

A-POSTERIORI KAM THEORY WITH OPTIMAL ESTIMATES FOR PARTIALLY INTEGRABLE SYSTEMS

ALEX HARO[†] AND ALEJANDRO LUQUE[‡]

ABSTRACT. In this paper we present a-posteriori KAM results for existence of d -dimensional isotropic invariant tori for n -DOF Hamiltonian systems with additional $n - d$ independent first integrals in involution. We carry out a covariant formulation that does not require the use of action-angle variables nor symplectic reduction techniques. The main advantage is that we overcome the curse of dimensionality avoiding the practical shortcomings produced by the use of reduced coordinates, which may cause difficulties and underperformance when quantifying the hypotheses of the KAM theorem in such reduced coordinates. The results include ordinary and (generalized) iso-energetic KAM theorems. The approach is suitable to perform numerical computations and computer assisted proofs.

CONTENTS

1. Introduction	2
2. Background and elementary constructions	4
2.1. Basic notation	4
2.2. Hamiltonian systems and invariant tori	6
2.3. Conserved quantities and families of invariant tori	8
2.4. Linearized dynamics and reducibility	9
2.5. Construction of a geometrically adapted frame	11
2.6. Univocal determination of an invariant torus of the family	12
3. A-posteriori KAM theory for partially integrable Hamiltonian systems	14
3.1. Analytic functions and norms	14
3.2. Ordinary KAM theorem	16
3.3. Generalized iso-energetic KAM theorem	18
4. Some lemmas to control approximate geometric properties	20
4.1. Estimates for cohomological equations	20
4.2. Approximate conserved quantities	21
4.3. Approximate isotropicity of tangent vectors	22
4.4. Approximate symplectic frame	24
4.5. Sharp control of the torsion matrix	26
4.6. Approximate reducibility	29
5. Proof of the ordinary KAM theorem	31
5.1. The quasi-Newton method	31

(†) DEPARTAMENT DE MATEMÀTIQUES I INFORMÀTICA, UNIVERSITAT DE BARCELONA, GRAN VIA 585, 08007 BARCELONA, SPAIN.

(‡) DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY, BOX 480, 751 06 UPPSALA, SWEDEN
E-mail addresses: alex@maia.ub.es, alejandro.luque@math.uu.se.

Date: 2nd August 2018.

5.2. One step of the iterative procedure	32
5.3. Convergence of the iterative process	42
6. Proof of the generalized iso-energetic KAM theorem	45
6.1. The quasi-Newton method	45
6.2. One step of the iterative procedure	47
6.3. Convergence of the iterative process	52
7. Acknowledgements	53
References	53
Appendix A. An auxiliary lemma to control the inverse of a matrix	55
Appendix B. Compendium of constants involved in the KAM theorem	56

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. Classical perturbation theory experienced a breakthrough around sixty years ago, with the work of Kolmogorov [35], Arnold [1] and Moser [40], the founders of what is nowadays known as KAM theory. They overcame the so called small divisors problem, that might prevent the convergence of the series appearing in perturbative methods. Since then, KAM theory has become a full body of knowledge that connects fundamental mathematical ideas, and the literature contains eminent contributions and applications in different contexts (e.g. unfoldings and bifurcations of invariant tori [7, 8], quasi-periodic solutions in partial differential equations [4, 24, 45], hard implicit function theorems [42, 43, 51, 52], stability of perturbed planetary problems [2, 18, 26], conformally symplectic systems and the spin-orbit problem [9, 17], reversible systems [49], or existence of vortex tubes in the Euler equation [25], just to mention a few).

The importance and significance of KAM theory in Mathematics, Physics and Science in general are accounted in the popular book [23]. But, although KAM theory holds for general dynamical systems under very mild technical assumptions, its application to concrete systems becomes a challenging problem. Moreover, the “threshold of validity” of the theory (the size of the perturbation strength for which KAM theorems can be applied) seemed to be absurdly small in applications to physical systems¹. With the advent of computers and new developed methodologies, the distance between theory and practice has been shortened (see e.g. [16] for a illuminating historical introduction and [14, 21] for pioneering computer-assisted applications). But new impulses have to be made in order to make KAM theory fully applicable to realistic physical systems. In this respect, recent years have witnessed a revival of this theory. One direction that have experienced a lot of progress is the a-posteriori approach based on the parameterization method [19, 20, 33, 50].

The main signatures of the parameterization method are the application of the a-posteriori approach and a covariant formulation, free of using any particular system of coordinates. In fact, the method was baptized as KAM theory *without action-angle variables* in [20]. Instead of performing canonical transformations, the

¹This source of skepticism was pointed out by the distinguished astronomer M. Hénon, who found a threshold of validity for the perturbation of the order of 10^{-333} in early KAM theorems.

strategy consists in solving the invariance equation for an invariant torus by correcting iteratively an approximately invariant one. If the initial approximation is good enough (relatively to some non-degeneracy conditions), then there is a true invariant torus nearby [16, 20] (see also [15, 19, 42, 47] for precedents). Hence one can consider non-perturbative problems in a natural way. The approach itself implies the traditional approach (of perturbative nature) and has been extended to different theoretical settings [12, 29, 31, 36]. A remarkable feature of the parameterization method is that it leads to very fast and efficient numerical methods for the approximation of quasi-periodic invariant tori (e.g. [10, 11, 30, 34]). We refer the reader to the recent monograph [33] for detailed discussions (beyond the KAM context) on the numerical implementations of the method, examples, and a more complete bibliography.

A recently reported success in KAM theory is the design of a general methodology to perform computer assisted proofs of existence of Lagrangian invariant tori in a non-perturbative setting [27]. The methodology has been applied to several low dimensional problems obtaining almost optimal results². The program is founded on an a-posteriori theorem with explicit estimates, whose hypotheses are checked using Fast Fourier Transform (with interval arithmetics) in combination with a sharp control of the discretization error. This fine control is crucial to estimate the norm of compositions and inverses of functions, outperforming the use of symbolic manipulations. An important consequence is that the rigorous computations are executed in a very fast way, thus allowing to manipulate millions of Fourier modes comfortably.

One of the typical obstacles to directly apply the above methodology (and KAM theory in general) to realistic problems in Mechanics is the presence of additional first integrals in involution. This degeneracy implies that quasi-periodic invariant tori appear in smooth families of lower dimensional tori. The constraints linked to these conserved quantities can be removed by using classical symplectic reduction techniques [13, 38], thus obtaining a lower dimensional Hamiltonian system in a quotient manifold, where Lagrangian tori can be computed. This approach has an undeniable theoretical importance (e.g. the moment map is an object of remarkable relevance [39]) and has been successfully used to obtain perturbative KAM results in significant examples [18, 26, 44]. However, the use of symplectic reduction presents serious difficulties when applying the a-posteriori KAM approach to the reduced system. This is obvious because, in the new set of coordinates, it may be difficult to quantify the required control of the norms for both global objects (e.g. Hamiltonian system or symplectic structure) and local objects (e.g. parameterization of the invariant torus, torsion matrix), or also, these estimates may become sub-optimal to apply a quantitative KAM theorem. The goal of this paper is overcoming these drawbacks. Instead of reducing the system, we will characterize a target lower dimensional torus, using the original coordinates of the problem, by constructing a geometrically adapted frame to suitably display the linearized dynamics on the full family of invariant tori.

²For example, for the Chirikov standard map, it is proved that the invariant curve with golden rotation persists up to $\varepsilon \leq 0.9716$, which corresponds to a threshold value with a defect of 0.004% with respect to numerical observations (e.g. [32, 37]).

In this paper we present two KAM theorems in a-posteriori format for existence of (families of) isotropic³ tori in Hamiltonian systems with first integrals in involution, avoiding the use of action-angle coordinates (in the spirit of [20]) and symplectic reduction. The first theorem is an ordinary KAM theorem (also known as *à la Kolmogorov*) on existence of an invariant torus with a fixed frequency vector. Of course, if no additional first integrals are present in the system, then we recover the results in [20], with the extra bonus of providing a geometrically improved scheme with explicit estimates. Our second theorem is a version of a non-perturbative iso-energetic KAM theorem (see [5, 22] for perturbative formulations), in which one fixes either the energy or one of the first integrals but modulates the frequency vector. Actually, we present a general version that allows us to consider any conserved quantity in involution. Even if there are no additional first integrals, the corollary of this result is an iso-energetic theorem for Lagrangian tori which is a novelty in this covariant formulation.

The statements of the results are written with an eye in the applications. Hence, we provide explicit estimates in the conditions of our theorems, so that the hypotheses can be checked using the computer assisted methodology in [27]. In particular, another novelty of this paper is that estimates are detailed taking advantage of the presence of additional geometric structures in phase space other than the symplectic structure, such as a Riemannian metric or a compatible triple, covering a gap in the literature [33]. We think researchers interested in applying these techniques to specific problems can benefit from these facts. It is worth mentioning that the required control of the first integrals is limited to estimate the norm of objects that depend only on elementary algebraic expressions and derivatives. Quantitative application of these theorems using computer assisted methods will be presented in a forthcoming work.

The paper is organized as follows. Section 2 introduces the background and the geometric constructions. We present the two main theorems of this paper in Section 3: the ordinary KAM theorem and the generalized iso-energetic KAM theorem (with presence of first integrals). Some common lemmas are given in Section 4, that control the approximation of different geometric properties. The proof of the ordinary KAM theorem is given in Section 5, while the proof of the generalized iso-energetic KAM theorem is done in Section 6. In order to collect the long list of expressions leading to the explicit estimates and conditions of the theorems, we include separate tables in the appendix.

2. BACKGROUND AND ELEMENTARY CONSTRUCTIONS

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

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

We consider the real and imaginary projections $\text{Re}, \text{Im} : \mathbb{C}^m \rightarrow \mathbb{R}^m$, and identify $\mathbb{R}^m \simeq \text{Im}^{-1}\{0\} \subset \mathbb{C}^m$. Given $U \subset \mathbb{R}^m$ and $\rho > 0$, the complex strip of size ρ is

³The invariant tori we consider in this paper are isotropic, but they are not lower dimensional invariant tori in strict meaning (see [6]). The invariant tori are neither partially hyperbolic nor elliptic (see the a-posteriori formulations in [29] and [36]), but partially parabolic tori that appear in families of invariant tori with the same frequency vector.

$U_\rho = \{\theta \in \mathbb{C}^m : \operatorname{Re} \theta \in U, |\operatorname{Im} \theta| < \rho\}$. Given two sets $X, Y \subset \mathbb{C}^m$, $\operatorname{dist}(X, Y)$ 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. We will consider the identifications $\mathbb{R}^m \simeq \mathbb{R}^{m \times 1}$ 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 will use the notation 0_n to represent the column vector $O_{n \times 1}$. 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

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

In particular, if v is an n_2 -dimensional vector, $|Mv| \leq |M||v|$. Moreover, M^\top denotes the transpose of the matrix M , so that

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

Given an analytic function $f : \mathcal{U} \subset \mathbb{C}^m \rightarrow \mathbb{C}$, defined in 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 (column) vectors $v_1, \dots, v_r \in \mathbb{C}^m$, with $v_k = (v_{1k}, \dots, v_{mk})$, is

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

where the indices ℓ_1, \dots, ℓ_r run from 1 to m .

The construction is extended to vector and matrix valued functions as follows: given a matrix valued function $M : \mathcal{U} \subset \mathbb{C}^m \rightarrow \mathbb{C}^{n_1 \times n_2}$ (whose components $M_{i,j}$ are analytic functions), a point $x \in \mathcal{U}$, and a collection of (column) vectors $v_1, \dots, v_r \in \mathbb{C}^m$, we obtain an $n_1 \times n_2$ matrix $D^r M(x)[v_1, \dots, v_r]$ such that

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

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

Notice that, given a function $f : \mathcal{U} \subset \mathbb{C}^m \rightarrow \mathbb{C}^n \simeq \mathbb{C}^{n \times 1}$, we can think of Df as a matrix function $Df : \mathcal{U} \rightarrow \mathbb{C}^{n \times m}$. Hence, $D^1 f(x)[v] = Df(x)v$ for $v \in \mathbb{C}^m$. Therefore, we can apply the transpose to obtain a matrix function $(Df)^\top$, which acts on n -dimensional vectors, while $Df^\top = D(f^\top)$ 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 $u : \mathbb{R}^d \rightarrow \mathbb{R}$ 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 $u : \mathbb{T}^d \rightarrow \mathbb{R}$, where $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ is the d -dimensional standard torus. Analogously, for $\rho > 0$, a function $u : \mathbb{R}_\rho^d \rightarrow \mathbb{C}$ is 1-periodic if $u(\theta + e) = u(\theta)$ for all $\theta \in \mathbb{R}_\rho^d$ and $e \in \mathbb{Z}^d$. We also abuse notation and write $u : \mathbb{T}_\rho^d \rightarrow \mathbb{R}$, where $\mathbb{T}_\rho^d = \{\theta \in \mathbb{C}^d / \mathbb{Z}^d : |\operatorname{Im} \theta| < \rho\}$ is the complex strip of \mathbb{T}^d of width $\rho > 0$. We will also write the Fourier expansion of a periodic function as

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

and introduce the notation $\langle u \rangle := \hat{u}_0$ for the average. The notation in this paragraph is extended to $n_1 \times n_2$ matrix valued periodic functions $M : \mathbb{T}_\rho^d \rightarrow \mathbb{C}^{n_1 \times n_2}$, 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. Hamiltonian systems and invariant tori. In this paper we consider an open set \mathcal{M} of \mathbb{R}^{2n} endowed with a symplectic form ω , that is, a closed ($d\omega = 0$) non-degenerate differential 2-form on \mathcal{M} . We will assume that ω is exact ($\omega = d\alpha$ for certain 1-form α called *action form*), so \mathcal{M} is endowed with an exact symplectic structure. The matrix representations of α and ω are given by the matrix valued functions

$$\begin{aligned} a : \mathcal{M} &\longrightarrow \mathbb{R}^{2n} \\ z &\longmapsto a(z), \end{aligned}$$

and

$$\begin{aligned} \Omega : \mathcal{M} &\longrightarrow \mathbb{R}^{2n \times 2n} \\ z &\longmapsto \Omega(z) = (Da(z))^\top - Da(z), \end{aligned}$$

respectively. The non-degeneracy of ω is equivalent to $\det \Omega(z) \neq 0$ for all $z \in \mathcal{M}$.

Remark 2.1. The prototype example of symplectic structure is the *standard symplectic structure* on $\mathcal{M} \subset \mathbb{R}^{2n}$: $\omega_0 = \sum_{i=1}^n dz_{n+i} \wedge dz_i$. An action form for ω_0 is $\alpha_0 = \sum_{i=1}^n z_{n+i} dz_i$. The matrix representations of α_0 and ω_0 are, respectively,

$$a_0(z) = \begin{pmatrix} O_n & I_n \\ O_n & O_n \end{pmatrix} z, \quad \Omega_0 = \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix}.$$

Another usual action form on $\mathcal{M} \subset \mathbb{R}^{2n}$ is $\alpha_0 = \frac{1}{2} \sum_{i=1}^n (z_{n+i} dz_i - z_i dz_{n+i})$, which is represented as

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

in coordinates.

We say that a vector field $X : \mathcal{M} \rightarrow \mathbb{R}^{2n}$ is *symplectic* (or *locally-Hamiltonian*) if $L_X \omega = 0$ where L_X stands for the Lie derivative with respect to X . Using Cartan's magic formula, and the fact that $d\omega = 0$, it turns out that X is symplectic if and only if $i_X \omega$ is closed. We say that X is *exact symplectic* (or *Hamiltonian*) if the contraction $i_X \omega$ is exact, i.e., if there exists a function $h : \mathcal{M} \rightarrow \mathbb{R}$ (globally defined) such that $i_X \omega = -dh$. In coordinates, an exact symplectic vector field satisfies

$$\Omega(z)X(z) = (Dh(z))^\top, \quad \text{i.e.,} \quad X(z) = \Omega(z)^{-1}(Dh(z))^\top.$$

Hence, we will use the notation $X = X_h$.

The Poisson bracket of two functions f, g is given by $\{f, g\} = -\omega(X_f, X_g)$. In coordinates,

$$\{f, g\}(z) = Df(z)\Omega(z)^{-1}(Dg(z))^\top.$$

Then, if φ_t is the flow of X_h , it follows that

$$\frac{d}{dt}(f \circ \varphi_t) = \{f, h\} \circ \varphi_t,$$

and so, f is a conserved quantity if $\{f, h\} = 0$.

In this context, given a Hamiltonian vector field X_h on \mathcal{M} and a frequency vector $\omega \in \mathbb{R}^d$, with $2 \leq d \leq n$, we are interested in finding a parameterization $K : \mathbb{T}^d \rightarrow \mathcal{M}$ satisfying

$$(2.1) \quad X_h(K(\theta)) = DK(\theta)\omega.$$

This means that the d -dimensional manifold $\mathcal{K} = K(\mathbb{T}^d)$ is invariant and the internal dynamics is given by the constant vector field ω . For obvious reasons, equation (2.1) is called *invariance equation for K* . Therefore, given a parameterization $K : \mathbb{T}^d \rightarrow \mathcal{M}$ and a frequency vector $\omega \in \mathbb{R}^d$, the *error of invariance* is the periodic function $E : \mathbb{T}^d \rightarrow \mathbb{R}^{2n}$ given by

$$(2.2) \quad E(\theta) := X_h(K(\theta)) - DK(\theta)\omega.$$

Roughly speaking, if we have a good enough approximation of a d -dimensional invariant torus \mathcal{K} (that is, the error (2.2) is small enough in certain norm), then one is interested in obtaining a true invariant torus, close to \mathcal{K} , satisfying (2.1).

Remark 2.2. Equation (2.1) is the infinitesimal version of the equation

$$\varphi_t(K(\theta)) = K(\theta + \omega t),$$

where φ_t is the flow of X_h . Accordingly, we can study the invariance of \mathcal{K} using a discrete version of a KAM theorem for symplectic maps. We refer the reader to [20, 27, 33] for such a-posteriori theorems, quantitative estimates, and applications. However, notice that obtaining the required estimates for the flow φ_t demands to integrate the equations of the motion up to second order variational equations.

If $\omega \in \mathbb{R}^d$ is *nonresonant* (i.e. if $\omega \cdot k \neq 0$ for every $k \in \mathbb{Z}^d \setminus \{0\}$) then $z(t) = K(\alpha + \omega t)$ is a quasi-periodic solution of X_h for every $\alpha \in \mathbb{T}^d$, and α is called the initial phase of the parameterization. It is well known that quasi-periodicity implies additional geometric properties of the torus (see [6, 41]). In particular, that the torus is *isotropic*. This means that the pullback $K^*\omega$ of the symplectic form on the torus \mathcal{K} vanishes. In matrix notation, the representation of $K^*\omega$ at a point $\theta \in \mathbb{T}^d$ is

$$(2.3) \quad \Omega_K(\theta) = (DK(\theta))^\top \Omega(K(\theta)) DK(\theta),$$

and so, \mathcal{K} is isotropic if $\Omega_K(\theta) = O_d, \forall \theta \in \mathbb{T}^d$. If $d = n$ then \mathcal{K} is *Lagrangian*. Moreover, quasi-periodicity implies that

$$h(K(\theta)) = \langle h \circ K \rangle, \quad \forall \theta \in \mathbb{T}^d,$$

which means that the torus is contained in an energy level of the Hamiltonian.

Remark 2.3. Other topologies can be considered for the ambient manifold. For example, we may have some information (e.g. from normal form analysis) that allows us to construct tubular coordinates around a torus. In this situation, it is interesting to look for an invariant torus inside an ambient manifold of the form $\mathcal{M} \subset \mathbb{T}^d \times U$, with $U \subset \mathbb{R}^{2n-d}$. Notice that both $a(z)$ and $h(z)$ are 1-periodic in the first d -variables, and, since the Poisson bracket preserves this property, also is X_h . Assume that we are interested in obtaining an invariant torus that preserves the topology of the manifold (typically called *primary torus*). Then we must choose a parameterization $K : \mathbb{T}^d \rightarrow \mathcal{M}$ such that $K(\theta) - (\theta, 0)$ is 1-periodic (that is, K is *homotopic to the zero-section*). Therefore, it turns out that the error function (2.2) is also 1-periodic, so the construction makes sense. All expressions and formulas presented in this paper remain valid, and one only needs to take into account that the elements of K and DK contain an additional term that comes from the topology. We refer to [33] for a detailed discussion of this case. The case of intermediate topologies $\mathcal{M} \subset \mathbb{T}^m \times U$ with $U \subset \mathbb{R}^{2n-m}$ is similar.

2.3. Conserved quantities and families of invariant tori. We will assume that the vector field X_h has $n - d$ first integrals in involution $p_1, \dots, p_{n-d} : \mathcal{M} \rightarrow \mathbb{R}$, that is to say:

$$(2.4) \quad \{h, p_j\} = 0, \quad 1 \leq j \leq n - d,$$

and

$$(2.5) \quad \{p_i, p_j\} = 0, \quad 1 \leq i, j \leq n - d.$$

Consequently, the Lie brackets of the corresponding Hamiltonian vector fields vanish and we have:

$$(2.6) \quad DX_h(z)X_{p_j}(z) = DX_{p_j}(z)X_h(z), \quad 1 \leq j \leq n - d,$$

and

$$(2.7) \quad DX_{p_i}(z)X_{p_j}(z) = DX_{p_j}(z)X_{p_i}(z), \quad 1 \leq i, j \leq n - d.$$

We will encode the $n - d$ first integrals in an only function $p : \mathcal{M} \rightarrow \mathbb{R}^{n-d}$, so that the involution conditions are rephrased as

$$Dh(z) \Omega(z)^{-1} (Dp(z))^\top = 0_{n-d}^\top,$$

and

$$Dp(z) \Omega(z)^{-1} (Dp(z))^\top = O_{n-d}.$$

Moreover, the corresponding $n - d$ Hamiltonian vector fields are the columns of the matrix function $X_p : \mathcal{M} \rightarrow \mathbb{R}^{2n \times (n-d)}$, with $(X_p)_{i,j} = (X_{p_j})_i$, and

$$X_p(z) = \Omega(z)^{-1} (Dp(z))^\top.$$

The commuting conditions are

$$(2.8) \quad DX_h(z)X_p(z) = DX_p(z)[X_h(z)],$$

and

$$(2.9) \quad DX_{p_i}(z)X_p(z) = DX_p(z)[X_{p_i}(z)], \quad 1 \leq i \leq n - d.$$

The above setting implies that p generates an $(n - d)$ -parameter family of local symplectomorphisms. In particular, we introduce

$$(2.10) \quad \Phi_s = \varphi_{s_1}^1 \circ \dots \circ \varphi_{s_{n-d}}^{n-d}$$

where $s = (s_1, \dots, s_{n-d})$ belongs to an open neighborhood of 0 in \mathbb{R}^{n-d} , and $\varphi_{s_i}^i$ is the flow of the Hamiltonian vector field X_{p_i} . Notice that this is a local group action of \mathbb{R}^{n-d} and, by the commutativity of the flows in (2.7), we have

$$\frac{\partial \Phi_s}{\partial s_i} = X_{p_i} \circ \Phi_s.$$

If the vector fields are linearly independent, then the local group action (2.10) defines a family of local diffeomorphisms $s \mapsto \Phi_s$ which commutes with the flow of X_h :

$$\Phi_s \circ \varphi_t = \varphi_t \circ \Phi_s.$$

The map Φ_s is usually called the *continuous family of symmetries* of X_h . A consequence of this is the following: if $\mathcal{K} = K(\mathbb{T}^d)$ is invariant for X_h , with frequency vector ω , then $\mathcal{K}_s = \Phi_s(\mathcal{K})$ is also invariant for X_h with the same frequency vector:

$$\varphi_t(\Phi_s(K(\theta))) = \Phi_s(\varphi_t(K(\theta))) = \Phi_s(K(\theta + \omega t)),$$

and so, $\Phi_s \circ K$ is a parameterization of \mathcal{K}_s .

An important observation is that all invariant tori of the family are contained in the submanifold

$$\{z \in \mathcal{M} : h(z) = h_0, p(z) = p_0\}.$$

Hence, once the frequency ω of the torus has been fixed, we cannot fix the values of h_0 or p_0 . This case is considered in Section 3.2, referred to as the ordinary case. If we are interested in obtaining an invariant torus on a target energy level h_0 , then we fix the direction of the frequency vector ω but adjust its modulus (iso-energetic case). Similarly, the same kind of adjustment of ω can be made if we are interested in obtaining an invariant torus having a prefixed value of one of the components of p (iso-momentum case). The previous scenarios can be generalized by considering any first integral $c : \mathcal{M} \rightarrow \mathbb{R}$ that commutes with h and $p = (p_1, \dots, p_{n-d})$, that is,

$$Dc(z)X_h(z) = 0, \quad Dc(z)X_p(z) = 0_{n-d}^\top.$$

For example, we can think that c is a function of h and p given by $c(z) = f_c(h(z), p(z))$, where $f_c : \mathbb{R} \times \mathbb{R}^{n-d} \rightarrow \mathbb{R}$ is known (we have the particular cases $c(z) = h(z)$ and $c(z) = p_j(z)$ with $j \in \{1, \dots, n-d\}$). In Section 3.3 we will establish sufficient conditions to obtain an invariant torus having a prefixed value of the target conserved quantity c (generalized iso-energetic case). We emphasize that selecting simultaneously the values of several conserved quantities is not possible in general, but makes sense for Cantor sets of frequencies.

2.4. Linearized dynamics and reducibility. In this section we describe the geometric construction of a suitable symplectic frame attached to an invariant torus \mathcal{K} of a Hamiltonian system X_h with conserved quantities $p : \mathcal{M} \rightarrow \mathbb{R}^{n-d}$. Indeed, given a parameterization $K : \mathbb{T}^d \rightarrow \mathcal{M}$ satisfying

$$(2.11) \quad X_h(K(\theta)) = DK(\theta)\omega,$$

and given any m -dimensional vector subbundle parameterized by $V : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times m}$, with $1 \leq m \leq 2n$, we introduce the operator

$$(2.12) \quad \mathcal{X}_V(\theta) := DX(K(\theta))V(\theta) - DV(\theta)[\omega],$$

which corresponds to the infinitesimal displacement of V , and we say that a bundle is invariant under the linearized equations if $\mathcal{X}_V(\theta) = O_{2n \times m}$ for every $\theta \in \mathbb{T}^d$.

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

$$L(\theta) = \begin{pmatrix} DK(\theta) & X_p(K(\theta)) \end{pmatrix},$$

and we will assume that $\text{rank } L(\theta) = n$ for every $\theta \in \mathbb{T}^d$. Then, it turns out that $L(\theta)$ satisfies

$$(2.13) \quad \mathcal{X}_L(\theta) = O_{2n \times n}, \quad \forall \theta \in \mathbb{T}^d.$$

This invariance follows from two observations. Firstly, taking derivatives at both sides of (2.11), we have

$$DX_h(K(\theta))DK(\theta) = D(DK(\theta))[\omega],$$

and secondly, from the commutation rule (2.6), we have:

$$\begin{aligned} D(X_p(K(\theta)))[\omega] &= DX_p(K(\theta))[DK(\theta)\omega] \\ &= DX_p(K(\theta))[X_h(K(\theta))] \\ &= DX_h(K(\theta))X_p(K(\theta)). \end{aligned}$$

Then, the property in (2.13) follows putting together both expressions.

By similar geometric properties (detailed computations will be presented in Section 4.4), it turns out that the subspace $L(\theta)$ is *Lagrangian*, i.e., we have

$$L(\theta)^\top \Omega(K(\theta))L(\theta) = O_{n \times n}$$

for every $\theta \in \mathbb{T}^d$. Then one can use the geometric structure of the problem to complement the above frame, thus obtaining linear coordinates on the full tangent bundle $\mathbb{T}\mathcal{M}$ that express $DX_h \circ K$ in a simple way. This is one of the main ingredients of recent KAM theorems presented in different contexts and structures (see e.g. [9, 20, 29, 31, 36]). The constructions have been summarized in [33] using a common framework that unifies the previous works and emphasizes the role of the symplectic properties. We briefly summarize this framework in Section 2.5.

Hence, we will obtain a map $N : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times n}$ such that the juxtaposed matrix

$$P(\theta) = \begin{pmatrix} L(\theta) & N(\theta) \end{pmatrix} ,$$

satisfies $\text{rank } P(\theta) = 2n$ for every $\theta \in \mathbb{T}^d$ and also

$$(2.14) \quad P(\theta)^\top \Omega(K(\theta))P(\theta) = \Omega_0 .$$

In this case, we say that $P : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times 2n}$ is a symplectic frame. The use of these linear coordinates on $\mathbb{T}_{\mathcal{K}}\mathcal{M}$ has several advantages. Among them, it produces a natural and geometrically meaningful non-degeneracy condition (twist condition) in the KAM theorem, it simplifies certain computations substantially (P^{-1} can be computed directly), but most importantly, it reduces the linearized equation to triangular form as follows:

$$DX_h(K(\theta))P(\theta) - DP(\theta)[\omega] = P(\theta)\Lambda(\theta) ,$$

with

$$\Lambda(\theta) = \begin{pmatrix} O_n & T(\theta) \\ O_n & O_n \end{pmatrix}$$

and

$$\begin{aligned} T(\theta) &= N(\theta)^\top \Omega(K(\theta)) (DX_h(K(\theta))N(\theta) - DN(\theta)[\omega]) \\ &= N(\theta)^\top \Omega(K(\theta)) \mathcal{X}_N(\theta) . \end{aligned}$$

The matrix T is usually called the *torsion matrix* and plays the role of Kolmogorov's non-degeneracy condition.

Remark 2.4. 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, and in fact, of the family of d -dimensional invariant tori.

The above setting allows us to approximate the solutions of the linearized equations around the torus by the solutions of a triangular system which is easier to handle. The fundamental idea is the following fact: if \mathcal{K} is approximately invariant, the above geometrical properties are still satisfied, modulo some error functions which can be controlled in terms of the error of invariance. The main ingredient is the fact that (under certain assumptions) the frame $\theta \mapsto L(\theta)$, associated to an $(n-d)$ -parameter family of approximately invariant tori, is also approximately Lagrangian. Hence, the linear dynamics around the torus is approximately reducible. This is enough to perform a quadratic scheme to correct the initial approximation. This will be discussed in Section 4.

2.5. Construction of a geometrically adapted frame. In this section we deal with the construction of a symplectic frame on the bundle $T_{\mathcal{K}}\mathcal{M}$ by complementing the column vectors of a map $L : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times n}$ that parameterizes a Lagrangian subbundle. To this end, we assume that we have a map $N^0 : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times n}$ such that

$$(2.15) \quad \text{rank}(L(\theta) \ N^0(\theta)) = 2n \quad \Leftrightarrow \quad \det(L(\theta)^\top \Omega(K(\theta))N^0(\theta)) \neq 0,$$

for every $\theta \in \mathbb{T}^d$. Then, we complement the Lagrangian subspace generated by $L(\theta)$ by means of a map $N : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times n}$ given by

$$N(\theta) = L(\theta)A(\theta) + N^0(\theta)B(\theta),$$

where

$$B(\theta) = -(L(\theta)^\top \Omega(K(\theta))N^0(\theta))^{-1}$$

and $A(\theta)$ is a solution of

$$A(\theta) - A(\theta)^\top = -B(\theta)^\top N^0(\theta)^\top \Omega(K(\theta))N^0(\theta)B(\theta).$$

The solution of this equation is given by

$$A(\theta) = -\frac{1}{2}(B(\theta)^\top N^0(\theta)^\top \Omega(K(\theta))N^0(\theta)B(\theta))$$

modulo the addition of any symmetric matrix. A direct computation shows that the juxtaposed matrix $P(\theta) = (L(\theta) \ N(\theta))$ satisfies (2.14).

The map N^0 can be obtained by directly complementing the tangent vectors of the initial (approximately invariant) torus, and after that, be fixed along the iterative procedure. This is called **Case I** in chapter 4 of [33]. It has the advantage of being more general and flexible, but requires some extra work to obtain optimal quantitative estimates (we refer the reader to [27], where additional geometric properties of N^0 are controlled).

A natural way to construct a map N^0 systematically is by using additional geometric information. In this paper, we assume that \mathcal{M} is also endowed with a Riemannian metric \mathbf{g} , represented in coordinates as the positive-definite symmetric matrix valued function $G : \mathcal{M} \rightarrow \mathbb{R}^{2n \times 2n}$. Then, we define the linear isomorphism $\mathbf{J} : T\mathcal{M} \rightarrow T\mathcal{M}$ such that $\omega_z(\mathbf{J}_z u, v) = \mathbf{g}_z(u, v)$, $\forall u, v \in T_z\mathcal{M}$. Observe also that \mathbf{J} is antisymmetric with respect to \mathbf{g} , that is, $\mathbf{g}_z(u, \mathbf{J}_z v) = -\mathbf{g}_z(\mathbf{J}_z u, v)$, $\forall u, v \in T_z\mathcal{M}$. The matrix representation of \mathbf{J} is given by the matrix valued function $J : \mathcal{M} \rightarrow \mathbb{R}^{2n \times 2n}$. Then, we have

$$(2.16) \quad \Omega^\top = -\Omega, \quad G^\top = G, \quad J^\top \Omega = G,$$

and we introduce the matrix valued function $\tilde{\Omega} : \mathcal{M} \rightarrow \mathbb{R}^{2n \times 2n}$, as

$$\tilde{\Omega} := J^\top \Omega J = GJ,$$

for the representation of the symplectic form in the frame given by J .

Then, we choose the map N^0 as follows

$$(2.17) \quad N^0(\theta) := J(K(\theta))L(\theta)$$

and condition (2.15) is equivalent to

$$\det(L(\theta)^\top G(K(\theta))L(\theta)) \neq 0 \quad \forall \theta \in \mathbb{T}^d.$$

Moreover, the matrices $A(\theta)$ and $B(\theta)$ are expressed as follows

$$A(\theta) = -\frac{1}{2}(B(\theta)^\top L(\theta)^\top \tilde{\Omega}(K(\theta))L(\theta)B(\theta)),$$

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

This is called **Case II** in [33].

Remark 2.5. We want to point out that the above construction differs slightly from the discussion in chapter 4 of [33], where the linear isomorphism \mathbf{J} is defined as $\omega_z(u, v) = \mathbf{g}_z(u, \mathbf{J}_z v)$. This choice results in the map

$$N^0(\theta) = -J(K(\theta))^{-1}L(\theta)$$

rather than (2.17). Both constructions are equivalent, but the construction described here is geometrically more natural and produces better quantitative estimates.

There is also the special case where the isomorphism \mathbf{J} is anti-involutive, that is, $\mathbf{J}^2 = -\mathbf{I}$. Then, we say that the triple $(\omega, \mathbf{g}, \mathbf{J})$ is compatible and that \mathbf{J} endows \mathcal{M} with a complex structure. This is called **Case III** in chapter 4 of [33]. In coordinates, we have the following properties

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

In this situation, we have

$$N^0(\theta) = J(K(\theta))L(\theta), \quad A(\theta) = O_n, \quad B(\theta) = (L(\theta)^\top G(K(\theta))L(\theta))^{-1}.$$

It is important to notice that the above constructions lead to different quantitative estimates in the KAM theorem. Selecting the best option depends on the particular problem under consideration. Since **Case I** has been fully reported in the references [27, 33], in this paper we will focus in obtaining sharp quantitative estimates for **Case II** and **Case III**. Hence, we cover a gap in the literature that could be valuable in future studies.

2.6. Univocal determination of an invariant torus of the family. In this section we describe suitable strategies to avoid the undeterminations observed in Sections 2.2 and 2.3. Let us recapitulate them:

- If $\mathcal{K} = K(\mathbb{T}^d)$ is a d -dimensional invariant torus of X_h of frequency ω , then $K^\alpha(\theta) = K(\theta + \alpha)$ also parameterizes \mathcal{K} for every $\alpha \in \mathbb{T}^d$.
- If $\mathcal{K} = K(\mathbb{T}^d)$ is a d -dimensional invariant torus of X_h of frequency ω , and we introduce $K_s = \Phi_s \circ K$ using the family of symmetries, then $\mathcal{K}_s = K_s(\mathbb{T}^d)$ is also an invariant torus of frequency ω for every $s \in \mathbb{R}^{n-d}$ in the domain of definition.

The first indetermination corresponds to the choice of the parameterization of the invariant object, and it can be avoided simply by fixing an initial phase of the torus. To this end, we consider a $(2n - d)$ -dimensional manifold given by the preimage of a map $Z : \mathbb{R}^{2n} \rightarrow \mathbb{R}^d$ and select the value of α such that

$$Z(K^\alpha(0)) = Z_0,$$

for some $Z_0 \in \mathbb{R}^d$. Notice that Z must be selected in such a way that the transversality condition

$$\det(DZ(K^\alpha(0))DK^\alpha(0)) \neq 0,$$

holds in an open set of values $\alpha \in \mathbb{R}^d$.

Remark 2.6. If $\mathcal{M} = \mathbb{T}^d \times U$ and we are considering a non-contractible invariant tori of the form $K(\theta) = (K^x(\theta), K^y(\theta)) \in \mathcal{M}$, then a typical way to determine the phase univocally is to ask the following average condition

$$\langle K^x - \text{id} \rangle = 0_d.$$

In this case, we select $\alpha = -\langle K^x - \text{id} \rangle$. Another possibility, in the spirit described above, is to select the transversal plane given by $Z(z) = (z_1, \dots, z_d)$.

The second indetermination corresponds to a choice of a given invariant torus inside the $(n-d)$ -parameter family described in Section 2.3. In this case, we need to fix additional $(n-d)$ conditions in order to define univocally a single torus of the family. For example, we may assume that there is a map $q : \mathcal{M} \rightarrow \mathbb{R}^{n-d}$ satisfying

$$(2.18) \quad Dq(z)X_p(z) = Dq(z)\Omega(z)^{-1}(Dp(z))^\top = I_{n-d},$$

which means that $\{q_i, p_j\} = \delta_{ij}$. For obvious reasons, p and q are referred to as the *generalized momentum* and the *generalized conjugated position*, respectively. Then, we can determine univocally a torus in the family by asking for the extra equations

$$q \circ K(0) = q_0 \in \mathbb{R}^{n-d}.$$

We can recover the full family by considering the map

$$s \mapsto q \circ \Phi_s \circ K = q \circ K + s,$$

where we used that

$$\begin{aligned} \frac{\partial}{\partial s_i}(q \circ \Phi_s \circ K) &= (Dq \circ \Phi_s \circ K)\partial_{s_i}(\Phi_s \circ K) \\ &= (Dq \circ \Phi_s \circ K)(X_{p_i} \circ \Phi_s \circ K) = e_i. \end{aligned}$$

Remark 2.7. The above construction can be readily generalized asking the map q to satisfy

$$\det(Dq(z)X_p(z)) \neq 0$$

instead of (2.18).

It is clear that both indeterminations can be addressed simultaneously by fixing n conditions. This can be done for example by asking for a transversality condition on the Lagrangian frame $\theta \mapsto L(\theta)$ described in Section 2.4 at a given point. To this end, we denote

$$K_{\alpha,s}(\theta) = \Phi_s(K(\theta + \alpha)), \quad \alpha \in \mathbb{R}^d, \quad s \in \mathbb{R}^{n-d},$$

consider a map $Q : \mathcal{M} \rightarrow \mathbb{R}^n$, and ask for the condition

$$Q(K_{\alpha,s}(0)) = Q_0$$

for a given point $Q_0 \in \mathbb{R}^n$. In this situation, the transversality condition reads

$$\det(DQ(K_{\alpha,s}(0))L_{\alpha,s}(0)) \neq 0$$

where

$$\theta \mapsto L_{\alpha,s}(\theta) = (DK_{\alpha,s}(\theta) \quad X_p(K_{\alpha,s}(\theta)))$$

is the Lagrangian frame associated with the torus $\mathcal{K}_{\alpha,s}$. For example, a natural choice would be

$$Q(z) = \begin{pmatrix} Z(z) \\ q(z) \end{pmatrix},$$

where Z is selected to fix the phase of the parameterizations and q are generalized positions associated with p . Depending on the topology of the ambient space, we may consider other choices (see Remark 2.6).

3. A-POSTERIORI KAM THEORY FOR PARTIALLY INTEGRABLE HAMILTONIAN SYSTEMS

In this section, we present two a-posteriori KAM theorems for d -dimensional quasi-periodic invariant tori in Hamiltonian systems with n degrees-of-freedom that have $n - d$ additional first integrals in involution. To this end, we will assume that the frequency vector ω satisfies Diophantine conditions. Specifically, we denote the set of Diophantine vectors as

$$(3.1) \quad \mathcal{D}_{\gamma,\tau} = \left\{ \omega \in \mathbb{R}^d : |k \cdot \omega| \geq \frac{\gamma}{|k|_1^\tau}, \forall k \in \mathbb{Z}^d \setminus \{0\}, |k|_1 = \sum_{i=1}^d |k_i| \right\},$$

for certain $\gamma > 0$ and $\tau \geq d - 1$.

In Section 3.1 we set some basic notation regarding Banach spaces and norms of analytic functions. In Section 3.2 we present the statement of a KAM theorem for existence (and persistence) of d -dimensional invariant tori having fixed frequency vector $\omega \in \mathcal{D}_{\gamma,\tau}$. This corresponds to the so-called ordinary (à la Kolmogorov) KAM theorem. In Section 3.3 we present an adapted version of the theorem that generalizes the iso-energetic approach. Section 4 is devoted to the control of approximate geometric properties, the anteroom of the proofs of the main theorems in Sections 5 and 6. We will pay special attention in providing explicit and rather optimal bounds, with an eye in the application of the theorems and in computer assisted proofs. The constants have been collected in a series of tables in Appendix B.

3.1. Analytic functions and norms. In this paper we work with real analytic functions defined in complex neighborhoods of real domains. We will consider the sup-norms of (matrix valued) analytic functions and their derivatives (see the notation in Section 2.1). That is, for $f : \mathcal{U} \subset \mathbb{C}^m \rightarrow \mathbb{C}$, we consider

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

and

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

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

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

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

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

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

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

is also analytic, and we have

$$\begin{aligned}
\|D^r M[v_1, \dots, v_r]\|_{\mathcal{U}} &\leq \max_{i=1, \dots, n_1} \sum_{j=1}^{n_2} \|D^r M_{i,j}[v_1, \dots, v_r]\|_{\mathcal{U}} \\
&\leq \max_{i=1, \dots, n_1} \sum_{j=1}^{n_2} \left\| \sum_{\ell_1, \dots, \ell_r} \frac{\partial^r M_{i,j}}{\partial x_{\ell_1} \cdots \partial x_{\ell_r}} v_{\ell_1,1} \cdots v_{\ell_r,r} \right\|_{\mathcal{U}} \\
&\leq \max_{i=1, \dots, n_1} \sum_{j=1}^{n_2} \sum_{\ell_1, \dots, \ell_r} \left\| \frac{\partial^r M_{i,j}}{\partial x_{\ell_1} \cdots \partial x_{\ell_r}} \right\|_{\mathcal{U}} \max_{\ell=1, \dots, r} \|v_{\ell,1}\|_{\mathcal{U}} \cdots \max_{\ell=1, \dots, r} \|v_{\ell,r}\|_{\mathcal{U}} \\
&= \|D^r M\|_{\mathcal{U}} \|v_1\|_{\mathcal{U}} \cdots \|v_r\|_{\mathcal{U}}.
\end{aligned}$$

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

$$\begin{aligned}
\|(D^r M[v_1, \dots, v_r])^\top\|_{\mathcal{U}} &\leq \max_{j=1, \dots, n_2} \sum_{i=1}^{n_1} \|D^r M_{i,j}[v_1, \dots, v_r]\|_{\mathcal{U}} \\
&\leq \|D^r M^\top\|_{\mathcal{U}} \|v_1\|_{\mathcal{U}} \cdots \|v_r\|_{\mathcal{U}}.
\end{aligned}$$

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

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

and

$$\|D(M_1 M_2)\|_{\mathcal{U}} \leq \|DM_1\|_{\mathcal{U}} \|M_2\|_{\mathcal{U}} + \|M_1\|_{\mathcal{U}} \|DM_2\|_{\mathcal{U}}.$$

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

$$\|u\|_\rho = \|u\|_{\mathbb{T}_\rho^d} = \max_{|\operatorname{Im}\theta| \leq \rho} |u(\theta)|.$$

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

$$\left\| \frac{\partial u}{\partial x_\ell} \right\|_{\rho-\delta} \leq \frac{1}{\delta} \|u\|_\rho, \quad \|Du\|_{\rho-\delta} \leq \frac{d}{\delta} \|u\|_\rho, \quad \|(Du)^\top\|_{\rho-\delta} \leq \frac{1}{\delta} \|u\|_\rho.$$

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

$$\|DM\|_{\rho-\delta} = \max_{i=1, \dots, n_1} \sum_{j=1, \dots, n_2} \|DM_{i,j}\|_{\rho-\delta} \leq \frac{d}{\delta} \|M\|_\rho.$$

A direct consequence is that $\|DM^\top\|_{\rho-\delta} \leq \frac{d}{\delta} \|M^\top\|_\rho$.

As it was mentioned in Section 2.1, the operators D and $(\cdot)^\top$ do not commute. In particular, given a real analytic vector function $w : \mathbb{T}_\rho^d \rightarrow \mathbb{C}^n \simeq \mathbb{C}^{n \times 1}$, we have:

$$\|Dw\|_{\rho-\delta} \leq \frac{d}{\delta} \|w\|_\rho, \quad \|Dw^\top\|_{\rho-\delta} \leq \frac{d}{\delta} \|w^\top\|_\rho \leq \frac{nd}{\delta} \|w\|_\rho, \quad \|(Dw)^\top\|_{\rho-\delta} \leq \frac{n}{\delta} \|w\|_\rho.$$

3.2. Ordinary KAM theorem. At this point, we are ready to state sufficient conditions to guarantee the existence of a d -dimensional invariant torus with fixed frequency close to an approximately invariant one. Notice that the hypotheses in Theorem 3.1 are tailored to be verified with a finite amount of computations.

The result is written simultaneously for **Case II** and **Case III** (see Section 2.5). Estimates corresponding to **Case I** can be easily obtained without any remarkable difficulty (see e.g. [27] for details). Hence, given a parameterization of a torus $\mathcal{K} = K(\mathbb{T}^d)$ (not necessarily invariant) and a tangent frame $L : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times n}$, the normal frame $N : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times n}$ is constructed as follows:

$$(3.2) \quad N(\theta) := L(\theta)A(\theta) + N^0(\theta)B(\theta),$$

where

$$(3.3) \quad N^0(\theta) = J(K(\theta))L(\theta),$$

$$(3.4) \quad B(\theta) = (L(\theta)^\top G(K(\theta))L(\theta))^{-1},$$

$$(3.5) \quad A(\theta) = \begin{cases} -\frac{1}{2}(B(\theta)^\top L(\theta)^\top \tilde{\Omega}(K(\theta))L(\theta)B(\theta)), & \text{if Case II,} \\ 0, & \text{if Case III.} \end{cases}$$

The torsion matrix $T : \mathbb{T}^d \rightarrow \mathbb{R}^{n \times n}$, given by

$$(3.6) \quad T(\theta) = N(\theta)^\top \Omega(K(\theta))\mathcal{X}_N(\theta),$$

where

$$(3.7) \quad \mathcal{X}_N(\theta) = DX_h(K(\theta))N(\theta) - DN(\theta)[\omega],$$

measures the infinitesimal twist of the normal bundle. With this geometric ingredients we are ready to state our main theorem, in the ordinary case.

Theorem 3.1 (KAM theorem with first integrals). *Let us consider an exact symplectic structure $\omega = d\alpha$ and a Riemannian metric g on the open set $\mathcal{M} \subset \mathbb{R}^{2n}$. Let h be a Hamiltonian function, having $n - d$ first integrals in involution $p = (p_1, \dots, p_{n-d})$, with $2 \leq d \leq n$, and let c be any first integral in involution with (h, p) . Let $K : \mathbb{T}^d \rightarrow \mathcal{M}$ be a parameterization of an approximately invariant torus with frequency vector $\omega \in \mathbb{R}^d$, and consider the tangent frame $L : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times n}$ given by*

$$(3.8) \quad L(\theta) = (DK(\theta) \quad X_p(K(\theta))).$$

Then, we assume that the following hypotheses hold.

H_1 *The global objects can be analytically extended to the complex domain $\mathcal{B} \subset \mathbb{C}^{2n}$, and there are constants that quantify the control of their analytic norms.*

For the geometric structures $\omega, g, J, \tilde{\omega} = J^\omega$ in \mathcal{B} , the matrix functions $\Omega, G, J, \tilde{\Omega} : \mathcal{B} \rightarrow \mathbb{C}^{2n \times 2n}$ satisfy:*

$$\begin{aligned} \|\Omega\|_{\mathcal{B}} &\leq c_{\Omega,0}, & \|\mathrm{D}\Omega\|_{\mathcal{B}} &\leq c_{\Omega,1}, \\ \|\tilde{\Omega}\|_{\mathcal{B}} &\leq c_{\tilde{\Omega},0}, & \|\mathrm{D}\tilde{\Omega}\|_{\mathcal{B}} &\leq c_{\tilde{\Omega},1}, & \|\mathrm{D}^2\tilde{\Omega}\|_{\mathcal{B}} &\leq c_{\tilde{\Omega},2}, \\ \|G\|_{\mathcal{B}} &\leq c_{G,0}, & \|\mathrm{D}G\|_{\mathcal{B}} &\leq c_{G,1}, & \|\mathrm{D}^2G\|_{\mathcal{B}} &\leq c_{G,2}, \\ \|J\|_{\mathcal{B}} &\leq c_{J,0}, & \|\mathrm{D}J\|_{\mathcal{B}} &\leq c_{J,1}, & \|\mathrm{D}^2J\|_{\mathcal{B}} &\leq c_{J,2}, \\ \|J^\top\|_{\mathcal{B}} &\leq c_{J^\top,0}, & \|\mathrm{D}J^\top\|_{\mathcal{B}} &\leq c_{J^\top,1}. \end{aligned}$$

For the Hamiltonian $h : \mathcal{B} \rightarrow \mathbb{C}$ and its corresponding vector field $X_h : \mathcal{B} \rightarrow \mathbb{C}^{2n}$, we have :

$$\begin{aligned} \|Dh\|_{\mathcal{B}} &\leq c_{h,1}, \\ \|X_h\|_{\mathcal{B}} &\leq c_{X_h,0}, & \|DX_h\|_{\mathcal{B}} &\leq c_{X_h,1}, \\ \|D^2X_h\|_{\mathcal{B}} &\leq c_{X_h,2}, & \|DX_h^\top\|_{\mathcal{B}} &\leq c_{X_h^\top,1}. \end{aligned}$$

For the first integrals $p : \mathcal{B} \rightarrow \mathbb{C}^{n-d}$ and the corresponding vector fields $X_p : \mathcal{B} \rightarrow \mathbb{C}^{2n \times (n-d)}$, we have:

$$\begin{aligned} \|Dp\|_{\mathcal{B}} &\leq c_{p,1}, & \|Dp^\top\|_{\mathcal{B}} &\leq c_{p^\top,1}, \\ \|X_p\|_{\mathcal{B}} &\leq c_{X_p,0}, & \|DX_p\|_{\mathcal{B}} &\leq c_{X_p,1}, & \|D^2X_p\|_{\mathcal{B}} &\leq c_{X_p,2}, \\ \|X_p^\top\|_{\mathcal{B}} &\leq c_{X_p^\top,0}, & \|DX_p^\top\|_{\mathcal{B}} &\leq c_{X_p^\top,1}, & \|D^2X_p^\top\|_{\mathcal{B}} &\leq c_{X_p^\top,2}. \end{aligned}$$

For the first integral $c : \mathcal{B} \rightarrow \mathbb{C}$ we have

$$\|Dc\|_{\mathcal{B}} \leq c_{c,1}.$$

H_2 The parameterization K is real analytic in a complex strip \mathbb{T}_ρ^d , with $\rho > 0$, which is contained in the global domain:

$$\text{dist}(K(\mathbb{T}_\rho^d), \partial\mathcal{B}) > 0.$$

Moreover, the components of K and DK belong to $\mathcal{A}(\mathbb{T}_\rho^d)$, and there are constants σ_K and σ_{K^\top} such that

$$\|DK\|_\rho < \sigma_K, \quad \|(DK)^\top\|_\rho < \sigma_{K^\top}.$$

H_3 We assume that $L(\theta)$ given by (3.8) has maximum rank for every $\theta \in \bar{\mathbb{T}}_\rho^d$. Moreover, there exists a constant σ_B such that

$$\|B\|_\rho < \sigma_B,$$

where $B(\theta)$ is given by (3.4).

H_4 There exists a constant σ_T such that

$$|\langle T \rangle^{-1}| < \sigma_T,$$

where $T(\theta)$ is given by (3.6).

H_5 The frequency ω belongs to $\mathcal{D}_{\gamma,\tau}$, given by (3.1), for certain $\gamma > 0$ and $\tau \geq d-1$.

Under the above hypotheses, for each $0 < \rho_\infty < \rho$ there exists a constant \mathfrak{C}_1 such that, if the error of invariance

$$(3.9) \quad E(\theta) = X_h(K(\theta)) - DK(\theta)\omega,$$

satisfies

$$(3.10) \quad \frac{\mathfrak{C}_1 \|E\|_\rho}{\gamma^4 \rho^{4\tau}} < 1,$$

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

$$\|DK_\infty\|_{\rho_\infty} < \sigma_K, \quad \|(DK_\infty)^\top\|_{\rho_\infty} < \sigma_{K^\top}, \quad \text{dist}(K_\infty(\mathbb{T}_{\rho_\infty}^d), \partial\mathcal{B}) > 0.$$

Furthermore, the objects are close to the original ones: there exist constants \mathfrak{C}_2 and \mathfrak{C}_3 such that

$$(3.11) \quad \|K_\infty - K\|_{\rho_\infty} < \frac{\mathfrak{C}_2 \|E\|_\rho}{\gamma^2 \rho^{2\tau}}, \quad |\langle c \circ K_\infty \rangle - \langle c \circ K \rangle| < \frac{\mathfrak{C}_3 \|E\|_\rho}{\gamma^2 \rho^{2\tau}}.$$

The constants \mathfrak{C}_1 , \mathfrak{C}_2 and \mathfrak{C}_3 are given explicitly in Appendix B.

Remark 3.2. 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 $c_{p,1} = 0$, $c_{p^\top,1} = 0$, $c_{X_p,0} = 0$, $c_{X_p^\top,0} = 0$, $c_{X_p,1} = 0$, $c_{X_p^\top,1} = 0$, $c_{X_p,2} = 0$, and $c_{X_p^\top,2} = 0$. Thus, as a by-product, we obtain optimal quantitative estimates for the KAM theorem for flows stated in [20].

Remark 3.3. In the *canonical case*, we have $\Omega = \tilde{\Omega} = \Omega_0$, $G = I_{2n}$, and $J = \Omega_0$. Hence, we have $c_{\Omega,0} = 1$, $c_{\Omega,1} = 0$, $c_{\tilde{\Omega},0} = 1$, $c_{\tilde{\Omega},1} = 0$, $c_{\tilde{\Omega},2} = 0$, $c_{G,0} = 1$, $c_{G,1} = 0$, $c_{G,2} = 0$, $c_{J,0} = 1$, $c_{J,1} = 0$, $c_{J,2} = 0$, $c_{J^\top,0} = 1$, and $c_{J^\top,1} = 0$.

Remark 3.4. Notice that the condition $d \geq 2$ is optimal. For $d = 1$, not only the torus becomes a periodic orbit (and the result would follow from a standard implicit theorem without small divisors) but also X_h is completely integrable.

Remark 3.5. The existence of a d -dimensional invariant torus with frequency ω implies the existence of an $(n-d)$ -parameter family of invariant tori with frequency ω . The family is locally unique, meaning that if there is an invariant torus with frequency ω close enough to the family, then it is a member of the family. Notice also that Theorem 3.1 states the existence of the parameterization of an invariant torus, but that we can also change the phase to obtain a new parameterization. As mentioned in Section 2.6, both indeterminacies (the phase and the element of the family) could be fixed by adding n extra scalar equations to the invariance equation.

3.3. Generalized iso-energetic KAM theorem. Let us consider the setting presented in Section 3.2, and let us focus on the first integral $c : \mathcal{M} \rightarrow \mathbb{R}$ that commutes with h and $p = (p_1, \dots, p_{n-d})$. It is clear that if $\mathcal{K} = K(\mathbb{T}^d)$ is invariant under X_h then \mathcal{K} is contained in the hypersurface $c(z) = c_0$, for $c_0 \in \mathbb{R}$. In this section we are interested in finding an invariant torus by pre-fixing such hypersurface, that is, our aim is to obtain a parameterization satisfying

$$(3.12) \quad X_h(K(\theta)) = \mathfrak{L}_\omega K(\theta), \quad \langle c \circ K \rangle = c_0,$$

where $c_0 \in \mathbb{R}$ is fixed and we think in $\omega \in \mathbb{P}\mathbb{R}^d$.

The following result, which is an extension of Theorem 3.1 to this generalized iso-energetic context, establishes quantitative sufficient conditions for the existence of a solution of (3.12) close to an approximate one. For this reason, we refer the reader to Section 3.2 for a compendium of the objects involved in the result and we do not state the common hypotheses.

Theorem 3.6 (Iso-energetic KAM theorem with first integrals). *Let us consider the setting of Theorem 3.1, assume that the hypotheses H_2 and H_3 hold, and replace H_1 , H_4 and H_5 by*

H'_1 *Assume that all estimates in H_1 hold. In addition, for the first integral $c : \mathcal{B} \rightarrow \mathbb{C}$ we have*

$$\|D^2 c\|_{\mathcal{B}} \leq c_{c,2}.$$

H'_4 There exists a constant σ_{T_c} such that

$$|\langle T_c \rangle^{-1}| < \sigma_{T_c},$$

where $T_c : \mathbb{T}_\rho^d \rightarrow \mathbb{C}^{(n+1) \times (n+1)}$ is the extended torsion matrix

$$(3.13) \quad T_c(\theta) := \begin{pmatrix} T(\theta) & \hat{\omega} \\ \text{D}c(K(\theta))N(\theta) & 0 \end{pmatrix}, \quad \hat{\omega} := \begin{pmatrix} \omega \\ 0_{n-d} \end{pmatrix}.$$

H'_5 Let us consider a constant $\sigma_\omega > 1$ and a frequency vector ω_* in the set $\mathcal{D}_{\gamma,\tau}$, given by (3.1), for certain $\gamma > 0$ and $\tau \geq d-1$. Then, we assume that $\omega \in \mathbb{R}^d$ is contained in the ray

$$\Theta = \Theta(\omega_*, \sigma_\omega) = \{s\omega_* \in \mathbb{R}^d : 1 < s < \sigma_\omega\} \subset \mathcal{D}_{\gamma,\tau}.$$

Notice that, by definition, we have $\text{dist}(\omega, \partial\Theta) > 0$.

Under the above hypotheses, for each $0 < \rho_\infty < \rho$ there exists a constant \mathfrak{C}_1 such that, if the total error

$$(3.14) \quad E_c(\theta) = \begin{pmatrix} E(\theta) \\ E^\omega \end{pmatrix} = \begin{pmatrix} X_h(K(\theta)) - \text{D}K(\theta)\omega \\ \langle c \circ K \rangle - c_0 \end{pmatrix}$$

satisfies

$$(3.15) \quad \frac{\mathfrak{C}_1 \|E_c\|_\rho}{\gamma^4 \rho^{4\tau}} < 1,$$

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

$$\|\text{D}K_\infty\|_{\rho_\infty} < \sigma_K, \quad \|(\text{D}K_\infty)^\top\|_{\rho_\infty} < \sigma_{K^\top}, \quad \text{dist}(K_\infty(\mathbb{T}_{\rho_\infty}^d), \partial\mathcal{B}) > 0.$$

Furthermore, the objects are close to the initial ones: there exist constants \mathfrak{C}_2 and \mathfrak{C}_3 such that

$$(3.16) \quad \|K_\infty - K\|_{\rho_\infty} < \frac{\mathfrak{C}_2 \|E_c\|_\rho}{\gamma^2 \rho^{2\tau}}, \quad |\omega_\infty - \omega| < \frac{\mathfrak{C}_3 \|E_c\|_\rho}{\gamma^2 \rho^{2\tau}}.$$

The constants \mathfrak{C}_1 , \mathfrak{C}_2 and \mathfrak{C}_3 are given explicitly in Appendix B.

Remark 3.7. If we consider the case $c(z) = h(z)$ then we recover the classical iso-energetic situation. Notice that, if \mathcal{K} is invariant with frequency ω , then the frame $P(\theta)$ is symplectic, and

$$\begin{aligned} \text{D}h(K(\theta))N(\theta) &= X_h(K(\theta))^\top \Omega(K(\theta))N(\theta) = -\omega^\top \text{D}K(\theta)^\top \Omega(K(\theta))N(\theta) \\ &= (\omega^\top \quad 0_{n-d}^\top) = \hat{\omega}^\top. \end{aligned}$$

Hence, the extended torsion matrix for an invariant torus is

$$(3.17) \quad T_h(\theta) := \begin{pmatrix} T(\theta) & \hat{\omega} \\ \hat{\omega}^\top & 0 \end{pmatrix}.$$

Remark 3.8. If we consider the case $c(z) = p_j(z)$ then we have an iso-momentum situation. In this case, if \mathcal{K} is invariant with frequency ω , then

$$\text{D}p_j(K(\theta))N(\theta) = -X_{p_j}(K(\theta))^\top \Omega(K(\theta))N(\theta) = e_{d+j}^\top,$$

where e_{d+j} is the $(d+j)$ -th canonic vector of \mathbb{R}^n (it has 1 in the $(d+j)$ -th component and 0 elsewhere). Hence, the extended torsion matrix for an invariant torus is

$$(3.18) \quad T_{p_j}(\theta) := \begin{pmatrix} T(\theta) & \hat{\omega} \\ e_{d+j}^\top & 0 \end{pmatrix}.$$

4. SOME LEMMAS TO CONTROL APPROXIMATE GEOMETRIC PROPERTIES

In this section we present some estimates regarding the control of some geometric properties for an approximately invariant torus. For the sake of clarity, we reduce the repetition of hypotheses and present a unique setting for the whole section, consisting in the assumptions of the KAM Theorems in Section 3.2 and 3.3.

4.1. Estimates for cohomological equations. Let us first introduce some useful notation regarding the so-called *cohomological equations* that play an important role in KAM theory. Given $\omega \in \mathbb{R}^d$ and a periodic function v , we consider the cohomological equation

$$(4.1) \quad \mathfrak{L}_\omega u = v - \langle v \rangle, \quad \mathfrak{L}_\omega := - \sum_{i=1}^d \omega_i \frac{\partial}{\partial \theta_i}.$$

The notation \mathfrak{L} comes from “left-operator”.

Let us assume that v is continuous and ω is rationally independent (this implies that the flow $t \mapsto \omega t$ is quasi-periodic). If there exists a continuous zero-average solution of equation (4.1), then it is unique and will be denoted by $u = \mathfrak{R}_\omega v$. The notation \mathfrak{R} comes from “right-operator”.

Note that the formal solution of equation (4.1) is immediate. Actually, if v has the Fourier expansion $v(\theta) = \sum_{k \in \mathbb{Z}^d} \hat{v}_k e^{2\pi i k \cdot \theta}$ and the dynamics is quasi-periodic, then

$$(4.2) \quad \mathfrak{R}_\omega v(\theta) = \sum_{k \in \mathbb{Z}^d \setminus \{0\}} \hat{u}_k e^{2\pi i k \cdot \theta}, \quad \hat{u}_k = \frac{-\hat{v}_k}{2\pi i k \cdot \omega}.$$

In particular, this implies that $\mathfrak{R}_\omega v = 0$ if $v = 0$. The solutions of equation (4.1) differ by their average.

We point out that quasi-periodicity is not enough to ensure regularity of the solutions of cohomological equations. This is related to the effect of the small divisors $k \cdot \omega$ in equation (4.2). To deal with regularity, we require stronger non-resonant conditions on the vector of frequencies. In this paper, we consider the classic Diophantine conditions in H'_5 and H'_5 .

Lemma 4.1 (Rüssmann estimates). *Let $\omega \in \mathcal{D}_{\gamma, \tau}$ for some $\gamma > 0$ and $\tau \geq d - 1$. Then, for any $v \in \mathcal{A}(\mathbb{T}_\rho^d)$, with $\rho > 0$, there exists a unique zero-average solution of $\mathfrak{L}_\omega u = v - \langle v \rangle$, denoted by $u = \mathfrak{R}_\omega v$. Moreover, for any $0 < \delta < \rho$ we have that $u \in \mathcal{A}(\mathbb{T}_{\rho-\delta}^d)$ and the estimate*

$$\|u\|_{\rho-\delta} \leq \frac{c_R}{\gamma \delta^\tau} \|v\|_\rho,$$

where c_R is a constant that depends on d , τ and possibly on δ .

Proof. There is no need to reproduce here this classical result, and we refer the reader to the original references [46, 48], where a uniform bound (independent of δ) is obtained. We refer to [27] for sharp non-uniform computer-assisted estimates (in the discrete case) of the form $c_R = c_R(\delta)$, representing a substantial advantage in order to apply the result to particular problems. Adapting these estimates to the continuous case is straightforward. Also, we refer to [28] for a numerical quantification of these estimates and for an analysis of the different sources of over-estimation. \square

4.2. Approximate conserved quantities. If $\mathcal{K} = K(\mathbb{T}^d)$ is not an invariant torus with frequency ω , it is clear that a first integral in involution c (such as the energy h or any of the components of p) is not necessarily preserved along $z(t) = K(\theta_0 + \omega t)$, since this is not a true trajectory. However, we can “shadow” its evolution in terms of the error of invariance.

Lemma 4.2. *Let us consider the setting of Theorem 3.1 or Theorem 3.6. Then, for a conserved quantity c the following estimates hold:*

$$(4.3) \quad \|c \circ K - \langle c \circ K \rangle\|_{\rho-\delta} \leq \frac{c_{Rc,1}}{\gamma\delta^\tau} \|E\|_\rho,$$

$$(4.4) \quad \|\mathfrak{L}_\omega(D(c \circ K))\|_{\rho-\delta} \leq \frac{dc_{c,1}}{\delta} \|E\|_\rho, \quad \|\mathfrak{L}_\omega(D(c \circ K))^\top\|_{\rho-\delta} \leq \frac{c_{c,1}}{\delta} \|E\|_\rho,$$

$$(4.5) \quad \|D(c \circ K)\|_{\rho-2\delta} \leq \frac{c_{Rdc,1}}{\gamma\delta^{\tau+1}} \|E\|_\rho, \quad \|(D(c \circ K))^\top\|_{\rho-2\delta} \leq \frac{c_{Rc,1}}{\gamma\delta^{\tau+1}} \|E\|_\rho.$$

In particular,

$$(4.6) \quad \|p \circ K - \langle p \circ K \rangle\|_{\rho-\delta} \leq \frac{c_{Rp,1}}{\gamma\delta^\tau} \|E\|_\rho,$$

$$(4.7) \quad \|\mathfrak{L}_\omega(D(p \circ K))\|_{\rho-\delta} \leq \frac{dc_{p,1}}{\delta} \|E\|_\rho, \quad \|\mathfrak{L}_\omega(D(p \circ K))^\top\|_{\rho-\delta} \leq \frac{c_{p,1}}{\delta} \|E\|_\rho,$$

$$(4.8) \quad \|D(p \circ K)\|_{\rho-2\delta} \leq \frac{c_{Rdp,1}}{\gamma\delta^{\tau+1}} \|E\|_\rho, \quad \|(D(p \circ K))^\top\|_{\rho-2\delta} \leq \frac{c_{Rp,1}}{\gamma\delta^{\tau+1}} \|E\|_\rho.$$

Proof. We will prove the result first for the conserved quantity c . This case includes analogous estimates for each of the first integrals h and p_i , with $i = 1, \dots, n-d$. We apply $\mathfrak{L}_\omega(\cdot) = -D(\cdot)\omega$ in the expression $c \circ K$, thus obtaining

$$(4.9) \quad \begin{aligned} \mathfrak{L}_\omega(c(K(\theta))) &= Dc(K(\theta))\mathfrak{L}_\omega K(\theta) \\ &= Dc(K(\theta))(E(\theta) - X_h(K(\theta))) \\ &= Dc(K(\theta))E(\theta). \end{aligned}$$

In the second line we used the expression of the error of invariance (3.9), and in the third line we used that c is in involution with h . In particular,

$$\|\mathfrak{L}_\omega(c \circ K)\|_\rho \leq c_{c,1} \|E\|_\rho,$$

where we use that $\|Dc\|_{\mathcal{B}} \leq c_{c,1}$. Thus, we end up with

$$c(K(\theta)) - \langle c \circ K \rangle = \mathfrak{R}_\omega(Dc(K(\theta))E(\theta)),$$

and the estimate (4.3) follows applying Lemma 4.1.

In order to prove (4.4) and (4.5) we just differentiate with respect to θ_ℓ , for $\ell = 1, \dots, d$, both formulae (4.9) and (4.4) and apply Cauchy estimates. Firstly,

$$\left\| \frac{\partial}{\partial \theta_\ell} \mathfrak{L}_\omega(c \circ K) \right\|_{\rho-\delta} \leq \frac{c_{c,1}}{\delta} \|E\|_\rho,$$

so estimates in (4.4) follow immediately. Secondly,

$$\left\| \frac{\partial}{\partial \theta_\ell} (c \circ K) \right\|_{\rho-2\delta} \leq \frac{c_{Rc,1}}{\gamma\delta^{\tau+1}} \|E\|_\rho,$$

and then estimates in (4.5) follow.

In order to prove (4.6), (4.7) and (4.8) we just notice that (4.3), (4.4) and (4.5) work for any of the first integrals p_i , for $i = 1, \dots, n - d$. Then, we change the occurrences of c by p_i in the formulae, with $\|Dp_i\|_{\mathcal{B}} \leq c_{p_i,1}$, and use that

$$c_{p,1} = \max_{i=1,\dots,n-d} c_{p_i,1}, \quad c_{p^\top,1} = \sum_{i=1,\dots,n-d} c_{p_i,1}$$

to obtain the bounds. \square

4.3. Approximate isotropicity of tangent vectors. In this section we prove that if \mathcal{K} is approximately invariant, then $K^*\omega$ is small and can be controlled by the error of invariance. We refer the reader to [29, 31] for similar computations, using generic constants in the estimates.

Lemma 4.3. *Let us consider the setting of Theorem 3.1 or Theorem 3.6. Let us consider $\Omega_K : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times 2n}$, the matrix representation of the pull-back on \mathbb{T}^d of the symplectic form. We have*

$$(4.10) \quad \langle \Omega_K \rangle = O_d,$$

and the following estimate holds:

$$(4.11) \quad \|\Omega_K\|_{\rho-2\delta} \leq \frac{C_{\Omega_K}}{\gamma\delta^{\tau+1}} \|E\|_{\rho},$$

where the constant C_{Ω_K} is provided in Table 1.

Proof. Property (4.10) follows directly from the exact symplectic structure, since $K^*\omega = d(K^*\alpha)$. In more algebraic terms, we have

$$\begin{aligned} \Omega_K(\theta) &= (DK(\theta))^\top ((Da(K(\theta)))^\top - Da(K(\theta))) DK(\theta) \\ &= (D(a(K(\theta))))^\top DK(\theta) - (DK(\theta))^\top D(a(K(\theta))), \end{aligned}$$

and so, the components of $\Omega_K(\theta)$ are

$$\begin{aligned} (\Omega_K)_{i,j}(\theta) &= \sum_{m=1}^{2n} \left(\frac{\partial(a_m(K(\theta)))}{\partial\theta_i} \frac{\partial K_m(\theta)}{\partial\theta_j} - \frac{\partial(a_m(K(\theta)))}{\partial\theta_j} \frac{\partial K_m(\theta)}{\partial\theta_i} \right) \\ &= \sum_{m=1}^{2n} \left(\frac{\partial}{\partial\theta_i} \left(a_m(K(\theta)) \frac{\partial(K_m(\theta))}{\partial\theta_j} \right) - \frac{\partial}{\partial\theta_j} \left(a_m(K(\theta)) \frac{\partial(K_m(\theta))}{\partial\theta_i} \right) \right). \end{aligned}$$

Hence, the components of Ω_K are sums of derivatives of periodic functions, and we obtain $\langle \Omega_K \rangle = O_d$.

Now we will use two crucial geometric properties (following Appendix A in [31]). Using the fact that ω is closed, we first obtain the expression

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

for any triplet (r, s, t) . The second property is obtained by taking derivatives at both sides of $\Omega(z)X_h(z) = (Dh(z))^\top$, obtaining

$$\frac{\partial^2 h}{\partial z_i \partial z_j}(z) = \sum_{m=1}^{2n} \left(\frac{\partial\Omega_{j,m}(z)}{\partial z_i} X_m(z) + \Omega_{j,m}(z) \frac{\partial X_m(z)}{\partial z_i} \right),$$

for any i, j , where we use the notation $X_m = (X_h)_m$ for the components of X_h . Hence,

$$\begin{aligned}
0 &= \frac{\partial^2 h}{\partial z_j \partial z_i}(z) - \frac{\partial^2 h}{\partial z_i \partial z_j}(z) \\
&= \sum_{m=1}^{2n} \left(\frac{\partial \Omega_{i,m}(z)}{\partial z_j} X_m(z) + \Omega_{i,m}(z) \frac{\partial X_m(z)}{\partial z_j} \right) \\
&\quad - \sum_{m=1}^{2n} \left(\frac{\partial \Omega_{j,m}(z)}{\partial z_i} X_m(z) + \Omega_{j,m}(z) \frac{\partial X_m(z)}{\partial z_i} \right) \\
&= \sum_{m=1}^{2n} \left(\frac{\partial \Omega_{i,j}(z)}{\partial z_m} X_m(z) + \Omega_{i,m}(z) \frac{\partial X_m(z)}{\partial z_j} + \Omega_{m,j}(z) \frac{\partial X_m(z)}{\partial z_i} \right).
\end{aligned}$$

The above expressions yield the formula

$$(4.12) \quad D\Omega(z)[X_h(z)] + (DX_h(z))^\top \Omega(z) + \Omega(z)DX_h(z) = O_{2n}.$$

Then, we compute the action of \mathfrak{L}_ω on Ω_K , thus obtaining

$$(4.13) \quad \begin{aligned} \mathfrak{L}_\omega \Omega_K(\theta) &= \mathfrak{L}_\omega (DK(\theta))^\top \Omega(K(\theta)) DK(\theta) + (DK(\theta))^\top \mathfrak{L}_\omega (\Omega(K(\theta))) DK(\theta) \\ &\quad + (DK(\theta))^\top \Omega(K(\theta)) \mathfrak{L}_\omega DK(\theta), \end{aligned}$$

and we use the properties (obtained from the invariance equation (3.9))

$$\begin{aligned} \mathfrak{L}_\omega (DK(\theta)) &= DE(\theta) - DX_h(K(\theta))DK(\theta), \\ \mathfrak{L}_\omega (\Omega(K(\theta))) &= D\Omega(K(\theta))[\mathfrak{L}_\omega K(\theta)] = D\Omega(K(\theta)) [E(\theta) - X_h(K(\theta))], \end{aligned}$$

in combination with (4.12), thus ending up with

$$\begin{aligned} \mathfrak{L}_\omega \Omega_K(\theta) &= (DE(\theta))^\top \Omega(K(\theta)) DK(\theta) + (DK(\theta))^\top (D\Omega(K(\theta)) [E(\theta)]) DK(\theta) \\ &\quad + (DK(\theta))^\top \Omega(K(\theta)) DE(\theta). \end{aligned}$$

The expression $\mathfrak{L}_\omega \Omega_K(\theta)$ is controlled using H_1, H_2 , the Banach algebra properties and Cauchy estimates as follows

$$(4.14) \quad \begin{aligned} \|\mathfrak{L}_\omega \Omega_K\|_{\rho-\delta} &\leq \|(DE)^\top\|_{\rho-\delta} \|\Omega\|_{\mathcal{B}} \|DK\|_{\rho} + \|(DK)^\top\|_{\rho} \|D\Omega\|_{\mathcal{B}} \|E\|_{\rho} \|DK\|_{\rho} \\ &\quad + \|(DK)^\top\|_{\rho} \|\Omega\|_{\mathcal{B}} \|DE\|_{\rho-\delta} \\ &\leq \frac{2nc_{\Omega,0}\sigma_K + \sigma_K \tau c_{\Omega,1}\sigma_K \delta + d\sigma_{K^\top c_{\Omega,0}}}{\delta} \|E\|_{\rho} =: \frac{C_{\mathfrak{L}_\omega \Omega_K}}{\delta} \|E\|_{\rho}. \end{aligned}$$

In particular, we used that $\|(D\Omega \circ K)[E]\|_{\mathcal{B}} \leq \|D\Omega\|_{\mathcal{B}} \|E\|_{\rho}$ (see Section 3.1). The estimate in (4.11) is obtained as follows

$$(4.15) \quad \|\Omega_K\|_{\rho-2\delta} \leq \frac{c_R}{\gamma \delta^\tau} \|\mathfrak{L}_\omega \Omega_K\|_{\rho-\delta} \leq \frac{c_R C_{\mathfrak{L}_\omega \Omega_K}}{\gamma \delta^{\tau+1}} \|E\|_{\rho} =: \frac{C_{\Omega_K}}{\gamma \delta^{\tau+1}} \|E\|_{\rho},$$

where we used Lemma 4.1 and the estimate in (4.14). \square

Remark 4.4. Notice that, even though the KAM theorems 3.1 and 3.6 do not require a quantitative control on the 1-form α , the facts that the symplectic structure ω is exact and the vector field X_h is globally Hamiltonian are crucial to obtain the above result.

4.4. Approximate symplectic frame. In this section we prove that if \mathcal{K} is approximately invariant, then we can construct an adapted frame that is approximately symplectic. First step is finding an adapted approximately Lagrangian bundle, that contains the tangent bundle $T_{\mathcal{K}}\mathcal{M}$.

Lemma 4.5. *Let us consider the setting of Theorem 3.1 or Theorem 3.6. Then, the map $L(\theta)$ given by (3.8) satisfies*

$$(4.16) \quad \|L\|_{\rho} \leq C_L, \quad \|L^{\top}\|_{\rho} \leq C_{L^{\top}}, \quad \langle L^{\top}(\Omega \circ K)E \rangle = 0_n,$$

and defines an approximately Lagrangian bundle, i.e. the error map

$$(4.17) \quad \Omega_L(\theta) := L(\theta)^{\top} \Omega(K(\theta)) L(\theta)$$

is small in the following sense:

$$(4.18) \quad \|\Omega_L\|_{\rho-2\delta} \leq \frac{C_{\Omega_L}}{\gamma\delta^{\tau+1}} \|E\|_{\rho}.$$

Furthermore, the objects

$$(4.19) \quad G_L(\theta) := L(\theta)^{\top} G(K(\theta)) L(\theta),$$

$$(4.20) \quad \tilde{\Omega}_L(\theta) := L(\theta)^{\top} \tilde{\Omega}(K(\theta)) L(\theta),$$

are controlled as

$$(4.21) \quad \|G_L\|_{\rho} \leq C_{G_L}, \quad \|\tilde{\Omega}_L\|_{\rho} \leq C_{\tilde{\Omega}_L}.$$

The above constants are given explicitly in Table 1.

Proof. We first obtain the property of the average in (4.16) by computing

$$L(\theta)^{\top} \Omega(K(\theta)) E(\theta) = \begin{pmatrix} (\mathrm{DK}(\theta))^{\top} \Omega(K(\theta)) E(\theta) \\ X_p(K(\theta))^{\top} \Omega(K(\theta)) E(\theta) \end{pmatrix}.$$

The upper term satisfies

$$\begin{aligned} (\mathrm{DK}(\theta))^{\top} \Omega(K(\theta)) E(\theta) &= (\mathrm{DK}(\theta))^{\top} \Omega(K(\theta)) X_h(K(\theta)) - \Omega_K(\theta) \omega \\ &= (\mathrm{DK}(\theta))^{\top} (\mathrm{D}h(K(\theta)))^{\top} - \Omega_K(\theta) \omega \\ &= (\mathrm{D}(h(K(\theta))))^{\top} - \Omega_K(\theta) \omega, \end{aligned}$$

which has zero average, since the first term is the derivative of a periodic function and $\Omega_K(\theta)$ has zero average (see Lemma 4.3). Moreover, the lower term satisfies

$$(4.22) \quad X_p(K(\theta))^{\top} \Omega(K(\theta)) E(\theta) = \mathrm{D}p(K(\theta)) E(\theta) = -\mathrm{D}(p(K(\theta))) \omega.$$

Hence, it is clear that $\langle L^{\top}(\Omega \circ K)E \rangle = 0_n$.

We control the norm of the frame $L(\theta)$, using H_1 and H_2 , as follows

$$(4.23) \quad \|L\|_{\rho} \leq \|\mathrm{DK}\|_{\rho} + \|X_p \circ K\|_{\rho} \leq \sigma_K + c_{X_p,0} =: C_L,$$

$$(4.24) \quad \|L^{\top}\|_{\rho} \leq \max\{\|(\mathrm{DK})^{\top}\|_{\rho}, \|X_p^{\top} \circ K\|_{\rho}\} \leq \max\{\sigma_{K^{\top}}, c_{X_p^{\top},0}\} =: C_{L^{\top}}.$$

We have obtained the estimates in (4.16). Using again the expression of $L(\theta)$ we obtain that the anti-symmetric matrix (4.17) is written as

$$\Omega_L(\theta) = \begin{pmatrix} (\mathrm{DK}(\theta))^{\top} \Omega(K(\theta)) \mathrm{DK}(\theta) & (\mathrm{DK}(\theta))^{\top} \Omega(K(\theta)) X_p(K(\theta)) \\ X_p(K(\theta))^{\top} \Omega(K(\theta)) \mathrm{DK}(\theta) & X_p(K(\theta))^{\top} \Omega(K(\theta)) X_p(K(\theta)) \end{pmatrix}.$$

Using the expression $\Omega_K(\theta)$ in (2.3), performing similar computations as in (4.22), and using the involution of the first integrals, we end up with

$$\Omega_L(\theta) = \begin{pmatrix} \Omega_K(\theta) & (\mathbf{D}(p(K(\theta))))^\top \\ -\mathbf{D}(p(K(\theta))) & O_{n-d} \end{pmatrix}.$$

Then, we have

$$(4.25) \quad \begin{aligned} \|\Omega_L\|_{\rho-2\delta} &= \max\{\|\Omega_K\|_{\rho-2\delta} + \|(\mathbf{D}(p \circ K))^\top\|_{\rho-2\delta}, \|\mathbf{D}(p \circ K)\|_{\rho-2\delta}\} \\ &\leq \frac{c_R \max\{C_{\mathcal{L}\Omega_K} + c_{p^\top,1}, dc_{p,1}\}}{\gamma\delta^{\tau+1}} \|E\|_\rho =: \frac{C_{\Omega_L}}{\gamma\delta^{\tau+1}} \|E\|_\rho, \end{aligned}$$

where we used Lemma 4.2 and Lemma 4.3. Thus, we have obtained the estimate in (4.18). Finally, the estimates in (4.21) with

$$(4.26) \quad \|G_L\|_\rho \leq C_{L^\top} c_{G,0} C_L =: C_{G_L},$$

$$(4.27) \quad \|\tilde{\Omega}_L\|_\rho \leq C_{L^\top} c_{\tilde{\Omega},0} C_L =: C_{\tilde{\Omega}_L},$$

follow directly. \square

In the following lemma, we will see that the geometric constructions detailed in Section 2.5 lead, for an approximately invariant torus, to an approximately symplectic frame attached to the torus.

Lemma 4.6. *Let us consider the setting of Theorem 3.1 or Theorem 3.6. Then, the map $N : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times n}$ given by (3.2) satisfies*

$$(4.28) \quad \|N\|_\rho \leq C_N, \quad \|N^\top\|_\rho \leq C_{N^\top},$$

and the map $P : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times 2n}$ given by

$$P(\theta) = \begin{pmatrix} L(\theta) & N(\theta) \end{pmatrix},$$

induces an approximately symplectic vector bundle isomorphism, i.e., the error map

$$(4.29) \quad E_{\text{sym}}(\theta) := P(\theta)^\top \Omega(K(\theta)) P(\theta) - \Omega_0, \quad \Omega_0 = \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix},$$

is small in the following sense:

$$(4.30) \quad \|E_{\text{sym}}\|_{\rho-2\delta} \leq \frac{C_{\text{sym}}}{\gamma\delta^{\tau+1}} \|E\|_\rho.$$

The above constants are given explicitly in Table 1.

Proof. First we control the norm of $N^0(\theta)$ in (3.3), using H_1 and (4.16), as

$$(4.31) \quad \|N^0\|_\rho \leq \|J \circ K\|_\rho \|L\|_\rho \leq \|J\|_{\mathcal{B}} \|L\|_\rho \leq c_{J,0} C_L =: C_{N^0},$$

$$(4.32) \quad \|(N^0)^\top\|_\rho \leq \|L^\top\|_\rho \|(J \circ K)^\top\|_\rho \leq \|L^\top\|_\rho \|J^\top\|_{\mathcal{B}} \leq C_{L^\top} c_{J^\top,0} =: C_{N^{0,\top}},$$

and the norm of $A(\theta)$ in (3.5) as

$$(4.33) \quad \|A\|_\rho \leq \frac{1}{2} \|B^\top\|_\rho \|L^\top(\tilde{\Omega} \circ K)L\|_\rho \|B\|_\rho \leq \frac{1}{2} \sigma_B C_{\tilde{\Omega}_L} \sigma_B =: C_A,$$

where we also used H_3 , the second estimate in (4.21) and that $B(\theta)$ is symmetric. We also control the complementary normal vectors as

$$(4.34) \quad \|N\|_\rho \leq \|L\|_\rho \|A\|_\rho + \|N^0\|_\rho \|B\|_\rho \leq C_L C_A + C_{N^0} \sigma_B =: C_N,$$

$$(4.35) \quad \|N^\top\|_\rho \leq \|A^\top\|_\rho \|L^\top\|_\rho + \|B^\top\|_\rho \|(N^0)^\top\|_\rho \leq C_A C_{L^\top} + \sigma_B C_{N^{0,\top}} =: C_{N^\top},$$

where we used H_1 , the estimates (4.16), (4.31), (4.32), and (4.33), and the fact that $A(\theta)$ is anti-symmetric. Thus, we have obtained the estimates in (4.28).

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

$$(4.36) \quad E_{\text{sym}}(\theta) = \begin{pmatrix} L(\theta)^\top \Omega(K(\theta))L(\theta) & L(\theta)^\top \Omega(K(\theta))N(\theta) + I_n \\ N(\theta)^\top \Omega(K(\theta))L(\theta) - I_n & N(\theta)^\top \Omega(K(\theta))N(\theta) \end{pmatrix},$$

and we expand the components of this block matrix using (3.2), (3.4), and (4.17). For example, we have

$$(4.37) \quad \begin{aligned} & L(\theta)^\top \Omega(K(\theta))N(\theta) \\ &= L(\theta)^\top \Omega(K(\theta))L(\theta)A(\theta) + L(\theta)^\top \Omega(K(\theta))N^0(\theta)B(\theta) \\ &= \Omega_L(\theta)A(\theta) - L(\theta)^\top G(K(\theta))L(\theta)B(\theta) \\ &= \Omega_L(\theta)A(\theta) - I_n, \end{aligned}$$

where we used that $J^\top \Omega = G$ and the definition of $B(\theta)$. We also have

$$(4.38) \quad \begin{aligned} & N(\theta)^\top \Omega(K(\theta))N(\theta) \\ &= B(\theta)^\top N^0(\theta)^\top \Omega(K(\theta))L(\theta)A(\theta) + B(\theta)^\top N^0(\theta)^\top \Omega(K(\theta))N^0(\theta)B(\theta) \\ &\quad + A(\theta)^\top L(\theta)^\top \Omega(K(\theta))L(\theta)A(\theta) + A(\theta)^\top L(\theta)^\top \Omega(K(\theta))N^0(\theta)B(\theta) \\ &= A(\theta)^\top \Omega_L(\theta)A(\theta) + A(\theta) - A(\theta)^\top + B(\theta)^\top L(\theta)^\top \tilde{\Omega}(K(\theta))L(\theta)B(\theta) \\ &= A(\theta)^\top \Omega_L(\theta)A(\theta). \end{aligned}$$

Then, introducing the expressions (4.17), (4.37) and (4.38) into (4.36), we get

$$E_{\text{sym}}(\theta) = \begin{pmatrix} \Omega_L(\theta) & \Omega_L(\theta)A(\theta) \\ A(\theta)^\top \Omega_L(\theta) & A(\theta)^\top \Omega_L(\theta)A(\theta) \end{pmatrix},$$

which is controlled as

$$(4.39) \quad \|E_{\text{sym}}\|_{\rho-2\delta} \leq \frac{(1 + C_A) \max\{1, C_A\} C_{\Omega_L}}{\gamma \delta^{\tau+1}} \|E\|_\rho =: \frac{C_{\text{sym}}}{\gamma \delta^{\tau+1}} \|E\|_\rho,$$

thus obtaining the estimate (4.30). \square

Remark 4.7. The above estimates can be readily adapted to **Case III**, for which $A = 0$. In this case we have (computations are left as an exercise to the reader)

$$N(\theta) = N^0(\theta)B(\theta), \quad E_{\text{sym}}(\theta) = \begin{pmatrix} \Omega_L(\theta) & O_n \\ O_n & B(\theta)^\top \Omega_L(\theta)B(\theta) \end{pmatrix}.$$

The corresponding estimates are given explicitly in Table 1.

4.5. Sharp control of the torsion matrix. In this section we control the torsion matrix $T(\theta)$, given in (3.6). To do so, we could use directly Cauchy estimates to control $\mathfrak{L}_\omega N(\theta)$, resulting in an additional bite in the domain and an additional factor δ in the denominator (among other overestimations). Hence, in order to improve this estimate, thus enhancing the threshold of validity of the result, we perform a finer analysis of the expression for $T(\theta)$. To this end, it is convenient to include here an additional smallness condition for the error of invariance (see (4.40) below), that later on will turn out to be rather irrelevant.

Lemma 4.8. *Let us consider the setting of Theorem 3.1 or Theorem 3.6, and let us assume that*

$$(4.40) \quad \frac{\|E\|_\rho}{\delta} < \nu,$$

where ν is an independent constant. Then, the torsion matrix $T(\theta)$, given by (3.6), has components in $\mathcal{A}(\mathbb{T}_{\rho-\delta}^d)$ and satisfies the estimate

$$\|T\|_{\rho-\delta} \leq C_T,$$

where the constant C_T is provided in Table 1.

Proof. Recalling that $\mathfrak{L}_\omega(\cdot) = -D(\cdot)\omega$ and (2.12), we have

$$(4.41) \quad \mathcal{X}_N(\theta) = DX_h(K(\theta))N(\theta) + \mathfrak{L}_\omega N(\theta),$$

where

$$(4.42) \quad \mathfrak{L}_\omega N(\theta) = \mathfrak{L}_\omega L(\theta)A(\theta) + L(\theta)\mathfrak{L}_\omega A(\theta) + \mathfrak{L}_\omega N^0(\theta)B(\theta) + N^0(\theta)\mathfrak{L}_\omega B(\theta).$$

Then, we must estimate the terms $\mathfrak{L}_\omega L(\theta)$, $\mathfrak{L}_\omega A(\theta)$, $\mathfrak{L}_\omega N^0(\theta)$, and $\mathfrak{L}_\omega B(\theta)$, that appear above.

We start considering

$$(4.43) \quad \mathfrak{L}_\omega K(\theta) = E(\theta) - X_h(K(\theta))$$

which is controlled as

$$(4.44) \quad \|\mathfrak{L}_\omega K\|_\rho \leq \|E\|_\rho + \|X_h \circ K\|_\rho \leq \delta\nu + c_{X_h,0} =: C_{\mathfrak{L}K},$$

where we used H_1 and the assumption (4.40). Then we consider the object $\mathfrak{L}_\omega L(\theta)$, given by

$$\mathfrak{L}_\omega L(\theta) = \left(\mathfrak{L}_\omega(DK(\theta)) \quad \mathfrak{L}_\omega(X_p(K(\theta))) \right).$$

Notice that the left block in the above expression follows by taking derivatives at both sides of (4.43), i.e.

$$\mathfrak{L}_\omega(DK(\theta)) = DE(\theta) - DX_h(K(\theta))[DK(\theta)],$$

and the right block follows from

$$\mathfrak{L}_\omega(X_p(K(\theta))) = DX_p(K(\theta))[\mathfrak{L}_\omega K(\theta)].$$

Then, using Cauchy estimates, the Hypothesis H_1 and H_2 , the assumption (4.40) and the estimate (4.44), we have

$$(4.45) \quad \|\mathfrak{L}_\omega L\|_{\rho-\delta} \leq d\nu + c_{X_h,1}\sigma_K + c_{X_p,1}C_{\mathfrak{L}K} =: C_{\mathfrak{L}L}.$$

Similarly, we obtain the estimates

$$(4.46) \quad \|\mathfrak{L}_\omega L^\top\|_{\rho-\delta} \leq \max \left\{ 2n\nu + c_{X_h^\top,1}\sigma_K, c_{X_p^\top,1}C_{\mathfrak{L}K} \right\} =: C_{\mathfrak{L}L^\top}.$$

The term $\mathfrak{L}_\omega N^0(\theta)$ is controlled using that

$$(4.47) \quad \mathfrak{L}_\omega N^0(\theta) = \mathfrak{L}_\omega(J(K(\theta)))L(\theta) + J(K(\theta))\mathfrak{L}_\omega L(\theta)$$

and the chain rule. Actually, we have

$$(4.48) \quad \|\mathfrak{L}_\omega(J \circ K)\|_\rho \leq c_{J,1}C_{\mathfrak{L}K} =: C_{\mathfrak{L}J},$$

$$(4.49) \quad \|\mathfrak{L}_\omega(G \circ K)\|_\rho \leq c_{G,1}C_{\mathfrak{L}K} =: C_{\mathfrak{L}G},$$

$$(4.50) \quad \|\mathfrak{L}_\omega(\tilde{\Omega} \circ K)\|_\rho \leq c_{\tilde{\Omega},1}C_{\mathfrak{L}K} =: C_{\mathfrak{L}\tilde{\Omega}},$$

and then, using (4.48), the expression (4.47) is controlled as

$$(4.51) \quad \|\mathfrak{L}_\omega N^0\|_{\rho-\delta} \leq C_{\mathfrak{L}J}C_L + c_{J,0}C_{\mathfrak{L}L} =: C_{\mathfrak{L}N^0}.$$

Before controlling $\mathfrak{L}_\omega B(\theta)$ and $\mathfrak{L}_\omega A(\theta)$, we recall the notation for $G_L(\theta)$ and $\tilde{\Omega}_L(\theta)$, given by (4.19) and (4.20) respectively, and we control the action of \mathfrak{L}_ω on these objects. For example, we have

$$\begin{aligned} \mathfrak{L}_\omega G_L(\theta) &= \mathfrak{L}_\omega L(\theta)^\top G(K(\theta))L(\theta) \\ &\quad + L(\theta)^\top \mathfrak{L}_\omega G(K(\theta))L(\theta) + L(\theta)^\top G(K(\theta))\mathfrak{L}_\omega L(\theta) \end{aligned}$$

and, using (4.49), we obtain the estimate

$$(4.52) \quad \|\mathfrak{L}_\omega G_L\|_{\rho-\delta} \leq C_{\mathfrak{L}L^\top}c_{G,0}C_L + C_{L^\top}C_{\mathfrak{L}G}C_L + C_{L^\top}c_{G,0}C_{\mathfrak{L}L} =: C_{\mathfrak{L}G_L}.$$

Analogously, using (4.50), we obtain the estimate

$$(4.53) \quad \|\mathfrak{L}_\omega \tilde{\Omega}_L\|_{\rho-\delta} \leq C_{\mathfrak{L}L^\top}c_{\tilde{\Omega},0}C_L + C_{L^\top}C_{\mathfrak{L}\tilde{\Omega}}C_L + C_{L^\top}c_{\tilde{\Omega},0}C_{\mathfrak{L}L} =: C_{\mathfrak{L}\tilde{\Omega}_L}.$$

Now we obtain a suitable expression for $\mathfrak{L}_\omega B(\theta)$. To this end, we compute

$$O_n = \mathfrak{L}_\omega(B(\theta)^{-1}B(\theta)) = \mathfrak{L}_\omega(B(\theta)^{-1})B(\theta) + B(\theta)^{-1}\mathfrak{L}_\omega(B(\theta)),$$

which, recalling (3.4), yields the following expression

$$\mathfrak{L}_\omega B(\theta) = -B(\theta)\mathfrak{L}_\omega(L(\theta)^\top G(K(\theta))L(\theta))B(\theta) = -B(\theta)\mathfrak{L}_\omega G_L(\theta)B(\theta).$$

Then, using (4.52), we get the estimate

$$(4.54) \quad \|\mathfrak{L}_\omega B\|_{\rho-\delta} \leq (\sigma_B)^2 C_{\mathfrak{L}G_L} =: C_{\mathfrak{L}B}.$$

Since $B(\theta)$ is symmetric, we also have $\|\mathfrak{L}_\omega B^\top\|_{\rho-\delta} \leq C_{\mathfrak{L}B}$.

Finally, using the above notation, we expand $\mathfrak{L}_\omega A(\theta)$ as

$$\begin{aligned} \mathfrak{L}_\omega A(\theta) &= -\frac{1}{2}\mathfrak{L}_\omega B(\theta)^\top \tilde{\Omega}_L(\theta)B(\theta) - \frac{1}{2}B(\theta)^\top \mathfrak{L}_\omega \tilde{\Omega}_L(\theta)B(\theta) \\ &\quad - \frac{1}{2}B(\theta)^\top \tilde{\Omega}_L(\theta)\mathfrak{L}_\omega B(\theta), \end{aligned}$$

which, using the constants (4.27) and (4.53), yields the following estimate

$$(4.55) \quad \|\mathfrak{L}_\omega A\|_{\rho-\delta} \leq C_{\mathfrak{L}B}C_{\tilde{\Omega}_L}\sigma_B + \frac{1}{2}(\sigma_B)^2 C_{\mathfrak{L}\tilde{\Omega}_L} =: C_{\mathfrak{L}A}.$$

With the above objects, we can control (4.42) as follows

$$(4.56) \quad \|\mathfrak{L}_\omega N\|_{\rho-\delta} \leq C_{\mathfrak{L}L}C_A + C_L C_{\mathfrak{L}A} + C_{\mathfrak{L}N^0}\sigma_B + C_{N^0}C_{\mathfrak{L}B} =: C_{\mathfrak{L}N}.$$

This estimate will be used later in the proof of Lemma 5.2. Now, we could use (4.56) in equation (4.41) to obtain

$$\|\mathcal{X}_N\|_{\rho-\delta} \leq c_{X_h,1}C_N + C_{\mathfrak{L}N}.$$

However, we obtain a sharper estimate observing that

$$\begin{aligned} \mathcal{X}_N(\theta) &= \mathcal{X}_L(\theta)A(\theta) + DX_h(K(\theta))N^0(\theta)B(\theta) \\ &\quad + L(\theta)\mathfrak{L}_\omega A(\theta) + \mathfrak{L}_\omega N^0(\theta)B(\theta) + N^0(\theta)\mathfrak{L}_\omega B(\theta), \end{aligned}$$

where

$$\begin{aligned} \mathcal{X}_L(\theta) &= DX_h(K(\theta))L(\theta) + \mathfrak{L}_\omega L(\theta) \\ &= (DE(\theta) \quad DX_p(K(\theta))[E(\theta)]) . \end{aligned}$$

This last expression follows using the previous formula for $\mathfrak{L}_\omega L(\theta)$ and the fact that the vector field X_h commutes with the fields X_{p_i} for $1 \leq i \leq n-d$.

Then, the objects $\mathcal{X}_L(\theta)$ and $\mathcal{X}_L(\theta)^\top$ are controlled as follows:

$$(4.57) \quad \|\mathcal{X}_L\|_{\rho-\delta} \leq \frac{d + c_{X_p,1}\delta}{\delta} \|E\|_\rho =: \frac{C_{\mathcal{X}_L}}{\delta} \|E\|_\rho,$$

$$(4.58) \quad \|\mathcal{X}_L^\top\|_{\rho-\delta} \leq \frac{\max\{2n, c_{X_p^\top,1}\delta\}}{\delta} \|E\|_\rho =: \frac{C_{\mathcal{X}_L^\top}}{\delta} \|E\|_\rho,$$

and, using again the smallness condition (4.40), we obtain

$$(4.59) \quad \begin{aligned} \|\mathcal{X}_N\|_{\rho-\delta} &\leq C_{\mathcal{X}_L} \nu C_A + c_{X_h,1} C_{N^0} \sigma_B + C_L C_{\mathfrak{L}A} + C_{\mathfrak{L}N^0} \sigma_B + C_{N^0} C_{\mathfrak{L}B} \\ &=: C_{\mathcal{X}_N}. \end{aligned}$$

Finally, the torsion matrix satisfies

$$(4.60) \quad \|T\|_{\rho-\delta} \leq C_{N^\top} c_{\Omega,0} C_{\mathcal{X}_N} =: C_T,$$

which completes the proof. \square

Remark 4.9. The bound C_T of Lemma 4.8 could be improved for the particular problem at hand, since the expression for $T(\theta)$ can in many cases be obtained explicitly and may have cancellations.

4.6. Approximate reducibility. A crucial step in the proof of both KAM theorems is the resolution of the linearized equation arising from the application of the Newton method. The resolution is based on the (approximate) reduction of such linear system into a simpler form, in particular, block triangular form. This is the content of the following lemma.

Lemma 4.10. *Let us consider the setting of Theorem 3.1 or Theorem 3.6. Then, the map $P : \mathbb{T}^d \rightarrow \mathbb{R}^{2n \times 2n}$, characterized in Lemma 4.6, approximately reduces the linearized equation associated with the vector field $DX_h \circ K$ to a block-triangular matrix, i.e. the error map*

$$(4.61) \quad E_{\text{red}}(\theta) := -\Omega_0 P(\theta)^\top \Omega(K(\theta)) (DX_h(K(\theta))P(\theta) + \mathfrak{L}_\omega P(\theta)) - \Lambda(\theta),$$

with

$$(4.62) \quad \Lambda(\theta) = \begin{pmatrix} O_n & T(\theta) \\ O_n & O_n \end{pmatrix}$$

and $T(\theta)$ is given by (3.6), is small in the following sense:

$$\|E_{\text{red}}\|_{\rho-2\delta} \leq \frac{C_{\text{red}}}{\gamma \delta^{\tau+1}} \|E\|_\rho,$$

where the constant C_{red} is provided in Table 1.

Proof. Using the notation in (2.12) we write the block components of (4.61), denoted as $E_{\text{red}}^{i,j}(\theta)$, as follows:

$$(4.63) \quad E_{\text{red}}^{1,1}(\theta) = N(\theta)^\top \Omega(K(\theta)) \mathcal{X}_L(\theta),$$

$$(4.64) \quad E_{\text{red}}^{1,2}(\theta) = N(\theta)^\top \Omega(K(\theta)) \mathcal{X}_N(\theta) - T(\theta) = O_n,$$

$$(4.65) \quad E_{\text{red}}^{2,1}(\theta) = -L(\theta)^\top \Omega(K(\theta)) \mathcal{X}_L(\theta),$$

$$E_{\text{red}}^{2,2}(\theta) = -L(\theta)^\top \Omega(K(\theta)) \mathcal{X}_N(\theta).$$

Notice that we have used the definition of $T(\theta)$ to see that (4.64) vanishes. To gather a suitable expression for $E_{\text{red}}^{2,2}(\theta)$, we apply \mathfrak{L}_ω at both sides of the expression obtained in (4.37):

$$\mathfrak{L}_\omega(L(\theta)^\top \Omega(K(\theta)))N(\theta) + L(\theta)^\top \Omega(K(\theta))\mathfrak{L}_\omega N(\theta) = \mathfrak{L}_\omega(\Omega_L(\theta)A(\theta)).$$

Then, introducing this expression into $E_{\text{red}}^{2,2}(\theta)$, using (3.9) and the geometric property (4.12), we obtain

$$\begin{aligned} E_{\text{red}}^{2,2}(\theta) &= -L(\theta)^\top \Omega(K(\theta))DX_h(K(\theta))N(\theta) + \mathfrak{L}_\omega(L(\theta)^\top \Omega(K(\theta)))N(\theta) \\ &\quad - \mathfrak{L}_\omega(\Omega_L(\theta)A(\theta)) \\ (4.66) \quad &= L(\theta)^\top (D\Omega(K(\theta))[E(\theta)])N(\theta) + \mathcal{X}_L(\theta)^\top \Omega(K(\theta))N(\theta) - \mathfrak{L}_\omega(\Omega_L(\theta)A(\theta)). \end{aligned}$$

At this point, we could use Cauchy estimates in the expression

$$\mathfrak{L}_\omega(\Omega_L(\theta)A(\theta)) = -D(\Omega_L(\theta)A(\theta))[\omega]$$

and obtain an estimate controlled by $\|E\|_\rho$. However, this would give a control of the form $\|E_{\text{red}}\|_{\rho-3\delta}$, and we are interested in keeping the strip of analyticity $\rho-2\delta$. For this reason, we compute

$$(4.67) \quad \mathfrak{L}_\omega(\Omega_L(\theta)A(\theta)) = \mathfrak{L}_\omega \Omega_L(\theta) A(\theta) + \Omega_L(\theta) \mathfrak{L}_\omega A(\theta),$$

and consider the block components of

$$(4.68) \quad \mathfrak{L}_\omega \Omega_L(\theta) = \begin{pmatrix} \mathfrak{L}_\omega \Omega_K(\theta) & \mathfrak{L}_\omega(D(p(K(\theta))))^\top \\ -\mathfrak{L}_\omega(D(p(K(\theta)))) & O_{n-d} \end{pmatrix}.$$

Then, we control (4.68) as follows

$$\begin{aligned} \|\mathfrak{L}_\omega \Omega_L\|_{\rho-\delta} &\leq \max\{\|\mathfrak{L}_\omega \Omega_K\|_{\rho-\delta} + \|\mathfrak{L}_\omega(D(p \circ K))^\top\|_{\rho-\delta}, \|\mathfrak{L}_\omega(D(p \circ K))\|_{\rho-\delta}\} \\ (4.69) \quad &\leq \frac{\max\{C_{\mathfrak{L}\Omega_K} + c_{p^\top,1}, dc_{p,1}\}}{\delta} \|E\|_\rho =: \frac{C_{\mathfrak{L}\Omega_L}}{\delta} \|E\|_\rho, \end{aligned}$$

where we used estimates (4.7) and (4.8) from Lemma 4.2.

Finally, we estimate the norms of the block components of E_{red} using the expressions (4.63), (4.65), (4.66) and (4.67), and the previous estimates:

$$(4.70) \quad \|E_{\text{red}}^{1,1}\|_{\rho-2\delta} \leq \|N^\top\|_\rho \|\Omega\|_{\mathcal{B}} \|\mathcal{X}_L\|_{\rho-\delta} \leq \frac{C_{N^\top} c_{\Omega,0} C_{\mathcal{X}_L}}{\delta} \|E\|_\rho =: \frac{C_{\text{red}}^{1,1}}{\delta} \|E\|_\rho,$$

$$\|E_{\text{red}}^{1,2}\|_{\rho-2\delta} = 0,$$

$$(4.71) \quad \|E_{\text{red}}^{2,1}\|_{\rho-2\delta} \leq \|L^\top\|_\rho \|\Omega\|_{\mathcal{B}} \|\mathcal{X}_L\|_{\rho-\delta} \leq \frac{C_{L^\top} c_{\Omega,0} C_{\mathcal{X}_L}}{\delta} \|E\|_\rho =: \frac{C_{\text{red}}^{2,1}}{\delta} \|E\|_\rho,$$

$$\|E_{\text{red}}^{2,2}\|_{\rho-2\delta} \leq \left(C_{L^\top} c_{\Omega,1} C_N + \frac{C_{\mathcal{X}_L^\top} c_{\Omega,0} C_N}{\delta} + \frac{C_{\mathfrak{L}\Omega_L} C_A}{\delta} + \frac{C_{\Omega_L} C_{\mathfrak{L}A}}{\gamma \delta^{\tau+1}} \right) \|E\|_\rho$$

$$(4.72) \quad =: \frac{C_{\text{red}}^{2,2}}{\gamma \delta^{\tau+1}} \|E\|_\rho.$$

Then, we end up with

$$(4.73) \quad \|E_{\text{red}}\|_{\rho-2\delta} \leq \frac{\max\{C_{\text{red}}^{1,1} \gamma \delta^\tau, C_{\text{red}}^{2,1} \gamma \delta^\tau + C_{\text{red}}^{2,2}\}}{\gamma \delta^{\tau+1}} \|E\|_\rho =: \frac{C_{\text{red}}}{\gamma \delta^{\tau+1}} \|E\|_\rho,$$

thus completing the proof. \square

5. PROOF OF THE ORDINARY KAM THEOREM

In this section we present a fully detailed proof of Theorem 3.1. For convenience, we will start by outlining the scheme used to correct the parameterization of the torus. That is, in Section 5.1 we discuss the approximate solution of linearized equations in the symplectic frame constructed in Section 4.4. This establishes a quasi-Newton method to obtain a solution of the invariance equation. In Section 5.2 we produce quantitative estimates for the objects obtained when performing one iteration of the previous procedure. Finally, in Section 5.3 we discuss the convergence of the quasi-Newton method.

5.1. The quasi-Newton method. As it is usual in the a-posteriori approach to KAM theory, the argument consists in refining $K(\theta)$ by means of a quasi-Newton method. Let us consider the equations associated with the error of invariance

$$E(\theta) = X_h(K(\theta)) + \mathfrak{L}_\omega K(\theta).$$

Then, we obtain the new parameterization $\bar{K}(\theta) = K(\theta) + \Delta K(\theta)$ by considering the linearized equation

$$(5.1) \quad DX_h(K(\theta))\Delta K(\theta) + \mathfrak{L}_\omega \Delta K(\theta) = -E(\theta).$$

If we obtain a good enough approximation of the solution $\Delta K(\theta)$ of (5.1), then $\bar{K}(\theta)$ provides a parameterization of an approximately invariant torus of frequency ω , with a quadratic error in terms of $E(\theta)$.

To face the linearized equation (5.1), we resort to the approximately symplectic frame $P(\theta)$, defined on the full tangent space, which has been characterized in Section 4 (see Lemma 4.6). In particular, we introduce the linear change

$$(5.2) \quad \Delta K(\theta) = P(\theta)\xi(\theta),$$

where $\xi(\theta)$ is the new unknown. Taking into account this expression, the linearized equation becomes

$$(5.3) \quad (DX_h(K(\theta))P(\theta) + \mathfrak{L}_\omega P(\theta))\xi(\theta) + P(\theta)\mathfrak{L}_\omega \xi(\theta) = -E(\theta).$$

We now multiply both sides of (5.3) by $-\Omega_0 P(\theta)^\top \Omega(K(\theta))$, and we use the geometric properties in Lemma 4.6 and Lemma 4.10, thus obtaining the equivalent equations:

$$(5.4) \quad \begin{aligned} & (\Lambda(\theta) + E_{\text{red}}(\theta))\xi(\theta) + (I_{2n} - \Omega_0 E_{\text{sym}}(\theta))\mathfrak{L}_\omega \xi(\theta) \\ & = \Omega_0 P(\theta)^\top \Omega(K(\theta))E(\theta), \end{aligned}$$

where $\Lambda(\theta)$ is the triangular matrix-valued map given in (4.62).

Then, it turns out that the solution of (5.4) are approximated by the solutions of a triangular system that requires to solve two cohomological equations of the form (4.1) consecutively. Quantitative estimates for the solutions of such equations are obtained by applying Rüssmann estimates. This is summarized in the following standard statement.

Lemma 5.1 (Upper triangular equations). *Let $\omega \in \mathcal{D}_{\gamma,\tau}$ and let us consider a map $\eta = (\eta^L, \eta^N) : \mathbb{T}^d \rightarrow \mathbb{R}^{2n} \simeq \mathbb{R}^n \times \mathbb{R}^n$, with components in $\mathcal{A}(\mathbb{T}_\rho^d)$, and a map $T : \mathbb{T}^d \rightarrow \mathbb{R}^{n \times n}$, with components in $\mathcal{A}(\mathbb{T}_{\rho-\delta}^d)$. Assume that T satisfies the*

non-degeneracy condition $\langle T \rangle \neq 0$ and η satisfies the compatibility condition $\langle \eta^N \rangle = 0_n$. Then, for any $\xi_0^L \in \mathbb{R}^n$, the system of equations

$$(5.5) \quad \begin{pmatrix} O_n & T(\theta) \\ O_n & O_n \end{pmatrix} \begin{pmatrix} \xi^L(\theta) \\ \xi^N(\theta) \end{pmatrix} + \begin{pmatrix} \mathfrak{L}_\omega \xi^L(\theta) \\ \mathfrak{L}_\omega \xi^N(\theta) \end{pmatrix} = \begin{pmatrix} \eta^L(\theta) \\ \eta^N(\theta) \end{pmatrix}$$

has a solution of the form

$$\begin{aligned} \xi^N(\theta) &= \xi_0^N + \mathfrak{R}_\omega(\eta^N(\theta)), \\ \xi^L(\theta) &= \xi_0^L + \mathfrak{R}_\omega(\eta^L(\theta) - T(\theta)\xi^N(\theta)), \end{aligned}$$

where

$$\xi_0^N = \langle T \rangle^{-1} \langle \eta^L - T\mathfrak{R}_\omega(\eta^N) \rangle$$

and \mathfrak{R}_ω is given by (4.2). Moreover, we have the estimates

$$\begin{aligned} |\xi_0^N| &\leq |\langle T \rangle^{-1}| \left(\|\eta^L\|_\rho + \frac{c_R}{\gamma\delta^\tau} \|T\|_{\rho-\delta} \|\eta^N\|_\rho \right), \\ \|\xi^N\|_{\rho-\delta} &\leq |\xi_0^N| + \frac{c_R}{\gamma\delta^\tau} \|\eta^N\|_\rho, \\ \|\xi^L\|_{\rho-2\delta} &\leq |\xi_0^L| + \frac{c_R}{\gamma\delta^\tau} \left(\|\eta^L\|_{\rho-\delta} + \|T\|_{\rho-\delta} \|\xi^N\|_{\rho-\delta} \right). \end{aligned}$$

Proof. This triangular structure is classic in KAM theory and appears in any Kolmogorov scheme (see e.g. [3, 19, 35]). The Lemma is directly adapted from Lemma 4.14 in [33] and the estimates are directly obtained using Lemma 4.1. \square

To approximate the solutions of (5.4) we will invoke Lemma 5.1 taking

$$(5.6) \quad \eta^L(\theta) = -N(\theta)^\top \Omega(K(\theta))E(\theta), \quad \eta^N(\theta) = L(\theta)^\top \Omega(K(\theta))E(\theta),$$

and $T(\theta)$ given by (3.6). We recall from Lemma 4.6 that the compatibility condition $\langle \eta^N \rangle = 0_n$ is satisfied. Note that $\langle \xi^N \rangle = \xi_0^N$ and we have the freedom of choosing any value for $\langle \xi^L \rangle = \xi_0^L \in \mathbb{R}^n$. For convenience, we will select later the solution with $\xi_0^L = 0_n$, even though other choices can be selected according to the context (see Remark 3.5).

From Lemma 5.1 we read that $\|\xi\|_{\rho-2\delta} = \mathcal{O}(\|E\|_\rho)$ and, using the geometric properties characterized in Section 4, we have $\|E_{\text{red}}\|_{\rho-2\delta} = \mathcal{O}(\|E\|_\rho)$ and $\|E_{\text{sym}}\|_{\rho-2\delta} = \mathcal{O}(\|E\|_\rho)$. From these estimates, we conclude that the solution of equation (5.4) is approximated by the solution of the cohomological equation

$$(5.7) \quad \Lambda(\theta)\xi(\theta) + \mathfrak{L}_\omega \xi(\theta) = \eta(\theta).$$

This, together with other estimates, will be suitably quantified in the next section.

5.2. One step of the iterative procedure. In this section we apply one correction of the quasi-Newton method described in Section 5.1 and we obtain sharp quantitative estimates for the new approximately invariant torus and related objects. We set sufficient conditions to preserve the control of the previous estimates.

Lemma 5.2 (The Iterative Lemma in the ordinary case). *Let us consider the same setting and hypotheses of Theorem 3.1, and a constant $\nu > 0$. Then, there exist constants $C_{\Delta K}$, $C_{\Delta B}$, $C_{\Delta \langle T \rangle^{-1}}$ and C_E such that if the inequalities*

$$(5.8) \quad \frac{\hat{C}_\Delta \|E\|_\rho}{\gamma^2 \delta^{2\tau+1}} < 1, \quad \frac{C_E \|E\|_\rho}{\gamma^4 \delta^{4\tau}} < 1$$

hold for some $0 < \delta < \rho$, where

$$(5.9) \quad \hat{C}_\Delta := \max \left\{ \frac{\gamma^2 \delta^{2\tau}}{\nu}, 2C_{\text{sym}} \gamma \delta^\tau, \frac{dC_{\Delta K}}{\sigma_K - \|DK\|_\rho}, \frac{2nC_{\Delta K}}{\sigma_{K^\top} - \|(DK)^\top\|_\rho}, \right. \\ \left. \frac{C_{\Delta B}}{\sigma_B - \|B\|_\rho}, \frac{C_{\Delta \langle T \rangle^{-1}}}{\sigma_T - |\langle T \rangle^{-1}|}, \frac{C_{\Delta K} \delta}{\text{dist}(K(\mathbb{T}_\rho^d), \partial B)} \right\},$$

then we have an approximate torus of the same frequency ω given by $\bar{K} = K + \Delta K$, with components in $\mathcal{A}(\mathbb{T}_{\rho-2\delta}^d)$, that defines new objects \bar{B} and \bar{T} (obtained replacing K by \bar{K}) satisfying

$$(5.10) \quad \|D\bar{K}\|_{\rho-3\delta} < \sigma_K,$$

$$(5.11) \quad \|(D\bar{K})^\top\|_{\rho-3\delta} < \sigma_{K^\top},$$

$$(5.12) \quad \|\bar{B}\|_{\rho-3\delta} < \sigma_B,$$

$$(5.13) \quad |\langle \bar{T} \rangle^{-1}| < \sigma_T,$$

$$(5.14) \quad \text{dist}(\bar{K}(\mathbb{T}_{\rho-2\delta}^d), \partial B) > 0,$$

and

$$(5.15) \quad \|\bar{K} - K\|_{\rho-2\delta} < \frac{C_{\Delta K}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho,$$

$$(5.16) \quad \|\bar{B} - B\|_{\rho-3\delta} < \frac{C_{\Delta B}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho,$$

$$(5.17) \quad |\langle \bar{T} \rangle^{-1} - \langle T \rangle^{-1}| < \frac{C_{\Delta \langle T \rangle^{-1}}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.$$

The new error of invariance is given by

$$\bar{E}(\theta) = X_h(\bar{K}(\theta)) + \mathfrak{L}_\omega \bar{K}(\theta),$$

and satisfies

$$(5.18) \quad \|\bar{E}\|_{\rho-2\delta} < \frac{C_E}{\gamma^4 \delta^{4\tau}} \|E\|_\rho^2.$$

The above constants are collected in Table 2.

Proof. This result requires rather cumbersome computations, so we divide the proof into several steps.

Step 1: Control of the new parameterization. We start by considering the new parameterization $\bar{K}(\theta) = K(\theta) + \Delta K(\theta)$ obtained from the system (5.5), with $\eta(\theta)$ given by (5.6). We choose the solution that satisfies $\xi_0^L = 0$. Using the estimates obtained in Section 4 we have

$$\|\eta^L\|_\rho \leq C_{N^\top} c_{\Omega,0} \|E\|_\rho, \quad \|\eta^N\|_\rho \leq C_{L^\top} c_{\Omega,0} \|E\|_\rho.$$

In order to invoke Lemma 4.8 (we must fulfill condition (4.40)) we have included the inequality

$$(5.19) \quad \frac{\|E\|_\rho}{\delta} < \nu$$

into Hypothesis (5.8) (this corresponds to the first term in (5.9)). Hence, combining Lemma 4.8 and Lemma 5.1, we obtain estimates for the solution of the cohomological equations (we recall that $\xi_0^L = 0_n$)

$$(5.20) \quad \begin{aligned} |\xi_0^N| &\leq |\langle T \rangle^{-1}| \left(\|\eta^L\|_\rho + \frac{c_R}{\gamma\delta^\tau} \|T\|_{\rho-\delta} \|\eta^N\|_\rho \right) \\ &\leq \sigma_T \left(C_{N^\top} c_{\Omega,0} + \frac{c_R}{\gamma\delta^\tau} C_T C_{L^\top} c_{\Omega,0} \right) \|E\|_\rho =: \frac{C_{\xi_0^N}}{\gamma\delta^\tau} \|E\|_\rho, \end{aligned}$$

$$(5.21) \quad \begin{aligned} \|\xi^N\|_{\rho-\delta} &\leq |\xi_0^N| + \frac{c_R}{\gamma\delta^\tau} \|\eta^N\|_\rho \\ &\leq \frac{C_{\xi_0^N}}{\gamma\delta^\tau} \|E\|_\rho + \frac{c_R}{\gamma\delta^\tau} C_{L^\top} c_{\Omega,0} \|E\|_\rho =: \frac{C_{\xi^N}}{\gamma\delta^\tau} \|E\|_\rho, \end{aligned}$$

$$(5.22) \quad \begin{aligned} \|\xi^L\|_{\rho-2\delta} &\leq |\xi_0^L| + \frac{c_R}{\gamma\delta^\tau} \left(\|\eta^L\|_\rho + \|T\|_{\rho-\delta} \|\xi^N\|_{\rho-\delta} \right) \\ &\leq \frac{c_R}{\gamma\delta^\tau} \left(C_{N^\top} c_{\Omega,0} + C_T \frac{C_{\xi^N}}{\gamma\delta^\tau} \right) \|E\|_\rho =: \frac{C_{\xi^L}}{\gamma^2\delta^{2\tau}} \|E\|_\rho. \end{aligned}$$

The norm of the full vector $\xi(\theta)$, which satisfies (5.7), is controlled as

$$(5.23) \quad \begin{aligned} \|\xi\|_{\rho-2\delta} &\leq \max\{\|\xi^L\|_{\rho-2\delta}, \|\xi^N\|_{\rho-\delta}\} \\ &\leq \frac{\max\{C_{\xi^L}, C_{\xi^N}\gamma\delta^\tau\}}{\gamma^2\delta^{2\tau}} \|E\|_\rho =: \frac{C_\xi}{\gamma^2\delta^{2\tau}} \|E\|_\rho. \end{aligned}$$

The new parameterization $\bar{K}(\theta)$ and the related objects are controlled using standard computations. Estimate (5.15) follows directly from

$$\bar{K}(\theta) - K(\theta) = \Delta K(\theta) = P(\theta)\xi(\theta) = L(\theta)\xi^L(\theta) + N(\theta)\xi^N(\theta),$$

that is, using estimates in (5.21) and (5.22), we obtain

$$(5.24) \quad \|\bar{K} - K\|_{\rho-2\delta} \leq \frac{C_L C_{\xi^L} + C_N C_{\xi^N} \gamma\delta^\tau}{\gamma^2\delta^{2\tau}} \|E\|_\rho =: \frac{C_{\Delta K}}{\gamma^2\delta^{2\tau}} \|E\|_\rho.$$

To complete this step, we check that $\bar{K}(\theta)$ remains inside the domain \mathcal{B} where the global objects are defined. This is important because we need to estimate the new error $\bar{E}(\theta)$ before controlling the remaining geometrical objects. Then, we observe that

$$(5.25) \quad \begin{aligned} \text{dist}(\bar{K}(\mathbb{T}_{\rho-2\delta}^d), \partial\mathcal{B}) &\geq \text{dist}(K(\mathbb{T}_\rho^d), \partial\mathcal{B}) - \|\Delta K\|_{\rho-2\delta} \\ &\geq \text{dist}(K(\mathbb{T}_\rho^d), \partial\mathcal{B}) - \frac{C_{\Delta K}}{\gamma^2\delta^{2\tau}} \|E\|_\rho > 0, \end{aligned}$$

where the last inequality follows from Hypothesis (5.8) (this corresponds to the seventh term in (5.9)). We have obtained the control in (5.14).

Step 2: Control of the new error of invariance. To control the error of invariance of the corrected parameterization \bar{K} , we first consider the error in the solution of the linearized equation (5.4), that is, we control the quadratic terms that are neglected when considering the equation (5.7):

$$E_{\text{lin}}(\theta) = E_{\text{red}}(\theta)\xi(\theta) - \Omega_0 E_{\text{sym}}(\theta)\mathfrak{L}_\omega \xi(\theta).$$

The term $\mathfrak{L}_\omega \xi(\theta)$ is controlled using that $\xi(\theta)$ is precisely the solution of the cohomological equation (5.7):

$$(5.26) \quad \|\mathfrak{L}_\omega \xi^N\|_\rho = \|\eta^N\|_\rho \leq C_{L^\top} c_{\Omega,0} \|E\|_\rho =: C_{\mathfrak{L}_\omega \xi^N} \|E\|_\rho,$$

$$\begin{aligned}
\|\mathfrak{L}_\omega \xi^L\|_{\rho-\delta} &= \|\eta^L - T\xi^N\|_{\rho-\delta} \leq \left(C_{N^\top} c_{\Omega,0} + C_T \frac{C_{\xi^N}}{\gamma\delta^\tau} \right) \|E\|_\rho \\
(5.27) \quad &=: \frac{C_{\mathfrak{L}\xi^L}}{\gamma\delta^\tau} \|E\|_\rho,
\end{aligned}$$

$$(5.28) \quad \|\mathfrak{L}_\omega \xi\|_{\rho-\delta} \leq \max \left(\frac{C_{\mathfrak{L}\xi^L}}{\gamma\delta^\tau}, C_{\mathfrak{L}\xi^N} \right) \|E\|_\rho =: \frac{C_{\mathfrak{L}\xi}}{\gamma\delta^\tau} \|E\|_\rho.$$

Hence, we control $E_{\text{lin}}(\theta)$ by

$$\begin{aligned}
\|E_{\text{lin}}\|_{\rho-2\delta} &\leq \|E_{\text{red}}\|_{\rho-2\delta} \|\xi\|_{\rho-2\delta} + c_{\Omega,0} \|E_{\text{sym}}\|_{\rho-2\delta} \|\mathfrak{L}_\omega \xi\|_{\rho-2\delta} \\
(5.29) \quad &\leq \frac{C_{\text{red}}}{\gamma\delta^{\tau+1}} \frac{C_\xi}{\gamma^2\delta^{2\tau}} \|E\|_\rho^2 + c_{\Omega,0} \frac{C_{\text{sym}}}{\gamma\delta^{\tau+1}} \frac{C_{\mathfrak{L}\xi}}{\gamma\delta^\tau} \|E\|_\rho^2 =: \frac{C_{\text{lin}}}{\gamma^3\delta^{3\tau+1}} \|E\|_\rho^2.
\end{aligned}$$

We remark that this last estimate can be improved by considering the components of $\xi(\theta) = (\xi^L(\theta), \xi^N(\theta))$ separately, thus obtaining a divisor $\gamma^2\delta^{2\tau+1}$ in (5.29). Nevertheless, this improvement is irrelevant for practical purposes.

After performing the correction, the error of invariance associated with the new parameterization is given by

$$\begin{aligned}
\bar{E}(\theta) &= X_h(K(\theta) + \Delta K(\theta)) + \mathfrak{L}_\omega K(\theta) + \mathfrak{L}_\omega \Delta K(\theta) \\
&= X_h(K(\theta)) + DX_h(K(\theta))\Delta K(\theta) + \mathfrak{L}_\omega K(\theta) + \mathfrak{L}_\omega \Delta K(\theta) + \Delta^2 X(\theta) \\
(5.30) \quad &= DX_h(K(\theta))\Delta K(\theta) + \mathfrak{L}_\omega \Delta K(\theta) + E(\theta) + \Delta^2 X(\theta) \\
&= (DX_h(K(\theta))P(\theta) + \mathfrak{L}_\omega P(\theta))\xi(\theta) + P(\theta)\mathfrak{L}_\omega \xi(\theta) + E(\theta) + \Delta^2 X(\theta) \\
&= (-\Omega_0 P(\theta)^\top \Omega(K(\theta)))^{-1} E_{\text{lin}}(\theta) + \Delta^2 X(\theta) \\
&= P(\theta)(I_{2n} - \Omega_0 E_{\text{sym}}(\theta))^{-1} E_{\text{lin}}(\theta) + \Delta^2 X(\theta),
\end{aligned}$$

where

$$\begin{aligned}
\Delta^2 X(\theta) &= X_h(K(\theta) + \Delta K(\theta)) - X_h(K(\theta)) - DX_h(K(\theta))\Delta K(\theta) \\
(5.31) \quad &= \int_0^1 (1-t) D^2 X_h(K(\theta) + t\Delta K(\theta)) [\Delta K(\theta), \Delta K(\theta)] dt,
\end{aligned}$$

and we used (4.29). Notice that the above error function is well defined, due to the computations in (5.25), and we estimate its norm as follows

$$\|\bar{E}\|_{\rho-2\delta} \leq \|P\|_{\rho-2\delta} \|(I - \Omega_0 E_{\text{sym}})^{-1}\|_{\rho-2\delta} \|E_{\text{lin}}\|_{\rho-2\delta} + \|\Delta^2 X\|_{\rho-2\delta}.$$

Then, using a Neumann series argument, we obtain

$$\|(I - \Omega_0 E_{\text{sym}})^{-1}\|_{\rho-2\delta} \leq \frac{1}{1 - \|\Omega_0 E_{\text{sym}}\|_{\rho-2\delta}} < 2,$$

where we used the inequality

$$\frac{C_{\text{sym}}}{\gamma\delta^{\tau+1}} \|E\|_\rho \leq \frac{1}{2},$$

that corresponds to the second term in (5.9) (Hypothesis (5.8)). Putting together the above estimates, and applying the mean value theorem to control $\Delta^2 X(\theta)$, we obtain

$$(5.32) \quad \|\bar{E}\|_{\rho-2\delta} \leq \left(\frac{2(C_L + C_N)C_{\text{lin}}}{\gamma^3\delta^{3\tau+1}} + \frac{1}{2} c_{X_h,2} \frac{(C_{\Delta K})^2}{\gamma^4\delta^{4\tau}} \right) \|E\|_\rho^2 =: \frac{C_E}{\gamma^4\delta^{4\tau}} \|E\|_\rho^2.$$

We have obtained the estimate (5.18). Notice that the second assumption in (5.8) and (5.19) imply that

$$(5.33) \quad \|\bar{E}\|_{\rho-2\delta} < \|E\|_{\rho} < \delta\nu.$$

This will be used in Step 6.

Step 3: Control of the new frame $L(\theta)$. Combining (5.24) with Cauchy estimates, we obtain the control (5.10):

$$(5.34) \quad \|\mathbf{D}\bar{K}\|_{\rho-3\delta} \leq \|\mathbf{D}K\|_{\rho} + \|\mathbf{D}\Delta K\|_{\rho-3\delta} \leq \|\mathbf{D}K\|_{\rho} + \frac{dC_{\Delta K}}{\gamma^2\delta^{2\tau+1}}\|E\|_{\rho} < \sigma_K,$$

where the last inequality follows from Hypothesis (5.8) (this corresponds to the third term in (5.9)). The control (5.11) on the transposed object is analogous

$$(5.35) \quad \|(\mathbf{D}\bar{K})^{\top}\|_{\rho-3\delta} \leq \|(\mathbf{D}K)^{\top}\|_{\rho} + \frac{2nC_{\Delta K}}{\gamma^2\delta^{2\tau+1}}\|E\|_{\rho} < \sigma_{K^{\top}},$$

where the last inequality follows from Hypothesis (5.8) (this corresponds to the fourth term in (5.9)).

After obtaining the estimates (5.34) and (5.35), it is clear that

$$\|\bar{L}\|_{\rho-3\delta} \leq C_L, \quad \|\bar{L}^{\top}\|_{\rho-3\delta} \leq C_{L^{\top}}.$$

Indeed, we can control the norm of the corresponding corrections using Cauchy estimates, the mean value theorem and estimate (5.24):

$$(5.36) \quad \|\bar{L} - L\|_{\rho-3\delta} \leq \frac{dC_{\Delta K}}{\gamma^2\delta^{2\tau+1}}\|E\|_{\rho} + \frac{c_{X_p,1}C_{\Delta K}}{\gamma^2\delta^{2\tau}}\|E\|_{\rho} =: \frac{C_{\Delta L}}{\gamma^2\delta^{2\tau+1}}\|E\|_{\rho},$$

$$(5.37) \quad \|\bar{L}^{\top} - L^{\top}\|_{\rho-3\delta} \leq \frac{C_{\Delta K} \max\{2n, c_{X_p^{\top},1}\delta\}}{\gamma^2\delta^{2\tau+1}}\|E\|_{\rho} =: \frac{C_{\Delta L^{\top}}}{\gamma^2\delta^{2\tau+1}}\|E\|_{\rho}.$$

Step 4: Control of the new transversality condition. To control \bar{B} we use Lemma A.1 taking

$$\begin{aligned} M(\theta) &= G_L(\theta) = L(\theta)^{\top}G(K(\theta))L(\theta), \\ \bar{M}(\theta) &= G_{\bar{L}}(\theta) = \bar{L}(\theta)^{\top}G(\bar{K}(\theta))\bar{L}(\theta), \end{aligned}$$

where we have used the notation introduced in (4.19). First, we compute

$$(5.38) \quad \begin{aligned} \|G \circ \bar{K} - G \circ K\|_{\rho-2\delta} &\leq \|DG\|_{\mathcal{B}}\|\bar{K} - K\|_{\rho-2\delta} \leq \frac{c_{G,1}C_{\Delta K}}{\gamma^2\delta^{2\tau}}\|E\|_{\rho} \\ &=: \frac{C_{\Delta G}}{\gamma^2\delta^{2\tau}}\|E\|_{\rho}, \end{aligned}$$

and

$$\begin{aligned} \|G_{\bar{L}} - G_L\|_{\rho-3\delta} &\leq \|\bar{L}^{\top}(G \circ \bar{K})\bar{L} - \bar{L}^{\top}(G \circ \bar{K})L\|_{\rho-3\delta} \\ &\quad + \|\bar{L}^{\top}(G \circ \bar{K})L - \bar{L}^{\top}(G \circ K)L\|_{\rho-3\delta} \\ &\quad + \|\bar{L}^{\top}(G \circ K)L - L^{\top}(G \circ K)L\|_{\rho-3\delta} \\ &\leq \|\bar{L}^{\top}\|_{\rho}\|G\|_{\mathcal{B}}\|\bar{L} - L\|_{\rho-3\delta} \\ &\quad + \|\bar{L}^{\top}\|_{\rho}\|G \circ \bar{K} - G \circ K\|_{\rho-3\delta}\|L\|_{\rho} \\ &\quad + \|\bar{L}^{\top} - L^{\top}\|_{\rho-3\delta}\|G\|_{\mathcal{B}}\|L\|_{\rho} \end{aligned}$$

$$\begin{aligned}
&\leq \frac{C_{L^\top} c_{G,0} C_{\Delta L} + C_{L^\top} C_{\Delta G} C_L \delta + C_{\Delta L^\top} c_{G,0} C_L}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
(5.39) \quad &=: \frac{C_{\Delta G_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.
\end{aligned}$$

Then, we introduce the constant

$$C_{\Delta B} := 2(\sigma_B)^2 C_{\Delta G_L}$$

and check the condition (A.1) in Lemma A.1:

$$\begin{aligned}
(5.40) \quad \frac{2(\sigma_B)^2 \|G_{\bar{L}} - G_L\|_{\rho-3\delta}}{\sigma_B - \|B\|_\rho} &\leq \frac{2(\sigma_B)^2 C_{\Delta G_L}}{\sigma_B - \|B\|_\rho} \frac{\|E\|_\rho}{\gamma^2 \delta^{2\tau+1}} \\
&= \frac{C_{\Delta B}}{\sigma_B - \|B\|_\rho} \frac{\|E\|_\rho}{\gamma^2 \delta^{2\tau+1}} < 1,
\end{aligned}$$

where the last inequality follows from Hypothesis (5.8) (this corresponds to the fifth term in (5.9)). Hence, by invoking Lemma A.1, we conclude that

$$(5.41) \quad \|\bar{B}\|_{\rho-3\delta} < \sigma_B, \quad \|\bar{B} - B\|_{\rho-3\delta} \leq \frac{2(\sigma_B)^2 C_{\Delta G_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho = \frac{C_{\Delta B}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho,$$

and so, we obtain the estimates (5.12) and (5.16) on the new object.

Step 5: Control of the new frame $N(\theta)$. To control the new adapted normal frame $\bar{N}(\theta)$, it is convenient to recall the following notation:

$$\begin{aligned}
N(\theta) &= L(\theta)A(\theta) + N^0(\theta)B(\theta), \\
N^0(\theta) &= J(K(\theta))L(\theta), \\
A(\theta) &= -\frac{1}{2}(B(\theta)^\top L(\theta)^\top \tilde{\Omega}(K(\theta))L(\theta)B(\theta)), \\
B(\theta) &= (L(\theta)^\top G(K(\theta))L(\theta))^{-1},
\end{aligned}$$

where, as usual, the new objects $\bar{N}(\theta)$, $\bar{A}(\theta)$, $\bar{N}^0(\theta)$ and $\bar{B}(\theta)$ are obtained by replacing $K(\theta)$ by $\bar{K}(\theta)$. Note that the object $\bar{B}(\theta)$ has been controlled in Step 4.

Now, we recall the notation introduced in (4.20) and reproduce the computations in (5.38) and (5.39) for the matrix functions

$$\begin{aligned}
\tilde{\Omega}_L(\theta) &= L(\theta)^\top \tilde{\Omega}(K(\theta))L(\theta), \\
\tilde{\Omega}_{\bar{L}}(\theta) &= \bar{L}(\theta)^\top \tilde{\Omega}(\bar{K}(\theta))\bar{L}(\theta),
\end{aligned}$$

thus obtaining

$$(5.42) \quad \|\tilde{\Omega} \circ \bar{K} - \tilde{\Omega} \circ K\|_{\rho-2\delta} \leq \frac{c_{\tilde{\Omega},1} C_{\Delta K}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho =: \frac{C_{\Delta \tilde{\Omega}}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho,$$

and

$$\begin{aligned}
(5.43) \quad \|\tilde{\Omega}_{\bar{L}} - \tilde{\Omega}_L\|_{\rho-3\delta} &\leq \frac{C_{L^\top} c_{\tilde{\Omega},0} C_{\Delta L} + C_{L^\top} C_{\Delta \tilde{\Omega}} C_L \delta + C_{\Delta L^\top} c_{\tilde{\Omega},0} C_L}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
&=: \frac{C_{\Delta \tilde{\Omega}_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.
\end{aligned}$$

Now, we control the matrix $\bar{A}(\theta)$ as follows

$$\|\bar{A} - A\|_{\rho-3\delta} \leq \frac{1}{2} \|\bar{B}^\top \tilde{\Omega}_{\bar{L}} \bar{B} - \bar{B}^\top \tilde{\Omega}_L B\|_{\rho-3\delta}$$

$$\begin{aligned}
& + \frac{1}{2} \|\bar{B}^\top \tilde{\Omega}_L B - \bar{B}^\top \tilde{\Omega}_L B\|_{\rho-3\delta} + \frac{1}{2} \|\bar{B}^\top \tilde{\Omega}_L B - B^\top \tilde{\Omega}_L B\|_{\rho-3\delta} \\
& \leq \sigma_B C_{\tilde{\Omega}_L} \|\bar{B} - B\|_{\rho-3\delta} + \frac{1}{2} (\sigma_B)^2 \|\tilde{\Omega}_L - \tilde{\Omega}_L\|_{\rho-3\delta} \\
(5.44) \quad & \leq \frac{\sigma_B C_{\tilde{\Omega}_L} C_{\Delta B} + \frac{1}{2} (\sigma_B)^2 C_{\Delta \tilde{\Omega}_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho =: \frac{C_{\Delta A}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho,
\end{aligned}$$

where we used that $B(\theta)^\top = B(\theta)$, and the constants (4.27), (5.41) and (5.43). The same control holds for $\|\bar{A}^\top - A^\top\|_{\rho-3\delta}$, since we have that $A(\theta)^\top = -A(\theta)$. We notice that $C_{\Delta A} = 0$ in **Case III**.

Analogous computations yield to

$$(5.45) \quad \|J \circ \bar{K} - J \circ K\|_{\rho-2\delta} \leq \frac{c_{J,1} C_{\Delta K}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho =: \frac{C_{\Delta J}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho,$$

$$(5.46) \quad \|(J \circ \bar{K} - J \circ K)^\top\|_{\rho-2\delta} \leq \frac{c_{J^\top,1} C_{\Delta K}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho =: \frac{C_{\Delta J^\top}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho,$$

$$\begin{aligned}
& \|\bar{N}^0 - N^0\|_{\rho-3\delta} \leq \|J|_{\mathcal{B}} \|\bar{L} - L\|_{\rho-3\delta} + \|J \circ \bar{K} - J \circ K\|_{\rho-2\delta} \|L\|_\rho \\
& \leq \frac{c_{J,0} C_{\Delta L} + C_{\Delta J} C_L \delta}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
(5.47) \quad & =: \frac{C_{\Delta N^0}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho,
\end{aligned}$$

$$\begin{aligned}
& \|(\bar{N}^0)^\top - (N^0)^\top\|_{\rho-3\delta} \leq \frac{C_{\Delta L^\top} c_{J^\top,0} + C_{L^\top} C_{\Delta J^\top} \delta}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
(5.48) \quad & =: \frac{C_{\Delta N^{0\top}}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.
\end{aligned}$$

Finally, we control the correction of the adapted normal frame:

$$\begin{aligned}
& \|\bar{N} - N\|_{\rho-3\delta} \\
& \leq \|\bar{L}\bar{A} - \bar{L}A\|_{\rho-3\delta} + \|\bar{L}A - LA\|_{\rho-3\delta} \\
& \quad + \|\bar{N}^0 \bar{B} - \bar{N}^0 B\|_{\rho-3\delta} + \|\bar{N}^0 B - N^0 B\|_{\rho-3\delta} \\
& \leq \frac{C_L C_{\Delta A} + C_{\Delta L} C_A + C_{N^0} C_{\Delta B} + C_{\Delta N^0} \sigma_B}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
(5.49) \quad & =: \frac{C_{\Delta N}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho,
\end{aligned}$$

$$\begin{aligned}
& \|\bar{N}^\top - N^\top\|_{\rho-3\delta} \\
& \leq \frac{C_A C_{\Delta L^\top} + C_{\Delta A} C_{L^\top} + C_{\Delta B} C_{N^{0,\top}} + \sigma_B C_{\Delta N^{0\top}}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
(5.50) \quad & =: \frac{C_{\Delta N^\top}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.
\end{aligned}$$

Step 6: Control of the action of the left operator. It is worth mentioning that an important effort is made to obtain optimal estimates for the twist condition. As it was illustrated in the proof of Lemma 4.8, improved estimates are obtained by avoiding the use of Cauchy estimates when controlling the action of \mathfrak{L}_ω on different objects and their corrections.

Using the assumption (5.19) and the control (5.33) of the new error of invariance, we preserve the control for the new objects $\mathfrak{L}_\omega \bar{K}(\theta)$, $\mathfrak{L}_\omega \bar{L}(\theta)$ and $\mathfrak{L}_\omega \bar{L}(\theta)^\top$. Indeed,

using (5.33), we have

$$\begin{aligned}\|\mathfrak{L}_\omega \bar{K}\|_{\rho-2\delta} &\leq \|\bar{E}\|_\rho + \|X_h \circ \bar{K}\|_\rho \leq \delta\nu + c_{X_h,0} = C_{\mathfrak{L}K}, \\ \|\mathfrak{L}_\omega \bar{L}\|_{\rho-3\delta} &\leq d\nu + c_{X_h,1}\sigma_K + c_{X_p,1}C_{\mathfrak{L}K} = C_{\mathfrak{L}L}, \\ \|\mathfrak{L}_\omega \bar{L}^\top\|_{\rho-3\delta} &\leq \max\left\{2n\nu + c_{X_h^\top,1}\sigma_K, c_{X_p^\top,1}C_{\mathfrak{L}K}\right\} = C_{\mathfrak{L}L^\top}.\end{aligned}$$

Then, we also control the action of \mathfrak{L}_ω on the correction of the torus, using that

$$\begin{aligned}\mathfrak{L}_\omega \bar{K}(\theta) - \mathfrak{L}_\omega K(\theta) &= \mathfrak{L}_\omega \Delta K(\theta) \\ &= \mathfrak{L}_\omega L(\theta)\xi^L(\theta) + L(\theta)\mathfrak{L}_\omega \xi^L(\theta) + \mathfrak{L}_\omega N(\theta)\xi^N(\theta) + N(\theta)\mathfrak{L}_\omega \xi^N(\theta)\end{aligned}$$

and, recalling previous estimates, we obtain:

$$\begin{aligned}\|\mathfrak{L}_\omega \bar{K} - \mathfrak{L}_\omega K\|_{\rho-2\delta} &\leq \left(\frac{C_{\mathfrak{L}L}C_{\xi^L}}{\gamma^2\delta^{2\tau}} + \frac{C_L C_{\xi^L}}{\gamma\delta^\tau} + \frac{C_{\mathfrak{L}N}C_{\xi^N}}{\gamma\delta^\tau} + C_N C_{\xi^N}\right) \|E\|_\rho \\ (5.51) \quad &=: \frac{C_{\Delta \mathfrak{L}K}}{\gamma^2\delta^{2\tau}} \|E\|_\rho.\end{aligned}$$

The action of \mathfrak{L}_ω on the correction of $L(\theta)$ is similar. On the one hand, we have

$$\begin{aligned}\|\mathfrak{L}_\omega \bar{L} - \mathfrak{L}_\omega L\|_{\rho-3\delta} &\leq \|\mathfrak{L}_\omega(D\bar{K} - DK)\|_{\rho-3\delta} + \|\mathfrak{L}_\omega(X_p \circ \bar{K} - X_p \circ K)\|_{\rho-3\delta} \\ &\leq \|D(\mathfrak{L}_\omega \bar{K} - \mathfrak{L}_\omega K)\|_{\rho-3\delta} \\ &\quad + \|(DX_p \circ \bar{K})[\mathfrak{L}_\omega \bar{K}] - (DX_p \circ \bar{K})[\mathfrak{L}_\omega K]\|_{\rho-3\delta} \\ &\quad + \|(DX_p \circ \bar{K})[\mathfrak{L}_\omega K] - (DX_p \circ K)[\mathfrak{L}_\omega K]\|_{\rho-3\delta} \\ &\leq \frac{dC_{\Delta \mathfrak{L}K}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho + \frac{c_{X_p,1}C_{\Delta \mathfrak{L}K}}{\gamma^2\delta^{2\tau}} \|E\|_\rho + \frac{c_{X_p,2}C_{\Delta K}C_{\mathfrak{L}K}}{\gamma^2\delta^{2\tau}} \|E\|_\rho \\ (5.52) \quad &=: \frac{C_{\Delta \mathfrak{L}L}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho,\end{aligned}$$

and on the other hand, we have

$$\begin{aligned}\|\mathfrak{L}_\omega \bar{L}^\top - \mathfrak{L}_\omega L^\top\|_{\rho-3\delta} &\leq \max\{\|\mathfrak{L}_\omega(D\bar{K} - DK)^\top\|_{\rho-3\delta}, \|\mathfrak{L}_\omega(X_p \circ \bar{K} - X_p \circ K)^\top\|_{\rho-3\delta}\} \\ &\leq \max\left\{\frac{2nC_{\Delta \mathfrak{L}K}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho, \frac{c_{X_p^\top,1}C_{\Delta \mathfrak{L}K}}{\gamma^2\delta^{2\tau}} \|E\|_\rho + \frac{c_{X_p^\top,2}C_{\Delta K}C_{\mathfrak{L}K}}{\gamma^2\delta^{2\tau}} \|E\|_\rho\right\} \\ (5.53) \quad &=: \frac{C_{\Delta \mathfrak{L}L^\top}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho.\end{aligned}$$

To control the action of \mathfrak{L}_ω on the correction of the matrix $B(\theta)$ that provides the transversality condition, we first consider the correction

$$\begin{aligned}\|\mathfrak{L}_\omega(G \circ \bar{K}) - \mathfrak{L}_\omega(G \circ K)\|_{\rho-2\delta} &\leq \|DG\|_{\mathcal{B}} \|\mathfrak{L}_\omega \bar{K} - \mathfrak{L}_\omega K\|_{\rho-2\delta} \\ &\quad + \|D^2G\|_{\mathcal{B}} \|\bar{K} - K\|_{\rho-2\delta} \|\mathfrak{L}_\omega K\|_{\rho-\delta} \\ (5.54) \quad &\leq \frac{c_{G,1}C_{\Delta \mathfrak{L}K} + c_{G,2}C_{\Delta K}C_{\mathfrak{L}K}}{\gamma^2\delta^{2\tau}} \|E\|_\rho =: \frac{C_{\Delta \mathfrak{L}G}}{\gamma^2\delta^{2\tau}} \|E\|_\rho,\end{aligned}$$

and we control the correction of the adapted metric

$$\begin{aligned}
(5.55) \quad \mathfrak{L}_\omega G_{\bar{L}}(\theta) - \mathfrak{L}_\omega G_L(\theta) &= \mathfrak{L}_\omega \bar{L}^\top(\theta) G(\bar{K}(\theta)) \bar{L}(\theta) - \mathfrak{L}_\omega L^\top(\theta) G(K(\theta)) L(\theta) \\
&\quad + \bar{L}^\top(\theta) \mathfrak{L}_\omega G(\bar{K}(\theta)) \bar{L}(\theta) - L^\top(\theta) \mathfrak{L}_\omega G(K(\theta)) L(\theta) \\
&\quad + \bar{L}^\top(\theta) G(\bar{K}(\theta)) \mathfrak{L}_\omega \bar{L}(\theta) - L^\top(\theta) G(K(\theta)) \mathfrak{L}_\omega L(\theta)
\end{aligned}$$

as follows

$$\begin{aligned}
(5.56) \quad &\| \mathfrak{L}_\omega G_{\bar{L}} - \mathfrak{L}_\omega G_L \|_{\rho-3\delta} \\
&\leq \frac{C_{\mathfrak{L}L^\top c_{G,0} C_{\Delta L}} + C_{\mathfrak{L}L^\top c_{G,1} C_{\Delta K} C_L \delta} + C_{\Delta \mathfrak{L}L^\top c_{G,0} C_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
&\quad + \frac{C_{L^\top c_{G,1} C_{\mathfrak{L}K} C_{\Delta L}} + C_{L^\top C_{\Delta \mathfrak{L}G} C_L \delta} + C_{\Delta L^\top c_{G,1} C_{\mathfrak{L}K} C_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
&\quad + \frac{C_{L^\top c_{G,0} C_{\Delta \mathfrak{L}L}} + C_{L^\top c_{G,1} C_{\Delta K} C_{\mathfrak{L}L} \delta} + C_{\Delta L^\top c_{G,0} C_{\mathfrak{L}L}}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
&=: \frac{C_{\Delta \mathfrak{L}G_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.
\end{aligned}$$

Moreover, the following estimates (borrowed from Lemma 4.8) will be also useful

$$\| \mathfrak{L}_\omega G_L \|_{\rho-\delta} \leq C_{\mathfrak{L}G_L}, \quad \| \mathfrak{L}_\omega G_{\bar{L}} \|_{\rho-3\delta} \leq C_{\mathfrak{L}G_L}.$$

Again, it is clear that this control is preserved for the corrected objects. Using the previous estimates, we have

$$\begin{aligned}
(5.57) \quad &\| \mathfrak{L}_\omega \bar{B} - \mathfrak{L}_\omega B \|_{\rho-3\delta} \leq \| \bar{B} \mathfrak{L}_\omega G_{\bar{L}} \bar{B} - \bar{B} \mathfrak{L}_\omega G_L B \|_{\rho-3\delta} \\
&\quad + \| \bar{B} \mathfrak{L}_\omega G_{\bar{L}} B - \bar{B} \mathfrak{L}_\omega G_L B \|_{\rho-3\delta} + \| \bar{B} \mathfrak{L}_\omega G_L B - B \mathfrak{L}_\omega G_L B \|_{\rho-3\delta} \\
&\leq \frac{2\sigma_B C_{\mathfrak{L}G_L} C_{\Delta B} + (\sigma_B)^2 C_{\Delta \mathfrak{L}G_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho =: \frac{C_{\Delta \mathfrak{L}B}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.
\end{aligned}$$

By repeating the computations (5.54), (5.55) and (5.56), mutatis mutandis, we obtain the estimates

$$\begin{aligned}
(5.58) \quad &\| \mathfrak{L}_\omega(\tilde{\Omega} \circ \bar{K}) - \mathfrak{L}_\omega(\tilde{\Omega} \circ K) \|_{\rho-2\delta} \\
&\leq \frac{c_{\tilde{\Omega},1} C_{\Delta \mathfrak{L}K} + c_{\tilde{\Omega},2} C_{\Delta K} C_{\mathfrak{L}K}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho =: \frac{C_{\Delta \mathfrak{L}\tilde{\Omega}}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho,
\end{aligned}$$

and

$$\begin{aligned}
(5.59) \quad &\| \mathfrak{L}_\omega \tilde{\Omega}_{\bar{L}} - \mathfrak{L}_\omega \tilde{\Omega}_L \|_{\rho-3\delta} \\
&\leq \frac{C_{\mathfrak{L}L^\top c_{\tilde{\Omega},0} C_{\Delta L}} + C_{\mathfrak{L}L^\top c_{\tilde{\Omega},1} C_{\Delta K} C_L \delta} + C_{\Delta \mathfrak{L}L^\top c_{\tilde{\Omega},0} C_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
&\quad + \frac{C_{L^\top c_{\tilde{\Omega},1} C_{\mathfrak{L}K} C_{\Delta L}} + C_{L^\top C_{\Delta \mathfrak{L}\tilde{\Omega}} C_L \delta} + C_{\Delta L^\top c_{\tilde{\Omega},1} C_{\mathfrak{L}K} C_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
&\quad + \frac{C_{L^\top c_{\tilde{\Omega},0} C_{\Delta \mathfrak{L}L}} + C_{L^\top c_{\tilde{\Omega},1} C_{\Delta K} C_{\mathfrak{L}L} \delta} + C_{\Delta L^\top c_{\tilde{\Omega},0} C_{\mathfrak{L}L}}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
&=: \frac{C_{\Delta \mathfrak{L}\tilde{\Omega}_L}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.
\end{aligned}$$

We also recall the following controls

$$\| \mathfrak{L}_\omega \tilde{\Omega}_L \|_{\rho-\delta} \leq C_{\mathfrak{L}\tilde{\Omega}_L}, \quad \| \mathfrak{L}_\omega \tilde{\Omega}_{\bar{L}} \|_{\rho-3\delta} \leq C_{\mathfrak{L}\tilde{\Omega}_L}.$$

Finally, we obtain

$$\begin{aligned}
& \|\mathfrak{L}_\omega \bar{A} - \mathfrak{L}_\omega A\|_{\rho-3\delta} \\
& \leq \frac{C_{\mathfrak{L}B} C_{\tilde{\Omega}_L} C_{\Delta B} + C_{\mathfrak{L}B} C_{\Delta \tilde{\Omega}_L} \sigma_B + C_{\Delta \mathfrak{L}B} C_{\tilde{\Omega}_L} \sigma_B}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
& \quad + \frac{1}{2} \frac{\sigma_B C_{\mathfrak{L} \tilde{\Omega}_L} C_{\Delta B} + (\sigma_B)^2 C_{\Delta \mathfrak{L} \tilde{\Omega}_L} + C_{\Delta B} C_{\mathfrak{L} \tilde{\Omega}_L} \sigma_B}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
(5.60) \quad & =: \frac{C_{\Delta \mathfrak{L}A}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho,
\end{aligned}$$

$$\begin{aligned}
(5.61) \quad & \|\mathfrak{L}_\omega (J \circ \bar{K}) - \mathfrak{L}_\omega (J \circ K)\|_{\rho-2\delta} \\
& \leq \frac{c_{J,1} C_{\Delta \mathfrak{L}L} + c_{J,2} C_{\Delta K} C_{\mathfrak{L}K}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho =: \frac{C_{\Delta \mathfrak{L}J}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho,
\end{aligned}$$

$$\begin{aligned}
(5.62) \quad & \|\mathfrak{L}_\omega \bar{N}^0 - \mathfrak{L}_\omega N^0\|_{\rho-3\delta} \\
& \leq \frac{C_{\mathfrak{L}J} C_{\Delta L} + C_{\Delta \mathfrak{L}J} C_L \delta + c_{J,0} C_{\Delta \mathfrak{L}L} + C_{\Delta J} C_{\mathfrak{L}L} \delta}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
& =: \frac{C_{\Delta \mathfrak{L}N^0}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho,
\end{aligned}$$

$$\begin{aligned}
(5.63) \quad & \|\mathfrak{L}_\omega \bar{N} - \mathfrak{L}_\omega N\|_{\rho-3\delta} \\
& \leq \frac{C_{\mathfrak{L}L} C_{\Delta A} + C_{\Delta \mathfrak{L}L} C_A + C_L C_{\Delta \mathfrak{L}A} + C_{\Delta L} C_{\mathfrak{L}A}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
& \quad + \frac{C_{\mathfrak{L}N^0} C_{\Delta B} + C_{\Delta \mathfrak{L}N^0} \sigma_B + C_{N^0} C_{\Delta \mathfrak{L}B} + C_{\Delta N^0} C_{\mathfrak{L}B}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho \\
& =: \frac{C_{\Delta \mathfrak{L}N}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho,
\end{aligned}$$

where we used the constant (5.57).

Step 7: Control of the new torsion condition. Notice that this step could be replaced by the use of Cauchy estimates, thus obtaining a much pessimistic control. Now we control the non-degeneracy (twist) condition associated to the new torsion matrix

$$\bar{T}(\theta) = \bar{N}^\top \Omega(\bar{K}(\theta)) \mathcal{X}_{\bar{N}}(\theta),$$

where

$$\mathcal{X}_{\bar{N}}(\theta) = DX_h(\bar{K}(\theta)) \bar{N}(\theta) + \mathfrak{L}_\omega \bar{N}(\theta),$$

is the infinitesimal displacement of the normal subbundle for the linearized dynamics. At this point, intermediate computations will be skipped for convenience. Such details are left to the reader (they are analogous to previous computations).

We start by controlling the correction of the displacement. To this end, we observe that

$$\begin{aligned}
\mathcal{X}_{\bar{N}}(\theta) - \mathcal{X}_N(\theta) &= DX_h(\bar{K}(\theta)) \bar{N}(\theta) - DX_h(K(\theta)) N(\theta) \\
& \quad + \mathfrak{L}_\omega \bar{N}(\theta) - \mathfrak{L}_\omega N(\theta),
\end{aligned}$$

and we readily obtain

$$\|\mathcal{X}_{\bar{N}} - \mathcal{X}_N\|_{\rho-3\delta} \leq \frac{c_{X_h,1} C_{\Delta N} + c_{X_h,2} C_{\Delta K} C_N \delta + C_{\Delta \mathfrak{L}N}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho$$

$$(5.64) \quad =: \frac{C_{\Delta\mathcal{X}_N}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho.$$

Now, we can control the correction of the torsion matrix

$$(5.65) \quad \begin{aligned} \|\bar{T} - T\|_{\rho-3\delta} &\leq \|\bar{N}^\top\|_{\rho-3\delta} \|\Omega \circ \bar{K}\|_{\rho-2\delta} \|\mathcal{X}_{\bar{N}} - \mathcal{X}_N\|_{\rho-3\delta} \\ &\quad + \|\bar{N}^\top\|_{\rho-3\delta} \|\Omega \circ \bar{K} - \Omega \circ K\|_{\rho-2\delta} \|\mathcal{X}_N\|_{\rho-\delta} \\ &\quad + \|\bar{N}^\top - N^\top\|_{\rho-3\delta} \|\Omega \circ K\|_\rho \|\mathcal{X}_N\|_{\rho-\delta} \\ &\leq \frac{C_{N^\top} c_{\Omega,0} C_{\Delta\mathcal{X}_N} + C_{N^\top} c_{\Omega,1} C_{\Delta K} C_{\mathcal{X}_N} \delta + C_{\Delta N^\top} c_{\Omega,0} C_{\mathcal{X}_N}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho \\ &=: \frac{C_{\Delta T}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho. \end{aligned}$$

Then, we introduce the constant

$$C_{\Delta\langle T \rangle^{-1}} := 2(\sigma_T)^2 C_{\Delta T}$$

and check the condition (A.1) in Lemma A.1:

$$(5.66) \quad \begin{aligned} \frac{2(\sigma_T)^2 |\langle \bar{T} \rangle - \langle T \rangle|}{\sigma_T - |\langle T \rangle^{-1}|} &\leq \frac{2(\sigma_T)^2 \|\bar{T} - T\|_{\rho-3\delta}}{\sigma_T - |\langle T \rangle^{-1}|} \leq \frac{2(\sigma_T)^2 C_{\Delta T}}{\sigma_T - |\langle T \rangle^{-1}|} \frac{\|E\|_\rho}{\gamma^2\delta^{2\tau+1}} \\ &= \frac{C_{\Delta\langle T \rangle^{-1}}}{\sigma_B - \|B\|_\rho} \frac{\|E\|_\rho}{\gamma^2\delta^{2\tau+1}} < 1, \end{aligned}$$

where the last inequality follows from Hypothesis (5.8) (this corresponds to the sixth term in (5.9)). Hence, by invoking Lemma A.1, we conclude that

$$(5.67) \quad |\langle \bar{T} \rangle^{-1}| < \sigma_T, \quad |\langle \bar{T} \rangle^{-1} - \langle T \rangle^{-1}| \leq \frac{2(\sigma_T)^2 C_{\Delta T}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho = \frac{C_{\Delta\langle T \rangle^{-1}}}{\gamma^2\delta^{2\tau+1}} \|E\|_\rho,$$

and so, we obtain the estimates (5.12) and (5.16) on the new object. \square

5.3. Convergence of the iterative process. Now we are ready to proof our first KAM theorem with conserved quantities. Once the quadratic procedure has been established in Section 5.2, proving the convergence of the scheme follows standard arguments. Nevertheless, the required computations will be carefully detailed since we are interested in providing explicit conditions for the KAM theorem.

Proof of Theorem 3.1. Let us consider the approximate torus $K_0 := K$ with initial error $E_0 := E$. We also introduce $B_0 := B$ and $T_0 := T$ associated with the initial approximation. By applying Lemma 5.2 recursively, we obtain new objects $K_s = \bar{K}_{s-1}$, $E_s = \bar{E}_{s-1}$, $B_s = \bar{B}_{s-1}$ and $T_s = \bar{T}_{s-1}$. The domain of analyticity of these objects is reduced at every step. To characterize this fact, we introduce parameters $a_1 > 1$, $a_2 > 1$, $a_3 = 3\frac{a_1}{a_1-1}\frac{a_2}{a_2-1}$ and define

$$\rho_0 = \rho, \quad \delta_0 = \frac{\rho_0}{a_3}, \quad \rho_s = \rho_{s-1} - 3\delta_{s-1}, \quad \delta_s = \frac{\delta_0}{a_1^s}, \quad \rho_\infty = \lim_{s \rightarrow \infty} \rho_s = \frac{\rho_0}{a_2}.$$

We can select the above parameters (together with the parameter ν) to optimize the convergence of the KAM process for a particular problem (see [27]).

Let us assume that we have successfully applied s times Lemma 5.2 (the Iterative Lemma), and let K_s , E_s , B_s and T_s be the objects at the s -step of the quasi-Newton method. We observe that condition (5.8) is required at every step, but the construction has been performed in such a way that we control $\|DK_s\|_{\rho_s}$, $\|(DK_s)^\top\|_{\rho_s}$, $\|B_s\|_{\rho_s}$, $\text{dist}(K_s(\mathbb{T}_{\rho_s}^d), \partial\mathcal{B})$, and $|\langle T_s \rangle^{-1}|$ uniformly with respect to s ,

so the constants that appear in Lemma 5.2 (which are obtained in Table 1 and Table 2) are taken to be the same for all steps by considering the worst value of δ_s , that is, $\delta_0 = \rho_0/a_3$.

The first computation is tracking the sequence E_s of errors:

$$(5.68) \quad \begin{aligned} \|E_s\|_{\rho_s} &< \frac{C_E}{\gamma^4 \delta_{s-1}^{4\tau}} \|E_{s-1}\|_{\rho_{s-1}}^2 = \frac{C_E a_1^{4\tau(s-1)}}{\gamma^4 \delta_0^{4\tau}} \|E_{s-1}\|_{\rho_{s-1}}^2 \\ &< \left(\frac{a_1^{4\tau} C_E \|E_0\|_{\rho_0}}{\gamma^4 \delta_0^{4\tau}} \right)^{2^s - 1} a_1^{-4\tau s} \|E_0\|_{\rho_0} < a_1^{-4\tau s} \|E_0\|_{\rho_0}, \end{aligned}$$

where we used the sums $1 + 2 + \dots + 2^{s-1} = 2^s - 1$, and $1(s-1) + 2(s-2) + 2^2(s-3) + \dots + 2^{s-2} = 2^s - s - 1$. Notice that we also used the inequality

$$(5.69) \quad \frac{a_1^{4\tau} C_E \|E_0\|_{\rho_0}}{\gamma^4 \delta_0^{4\tau}} < 1,$$

which is included in (3.10). Now, using expression (5.68), we check the Hypothesis (5.8) of the iterative Lemma, so that we can perform the step $s + 1$. The required sufficient condition will be included in the hypothesis (3.10) of the KAM theorem.

The right inequality in (5.8) reads:

$$\frac{C_E \|E_s\|_{\rho_s}}{\gamma^4 \delta_s^{4\tau}} \leq \frac{C_E a_1^{-4\tau s} \|E_0\|_{\rho_0}}{\gamma^4 \delta_s^{4\tau}} = \frac{C_E \|E_0\|_{\rho_0}}{\gamma^4 \delta_0^{4\tau}} \leq a_1^{-4\tau} < 1,$$

where we used (5.68) and (5.69).

The left inequality in (5.8) has several terms (which correspond to the different components in (5.9)). The first of them, using again (5.68), is given by

$$(5.70) \quad \frac{\|E_s\|_{\rho_s}}{\delta_s} \leq \frac{a_1^s a_1^{-4\tau s} \|E_0\|_{\rho_0}}{\delta_0} \leq \frac{\|E_0\|_{\rho_0}}{\delta_0} < \nu.$$

We used that $\tau \geq d - 1 \geq 1$, so that $1 - 4\tau < 0$. The last inequality in (5.70) is included in (3.10). The second term is guaranteed by performing the following computation

$$(5.71) \quad \frac{2C_{\text{sym}} \|E_s\|_{\rho_s}}{\gamma \delta_s^{\tau+1}} \leq \frac{2C_{\text{sym}} a_1^{s(\tau+1)} a_1^{-4\tau s} \|E_0\|_{\rho_0}}{\gamma \delta_0^{\tau+1}} \leq \frac{2C_{\text{sym}} \|E_0\|_{\rho_0}}{\gamma \delta_0^{\tau+1}} < 1,$$

where we used (5.68), the fact that $1 - 3\tau < 0$, and we have included the last inequality in (3.10). The remaining conditions are similar. We only need to pay attention to the fact that they involve the objects DK_s , $(DK_s)^\top$, B_s and $\langle T_s \rangle^{-1}$, at the s^{th} step. Hence, we have to relate these conditions to the corresponding initial objects DK_0 , $(DK_0)^\top$, B_0 and $\langle T_0 \rangle^{-1}$. For example, the third term in (5.8) reads

$$\left(\frac{dC_{\Delta K}}{\sigma_K - \|DK_s\|_{\rho_s}} \right) \frac{\|E_s\|_{\rho_s}}{\gamma^2 \delta_s^{2\tau+1}} < 1,$$

and it is checked by performing the following computations

$$\begin{aligned} \|DK_s\|_{\rho_s} + \frac{dC_{\Delta K} \|E_s\|_{\rho_s}}{\gamma^2 \delta_s^{2\tau+1}} &< \|DK_0\|_{\rho_0} + \sum_{j=0}^s \frac{dC_{\Delta K} \|E_j\|_{\rho_j}}{\gamma^2 \delta_j^{2\tau+1}} \\ &< \|DK_0\|_{\rho_0} + \sum_{j=0}^{\infty} \frac{dC_{\Delta K} a_1^{(1-2\tau)j}}{\gamma^2 \delta_0^{2\tau+1}} \|E_0\|_{\rho_0} \end{aligned}$$

$$(5.72) \quad = \|DK_0\|_{\rho_0} + \frac{dC_{\Delta K}\|E_0\|_{\rho_0}}{\gamma^2\delta_0^{2\tau+1}} \left(\frac{1}{1-a_1^{1-2\tau}} \right) < \sigma_K.$$

Again, the last inequality is included into (3.10). The fourth, fifth and sixth terms in (5.8) (associated to $(DK_s)^\top$, B_s and $\langle T_s \rangle^{-1}$, respectively) follow by reproducing the same computations. Finally, the seventh term in (5.8) is checked as follows

$$(5.73) \quad \begin{aligned} \text{dist}(K_s(\mathbb{T}_{\rho_s}^d), \partial\mathcal{B}) - \frac{C_{\Delta K}}{\gamma^2\delta_s^{2\tau}} \|E_s\|_{\rho_s} &> \text{dist}(K_0(\mathbb{T}_{\rho_0}^d), \partial\mathcal{B}) - \sum_{j=0}^{\infty} \frac{C_{\Delta K} a_1^{-2\tau j} \|E_0\|_{\rho_0}}{\gamma^2\delta_0^{2\tau}} \\ &= \text{dist}(K_0(\mathbb{T}_{\rho_0}^d), \partial\mathcal{B}) - \frac{C_{\Delta K}\|E_0\|_{\rho_0}}{\gamma^2\delta_0^{2\tau}} \frac{1}{1-a_1^{-2\tau}} > 0, \end{aligned}$$

where the last inequality is included into (3.10).

Having guaranteed all hypothesis of Lemma 5.2, we collect the inequalities (5.69), (5.70), (5.71), (5.72) and (5.73) that are included into hypothesis (3.10). This follows by introducing the constant \mathfrak{C}_1 as

$$(5.74) \quad \mathfrak{C}_1 := \max \{ (a_1 a_3)^{4\tau} C_E, (a_3)^{2\tau+1} \gamma^2 \rho^{2\tau-1} C_\Delta \}$$

where

$$(5.75) \quad C_{\Delta,1} := \max \left\{ \frac{dC_{\Delta K}}{\sigma_K - \|DK\|_\rho}, \frac{2nC_{\Delta K}}{\sigma_{K^\top} - \|(DK)^\top\|_\rho}, \frac{C_{\Delta B}}{\sigma_B - \|B\|_\rho}, \frac{C_{\Delta\langle T \rangle^{-1}}}{\sigma_T - |\langle T \rangle^{-1}|} \right\},$$

$$(5.76) \quad C_{\Delta,2} := \frac{C_{\Delta K} \delta}{\text{dist}(K(\mathbb{T}_\rho^d), \partial\mathcal{B})},$$

$$(5.77) \quad C_\Delta := \max \left\{ \frac{\gamma^2 \delta^{2\tau}}{\nu}, 2C_{\text{sym}} \gamma \delta^\tau, \frac{C_{\Delta,1}}{1-a_1^{1-2\tau}}, \frac{C_{\Delta,2}}{1-a_1^{-2\tau}} \right\}.$$

Note that we recovered the original notation $K = K_0$, $B = B_0$, $T = T_0$, $\rho = \rho_0$ and $\delta = \delta_0$ for the original objects.

Therefore, by induction, we can apply the iterative process infinitely many times. Indeed, we have

$$\|E_s\|_{\rho_s} < a_1^{-4\tau s} \|E\|_\rho \longrightarrow 0 \quad \text{when } s \rightarrow 0$$

so the iterative scheme converges to a true quasi-periodic torus K_∞ . As a result of the output of Lemma 5.2, this object satisfies $K_\infty \in \mathcal{A}(\mathbb{T}_{\rho_\infty}^d)$ and

$$\|DK_\infty\|_{\rho_\infty} < \sigma_K, \quad \|(DK_\infty)^\top\|_{\rho_\infty} < \sigma_{K^\top}, \quad \text{dist}(K_\infty(\mathbb{T}^d), \partial\mathcal{B}) > 0.$$

Furthermore, we control the limit objects by repeating the computations in (5.73) as follows

$$(5.78) \quad \|K_\infty - K\|_{\rho_\infty} \leq \sum_{j=0}^{\infty} \|K_{j+1} - K_j\|_{\rho_{j+1}} < \frac{C_{\Delta K}\|E\|_\rho}{\gamma^2\delta^{2\tau}} \frac{1}{1-a_1^{-2\tau}} =: \frac{\mathfrak{C}_2\|E\|_\rho}{\gamma^2\rho^{2\tau}},$$

$$(5.79) \quad |\langle c \circ K_\infty \rangle - \langle c \circ K \rangle| \leq \|Dc\|_{\mathcal{B}} \|K_\infty - K\|_{\rho_\infty} < \frac{c_{c,1}\mathfrak{C}_2\|E\|_\rho}{\gamma^2\rho^{2\tau}} =: \frac{\mathfrak{C}_3\|E\|_\rho}{\gamma^2\rho^{2\tau}},$$

thus obtaining the estimates in (3.11). This completes the proof of the ordinary KAM theorem. \square

6. PROOF OF THE GENERALIZED ISO-ENERGETIC KAM THEOREM

In this section we present a proof of Theorem 3.6 following the same structure of the proof of Theorem 3.1 in Section 5. We will emphasize the differences between both proofs, and the computations that are analogous will be conveniently omitted for the sake of brevity. In Section 6.1 we discuss the approximate solution of linearized equations in the symplectic frame constructed in Section 4. In Section 6.2 we produce quantitative estimates for the objects obtained when performing one iteration of the previous procedure. Finally, in Section 6.3 we discuss the convergence of the quasi-Newton method.

6.1. The quasi-Newton method. The proof of Theorem 3.6 consists again on refining $K(\theta)$ and ω by means of a quasi-Newton method. In this case, the total error is associated with the error of invariance and the target energy level:

$$E_c(\theta) = \begin{pmatrix} E(\theta) \\ E^\omega \end{pmatrix} = \begin{pmatrix} X_h(K(\theta)) + \mathfrak{L}_\omega K(\theta) \\ \langle c \circ K \rangle - c_0 \end{pmatrix}.$$

Then, we look for a corrected parameterization $\bar{K}(\theta) = K(\theta) + \Delta K(\theta)$ and a corrected frequency $\bar{\omega} = \omega + \Delta\omega$ by considering the linearized system

$$(6.1) \quad \begin{aligned} DX_h(K(\theta))\Delta K(\theta) + \mathfrak{L}_\omega \Delta K(\theta) + \mathfrak{L}_{\Delta\omega} K(\theta) &= -E(\theta), \\ \langle (Dc \circ K)\Delta K \rangle &= -E^\omega. \end{aligned}$$

If we obtain a good enough approximation of the solution $(\Delta K(\theta), \Delta\omega)$ of (6.1), then $\bar{K}(\theta)$ provides a parameterization of an approximately invariant torus of frequency $\bar{\omega}$, with a quadratic error in terms of $E_c(\theta) = (E(\theta), E^\omega)$.

To face the linearized equations (6.1), we introduce again the linear change

$$(6.2) \quad \Delta K(\theta) = P(\theta)\xi(\theta),$$

where $P(\theta)$ is the approximately symplectic frame characterized in Lemma 4.6. In addition, to ensure Diophantine properties for $\bar{\omega}$, we select a parallel correction of the frequency:

$$(6.3) \quad \Delta\omega = -\omega \xi^\omega,$$

where ξ^ω is a real number. This guarantees the solvability of system (6.1) along the iterative procedure. The following notation for the new unknowns will be useful

$$\xi_c(\theta) = \begin{pmatrix} \xi(\theta) \\ \xi^\omega \end{pmatrix}, \quad \xi(\theta) = \begin{pmatrix} \xi^L(\theta) \\ \xi^N(\theta) \end{pmatrix}.$$

Then, taking into account the expressions (6.2) and (6.3), the system of equations in (6.1) becomes

$$(6.4) \quad \begin{aligned} (DX_h(K(\theta))P(\theta) + \mathfrak{L}_\omega P(\theta))\xi(\theta) + P(\theta)\mathfrak{L}_\omega \xi(\theta) - \mathfrak{L}_\omega K(\theta)\xi^\omega &= -E(\theta), \\ \langle (Dc \circ K)P\xi \rangle &= -E^\omega. \end{aligned}$$

We now multiply both sides of (6.4) by $-\Omega_0 P(\theta)^\top \Omega(K(\theta))$, and we use the geometric properties in Lemma 4.6 and Lemma 4.10, thus obtaining the equivalent equations:

$$(6.5) \quad \begin{aligned} (\Lambda(\theta) + E_{\text{red}}(\theta))\xi(\theta) + (I_{2n} - \Omega_0 E_{\text{sym}}(\theta))\mathfrak{L}_\omega \xi(\theta) \\ + \Omega_0 P(\theta)^\top \Omega(K(\theta))\mathfrak{L}_\omega K(\theta)\xi^\omega = \Omega_0 P(\theta)^\top \Omega(K(\theta))E(\theta), \end{aligned}$$

$$(6.6) \quad \langle (Dc \circ K)N\xi^N \rangle + \langle (Dc \circ K)L\xi^L \rangle = -E^\omega.$$

We observe that

$$\begin{aligned}\Omega_0 P(\theta)^\top \Omega(K(\theta)) \mathfrak{L}_\omega K(\theta) &= -\Omega_0 P(\theta)^\top \Omega(K(\theta)) L(\theta) \hat{\omega} \\ &= \begin{pmatrix} \hat{\omega} \\ 0_n \end{pmatrix} + \begin{pmatrix} A(\theta)^\top \Omega_L(\theta) \hat{\omega} \\ -\Omega_L(\theta) \hat{\omega} \end{pmatrix},\end{aligned}$$

where we used the notation for $\hat{\omega}$ introduced in the statement of the theorem (e.g. equation (3.13)). Moreover, recalling the computations in Lemma 4.2, we obtain

$$\begin{aligned}\text{Dc}(K(\theta))L(\theta) &= (\text{Dc}(K(\theta))\text{D}K(\theta) \quad \text{Dc}(K(\theta))X_p(K(\theta))) \\ &= (\text{D}(c(K(\theta))) \quad 0_{n-d}^\top) \\ &= (\text{D}(\mathfrak{R}_\omega(\text{Dc}(K(\theta))E(\theta))) \quad 0_{n-d}^\top).\end{aligned}$$

From the above expressions, we conclude that equations (6.5) and (6.6), are approximated by a triangular system that requires to solve two cohomological equations of the form (4.1) consecutively. Quantitative estimates for the solutions of such equations are provided in the following statement.

Lemma 6.1 (Upper triangular equations for the iso-energetic case). *Let $\omega \in \mathcal{D}_{\gamma,\tau}$, $\eta^\omega \in \mathbb{R}$, and let us consider the map $\eta = (\eta^L, \eta^N) : \mathbb{T}^d \rightarrow \mathbb{R}^{2n} \simeq \mathbb{R}^n \times \mathbb{R}^n$, with components in $\mathcal{A}(\mathbb{T}_\rho^d)$, and the maps $T : \mathbb{T}^d \rightarrow \mathbb{R}^{n \times n}$ and $H : \mathbb{T}^d \rightarrow \mathbb{R}^{1 \times n}$, with components in $\mathcal{A}(\mathbb{T}_{\rho-\delta}^d)$. Let us introduce the notation*

$$T_c(\theta) = \begin{pmatrix} T(\theta) & \hat{\omega} \\ H(\theta) & 0 \end{pmatrix}, \quad \hat{\omega} = \begin{pmatrix} \omega \\ 0_{n-d} \end{pmatrix}.$$

Let us assume that T_c satisfies the non-degeneracy condition $\det \langle T_c \rangle \neq 0$, and η satisfies the compatibility condition $\langle \eta^N \rangle = 0_n$. Then, for any $\xi_0^L \in \mathbb{R}^n$, the system of equations

$$(6.7) \quad \begin{pmatrix} O_n & T(\theta) \\ O_n & O_n \end{pmatrix} \begin{pmatrix} \xi^L(\theta) \\ \xi^N(\theta) \end{pmatrix} + \begin{pmatrix} \mathfrak{L}_\omega \xi^L(\theta) \\ \mathfrak{L}_\omega \xi^N(\theta) \end{pmatrix} + \begin{pmatrix} \hat{\omega} \xi^\omega \\ 0_n \end{pmatrix} = \begin{pmatrix} \eta^L(\theta) \\ \eta^N(\theta) \end{pmatrix}$$

$$\langle H \xi^N \rangle = \eta^\omega,$$

has a solution given by

$$(6.8) \quad \xi^N(\theta) = \xi_0^N + \mathfrak{R}_\omega(\eta^N(\theta)),$$

$$(6.9) \quad \xi^L(\theta) = \xi_0^L + \mathfrak{R}_\omega(\eta^L(\theta) - T(\theta)\xi^N(\theta)),$$

where

$$(6.10) \quad \begin{pmatrix} \xi_0^N \\ \xi^\omega \end{pmatrix} = \langle T_c \rangle^{-1} \begin{pmatrix} \langle \eta^L - T \mathfrak{R}_\omega(\eta^N) \rangle \\ \langle \eta^\omega - H \mathfrak{R}_\omega(\eta^N) \rangle \end{pmatrix}$$

and \mathfrak{R}_ω is given by (4.2).

Moreover, we have the estimates

$$\begin{aligned}|\xi_0^N|, |\xi^\omega| &\leq |\langle T_c \rangle^{-1}| \max \left\{ \|\eta^L\|_\rho + \frac{c_R}{\gamma \delta^\tau} \|T\|_{\rho-\delta} \|\eta^N\|_\rho, \right. \\ &\quad \left. \|\eta^\omega\| + \frac{c_R}{\gamma \delta^\tau} \|H\|_{\rho-\delta} \|\eta^N\|_\rho \right\}, \\ \|\xi^N\|_{\rho-\delta} &\leq |\xi_0^N| + \frac{c_R}{\gamma \delta^\tau} \|\eta^N\|_\rho, \\ \|\xi^L\|_{\rho-2\delta} &\leq |\xi_0^L| + \frac{c_R}{\gamma \delta^\tau} (\|\eta^L\|_{\rho-\delta} + \|T\|_{\rho-\delta} \|\xi^N\|_{\rho-\delta}).\end{aligned}$$

Proof. It is analogous to Lemma 5.1. After solving $\xi^N = \xi_0^N + \mathfrak{R}_\omega(\eta^N)$, $\xi_0^N \in \mathbb{R}^n$ from the triangular system, we observe that

$$\langle H\xi^N \rangle = \langle H \rangle \xi_0^N + \langle H\mathfrak{R}_\omega(\eta^N) \rangle,$$

so that the last equation in (6.7) becomes

$$\langle H \rangle \xi_0^N = \eta^\omega - \langle H\mathfrak{R}_\omega(\eta^N) \rangle = \langle \eta^\omega - H\mathfrak{R}_\omega(\eta^N) \rangle.$$

This equation, together with the compatibility condition required to solve the equation for ξ^L , yields to a linear system which can be solved using that $\det \langle T_c \rangle \neq 0$, thus obtaining (6.10). The estimates are obtained using Lemma 4.1. \square

To approximate the solutions of (6.5)–(6.6), we will invoke Lemma 6.1 taking

$$(6.11) \quad \eta^L(\theta) = -N(\theta)^\top \Omega(K(\theta))E(\theta),$$

$$(6.12) \quad \eta^N(\theta) = L(\theta)^\top \Omega(K(\theta))E(\theta),$$

$$(6.13) \quad \eta^\omega = -E^\omega,$$

$$(6.14) \quad H(\theta) = \text{Dc}(K(\theta))N(\theta),$$

and $T(\theta)$ given by (3.6). We recall from Lemma 4.6 that the compatibility condition $\langle \eta^N \rangle = 0_n$ is satisfied. We will select the solution that satisfies $\xi_0^L = \langle \xi^L \rangle = 0_n$, even though other choices can be selected according to the context (see Remark 3.5).

Then, from Lemma 6.1 we have $\|\xi\|_{\rho-2\delta} = \mathcal{O}(\|E_c\|_\rho)$ and $|\xi^\omega| = \mathcal{O}(\|E_c\|_\rho)$, and using the geometric properties characterized in Section 4, we have $\|E_{\text{red}}\|_{\rho-2\delta} = \mathcal{O}(\|E\|_\rho)$ and $\|E_{\text{sym}}\|_{\rho-2\delta} = \mathcal{O}(\|E\|_\rho)$. Hence, the error produced when approximating the solutions of (6.5)–(6.6) using the solutions of (6.7) will be controlled by $\mathcal{O}(\|E_c\|_\rho^2)$. This, together with other estimates, will be suitably quantified in the next section.

6.2. One step of the iterative procedure. In this section we apply one correction of the quasi-Newton method described in Section 6.1 and we obtain sharp quantitative estimates for the new approximately invariant torus and related objects. We set sufficient conditions to preserve the control of the previous estimates.

Lemma 6.2 (The Iterative Lemma in the iso-energetic case). *Let us consider the same setting and hypotheses of Theorem 3.6, and a constant $\nu > 0$. Then, there exist constants $C_{\Delta K}$, $C_{\Delta\omega}$, $C_{\Delta B}$, $C_{\Delta\langle T_c \rangle^{-1}}$ and C_{E_c} such that if the inequalities*

$$(6.15) \quad \frac{\hat{C}_\Delta \|E_c\|_\rho}{\gamma^2 \delta^{2\tau+1}} < 1, \quad \frac{C_{E_c} \|E_c\|_\rho}{\gamma^4 \delta^{4\tau}} < 1$$

hold for some $0 < \delta < \rho$, where

$$(6.16) \quad \hat{C}_\Delta := \max \left\{ \frac{\gamma^2 \delta^{2\tau}}{\nu}, 2C_{\text{sym}} \gamma \delta^\tau, \frac{dC_{\Delta K}}{\sigma_K - \|DK\|_\rho}, \frac{2nC_{\Delta K}}{\sigma_{K^\top} - \|(DK)^\top\|_\rho}, \right. \\ \left. \frac{C_{\Delta B}}{\sigma_B - \|B\|_\rho}, \frac{C_{\Delta\langle T_c \rangle^{-1}}}{\sigma_{T_c} - |\langle T_c \rangle^{-1}|}, \frac{C_{\Delta K} \delta}{\text{dist}(K(\mathbb{T}_\rho^d), \partial B)}, \frac{C_{\Delta\omega} \gamma \delta^{\tau+1}}{\text{dist}(\omega, \partial\Theta)} \right\},$$

then we have an approximate torus of frequency $\bar{\omega} = \omega + \Delta\omega$ given by $\bar{K} = K + \Delta K$, with components in $\mathcal{A}(\mathbb{T}_{\rho-2\delta}^d)$, that defines new objects \bar{B} and \bar{T}_c (obtained replacing K by \bar{K}) satisfying

$$(6.17) \quad \|D\bar{K}\|_{\rho-3\delta} < \sigma_K,$$

$$(6.18) \quad \|(D\bar{K})^\top\|_{\rho-3\delta} < \sigma_{K^\top},$$

$$(6.19) \quad \|\bar{B}\|_{\rho-3\delta} < \sigma_B,$$

$$(6.20) \quad |\langle \bar{T}_c \rangle^{-1}| < \sigma_{T_c},$$

$$(6.21) \quad \text{dist}(\bar{K}(\mathbb{T}_{\rho-2\delta}^d), \partial\mathcal{B}) > 0,$$

$$(6.22) \quad \text{dist}(\bar{\omega}, \partial\Theta) > 0,$$

and

$$(6.23) \quad \|\bar{K} - K\|_{\rho-2\delta} < \frac{C_{\Delta K}}{\gamma^2 \delta^{2\tau}} \|E_c\|_{\rho},$$

$$(6.24) \quad |\bar{\omega} - \omega| < \frac{C_{\Delta\omega}}{\gamma \delta^{\tau}} \|E_c\|_{\rho},$$

$$(6.25) \quad \|\bar{B} - B\|_{\rho-3\delta} < \frac{C_{\Delta B}}{\gamma^2 \delta^{2\tau+1}} \|E_c\|_{\rho},$$

$$(6.26) \quad |\langle \bar{T}_c \rangle^{-1} - \langle T_c \rangle^{-1}| < \frac{C_{\Delta(T_c)^{-1}}}{\gamma^2 \delta^{2\tau+1}} \|E_c\|_{\rho}.$$

The new total error is given by

$$\bar{E}_c(\theta) = \begin{pmatrix} \bar{E}(\theta) \\ \bar{E}^{\omega} \end{pmatrix} = \begin{pmatrix} X_h(\bar{K}(\theta)) + \mathfrak{L}_{\bar{\omega}} \bar{K}(\theta) \\ \langle c \circ \bar{K} \rangle - c_0 \end{pmatrix},$$

and satisfies

$$(6.27) \quad \|\bar{E}_c\|_{\rho-2\delta} < \frac{C_{E_c}}{\gamma^4 \delta^{4\tau}} \|E_c\|_{\rho}^2.$$

The above constants are collected in Table 4.

Proof. The proof of this result is parallel to the ordinary situation. On the one hand, those constants that must be changed in this result (e.g. $C_{\xi_0^N}$) will be redefined using the symbol “:=” and will be included in Table 4. On the other hand, those constants that are not redefined (e.g. $C_{\Delta K}$) will have the same expression that in the proof of Lemma 5.2 and the reader is referred to Table 2 for the corresponding label in the text.

Step 1: Control of the new parameterization. We start by considering the new objects $\bar{K}(\theta) = K(\theta) + \Delta K(\theta)$ and $\bar{\omega} = \omega + \Delta\omega$, obtained from the expressions (6.2) and (6.3), using the solutions of the system (6.7) taking the objects (6.11) (6.12) (6.13) and (6.14). We have

$$\|\eta^L\|_{\rho} \leq C_{N^{\top} c_{\Omega,0}} \|E\|_{\rho}, \quad \|\eta^N\|_{\rho} \leq C_{L^{\top} c_{\Omega,0}} \|E\|_{\rho}, \quad |\eta^{\omega}| \leq |E^{\omega}|,$$

and

$$\|H\|_{\rho-\delta} = \|(\text{Dc} \circ K)N\|_{\rho-\delta} \leq \|\text{Dc}\|_{\mathcal{B}} \|N\|_{\rho} \leq c_{c,1} C_N.$$

Notice that

$$\|E_c\|_{\rho} = \max\{\|E\|_{\rho}, |E^{\omega}|\},$$

so, from now on, we will use that $\|E\|_{\rho} \leq \|E_c\|_{\rho}$ and $|E^{\omega}| \leq \|E_c\|_{\rho}$.

In order to invoke Lemma 4.8 (we must fulfill condition (4.40)) we include the inequality

$$(6.28) \quad \frac{\|E_c\|_{\rho}}{\delta} < \nu$$

into Hypothesis (6.15) (this corresponds to the first term in (6.16)). Hence, combining Lemma 4.8 and Lemma 6.1, we obtain estimates for the solution of the cohomological equations (we recall that $\xi_0^L = 0_n$)

$$(6.29) \quad \begin{aligned} |\xi_0^N|, |\xi^\omega| &\leq \sigma_{T_c} \max \left\{ C_{N^\top} c_{\Omega,0} + \frac{c_R}{\gamma \delta^\tau} C_T C_{L^\top} c_{\Omega,0}, 1 + \frac{c_R}{\gamma \delta^\tau} c_{c,1} C_N C_{L^\top} c_{\Omega,0} \right\} \|E_c\|_\rho \\ &=: \frac{C_{\xi_0^N}}{\gamma \delta^\tau} \|E_c\|_\rho =: \frac{C_{\xi^\omega}}{\gamma \delta^\tau} \|E_c\|_\rho, \end{aligned}$$

and we observe that $\xi^L(\theta)$ and $\xi^N(\theta)$ are controlled as in Lemma 5.2, and so are the objects $\xi(\theta)$, $\bar{K}(\theta)$ and $\bar{K}(\theta) - K(\theta)$, thus obtaining the estimate in (6.23). Observing that

$$(6.30) \quad |\omega_*| < |\omega| < \sigma_\omega |\omega_*| =: C_\omega,$$

we get the estimate (6.24) as follows:

$$(6.31) \quad |\bar{\omega} - \omega| = |\omega \xi^\omega| < \frac{C_\omega C_{\xi^\omega}}{\gamma \delta^\tau} \|E_c\|_\rho =: \frac{C_{\Delta\omega}}{\gamma \delta^\tau} \|E_c\|_\rho.$$

To obtain (6.21), we repeat the computations in (5.25) using Hypothesis (6.15) (this corresponds to the seventh term in (6.16)). Similarly, we obtain (6.22):

$$\text{dist}(\bar{\omega}, \partial\Theta) \geq \text{dist}(\omega, \partial\Theta) - |\Delta\omega| \geq \text{dist}(\omega, \partial\Theta) - \frac{C_{\Delta\omega}}{\gamma \delta^\tau} \|E_c\|_\rho > 0,$$

where we used Hypothesis (6.15) (this corresponds to the eight term in (6.16)). A direct consequence of the fact that the new frequency $\bar{\omega}$ is strictly contained in Θ , and so $|\omega_*| < |(1 - \xi^\omega)\omega| < \sigma_\omega |\omega_*|$, is that we have

$$(6.32) \quad \frac{1}{\sigma_\omega} < 1 - \xi^\omega < \sigma_\omega, \quad \frac{1}{\sigma_\omega} < \frac{1}{1 - \xi^\omega} < \sigma_\omega.$$

Step 2: Control of the new error of invariance. To control the error of invariance of the corrected parameterization \bar{K} , we first consider the error in the solution of the linearized equation (6.1), that is

$$\begin{aligned} E_{\text{lin}}(\theta) &:= E_{\text{red}}(\theta)\xi(\theta) - \Omega_0 E_{\text{sym}}(\theta)\mathfrak{L}_\omega \xi(\theta) + \begin{pmatrix} A(\theta)^\top \Omega_L(\theta) \\ -\Omega_L(\theta) \end{pmatrix} \hat{\omega} \xi^\omega, \\ E_{\text{lin}}^\omega &:= \langle (D(\mathfrak{R}_\omega((Dc \circ K)E)) \quad 0_{n-d}^\top) \xi^L \rangle. \end{aligned}$$

First, we control $\mathfrak{L}_\omega \xi(\theta)$ in a similar fashion as in Step 2 of the ordinary case, but using now that $\xi(\theta)$ is the solution of (6.7). The only difference is that

$$(6.33) \quad \begin{aligned} \|\mathfrak{L}_\omega \xi^L\|_{\rho-\delta} &= \|\eta^L - T\xi^N - \hat{\omega} \xi^\omega\|_{\rho-\delta} \\ &\leq \left(C_{N^\top} c_{\Omega,0} + C_T \frac{C_{\xi^N}}{\gamma \delta^\tau} + C_\omega \frac{C_{\xi^\omega}}{\gamma \delta^\tau} \right) \|E\|_\rho =: \frac{C_{\mathfrak{L}\xi^L}}{\gamma \delta^\tau} \|E\|_\rho. \end{aligned}$$

Then, we control $E_{\text{lin}}(\theta)$ and E_{lin}^ω by

$$(6.34) \quad \begin{aligned} \|E_{\text{lin}}\|_{\rho-2\delta} &\leq \|E_{\text{red}}\|_{\rho-2\delta} \|\xi\|_{\rho-2\delta} + c_{\Omega,0} \|E_{\text{sym}}\|_{\rho-2\delta} \|\mathfrak{L}_\omega \xi\|_{\rho-2\delta} \\ &\quad + \max\{\|A\|_{\rho-2\delta}, 1\} \|\Omega_L\|_{\rho-2\delta} |\hat{\omega}| |\xi^\omega| \\ &\leq \frac{C_{\text{red}}}{\gamma \delta^{\tau+1}} \frac{C_\xi}{\gamma^2 \delta^{2\tau}} \|E_c\|_\rho^2 + c_{\Omega,0} \frac{C_{\text{sym}}}{\gamma \delta^{\tau+1}} \frac{C_{\mathfrak{L}\xi}}{\gamma \delta^\tau} \|E_c\|_\rho^2 \\ &\quad + \frac{\max\{C_A, 1\} C_{\Omega_L} C_\omega C_{\xi_0^N}}{\gamma^2 \delta^{2\tau+1}} \|E_c\|_\rho^2 =: \frac{C_{\text{lin}}}{\gamma^3 \delta^{3\tau+1}} \|E_c\|_\rho^2, \end{aligned}$$

$$(6.35) \quad |E_{\text{lin}}^\omega| \leq \frac{dc_R c_{c,1}}{\gamma \delta^{\tau+1}} \|E_c\|_\rho \frac{C_{\xi^L}}{\gamma^2 \delta^{2\tau}} \|E_c\|_\rho =: \frac{C_{\text{lin}}^\omega}{\gamma^3 \delta^{3\tau+1}} \|E_c\|_\rho^2.$$

After performing the correction, the total error associated with the new parameterization is given by (the computation is analogous to (5.30))

$$\bar{E}_c(\theta) = \begin{pmatrix} \bar{E}(\theta) \\ \bar{E}^\omega \end{pmatrix} = \begin{pmatrix} P(\theta)(I - \Omega_0 E_{\text{sym}}(\theta))^{-1} E_{\text{lin}}(\theta) + \Delta^2 X(\theta) + \mathfrak{L}_{\Delta\omega} \Delta K(\theta) \\ E_{\text{lin}}^\omega + \langle \Delta^2 c(\theta) \rangle \end{pmatrix},$$

where $\Delta^2 X(\theta)$ is given by (5.31) and

$$\Delta^2 c(\theta) = \int_0^1 (1-t) D^2 c(K(\theta) + t \Delta K(\theta)) [\Delta K(\theta), \Delta K(\theta)] dt.$$

Now we observe that

$$\mathfrak{L}_{\Delta\omega} \Delta K(\theta) = -\xi^\omega \mathfrak{L}_\omega \Delta K(\theta),$$

where $\mathfrak{L}_\omega \Delta K(\theta)$ is controlled using the expression

$$\mathfrak{L}_\omega \Delta K(\theta) = \mathfrak{L}_\omega L(\theta) \xi^L(\theta) + L(\theta) \mathfrak{L}_\omega \xi^L(\theta) + \mathfrak{L}_\omega N(\theta) \xi^N(\theta) + N(\theta) \mathfrak{L}_\omega \xi^N(\theta),$$

thus obtaining

$$(6.36) \quad \begin{aligned} \|\mathfrak{L}_\omega \Delta K\|_{\rho-2\delta} &\leq \left(\frac{C_{\mathfrak{L}L} C_{\xi^L}}{\gamma^2 \delta^{2\tau}} + \frac{C_L C_{\mathfrak{L}\xi^L}}{\gamma \delta^\tau} + \frac{C_{\mathfrak{L}N} C_{\xi^N}}{\gamma \delta^\tau} + C_N C_{\mathfrak{L}\xi^N} \right) \|E\|_\rho \\ &=: \frac{C_{\mathfrak{L}\Delta K}}{\gamma^2 \delta^{2\tau}} \|E\|_\rho. \end{aligned}$$

Hence, we have:

$$(6.37) \quad \begin{aligned} \|\bar{E}\|_{\rho-2\delta} &\leq \left(\frac{2(C_L + C_N) C_{\text{lin}}}{\gamma^3 \delta^{3\tau+1}} + \frac{1}{2} c_{X,h,2} \frac{(C_{\Delta K})^2}{\gamma^4 \delta^{4\tau}} + \frac{C_{\xi^\omega} C_{\mathfrak{L}\Delta K}}{\gamma^3 \delta^{3\tau}} \right) \|E_c\|_\rho^2 \\ &=: \frac{C_E}{\gamma^4 \delta^{4\tau}} \|E_c\|_\rho^2, \end{aligned}$$

where we used the second term in (6.16) (Hypothesis (6.15)). Moreover, we get

$$(6.38) \quad |\bar{E}^\omega| \leq \left(\frac{C_{\text{lin}}^\omega}{\gamma^3 \delta^{3\tau+1}} + \frac{1}{2} c_{c,2} \frac{(C_{\Delta K})^2}{\gamma^4 \delta^{4\tau}} \right) \|E_c\|_\rho^2 =: \frac{C_{E^\omega}}{\gamma^4 \delta^{4\tau}} \|E_c\|_\rho^2$$

and, finally,

$$(6.39) \quad \|\bar{E}_c\|_{\rho-2\delta} \leq \frac{\max\{C_E, C_{E^\omega}\}}{\gamma^4 \delta^{4\tau}} \|E_c\|_\rho^2 =: \frac{C_{E_c}}{\gamma^4 \delta^{4\tau}} \|E_c\|_\rho^2.$$

We have obtained the estimate (6.27). Notice that the second assumption in (6.15) and (6.28) imply that

$$(6.40) \quad \|\bar{E}_c\|_{\rho-2\delta} < \|E_c\|_\rho < \delta\nu.$$

Step 3, 4 and 5. All arguments and computations presented in these steps depend only on the invariance equation. Hence, the control of the new frames $L(\theta)$, $N(\theta)$ and the new transversality condition is exactly the same as in Lemma 5.2, but replacing $\|E\|_\rho$ by $\|E_c\|_\rho$. Specifically, we obtain the estimates (6.17) and (6.18) using Hypothesis (6.15) (they correspond to the third and fourth term in (6.16), respectively). We obtain the estimates (6.19) and (6.25) following the computations in (5.40) and (5.41) (using the fifth term in (6.16)).

Step 6: Control of the action of the left operator. Notice that the action of \mathfrak{L}_ω is affected by the change of the frequency, since now the natural operator to control is $\mathfrak{L}_{\bar{\omega}}$. From now on, given any object X , we introduce the operator

$$(6.41) \quad \Delta \mathfrak{L}X(\theta) := \mathfrak{L}_{\bar{\omega}} \bar{X}(\theta) - \mathfrak{L}_\omega X(\theta)$$

for the convenience of notation.

The control of $\mathfrak{L}_{\bar{\omega}} \bar{K}$ is straightforward, since

$$\|\mathfrak{L}_{\bar{\omega}} \bar{K}\|_{\rho-2\delta} \leq \|\bar{E}_c\|_{\rho-2\delta} + \|X_h \circ \bar{K}\|_{\rho-2\delta} \leq \delta\nu + c_{X_h,0} = C_{\mathfrak{L}K},$$

and similarly we obtain $\|\mathfrak{L}_{\bar{\omega}} \bar{L}\|_{\rho-3\delta} \leq C_{\mathfrak{L}L}$ and $\|\mathfrak{L}_{\bar{\omega}} \bar{L}^\top\|_{\rho-3\delta} \leq C_{\mathfrak{L}L^\top}$. However, to control increments of the form $\|\Delta \mathfrak{L}K\|_{\rho-2\delta}$ we need to include an additional term. More specifically,

$$\Delta \mathfrak{L}K(\theta) = \mathfrak{L}_{\Delta\omega} \bar{K}(\theta) + \mathfrak{L}_\omega \Delta K(\theta) = -\frac{\xi^\omega}{1-\xi^\omega} \mathfrak{L}_{\bar{\omega}} \bar{K}(\theta) + \mathfrak{L}_\omega \Delta K(\theta),$$

where we used that $\bar{\omega} = (1-\xi^\omega)\omega$, and, from bounds (6.32) and (6.36),

$$(6.42) \quad \|\Delta \mathfrak{L}K\|_{\rho-2\delta} \leq \left(\frac{\sigma_\omega C_{\xi^\omega} C_{\mathfrak{L}K}}{\gamma \delta^\tau} + \frac{C_{\mathfrak{L}\Delta K}}{\gamma^2 \delta^{2\tau}} \right) \|E_c\|_\rho =: \frac{C_{\Delta \mathfrak{L}K}}{\gamma^2 \delta^{2\tau}} \|E_c\|_\rho.$$

We now observe that this is the only estimate that must be updated, since it is the only place where cohomological equations play a role. For example, we have

$$\begin{aligned} \Delta \mathfrak{L}L(\theta) &= (\Delta \mathfrak{L}DK(\theta) \quad \Delta \mathfrak{L}(X_p \circ K)) \\ &= (D(\Delta \mathfrak{L}K(\theta)) \quad DX_p \circ \bar{K} \Delta \mathfrak{L}K + (DX_p \circ \bar{K} - DX_p \circ K) \mathfrak{L}_\omega K) \end{aligned}$$

and this expression yields formally to the same estimate in (5.52), but using the constants $C_{\Delta \mathfrak{L}K}$, $C_{\Delta K}$ and $C_{\mathfrak{L}K}$ defined in this section (and replacing E by E_c). This also affects to the control of the objects $\Delta \mathfrak{L}L^\top$, $\Delta \mathfrak{L}(G \circ K)$, $\Delta \mathfrak{L}G_L$, $\Delta \mathfrak{L}(\Omega \circ K)$, $\Delta \mathfrak{L}\tilde{\Omega}_L$, $\Delta \mathfrak{L}A$, $\Delta \mathfrak{L}(J \circ K)$, $\Delta \mathfrak{L}N^0$ and $\Delta \mathfrak{L}N$.

Step 7: Control of the new torsion condition. Now we consider the control of the extended torsion matrix $T_c(\theta)$ and the corresponding non-degeneracy condition. First, we observe that the upper-left block $T(\theta)$ is controlled as in Lemma 5.2. Thus, we control the extended torsion as

$$(6.43) \quad \begin{aligned} \|\bar{T}_c - T_c\|_{\rho-3\delta} &\leq \max \left\{ \frac{C_{\Delta T}}{\gamma^2 \delta^{2\tau+1}} + \frac{C_{\Delta\omega}}{\gamma \delta^\tau}, \frac{c_{c,1} C_{\Delta N}}{\gamma^2 \delta^{2\tau+1}} + \frac{c_{c,2} C_{\Delta K} C_N}{\gamma^2 \delta^{2\tau}} \right\} \|E_c\|_\rho \\ &=: \frac{C_{\Delta T_c}}{\gamma^2 \delta^{2\tau+1}} \|E_c\|_\rho. \end{aligned}$$

Finally, we obtain the estimates (6.20) and (6.26) by adapting the computations in (5.66) and (5.67). We use the second term in (6.16) (Hypothesis (6.15)) to get the estimate

$$(6.44) \quad |\langle \bar{T}_c \rangle^{-1} - \langle T_c \rangle^{-1}| \leq \frac{2(\sigma_{T_c})^2 C_{\Delta T_c}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho =: \frac{C_{\Delta \langle T_c \rangle^{-1}}}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho.$$

This completes the proof of the lemma. \square

6.3. Convergence of the iterative process. Now we are ready to prove our second KAM theorem with conserved quantities. Again, we comment the differences with respect to Theorem 3.1 and omit the common parts.

Proof of Theorem 3.6. Let us consider the approximate torus $K_0 := K$ with frequency $\omega_0 := \omega$ and with initial errors $E_0 := E$ and $E_0^\omega = E^\omega$. We also introduce $B_0 := B$, $T_0 := T$, $T_{c,0} = T_c$ and $E_{c,0} = E_c$ associated with the initial approximation. We reproduce the iterative construction in the proof of Theorem 3.1, but applying Lemma 6.2 recursively, and taking into account the evolution of the error $E_{c,s}$ at the s -step of the quasi-Newton method.

Computations are the same *mutatis mutandis*. In this case, we need to consider additional computations regarding the correction of the frequency. In particular, the eight term in (6.15) is checked as follows

$$\text{dist}(\omega_s, \partial\Theta) - \frac{C_{\Delta\omega} \|E_{c,s}\|_{\rho_s}}{\gamma\delta_s^\tau} > \text{dist}(\omega_0, \partial\Theta) - \frac{C_{\Delta\omega} \|E_{c,0}\|_{\rho_0}}{\gamma\delta_0^\tau} \frac{1}{1 - a_1^{-3\tau}} > 0,$$

where the last inequality is included into (3.15).

Having guaranteed all hypothesis of Lemma 6.2, we collect the assumptions by introducing the constant \mathfrak{C}_1 as

$$(6.45) \quad \mathfrak{C}_1 := \max \{ (a_1 a_3)^{4\tau} C_{E_c}, (a_3)^{2\tau+1} \gamma^2 \rho^{2\tau-1} C_\Delta \}$$

where

$$(6.46) \quad C_{\Delta,1} := \max \left\{ \frac{dC_{\Delta K}}{\sigma_K - \|\text{DK}\|_\rho}, \frac{2nC_{\Delta K}}{\sigma_{K^\top} - \|(\text{DK})^\top\|_\rho}, \frac{C_{\Delta B}}{\sigma_B - \|B\|_\rho}, \frac{C_{\Delta \langle T_c \rangle^{-1}}}{\sigma_{T_c} - |\langle T_c \rangle^{-1}|} \right\},$$

$$(6.47) \quad C_{\Delta,2} := \frac{C_{\Delta K} \delta}{\text{dist}(K(\mathbb{T}_\rho^d), \partial\mathcal{B})},$$

$$(6.48) \quad C_{\Delta,3} := \frac{C_{\Delta\omega} \gamma \delta^{\tau+1}}{\text{dist}(\omega, \partial\Theta)},$$

$$(6.49) \quad C_\Delta := \max \left\{ \frac{\gamma^2 \delta^{2\tau}}{\nu}, 2C_{\text{sym}} \gamma \delta^\tau, \frac{C_{\Delta,1}}{1 - a_1^{1-2\tau}}, \frac{C_{\Delta,2}}{1 - a_1^{-2\tau}}, \frac{C_{\Delta,3}}{1 - a_1^{-3\tau}} \right\}.$$

Note that we recovered the original notation $K = K_0$, $\omega = \omega_0$, $B = B_0$, $T_c = T_{c,0}$, $\rho = \rho_0$ and $\delta = \delta_0$ for the original objects.

Therefore, by induction, we can apply the iterative process infinitely many times. Indeed, we have

$$\|E_{c,s}\|_{\rho_s} < a_1^{-4\tau s} \|E_c\|_\rho \longrightarrow 0 \quad \text{when } s \rightarrow 0$$

so the iterative scheme converges to a true quasi-periodic torus K_∞ with frequency ω_∞ . As a result of the output of Lemma 6.2, these objects satisfy $K_\infty \in \mathcal{A}(\mathbb{T}_{\rho_\infty}^d)$, $\omega_\infty \in \Theta$ and

$$\|\text{DK}_\infty\|_{\rho_\infty} < \sigma_K, \quad \|(\text{DK}_\infty)^\top\|_{\rho_\infty} < \sigma_{K^\top}, \quad \text{dist}(K_\infty(\mathbb{T}^d), \partial\mathcal{B}) > 0.$$

Furthermore, we control the limit objects as follows:

$$(6.50) \quad \|K_\infty - K\|_{\rho_\infty} < \frac{C_{\Delta K} \|E_c\|_\rho}{\gamma^2 \delta^{2\tau}} \frac{1}{1 - a_1^{-2\tau}} =: \frac{\mathfrak{C}_2 \|E_c\|_\rho}{\gamma^2 \rho^{2\tau}},$$

$$(6.51) \quad |\omega_\infty - \omega| < \frac{C_{\Delta\omega} \|E_c\|_\rho}{\gamma\delta^\tau} \frac{1}{1 - a_1^{-3\tau}} =: \frac{\mathfrak{C}_3 \|E_c\|_\rho}{\gamma\rho^\tau},$$

thus obtaining the estimates in (3.16). This completes the proof of the generalized iso-energetic KAM theorem. \square

7. ACKNOWLEDGEMENTS

A. H. acknowledges the Spanish grants MTM2015-67724-P (MINECO/FEDER), MDM-2014-0445 (MINECO) and 2014 SGR 1145 (AGAUR), and the European Union's Horizon 2020 research and innovation programme MSCA 734557. A. L. acknowledges the Spanish grants MTM2016-76072-P (MINECO/FEDER) and SEV-2015-0554 (MINECO, Severo Ochoa Programme for Centres of Excellence in R&D), and the ERC Starting Grant 335079. We are also grateful to D. Peralta-Salas and J.-Ll. Figueras for fruitful discussions.

REFERENCES

- [1] 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. *Uspehi Mat. Nauk*, 18(5 (113)):13–40, 1963.
- [2] V.I. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russ. Math. Surveys*, 18:85–192, 1963.
- [3] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. *Nuovo Cimento B (11)*, 79(2):201–223, 1984.
- [4] M. Berti, L. Biasco, and M. Procesi. Kam for reversible derivative wave equations. *Archive for Rational Mechanics and Analysis*, 212(3):905–955, 2014.
- [5] H.W. Broer and G.B. Huitema. A proof of the iso-energetic kam-theorem from the ordinary one. *J. Differential Equations*, 90(1):52–60, 1991.
- [6] H.W. Broer, G.B. Huitema, and M.B. Sevryuk. *Quasi-periodic motions in families of dynamical systems. Order amidst chaos*. Lecture Notes in Math., Vol 1645. Springer-Verlag, Berlin, 1996.
- [7] H.W. Broer, G.B. Huitema, F. Takens, and B.L.J. Braaksma. Unfoldings and bifurcations of quasi-periodic tori. *Mem. Am. Math. Soc.*, 421:175 p., 1990.
- [8] H.W. Broer and F.O.O. Wagener. Quasi-periodic stability of subfamilies of an unfolded skew hopf bifurcation. *Arch. Rational Mechanics Analysis*, 152:283–326, 2000.
- [9] R. Calleja, A. Celletti, and R. de la Llave. A KAM theory for conformally symplectic systems: efficient algorithms and their validation. *J. Differential Equations*, 255(5):978–1049, 2013.
- [10] R. Calleja and R. de la Llave. A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification. *Nonlinearity*, 23(9):2029–2058, 2010.
- [11] M. Canadell and À. Haro. Computation of Quasi-Periodic Normally Hyperbolic Invariant Tori: Algorithms, Numerical Explorations and Mechanisms of Breakdown. *J. Nonlinear Sci.*, 27(6):1829–1868, 2017.
- [12] M. Canadell and À. Haro. Computation of Quasiperiodic Normally Hyperbolic Invariant Tori: Rigorous Results. *J. Nonlinear Sci.*, 27(6):1869–1904, 2017.
- [13] A. Cannas da Silva. *Lectures on symplectic geometry*, volume 1764 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.
- [14] A. Celletti and L. Chierchia. Construction of Analytic KAM Surfaces and Effective Stability Bounds. *Comm. Math. Phys.*, 118(1):199–161, 1988.
- [15] A. Celletti and L. Chierchia. A constructive theory of Lagrangian tori and computer-assisted applications. In *Dynamics Reported*, pages 60–129. Springer, Berlin, 1995.
- [16] A. Celletti and L. Chierchia. KAM stability and celestial mechanics. *Mem. Amer. Math. Soc.*, 187(878):viii+134, 2007.
- [17] A. Celletti and L. Chierchia. Quasi-periodic attractors in Celestial Mechanics. *Arch. Rational Mech. Anal.*, 191(2):311–345, 2009.

- [18] L. Chierchia and G. Pinzari. The planetary N -body problem: symplectic foliation, reductions and invariant tori. *Invent. Math.*, 186(1):1–77, 2011.
- [19] R. de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001.
- [20] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.
- [21] R. de la Llave and D. Rana. Accurate strategies for small divisor problems. *Bull. Amer. Math. Soc. (N.S.)*, 22(1):85–90, 1990.
- [22] A. Delshams and P. Gutiérrez. Effective Stability and KAM Theory. *J. Differential Equations*, 128:415–490, 1996.
- [23] 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.
- [24] L.H. Eliasson, B. Fayad, and R. Krikorian. Around the stability of KAM tori. *Duke Math. J.*, 164(9):1733–1775, 2015.
- [25] A. Enciso and D. Peralta-Salas. Existence of knotted vortex tubes in steady euler flows. *Acta Math.*, 214:61–134, 2015.
- [26] J. Féjoz. Démonstration du ‘théorème d’Arnold’ sur la stabilité du système planétaire (d’après Herman). *Ergodic Theory Dynam. Systems*, 24(5):1521–1582, 2004.
- [27] J.-Ll. Figueras, A. Haro, and A. Luque. Rigorous Computer-Assisted Application of KAM Theory: A Modern Approach. *Found. Comput. Math.*, 17(5):1123–1193, 2017.
- [28] J.-Ll. Figueras, A. Haro, and A. Luque. On the sharpness of the Rüssmann estimates. *Commun. Nonlinear Sci. Numer. Simul.*, 55:42–55, 2018.
- [29] E. Fontich, R. de la Llave, and Y. Sire. Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions. *J. Differential Equations*, 246(8):3136–3213, 2009.
- [30] A.M. Fox and J.D. Meiss. Critical invariant circles in asymmetric and multiharmonic generalized standard maps. *Commun. Nonlinear Sci. Numer. Simul.*, 19(4):1004–1026, 2014.
- [31] A. González, A. Haro, and R. de la Llave. Singularity theory for non-twist KAM tori. *Mem. Amer. Math. Soc.*, 227(1067):vi+115, 2014.
- [32] J.M. Greene. A method for determining a stochastic transition. *J. Math. Phys*, 20(6):1183–1201, 1975.
- [33] À. Haro, M. Canadell, J.-Ll. Figueras, A. Luque, and J.-M. Mondelo. *The parameterization method for invariant manifolds*, volume 195 of *Applied Mathematical Sciences*. Springer, [Cham], 2016. From rigorous results to effective computations.
- [34] G. Huguet, R. de la Llave, and Y. Sire. Computation of whiskered invariant tori and their associated manifolds: new fast algorithms. *Discrete Contin. Dyn. Syst.*, 32(4):1309–1353, 2012.
- [35] A.N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton’s function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954. 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.
- [36] A. Luque and J. Villanueva. A KAM theorem without action-angle variables for elliptic lower dimensional tori. *Nonlinearity*, 24(4):1033–1080, 2011.
- [37] R.S. MacKay. *Renormalisation in area-preserving maps*, volume 6 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co. Inc., River Edge, NJ, 1993.
- [38] J. Marsden and A. Weinstein. Reduction of symplectic manifolds with symmetry. *Rep. Mathematical Phys.*, 5(1):121–130, 1974.
- [39] J.E. Marsden, G. Misiołek, J.P. Ortega, M. Perlmutter, and Ratiu. T.S. *Hamiltonian reduction by stages*, volume 1913 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.
- [40] J. Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*, 1962:1–20, 1962.
- [41] J. Moser. On the theory of quasiperiodic motions. *SIAM Rev.*, 8(2):145–172, 1966.
- [42] J. Moser. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:499–535, 1966.
- [43] J. Moser. A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:265–315, 1966.

- [44] J.F. Palacián, F. Sayas, and P. Yanguas. Flow reconstruction and invariant tori in the spatial three-body problem. *Journal of Differential Equations*, 258:2114–2159, 2015.
- [45] J. Pöschel. A kam-theorem for some nonlinear partial differential equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 23(1):119–148, 1996.
- [46] 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)*, pages 598–624. Lecture Notes in Phys., Vol. 38. Springer, Berlin, 1975.
- [47] H. Rüssmann. On a new proof of Moser’s twist mapping theorem. In *Proceedings of the Fifth Conference on Mathematical Methods in Celestial Mechanics (Oberwolfach, 1975), Part I. Celestial Mech.*, 14(1):19–31, 1976.
- [48] 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. Celestial Mech.*, volume 14, 1976.
- [49] M.B. Sevryuk. Invariant tori in quasi-periodic non-autonomous dynamical systems via Herman’s method. *Discrete Contin. Dyn. Syst.*, 18(2-3):569–595, 2007.
- [50] M.B. Sevryuk. The classical KAM theory in the last decade: a slow progress. http://www.mathnet.ru/php/presentation.phtml?option_lang=eng&presentid=10271, 2014. The Seventh International Conference on Differential and Functional Differential Equations.
- [51] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.
- [52] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. II. *Comm. Pure Appl. Math.*, 29(1):49–111, 1976.

APPENDIX A. AN AUXILIARY LEMMA TO CONTROL THE INVERSE OF A MATRIX

To prove Lemma 5.2 and Lemma 6.2, we control the correction of inverses of matrices several times using Neumann series. This affects to the estimates in (5.16), (5.17), (6.25) and (6.26). 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 3.1).

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*

$$(A.1) \quad \frac{2\sigma^2|\bar{M} - M|}{\sigma - |M^{-1}|} \leq 1.$$

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

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

Proof. A direct computation shows that

$$(A.2) \quad \bar{M}^{-1} = (I_n + M^{-1}(\bar{M} - M))^{-1}M^{-1}.$$

By hypothesis (A.1) we obtain

$$(A.3) \quad |M^{-1}||\bar{M} - M| < \frac{\sigma^2|\bar{M} - M|}{\sigma - |M^{-1}|} < \frac{1}{2}.$$

Then a Neumann series argument in (A.2), using (A.3) and $|M^{-1}| < \sigma$, yields the estimate

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

Finally, we conclude that

$$\begin{aligned} |\bar{M}^{-1}| &\leq |M^{-1}| + |\bar{M}^{-1} - M^{-1}| \leq |M^{-1}| + 2\sigma^2|\bar{M} - M| \\ &< |M^{-1}| + \sigma - |M^{-1}| = \sigma, \end{aligned}$$

where we used again (A.1). \square

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 paper. Keeping track of these constants is crucial to apply the presented KAM theorems in particular problems and for concrete values of parameters. In addition, we think that the labels included in the tables will be of valuable help assisting the reading of the paper. Thus, the reader can find the place where a particular object is estimated.

Given an object $X : \mathbb{T}_{\rho_*}^n \rightarrow \mathbb{C}^{n_1 \times n_2}$, the following tables code an estimate of the form

$$\|X\|_{\rho_*} \leq \frac{C_X}{\gamma^{a_*} \delta^{b_*}} \|E_*\|_{\rho_*}^{c_*}.$$

Notice that the strip ρ_* and the exponents a_*, b_*, c_* can be tracked following the corresponding label; and E_* is the target error ($E_* = E$ for Theorem 3.1 and $E_* = E_c$ for Theorem 3.6).

Let us remark that, as it becomes clear in the proof, the numbers a_1, a_2, a_3 and ν are independent parameters, that can be selected in order to optimize the applicability of the theorems depending on the particular problem at hand.

Remark B.1. Table 1 corresponds to the geometric construction that is common to both theorems. The constants associated to the ordinary KAM theorem 3.1 are presented in Tables 2 and 3. The constants associated to the iso-energetic KAM theorem 3.6 are presented in Tables 4 and 5. To reduce the length of the tables, in the iso-energetic situation we have omitted those constants that have the same formula that in the ordinary case.

Table 1: Constants introduced in Section 4. Constants with * in the label are 0 in **Case III**. We denote by χ_0 the characteristic function of the set $\{0\}$.

Object	Constant	Label	Result
$\mathfrak{L}_\omega \Omega_K$	$C_{\mathfrak{L}\Omega_K} = 2nc_{\Omega,0}\sigma_K + \sigma_{K^\top} c_{\Omega,1}\sigma_K \delta + d\sigma_{K^\top} c_{\Omega,0}$	(4.14)	Lemma 4.3
Ω_K	$C_{\Omega_K} = c_R C_{\mathfrak{L}\Omega_K}$	(4.15)	Lemma 4.3
L	$C_L = \sigma_K + c_{X_p,0}$	(4.23)	Lemma 4.5
L^\top	$C_{L^\top} = \max\{\sigma_{K^\top}, c_{X_p^\top,0}\}$	(4.24)	Lemma 4.5
Ω_L	$C_{\Omega_L} = c_R \max\{C_{\mathfrak{L}\Omega_K} + c_{p^\top,1}, dc_{p,1}\}$	(4.25)	Lemma 4.5
G_L	$C_{G_L} = C_{L^\top} c_{G,0} C_L$	(4.26)	Lemma 4.5
$\tilde{\Omega}_L$	$C_{\tilde{\Omega}_L} = C_{L^\top} c_{\tilde{\Omega},0} C_L$	(4.27)	Lemma 4.5
N^0	$C_{N^0} = c_{J,0} C_L$	(4.31)	Lemma 4.6
$(N^0)^\top$	$C_{N^0,\top} = C_{L^\top} c_{J^\top,0}$	(4.32)	Lemma 4.6
A	$C_A = \frac{1}{2}(\sigma_B)^2 C_{\tilde{\Omega}_L}$	(4.33)*	Lemma 4.6
N	$C_N = C_L C_A + C_{N^0} \sigma_B$	(4.34)	Lemma 4.6
N^\top	$C_{N^\top} = C_A C_{L^\top} + \sigma_B C_{N^0,\top}$	(4.35)	Lemma 4.6
E_{sym}	$C_{\text{sym}} = (1 + C_A) \max\{1, C_A + (\sigma_B)^2 \chi_0(C_A)\} C_{\tilde{\Omega}_L}$	(4.39)	Lemma 4.6
$\mathfrak{L}_\omega K$	$C_{\mathfrak{L}K} = \delta\nu + c_{X_h,0}$	(4.44)	Lemma 4.8
$\mathfrak{L}_\omega L$	$C_{\mathfrak{L}L} = d\nu + c_{X_h,1}\sigma_K + c_{X_p,1} C_{\mathfrak{L}K}$	(4.45)	Lemma 4.8
$\mathfrak{L}_\omega L^\top$	$C_{\mathfrak{L}L^\top} = \max\{2n\nu + c_{X_h^\top,1}\sigma_K, c_{X_p^\top,1} C_{\mathfrak{L}K}\}$	(4.46)	Lemma 4.8
$\mathfrak{L}_\omega (J \circ K)$	$C_{\mathfrak{L}J} = c_{J,1} C_{\mathfrak{L}K}$	(4.48)	Lemma 4.8
$\mathfrak{L}_\omega (G \circ K)$	$C_{\mathfrak{L}G} = c_{G,1} C_{\mathfrak{L}K}$	(4.49)	Lemma 4.8
$\mathfrak{L}_\omega (\tilde{\Omega} \circ K)$	$C_{\mathfrak{L}\tilde{\Omega}} = c_{\tilde{\Omega},1} C_{\mathfrak{L}K}$	(4.50)	Lemma 4.8
$\mathfrak{L}_\omega N^0$	$C_{\mathfrak{L}N^0} = C_{\mathfrak{L}J} C_L + c_{J,0} C_{\mathfrak{L}L}$	(4.51)	Lemma 4.8
$\mathfrak{L}_\omega G_L$	$C_{\mathfrak{L}G_L} = C_{\mathfrak{L}L^\top} c_{G,0} C_L + C_{L^\top} C_{\mathfrak{L}G} C_L + C_{L^\top} c_{G,0} C_{\mathfrak{L}L}$	(4.52)	Lemma 4.8
$\mathfrak{L}_\omega \tilde{\Omega}_L$	$C_{\mathfrak{L}\tilde{\Omega}_L} = C_{\mathfrak{L}L^\top} c_{\tilde{\Omega},0} C_L + C_{L^\top} C_{\mathfrak{L}\tilde{\Omega}} C_L + C_{L^\top} c_{\tilde{\Omega},0} C_{\mathfrak{L}L}$	(4.53)	Lemma 4.8
$\mathfrak{L}_\omega B$	$C_{\mathfrak{L}B} = (\sigma_B)^2 C_{\mathfrak{L}G_L}$	(4.54)	Lemma 4.8
$\mathfrak{L}_\omega A$	$C_{\mathfrak{L}A} = C_{\mathfrak{L}B} C_{\tilde{\Omega}_L} \sigma_B + \frac{1}{2}(\sigma_B)^2 C_{\mathfrak{L}\tilde{\Omega}_L}$	(4.55)*	Lemma 4.8
$\mathfrak{L}_\omega N$	$C_{\mathfrak{L}N} = C_{\mathfrak{L}L} C_A + C_L C_{\mathfrak{L}A} + C_{\mathfrak{L}N^0} \sigma_B + C_{N^0} C_{\mathfrak{L}B}$	(4.56)	Lemma 4.8
\mathcal{X}_L	$C_{\mathcal{X}_L} = d + c_{X_p,1} \delta$	(4.57)	Lemma 4.8
\mathcal{X}_L^\top	$C_{\mathcal{X}_L^\top} = \max\{2n, c_{X_p^\top,1} \delta\}$	(4.58)	Lemma 4.8
\mathcal{X}_N	$C_{\mathcal{X}_N} = C_{\mathcal{X}_L} \nu C_A + c_{X_h,1} C_{N^0} \sigma_B + C_L C_{\mathfrak{L}A}$ $+ C_{\mathfrak{L}N^0} \sigma_B + C_{N^0} C_{\mathfrak{L}B}$	(4.59)	Lemma 4.8
T	$C_T = C_{N^\top} c_{\Omega,0} C_{\mathcal{X}_N}$	(4.60)	Lemma 4.8
$\mathfrak{L}_\omega \Omega_L$	$C_{\mathfrak{L}\Omega_L} = \max\{C_{\mathfrak{L}\Omega_K} + c_{p^\top,1}, dc_{p,1}\}$	(4.69)	Lemma 4.10
$E_{\text{red}}^{1,1}$	$C_{\text{red}}^{1,1} = C_{N^\top} c_{\Omega,0} C_{\mathcal{X}_L}$	(4.70)	Lemma 4.10
$E_{\text{red}}^{2,1}$	$C_{\text{red}}^{2,1} = C_{L^\top} c_{\Omega,0} C_{\mathcal{X}_L}$	(4.71)	Lemma 4.10
$E_{\text{red}}^{2,2}$	$C_{\text{red}}^{2,2} = (C_{L^\top} c_{\Omega,1} C_N \delta + C_{\mathcal{X}_L^\top} c_{\Omega,0} C_N + C_{\mathfrak{L}\Omega_L} C_A) \gamma \delta^\tau$ $+ C_{\Omega_L} C_{\mathfrak{L}A}$	(4.72)	Lemma 4.10
E_{red}	$C_{\text{red}} = \max\{C_{\text{red}}^{1,1} \gamma \delta^\tau, C_{\text{red}}^{2,1} \gamma \delta^\tau + C_{\text{red}}^{2,2}\}$	(4.73)	Lemma 4.10

Table 2: Constants introduced in Lemma 5.2. Given an object X , we use the notation $\Delta \mathfrak{L}X = \mathfrak{L}_\omega \bar{X} - \mathfrak{L}_\omega X$. Constants with * in the label are 0 in **Case III**.

Object	Constant	Label
ξ_0^N	$C_{\xi_0^N} = \sigma_T(C_{N^\top} c_{\Omega,0} \gamma \delta^\tau + c_R C_T C_{L^\top} c_{\Omega,0})$	(5.20)
ξ^N	$C_{\xi^N} = C_{\xi_0^N} + c_R C_{L^\top} c_{\Omega,0}$	(5.21)
ξ^L	$C_{\xi^L} = c_R(C_{N^\top} c_{\Omega,0} \gamma \delta^\tau + C_T C_{\xi^N})$	(5.22)
ξ	$C_\xi = \max\{C_{\xi^L}, C_{\xi^N} \gamma \delta^\tau\}$	(5.23)
ΔK	$C_{\Delta K} = C_L C_{\xi^L} + C_N C_{\xi^N} \gamma \delta^\tau$	(5.24)
$\mathfrak{L}_\omega \xi^N$	$C_{\mathfrak{L}_\omega \xi^N} = C_{L^\top} c_{\Omega,0}$	(5.26)
$\mathfrak{L}_\omega \xi^L$	$C_{\mathfrak{L}_\omega \xi^L} = C_{N^\top} c_{\Omega,0} \gamma \delta^\tau + C_T C_{\xi^N}$	(5.27)
$\mathfrak{L}_\omega \xi$	$C_{\mathfrak{L}_\omega \xi} = \max\{C_{\mathfrak{L}_\omega \xi^L}, C_{\mathfrak{L}_\omega \xi^N} \gamma \delta^\tau\}$	(5.28)
E_{lin}	$C_{\text{lin}} = C_{\text{red}} C_\xi + c_{\Omega,0} C_{\text{sym}} C_{\mathfrak{L}_\omega \xi} \gamma \delta^\tau$	(5.29)
\bar{E}	$C_E = 2(C_L + C_N) C_{\text{lin}} \gamma \delta^{\tau-1} + \frac{1}{2} c_{X_h,2} (C_{\Delta K})^2$	(5.32)
ΔL	$C_{\Delta L} = (d + c_{X_p,1} \delta) C_{\Delta K}$	(5.36)
ΔL^\top	$C_{\Delta L^\top} = \max\{2n, c_{X_p^\top,1} \delta\} C_{\Delta K}$	(5.37)
ΔG	$C_{\Delta G} = c_{G,1} C_{\Delta K}$	(5.38)
ΔG_L	$C_{\Delta G_L} = C_{L^\top} c_{G,0} C_{\Delta L} + C_{L^\top} C_{\Delta G} C_L \delta + C_{\Delta L^\top} c_{G,0} C_L$	(5.39)
ΔB	$C_{\Delta B} = 2(\sigma_B)^2 C_{\Delta G_L}$	(5.41)
$\Delta \bar{\Omega}$	$C_{\Delta \bar{\Omega}} = c_{\bar{\Omega},1} C_{\Delta K}$	(5.42)
$\Delta \bar{\Omega}_L$	$C_{\Delta \bar{\Omega}_L} = C_{L^\top} c_{\bar{\Omega},0} C_{\Delta L} + C_{L^\top} C_{\Delta \bar{\Omega}} C_L \delta + C_{\Delta L^\top} c_{\bar{\Omega},0} C_L$	(5.43)
ΔA	$C_{\Delta A} = \sigma_B C_{\bar{\Omega}_L} C_{\Delta B} + \frac{1}{2} (\sigma_B)^2 C_{\Delta \bar{\Omega}_L}$	(5.44)*
ΔJ	$C_{\Delta J} = c_{J,1} C_{\Delta K}$	(5.45)
ΔJ^\top	$C_{\Delta J^\top} = c_{J^\top,1} C_{\Delta K}$	(5.46)
ΔN^0	$C_{\Delta N^0} = c_{J,0} C_{\Delta L} + C_{\Delta J} C_L \delta$	(5.47)
ΔN^{0^\top}	$C_{\Delta N^{0^\top}} = C_{\Delta L^\top} c_{J^\top,0} + C_{L^\top} C_{\Delta J^\top} \delta$	(5.48)
ΔN	$C_{\Delta N} = C_L C_{\Delta A} + C_{\Delta L} C_A + C_{N^0} C_{\Delta B} + C_{\Delta N^0} \sigma_B$	(5.49)
ΔN^\top	$C_{\Delta N^\top} = C_A C_{\Delta L^\top} + C_{\Delta A} C_{L^\top} + C_{\Delta B} C_{N^0, \top} + \sigma_B C_{\Delta N^0 \top}$	(5.50)
$\mathfrak{L}_\omega \Delta K$	$C_{\mathfrak{L}_\omega \Delta K} = C_{\mathfrak{L}_\omega} C_{\xi^L} + (C_L C_{\mathfrak{L}_\omega \xi^L} + C_{\mathfrak{L}_\omega} C_{\xi^N}) \gamma \delta^\tau + C_N C_{\mathfrak{L}_\omega \xi^N} \gamma^2 \delta^{2\tau}$	(5.51)
$\Delta \mathfrak{L}L$	$C_{\Delta \mathfrak{L}L} = d C_{\Delta \mathfrak{L}K} + c_{X_p,1} C_{\Delta \mathfrak{L}K} \delta + c_{X_p,2} C_{\Delta K} C_{\mathfrak{L}K} \delta$	(5.52)
$\Delta \mathfrak{L}L^\top$	$C_{\Delta \mathfrak{L}L^\top} = \max\{2n C_{\Delta \mathfrak{L}K}, c_{X_p^\top,1} C_{\Delta \mathfrak{L}K} \delta + c_{X_p^\top,2} C_{\Delta K} C_{\mathfrak{L}K} \delta\}$	(5.53)
$\Delta \mathfrak{L}G$	$C_{\Delta \mathfrak{L}G} = c_{G,1} C_{\Delta \mathfrak{L}K} + c_{G,2} C_{\Delta K} C_{\mathfrak{L}K}$	(5.54)
$\Delta \mathfrak{L}G_L$	$C_{\Delta \mathfrak{L}G_L} = C_{\mathfrak{L}L^\top} c_{G,0} C_{\Delta L} + C_{\mathfrak{L}L^\top} c_{G,1} C_{\Delta K} C_L \delta + C_{\Delta \mathfrak{L}L^\top} c_{G,0} C_L$ $+ C_{L^\top} c_{G,1} C_{\mathfrak{L}K} C_{\Delta L} + C_{L^\top} C_{\Delta \mathfrak{L}G} C_L \delta + C_{\Delta L^\top} c_{G,1} C_{\mathfrak{L}K} C_L$ $+ C_{L^\top} c_{G,0} C_{\Delta \mathfrak{L}L} + C_{L^\top} c_{G,1} C_{\Delta K} C_{\mathfrak{L}L} \delta + C_{\Delta L^\top} c_{G,0} C_{\mathfrak{L}L},$	(5.56)
$\Delta \mathfrak{L}B$	$C_{\Delta \mathfrak{L}B} = 2\sigma_B C_{\mathfrak{L}G_L} C_{\Delta B} + (\sigma_B)^2 C_{\Delta \mathfrak{L}G_L}$	(5.57)
$\Delta \mathfrak{L}\bar{\Omega}$	$C_{\Delta \mathfrak{L}\bar{\Omega}} = c_{\bar{\Omega},1} C_{\Delta \mathfrak{L}K} + c_{\bar{\Omega},2} C_{\Delta K} C_{\mathfrak{L}K}$	(5.58)
$\Delta \mathfrak{L}\bar{\Omega}_L$	$C_{\Delta \mathfrak{L}\bar{\Omega}_L} = C_{\mathfrak{L}L^\top} c_{\bar{\Omega},0} C_{\Delta L} + C_{\mathfrak{L}L^\top} c_{\bar{\Omega},1} C_{\Delta K} C_L \delta + C_{\Delta \mathfrak{L}L^\top} c_{\bar{\Omega},0} C_L$ $+ C_{L^\top} c_{\bar{\Omega},1} C_{\mathfrak{L}K} C_{\Delta L} + C_{L^\top} C_{\Delta \mathfrak{L}\bar{\Omega}} C_L \delta + C_{\Delta L^\top} c_{\bar{\Omega},1} C_{\mathfrak{L}K} C_L$ $+ C_{L^\top} c_{\bar{\Omega},0} C_{\Delta \mathfrak{L}L} + C_{L^\top} c_{\bar{\Omega},1} C_{\Delta K} C_{\mathfrak{L}L} \delta + C_{\Delta L^\top} c_{\bar{\Omega},0} C_{\mathfrak{L}L},$	(5.59)
$\Delta \mathfrak{L}A$	$C_{\Delta \mathfrak{L}A} = C_{\mathfrak{L}B} C_{\bar{\Omega}_L} C_{\Delta B} + C_{\mathfrak{L}B} C_{\Delta \bar{\Omega}_L} \sigma_B + C_{\Delta \mathfrak{L}B} C_{\bar{\Omega}_L} \sigma_B$ $+ \sigma_B C_{\mathfrak{L}\bar{\Omega}_L} C_{\Delta B} + \frac{1}{2} (\sigma_B)^2 C_{\Delta \mathfrak{L}\bar{\Omega}_L}$	(5.60)*
$\Delta \mathfrak{L}J$	$C_{\Delta \mathfrak{L}J} = c_{J,1} C_{\Delta \mathfrak{L}L} + c_{J,2} C_{\Delta K} C_{\mathfrak{L}K}$	(5.61)
$\Delta \mathfrak{L}N^0$	$C_{\Delta \mathfrak{L}N^0} = C_{\mathfrak{L}J} C_{\Delta L} + C_{\Delta \mathfrak{L}J} C_L \delta + c_{J,0} C_{\Delta \mathfrak{L}L} + C_{\Delta J} C_{\mathfrak{L}L} \delta$	(5.62)
$\Delta \mathfrak{L}N$	$C_{\Delta \mathfrak{L}N} = C_{\mathfrak{L}L} C_{\Delta A} + C_{\Delta \mathfrak{L}L} C_A + C_L C_{\Delta \mathfrak{L}A} + C_{\Delta L} C_{\mathfrak{L}A}$ $+ C_{\mathfrak{L}N^0} C_{\Delta B} + C_{\Delta \mathfrak{L}N^0} \sigma_B + C_{N^0} C_{\Delta \mathfrak{L}B} + C_{\Delta N^0} C_{\mathfrak{L}B}$	(5.63)
$\Delta \mathfrak{L}X_N$	$C_{\Delta \mathfrak{L}X_N} = c_{X_p,1} C_{\Delta N} + c_{X_p,2} C_{\Delta K} C_N \delta + C_{\Delta \mathfrak{L}N}$	(5.64)
ΔT	$C_{\Delta T} = C_{N^\top} c_{\Omega,0} C_{\Delta X_N} + C_{N^\top} c_{\Omega,1} C_{\Delta K} C_{X_N} \delta + C_{\Delta N^\top} c_{\Omega,0} C_{X_N}$	(5.65)
$\Delta(T)^{-1}$	$C_{\Delta(T)^{-1}} = 2(\sigma_T)^2 C_{\Delta T}$	(5.67)

Table 3: Constants introduced in Section 5.3 associated with the convergence of the quasi-Newton method that yields Theorem 3.1, i.e., the ordinary formulation.

Constant	Label
$\mathfrak{C}_1 = \max \{ (a_1 a_3)^{4\tau} C_E, (a_3)^{2\tau+1} \gamma^2 \rho^{2\tau-1} C_\Delta \}$	(5.74)
$C_{\Delta,1} = \max \left\{ \frac{dC_{\Delta K}}{\sigma_K - \ \mathbb{D}K\ _\rho}, \frac{2n C_{\Delta K}}{\sigma_{K^\top} - \ \mathbb{D}K^\top\ _\rho}, \frac{C_{\Delta B}}{\sigma_B - \ B\ _\rho}, \frac{C_{\Delta(T)}^{-1}}{\sigma_T - (T)^{-1} } \right\}$	(5.75)
$C_{\Delta,2} = \frac{C_{\Delta K} \delta}{\text{dist}(K(\mathbb{T}_\rho^d), \partial \mathcal{B})}$	(5.76)
$C_\Delta = \max \left\{ \frac{\gamma^2 \delta^{2\tau}}{\nu}, 2C_{\text{sym}} \gamma \delta^\tau, \frac{C_{\Delta,1}}{1 - a_1^{-2\tau}}, \frac{C_{\Delta,2}}{1 - a_1^{-2\tau}} \right\}$	(5.77)
$\mathfrak{C}_2 = a_3^{2\tau} C_{\Delta K} / (1 - a_1^{-2\tau})$	(5.78)
$\mathfrak{C}_3 = c_{c,1} \mathfrak{C}_2$	(5.79)

Table 4: Constants introduced in Lemma 6.2, i.e. in the Iterative Lemma for the generalized iso-energetic KAM theorem. In this table, given an object X , we use the notation $\Delta \mathfrak{L}X = \mathfrak{L}_{\bar{\omega}} \bar{X} - \mathfrak{L}_\omega X$.

Object	Constant	Label
ξ_0^N	$C_{\xi_0^N} = \sigma_{T_c} \max \{ C_{N^\top} c_{\Omega,0} \gamma \delta^\tau + c_R C_T C_{L^\top} c_{\Omega,0}, \gamma \delta^\tau + c_R c_{c,1} C_N C_{L^\top} c_{\Omega,0} \}$	(6.29)
ξ^ω	$C_{\xi^\omega} = C_{\xi_0^N}$	(6.29)
ω	$C_\omega = \sigma_\omega \omega_* $	(6.30)
$\Delta \omega$	$C_{\Delta \omega} = C_\omega C_{\xi_0^N}$	(6.31)
$\mathfrak{L}_\omega \xi^L$	$C_{\mathfrak{L}_\omega \xi^L} = C_{N^\top} c_{\Omega,0} \gamma \delta^\tau + C_T C_{\xi^N} + C_\omega C_{\xi^\omega}$	(6.33)
E_{lin}	$C_{\text{lin}} = C_{\text{red}} C_\xi + c_{\Omega,0} C_{\text{sym}} C_{\mathfrak{L}_\omega \xi} \gamma \delta^\tau + \max\{C_A, 1\} C_{\Omega_L} C_\omega C_{\xi^\omega}$	(6.34)
E_{lin}^ω	$C_{\text{lin}}^\omega = d c_R c_{c,1} C_{\xi^L}$	(6.35)
$\mathfrak{L}_\omega \Delta K$	$C_{\mathfrak{L}_\omega \Delta K} = C_{\mathfrak{L}_\omega \xi^L} + (C_L C_{\mathfrak{L}_\omega \xi^L} + C_{\mathfrak{L}_\omega \xi^N}) \gamma \delta^\tau + C_N C_{\mathfrak{L}_\omega \xi^N} \gamma^2 \delta^{2\tau}$	(6.36)
\bar{E}	$C_E = 2(C_L + C_N) C_{\text{lin}} \gamma \delta^{\tau-1} + \frac{1}{2} c_{X_h,2} (C_{\Delta K})^2 + C_{\xi^\omega} C_{\mathfrak{L}_\omega \Delta K} \gamma \delta^\tau$	(6.37)
\bar{E}^ω	$C_{E^\omega} = C_{\text{lin}}^\omega \gamma \delta^{\tau-1} + \frac{1}{2} c_{c,2} (C_{\Delta K})^2$	(6.38)
\bar{E}_c	$C_{E_c} = \max\{C_E, C_{E^\omega}\}$	(6.39)
$\Delta \mathfrak{L}K$	$C_{\Delta \mathfrak{L}K} = \sigma_\omega C_{\xi^\omega} C_{\mathfrak{L}_\omega \xi} \gamma \delta^\tau + C_{\mathfrak{L}_\omega \Delta K}$	(6.42)
ΔT_c	$C_{\Delta T_c} = \max\{C_{\Delta T} + C_{\Delta \omega} \gamma \delta^{\tau+1}, c_{c,1} C_{\Delta N} + c_{c,2} C_{\Delta K} C_N \delta\}$	(6.43)
$\Delta(T_c)^{-1}$	$C_{\Delta(T_c)^{-1}} = 2(\sigma_{T_c})^2 C_{\Delta T_c}$	(6.44)

Table 5: Constants introduced in Section 5.3 associated with the convergence of the quasi-Newton method that yields Theorem 3.6, i.e., the iso-energetic formulation.

Constant	Label
$\mathfrak{C}_1 = \max \{ (a_1 a_3)^{4\tau} C_{E_c}, (a_3)^{2\tau+1} \gamma^2 \rho^{2\tau-1} C_\Delta \}$	(6.45)
$C_{\Delta,1} = \max \left\{ \frac{dC_{\Delta K}}{\sigma_K - \ \mathbb{D}K\ _\rho}, \frac{2n C_{\Delta K}}{\sigma_{K^\top} - \ \mathbb{D}K^\top\ _\rho}, \frac{C_{\Delta B}}{\sigma_B - \ B\ _\rho}, \frac{C_{\Delta(T_c)}^{-1}}{\sigma_{T_c} - (T_c)^{-1} } \right\}$	(6.46)
$C_{\Delta,2} = \frac{C_{\Delta K} \delta}{\text{dist}(K(\mathbb{T}_\rho^d), \partial \mathcal{B})}$	(6.47)
$C_{\Delta,3} = \frac{C_{\Delta \omega} \gamma \delta^{\tau+1}}{\text{dist}(\omega, \partial \Theta)}$	(6.48)
$C_\Delta = \max \left\{ \frac{\gamma^2 \delta^{2\tau}}{\nu}, 2C_{\text{sym}} \gamma \delta^\tau, \frac{C_{\Delta,1}}{1 - a_1^{-2\tau}}, \frac{C_{\Delta,2}}{1 - a_1^{-2\tau}}, \frac{C_{\Delta,3}}{1 - a_1^{-3\tau}} \right\}$	(6.49)
$\mathfrak{C}_2 = a_3^{2\tau} C_{\Delta K} / (1 - a_1^{-2\tau})$	(6.50)
$\mathfrak{C}_3 = a_3^\tau C_{\Delta \omega} / (1 - a_1^{-3\tau})$	(6.51)