\,\|w\|_\rho.$$

(Recall that $\mathrm{D}w^\top$ and $(\mathrm{D}w)^\top$ are different, since operators D and $(\cdot)^\top$ do not commute. See Section 2.1.)

Small divisors equations and Diophantine vectors

Another ingredient in KAM theory are the estimates of the solutions of the *small divisors equations*. Given $\omega \in \mathbb{R}^d$, ergodic, and a real-analytic periodic function $v \in \mathcal{A}(\mathbb{T}^d_\rho)$, we consider the equation

$$\mathfrak{L}_\omega u = v - \langle v\rangle, \tag{2.20}$$

where $\mathfrak{L}_\omega$ is defined in (2.1). Expanding in Fourier series, the only zero-average solution $u = \mathfrak{R}_\omega v$ of (2.20) is be

$$\mathfrak{R}_\omega v(\theta) = \sum_{k\in\mathbb{Z}^d\backslash\{0\}} \hat{u}_k e^{2\pi i k\cdot\theta}, \qquad \hat{u}_k = \frac{-\hat{v}_k}{2\pi i\, k\cdot\omega}, \tag{2.21}$$

and all the other solutions of (2.20) are of the form $u = \hat{u}_0 + \mathfrak{R}_\omega v$ with $\hat{u}_0 \in \mathbb{R}$.

The convergence of the expansion (2.21) is implied by a Diophantine condition on ω. Specifically, for given $\gamma > 0$ and $\tau \ge d-1$, we denote the set of Diophantine vectors

$$\mathcal{D}^d_{\gamma,\tau} = \left\{\omega \in \mathbb{R}^d \ :\ |k\cdot\omega| \ge \frac{\gamma}{|k|_1^\tau}, \forall k \in \mathbb{Z}^d\backslash\{0\}\right\},$$

where $|k|_1 = \sum_{i=1}^d |k_i|$, and $\omega \in \mathcal{D}^d_{\gamma,\tau}$.

Classical sharp estimates were obtained in [27,28], and a refinement on the constants in [14] for the discrete version of Eq. (2.20). An adaptation of this last result into to the continuous case is the following lemma.

Lemma 2.14 (*Rüssmann Estimates*). *Let $\omega \in \mathcal{D}^d_{\gamma,\tau}$ for some $\gamma > 0$ and $\tau \ge d-1$. Then, for any $v \in \mathcal{A}(\mathbb{T}^d_\rho)$, with $\rho > 0$, there exists a unique zero-average solution of $\mathfrak{L}_\omega u = v - \langle v\rangle$, denoted by $u = \mathfrak{R}_\omega v$, such that $u \in \mathcal{A}(\mathbb{T}^d_{\rho-\delta})$ for any $\delta \in]0,\rho]$. Moreover,*

$$\|\mathfrak{R}_\omega v\|_{\rho-\delta} \le \frac{c_{\mathfrak{R}}(\delta)}{\gamma\delta^\tau}\,\|v\|_\rho,$$

where $c_{\mathfrak{R}}(\delta)$ *depends on* ω *and* δ*, and it is bounded from above by a constant* $\hat{c}_{\mathfrak{R}}$ *than depends only on* d *and* τ*. More concretely, for any* $m > 0$*,*

$$\begin{aligned}
c_{\mathfrak{R}}(\delta) &:= \sqrt{\frac{\gamma^2\delta^{2\tau}2^d}{(2\pi)^2}\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\frac{e^{-4\pi|k|_1\delta}}{|k\cdot\omega|^2}} \\
&\leq \sqrt{\frac{\gamma^2\delta^{2\tau}2^d}{(2\pi)^2}\sum_{0<|k|_1\leq m}\frac{e^{-4\pi|k|_1\delta}}{|k\cdot\omega|^2}+\frac{2^{d+1}\zeta(2,2^\tau)}{(2\pi)^{2\tau+2}}\int_{4\pi\delta(m+1)}^{\infty}u^{2\tau}e^{-u}du} \\
&=: c_{\mathfrak{R}}(\delta,m) \\
&\leq \sqrt{2^{d+1}\zeta(2,2^\tau)(2\pi)^{-2\tau-2}\Gamma(2\tau+1)} \quad =: \hat{c}_{\mathfrak{R}},
\end{aligned}$$

where $\zeta(a,b)=\sum_{j\geq 0}(b+j)^{-a}$ *is the Hurwitz zeta function.*

Along the proof of the main theorem of this paper, we encounter situations in which we have to combine Cauchy and Rüssmann estimates. The following lemma gives sharp width reductions to perform such combined bounds.

Corollary 2.15 (*Rüssmann–Cauchy Estimates*)**.** *Let* $\omega \in \mathcal{D}^d_{\gamma,\tau}$ *for some* $\gamma > 0$ *and* $\tau \geq d-1$*. Then, for any* $v, w \in \mathcal{A}(\mathbb{T}^d_\rho)$*, with* $\rho > 0$*, for any* $\ell = 1,\dots,d$*,* $\delta \in]0,\rho]$ *and* $m > 0$*:*

$$\left\|\frac{\partial}{\partial\theta_\ell}(v\,\mathfrak{R}_\omega w)\right\|_{\rho-\delta} \leq \frac{c^1_{\mathfrak{R}}(\delta)}{\gamma\delta^{\tau+1}}\,\|v\|_\rho\,\|w\|_\rho\,,$$

where

$$\begin{aligned}
c^1_{\mathfrak{R}}(\delta) &:= \frac{(\tau+1)^{\tau+1}}{\tau^\tau}\,c_{\mathfrak{R}}\left(\frac{\tau}{\tau+1}\delta\right) \\
&\leq \frac{(\tau+1)^{\tau+1}}{\tau^\tau}\,c_{\mathfrak{R}}\left(\frac{\tau}{\tau+1}\delta, m\right) \quad =: c^1_{\mathfrak{R}}(\delta,m) \\
&\leq \frac{(\tau+1)^{\tau+1}}{\tau^\tau}\,\hat{c}_{\mathfrak{R}} \quad =: \hat{c}^1_{\mathfrak{R}}.
\end{aligned} \tag{2.22}$$

Proof. The estimates for the derivatives of the solution u follow from applying Cauchy and Rüssmann estimates with width reductions $\delta-\hat{\delta}$ and $\hat{\delta}$, respectively, and choosing $\hat{\delta}$ to maximize $(\delta-\hat{\delta})\hat{\delta}^\tau$. This happens for $\hat{\delta}=\frac{\tau}{1+\tau}\delta$. ☐

Remark 2.16. If one applies Rüssmann and Cauchy estimates with width reductions $\delta/2$ to get the upper bound (2.22), then one obtains $c^1_{\mathfrak{R}}(\delta) = 2^{\tau+1}\hat{c}_{\mathfrak{R}}$. Notice the factor in (2.22) is

$$\frac{(\tau+1)^{\tau+1}}{\tau^\tau} \leq 2^{\tau+1}.$$

Remark 2.17. In applications, for a given $\delta \in]0,\rho]$ one could select m big enough so that the integral term in $c_{\mathfrak{R}}(\delta,m)$ or $c^1_{\mathfrak{R}}(\delta,m)$ is small compared with the preceding sum of terms up to order m. Also, we refer to [29] for a numerical quantification of these estimates and for an analysis of the different sources of overestimation.

Again, the above definitions, constructs and estimates extend naturally to matrix-valued maps.

2.8. The KAM theorem

In this subsection, we present an a-posteriori KAM theorem for d-dimensional quasi-periodic invariant tori in Hamiltonian systems with n degrees-of-freedom that have $n-d$ additional first integrals in involution. The hypotheses in Theorem 2.18 are tailored to be verified with a finite amount of computations.

Theorem 2.18 (*KAM Theorem with First Integrals*)**.** *Let* $(\boldsymbol{\omega},\boldsymbol{g},\boldsymbol{J})$ *be a compatible triple on the open set* $U \subset \mathbb{R}^{2n}$*, where the symplectic form is exact:* $\boldsymbol{\omega} = d\boldsymbol{\alpha}$*. Let* $h : U \to \mathbb{R}$ *be a Hamiltonian for which there exists a moment map* $p : U \to \mathbb{R}^{n-d}$ *whose components and* h *are pairwise in involution functionally independent functions, being* $\Phi : D \subset \mathbb{R}^{n-d}\times U \to U$ *the corresponding moment flow. Let* $\omega \in \mathcal{D}^d_{\gamma,\tau}$ *be a Diophantine vector, for some constants* $\gamma > 0$ *and* $\tau \geq d-1$*. Let* $K : \mathbb{T}^d \to U$ *be a parameterization. We assume that the following hypotheses hold.*

H_1 *The geometric objects* $\boldsymbol{\omega},\boldsymbol{g},\boldsymbol{J},\boldsymbol{\alpha}$*, the Hamiltonian* h*, and the moment map* p *can be analytically extended to an open complex set* $\mathcal{U} \subset \mathbb{C}^{2n}$ *covering* U*, and the moment flow to an open complex set* $\mathcal{D} \subset \mathbb{C}^{n-d}\times\mathcal{U}$ *covering* D*. Moreover, there exist constants* c_Ω*,* c_G*,* c_J*,* $c_{J^\top}$*,* $c_{\mathrm{D}\Omega}$*,* $c_{\mathrm{D}G}$*,* c_{DJ}*,* $c_{DJ^\top}$*,* c_{X_h}*,* $c_{\mathrm{D}X_h}$*,* $c_{(\mathrm{D}X_h)^\top}$*,* $c_{\mathrm{D}^2X_h}$*,* c_{T_h}*,* c_{DT_h}*,* c_{X_p}*,* $c_{\mathrm{D}X_p}$*,* $c_{X_p^\top}$*,* $c_{\mathrm{D}X_p^\top}$*, and* $c_{D\Phi}$ *such that:*

- *the matrix representations* $\Omega, G, J : \mathcal{U} \to \mathbb{C}^{2n\times 2n}$ *of* $\boldsymbol{\omega}$*,* $\boldsymbol{g}$*,* $\boldsymbol{J}$ *satisfy:*

$$\begin{aligned}
&\|\Omega\|_{\mathcal{U}} \leq c_\Omega, \quad \|G\|_{\mathcal{U}} \leq c_G, \quad \|J\|_{\mathcal{U}} \leq c_J, \quad \left\|J^\top\right\|_{\mathcal{U}} \leq c_{J^\top}, \\
&\|D\Omega\|_{\mathcal{U}} \leq c_{\mathrm{D}\Omega}, \quad \|DG\|_{\mathcal{U}} \leq c_{\mathrm{D}G}, \quad \|DJ\|_{\mathcal{U}} \leq c_{DJ}, \quad \left\|DJ^\top\right\|_{\mathcal{U}} \leq c_{DJ^\top};
\end{aligned}$$

- *the Hamiltonian vector field* $X_h : \mathcal{U} \to \mathbb{C}^{2n}$ *and its torsion* $T_h : \mathcal{U} \to \mathbb{C}^{2n\times 2n}$*, satisfy:*

$$\begin{aligned}
&\|X_h\|_{\mathcal{U}} \leq c_{X_h}, \quad \|DX_h\|_{\mathcal{U}} \leq c_{\mathrm{D}X_h}, \quad \left\|(DX_h)^\top\right\|_{\mathcal{U}} \leq c_{(\mathrm{D}X_h)^\top} \\
&\left\|D^2X_h\right\|_{\mathcal{U}} \leq c_{\mathrm{D}^2X_h} \\
&\|T_h\|_{\mathcal{U}} \leq c_{T_h}, \quad \|DT_h\|_{\mathcal{U}} \leq c_{DT_h};
\end{aligned}$$

- *the moment vector fields* $X_p : \mathcal{U} \to \mathbb{C}^{2n\times(n-d)}$ *and the moment flow* $\Phi : \mathcal{D} \to \mathcal{U}$*, satisfy:*

$$\begin{aligned}
&\left\|X_p\right\|_{\mathcal{U}} \leq c_{X_p}, \quad \left\|DX_p\right\|_{\mathcal{U}} \leq c_{\mathrm{D}X_p}, \\
&\left\|X_p^\top\right\|_{\mathcal{U}} \leq c_{X_p^\top}, \quad \left\|DX_p^\top\right\|_{\mathcal{U}} \leq c_{\mathrm{D}X_p^\top} \quad \|D_z\Phi\|_{\mathcal{D}} \leq c_{D\Phi}.
\end{aligned}$$

H_2 *There are* $r > 0$*, an open subset* $\mathcal{U}_0 \subset \mathcal{U}$*, and condition numbers* $\sigma_{\mathrm{D}K}$*,* $\sigma_{(\mathrm{D}K)^\top}$*,* σ_B*,* σ_N*,* $\sigma_{N^\top}$*, and* $\sigma_{\langle T\rangle^{-1}}$ *such that:*

- $K \in (\mathcal{A}_{C^1}(\mathbb{T}^d_\rho))^{2n}$*, with* $0 < \rho < r$*, is an embedding with* $K(\bar{\mathbb{T}}^d_\rho) \subset \mathcal{U}_0$*, whose averaged torsion* $\langle T\rangle$ *is invertible and, moreover:*

$$\|DK\|_\rho < \sigma_{\mathrm{D}K}, \quad \left\|(DK)^\top\right\|_\rho < \sigma_{(\mathrm{D}K)^\top}, \quad \|B\|_\rho < \sigma_B,$$

$$\|N\|_\rho < \sigma_N, \quad \left\|N^\top\right\|_\rho < \sigma_{N^\top}, \quad |\langle T\rangle^{-1}| < \sigma_{\langle T\rangle^{-1}};$$

  We define $\sigma_L = \sigma_{\mathrm{D}K} + c_{X_p}$ *and* $\sigma_{L^\top} = \max\{\sigma_{(\mathrm{D}K)^\top}, c_{X_p^\top}\}$*, so that*

$$\|L\|_\rho < \sigma_L, \quad \left\|L^\top\right\|_\rho < \sigma_{L^\top},$$

- $\mathcal{D}_0 := \{s \in \mathbb{C}^{n-d} \mid |s| < r\}\times\mathcal{U}_0 \subset \mathcal{D}$*.*

Under the above hypotheses, for each $\rho_\infty \in]0,\rho[$ *and* $\delta \in]0,(\rho-\rho_\infty)/3[$*, there exists a constant* $\mathfrak{C}$ *depending on* ρ,ρ_∞,δ *and the constants and objects introduced above such that, if the error of invariance*

$$E = X_h\circ K + \mathfrak{L}_\omega K, \tag{2.23}$$

satisfies[1]

$$\frac{\mathfrak{C}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| < 1, \tag{2.24}$$

where

$$\eta^L = -N^\top\Omega\circ K\,E, \quad \eta^N = \ L^\top\Omega\circ K\,E,$$

then there exists an invariant torus $\mathcal{K}_\infty = K_\infty(\mathbb{T}^d)$ *with frequency* ω*, satisfying* $K_\infty \in \mathcal{A}(\mathbb{T}^d_{\rho_\infty})^{2n}$ *and*

$$\operatorname{dist}(K_\infty(\mathbb{T}^d_{\rho_\infty}), \partial\mathcal{U}_0) > 0, \tag{2.25}$$

$$\|DK_\infty\|_{\rho_\infty} < \sigma_{\mathrm{D}K}, \tag{2.26}$$

---

[1] We recall that $|a,b| := \max\{|a|,|b|\}$..

$$\left\|(DK)_\infty^\top\right\|_{\rho_\infty} < \sigma_{(\mathrm{D}K)^\top}, \tag{2.27}$$

$$\|B_\infty\|_{\rho_\infty} < \sigma_B, \tag{2.28}$$

$$\|N_\infty\|_{\rho_\infty} < \sigma_N, \tag{2.29}$$

$$\left\|N_\infty^\top\right\|_{\rho_\infty} < \sigma_{N^\top}, \tag{2.30}$$

$$|\langle T_\infty\rangle^{-1}| < \sigma_{\langle T\rangle^{-1}}. \tag{2.31}$$

Furthermore, the objects are close to the original ones: there exist constants $\mathfrak{C}_{\Delta K}$, $\mathfrak{C}_{\Delta L}$, $\mathfrak{C}_{\Delta L^\top}$, $\mathfrak{C}_{\Delta B}$, $\mathfrak{C}_{\Delta N}$, $\mathfrak{C}_{\Delta N^\top}$ *and* $\mathfrak{C}_{\Delta\langle T\rangle^{-1}}$ *(like* $\mathfrak{C}$*, given explicitly throughout the proof and summarized in Appendix* B*) such that*

$$\|K_\infty - K\|_{\rho_\infty} \le \frac{\mathfrak{C}_{\Delta K}}{\gamma\delta^\tau}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \tag{2.32}$$

$$\|DK_\infty - DK\|_{\rho_\infty} \le \frac{\mathfrak{C}_{\Delta \mathrm{D}K}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \tag{2.33}$$

$$\left\|(DK)_\infty^\top - (DK)^\top\right\|_{\rho_\infty} \le \frac{\mathfrak{C}_{\Delta (\mathrm{D}K)^\top}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \tag{2.34}$$

$$\|B_\infty - B\|_{\rho_\infty} \le \frac{\mathfrak{C}_{\Delta B}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \tag{2.35}$$

$$\|N_\infty - N\|_{\rho_\infty} \le \frac{\mathfrak{C}_{\Delta N}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \tag{2.36}$$

$$\left\|N_\infty^\top - N^\top\right\|_{\rho_\infty} \le \frac{\mathfrak{C}_{\Delta N^\top}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \tag{2.37}$$

$$|\langle T_\infty\rangle^{-1} - \langle T\rangle^{-1}| \le \frac{\mathfrak{C}_{\Delta\langle T\rangle^{-1}}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \tag{2.38}$$

Remark 2.19. Notice that if $\|L\|_\rho < \sigma_L$, $\left\|L^\top\right\|_\rho < \sigma_{L^\top}$ and $\|B\|_\rho < \sigma_B$, then $\|N\|_\rho < c_J\sigma_L\sigma_B$ and $\left\|N^\top\right\|_\rho < \sigma_B\sigma_{L^\top}c_{J^\top}$ and, hence, if one takes $\sigma_N \ge c_J\sigma_L\sigma_B$ and $\sigma_{N^\top} \ge \sigma_B\sigma_{L^\top}c_{J^\top}$ the conditions for $\|N\|_\rho$ and $\left\|N^\top\right\|_\rho$ follow immediately. Our point is to provide maximum flexibility of the results to be applied to specific problems. Similar controls could be also do for other objects, such as $X_p{\circ}K$, $(X_p{\circ}K)^\top$, G_L or T, leading to similar formulae.

Remark 2.20. If $d = n$ then there are no additional first integrals and we recover the classical KAM theorem for Lagrangian tori. The corresponding estimates follow by taking zero the constants $c_{X_p} = 0$, $c_{X_p^\top} = 0$, $c_{\mathrm{D}X_p} = 0$, $c_{\mathrm{D}X_p^\top} = 0$, $c_{\mathrm{D}\Phi} = 1$.

Remark 2.21. In the *canonical case* we have $\Omega = \Omega_0$, $G = I_{2n}$, and $J = \Omega_0$ and, hence, $c_\Omega = 1$, $c_{\mathrm{D}\Omega} = 0$, $c_G = 1$, $c_{\mathrm{D}G} = 0$, $c_J = 1$, $c_{\mathrm{D}J} = 0$, $c_{J^\top} = 1$, and $c_{\mathrm{D}J^\top} = 0$.

Remark 2.22. Theorem 2.18 produces a d-dimensional isotropic invariant torus with frequency $\omega \in \mathcal{D}^d_{\gamma,\tau}$, that generates an $(n-d)$-parameters family of d-dimensional isotropic invariant tori with such a frequency, foliating an n-dimensional invariant cylinder. With the aid of discounted Hamiltonians, one can produce also n-dimensional invariant cylinders, see Lemma 2.5 or, if the moment map p induces a Hamiltonian torus action, one can produce n-dimensional invariant tori, see Remark 2.8. These tori have frequencies $(\omega, \omega_p) \in \mathcal{D}^d_{\gamma,\tau} \times \mathbb{R}^{n-d}$, thus one obtains analytic families of Lagrangian invariant tori.

Remark 2.23. The invariant d-dimensional tori are locally unique, meaning that if there is another d-dimensional invariant torus with the same frequency nearby, then both generate the same invariant cylinder. More specifically, the corresponding parameterizations K and K', say, are related by

$$K'(\theta) = \Phi_\beta(K(\theta+\alpha)),$$

for suitable $\alpha \in \mathbb{R}^d$, $\beta \in \mathbb{R}^{n-d}$ small. As mentioned in Remark 2.6, both indeterminacies (the phase α and the displacement β) could be fixed by adding n extra scalar equations to the invariance equation.

Remark 2.24. Theorem 2.18 gives the convergence to a parameterization of an invariant torus defined in a complex strip of size ρ_∞ from a parameterization of an approximately invariant torus defined in a complex strip of size ρ, through a sequence of approximations (given by a Newton-like method) whose complex strips sizes are determined by the initial width reduction 3δ (in the proof, the width reductions are given by a geometric sequence). In practical situations, these are parameters that can be adjusted appropriately. Heuristically, see Remark 3.7, a good choice is $\delta = (\rho - \rho_\infty)/6$. Also, if one is not interested in controlling the domain of analyticity of the invariant torus, can take $\rho_\infty = 0$. Since the conditions on the initial parameterization are given by strict inequalities, and the final constants depend continuously on all constants in the hypothesis (including the sizes), then it follows that for a small enough final strip size the conditions hold.

3. Proof of the KAM theorem

In this section we present a fully detailed proof of Theorem 2.18. Hence, from now on we assume the setting and hypotheses of Theorem 2.18. The proof consists in demonstrating the convergence of the (modified) quasi-Newton method outlined in Section 2.6. In Section 3.1 we present some estimates regarding the control of some geometric and dynamical properties for an approximately invariant torus. In Section 3.2 we produce quantitative estimates for the objects obtained when performing one iteration of the procedure. Finally, in Section 3.3 we discuss the convergence of the (modified) quasi-Newton method.

3.1. Some lemmas to control approximate geometric properties

Here we present some estimates regarding the control of some geometric and dynamical properties for an approximately invariant torus, including approximate symplecticity of the corresponding frame, the control of the total error of invariance by its tangent and normal projections, and the approximate reducibility of the linearized dynamics. We collect all constants appearing in the bounds in Appendix B, Table 1.

3.1.1. Approximate symplecticity of the adapted frame

We prove here that the adapted frame $P : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{2n\times 2n}$ attached to the torus $\mathcal{K}$ parameterized by $K : \bar{\mathbb{T}}^d_\rho \to \mathcal{U}_0$, defined in (2.8), induces an approximately symplectic vector bundle isomorphism and, in particular, that the bundle $\mathcal{L}$ framed by $L : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{2n\times n}$ given in (2.7) is approximately Lagrangian. See e.g. [10,23] for similar considerations. An extra ingredient is that, following [18], the errors in the symplecticity of P and Lagragianity of L are controlled by the normal component of the invariance error, η^N.

The symplectic form on the bundle $\mathcal{L}$, is represented by the anti-symmetric matrix-valued map $\Omega_L : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{n\times n}$, which is

$$\Omega_L = \begin{pmatrix} \Omega_{\mathrm{D}K} & (\mathrm{D}(p{\circ}K))^\top \\ -\mathrm{D}(p{\circ}K) & O_{n-d} \end{pmatrix}, \tag{3.1}$$

where we use the pairwise involution of the first integrals,

$$(X_p{\circ}K)^\top\,\Omega{\circ}K\;X_p{\circ}K = O_{n-d},$$

and the corresponding Hamiltonian vector fields to get

$$(X_p{\circ}K)^\top\,\Omega{\circ}K\;\mathrm{D}K = -(\mathrm{D}p){\circ}K\;\mathrm{D}K = -\mathrm{D}(p{\circ}K).$$

Lemma 3.1. *Let* $\Omega_L : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{n\times n}$ *be the matrix-valued map given by* (3.1).

Then, $\langle\Omega_L\rangle = O_n$ *and, in* $\mathbb{T}^d_\rho$,

$$\Omega_L = \begin{pmatrix} \mathfrak{R}_\omega(D\eta^N_{DK} - (D\eta^N_{DK})^\top) & -\mathfrak{R}_\omega(D\eta^N_{X_p})^\top \\ \mathfrak{R}_\omega D\eta^N_{X_p} & O_{n-d} \end{pmatrix}.$$

Moreover, for any $\delta \in]0, \rho]$*:*

$$\|\Omega_L\|_{\rho-\delta} \leq \frac{C^N_{\Omega_L}}{\gamma\delta^{\tau+1}} \left\|\eta^N\right\|_\rho \tag{3.2}$$

and

$$\|\Omega_N\|_{\rho-\delta} \leq \frac{C^N_{\Omega_N}}{\gamma\delta^{\tau+1}} \left\|\eta^N\right\|_\rho . \tag{3.3}$$

Proof. The fact that $\langle \Omega_{\mathrm{DK}} \rangle = O_d$ follows directly from the exact symplectic structure, since $K^*\boldsymbol{\omega} = \mathrm{d}(K^*\boldsymbol{\alpha})$. The fact that $\langle (\mathrm{D}(p \circ K)) \rangle = O_{(n-d)\times d}$ is straightforward. Hence $\langle \Omega_L \rangle = O_n$ follows.

We claim that

$$\mathfrak{L}_\omega \Omega_L = \begin{pmatrix} \mathrm{D}\eta^N_{\mathrm{DK}} - (\mathrm{D}\eta^N_{\mathrm{DK}})^\top & -(\mathrm{D}\eta^N_{X_p})^\top \\ \mathrm{D}\eta^N_{X_p} & O_{n-d} \end{pmatrix}. \tag{3.4}$$

To do so, we first compute the action of $\mathfrak{L}_\omega$ on Ω_{DK} and, using (2.2) and (2.23) we get

$$\mathfrak{L}_\omega \Omega_{\mathrm{DK}} = (\mathrm{D}E)^\top \Omega \circ K\, \mathrm{D}K + (\mathrm{D}K)^\top (\mathrm{D}\Omega) \circ K[E]\, \mathrm{D}K + (\mathrm{D}K)^\top \Omega \circ K\, \mathrm{D}E. \tag{3.5}$$

See [17]. Inspired by [18], we obtain formula

$$\mathfrak{L}_\omega \Omega_{\mathrm{DK}} = \mathrm{D}\eta^N_{\mathrm{DK}} - (\mathrm{D}\eta^N_{\mathrm{DK}})^\top \tag{3.6}$$

from differentiating the projected error

$$\eta^N_{\mathrm{DK}} = (\mathrm{D}K)^\top \Omega \circ K\ E,$$

and symplecticity of $\boldsymbol{\omega}$. To do so using the matrix components, first notice that (3.5) reads

$$\begin{aligned}
\left(\mathfrak{L}_\omega \Omega_{\mathrm{DK}}\right)_{i,j} = {} & \sum_{r,s} \left(\frac{\partial E_r}{\partial \theta_i} \Omega_{r,s} \circ K \frac{\partial K_s}{\partial \theta_j} + \frac{\partial K_r}{\partial \theta_i} \Omega_{r,s} \circ K \frac{\partial E_s}{\partial \theta_j} \right) \\
& + \sum_{r,s,t} \frac{\partial K_r}{\partial \theta_i} \frac{\partial \Omega_{r,s}}{\partial z_t} \circ K E_t \frac{\partial K_s}{\partial \theta_j}
\end{aligned}$$

and also

$$\begin{aligned}
\left(\mathrm{D}\eta^N_{\mathrm{DK}}\right)_{i,j} = {} & \sum_{r,s} \left(\frac{\partial^2 K_r}{\partial \theta_i \partial \theta_j} \Omega_{r,s} \circ K E_s + \frac{\partial K_r}{\partial \theta_i} \Omega_{r,s} \circ K \frac{\partial E_s}{\partial \theta_j} \right) \\
& + \sum_{r,s,t} \frac{\partial K_r}{\partial \theta_i} \frac{\partial \Omega_{r,s}}{\partial z_t} \circ K \frac{\partial K_t}{\partial \theta_j} E_s,
\end{aligned}$$

for $i, j = 1, \dots, d$, where the indices r, s, t run in $1, \dots, 2n$. Hence,

$$\begin{aligned}
\left(\mathrm{D}\eta^N_{\mathrm{DK}}\right)_{i,j} - \left(\mathrm{D}\eta^N_{\mathrm{DK}}\right)_{j,i} = {} & \sum_{r,s} \left(\frac{\partial K_r}{\partial \theta_i} \Omega_{r,s} \circ K \frac{\partial E_s}{\partial \theta_j} - \frac{\partial K_r}{\partial \theta_j} \Omega_{r,s} \circ K \frac{\partial E_s}{\partial \theta_i} \right) \\
& + \sum_{r,s,t} \left(\frac{\partial K_r}{\partial \theta_i} \frac{\partial \Omega_{r,s}}{\partial z_t} \circ K \frac{\partial K_t}{\partial \theta_j} E_s - \frac{\partial K_r}{\partial \theta_j} \frac{\partial \Omega_{r,s}}{\partial z_t} \circ K \frac{\partial K_t}{\partial \theta_i} E_s \right) \\
= {} & \sum_{r,s} \left(\frac{\partial K_r}{\partial \theta_i} \Omega_{r,s} \circ K \frac{\partial E_s}{\partial \theta_j} + \frac{\partial K_s}{\partial \theta_j} \Omega_{r,s} \circ K \frac{\partial E_r}{\partial \theta_i} \right) \\
& + \sum_{r,s,t} \left(- \frac{\partial K_t}{\partial \theta_i} \frac{\partial \Omega_{s,t}}{\partial z_r} \circ K \frac{\partial K_r}{\partial \theta_j} E_s - \frac{\partial K_r}{\partial \theta_j} \frac{\partial \Omega_{r,s}}{\partial z_t} \circ K \frac{\partial K_t}{\partial \theta_i} E_s \right) \\
= {} & \sum_{r,s} \left(\frac{\partial K_r}{\partial \theta_i} \Omega_{r,s} \circ K \frac{\partial E_s}{\partial \theta_j} + \frac{\partial K_s}{\partial \theta_j} \Omega_{r,s} \circ K \frac{\partial E_r}{\partial \theta_i} \right) \\
& + \sum_{r,s,t} \frac{\partial K_t}{\partial \theta_i} \frac{\partial \Omega_{t,r}}{\partial z_s} \circ K \frac{\partial K_r}{\partial \theta_j} E_s \\
= {} & \left(\mathfrak{L}_\omega \Omega_{\mathrm{DK}}\right)_{i,j},
\end{aligned}$$

from where we obtain (3.6).

Notice also that, since

$$\eta^N_{X_p} = (X_p \circ K)^\top \Omega \circ K\ E = -(\mathrm{D}p) \circ K\ (X_h + \mathfrak{L}_\omega K) = -\mathfrak{L}_\omega (p \circ K),$$

then

$$\mathfrak{L}_\omega((X_p \circ K)^\top\ \Omega \circ K\ \mathrm{D}K) = -\mathfrak{L}_\omega \mathrm{D}(p \circ K) = \mathrm{D}\eta^N_{X_p},$$

thus completing the proof of formula (3.4).

Finally, the quantitative estimate (3.2) follows from Rüssmann and Cauchy estimates from Corollary 2.15 applied to the components of (3.4). In particular,

$$\|\Omega_{\mathrm{DK}}\|_{\rho-\delta} \leq \frac{2(d-1)c^1_{\mathfrak{R}}(\delta)}{\gamma\delta^{\tau+1}} \left\|\eta^N\right\|_\rho . \quad \square$$

With the previous lemma we control the approximate symplecticity of the frame P.

Lemma 3.2. *The matrix-valued map* $P : \mathbb{T}^d_\rho \to \mathbb{C}^{2n\times 2n}$*, defined in* (2.8)*, is approximately symplectic, i.e., the symplecticity error map*

$$E_{sym} := P^\top \Omega \circ K\ P - \Omega_0 , \qquad \Omega_0 = \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix},$$

is small in the sense that, for any $\delta \in]0, \rho]$*:*

$$\left\|E_{sym}\right\|_{\rho-\delta} \leq \frac{C_{\mathrm{sym}}}{\gamma\delta^{\tau+1}} \left\|\eta^N\right\|_\rho . \tag{3.7}$$

Proof. To characterize the error in the symplectic character of the frame, we compute

$$E_{\mathrm{sym}} = \begin{pmatrix} L^\top \Omega \circ K\ L & L^\top \Omega \circ K\ N + I_n \\ N^\top \Omega \circ K\ L - I_n & N^\top \Omega \circ K\ N \end{pmatrix} = \begin{pmatrix} \Omega_L & O_n \\ O_n & \Omega_N \end{pmatrix},$$

from which the result follows immediately. $\square$

3.1.2. Relations between the invariance error and their tangent and normal components

From the definitions of η^L and η^N, we obtain easily their bounds controlled by E:

$$\left\|\eta^L\right\|_\rho \leq \sigma_{N^\top} c_\Omega \|E\|_\rho ,$$

$$\left\|\eta^N\right\|_\rho \leq \sigma_{L^\top} c_\Omega \|E\|_\rho .$$

In order to control E in terms of η^L and η^N we have to assume the invertibility of the frame P, which is a consequence of the approximate symplecticity, that is controlled in a narrower strip. We obtain the following lemma.

Lemma 3.3. *Assume that*

$$\nu := \frac{C_{\mathrm{sym}}}{\delta} \left| \left\|\eta^L\right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right| < 1. \tag{3.8}$$

Then, for any $\delta \in]0, \rho]$*:*

$$\|E\|_{\rho-\delta} \leq C^L_E \left\|\eta^L\right\|_\rho + C^N_E \left\|\eta^N\right\|_\rho \leq C_E \left| \left\|\eta^L\right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.9}$$

$$\left\|E^\top\right\|_{\rho-\delta} \leq C^L_{E^\top} \left\|\eta^L\right\|_\rho + C^N_{E^\top} \left\|\eta^N\right\|_\rho \leq C_{E^\top} \left| \left\|\eta^L\right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.10}$$

and

$$\|DE\|_{\rho-\delta} \leq \frac{C^L_{\mathrm{DE}}}{\delta} \left\|\eta^L\right\|_\rho + \frac{C^N_{\mathrm{DE}}}{\delta} \left\|\eta^N\right\|_\rho \leq \frac{C_{\mathrm{DE}}}{\delta} \left| \left\|\eta^L\right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.11}$$

$$\left\|(DE)^\top\right\|_{\rho-\delta} \leq \frac{C^L_{(\mathrm{DE})^\top}}{\delta} \left\|\eta^L\right\|_\rho + \frac{C^N_{(\mathrm{DE})^\top}}{\delta} \left\|\eta^N\right\|_\rho \leq \frac{C_{(\mathrm{DE})^\top}}{\delta} \left| \left\|\eta^L\right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.12}$$

Proof. The hypothesis implies that $\|\Omega_L\|_{\rho-\delta} \leq \nu < 1$ and $\|\Omega_N\|_{\rho-\delta} \leq \nu < 1$, so the matrices $(I_n + \Omega_N \Omega_L)$ and $(I_n + \Omega_L \Omega_N)$ are invertible and

$$\left\|(I_n + \Omega_N \Omega_L)^{-1}\right\|_{\rho-\delta} \leq \frac{1}{1-\nu^2}, \quad \left\|(I_n + \Omega_L \Omega_N)^{-1}\right\|_{\rho-\delta} \leq \frac{1}{1-\nu^2}.$$

Then,

$$I_{2n} + \Omega_0^{-1} E_{\mathrm{sym}} = \begin{pmatrix} I_n & \Omega_N \\ -\Omega_L & I_n \end{pmatrix}$$

is invertible, and

$$(I_{2n} + \Omega_0^{-1} E_{\mathrm{sym}})^{-1} = \begin{pmatrix} I_n & -\Omega_N \\ \Omega_L & I_n \end{pmatrix} \begin{pmatrix} (I_n + \Omega_N \Omega_L)^{-1} & O_n \\ O_n & (I_n + \Omega_L \Omega_N)^{-1} \end{pmatrix}.$$

Notice that, since $I_{2n} + \Omega_0^{-1} E_{\mathrm{sym}} = \Omega_0^{-1} P^\top \Omega\circ K\ P$, then both P and $P^\top$ are invertible in $\bar{\mathbb{T}}_{\rho-\delta}$.

From the definition (2.16) of η we obtain

$$E = -P\,(I_{2n} + \Omega_0^{-1} E_{\mathrm{sym}})^{-1}\,\eta, \quad E^\top = -\eta^\top\,(I_{2n} + E_{\mathrm{sym}} \Omega_0^{-1})^{-1}\ P^\top,$$

from where we could obtain easily bounds for $\|E\|_{\rho-\delta}$ and $\|E^\top\|_{\rho-\delta}$, but it is better to keep track the dependences with respect to $\|\eta^L\|_\rho$ and $\|\eta^N\|_\rho$ separately. Since

$$E = -(L + N\Omega_L)(I_n + \Omega_N \Omega_L)^{-1} \eta^L - (N - L\Omega_N)(I_n + \Omega_L \Omega_N)^{-1} \eta^N$$

and

$$E^\top = -(\eta^L)^\top (I_n + \Omega_L \Omega_N)^{-1}(L^\top - \Omega_L N^\top) - (\eta^N)^\top (I_n + \Omega_N \Omega_L)^{-1}(N^\top + \Omega_N L^\top),$$

the bounds (3.9) and (3.10) follow.

We want to avoid using extra Cauchy estimates for $\mathrm{D}E$ and $(\mathrm{D}E)^\top$, and then lose more analyticity strip. To do so, first, since

$$\mathrm{D}\eta = -\mathrm{D}(\Omega_0^{-1}\ P^\top\ \Omega\circ K\,)E - \Omega_0^{-1}\ P^\top\ \Omega\circ K\ \mathrm{D}E$$

then

$$\mathrm{D}E = -P\,(I_{2n} + \Omega_0^{-1} E_{\mathrm{sym}})^{-1}\left(\mathrm{D}\eta + \mathrm{D}(\Omega_0^{-1}\ P^\top\ \Omega\circ K\,)E\right)$$

and

$$(\mathrm{D}E)^\top = -\left((\mathrm{D}\eta)^\top + (\mathrm{D}(\Omega_0^{-1}\ P^\top\ \Omega\circ K\,)E)^\top\right)(I_{2n} + E_{\mathrm{sym}} \Omega_0^{-1})^{-1}\ P^\top \quad ,$$

from where the bounds (3.11) and (3.12) follow. □

3.1.3. *Control of the action of the Lie operator*

Here we control the action of the operator $\mathfrak{L}_\omega$ on K, L, $L^\top$, G_L, B, and N, avoiding the dependence of the estimates on ω.

Lemma 3.4. *Assume the condition* (3.8) *in* Lemma 3.3. *Then, for any* $\delta \in]0, \rho]$*:*

$$\|\mathfrak{L}_\omega K\|_{\rho-\delta} \le C_{\mathfrak{L}K}, \tag{3.13}$$

$$\|\mathfrak{L}_\omega L\|_{\rho-\delta} \le C_{\mathfrak{L}L}, \tag{3.14}$$

$$\left\|\mathfrak{L}_\omega L^\top\right\|_{\rho-\delta} \le C_{\mathfrak{L}L^\top}, \tag{3.15}$$

$$\|\mathfrak{L}_\omega G_L\|_{\rho-\delta} \le C_{\mathfrak{L}G_L}, \tag{3.16}$$

$$\|\mathfrak{L}_\omega B\|_{\rho-\delta} \le C_{\mathfrak{L}B}, \tag{3.17}$$

$$\|\mathfrak{L}_\omega N\|_{\rho-\delta} \le C_{\mathfrak{L}N}. \tag{3.18}$$

Proof. By using $\varepsilon = \left|\|\eta^L\|_\rho, \frac{1}{\gamma\delta^\tau}\|\eta^N\|_\rho\right|$, $\varepsilon' = \delta^{-1}\varepsilon$, bounds of Lemma 3.3 summarize as $\|E\|_{\rho-\delta} \le C_E\ \varepsilon$, $\|E^\top\|_{\rho-\delta} \le C_{E^\top}\varepsilon$, $\|\mathrm{D}E\|_{\rho-\delta} \le C_{\mathrm{D}E}\ \varepsilon'$ and $\|(\mathrm{D}E)^\top\|_{\rho-\delta} \le C_{(\mathrm{D}E)^\top}\varepsilon'$.

Then, estimate (3.13) follows from the identity $\mathfrak{L}_\omega K = E - X_h\circ K$, and the bound (3.9).

Now, we consider the objects $\mathfrak{L}_\omega L$ and $\mathfrak{L}_\omega L^\top$, given by

$$\begin{aligned}\mathfrak{L}_\omega L &= \begin{pmatrix}\mathfrak{L}_\omega \mathrm{D}K & \mathfrak{L}_\omega(X_p\circ K)\end{pmatrix} = \begin{pmatrix}\mathrm{D}E - (\mathrm{D}X_h)\circ K\ \mathrm{D}K & (\mathrm{D}X_p)\circ K\ [\mathfrak{L}_\omega K]\end{pmatrix}\\ &= \begin{pmatrix}\mathrm{D}E - (\mathrm{D}X_h)\circ K\ \mathrm{D}K & (\mathrm{D}X_p)\circ K\ [E - X_h\circ K]\end{pmatrix}\\ &= \begin{pmatrix}\mathrm{D}E & (\mathrm{D}X_p)\circ K\ [E]\end{pmatrix} - (\mathrm{D}X_h)\circ K\ L\end{aligned}$$

and

$$\mathfrak{L}_\omega L^\top = \begin{pmatrix}(\mathrm{D}E)^\top \\ ((\mathrm{D}X_p)\circ K\ [E])^\top\end{pmatrix} - ((\mathrm{D}X_h)\circ K\ L)^\top$$

from where (3.14) and (3.15) follow.

Bound (3.16) follows from

$$\mathfrak{L}_\omega G_L = \mathfrak{L}_\omega L^\top G\circ K\ L + L^\top(\mathrm{D}G)\circ K[\mathfrak{L}_\omega K]\ L + L^\top G\circ K\ \mathfrak{L}_\omega L.$$

Then, from the identity $G_L\ B = I_n$, we obtain

$$\mathfrak{L}_\omega B = -B\ \mathfrak{L}_\omega G_L\ B,$$

from where we get the estimate (3.17). Notice that G_L and B are symmetric, so we obtain the same bounds for their transposes.

Finally, we consider the object $\mathfrak{L}_\omega N$, given by

$$\mathfrak{L}_\omega N = (\mathrm{D}J)\circ K[\mathfrak{L}_\omega K]\ L\ B + J\circ K\ \mathfrak{L}_\omega L\ B + J\circ K\ L\ \mathfrak{L}_\omega B,$$

and then we get (3.18). □

3.1.4. *Approximate reducibility*

A crucial step in the proof Theorem 2.18 is solving the linearized equation arising from the application of the Newton method. This is based on the (approximate) reduction of such linear system into a block triangular form. This is the content of the following lemma.

Lemma 3.5. *Assume the condition* (3.8) *in* Lemma 3.3. *Then, the linearized dynamics* $\mathrm{D}X_h\circ K : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{2n\times 2n}$ *is approximately reducible via the frame* $P : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{2n\times 2n}$ *defined in* (2.8)*, to the block-triangular matrix-valued map* $\Lambda : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{2n\times 2n}$ *defined in* (2.10)*, where* $T : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{n\times n}$*, defined in* (2.9)*, is the torsion, for which*

$$\|T\|_\rho \le C_T. \tag{3.19}$$

Specifically, the reducibility error map $E_{red} : \bar{\mathbb{T}}^d_\rho \to \mathbb{C}^{2n\times 2n}$ *defined as*

$$E_{red} := \Omega_0^{-1} P^\top \Omega\circ K\ \mathcal{X}_P - \Lambda = \begin{pmatrix} E^{LL}_{red} & E^{LN}_{red} \\ E^{NL}_{red} & E^{NN}_{red}\end{pmatrix} \tag{3.20}$$

satisfies, for any $\delta \in]0, \rho]$*:*

$$\|E^{LL}_{red}\|_{\rho-\delta} \le \frac{C^L_{E^{LL}_{\mathrm{red}}}}{\delta}\left\|\eta^L\right\|_\rho + \frac{C^N_{E^{LL}_{\mathrm{red}}}}{\delta}\left\|\eta^N\right\|_\rho, \tag{3.21}$$

$$\|E^{NL}_{red}\|_{\rho-\delta} \le \frac{C^L_{E^{NL}_{\mathrm{red}}}}{\delta}\left\|\eta^L\right\|_\rho + \frac{C^N_{E^{NL}_{\mathrm{red}}}}{\delta}\left\|\eta^N\right\|_\rho, \tag{3.22}$$

$$\|E^{NN}_{red}\|_{\rho-\delta} \le \frac{C^L_{E^{NN}_{\mathrm{red}}}}{\delta}\left\|\eta^L\right\|_\rho + \frac{C^N_{E^{NN}_{\mathrm{red}}}}{\delta}\left\|\eta^N\right\|_\rho \le \frac{C_{E^{NN}_{\mathrm{red}}}}{\delta}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \tag{3.23}$$

and

$$\|E^{LN}_{red}\|_{\rho-\delta} \le \frac{C^L_{E^{LN}_{\mathrm{red}}}}{\delta}\left\|\eta^L\right\|_\rho + \frac{C^N_{E^{LN}_{\mathrm{red}}}}{\gamma\delta^{\tau+1}}\left\|\eta^N\right\|_\rho \le \frac{C_{E^{LN}_{\mathrm{red}}}}{\delta}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \tag{3.24}$$

Proof. Using the notation in (2.6) we write the block components of (3.20):

$$E_{\mathrm{red}} = \begin{pmatrix} E^{LL}_{\mathrm{red}} & E^{LN}_{\mathrm{red}} \\ E^{NL}_{\mathrm{red}} & E^{NN}_{\mathrm{red}}\end{pmatrix} = \begin{pmatrix} N^\top\ \Omega\circ K\ \mathcal{X}_L & N^\top\ \Omega\circ K\ \mathcal{X}_N - T \\ -L^\top\ \Omega\circ K\ \mathcal{X}_L & -L^\top\ \Omega\circ K\ \mathcal{X}_N\end{pmatrix}$$

First, since

$$E^{LL}_{\mathrm{red}} = \begin{pmatrix} N^\top \Omega\circ K\ \mathrm{D}E & N^\top \Omega\circ K\ \mathrm{D}X_p\circ K[E]\end{pmatrix}$$

and

$$N^\top \Omega\circ K\ \mathrm{D}E = -\mathrm{D}(N^\top \Omega\circ K)E - \mathrm{D}\eta^L,$$

then we obtain (3.21). Analogously,

$$E^{NL}_{\mathrm{red}} = \begin{pmatrix} -L^\top \Omega\circ K\ \mathrm{D}E & -L^\top \Omega\circ K\ \mathrm{D}X_p\circ K[E]\end{pmatrix}$$

and

$$-L^\top \Omega\circ K\ \mathrm{D}E = \mathrm{D}(L^\top \Omega\circ K)E - \mathrm{D}\eta^N,$$

from which we obtain (3.22).

In order to bound E_{red}^{NN}, we apply the left operator $\mathfrak{L}_\omega$ to the identity $L^\top \Omega{\circ}K\, N = -I_n$, and obtain

$$O_n = \mathfrak{L}_\omega L^\top\, \Omega{\circ}K\, N + L^\top(\mathrm{D}\Omega){\circ}K\,[E - X_h{\circ}K] + L^\top \Omega{\circ}K\, \mathfrak{L}_\omega N.$$

Then, using the geometric property (2.2), we obtain

$$E_{\mathrm{red}}^{NN} = L^\top(\mathrm{D}\Omega{\circ}K\,[E])N + \mathcal{X}_L^\top \Omega{\circ}K\, N = L^\top(\mathrm{D}\Omega{\circ}K\,[E])N - (E_{\mathrm{red}}^{LL})^\top,$$

where

$$(E_{\mathrm{red}}^{LL})^\top = \begin{pmatrix} (\mathrm{D}(\Omega{\circ}K\, N)E)^\top - (\mathrm{D}\eta^L)^\top \\ -(\mathrm{D}X_p{\circ}K[E])^\top \Omega{\circ}K\, N \end{pmatrix},$$

from which we obtain (3.23).

Finally,

$$\begin{aligned} E_{\mathrm{red}}^{LN} &= B\, L^\top\, G{\circ}K\, (\mathrm{D}J){\circ}K[E]\, L\, B + B\, L^\top\, \Omega{\circ}K\, \mathcal{X}_L\, B + B\, \Omega_L\, \mathfrak{L}_\omega B \\ &= B\, L^\top\, G{\circ}K\, (\mathrm{D}J){\circ}K[E]\, L\, B - B\, E_{\mathrm{red}}^{NL}\, B + B\, \Omega_L\, \mathfrak{L}_\omega B. \end{aligned}$$

from which we obtain (3.24). ☐

3.2. One step of the iterative procedure

Here we apply one correction of the (modified) quasi-Newton method described in Section 2.6 and we obtain quantitative estimates for the new approximately invariant torus and related objects. We set sufficient conditions to preserve the control of the previous estimates. The constants that appear along the proof are collected in Appendix B, Tables 2 and 3.

Lemma 3.6 (*The Iterative Lemma*). *For any $\delta \in]0, \rho/3]$, there exist constants C_{sym}, C_{ξ^L}, $C_{\Delta\bar{\bar{K}}}$, $C_{\Delta\bar{\bar{L}}}$, $C_{\Delta\bar{\bar{L}}^\top}$, $C_{\Delta\bar{\bar{B}}}$, $C_{\Delta\bar{\bar{N}}}$, $C_{\Delta\bar{\bar{N}}^\top}$, $C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}$ and $Q_{\bar{\bar{\eta}}^L}$, $Q_{\bar{\bar{\eta}}^N}$, such that if*

$$\frac{\hat{C}_\Delta}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right| < 1, \tag{3.25}$$

where

$$\begin{aligned} \hat{C}_\Delta := \max \Bigg\{ & \gamma\delta^\tau C_{\mathrm{sym}}, C_{\xi^L}, \frac{\delta\, C_{\Delta\bar{\bar{K}}}}{\mathrm{dist}(K(\mathbb{T}_\rho^d), \partial\mathcal{U}_0)}, \\ & \frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\sigma_{\mathrm{D}K} - \|DK\|_\rho}, \frac{C_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}}{\sigma_{(\mathrm{D}K)^\top} - \|(DK)^\top\|_\rho}, \\ & \frac{C_{\Delta\bar{\bar{B}}}}{\sigma_B - \|B\|_\rho}, \frac{C_{\Delta\bar{\bar{N}}}}{\sigma_N - \|N\|_\rho}, \frac{C_{\Delta\bar{\bar{N}}^\top}}{\sigma_{N^\top} - \|N^\top\|_\rho}, \frac{C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}}{\sigma_{\langle T\rangle^{-1}} - |\langle T\rangle^{-1}|} \Bigg\}, \end{aligned} \tag{3.26}$$

then we have a new parameterization $\bar{\bar{K}} \in (\mathcal{A}(\mathbb{T}_{\rho-2\delta}^d))^{2n}$, that defines new objects $\bar{\bar{L}}$, $\bar{\bar{B}}$, $\bar{\bar{N}}$ and $\bar{\bar{T}}$ (obtained replacing K by $\bar{\bar{K}}$ in the corresponding definitions) satisfying

$$\mathrm{dist}(\bar{\bar{K}}(\mathbb{T}_{\rho-2\delta}^d), \partial\mathcal{U}_0) > 0, \tag{3.27}$$

$$\left\|D\bar{\bar{K}}\right\|_{\rho-3\delta} < \sigma_{\mathrm{D}K}, \tag{3.28}$$

$$\left\|(D\bar{\bar{K}})^\top\right\|_{\rho-3\delta} < \sigma_{(\mathrm{D}K)^\top}, \tag{3.29}$$

$$\left\|\bar{\bar{B}}\right\|_{\rho-3\delta} < \sigma_B, \tag{3.30}$$

$$\left\|\bar{\bar{N}}\right\|_{\rho-3\delta} < \sigma_N, \tag{3.31}$$

$$\left\|\bar{\bar{N}}^\top\right\|_{\rho-3\delta} < \sigma_{N^\top}, \tag{3.32}$$

$$|\langle\bar{\bar{T}}\rangle^{-1}| < \sigma_{\langle T\rangle^{-1}}, \tag{3.33}$$

and

$$\left\|\bar{\bar{K}} - K\right\|_{\rho-2\delta} \le \frac{C_{\Delta\bar{\bar{K}}}}{\gamma\delta^\tau} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.34}$$

$$\left\|D\bar{\bar{K}} - DK\right\|_{\rho-3\delta} \le \frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.35}$$

$$\left\|(D\bar{\bar{K}})^\top - (DK)^\top\right\|_{\rho-3\delta} \le \frac{C_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.36}$$

$$\left\|\bar{\bar{B}} - B\right\|_{\rho-3\delta} \le \frac{C_{\Delta\bar{\bar{B}}}}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.37}$$

$$\left\|\bar{\bar{N}} - N\right\|_{\rho-3\delta} \le \frac{C_{\Delta\bar{\bar{N}}}}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.38}$$

$$\left\|\bar{\bar{N}}^\top - N^\top\right\|_{\rho-3\delta} \le \frac{C_{\Delta\bar{\bar{N}}^\top}}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|, \tag{3.39}$$

$$|\langle\bar{\bar{T}}\rangle^{-1} - \langle T\rangle^{-1}| \le \frac{C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.40}$$

Moreover, the tangent and normal components of the new error of invariance

$$\bar{\bar{E}} = X_h{\circ}\bar{\bar{K}} + \mathfrak{L}_\omega \bar{\bar{K}},$$

satisfy

$$\left\|\bar{\bar{\eta}}^N\right\|_{\rho-3\delta} \le \frac{Q_{\bar{\bar{\eta}}^N}}{\delta} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|^2. \tag{3.41}$$

and

$$\left\|\bar{\bar{\eta}}^L\right\|_{\rho-3\delta} \le \frac{Q_{\bar{\bar{\eta}}^L}}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|^2. \tag{3.42}$$

Proof. We divide the proof into several steps. Starting from the initial parameterization, K, we first consider an intermediate parameterization,

$$\bar{K} = K + N\xi^N \tag{3.43}$$

and then compute the new parameterization as

$$\bar{\bar{K}} = \Phi_{\xi^L_{X_p}} \circ \bar{K} \circ (\mathrm{id} + \xi_{\mathrm{D}K}^L).$$

In the following, we will invoke Lemmas 3.3–3.5, whose condition (3.8) is included into the hypothesis (3.25) (this corresponds to the first term in (3.26)).

Step 1: Control of the intermediate parameterization. We recall that, in (3.43),

$$\xi^N = \hat{\xi}_0^N + \mathfrak{R}_\omega \eta^N.$$

where

$$\hat{\xi}_0^N = \langle T\rangle^{-1} \langle \eta^L - T\mathfrak{R}_\omega \eta^N \rangle.$$

Hence, from Rüssmann estimates in Lemma 2.14 with width reduction ρ to $\mathfrak{R}_\omega \eta^N$ we obtain

$$|\hat{\xi}_0^N| \le C_{\hat{\xi}_0^N}^L \left\|\eta^L\right\|_\rho + \frac{C_{\hat{\xi}_0^N}^N}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \le C_{\hat{\xi}_0^N} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.44}$$

and with width reduction δ we obtain

$$\left\|\xi^N\right\|_{\rho-\delta} \le C_{\xi^N}^L \left\|\eta^L\right\|_\rho + \frac{C_{\xi^N}^N}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \le C_{\xi^N} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.45}$$

Then

$$\bar{K} - K = N\xi^N$$

and

$$\|\bar{K} - K\|_{\rho-\delta} \le C_{\Delta\bar{K}}^L \left\|\eta^L\right\|_\rho + \frac{C_{\Delta\bar{K}}^N}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \le C_{\Delta\bar{K}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right| \tag{3.46}$$

The intermediate torus should be included in the domain $\mathcal{U}_0$, more specifically $\bar{K}(\mathbb{T}_{\rho-\delta}^d) \subset \mathcal{U}_0$. To verify that, notice that

$$\mathrm{dist}(\bar{K}(\mathbb{T}_{\rho-\delta}^d), \partial\mathcal{U}_0) \ge \mathrm{dist}(K(\mathbb{T}_\rho^d), \partial\mathcal{U}_0) - \|\bar{K} - K\|_{\rho-\delta}$$

$$\geq \operatorname{dist}(K(\mathbb{T}^d_\rho), \partial\mathcal{U}_0) - C_{\Delta\bar{K}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right| > 0,$$

where the last inequality follows if

$$\frac{C_{\Delta\bar{K}}}{\operatorname{dist}(K(\mathbb{T}^d_\rho), \partial\mathcal{U}_0)} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right| < 1. \tag{3.47}$$

As we will see, this condition is implied by the third term in (3.26).

Using Rüssmann and Cauchy estimates, see Corollary 2.15, on

$$\bar{K} - K = N\hat{\xi}_0^N + N\mathfrak{R}_\omega\eta^N,$$

we get

$$\left\|\mathrm{D}\bar{K} - \mathrm{D}K\right\|_{\rho-\delta} \leq \frac{C^L_{\Delta\mathrm{D}\bar{K}}}{\delta} \left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\mathrm{D}\bar{K}}}{\gamma\delta^{\tau+1}} \left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\mathrm{D}\bar{K}}}{\delta} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.48}$$

and, analogously,

$$\begin{aligned} \left\|(\mathrm{D}\bar{K})^\top - (\mathrm{D}K)^\top\right\|_{\rho-\delta} &\leq \frac{C^L_{\Delta(\mathrm{D}\bar{K})^\top}}{\delta} \left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta(\mathrm{D}\bar{K})^\top}}{\gamma\delta^{\tau+1}} \left\|\eta^N\right\|_\rho \\ &\leq \frac{C_{\Delta(\mathrm{D}\bar{K})^\top}}{\delta} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \end{aligned} \tag{3.49}$$

Notice that, in particular,

$$\left\|\mathrm{D}\bar{K}\right\|_{\rho-\delta} < \sigma_{\mathrm{D}K}, \quad \left\|(\mathrm{D}\bar{K})^\top\right\|_{\rho-\delta} < \sigma_{(\mathrm{D}K)^\top} \tag{3.50}$$

provided that

$$\frac{C_{\Delta\mathrm{D}\bar{K}}}{\delta(\sigma_{\mathrm{D}K} - \|\mathrm{D}K\|_\rho)} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right| < 1 \tag{3.51}$$

and

$$\frac{C_{\Delta(\mathrm{D}\bar{K})^\top}}{\delta(\sigma_{(\mathrm{D}K)^\top} - \|(\mathrm{D}K)^\top\|_\rho)} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right| < 1, \tag{3.52}$$

conditions that, as we will see, are implied by the fourth and fifth terms in (3.26).

Let $\bar{L}$ be the L object associated to $\bar{K}$. Notice that, from (3.51),

$$\left\|\bar{L}\right\|_{\rho-\delta} < \sigma_L, \quad \left\|\bar{L}^\top\right\|_{\rho-\delta} < \sigma_{L^\top}.$$

Then, since

$$\bar{L} - L = \left(\mathrm{D}\bar{K} - \mathrm{D}K \quad \int_0^1 (\mathrm{D}X_p)\circ(K + \lambda N\xi^N)\, d\lambda\, [\bar{K} - K]\right),$$

we obtain

$$\left\|\bar{L} - L\right\|_{\rho-\delta} \leq \frac{C^L_{\Delta\bar{L}}}{\delta} \left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{L}}}{\gamma\delta^{\tau+1}} \left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{L}}}{\delta} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.53}$$

and

$$\left\|\bar{L}^\top - L^\top\right\|_{\rho-\delta} \leq \frac{C^L_{\Delta\bar{L}^\top}}{\delta} \left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{L}^\top}}{\gamma\delta^{\tau+1}} \left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{L}^\top}}{\delta} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.54}$$

Step 2: Computation of the intermediate invariance errors. The error of invariance in the intermediate step is

$$\bar{E} = X_h\circ\bar{K} + \mathfrak{L}_\omega\bar{K}$$

which can be written as

$$\begin{aligned} \bar{E} &= E + \Delta^1 X_h + \mathfrak{L}_\omega N\, \xi^N + N\, \mathfrak{L}_\omega\xi^N \\ &= E + \Delta^1 X_h + \mathfrak{L}_\omega N\, \xi^N + N\, \eta^N \end{aligned}$$

where

$$\Delta^1 X_h = \int_0^1 (\mathrm{D}X_h)\circ(K + \lambda N\xi^N)\, d\lambda\, [N\xi^N],$$

from where we obtain the bound

$$\left\|\bar{E}\right\|_{\rho-\delta} \leq C^L_{\bar{E}} \left\|\eta^L\right\|_\rho + \frac{C^N_{\bar{E}}}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \leq C_{\bar{E}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.55}$$

Also

$$\begin{aligned} \bar{E} &= E + (\mathrm{D}X_h)\circ K N\xi^N + \Delta^2 X_h + \mathfrak{L}_\omega N\, \xi^N + N\, \mathfrak{L}_\omega\xi^N \\ &= E + \mathcal{X}_N\, \xi^N + \Delta^2 X_h + N\, \mathfrak{L}_\omega\xi^N, \end{aligned} \tag{3.56}$$

where

$$\Delta^2 X_h = \int_0^1 (1-\lambda)(\mathrm{D}^2 X_h)\circ(K + \lambda(\bar{K} - K))\, d\lambda\, [\bar{K} - K, \bar{K} - K],$$

from where we obtain that the *intermediate normal error* is

$$\begin{aligned} \bar{\eta}^N &= \bar{L}^\top\, \Omega\circ\bar{K}\, \bar{E} \\ &= (\bar{L}^\top - L^\top)\, \Omega\circ\bar{K}\, \bar{E} + L^\top\, (\Omega\circ\bar{K} - \Omega\circ K)\, \bar{E} \\ &\quad + L^\top \Omega\circ K(E + \mathcal{X}_N\, \xi^N + \Delta^2 X_h + N\, \mathfrak{L}_\omega\xi^N) \\ &= (\bar{L}^\top - L^\top)\, \Omega\circ\bar{K}\, \bar{E} + L^\top\, (\Omega\circ\bar{K} - \Omega\circ K)\, \bar{E} + \eta^N \\ &\quad - E^{NN}_{\mathrm{red}}\xi^N + L^\top \Omega\circ K\Delta^2 X_h - \mathfrak{L}_\omega\xi^N \\ &= (\bar{L}^\top - L^\top)\, \Omega\circ\bar{K}\, \bar{E} + L^\top\, (\Omega\circ\bar{K} - \Omega\circ K)\, \bar{E} - E^{NN}_{\mathrm{red}}\xi^N + L^\top \Omega\circ K\, \Delta^2 X_h, \end{aligned} \tag{3.57}$$

where we emphasize that $\mathfrak{L}_\omega\xi^N = \eta^N$ in $\mathbb{T}^d_\rho$. Notice that the intermediate normal error is quadratically small. Quantitatively,

$$\left\|\bar{\eta}^N\right\|_{\rho-\delta} \leq \frac{Q_{\bar{\eta}^N}}{\delta} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|^2. \tag{3.58}$$

Step 3: Control of the new parameterization. The new approximation is

$$\bar{\bar{K}} = \Phi_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}), \tag{3.59}$$

where $\xi^L = (\xi^L_{\mathrm{D}K}, \xi^L_{X_p})$ is given by

$$\xi^L = \mathfrak{R}_\omega(\eta^L - T\xi^N).$$

Using Rüssmann estimates we obtain

$$\left\|\xi^L\right\|_{\rho-2\delta} \leq \frac{C^L_{\xi^L}}{\gamma\delta^\tau} \left\|\eta^L\right\|_\rho + \frac{C^N_{\xi^L}}{\gamma^2\delta^{2\tau}} \left\|\eta^N\right\|_\rho \leq \frac{C_{\xi^L}}{\gamma\delta^\tau} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \tag{3.60}$$

Notice that $\bar{K}(\mathbb{T}^d_{\rho-\delta}) \subset \mathcal{U}_0$, and the components of ξ^L are in $\mathcal{A}(\mathbb{T}^d_{\rho-2\delta})$. Hence, the computation of $\bar{\bar{K}}$ in (3.59) could be done if $\|\xi^L_{\mathrm{D}K}\|_{\rho-2\delta} < \delta$ and $\left\|\xi^L_{X_p}\right\|_{\rho-2\delta} < r$, where we recall r is the width of complex 'time-domain" of Φ_s. Since $\rho < r$, these are conditions that are implied by the hypothesis

$$\frac{C_{\xi^L}}{\gamma\delta^{\tau+1}} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right| < 1, \tag{3.61}$$

which corresponds to the second term in (3.26).

Then,

$$\begin{aligned} \bar{\bar{K}} - K &= \Phi_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) - \bar{K} + \bar{K} - K \\ &= \int_0^1 \frac{\partial}{\partial\lambda} \left(\Phi_{\lambda\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \lambda\xi^L_{\mathrm{D}K}) \right) \mathrm{d}\lambda\ + \bar{K} - K \\ &= \int_0^1 \mathrm{D}_z\Phi_{\lambda\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \lambda\xi^L_{\mathrm{D}K})\, \bar{L}\circ(\mathrm{id} + \lambda\xi^L_{\mathrm{D}K})\, \mathrm{d}\lambda\, \xi^L\ + \bar{K} - K \end{aligned} \tag{3.62}$$

from where

$$\begin{aligned} \left\|\bar{\bar{K}} - K\right\|_{\rho-2\delta} &\leq c_{\mathrm{D}\Phi}\sigma_L \left\|\xi^L\right\|_{\rho-2\delta} + \left\|\bar{K} - K\right\|_{\rho-\delta} + \\ &\leq \frac{C^L_{\Delta\bar{\bar{K}}}}{\gamma\delta^\tau} \left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{\bar{K}}}}{\gamma^2\delta^{2\tau}} \left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{\bar{K}}}}{\gamma\delta^\tau} \left| \left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau} \left\|\eta^N\right\|_\rho \right|. \end{aligned} \tag{3.63}$$

The new approximation should remain in $\mathcal{U}_0$. In particular, we observe that

$$\begin{aligned}\operatorname{dist}(\bar{\bar{K}}(\mathbb{T}^d_{\rho-2\delta}), \partial\mathcal{U}_0) &\geq \operatorname{dist}(K(\mathbb{T}^d_\rho), \partial\mathcal{U}_0) - \left\|\bar{\bar{K}} - K\right\|_{\rho-2\delta} \\ &\geq \operatorname{dist}(K(\mathbb{T}^d_\rho), \partial\mathcal{U}_0) - \frac{C_{\Delta\bar{\bar{K}}}}{\gamma\delta^\tau}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| > 0,\end{aligned}$$

where the last inequality follows from hypothesis (3.25) (this corresponds to the third term in (3.26)). We emphasize that this control includes the fact that $\operatorname{dist}(\bar{K}(\mathbb{T}^d_{\rho-\delta}), \partial\mathcal{U}_0) > 0$ (see (3.47)). Moreover, we get (3.27) and (3.34) in Lemma 3.6.

By directly applying Cauchy estimates to the first part of (3.62) and bounds (3.48) and (3.49) we get

$$\begin{aligned}\left\|\mathrm{D}\bar{\bar{K}} - \mathrm{D}K\right\|_{\rho-3\delta} &\leq \frac{C^L_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \\ &\leq \frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|,\end{aligned} \tag{3.64}$$

and

$$\begin{aligned}\left\|(\mathrm{D}\bar{\bar{K}})^\top - (\mathrm{D}K)^\top\right\|_{\rho-3\delta} &\leq \frac{C^L_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \\ &\leq \frac{C_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|,\end{aligned} \tag{3.65}$$

that correspond to (3.35) and (3.36). Then (3.28) and (3.29) follow from hypotheses

$$\frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma\delta^{\tau+1}(\sigma_{\mathrm{D}K} - \|\mathrm{D}K\|_\rho)}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| < 1$$

and

$$\frac{C_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}}{\gamma\delta^{\tau+1}(\sigma_{(\mathrm{D}K)^\top} - \left\|(\mathrm{D}K)^\top\right\|_\rho)}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| < 1,$$

which corresponds to the fourth and fifth terms in (3.26), and imply conditions (3.51) and (3.52).

In the following, we write $\bar{\bar{L}}, \bar{\bar{N}}, \ldots$ for the corresponding $L, N, \ldots$ objects associated to $\bar{\bar{K}}$. In particular, the new tangent frame is

$$\bar{\bar{L}} = \begin{pmatrix}\mathrm{D}\bar{\bar{K}} & X_p\circ\bar{\bar{K}}\end{pmatrix},$$

and

$$\left\|\bar{\bar{L}}\right\|_{\rho-3\delta} < \sigma_L, \quad \left\|\bar{\bar{L}}^\top\right\|_{\rho-3\delta} < \sigma_{L^\top}.$$

Moreover, since

$$\bar{\bar{L}} - L = \begin{pmatrix}\mathrm{D}\bar{\bar{K}} - \mathrm{D}K & \int_0^1(\mathrm{D}X_p)\circ(K + \lambda(\bar{\bar{K}} - K))\, d\lambda\,(\bar{\bar{K}} - K)\end{pmatrix},$$

we get

$$\left\|\bar{\bar{L}} - L\right\|_{\rho-3\delta} \leq \frac{C^L_{\Delta\bar{\bar{L}}}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{\bar{L}}}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{\bar{L}}}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| \tag{3.66}$$

and

$$\left\|\bar{\bar{L}}^\top - L^\top\right\|_{\rho-3\delta} \leq \frac{C^L_{\Delta\bar{\bar{L}}^\top}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{\bar{L}}^\top}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{\bar{L}}^\top}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \tag{3.67}$$

The new restricted metric is

$$G_{\bar{\bar{L}}} = \bar{\bar{L}}^\top G\circ\bar{\bar{K}}\;\bar{\bar{L}},$$

from where

$$G_{\bar{\bar{L}}} - G_L = (\bar{\bar{L}}^\top - L^\top)\,G\circ\bar{\bar{K}}\;\bar{\bar{L}} + L^\top\,(G\circ\bar{\bar{K}} - G\circ K)\;\bar{\bar{L}} + L^\top\,G\circ K\,(\bar{\bar{L}} - L)$$

and, then,

$$\|G_{\bar{\bar{L}}} - G_L\|_{\rho-3\delta} \leq \frac{C^L_{\Delta G_{\bar{\bar{L}}}}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta G_{\bar{\bar{L}}}}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta G_{\bar{\bar{L}}}}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \tag{3.68}$$

We know that G_L is invertible (in $\bar{\mathbb{T}}^d_\rho$), and $B = G_L^{-1}$ with $\|B\|_\rho < \sigma_B$. We introduce now the constants

$$C^L_{\Delta\bar{\bar{B}}} := (\sigma_B)^2 C^L_{\Delta G_{\bar{\bar{L}}}}, \quad C^N_{\Delta\bar{\bar{B}}} := (\sigma_B)^2 C^N_{\Delta G_{\bar{\bar{L}}}}, \quad C_{\Delta\bar{\bar{B}}} := (\sigma_B)^2 C_{\Delta G_{\bar{\bar{L}}}}.$$

Then, since

$$\frac{C_{\Delta\bar{\bar{B}}}}{\gamma\delta^{\tau+1}(\sigma_B - \|B\|_\rho)}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| < 1,$$

that corresponds to the sixth term in (3.26), from Lemma A.1 we obtain that $G_{\bar{\bar{L}}}$ is invertible (in $\bar{\mathbb{T}}^d_{\rho-3\delta}$) and

$$\left\|\bar{\bar{B}} - B\right\|_{\rho-3\delta} \leq (\sigma_B)^2\left\|G_{\bar{\bar{L}}} - G_L\right\|_{\rho-3\delta}, \quad \left\|\bar{\bar{B}}\right\|_{\rho-3\delta} < \sigma_B,$$

from where we obtain that

$$\left\|\bar{\bar{B}} - B\right\|_{\rho-3\delta} \leq \frac{C^L_{\Delta\bar{\bar{B}}}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{\bar{B}}}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{\bar{B}}}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \tag{3.69}$$

We obtain estimates (3.30) and (3.37) in Lemma 3.6.

The new normal frame is

$$\bar{\bar{N}} = J\circ\bar{\bar{K}}\,\bar{\bar{L}}\,\bar{\bar{B}}.$$

Since

$$\bar{\bar{N}} - N = (J\circ\bar{\bar{K}} - J\circ K)\bar{\bar{L}}\,\bar{\bar{B}} + J\circ K(\bar{\bar{L}} - L)\bar{\bar{B}} + J\circ K\,L(\bar{\bar{B}} - B),$$

then

$$\left\|\bar{\bar{N}} - N\right\|_{\rho-3\delta} \leq \frac{C^L_{\Delta\bar{\bar{N}}}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{\bar{N}}}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{\bar{N}}}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| \tag{3.70}$$

and

$$\left\|\bar{\bar{N}}^\top - N^\top\right\|_{\rho-3\delta} \leq \frac{C^L_{\Delta\bar{\bar{N}}^\top}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{\bar{N}}^\top}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{\bar{N}}^\top}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \tag{3.71}$$

Then, (3.31) and (3.32) follow from hypotheses

$$\frac{C_{\Delta\bar{\bar{N}}}}{\gamma\delta^{\tau+1}(\sigma_N - \|N\|_\rho)}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| < 1$$

and

$$\frac{C_{\Delta\bar{\bar{N}}^\top}}{\gamma\delta^{\tau+1}(\sigma_{N^\top} - \left\|N^\top\right\|_\rho)}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| < 1,$$

which corresponds to the seventh and eighth terms in (3.26). Hence, we obtain estimates (3.31), (3.32) and (3.38), (3.39) in Lemma 3.6.

Finally, the new torsion is

$$\bar{\bar{T}} = \bar{\bar{N}}^\top\;T_h\circ\bar{\bar{K}}\;\bar{\bar{N}},$$

and satisfies

$$\left\|\bar{\bar{T}}\right\|_{\rho-3\delta} \leq C_T.$$

Moreover, since

$$\bar{\bar{T}} - T = (\bar{\bar{N}}^\top - N^\top)T_h\circ\bar{\bar{K}}\;\bar{\bar{N}} + N^\top(T_h\circ\bar{\bar{K}} - T_h\circ K)\;\bar{\bar{N}} + N^\top T_h\circ K\,(\bar{\bar{N}} - N),$$

we obtain

$$\left\|\bar{\bar{T}} - T\right\|_{\rho-3\delta} \leq \frac{C^L_{\Delta\bar{\bar{T}}}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\bar{\bar{T}}}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \leq \frac{C_{\Delta\bar{\bar{T}}}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \tag{3.72}$$

We know that $\langle T\rangle$ is invertible and $|\langle T\rangle^{-1}| < \sigma_{\langle T\rangle^{-1}}$. Mimicking the arguments made above for G_L and $B = G_L^{-1}$, we introduce the constants

$$C^L_{\Delta\langle\bar{\bar{T}}\rangle^{-1}} := (\sigma_{\langle T\rangle^{-1}})^2 C^L_{\Delta\bar{\bar{T}}}, \quad C^N_{\Delta\langle\bar{\bar{T}}\rangle^{-1}} := (\sigma_{\langle T\rangle^{-1}})^2 C^N_{\Delta\bar{\bar{T}}}, \quad C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}} := (\sigma_{\langle T\rangle^{-1}})^2 C_{\Delta\bar{\bar{T}}}.$$

Then, since

$$\frac{C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}}{\gamma\delta^{\tau+1}(\sigma_{\langle T\rangle^{-1}} - |\langle T\rangle^{-1}|)} \left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| < 1,$$

which corresponds to the ninth term in (3.26), we obtain that $\langle\bar{\bar{T}}\rangle$ is invertible and

$$|\langle\bar{\bar{T}}\rangle^{-1} - \langle T\rangle^{-1}| \leq (\sigma_{\langle T\rangle^{-1}})^2 \left\|\bar{\bar{T}} - T\right\|_{\rho-3\delta}, \quad |\langle\bar{\bar{T}}\rangle^{-1}|_{\rho-3\delta} < \sigma_{\langle T\rangle^{-1}},$$

from where we obtain that

$$\begin{aligned} |\langle\bar{\bar{T}}\rangle^{-1} - \langle T\rangle^{-1}| &\leq \frac{C^L_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}}{\gamma\delta^{\tau+1}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}}{\gamma^2\delta^{2\tau+1}}\left\|\eta^N\right\|_\rho \\ &\leq \frac{C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \end{aligned} \tag{3.73}$$

We obtain then the estimates (3.33) and (3.40) on the new object.

Step 4: Computation of the new invariance errors. In order to compute the new invariance error $\bar{\bar{E}}$, we first compute

$$\begin{aligned} \mathrm{D}\bar{\bar{K}} &= \mathrm{D}_s\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathrm{D}\xi^L_{X_p} \\ &\quad + \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathrm{D}\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,(I_d + \mathrm{D}\xi^L_{\mathrm{D}K}) \\ &= X_p\circ\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathrm{D}\xi^L_{X_p} \\ &\quad + \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathrm{D}\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,(I_d + \mathrm{D}\xi^L_{\mathrm{D}K}) \\ &= \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\Big(X_p\circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathrm{D}\xi^L_{X_p} \\ &\quad + \mathrm{D}\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,(I_d + \mathrm{D}\xi^L_{\mathrm{D}K})\Big), \end{aligned} \tag{3.74}$$

then

$$\begin{aligned} \mathfrak{L}_\omega\bar{\bar{K}} &= \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,(X_p\circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\mathfrak{L}_\omega\xi^L_{X_p} + \mathfrak{L}_\omega\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})) \\ &\quad + \mathrm{D}\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L_{\mathrm{D}K} \\ &= \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,(\mathfrak{L}_\omega\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L) \end{aligned}$$

and

$$\begin{aligned} X_h\circ\bar{\bar{K}} &= X_h\circ\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) \\ &= \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,X_h\circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}). \end{aligned}$$

As a result,

$$\begin{aligned} \bar{\bar{E}} &= \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\big(\bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L\big) \\ &= \bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L + \\ &\quad \left(\mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) - I_{2n}\right)\big(\bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L\big). \end{aligned} \tag{3.75}$$

In order to get a crude bound of $\bar{\bar{E}}$ from the first line in (3.75), we first bound

$$\begin{aligned} \left\|\mathfrak{L}_\omega\xi^L\right\|_{\rho-\delta} &= \left\|\eta^L - T\xi^N\right\|_{\rho-\delta} \\ &\leq C^L_{\mathfrak{L}\xi^L}\left\|\eta^L\right\|_\rho + \frac{C^N_{\mathfrak{L}\xi^L}}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho \leq C_{\mathfrak{L}\xi^L}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| \end{aligned} \tag{3.76}$$

(notice that $\mathfrak{L}_\omega\xi^L = \eta^L - T\xi^N$ is in fact real-analytic in $\mathbb{T}^d_\rho$), and, then,

$$\begin{aligned} \left\|\bar{\bar{E}}\right\|_{\rho-2\delta} &\leq c_{\mathrm{D}\Phi}(\|\bar{E}\|_{\rho-\delta} + \sigma_L\left\|\mathfrak{L}_\omega\xi^L\right\|_{\rho-\delta}) \\ &\leq C^L_{\bar{\bar{E}}}\left\|\eta^L\right\|_\rho + \frac{C^N_{\bar{\bar{E}}}}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho \leq C_{\bar{\bar{E}}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|. \end{aligned} \tag{3.77}$$

A posteriori, we will see that $\bar{\bar{E}}$ is quadratically small.

To compute the new normal error $\bar{\bar{\eta}}^N$ we first compute the new tangent frame $\bar{\bar{L}}$. To do so, we first obtain

$$\begin{aligned} X_p\circ\bar{\bar{K}} &= X_p\circ\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) \\ &= \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,X_p\circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) \end{aligned}$$

and, using (3.74), we get

$$\bar{\bar{L}} = \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,(I_n + \hat{\mathrm{D}}\xi^L), \tag{3.78}$$

where

$$\hat{\mathrm{D}}\xi^L := \begin{pmatrix}\mathrm{D}\xi^L & O_{n\times n-d}\end{pmatrix}.$$

From (3.75) and (3.78) we get

$$\begin{aligned} \bar{\bar{\eta}}^N &= \bar{\bar{L}}^\top\Omega\circ\bar{\bar{K}}\,\bar{\bar{E}} \\ &= (I_n + (\hat{\mathrm{D}}\xi^L)^\top)\,(\bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}))^\top\left(\mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\right)^\top \\ &\quad \times\Omega\circ\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) \\ &\quad \mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\big(\bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L\big) \\ &= (I_n + (\hat{\mathrm{D}}\xi^L)^\top)\,(\bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}))^\top\,\Omega\circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\big(\bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) \\ &\quad + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L\big) \\ &= (I_n + (\hat{\mathrm{D}}\xi^L)^\top)\,\big(\bar{\eta}^N\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) + \Omega_{\bar{L}}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L\big). \end{aligned}$$

See (3.57). We observe that $\bar{\bar{\eta}}^N$ is quadratically small. Quantitatively, since

$$\left\|(\mathrm{D}\xi^L)^\top\right\|_{\rho-3\delta} \leq \frac{n}{\delta}\left\|\xi^L\right\|_{\rho-2\delta} \leq \frac{nC_{\xi^L}}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right| < n,$$

where we use condition (3.61),

$$\begin{aligned} \left\|\Omega_{\bar{L}}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\right\|_{\rho-3\delta} \leq \|\Omega_{\bar{L}}\|_{\rho-2\delta} &\leq \frac{C^{\mathrm{N}}_{\Omega_L}}{\gamma\delta^{\tau+1}}\left\|\bar{\eta}^N\right\|_{\rho-\delta} \\ &\leq \frac{C^{\mathrm{N}}_{\Omega_L}}{\gamma\delta^{\tau+1}}\frac{Q_{\bar{\eta}^N}}{\delta}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|^2, \end{aligned}$$

where we apply Lemma 3.1 to $\bar{K}$ and bound (3.58), and, finally, applying (3.76) and defining

$$\varkappa := \frac{1}{\gamma\delta^{\tau+1}}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \tag{3.79}$$

we get

$$\left\|\bar{\bar{\eta}}^N\right\|_{\rho-3\delta} \leq \frac{Q_{\bar{\bar{\eta}}^N}}{\delta}\left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|^2, \tag{3.80}$$

which corresponds to the bound (3.41) in Lemma 3.6.

The new tangent error is

$$\begin{aligned} \bar{\bar{\eta}}^L &= -\bar{\bar{N}}^\top\,\Omega\circ\bar{\bar{K}}\,\bar{\bar{E}} \\ &= -N^\top\Omega\circ K\,\bar{\bar{E}} - N^\top\,(\Omega\circ\bar{\bar{K}} - \Omega\circ K)\,\bar{\bar{E}} - (\bar{\bar{N}}^\top - N^\top)\,\Omega\circ\bar{\bar{K}}\,\bar{\bar{E}} \\ &= -N^\top\,\Omega\circ K\,\big(\bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L\big) \\ &\quad - N^\top\,\Omega\circ K\left(\mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) - I_{2n}\right)\big(\bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) \\ &\quad + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L\big) \\ &\quad - N^\top\,(\Omega\circ\bar{\bar{K}} - \Omega\circ K)\,\bar{\bar{E}} - (\bar{\bar{N}}^\top - N^\top)\,\Omega\circ\bar{\bar{K}}\,\bar{\bar{E}} \\ &= -N^\top\,\Omega\circ K E - N^\top\,\Omega\circ K(\bar{E} - E) - N^\top\,\Omega\circ K\,L\,\mathfrak{L}_\omega\xi^L \\ &\quad - N^\top\,\Omega\circ K(\bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) - \bar{E}) \\ &\quad - N^\top\,\Omega\circ K(\bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) - L)\,\mathfrak{L}_\omega\xi^L \\ &\quad - N^\top\,\Omega\circ K\,(\mathrm{D}_z\boldsymbol{\Phi}_{\xi^L_{X_p}} \circ\bar{K}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) - I_{2n})\,(\bar{E}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K}) \\ &\quad + \bar{L}\circ(\mathrm{id} + \xi^L_{\mathrm{D}K})\,\mathfrak{L}_\omega\xi^L) \\ &\quad - (\bar{\bar{N}}^\top - N^\top)\,\Omega\circ\bar{\bar{K}}\,\bar{\bar{E}} - N^\top\,(\Omega\circ\bar{\bar{K}} - \Omega\circ K)\,\bar{\bar{E}} \\ &=: \bar{\bar{\eta}}^L_1 + \bar{\bar{\eta}}^L_2 + \bar{\bar{\eta}}^L_3 + \bar{\bar{\eta}}^L_4 + \bar{\bar{\eta}}^L_5, \end{aligned} \tag{3.81}$$

where the addends are numbered in order. In the previous expression, all addends, but the first, are trivially quadratically small. But, in fact, $\mathfrak{L}_\omega \xi^L$ is selected in such a way that $\bar{\bar{\eta}}_1^L$ *is* quadratically small, since

$$
\begin{aligned}
\bar{\bar{\eta}}_1^L &= \eta^L - N^\top\, \Omega{\circ}K\, (\mathcal{X}_N\, \xi^N + \Delta^2 X_h + N\, \mathfrak{L}_\omega \xi^N) - \mathfrak{L}_\omega \xi^L \\
&= \eta^L - T\xi^N - E_{\mathrm{red}}^{LN}\xi^N - N^\top\, \Omega{\circ}K\, \Delta^2 X_h - \Omega_N\, \mathfrak{L}_\omega \xi^N - \mathfrak{L}_\omega \xi^L \\
&= -E_{\mathrm{red}}^{LN}\xi^N - N^\top\, \Omega{\circ}K\, \Delta^2 X_h - \Omega_N\, \eta^N,
\end{aligned}
$$

where we apply that $N^\top \Omega{\circ}K\, L = I_n$, (3.56), and $\mathfrak{L}_\omega \xi^L = \eta^L - T\xi^N$. From this, we get a bound for the first addend of (3.81):

$$
\left\| \bar{\bar{\eta}}_1^L \right\|_{\rho-3\delta} \leq \frac{Q_{\bar{\bar{\eta}}_1^L}}{\delta} \left| \left\| \eta^L \right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\| \eta^N \right\|_\rho \right|^2 . \tag{3.82}
$$

For the second addend, since

$$
\bar{E}{\circ}(\mathrm{id} + \xi_{\mathrm{D}K}^L) - \bar{E} = \int_0^1 \mathrm{D}\bar{E}{\circ}(\mathrm{id} + \lambda \xi_{\mathrm{D}K}^L) d\lambda\, \xi_{\mathrm{D}K}^L ,
$$

we get

$$
\left\| \bar{\bar{\eta}}_2^L \right\|_{\rho-3\delta} \leq \sigma_{N^\top} c_\Omega \left\| \mathrm{D}\bar{E} \right\|_{\rho-2\delta} \left\| \xi_{\mathrm{D}K}^L \right\|_{\rho-2\delta} \leq \frac{Q_{\bar{\bar{\eta}}_2^L}}{\gamma\delta^{\tau+1}} \left| \left\| \eta^L \right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\| \eta^N \right\|_\rho \right|^2 , \tag{3.83}
$$

applying Cauchy estimates. The third addend is

$$
\bar{\bar{\eta}}_3^L = -N^\top\, \Omega{\circ}K (\bar{L}{\circ}(\mathrm{id} + \xi_{\mathrm{D}K}^L) - \bar{L} + \bar{L} - L)\, \mathfrak{L}_\omega \xi^L ,
$$

from which, proceeding analogously, we obtain the bound

$$
\begin{aligned}
\left\| \bar{\bar{\eta}}_3^L \right\|_{\rho-3\delta} &\leq \sigma_{N^\top} c_\Omega \left(\left\| \mathrm{D}\bar{L} \right\|_{\rho-2\delta} \left\| \xi_{\mathrm{D}K}^L \right\|_{\rho-2\delta} + \left\| \bar{L} - L \right\|_{\rho-2\delta} \right) \left\| \mathfrak{L}_\omega \xi^L \right\|_{\rho-\delta} \\
&\leq \frac{Q_{\bar{\bar{\eta}}_3^L}}{\gamma\delta^{\tau+1}} \left| \left\| \eta^L \right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\| \eta^N \right\|_\rho \right|^2 .
\end{aligned} \tag{3.84}
$$

In order to bound the fourth addend, we first observe that

$$
\mathrm{D}_z \boldsymbol{\Phi}_s - I_{2n} = \int_0^1 \mathrm{D}_z X_p {\circ} \boldsymbol{\Phi}_{\lambda s}\, \mathrm{D}_z \boldsymbol{\Phi}_s\, d\lambda\, s.
$$

Then

$$
\begin{aligned}
\left\| \bar{\bar{\eta}}_4^L \right\|_{\rho-3\delta} &\leq \sigma_{N^\top} c_\Omega c_{\mathrm{D}X_p} c_{\mathrm{D}\boldsymbol{\Phi}} \left\| \xi_{X_p}^L \right\|_{\rho-2\delta} \left(\left\| \bar{E} \right\|_{\rho-\delta} + \sigma_L \left\| \mathfrak{L}_\omega \xi^L \right\|_{\rho-\delta} \right) \\
&\leq \frac{Q_{\bar{\bar{\eta}}_4^L}}{\gamma\delta^{\tau}} \left| \left\| \eta^L \right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\| \eta^N \right\|_\rho \right|^2 .
\end{aligned} \tag{3.85}
$$

For the fifth addend, we first observe that

$$
\Omega{\circ}\boldsymbol{\Phi}_s - \Omega = \int_0^1 \mathrm{D}\Omega{\circ}\boldsymbol{\Phi}_{\lambda s} X_p {\circ} \boldsymbol{\Phi}_{\lambda s}\, d\lambda\, s.
$$

Then,

$$
\left\| \bar{\bar{\eta}}_5^L \right\|_{\rho-3\delta} \leq \frac{Q_{\bar{\bar{\eta}}_5^L}}{\gamma\delta^{\tau+1}} \left| \left\| \eta^L \right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\| \eta^N \right\|_\rho \right|^2 . \tag{3.86}
$$

Hence, collecting (3.82), (3.83), (3.84), (3.85), and (3.86) we get

$$
\left\| \bar{\bar{\eta}}^L \right\|_{\rho-3\delta} \leq \frac{Q_{\bar{\bar{\eta}}^L}}{\gamma\delta^{\tau+1}} \left| \left\| \eta^L \right\|_\rho , \frac{1}{\gamma\delta^\tau} \left\| \eta^N \right\|_\rho \right|^2 , \tag{3.87}
$$

which is bound (3.42). With this we finish the proof of Lemma 3.6. □

3.3. Convergence of the iterative process

Once the quadratic procedure has been established in Section 3.2, proving the convergence of the scheme follows standard arguments, that we will detail for providing explicit conditions for the KAM theorem.

Proof of Theorem 2.18. Let us consider the parameterization $K_0 := K$ with initial invariance error $E_0 := E$, whose tangent and normal projections are η_0^L and η_0^N, respectively. We also introduce $B_0 := B$, $N_0 := N$, $N_0^\top := N^\top$, and $T_0 := T$ associated to the initial parameterization. By applying Lemma 3.6 recursively, at the step j we obtain new objects $K_j := \bar{\bar{K}}_{j-1}$, $E_j := \bar{\bar{E}}_{j-1}$, $\eta_j^L := \bar{\bar{\eta}}_{j-1}^L$, $\eta_j^N := \bar{\bar{\eta}}_{j-1}^N$, $B_j := \bar{\bar{B}}_{j-1}$, $N_j := \bar{\bar{N}}_{j-1}$, $N_j^\top := \bar{\bar{N}}_{j-1}^\top$, and $T_j := \bar{\bar{T}}_{j-1}$.

The domain of analyticity of these objects is reduced at every step, from the initial value $\rho_0 = \rho$ to a limiting value ρ_∞. At the step j, the parameterization K_j and associated objects are defined in a strip of width ρ_j, and have been produced from the parameterization K_{j-1}, which is defined in a strip of width ρ_{j-1}, throughout computations (involving small divisors equations and derivatives) that produce three width reductions of size δ_{j-1} to the width ρ_{j-1}, so then

$$
\rho_j = \rho_{j-1} - 3\delta_{j-1} .
$$

If we select a geometric sequence of width reductions

$$
\delta_j = \frac{\delta_0}{a^j}
$$

with $a > 1$, then, from the identity

$$
\rho_\infty = \rho_0 - 3 \sum_{j=0}^{\infty} \delta_j = \rho_0 - 3\delta_0 \frac{a}{a-1} ,
$$

we get

$$
a = \frac{\rho_0 - \rho_\infty}{\rho_0 - 3\delta_0 - \rho_\infty} = \frac{\rho_0 - \rho_\infty}{\rho_1 - \rho_\infty} . \tag{3.88}
$$

Let us assume that we have successfully applied j times Lemma 3.6 We observe that condition (3.25) is required at every step, but the construction has been performed in such a way that we control $\mathrm{dist}(K_j(\mathbb{T}_{\rho_j}^d), \partial\mathcal{U}_0)$, $\mathrm{D}K_j$, $(\mathrm{D}K_j)^\top$, B_j, N_j, $N_j^\top$, and $\langle T_j \rangle^{-1}$ uniformly with respect to j, so the constants that appear in Section 3.1, displayed in Table 1, and in Lemma 3.6, displayed in Tables 2 and 3, are taken to be the same for all steps by considering the worst value of δ_j, that is, δ_0. We also take the upper bounds $\hat{c}_{\mathfrak{R}}$ and $\hat{c}_{\mathfrak{R}}^1$ for $c_{\mathfrak{R}}(\delta_j)$ and $\hat{c}_{\mathfrak{R}}^1(\delta_j)$.

The first computation is tracking the sequences ε_j and $\varkappa_j$ of weighted errors, defined to be

$$
\varepsilon_j = \left| \left\| \eta_j^L \right\|_{\rho_j} , \frac{1}{\gamma\delta_j^\tau} \left\| \eta_j^N \right\|_{\rho_j} \right| , \qquad \varkappa_j = \frac{1}{\gamma\delta_j^{\tau+1}} \varepsilon_j .
$$

By defining

$$
Q_{\bar{\bar{\eta}}} := \left| Q_{\bar{\bar{\eta}}^L} , a^\tau Q_{\bar{\bar{\eta}}^N} \right| , \tag{3.89}
$$

and writing $Q_{\bar{\bar{\eta}}}(\delta_j, \varkappa_j)$ in order to emphasize its dependence on δ_j and $\varkappa_j$ at the step j, we get

$$
\varepsilon_j \leq \frac{Q_{\bar{\bar{\eta}}}(\delta_{j-1}, \varkappa_{j-1})}{\gamma\delta_{j-1}^{\tau+1}} \varepsilon_{j-1}^2 , \qquad \varkappa_j \leq a^{\tau+1} Q_{\bar{\bar{\eta}}}(\delta_{j-1}, \varkappa_{j-1}) \varkappa_{j-1}^2 . \tag{3.90}
$$

By imposing

$$
\kappa := a^{\tau+1} Q_{\bar{\bar{\eta}}}(\delta_0, \varkappa_0) \varkappa_0 = \frac{Q_{\bar{\bar{\eta}}} a^{\tau+1}}{\gamma\delta_0^{\tau+1}} \varepsilon_0 < 1 , \tag{3.91}
$$

which is included in (2.24), and starting at $\varkappa_0$, $\kappa_0 = \kappa$, we get decreasing sequences

$$
\varkappa_j \leq k_{j-1} \varkappa_{j-1} , \qquad \kappa_j = a^{\tau+1} Q_{\bar{\bar{\eta}}}(\delta_j, \varkappa_j) \varkappa_j \leq \kappa_{j-1}^2
$$

that, moreover, satisfy

$$
\varkappa_j \leq \kappa^{2^j - 1} \varkappa_0 , \qquad \kappa_j \leq \kappa^{2^j} .
$$

We also have

$$
\varepsilon_j \leq \kappa^{2^j - 1} a^{-j(\tau+1)} \varepsilon_0 \leq (\kappa / a^{\tau+1})^j \varepsilon_0 . \tag{3.92}
$$

Now, using expression (3.92), we check Hypothesis (3.25) of the iterative lemma, Lemma 3.6, so that we can perform the step $j+1$. The required sufficient condition is included in the hypothesis (2.24) of the KAM theorem, whose inequality has several terms that correspond to the different components in (3.26).

The first condition, using (3.92), is given by

$$\nu_j := \frac{C_{\mathrm{sym}}}{\delta_j}\varepsilon_j \le \frac{C_{\mathrm{sym}}}{\delta_0}(\kappa/a^\tau)^j\varepsilon_0 \le \frac{C_{\mathrm{sym}}}{\delta_0}\varepsilon_0 = \nu_0 < 1, \tag{3.93}$$

where the last inequality is included in (2.24). Checking the second conditions is analogous, and it is

$$\frac{C_{\xi L}}{\gamma\delta_j^{\tau+1}}\varepsilon_j \le \frac{C_{\xi L}}{\gamma\delta_0^{\tau+1}}\kappa^j\varepsilon_0 \le \frac{C_{\xi L}}{\gamma\delta_0^{\tau+1}}\varepsilon_0 < 1, \tag{3.94}$$

where the last inequality is again included in (2.24).

In order to check the rest of conditions for K_j, $\mathrm{D}K_j$, $(\mathrm{D}K_j)^\top$, B_j, N_j, $N_j^\top$ and $\langle T_j\rangle^{-1}$, we have to relate them to the conditions corresponding to the initial objects K_0, $\mathrm{D}K_0$, $(\mathrm{D}K_0)^\top$, B_0, N_0, $N_0^\top$ and $\langle T_0\rangle^{-1}$.

The third condition in (3.25) is checked as follows. We recursively obtain that

$$\begin{aligned}\mathrm{dist}(K_j(\mathbb{T}^d_{\rho_j}),\partial\mathcal{U}_0) - \frac{C_{\Delta\bar{\bar{K}}}}{\gamma\delta_j^\tau}\varepsilon_j &\ge \mathrm{dist}(K_0(\mathbb{T}^d_{\rho_0}),\partial\mathcal{U}_0) - \sum_{i=0}^{j}\frac{C_{\Delta\bar{\bar{K}}}}{\gamma\delta_i^\tau}\varepsilon_i\\ &\ge \mathrm{dist}(K_0(\mathbb{T}^d_{\rho_0}),\partial\mathcal{U}_0) - \sum_{i=0}^{j}\frac{C_{\Delta\bar{\bar{K}}}}{\gamma\delta_0^\tau}(\kappa/a)^i\varepsilon_0\\ &\ge \mathrm{dist}(K_0(\mathbb{T}^d_{\rho_0}),\partial\mathcal{U}_0) - \frac{C_{\Delta\bar{\bar{K}}}}{\gamma\delta_0^\tau}\frac{a}{a-\kappa}\varepsilon_0 > 0\end{aligned}$$

and the last inequality reads

$$\frac{a}{a-\kappa}\,\frac{C_{\Delta\bar{\bar{K}}}}{\gamma\delta_0^\tau\,\mathrm{dist}(K_0(\mathbb{T}^d_{\rho_0}),\partial\mathcal{U}_0)}\varepsilon_0 < 1 \tag{3.95}$$

which is included as a condition into (2.24)

$$\frac{1}{\gamma\delta_0^{\tau+1}}\left(\delta_0 C_{\mathrm{dist}} + Q_{\bar{\bar{\eta}}}a^\tau\right)\varepsilon_0 < 1.$$

The rest of conditions in (3.25), associated to $\mathrm{D}K_j$, $(\mathrm{D}K_j)^\top$, B_j, N_j, $N_j^\top$ and $\langle T_j\rangle^{-1}$, follow by reproducing similar computations. For instance, from

$$\begin{aligned}\left\|\mathrm{D}K_j\right\|_{\rho_j} + \frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma\delta_j^{\tau+1}}\varepsilon_j &\le \|\mathrm{D}K_0\|_{\rho_0} + \sum_{i=0}^{j}\frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma\delta_i^{\tau+1}}\varepsilon_i\\ &\le \|\mathrm{D}K_0\|_{\rho_0} + \sum_{i=0}^{j}\frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma\delta_0^{\tau+1}}\kappa^i\varepsilon_0\\ &\le \|\mathrm{D}K_0\|_{\rho_0} + \frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\gamma\delta_0^{\tau+1}}\frac{1}{1-\kappa}\varepsilon_0 < \sigma_{\mathrm{D}K},\end{aligned}$$

we include the last inequality as

$$\frac{1}{\gamma\delta_0^{\tau+1}}\left(\frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\sigma_{\mathrm{D}K} - \|\mathrm{D}K_0\|_{\rho_0}} + Q_{\bar{\bar{\eta}}}a^{\tau+1}\right)\varepsilon_0 < 1 \tag{3.96}$$

into (2.24), and the other conditions follow *mutatis mutandis*. Notice that condition (3.96) includes condition (3.92).

Having guaranteed all hypotheses in Lemma 3.6, we collect the inequalities (3.93), (3.94), (3.95), (3.96) for DK and the corresponding inequalities for the other objects, that are included into hypothesis (2.24). This follows by introducing the constant $\mathfrak{C}$ as

$$\mathfrak{C} := \max\left\{\gamma\delta^\tau C_{\mathrm{sym}},\ C_{\xi L},\ \delta C_{\mathrm{dist}} + Q_{\bar{\bar{\eta}}}a^\tau,\ C_\Delta + Q_{\bar{\bar{\eta}}}a^{\tau+1}\right\}. \tag{3.97}$$

with

$$C_{\mathrm{dist}} := \frac{C_{\Delta\bar{\bar{K}}}}{\mathrm{dist}(K(\mathbb{T}^d_\rho),\partial\mathcal{U}_0)}, \tag{3.98}$$

and

$$\begin{aligned}C_\Delta := \max\Bigg\{&\frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\sigma_{\mathrm{D}K} - \|\mathrm{D}K\|_\rho}, \frac{C_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}}{\sigma_{(\mathrm{D}K)^\top} - \left\|(\mathrm{D}K)^\top\right\|_\rho}, \frac{C_{\Delta\bar{\bar{B}}}}{\sigma_B - \|B\|_\rho},\\ &\frac{C_{\Delta\bar{\bar{N}}}}{\sigma_N - \|N\|_\rho}, \frac{C_{\Delta\bar{\bar{N}}^\top}}{\sigma_{N^\top} - \left\|N^\top\right\|_\rho}, \frac{C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}}{\sigma_{\langle T\rangle^{-1}} - |\langle T\rangle^{-1}|}\Bigg\}.\end{aligned} \tag{3.99}$$

Finally, since

$$\frac{\mathfrak{C}}{\gamma\delta^{\tau+1}}\varepsilon < 1,$$

which is (2.24), we can apply the iterative process infinitely many times. Indeed, since $\varepsilon_j \to 0$ when $j \to 0$, the iterative process converges to a true quasi-periodic torus K_∞. Moreover, from the computations above this object satisfies $K_\infty \in \mathcal{A}(\mathbb{T}^d_{\rho_\infty})$ and the controls (2.25), (2.26), (2.27), (2.28), (2.29), (2.30), and (2.31). Furthermore, we also get the distances between the initial and the limiting objects, (2.32), (2.33), (2.34), (2.35), (2.36), (2.37), (2.38). This completes the proof of the KAM theorem. □

Remark 3.7. We can argue from (3.93) that, once one fixes the initial and final strip sizes ρ_0 and ρ_∞, an (almost optimal) choice for the ratio a (or the initial width reduction $\delta = \delta_0$) is the one that minimizes $\frac{a}{\delta_0}$. Since

$$\frac{a}{\delta} = \frac{\rho - \rho_\infty}{\delta(\rho - 3\delta - \rho_\infty)}$$

the best choice is for $\delta = \frac{\rho-\rho_\infty}{6}$, for which $a = 2$ and, hence

$$\delta_j = \frac{\Delta}{2^{j+1}},$$

where $\Delta = (\rho_\infty - \rho_0)/3$. This choice is empirically supported by the computations in [14].

Remark 3.8. The choice of the δ_j above as the geometric series with ratio 1/2 is justified by the following rationale. Let us assume that the constants involved in the theorem do not depend on δ (their dependence on it is very mild) then, (3.90) can be written as

$$\varepsilon_j \le \left(\frac{Q_{\bar{\bar{\eta}}}}{\gamma}\varepsilon_0\right)^{2^j-1}\left(\frac{1}{\delta_{j-1}^{1/2^{j-1}}\,\delta_{j-2}^{1/2^{j-2}}\ldots\delta_0}\right)^{2^{j-1}(\tau+1)}\varepsilon_0.$$

Then, one needs to minimize

$$\prod_{j=0}^{\infty}\delta_j^{-\frac{1}{2^j}}$$

or, equivalently

$$-\sum_{j=0}^{\infty}\frac{1}{2^j}\log\delta_j,$$

under the constraint

$$\sum_{j=0}^{\infty}\delta_j = \Delta, \tag{3.100}$$

where $\delta_j > 0$ for all $j \ge 0$.

Then, if $(\delta_j^0 = 2^{-j-1}\Delta)_{j\ge0}$ is the target geometric sequence obtained in Remark 3.7, for any other positive sequence $(\delta_j)_{j\ge0}$ satisfying (3.100), the function $f : [0,1] \to \mathbb{R}$ defined as

$$f(t) = -\sum_{j=0}^{\infty}\frac{1}{2^j}\log((1-t)\delta_j^0 + t\delta_j)$$

is strictly convex and $f'(0) = 0$.

Table 1
Constants in Section 3.1, Lemmas 3.1–3.5.

Object	Constant	Label
η^L, η^N	$\varepsilon = \left\lvert \lVert \eta^L \rVert_\rho , \frac{1}{\gamma\delta^\tau} \lVert \eta^N \rVert_\rho \right\rvert, \quad \varepsilon' = \frac{1}{\delta}\varepsilon, \quad \varkappa = \frac{1}{\gamma\delta^{\tau+1}}\varepsilon$	(B.1)
Ω_L	$C^{\mathrm{N}}_{\Omega_L} = (d+n-2)c^1_{\mathfrak{R}}(\delta)$	(3.2)
Ω_N	$C^{\mathrm{N}}_{\Omega_{\mathrm{N}}} = (\sigma_B)^2 C^{\mathrm{N}}_{\Omega_L}$	(3.3)
E_{sym}	$C_{\mathrm{sym}} = \max\{C^{\mathrm{N}}_{\Omega_L}, C^{\mathrm{N}}_{\Omega_{\mathrm{N}}}\}, \quad \nu = C_{\mathrm{sym}}\ \varepsilon'$	(3.7)
E	$C^L_E = \frac{\sigma_L + \sigma_N \nu}{1-\nu^2}, \quad C^{\mathrm{N}}_E = \frac{\sigma_N + \sigma_L \nu}{1-\nu^2}, \quad C_E = C^L_E + \gamma\delta^\tau C^{\mathrm{N}}_E$	(3.9)
$E^\top$	$C^L_{E^\top} = n\frac{\sigma_{L^\top} + \nu\sigma_{N^\top}}{1-\nu^2}, \quad C^{\mathrm{N}}_{E^\top} = n\frac{\sigma_{N^\top} + \nu\sigma_{L^\top}}{1-\nu^2}, \quad C_E = C^L_{E^\top} + \gamma\delta^\tau C^{\mathrm{N}}_{E^\top}$	(3.10)
$\mathrm{D}E$	$C^L_{\mathrm{D}E} = d(1 + c_\Omega(\sigma_{N^\top} C^L_E + \sigma_{L^\top} C^{\mathrm{N}}_E))C^L_E$ $C^{\mathrm{N}}_{\mathrm{D}E} = d(1 + c_\Omega(\sigma_{N^\top} C^L_E + \sigma_{L^\top} C^{\mathrm{N}}_E))C^{\mathrm{N}}_E, \quad C_{\mathrm{D}E} = C^L_{\mathrm{D}E} + \gamma\delta^\tau C^{\mathrm{N}}_{\mathrm{D}E}$	(3.11)
$(\mathrm{D}E)^\top$	$C^L_{(\mathrm{D}E)^\top} = n(1 + c_\Omega(\sigma_N C^L_{E^\top} + \sigma_L C^{\mathrm{N}}_{E^\top}))C^L_{E^\top}$ $C^{\mathrm{N}}_{(\mathrm{D}E)^\top} = n(1 + c_\Omega(\sigma_N C^L_{E^\top} + \sigma_L C^{\mathrm{N}}_{E^\top}))C^{\mathrm{N}}_{E^\top}, \quad C_{(\mathrm{D}E)^\top} = C^L_{(\mathrm{D}E)^\top} + \gamma\delta^\tau C^{\mathrm{N}}_{(\mathrm{D}E)^\top}$	(3.12)
$\mathfrak{L}_\omega K$	$C_{\mathfrak{L}K} = C_E\varepsilon + c_{X_h}$	(3.13)
$\mathfrak{L}_\omega L$	$C_{\mathfrak{L}L} = C_{\mathrm{D}E}\varepsilon' + c_{\mathrm{D}X_p} C_E \varepsilon + c_{\mathrm{D}X_h}\sigma_L$	(3.14)
$\mathfrak{L}_\omega L^\top$	$C_{\mathfrak{L}L^\top} = \max\{C_{(\mathrm{D}E)^\top}\varepsilon', c_{\mathrm{D}X_p^\top} C_E\varepsilon\} + \sigma_{L^\top} c_{(\mathrm{D}X_h)^\top}$	(3.15)
$\mathfrak{L}_\omega G_L$	$C_{\mathfrak{L}G_L} = C_{\mathfrak{L}L^\top} c_G \sigma_L + \sigma_{L^\top} c_{\mathrm{D}G} C_{\mathfrak{L}K}\sigma_L + \sigma_{L^\top} c_G C_{\mathfrak{L}L}$	(3.16)
$\mathfrak{L}_\omega B$	$C_{\mathfrak{L}B} = (\sigma_B)^2 C_{\mathfrak{L}G_L}$	(3.17)
$\mathfrak{L}_\omega N$	$C_{\mathfrak{L}N} = c_{\mathrm{D}J} C_{\mathfrak{L}K}\sigma_L\sigma_B + c_J C_{\mathfrak{L}L}\sigma_B + c_J \sigma_L C_{\mathfrak{L}B}$	(3.18)
T	$C_T = \sigma_{N^\top} c_{T_h} \sigma_N$	(3.19)
E^{LL}_{red}	$C^L_{E^{\mathrm{LL}}_{\mathrm{red}}} = d + (d + \delta c_{\mathrm{D}X_p})\sigma_{N^\top} c_\Omega C^L_E, \quad C^N_{E^{\mathrm{LL}}_{\mathrm{red}}} = (d + \delta c_{\mathrm{D}X_p})\sigma_{N^\top} c_\Omega C^{\mathrm{N}}_E$	(3.21)
E^{NL}_{red}	$C^L_{E^{\mathrm{NL}}_{\mathrm{red}}} = (d + \delta c_{\mathrm{D}X_p})\sigma_{L^\top} c_\Omega C^L_E, \quad C^N_{E^{\mathrm{NL}}_{\mathrm{red}}} = d + (d + \delta c_{\mathrm{D}X_p})\sigma_{L^\top} c_\Omega C^{\mathrm{N}}_E$	(3.22)
E^{NN}_{red}	$C^L_{E^{\mathrm{NN}}_{\mathrm{red}}} = \delta\ \sigma_{L^\top} c_{\mathrm{D}\Omega} C^L_E \sigma_N + \max\{n + dC^L_E c_\Omega \sigma_{N^\top}, \delta c_{\mathrm{D}X_p^\top} C^L_E c_\Omega \sigma_N\}$ $C^N_{E^{\mathrm{NN}}_{\mathrm{red}}} = \delta\ \sigma_{L^\top} c_{\mathrm{D}\Omega} C^{\mathrm{N}}_E \sigma_N + \max\{dC^{\mathrm{N}}_E c_\Omega \sigma_{N^\top}, \delta c_{\mathrm{D}X_p^\top} C^{\mathrm{N}}_E c_\Omega \sigma_N\}$ $C_{E^{\mathrm{NN}}_{\mathrm{red}}} = C^L_{E^{\mathrm{NN}}_{\mathrm{red}}} + \gamma\delta^\tau C^N_{E^{\mathrm{NN}}_{\mathrm{red}}}$	(3.23)
E^{LN}_{red}	$C^L_{E^{\mathrm{LN}}_{\mathrm{red}}} = \delta(\sigma_B)^2 \sigma_{L^\top} c_G c_{\mathrm{D}J} \sigma_L C^L_E + (\sigma_B)^2 C^L_{E^{\mathrm{NL}}_{\mathrm{red}}}$ $C^N_{E^{\mathrm{LN}}_{\mathrm{red}}} = \gamma\delta^{\tau+1}(\sigma_B)^2 \sigma_{L^\top} c_G c_{\mathrm{D}J} \sigma_L C^{\mathrm{N}}_E + \gamma\delta^\tau(\sigma_B)^2 C^N_{E^{\mathrm{NL}}_{\mathrm{red}}} + \sigma_B C^{\mathrm{N}}_{\Omega_L} C_{\mathfrak{L}B}$ $C_{E^{\mathrm{LN}}_{\mathrm{red}}} = C^L_{E^{\mathrm{LN}}_{\mathrm{red}}} + C^N_{E^{\mathrm{LN}}_{\mathrm{red}}}$	(3.24)

Table 2
Constants in the proof of Lemma 3.6, steps 1 and 2.

Object	Constant	Label
$\hat{\xi}^N_0$	$C^L_{\hat{\xi}^N_0} = \sigma_{\langle T\rangle^{-1}}, \quad C^N_{\hat{\xi}^N_0} = \left(\frac{\delta}{\rho}\right)^\tau \sigma_{\langle T\rangle^{-1}} C_T\ c_{\mathfrak{R}}(\rho), \quad C_{\hat{\xi}^N_0} = C^L_{\hat{\xi}^N_0} + C^N_{\hat{\xi}^N_0}$	(3.44)
ξ^N	$C^L_{\xi^N} = C^L_{\hat{\xi}^N_0}, \quad C^N_{\xi^N} = C^N_{\hat{\xi}^N_0} + c_{\mathfrak{R}}(\delta), \quad C_{\xi^N} = C^L_{\xi^N} + C^N_{\xi^N}$	(3.45)
$\bar{K} - K$	$C^L_{\Delta\bar{K}} = \sigma_N C^L_{\xi^N}, \quad C^N_{\Delta\bar{K}} = \sigma_N C^N_{\xi^N}, \quad C_{\Delta\bar{K}} = C^L_{\Delta\bar{K}} + C^N_{\Delta\bar{K}}$	(3.46)
$\mathrm{D}\bar{K} - \mathrm{D}K$	$C^L_{\Delta\mathrm{D}\bar{K}} = d\sigma_N C^L_{\hat{\xi}^N_0}, \quad C^N_{\Delta\mathrm{D}\bar{K}} = d\sigma_N (C^N_{\hat{\xi}^N_0} + c^1_{\mathfrak{R}}(\delta)), \quad C_{\Delta\mathrm{D}\bar{K}} = C^L_{\Delta\mathrm{D}\bar{K}} + C^N_{\Delta\mathrm{D}\bar{K}}$	(3.48)
$(\mathrm{D}\bar{K})^\top - (\mathrm{D}K)^\top$	$C^L_{\Delta(\mathrm{D}\bar{K})^\top} = n\sigma_{N^\top} C^L_{\hat{\xi}^N_0},$ $C^N_{\Delta(\mathrm{D}\bar{K})^\top} = n\sigma_{N^\top} (C^N_{\hat{\xi}^N_0} + c^1_{\mathfrak{R}}(\delta)), \quad C_{\Delta(\mathrm{D}\bar{K})^\top} = C^L_{\Delta(\mathrm{D}\bar{K})^\top} + C^N_{\Delta(\mathrm{D}\bar{K})^\top}$	(3.49)
$\bar{L} - L$	$C^L_{\Delta\bar{L}} = C^L_{\Delta\mathrm{D}\bar{K}} + \delta c_{\mathrm{D}X_p} C^L_{\Delta\bar{K}}, \quad C^N_{\Delta\bar{L}} = C^N_{\Delta\mathrm{D}\bar{K}} + \delta c_{\mathrm{D}X_p} C^N_{\Delta\bar{K}}, \quad C_{\Delta\bar{L}} = C^L_{\Delta\bar{L}} + C^N_{\Delta\bar{L}}$	(3.53)
$\bar{L}^\top - L^\top$	$C^L_{\Delta\bar{L}^\top} = \max\{C^L_{\Delta(\mathrm{D}\bar{K})^\top}, \delta c_{\mathrm{D}X_p^\top} C^L_{\Delta\bar{K}}\},$ $C^N_{\Delta\bar{L}^\top} = \max\{C^N_{\Delta(\mathrm{D}\bar{K})^\top}, \delta c_{\mathrm{D}X_p^\top} C^N_{\Delta\bar{K}}\}, \quad C_{\Delta\bar{L}^\top} = C^L_{\Delta\bar{L}^\top} + C^N_{\Delta\bar{L}^\top}$	(3.54)
$\bar{E}$	$C^L_{\bar{E}} = C^L_E + c_{\mathrm{D}X_h} C^L_{\Delta\bar{K}} + C_{\mathfrak{L}N} C^L_{\xi^N},$ $C^N_{\bar{E}} = \gamma\delta^\tau(C^{\mathrm{N}}_E + \sigma_N) + c_{\mathrm{D}X_h} C^N_{\Delta\bar{K}} + C_{\mathfrak{L}N} C^N_{\xi^N}, \quad C_{\bar{E}} = C^L_{\bar{E}} + C^N_{\bar{E}}$	(3.55)
$\bar{\eta}^N$	$Q_{\bar{\eta}^N} = (C_{\Delta\bar{L}^\top} c_\Omega + \delta\sigma_{L^\top} c_{\mathrm{D}\Omega} C_{\Delta\bar{K}})C_{\bar{E}} + C_{E^{\mathrm{NN}}_{\mathrm{red}}} C_{\xi^N} + \delta\sigma_{L^\top} c_\Omega \frac{1}{2} c_{\mathrm{D}^2 X_h} (C_{\Delta\bar{K}})^2$	(3.58)

Table 3
Constants in the proof of Lemma 3.6, steps 3 and 4.

Object	Constant	Label
ξ^L	$C^L_{\xi^L} = c_{\mathfrak{R}}(\delta)(1 + C_T C^L_{\xi^N}), \quad C^N_{\xi^L} = c_{\mathfrak{R}}(\delta) C_T C^N_{\xi^N}, \quad C_{\xi^L} = C^L_{\xi^L} + C^N_{\xi^L}$	(3.60)
$\bar{\bar{K}} - K$	$C^L_{\Delta\bar{\bar{K}}} = c_{\mathrm{D}\Phi}\sigma_L C^L_{\xi^L} + \gamma\delta^\tau C^L_{\Delta\bar{K}},$ $C^N_{\Delta\bar{\bar{K}}} = c_{\mathrm{D}\Phi}\sigma_L C^N_{\xi^L} + \gamma\delta^\tau C^N_{\Delta\bar{K}}, \quad C_{\Delta\bar{\bar{K}}} = C^L_{\Delta\bar{\bar{K}}} + C^N_{\Delta\bar{\bar{K}}}$	(3.63)
$\mathrm{D}\bar{\bar{K}} - \mathrm{D}K$	$C^L_{\Delta\mathrm{D}\bar{\bar{K}}} = d c_{\mathrm{D}\Phi}\sigma_L C^L_{\xi^L} + \gamma\delta^\tau C^L_{\Delta\mathrm{D}\bar{K}}, \quad C^N_{\Delta\mathrm{D}\bar{\bar{K}}} = d c_{\mathrm{D}\Phi}\sigma_L C^N_{\xi^L} + \gamma\delta^\tau C^N_{\Delta\mathrm{D}\bar{K}},$ $C_{\Delta\mathrm{D}\bar{\bar{K}}} = C^L_{\Delta\mathrm{D}\bar{\bar{K}}} + C^N_{\Delta\mathrm{D}\bar{\bar{K}}}$	(3.64)
$(\mathrm{D}\bar{\bar{K}})^\top - (\mathrm{D}K)^\top$	$C^L_{\Delta(\mathrm{D}\bar{\bar{K}})^\top} = 2n c_{\mathrm{D}\Phi}\sigma_L C^L_{\xi^L} + \gamma\delta^\tau C^L_{\Delta(\mathrm{D}\bar{K})^\top}, \quad C^N_{\Delta(\mathrm{D}\bar{\bar{K}})^\top} = 2n c_{\mathrm{D}\Phi}\sigma_L C^N_{\xi^L} + \gamma\delta^\tau C^N_{\Delta(\mathrm{D}\bar{K})^\top},$ $C_{\Delta(\mathrm{D}\bar{\bar{K}})^\top} = C^L_{\Delta(\mathrm{D}\bar{\bar{K}})^\top} + C^N_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}$	(3.65)
$\bar{\bar{L}} - L$	$C^L_{\Delta\bar{\bar{L}}} = C^L_{\Delta\mathrm{D}\bar{\bar{K}}} + \delta c_{\mathrm{D}X_p} C^L_{\Delta\bar{\bar{K}}},$ $C^N_{\Delta\bar{\bar{L}}} = C^N_{\Delta\mathrm{D}\bar{\bar{K}}} + \delta c_{\mathrm{D}X_p} C^N_{\Delta\bar{\bar{K}}}, \quad C_{\Delta\bar{\bar{L}}} = C^L_{\Delta\bar{\bar{L}}} + C^N_{\Delta\bar{\bar{L}}}$	(3.66)
$\bar{\bar{L}}^\top - L^\top$	$C^L_{\Delta\bar{\bar{L}}^\top} = \max\{C^L_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}, \delta c_{\mathrm{D}X_p^\top} C^L_{\Delta\bar{\bar{K}}}\},$ $C^N_{\Delta\bar{\bar{L}}^\top} = \max\{C^N_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}, \delta c_{\mathrm{D}X_p^\top} C^N_{\Delta\bar{\bar{K}}}\}, \quad C_{\Delta\bar{\bar{L}}^\top} = C^L_{\Delta\bar{\bar{L}}^\top} + C^N_{\Delta\bar{\bar{L}}^\top}$	(3.67)
$G_{\bar{\bar{L}}} - G_L$	$C^L_{\Delta G_{\bar{\bar{L}}}} = C^L_{\Delta\bar{\bar{L}}^\top} c_G \sigma_L + \delta\sigma_{L^\top} c_{\mathrm{D}G} C^L_{\Delta\bar{\bar{K}}}\sigma_L + \sigma_{L^\top} c_G C^L_{\Delta\bar{\bar{L}}},$ $C^N_{\Delta G_{\bar{\bar{L}}}} = C^N_{\Delta\bar{\bar{L}}^\top} c_G \sigma_L + \delta\sigma_{L^\top} c_{\mathrm{D}G} C^N_{\Delta\bar{\bar{K}}}\sigma_L + \sigma_{L^\top} c_G C^N_{\Delta\bar{\bar{L}}}, \quad C_{\Delta G_{\bar{\bar{L}}}} = C^L_{\Delta G_{\bar{\bar{L}}}} + C^N_{\Delta G_{\bar{\bar{L}}}}$	(3.68)
$\bar{\bar{B}} - B$	$C^L_{\Delta\bar{\bar{B}}} = (\sigma_B)^2 C^L_{\Delta G_{\bar{\bar{L}}}}, \quad C^N_{\Delta\bar{\bar{B}}} = (\sigma_B)^2 C^N_{\Delta G_{\bar{\bar{L}}}}, \quad C_{\Delta\bar{\bar{B}}} = (\sigma_B)^2 C_{\Delta G_{\bar{\bar{L}}}}$	(3.69)
$\bar{\bar{N}} - N$	$C^L_{\Delta\bar{\bar{N}}} = \delta c_{\mathrm{D}J} C^L_{\Delta\bar{\bar{K}}}\sigma_L\sigma_B + c_J C^L_{\Delta\bar{\bar{L}}}\sigma_B + c_J\sigma_L C^L_{\Delta\bar{\bar{B}}},$ $C^N_{\Delta\bar{\bar{N}}} = \delta c_{\mathrm{D}J} C^N_{\Delta\bar{\bar{K}}}\sigma_L\sigma_B + c_J C^N_{\Delta\bar{\bar{L}}}\sigma_B + c_J\sigma_L C^N_{\Delta\bar{\bar{B}}}, \quad C_{\Delta\bar{\bar{N}}} = C^L_{\Delta\bar{\bar{N}}} + C^N_{\Delta\bar{\bar{N}}}$	(3.70)
$\bar{\bar{N}}^\top - N^\top$	$C^L_{\Delta\bar{\bar{N}}^\top} = \delta\sigma_B\sigma_L c_{\mathrm{D}J^\top} C^L_{\Delta\bar{\bar{K}}} + \sigma_B C^L_{\Delta\bar{\bar{L}}} c_{J^\top} + C^L_{\Delta\bar{\bar{B}}}\sigma_L c_{J^\top},$ $C^N_{\Delta\bar{\bar{N}}^\top} = \delta\sigma_B\sigma_L c_{\mathrm{D}J^\top} C^N_{\Delta\bar{\bar{K}}} + \sigma_B C^N_{\Delta\bar{\bar{L}}} c_{J^\top} + C^N_{\Delta\bar{\bar{B}}}\sigma_L c_{J^\top}, \quad C_{\Delta\bar{\bar{N}}^\top} = C^L_{\Delta\bar{\bar{N}}^\top} + C^N_{\Delta\bar{\bar{N}}^\top}$	(3.71)
$\bar{\bar{T}} - T$	$C^L_{\Delta\bar{\bar{T}}} = C^L_{\Delta\bar{\bar{N}}^\top} c_{T_h}\sigma_N + \delta\sigma_{N^\top} c_{\mathrm{D}T_h} C^L_{\Delta\bar{\bar{K}}}\sigma_N + \sigma_{N^\top} c_{T_h} C^L_{\Delta\bar{\bar{N}}},$ $C^N_{\Delta\bar{\bar{T}}} = C^N_{\Delta\bar{\bar{N}}^\top} c_{T_h}\sigma_N + \delta\sigma_{N^\top} c_{\mathrm{D}T_h} C^N_{\Delta\bar{\bar{K}}}\sigma_N + \sigma_{N^\top} c_{T_h} C^N_{\Delta\bar{\bar{N}}}, \quad C_{\Delta\bar{\bar{T}}} = C^L_{\Delta\bar{\bar{T}}} + C^N_{\Delta\bar{\bar{T}}}$	(3.72)
$\langle\bar{\bar{T}}\rangle^{-1} - \langle T\rangle^{-1}$	$C^L_{\Delta\langle\bar{\bar{T}}\rangle^{-1}} = (\sigma_{\langle T\rangle^{-1}})^2 C^L_{\Delta\bar{\bar{T}}}, \quad C^N_{\Delta\langle\bar{\bar{T}}\rangle^{-1}} = (\sigma_{\langle T\rangle^{-1}})^2 C^N_{\Delta\bar{\bar{T}}}, \quad C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}} = (\sigma_{\langle T\rangle^{-1}})^2 C_{\Delta\bar{\bar{T}}}$	(3.73)
$\mathfrak{L}\xi^L$	$C^L_{\mathfrak{L}\xi^L} = 1 + C_T C^L_{\xi^N}, \quad C^N_{\mathfrak{L}\xi^L} = C_T C^N_{\xi^N}, \quad C_{\mathfrak{L}\xi^L} = C^L_{\mathfrak{L}\xi^L} + C^N_{\mathfrak{L}\xi^L}$	(3.76)
$\bar{\bar{E}}$	$C^L_{\bar{\bar{E}}} = c_{\mathrm{D}\Phi}(C^L_{\bar{E}} + \sigma_L C^L_{\mathfrak{L}\xi^L}), \quad C^N_{\bar{\bar{E}}} = c_{\mathrm{D}\Phi}(C^N_{\bar{E}} + \sigma_L C^N_{\mathfrak{L}\xi^L}), \quad C_{\bar{\bar{E}}} = C^L_{\bar{\bar{E}}} + C^N_{\bar{\bar{E}}}$	(3.77)
$\bar{\bar{\eta}}^N$	$Q_{\bar{\bar{\eta}}^N} = (1+n)(1 + C^{\mathrm{N}}_{\Omega_L} C_{\mathfrak{L}\xi^L}\varkappa) Q_{\bar{\eta}^N}$	(3.80)
$\bar{\bar{\eta}}^L_1$	$Q_{\bar{\bar{\eta}}^L_1} = C_{E^{\mathrm{LN}}_{\mathrm{red}}} C_{\xi^N} + \delta\sigma_{N^\top} c_\Omega \frac{1}{2} c_{\mathrm{D}^2 X_h} (C_{\Delta\bar{K}})^2 + \gamma\delta^\tau C^{\mathrm{N}}_{\Omega_{\mathrm{N}}}$	(3.82)
$\bar{\bar{\eta}}^L_2$	$Q_{\bar{\bar{\eta}}^L_2} = \sigma_{N^\top} c_\Omega\, d\, C_{\bar{E}} C_{\xi^L}$	(3.83)
$\bar{\bar{\eta}}^L_3$	$Q_{\bar{\bar{\eta}}^L_3} = \sigma_{N^\top} c_\Omega\, (d\, \sigma_L C_{\xi^L} + \gamma\delta^\tau C_{\Delta\bar{L}}) C_{\mathfrak{L}\xi^L}$	(3.84)
$\bar{\bar{\eta}}^L_4$	$Q_{\bar{\bar{\eta}}^L_4} = \sigma_{N^\top} c_\Omega c_{\mathrm{D}X_p} c_{\mathrm{D}\Phi} C_{\xi^L} (C_{\bar{E}} + \sigma_L C_{\mathfrak{L}\xi^L})$	(3.85)
$\bar{\bar{\eta}}^L_5$	$Q_{\bar{\bar{\eta}}^L_5} = (C_{\Delta\bar{\bar{N}}^\top} c_\Omega + \delta\sigma_{N^\top} c_{\mathrm{D}\Omega} C_{\Delta\bar{\bar{K}}}) C_{\bar{\bar{E}}}$	(3.86)
$\bar{\bar{\eta}}^L$	$Q_{\bar{\bar{\eta}}^L} = \gamma\delta^\tau Q_{\bar{\bar{\eta}}^L_1} + Q_{\bar{\bar{\eta}}^L_2} + Q_{\bar{\bar{\eta}}^L_3} + \delta Q_{\bar{\bar{\eta}}^L_4} + Q_{\bar{\bar{\eta}}^L_5}$	(3.87)

Table 4
Constants in Theorem Theorem 2.18, defined in Section 3.3.

Constant	Label
$a = \frac{\rho-\rho_\infty}{\rho-3\delta-\rho_\infty}$	(3.88)
$Q_{\bar{\bar{\eta}}} = \left\lvert Q_{\bar{\bar{\eta}}^L}, a^\tau Q_{\bar{\bar{\eta}}^N} \right\rvert$	(3.89)
$C_{\mathrm{dist}} = \frac{C_{\Delta\bar{\bar{K}}}}{\mathrm{dist}(K(\mathbb{T}^d_\rho), \partial\mathcal{U}_0)}$	(3.98)
$C_\Delta = \max\left\{ \frac{C_{\Delta\mathrm{D}\bar{\bar{K}}}}{\sigma_{\mathrm{D}K} - \lVert\mathrm{D}K\rVert_\rho}, \frac{C_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}}{\sigma_{(\mathrm{D}K)^\top} - \lVert(\mathrm{D}K)^\top\rVert_\rho}, \frac{C_{\Delta\bar{\bar{B}}}}{\sigma_B - \lVert B\rVert_\rho}, \frac{C_{\Delta\bar{\bar{N}}}}{\sigma_N - \lVert N\rVert_\rho}, \frac{C_{\Delta\bar{\bar{N}}^\top}}{\sigma_{N^\top} - \lVert N^\top\rVert_\rho}, \frac{C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}}{\sigma_{\langle T\rangle^{-1}} - \lvert\langle T\rangle^{-1}\rvert} \right\}$	(3.99)
$\mathfrak{C} = \max\left\{ \gamma\delta^\tau C_{\mathrm{sym}}, \; C_{\xi^L}, \delta\, C_{\mathrm{dist}} + Q_{\bar{\bar{\eta}}} a^\tau, \; C_\Delta + Q_{\bar{\bar{\eta}}} a^{\tau+1} \right\}$	(3.97)
$\kappa = a^{\tau+1} Q_{\bar{\bar{\eta}}} \varkappa$	(3.91)
$\mathfrak{C}_{\Delta K} = \frac{a}{a-\kappa} C_{\Delta\bar{\bar{K}}}$	(2.32)
$\mathfrak{C}_{\Delta\mathrm{D}K} = \frac{1}{1-\kappa} C_{\Delta\mathrm{D}\bar{\bar{K}}}$	(2.33)
$\mathfrak{C}_{\Delta(\mathrm{D}K)^\top} = \frac{1}{1-\kappa} C_{\Delta(\mathrm{D}\bar{\bar{K}})^\top}$	(2.34)
$\mathfrak{C}_{\Delta B} = \frac{1}{1-\kappa} C_{\Delta\bar{\bar{B}}}$	(2.35)
$\mathfrak{C}_{\Delta N} = \frac{1}{1-\kappa} C_{\Delta\bar{\bar{N}}}$	(2.36)
$\mathfrak{C}_{\Delta N^\top} = \frac{1}{1-\kappa} C_{\Delta\bar{\bar{N}}^\top}$	(2.37)
$\mathfrak{C}_{\Delta\langle T\rangle^{-1}} = \frac{1}{1-\kappa} C_{\Delta\langle\bar{\bar{T}}\rangle^{-1}}$	(2.38)

CRediT authorship contribution statement

Jordi-Lluís Figueras: Investigation, Writing – original draft, Writing – review & editing. **Alex Haro:** Investigation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jordi-Lluis Figueras Romero reports financial support was provided by Swedish Research Council.

Data availability

No data was used for the research described in the article.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) have not used any AI tool for the generation of the text nor for the creation of intellectual content.

Acknowledgments

The authors are grateful to Alejandro Luque, Benjamin Meco, Christian Bjerklöv and Gerard Farré for fruitful discussions.

J.-Ll.F. has been partially supported by the Swedish VR Grant 2019-04591, and A.H. has been supported by the Spanish grant PID2021-125535NB-I00 (MCIU/AEI/FEDER, UE). This work has also being supported by the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M).

Appendix A. An auxiliary lemma to control the inverse of a matrix

In several instances in the proofs performed in this paper we control the correction of inverses of matrices using Neumann series argument. For instance, this affects to the estimates in (3.37), (3.40). For convenience, we present the following auxiliary result separately. Notice that the result is presented for matrices but it is directly extended for matrix-valued maps with the corresponding norm (see Section 2.7).

Lemma A.1. *Let $M \in \mathbb{C}^{n\times n}$ be an invertible matrix satisfying $|M^{-1}| < \sigma$. Assume that $\bar{M} \in \mathbb{C}^{n\times n}$ satisfies*

$$\frac{\sigma|M^{-1}||\bar{M}-M|}{\sigma-|M^{-1}|} < 1\,. \tag{A.1}$$

Then, we have that $\bar{M}$ is invertible and

$$|\bar{M}^{-1}| < \sigma \qquad |\bar{M}^{-1}-M^{-1}| \le \sigma|M^{-1}||\bar{M}-M|.$$

Proof. For notational convenience, let us denote

$$x = \frac{|M^{-1}|}{\sigma}, \quad \lambda = \frac{\sigma|M^{-1}||\bar{M}-M|}{\sigma-|M^{-1}|},$$

so that $x \in \,]0,1[$ and $\lambda \in [0,1[$. Since

$$|M^{-1}||\bar{M}-M| = \lambda(1-x) < 1,$$

the matrix

$$\bar{M} = M(I + M^{-1}(\bar{M}-M))$$

is invertible and

$$|\bar{M}^{-1}| \le \frac{\sigma x}{1-\lambda(1-x)} < \sigma,$$

where in the last inequality we use that, since $\lambda < 1$, the function $f(x) = x/(1-\lambda(1-x))$ is continuous and strictly increasing in the interval $[0,1]$, and $f(1) = 1$. Moreover,

$$|\bar{M}^{-1}-M^{-1}| = |\bar{M}^{-1}(M-\bar{M})M^{-1}| \;\le\; \sigma|M^{-1}||\bar{M}-M|. \quad \square$$

Remark A.2. In particular, notice that, if $|M^{-1}| < \sigma$, hypothesis (A.1) holds if

$$\frac{\sigma^2|\bar{M}-M|}{\sigma-|M^{-1}|} < 1.$$

Appendix B. Compendium of constants involved in the KAM theorem

In this appendix we collect the recipes to compute all constants involved in the different estimates presented in the proof of the main results of this paper. Keeping track of these constants is crucial to apply Theorem 2.18 in particular problems and for concrete values of parameters. In the following, Table 1 corresponds to the geometric constructions outlined in Section 3.1, and Tables 2 and 3 correspond to the iterative lemma in Section 3.2. Finally, Table 4 presents the constants in Theorem 2.18, that are defined at the end of its proof, in Section 3.3.

The input values are (see Theorem 2.18) :

- the global bounds c_Ω, c_G, c_J, $c_{J^\top}$, $c_{\mathrm{D}\Omega}$, $c_{\mathrm{D}G}$, $c_{\mathrm{D}J}$, $c_{\mathrm{D}J^\top}$, c_{X_h}, $c_{\mathrm{D}X_h}$, $c_{(\mathrm{D}X_h)^\top}$, $c_{\mathrm{D}^2X_h}$, c_{T_h}, $c_{\mathrm{D}T_h}$, c_{X_p}, $c_{X_p^\top}$, $c_{\mathrm{D}X_p}$, $c_{\mathrm{D}X_p^\top}$, and $c_{\mathrm{D}\Phi}$;
- the condition numbers $\sigma_{\mathrm{D}K}$, $\sigma_{(\mathrm{D}K)^\top}$, σ_L, $\sigma_{L^\top}$, σ_B, σ_N, $\sigma_{N^\top}$, and $\sigma_{\langle T\rangle^{-1}}$;
- the strip sizes $\rho_\infty < \rho < r$ and width reduction $\delta < \frac{\rho-\rho_\infty}{3}$;
- the weighted errors

$$\varepsilon = \left|\left\|\eta^L\right\|_\rho, \frac{1}{\gamma\delta^\tau}\left\|\eta^N\right\|_\rho\right|, \quad \varepsilon' = \frac{1}{\delta}\varepsilon, \quad \varkappa = \frac{1}{\gamma\delta^{\tau+1}}\varepsilon. \tag{B.1}$$

Constants $c_{\mathfrak{R}}(\delta)$ and $c^1_{\mathfrak{R}}(\delta)$ from Lemma 2.14 and Corollary 2.15 depend on the width reduction on the strip, δ, and can substituted by the corresponding upper bounds $\hat{c}_{\mathfrak{R}}$ and $\hat{c}^1_{\mathfrak{R}}$.

References

[1] A.N. Kolmogorov, On conservation of conditionally periodic motions for a small change in Hamilton's function, Dokl. Akad. Nauk SSSR (N.S.) 98 (1954) 527–530, Translated in p. 51–56 of Stochastic Behavior in Classical and Quantum Hamiltonian Systems, Como 1977 (eds. G. Casati and J. Ford) Lect. Notes Phys. 93, Springer, Berlin, 1979.

[2] V.I. Arnold, Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian, Uspekhi Mat. Nauk 18 (5 (113)) (1963) 13–40.

[3] J. Moser, On invariant curves of area-preserving mappings of an annulus, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II 1962 (1962) 1–20.

[4] H.S. Dumas, The KAM Story, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014, A friendly introduction to the content, history, and significance of classical Kolmogorov-Arnold-Moser theory.

[5] H.W. Broer, G.B. Huitema, M.B. Sevryuk, Quasi-periodic motions in families of dynamical systems, in: Order Amidst Chaos, in: Lecture Notes in Math., vol 1645, Springer-Verlag, Berlin, 1996.

[6] R. de la Llave, A tutorial on KAM theory, in: Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), in: Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc. Providence, RI, 2001, pp. 175–292.

[7] A. Celletti, L. Chierchia, Construction of analytic KAM surfaces and effective stability bounds, Comm. Math. Phys. 118 (1) (1988) 161–199.

[8] R. de la Llave, D. Rana, Accurate strategies for small divisor problems, Bull. Amer. Math. Soc. (N.S.) 22 (1) (1990) 85–90.

[9] A. Celletti, L. Chierchia, KAM stability and celestial mechanics, Mem. Amer. Math. Soc. 187 (878) (2007) viii+134.

[10] R. de la Llave, A. González, À. Jorba, J. Villanueva, KAM theory without action–angle variables, Nonlinearity 18 (2) (2005) 855–895.

[11] A. Haro, M. Canadell, J.-Ll. Figueras, A. Luque, J.-M. Mondelo, The Parameterization Method for Invariant Manifolds, in: Applied Mathematical Sciences, vol. 195, Springer, [Cham], 2016, From rigorous results to effective computations.

[12] M.B. Sevryuk, The classical KAM theory in the last decade: A slow progress, 2014, http://www.mathnet.ru/php/presentation.phtml?option_lang=eng&presentid=10271, The Seventh International Conference on Differential and Functional Differential Equations..

[13] Jordi-Lluís Figueras, Alex Haro, Alejandro Luque, Effective bounds for the measure of rotations, Nonlinearity 33 (2) (2020) 700–741.
[14] J.-Ll. Figueras, A. Haro, A. Luque, Rigorous computer-assisted application of KAM theory: A modern approach, Found. Comput. Math. 17 (5) (2017) 1123–1193.
[15] A. Cannas da Silva, Lectures on Symplectic Geometry, in: Lecture Notes in Mathematics, vol. 1764, Springer-Verlag, Berlin, 2001.
[16] J. Marsden, A. Weinstein, Reduction of symplectic manifolds with symmetry, Rep. Math. Phys. 5 (1) (1974) 121–130.
[17] A. Haro, A. Luque, A-posteriori KAM theory with optimal estimates for partially integrable systems, J. Differential Equations 266 (2–3) (2019) 1605–1674.
[18] J. Villanueva, A new approach to the parameterization method for Lagrangian tori of Hamiltonian systems, J. Nonlinear Sci. 27 (2017) 495–530.
[19] Jordi. Villanueva, A parameterization method for Lagrangian tori of exact symplectic maps of $\mathbb{R}^{2r}$, SIAM J. Appl. Dyn. Syst. 17 (3) (2018) 2289–2331.
[20] J. Villanueva, Kolmogorov theorem revisited, J. Differential Equations 244 (9) (2008) 2251–2276.
[21] A.I. Neishtadt, Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions, J. Appl. Math. Mech. 45 (6) (1981) 1016–1025.
[22] J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, Comm. Pure Appl. Math. 35 (5) (1982) 653–696.
[23] A. González, A. Haro, R. de la Llave, Singularity theory for non-twist KAM tori, Mem. Amer. Math. Soc. 227 (1067) (2014) vi+115.
[24] E. Fontich, R. de la Llave, Y. Sire, Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions, J. Differ. Equ. 246 (8) (2009) 3136–3213.
[25] M.-R. Herman, Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 2, Astérisque (144) (1986) 248, With a correction to: On the curves invariant under diffeomorphisms of the annulus, Vol. 1 (French) [Astérisque No. 103-104, Soc. Math. France, Paris, 1983; MR 85 m:58062].
[26] J. Moser, On the theory of quasiperiodic motions, SIAM Rev. 8 (2) (1966) 145–172.
[27] H. Rüssmann, On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the Torus, in: Dynamical Systems, Theory and Applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974), in: Lecture Notes in Phys., vol. 38, Springer, Berlin, 1975, pp. 598–624.
[28] H. Rüssmann, On optimal estimates for the solutions of linear difference equations on the circle, in: Proceedings of the Fifth Conference on Mathematical Methods in Celestial Mechanics, (Oberwolfach, 1975), Part I, in: Celestial Mech., vol. 14, 1976.
[29] J.-Ll. Figueras, A. Haro, A. Luque, On the sharpness of the Rüssmann estimates, Commun. Nonlinear Sci. Numer. Simul. 55 (2018) 42–55.