

Rigorous computer assisted application of KAM theory: a modern approach

Jordi-Lluís Figueras*

Department of Mathematics
Uppsala University
Box 480, 751 06 Uppsala (Sweden).

Alex Haro[†]

Departament de Matemàtica Aplicada i Anàlisi
Universitat de Barcelona
Gran Via 585, 08007 Barcelona (Spain).

Alejandro Luque[‡]

Instituto de Ciencias Matemáticas
Consejo Superior de Investigaciones Científicas
C/ Nicolás Cabrera 13-15, 28049 Madrid (Spain).

26th August 2016

Abstract

In this paper we present and illustrate a general methodology to apply KAM theory in particular problems, based on an *a posteriori* approach. We focus on the existence of real-analytic quasi-periodic Lagrangian invariant tori for symplectic maps. The purpose is to verify the hypotheses of a KAM theorem in an *a posteriori* format: given a parameterization of an approximately invariant torus, we have to check non-resonance (Diophantine) conditions, non-degeneracy conditions and certain inequalities to hold. To check such inequalities we require to control the analytic norm of some functions that depend on the map, the ambient structure and the parameterization. To this end, we propose an efficient computer assisted methodology, using fast Fourier transform, having the same asymptotic cost of using the parameterization method for obtaining numerical approximations of invariant tori. We illustrate our methodology by proving the existence of invariant curves for the standard map (up to $\varepsilon = 0.9716$), meandering curves for the non-twist standard map and 2-dimensional tori for the Froeschlé map.

Mathematics Subject Classification: 37J40; 65G20; 65G40; 65T50;

Keywords: a posteriori KAM theory; computer-assisted proofs; Rüssmann estimates; fast Fourier transform.

*figueras@math.uu.se

†alex@maia.ub.es

‡luque@icmat.es

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 2 | A KAM theorem for Lagrangian invariant tori of exact symplectic maps | 4 |
| 2.1 | Geometric setting and invariant tori | 5 |
| 2.2 | Analytic functions and norms | 7 |
| 2.3 | Statement of the KAM theorem | 8 |
| 2.4 | Proof of the KAM theorem | 9 |
| 3 | On the approximation of periodic functions using discrete Fourier transform | 18 |
| 3.1 | Notation regarding discretization of the torus and Fourier transforms | 18 |
| 3.2 | Error estimates on the approximation of analytic periodic functions | 18 |
| 3.3 | Comments on the 1-dimensional case | 22 |
| 3.4 | Matrices of periodic functions | 23 |
| 4 | Dealing with the small divisors | 24 |
| 4.1 | On the characterization of Diophantine constants | 24 |
| 4.2 | On the Rüssmann estimates | 27 |
| 5 | Validation algorithm to apply the KAM theorem | 28 |
| 5.1 | A validation algorithm | 30 |
| 5.2 | Implementation details of Step 0 | 32 |
| 5.3 | Implementation details of Step 1 | 32 |
| 5.4 | Implementation details of Step 2 | 33 |
| 5.5 | Implementation details of Step 3 | 34 |
| 5.6 | Implementation details of Step 4 | 34 |
| 6 | Application of the KAM theorem in some examples | 35 |
| 6.1 | Standard map | 36 |
| 6.2 | Non-twist standard map | 39 |
| 6.3 | Froeschlé map | 42 |
| 7 | Final remarks and perspectives | 45 |
| | References | 46 |
| A | A heuristic selection of parameters to validate invariant tori | 50 |

1 Introduction

KAM theory is concerned with the existence and stability of quasi-periodic motions in different contexts of dynamical systems such as symplectic maps, Hamiltonian systems, reversible systems or volume-preserving systems, just to mention a few. The foundations of the theory started with the celebrated works of A.N. Kolmogorov [46], V.I. Arnold [2], and J.K. Moser [59], so that the acronym KAM is used in their honor. These pioneer papers sowed the seed of a subject of remarkable importance in dynamical systems and, nowadays, KAM theory is a vast area of research that involves a large collection of methods. Actually, there are many excellent surveys covering different points of view in the theory (e.g. [3, 5, 6, 19, 33, 62]).

Classic KAM methods typically deal with a perturbative setting in such a way that the problem is written as a perturbation of an integrable system (in the sense that it has a continuous family of invariant tori). They are based on

the use of canonical transformations to simplify the expression of the problem. To this end, one takes advantage of the existence of action-angle-like coordinates for the unperturbed system. This is a source of different shortcomings and limitations in the study of particular problems, mainly related to the fact that many systems are non-perturbative. For example, in some cases it is possible to identify an integrable approximation of a given system but the remaining part cannot be considered as an arbitrarily small perturbation. Moreover, given a particular perturbative problem, in general it is very complicated to establish action-angle variables for the unperturbed system. Such action-angle variables can be defined implicitly, become singular or introduce problems of regularity.

In spite of the previous difficulties, classic KAM methods have been successfully applied in several problems. The interested reader is referred to Section 1.4 in [16] for a brief history and references of the application of KAM theory, and to [15, 48, 49] for computer assisted proofs in problems of celestial mechanics. A prominent example is the persistence of the golden invariant curve of the standard map (c.f. Section 6.1 and notation therein). From the numerical point of view, the persistence of this golden curve has been considered for example in [18, 35, 53] observing that the breakdown takes place around $\varepsilon_c \simeq 0.97163540324$. Upper bounds for ε_c were provided in [55, 56] by Converse KAM Theory. A quite sharp non-existence result was reported in [43], proving that the standard map has no rotational invariant circles for several parameter values including $\varepsilon = 0.9718$. KAM theory provides lower bounds for the critical value ε_c . This was already considered by Herman [40], obtaining that the golden curve persists for $\varepsilon \leq 0.029$. Later, a computer assisted proof was given in [13] proving existence of the invariant curve for $\varepsilon < 0.68$ using Lindstedt series. This lower bound was improved in [22, 23] extending the result up to $\varepsilon \leq 0.91$.

An alternative to the classic approach is the use of the parameterization method. Instead of performing canonical transformations, the strategy consists in solving the invariance equation directly by correcting an approximately invariant object. Such correction is obtained iteratively by considering the linearized equation around the previous approximately invariant torus. The parameterization method is suitable for studying existence of invariant tori without using neither action-angle variables nor being in a perturbative setting. We point out that the geometry of the problem plays an important role in the study of these equations. Such geometric approach, also referred to as KAM theory *without action-angle variables* or *a posteriori* KAM theory, was suggested by R. de la Llave in [19] (following long-time developed ideas, e.g. [14, 15, 42, 61, 68, 70, 78]) and a complete proof was presented in [20]. This approach has been later extended to other contexts, such as the study of lower dimensional (isotropic) invariant tori that are hyperbolic [30] or elliptic [51], the case of non-twist invariant tori in degenerate systems [34] or, more recently, dissipative systems [7, 11]. A remarkable advantage of the parameterization method is that the steps of the proof allow us to obtain very fast and efficient numerical methods for the approximation of quasi-periodic invariant tori (e.g. [8, 31, 41]). We refer the reader to the recent survey [37] for a detailed discussion on the numerical implementation of the method and examples.

The goal of this paper is to present and illustrate a general methodology to apply the KAM theorem in specific problems. We focus on the existence of analytic quasi-periodic Lagrangian invariant tori of symplectic maps. We resort to a revisited version of the *a posteriori* KAM result presented in [20]. As usual in KAM theory, the main hypotheses consist in checking non-resonance (Diophantine) conditions, non-degeneracy conditions and also asking certain inequalities to hold. To check such inequalities we require to control the analytic norm of some functions that depend on the known objects (the map, the ambient structure and the initial parameterization). To this end, we propose a rigorous computer assisted methodology based on the use of fast Fourier transform. An important consequence of our methodology is that the application of the KAM theorem is performed in a very fast way. Indeed, with the same asymptotic cost of using the parameterization method to obtain numerical approximations of invariant tori (which is of the same order of the fast Fourier transform: $O(N \log(N))$, N being the total number of used Fourier coefficients). Of course, the asymptotic constant depends on the dimension of the phase space, the map, and the symplectic form.

It is worth mentioning that computer assisted analysis has played an important role to achieve remarkable results in the literature. Among them, we highlight: the proof of the Feigenbaum conjecture [45, 47]; Rigorous interval methods in quantum mechanics [28]; the proof that the Lorenz attractor exists [74]; the proof of the double bubbling conjecture [39], and the existence of singular solutions in fluid dynamics [12]. Computer assisted methods in analysis rely on the fact that one can define a rigorous interval arithmetic on computers. These intervals have computer representable floating point numbers as end-points and all basic operations as addition, subtraction, multiplication, division and

composition of standard functions (e.g. \exp , \cos , \sin , \log) satisfy the *isotonicity inclusion principle* (the image of any two nested intervals is nested) and the *range enclosure principle* (the range of any function is enclosed with the image of the domain under the action of the natural interval extension). We refer the reader to [75] for more details.

Finally, we describe the organization of the paper and we briefly summarize the content of each section:

- In Section 2 we introduce some elementary geometric objects (Section 2.1), set up notation and norms (Section 2.2), and present a detailed statement of the KAM theorem (Section 2.3). Then we present a full proof of this result (Section 2.4). This is necessary in order to link the different expressions that appear in the constants of the theorem with their corresponding geometric object or equation. We give explicit and sharp estimates for all constants quantifying the hypotheses of the theorem.
- In Section 3 we control the difference between an analytic function f on the torus and its discrete Fourier approximation \tilde{f} , and we present several technical results that allow us to control (with explicit constants) the analytic norm of $\tilde{f} - f$. More concretely, if f is an analytic and bounded function on the complex strip of size $\hat{\rho} > 0$, then

$$\|\tilde{f} - f\|_{\rho} \leq C_{N_F}(\rho, \hat{\rho}) \|f\|_{\hat{\rho}},$$

for every $0 \leq \rho < \hat{\rho}$, where $C_{N_F}(\rho, \hat{\rho})$ is an explicit constant that depends also on the dimension and the number of Fourier coefficients. This result is motivated by the previous work [27] and related ideas have been used in [57, 71]. Our aim is that this section can be read independently, in spite of some notation introduced in Section 2.2 regarding Fourier series and norms.

- In Section 4 we consider additional technical results that allow us to apply the KAM theorem in an effective and efficient way. On the one hand, we present an approach to obtain a positive measure set of Diophantine vectors close to a given vector, possibly obtained numerically (Section 4.1). On the other hand, we present an improvement of the classic Rüssmann estimates (Section 4.2). To take into account the effect of small divisors, we compute the first elements explicitly and then we control the remaining tail analytically.
- In Section 5 we present the main methodology to apply the KAM result: the validation algorithm. The validation procedure is performed on a sampling of an approximately invariant torus obtained numerically. We also require a finite amount of input data characterizing the geometric information of the problem. Suitable values for these parameters can be obtained following an heuristic method explained in Appendix A.
- In Section 6, we apply the validation algorithm to several examples, thus highlighting different features of our approach.

In order to illustrate the reliability of the method, we consider the standard map (Section 6.1) and prove that the golden invariant curve exists up to $\varepsilon = 0.9716$, thus establishing a new lower bound to the so-called Greene critical value [35] $\varepsilon_c \simeq 0.97163540324$. Moreover, we have also proved the existence of other rotational invariant curves with different rotation numbers.

An important feature of the method is that it can be applied to invariant curves that are not graphs over the angular coordinate. We consider the non-twist standard map and prove the existence of so-called meandering invariant curves (Section 6.2).

We finally consider a higher dimensional example. We prove the existence of 2-dimensional invariant tori for the Froeschlé map, a 4-dimensional symplectic map consisting in two coupled standard maps (Section 6.3).

2 A KAM theorem for Lagrangian invariant tori of exact symplectic maps

In this section we present an *a posteriori* KAM theorem for Lagrangian invariant tori of exact symplectic maps. The result was first obtained in [20] and it is a version of Kolmogorov theorem [46] using neither action-angle coordinates nor a perturbative setting. The specific statement given here, with explicit and sharp estimates, is a slightly modified

version of the KAM Theorem discussed in Chapter 4 of [37]. Roughly speaking, the *a posteriori* result reads as follows: *if we have a good enough approximation of an invariant torus with frequency ω , then, under certain non-degeneracy and non-resonance conditions, there exists a true invariant torus nearby.*

After setting the problem and geometrical background in Section 2.1, we introduce some basic notation regarding Banach spaces, norms and cohomological equations in Section 2.2. In Section 2.3 we present the statement of a KAM theorem for existence (and persistence) of Lagrangian invariant tori having Diophantine frequencies. In the proof of the main result, given in Section 2.4, we pay special attention to compute explicitly all constants appearing during the process and to obtain optimal and sharp estimates.

2.1 Geometric setting and invariant tori

We denote $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$ the n -dimensional torus with covering space. The ambient manifold is a $2n$ -dimensional annulus $\mathcal{A} \subset \mathbb{T}^n \times \mathbb{R}^n$ with covering space $\tilde{\mathcal{A}} \subset \mathbb{R}^{2n}$. The coordinates on \mathcal{A} are denoted by $z = (z_1, \dots, z_{2n}) = (x, y)$, with $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$. A function $u : \mathbb{R}^n \rightarrow \mathbb{R}$ is 1-periodic if $u(\theta + e) = u(\theta)$ for all $\theta \in \mathbb{R}^n$ and $e \in \mathbb{Z}^n$. We abuse notation and denote it as $u : \mathbb{T}^n \rightarrow \mathbb{R}$. Similarly, a function $g : \tilde{\mathcal{A}} \rightarrow \mathbb{R}$ is 1-periodic in x if $g(x + e, y) = g(x, y)$ for all $x, y \in \mathbb{R}^n$ and $e \in \mathbb{Z}^n$. We abuse notation and denote it as $g : \mathcal{A} \rightarrow \mathbb{R}$.

In the following we assume that \mathcal{A} is endowed with an exact symplectic form $\omega = d\alpha$ for a certain 1-form α . For any point $z \in \mathcal{A}$, we write the matrix representation of the 1-form α_z and the 2-form ω_z as

$$a(z) = (a_1(z) \dots a_{2n}(z))^\top, \quad \text{and} \quad \Omega(z) = Da(z)^\top - Da(z), \quad (1)$$

respectively. Notice that $\det \Omega(z) \neq 0$.

Remark 2.1. *The prototype example of symplectic structure is the standard symplectic structure on $\mathbb{T}^n \times \mathbb{R}^n$: $\omega_0 = \sum_{i=1}^n dz_{n+i} \wedge dz_i$. 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,*

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

Remark 2.2. *The fact that the phase space is chosen to be the annulus \mathcal{A} is not related to the use of action-angle coordinates. We make this choice for the sake of simplicity. All the arguments, computations and ideas can be straightforwardly adapted to consider that the phase-space is an open set of \mathbb{R}^{2n} (in fact this is the setting in [20]). What is important is that the presented approach does not require to implement any sequence of canonical transformations in the problem.*

A map $F : \mathcal{A} \rightarrow \mathcal{A}$ is *symplectic* if $F^*\omega = \omega$. A symplectic map $F : \mathcal{A} \rightarrow \mathcal{A}$ is *exact* if there is a smooth function $S : \mathcal{A} \rightarrow \mathbb{R}$, called *primitive function of F* , such that $F^*\alpha - \alpha = dS$. In coordinates, the symplectic and the exact symplectic properties of a map F are equivalent to

$$DF(z)^\top \Omega(F(z)) DF(z) = \Omega(z), \quad \forall z \in \mathcal{A}, \quad (2)$$

and

$$DS(z) = a(F(z))^\top DF(z) - a(z)^\top, \quad \forall z \in \mathcal{A}, \quad (3)$$

respectively. A map $F : \mathcal{A} \rightarrow \mathcal{A}$ is *homotopic to the identity* if $F(x, y) - (x, 0)$ is 1-periodic in x .

Given an embedding $K : \mathbb{T}^n \rightarrow \mathcal{A}$, called *parameterization* from now on, we say that $K(\mathbb{T}^n)$ is an *F -invariant torus* with frequency $\omega \in \mathbb{R}^n$ if

$$F(K(\theta)) = K(\theta + \omega). \quad (4)$$

For convenience, we denote $R_\omega(\theta) = \theta + \omega$ the rigid rotation of frequency ω . In case that ω is rationally independent (i.e., $k \cdot \omega \notin \mathbb{Z}$ for all $k \in \mathbb{Z}^d \setminus \{0\}$) then the rotation R_ω is *ergodic* and the invariant torus $K(\mathbb{T}^n)$ is *quasi-periodic*. Finally, the parameterization $K : \mathbb{T}^n \rightarrow \mathcal{A}$ is *homotopic to the zero section* if $K(\theta) - (\theta, 0)$ is 1-periodic in θ .

Remark 2.3. *If K is homotopic to the zero section, then $K(\mathbb{T}^n)$ is called primary tori. In the classic KAM perturbative setting, these objects correspond to continuation of the planar tori that are present in the unperturbed problem. The methodology presented in this paper can be adapted to deal with invariant tori having other relative homotopies in a straightforward way. In the context of Remark 2.2, if the phase space is an open set of \mathbb{R}^{2n} (there is no angular variables in phase space) then one simply looks for a parameterization $K(\theta)$ that is 1-periodic in θ .*

By taking derivatives at both sides of Equation (4) we observe that the tangent bundle is invariant. Indeed,

$$DF(K(\theta))DK(\theta) = DK(\theta + \omega). \quad (5)$$

Given a parameterization K as described above, we consider the pullback $K^*\omega$. Its matrix representation at a point $K(\theta)$ is

$$\Omega_K(\theta) = DK(\theta)^\top \Omega(K(\theta)) DK(\theta). \quad (6)$$

It is well known (c.f. [60]) that if $K(\mathbb{T}^n)$ is a quasi-periodic F -invariant torus, then $\Omega_K(\theta) = O_n$ for every $\theta \in \mathbb{T}^n$. In combination with Equation (5), this means that we have a Lagrangian invariant subbundle.

Roughly speaking, the parameterization method for proving existence of quasi-periodic F -invariant tori consists in studying the linearized invariance equation around an approximate solution. The invariance of $DK(\theta)$ in Equation (5) suggests that these vectors will help us to obtain a suitable frame to write the map DF . The fact that every Lagrangian subspace has a Lagrangian complementary is the starting point of the general construction discussed in [37] which is followed in this paper (for previous constructions we refer to [20, 30, 51, 34]). This construction goes as follows: given a map $N_0 : \mathbb{T}^n \rightarrow \mathbb{R}^{2n \times n}$ such that

$$\det(DK(\theta)^\top \Omega(K(\theta))N_0(\theta)) \neq 0 \quad (7)$$

it turns out that the *frame map* $P : \mathbb{T}^n \rightarrow \mathbb{R}^{2n \times 2n}$, given by

$$P(\theta) = \begin{pmatrix} DK(\theta) & N_0(\theta) \end{pmatrix}, \quad (8)$$

with

$$N(\theta) = DK(\theta)A(\theta) + N_0(\theta)B(\theta), \quad (9)$$

$$B(\theta) = - (DK(\theta)^\top \Omega(K(\theta))N_0(\theta))^{-1}, \text{ and} \quad (10)$$

$$A(\theta) = - \frac{1}{2} (B(\theta)^\top N_0(\theta)^\top \Omega(K(\theta))N_0(\theta)B(\theta)), \quad (11)$$

is a symplectic frame (N is the Lagrangian complement of DK). Since the dynamics on the torus is ergodic, it follows that the symplectic frame in Equation (8) reduces the linearized dynamics $DF \circ K$ to a block-triangular matrix

$$P(\theta + \omega)^{-1}DF(K(\theta))P(\theta) = \Lambda(\theta), \quad \Lambda(\theta) = \begin{pmatrix} I_n & T(\theta) \\ O_n & I_n \end{pmatrix}, \quad (12)$$

where the *torsion matrix* $T : \mathbb{T}^n \rightarrow \mathbb{R}^{n \times n}$ is given by

$$T(\theta) = N(\theta + \omega)^\top \Omega(K(\theta + \omega)) DF(K(\theta)) N(\theta). \quad (13)$$

Remark 2.4. *Of course, if we endow the annulus with additional structure (e.g. a Riemannian metric) we can obtain N_0 in a natural way according to this structure. A summary of different approaches used in the literature can be found in Chapter 4 of [37].*

2.2 Analytic functions and norms

In this paper we work with Banach spaces of real analytic functions in complex neighborhoods of real domains. A complex strip of \mathbb{T}^n of width $\rho > 0$ is defined as

$$\mathbb{T}_\rho^n = \{\theta \in \mathbb{C}^n / \mathbb{Z}^n : |\operatorname{Im} \theta_i| < \rho, i = 1, \dots, n\}.$$

A function defined on \mathbb{T}^n is *real analytic* if it can be analytically extended to a complex strip \mathbb{T}_ρ^n .

We consider analytic functions $u : \mathbb{T}_\rho^n \rightarrow \mathbb{C}$ such that they can be continuously extended up to the boundary of \mathbb{T}_ρ^n . We endow these functions with the norm

$$\|u\|_\rho = \sup_{\theta \in \mathbb{T}_\rho^n} |u(\theta)|. \quad (14)$$

Moreover, we write the Fourier expansion

$$u(\theta) = \sum_{k \in \mathbb{Z}^n} u_k e^{2\pi i k \cdot \theta}, \quad u_k = \int_{[0,1]^n} u(\theta) e^{-2\pi i k \cdot \theta} d\theta,$$

and we denote the average of u as $\langle u \rangle = u_0 = \int_{[0,1]^n} u(\theta) d\theta$. Then, we consider the Fourier norm

$$\|u\|_{F,\rho} = \sum_{k \in \mathbb{Z}^n} |u_k| e^{2\pi |k|_1 \rho}, \quad (15)$$

where $|k|_1 = \sum_{i=1}^n |k_i|$. We observe that $\|u\|_\rho \leq \|u\|_{F,\rho}$ for every $\rho > 0$.

A complex strip of \mathcal{A} is a complex connected open neighborhood $\mathcal{B} \subset (\mathbb{C}^n / \mathbb{Z}^n) \times \mathbb{C}^n$ of \mathcal{A} that projects surjectively on \mathbb{T}^n . A function defined on \mathcal{A} is *real analytic* if it can be analytically extended to a complex strip \mathcal{B} . Given an analytic function $u : \mathcal{B} \rightarrow \mathbb{C}$ we introduce the norm

$$\|u\|_{\mathcal{B}} = \sup_{z \in \mathcal{B}} |u(z)|. \quad (16)$$

The previous definitions extend naturally to matrices. If A is an $n_1 \times n_2$ matrix of analytic functions on \mathbb{T}_ρ^n (resp. on \mathcal{B}), we extend the norms in Equations (14) and (15) (resp. Equation (16)) as follows

$$\|A\|_\rho = \max_{i=1, \dots, n_1} \sum_{j=1}^{n_2} \|A_{i,j}\|_\rho, \quad \|A\|_{F,\rho} = \max_{i=1, \dots, n_1} \sum_{j=1}^{n_2} \|A_{i,j}\|_{F,\rho}, \quad (\text{resp. } \|A\|_{\mathcal{B}}). \quad (17)$$

Notice that, if $F : \mathcal{A} \rightarrow \mathcal{A}$, Ω is the matrix representation of ω and a is the matrix representation of α , then

$$\begin{aligned} \|DF\|_{\mathcal{B}} &= \max_{i=1, \dots, 2n} \sum_{j=1}^{2n} \left\| \frac{\partial F_i}{\partial z_j} \right\|_{\mathcal{B}}, & \|D^2 F\|_{\mathcal{B}} &= \max_{i=1, \dots, 2n} \sum_{j,k=1}^{2n} \left\| \frac{\partial^2 F_i}{\partial z_j \partial z_k} \right\|_{\mathcal{B}}, \\ \|\Omega\|_{\mathcal{B}} &= \max_{i=1, \dots, 2n} \sum_{j=1}^{2n} \|\Omega_{i,j}\|_{\mathcal{B}}, & \|D\Omega\|_{\mathcal{B}} &= \max_{i=1, \dots, 2n} \sum_{j,k=1}^{2n} \left\| \frac{\partial \Omega_{i,j}}{\partial z_k} \right\|_{\mathcal{B}}, \\ \|Da\|_{\mathcal{B}} &= \max_{i=1, \dots, 2n} \sum_{j=1}^{2n} \left\| \frac{\partial a_i}{\partial z_j} \right\|_{\mathcal{B}}, & \|D^2 a\|_{\mathcal{B}} &= \max_{i=1, \dots, 2n} \sum_{j,k=1}^{2n} \left\| \frac{\partial^2 a_i}{\partial z_j \partial z_k} \right\|_{\mathcal{B}}, \end{aligned}$$

Finally, we introduce some useful notation regarding the so-called *cohomological equations* that play an important role in KAM theory. Given $\omega \in \mathbb{R}^n$, we define the cohomology operator \mathcal{L} on functions $u : \mathbb{T}^n \rightarrow \mathbb{R}$ as follows:

$$\mathcal{L}u = u - u \circ R_\omega. \quad (18)$$

Then, the core of KAM theory is the cohomological equation

$$\mathcal{L}u = v - \langle v \rangle, \quad (19)$$

for a given periodic function v .

Let us assume that v is a continuous function and R_ω is ergodic. If there exists a continuous zero-average solution of Equation (19), then it is unique and will be denoted by $u = \mathcal{R}v$. Note that the formal solution of Equation (19) is immediate. Actually, if v has the Fourier expansion $v(\theta) = \sum_{k \in \mathbb{Z}^n} \hat{v}_k e^{2\pi i k \cdot \theta}$ and the dynamics is ergodic, then

$$\mathcal{R}v(\theta) = \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \hat{u}_k e^{2\pi i k \cdot \theta}, \quad \hat{u}_k = \frac{\hat{v}_k}{1 - e^{2\pi i k \cdot \omega}}. \quad (20)$$

In particular, this implies that $\mathcal{R}v = 0$ if $v = 0$. The solutions of Equation (19) differ by a constant (the average).

We point out that ergodicity is not enough to ensure regularity of the solutions of cohomological equations. This is related to the effect of the small divisors $1 - e^{2\pi i k \cdot \omega}$ in Equation (20). To deal with regularity, we require stronger non-resonant conditions on the vector of frequencies. In this paper, we consider the following classic condition:

Definition 2.5. *Given $\gamma > 0$ and $\tau \geq n$, we say that $\omega \in \mathbb{R}^n$ is a (γ, τ) -Diophantine vector of frequencies if*

$$|k \cdot \omega - m| \geq \gamma |k|_1^{-\tau}, \quad \forall k \in \mathbb{Z}^n \setminus \{0\}, m \in \mathbb{Z}, \quad (21)$$

where $|k|_1 = \sum_{i=1}^n |k_i|$.

Finally, we recall the so-called Rüssmann estimates to control the regularity of the solutions of Equation (19) (we refer the reader to [67]). If $v : \mathbb{T}^n \rightarrow \mathbb{R}$ is analytic, with $\|v\|_\rho < \infty$ and ω satisfies (21), then

$$\|\mathcal{R}v\|_{\rho-\delta} \leq \frac{c_R}{\gamma \delta^\tau} \|v\|_\rho \quad (22)$$

for $0 < \delta \leq \rho$. In Lemma 4.3 we present an improvement of the classic Rüssmann constant c_R with the help of the computer.

The above definitions for \mathcal{L} and \mathcal{R} extend component-wise to vector and matrix-valued functions. These extensions also satisfy the Rüssmann estimates.

2.3 Statement of the KAM theorem

At this point, we are ready to state sufficient conditions to guarantee the existence of an F -invariant torus with fixed frequency close to an approximately F -invariant torus. Theorems of this type are often called *a posteriori* results. Notice that the hypotheses in Theorem 2.6 are tailored to be verified with a finite amount of computations.

Theorem 2.6. *Let us consider an exact symplectic structure $\omega = d\alpha$ on the n -dimensional annulus \mathcal{A} , an exact symplectic map $F : \mathcal{A} \rightarrow \mathcal{A}$ homotopic to the identity and a frequency vector $\omega \in \mathbb{R}^n$. Let us assume that the following hypotheses hold:*

H_1 *The map F , the 1-form α and the 2-form ω are real analytic and can be analytically extended to some complex strip \mathcal{B} and continuously up to the boundary. Moreover, there are constants $c_{DF}, c_{D^2F}, c_\Omega, c_{D\Omega}, c_{Da}$ and c_{D^2a} such that $\|DF\|_{\mathcal{B}} \leq c_{DF}$, $\|D^2F\|_{\mathcal{B}} \leq c_{D^2F}$, $\|\Omega\|_{\mathcal{B}} \leq c_\Omega$, $\|D\Omega\|_{\mathcal{B}} \leq c_{D\Omega}$, $\|Da\|_{\mathcal{B}} \leq c_{Da}$, and $\|D^2a\|_{\mathcal{B}} \leq c_{D^2a}$.*

H_2 *There exists $K : \mathbb{T}^n \rightarrow \mathcal{A}$, homotopic to the zero section, that can be analytically extended to \mathbb{T}_ρ^n with $\rho > 0$, and continuously up to the boundary, with $K(\mathbb{T}_\rho^n) \subset \mathcal{B}$. Moreover, there exist constants σ_{DK} and σ_{DK^\top} such that*

$$\|DK\|_\rho < \sigma_{DK}, \quad \|DK^\top\|_\rho < \sigma_{DK^\top}, \quad \text{dist}(K(\mathbb{T}_\rho^n), \partial\mathcal{B}) > 0.$$

Given two subsets $X, Y \in \mathbb{C}^{2n}$, $\text{dist}(X, Y)$ is defined as $\inf\{|x - y| : x \in X, y \in Y\}$, where $|\cdot|$ is the maximum norm.

H_3 There exists a map $N_0 : \mathbb{T}^n \rightarrow \mathbb{R}^{2n \times n}$ that is real analytic and can be analytically extended to \mathbb{T}_ρ^n , and continuously up to the boundary. Moreover, there exist constants c_{N_0} , $c_{N_0^\top}$, $\sigma_{N_0^\top(\Omega \circ K)N_0}$, and σ_B such that

$$\|N_0\|_\rho \leq c_{N_0}, \quad \|N_0^\top\|_\rho \leq c_{N_0^\top}, \quad \|N_0^\top(\Omega \circ K)N_0\|_\rho < \sigma_{N_0^\top(\Omega \circ K)N_0}, \quad \|B\|_\rho < \sigma_B,$$

where $B(\theta) = -(\mathrm{DK}(\theta)^\top \Omega(K(\theta))N_0(\theta))^{-1}$.

H_4 There exists σ_T such that the matrix-valued map

$$T(\theta) = N(\theta + \omega)^\top \Omega(K(\theta + \omega)) \mathrm{DF}(K(\theta)) N(\theta)$$

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

$$N(\theta) = \mathrm{DK}(\theta)A(\theta) + N_0(\theta)B(\theta),$$

with $A(\theta) = -\frac{1}{2}(B(\theta)^\top N_0(\theta)^\top \Omega(K(\theta))N_0(\theta)B(\theta))$.

H_5 The frequency vector ω is (γ, τ) -Diophantine for certain $\gamma > 0$ and $\tau \geq n$.

Under the above hypotheses, for each $0 < \rho_\infty < \rho$ there exists a constant \mathfrak{C}_1 (see Remark 2.8) such that, if the following condition holds

$$\frac{\mathfrak{C}_1 \|E\|_\rho}{\gamma^4 \rho^{4\tau}} < 1, \quad (23)$$

where $E(\theta) = F(K(\theta)) - K(\theta + \omega)$, then there exists an F -invariant torus $K_\infty(\mathbb{T}^n)$ with frequency ω . The map K_∞ is an embedding, homotopic to the zero section, analytic in $\mathbb{T}_{\rho_\infty}^n$, and satisfies

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

Furthermore, the map K_∞ is close to K : there exists a constant \mathfrak{C}_2 (see Remark 2.8) such that

$$\|K_\infty - K\|_{\rho_\infty} < \frac{\mathfrak{C}_2 \|E\|_\rho}{\gamma^2 \rho^{2\tau}}. \quad (24)$$

Remark 2.7. We use the symbol σ to denote those constants that control objects that are corrected iteratively, and so, we have to ensure that the control prevails along the proof.

Remark 2.8. Constants \mathfrak{C}_1 and \mathfrak{C}_2 , which are given explicitly later in Equations (72) and (73), depend explicitly on the initial data. Concretely, they depend polynomially on c_{DF} , c_{D^2F} , c_Ω , $c_{\mathrm{D}\Omega}$, $c_{\mathrm{D}a}$, c_{D^2a} , c_{N_0} , and $c_{N_0^\top}$. They also depend polynomially on $(\sigma_{\mathrm{DK}} - \|\mathrm{DK}\|_\rho)^{-1}$, $(\sigma_{\mathrm{DK}^\top} - \|\mathrm{DK}^\top\|_\rho)^{-1}$, $(\sigma_{N_0^\top(\Omega \circ K)N_0} - \|N_0^\top(\Omega \circ K)N_0\|)^{-1}$, $(\sigma_B - \|B\|_\rho)^{-1}$, $(\sigma_T - |\langle T \rangle^{-1}|)^{-1}$ and $\mathrm{dist}(K(\mathbb{T}_\rho^n), \partial\mathcal{B})^{-1}$, and on the strict estimations σ_{DK} , $\sigma_{\mathrm{DK}^\top}$, $\sigma_{N_0^\top(\Omega \circ K)N_0}$, σ_B , and σ_T , respectively. If we fix $c_R > 0$ then \mathfrak{C}_1 and \mathfrak{C}_2 depend polynomially on n , c_R , γ and powers of ρ . These constants can be optimized by selecting a suitable value of ρ_∞ and adjusting the rate of converge of the iterative scheme.

2.4 Proof of the KAM theorem

The proof follows from a standard KAM scheme. Although a detailed proof of a very similar statement is given in Chapter 4 of [37], for the sake of completeness we present here a compact self-contained exposition. This allows us to describe the main geometric objects so that the reader can relate them to a corresponding constant that contributes to the computation of \mathfrak{C}_1 and \mathfrak{C}_2 .

The argument consists in refining $K(\theta)$ by means of a Newton method. At every step, we add to $K(\theta)$ a correction $\Delta K(\theta)$, given by an approximate solution of the linearized equation

$$\mathrm{DF}(K(\theta))\Delta K(\theta) - \Delta K(\theta + \omega) = -E(\theta). \quad (25)$$

To face this equation we consider a suitable frame on the full tangent space. The main ingredient is the fact that (under certain assumptions) an approximately F -invariant torus is also approximately Lagrangian. Hence, the linear dynamics around the torus is approximately reducible. Specifically, it turns out that we have a behavior similar to Equation (12) but with an error of order $\|E\|_\rho$. This is enough to perform a quadratic scheme to correct the initial approximation.

Lemma 2.9 (The Iterative Lemma). *Let us consider the same setting and hypotheses of Theorem 2.6. Then, there exist constants $\hat{C}_1, \hat{C}_2, \hat{C}_3, \hat{C}_4,$ and \hat{C}_5 (depending explicitly on the constants defined in the hypotheses) such that if*

$$\frac{\hat{\mathcal{C}}_1 \|E\|_\rho}{\gamma^2 \delta^{2\tau+1}} < 1 \quad (26)$$

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

$$\hat{\mathcal{C}}_1 := \max \left\{ \hat{C}_1 \gamma \delta^\tau, \frac{n \hat{C}_2}{\sigma_{\text{DK}} - \|\text{DK}\|_\rho}, \frac{2n \hat{C}_2}{\sigma_{\text{DK}^\top} - \|\text{DK}^\top\|_\rho}, \right. \\ \left. \frac{c_{N_0^\top} c_{\text{D}\Omega} c_{N_0} \delta \hat{C}_2}{\sigma_{N_0^\top(\Omega \circ K)N_0} - \|N_0^\top(\Omega \circ K)N_0\|_\rho}, \frac{\hat{C}_3}{\sigma_B - \|B\|_\rho}, \frac{\hat{C}_4}{\sigma_T - |\langle T \rangle^{-1}|}, \frac{\hat{C}_2 \delta}{\text{dist}(K(\mathbb{T}_\rho^n), \partial \mathcal{B})} \right\}, \quad (27)$$

then we have an approximate F -invariant torus of the same frequency ω given by $\bar{K} = K + \Delta K$, that defines new objects \bar{B} and \bar{T} (obtained replacing K by \bar{K}) satisfying

$$\|\text{D}\bar{K}\|_{\rho-3\delta} < \sigma_{\text{DK}}, \quad \|\text{D}\bar{K}^\top\|_{\rho-3\delta} < \sigma_{\text{DK}^\top}, \quad \text{dist}(\bar{K}(\mathbb{T}_{\rho-2\delta}^n), \partial \mathcal{B}) > 0, \quad (28)$$

$$\|\bar{B}\|_{\rho-3\delta} < \sigma_B, \quad |\langle \bar{T} \rangle^{-1}| < \sigma_T, \quad \|N_0^\top(\Omega \circ \bar{K})N_0\|_{\rho-2\delta} < \sigma_{N_0^\top(\Omega \circ K)N_0} \quad (29)$$

and

$$\|\bar{K} - K\|_{\rho-2\delta} < \frac{\hat{C}_2}{\gamma^2 \delta^{2\tau}} \|E\|_\rho, \quad \|\bar{B} - B\|_{\rho-3\delta} < \frac{\hat{C}_3}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho, \quad |\langle \bar{T} \rangle^{-1} - \langle T \rangle^{-1}| < \frac{\hat{C}_4}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho. \quad (30)$$

The new error of invariance is given by

$$\bar{E}(\theta) = F(\bar{K}(\theta)) - \bar{K}(\theta + \omega), \quad \|\bar{E}\|_{\rho-2\delta} < \frac{\hat{C}_5}{\gamma^4 \delta^{4\tau}} \|E\|_\rho^2. \quad (31)$$

Before proving Lemma 2.9, we present two auxiliary results.

Lemma 2.10. *Let us consider vector-valued maps $\eta = (\eta^{\text{DK}}, \eta^N) : \mathbb{T}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ and a matrix-valued map $T : \mathbb{T}^n \rightarrow \mathbb{R}^{n \times n}$. Assume that T satisfies the non-degeneracy condition $\det \langle T(\theta) \rangle \neq 0, \forall \theta \in \mathbb{T}^n$. Then, the system of equations*

$$\begin{pmatrix} I_n & T(\theta) \\ O_n & I_n \end{pmatrix} \begin{pmatrix} \xi^{\text{DK}}(\theta) \\ \xi^N(\theta) \end{pmatrix} - \begin{pmatrix} \xi^{\text{DK}}(\theta + \omega) \\ \xi^N(\theta + \omega) \end{pmatrix} = \begin{pmatrix} \eta^{\text{DK}}(\theta) \\ \eta^N(\theta) - \langle \eta^N \rangle \end{pmatrix}$$

has a (formal) solution $\xi = (\xi^{\text{DK}}, \xi^N) : \mathbb{T}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ given by

$$\xi^N(\theta) = \mathcal{R}(\eta^N(\theta)) + \xi_0^N, \quad (32)$$

$$\xi^{\text{DK}}(\theta) = \mathcal{R}(\eta^{\text{DK}}(\theta) - T(\theta)\xi^N(\theta)) + \xi_0^{\text{DK}}, \quad (33)$$

for every $\xi_0^{\text{DK}} \in \mathbb{R}^n$, and

$$\xi_0^N = \langle T \rangle^{-1} \langle \eta^{\text{DK}} - T\mathcal{R}(\eta^N) \rangle. \quad (34)$$

Note that \mathcal{R} gives the zero-average solution of the one bite cohomological equation (see Equation (19)).

Proof. The triangular form of this system allows us to face first the equation $\mathcal{L}\xi^N(\theta) = \eta^N(\theta) - \langle \eta^N \rangle$, where \mathcal{L} is given by Equation (18). The right hand side of this equation has already zero average, so we obtain the solution in (32), where $\xi_0^N = \langle \eta^N \rangle \in \mathbb{R}^n$. Then, the upper equation is $\mathcal{L}\eta^{\text{DK}}(\theta) = \eta^N(\theta) - T(\theta)\xi^N(\theta)$ and the vector ξ_0^N selected in (34) allows us to guarantee that $\langle \eta^{\text{DK}} - T\xi^N \rangle = 0$. In this way, we obtain the solution in (33). \square

Lemma 2.11. *If $K(\theta)$ is an approximately F -invariant torus with error $E(\theta)$, then*

$$\left\langle \text{DK}(\theta + \omega)^\top \Omega(K(\theta + \omega))E(\theta) \right\rangle = \left\langle \text{DE}(\theta)^\top \Delta a(\theta) + \text{DK}(\theta + \omega)^\top \Delta^2 a(\theta) \right\rangle,$$

where

$$\begin{aligned} \Delta a(\theta) &= a(F(K(\theta))) - a(K(\theta + \omega)) = \int_0^1 \text{Da}(K(\theta + \omega) + tE(\theta))E(\theta)dt, \\ \Delta^2 a(\theta) &= a(F(K(\theta))) - a(K(\theta + \omega)) - \text{Da}(K(\theta + \omega))E(\theta) \\ &= \int_0^1 (1-t)\text{D}^2 a(K(\theta + \omega) + tE(\theta))(E(\theta), E(\theta))dt. \end{aligned}$$

Proof. From the definition of Ω in (1), and some easy computations,

$$\begin{aligned} &\text{DK}(\theta + \omega)^\top \Omega(K(\theta + \omega))E(\theta) \\ &= \text{DK}(\theta + \omega)^\top \text{Da}(K(\theta + \omega))^\top E(\theta) - \text{DK}(\theta + \omega)^\top \text{Da}(K(\theta + \omega))E(\theta) \\ &= (\text{D}(a(K(\theta + \omega))))^\top E(\theta) + \text{DK}(\theta + \omega)^\top (\Delta^2 a(\theta) - a(F(K(\theta))) + a(K(\theta + \omega))) \\ &= \left(\text{D}(a(K(\theta + \omega))^\top E(\theta)) \right)^\top - (\text{DE}(\theta))^\top a(K(\theta + \omega)) + \text{DK}(\theta + \omega)^\top \Delta^2 a(\theta) \\ &\quad - (\text{DF}(K(\theta))\text{DK}(\theta) - \text{DE}(\theta))^\top a(F(K(\theta))) + \text{DK}(\theta + \omega)^\top a(K(\theta + \omega)) \\ &= \left(\text{D}(a(K(\theta + \omega))^\top E(\theta)) \right)^\top + (\text{DE}(\theta))^\top \Delta a(\theta) + \text{DK}(\theta + \omega)^\top \Delta^2 a(\theta) \\ &\quad - (\text{D}(S(K(\theta))))^\top - \text{DK}(\theta)^\top a(K(\theta)) + \text{DK}(\theta + \omega)^\top a(K(\theta + \omega)), \end{aligned}$$

where in the last identity we use that S is the primitive function of F , see Equation (3). The result follows by taking averages and realizing that $\text{D}(a(K(\theta + \omega))^\top E(\theta))$, $\text{D}(S(K(\theta)))$ and $a(K(\theta + \omega))^\top \text{DK}(\theta + \omega) - a(K(\theta))^\top \text{DK}(\theta)$ have zero average. \square

Proof of Lemma 2.9. In the first part of the proof we see that, since $K(\mathbb{T}^n)$ is approximately F -invariant, the frame $P(\theta)$ is symplectic up to an error controlled by $E(\theta)$.

We start by controlling the objects N , B and A , given in Equations (9), (10), and (11), respectively. By hypothesis, we have $\|\text{DK}\|_\rho < \sigma_{\text{DK}}$ and $\|B\|_\rho < \sigma_B$. Then, we obtain

$$\|A\|_\rho = \|A^\top\|_\rho \leq \frac{1}{2}\|B^\top N_0^\top (\Omega \circ K) N_0 B\|_\rho \leq \frac{1}{2}n\sigma_{N_0^\top (\Omega \circ K) N_0}(\sigma_B)^2 =: c_A. \quad (35)$$

where the constant c_A is introduced in order to simplify subsequent computations. We use small letters (c_A , c_N , etc.) when the constant is related to an estimation of a geometric object, using the subscript to identify the corresponding object. We use capital letters (C_1 , C_2 , etc.) for constants that appear in estimates that depend on the error $\|E\|_\rho$ (divisors are considered separately).

We estimate the norm of N as

$$\|N\|_\rho \leq \|\text{DK}\|_\rho \|A\|_\rho + \|N_0\|_\rho \|B\|_\rho \leq \sigma_{\text{DK}}c_A + c_{N_0}\sigma_B =: c_N$$

and

$$\|N^\top\|_\rho \leq c_A\sigma_{\text{DK}^\top} + n\sigma_B c_{N_0^\top} =: c_{N^\top}.$$

The frame $P(\theta)$, given by Equation (8), satisfies

$$\|P\|_\rho \leq \|DK\|_\rho + \|N\|_\rho \sigma_{DK} + c_N =: c_P$$

and the torsion $T(\theta)$, given by Equation (13), is controlled by

$$\|T\|_\rho \leq \|N^\top\|_\rho \|\Omega\|_{\mathcal{B}} \|DF\|_{\mathcal{B}} \|N\|_\rho \leq c_N^\top c_\Omega c_{DF} c_N =: c_T.$$

Now we control the approximate Lagrangian character of $K(\mathbb{T}^n)$. Taking derivatives at both sides of $E(\theta) = F(K(\theta)) - K(\theta + \omega)$ we have

$$DF(K(\theta))DK(\theta) = DK(\theta + \omega) + DE(\theta). \quad (36)$$

Then, a direct computation of $\mathcal{L}\Omega_K(\theta)$, using Equations (18) and (36), leads to

$$\begin{aligned} \mathcal{L}\Omega_K(\theta) = & DK(\theta + \omega)^\top \Delta\Omega(\theta) DK(\theta + \omega) + DK(\theta + \omega)^\top \Omega(F(K(\theta))) DE(\theta) \\ & + DE(\theta)^\top \Omega(F(K(\theta))) DF(K(\theta)) DK(\theta), \end{aligned} \quad (37)$$

where

$$\Delta\Omega(\theta) = \Omega(F(K(\theta))) - \Omega(K(\theta + \omega)) = \int_0^1 D\Omega(K(\theta + \omega) + tE(\theta))E(\theta)dt. \quad (38)$$

Using the Mean Value Theorem for integrals and properties of Banach algebras we obtain $\|\Delta\Omega\|_\rho \leq c_{D\Omega}\|E\|_\rho$, and introducing this expression into Equation (37) we control $\|\mathcal{L}\Omega_K\|_{\rho-\delta}$ as follows (we use Cauchy estimates)

$$\|\mathcal{L}\Omega_K\|_{\rho-\delta} \leq \left(\sigma_{DK}^\top \sigma_{DK} c_{D\Omega} \delta + n \sigma_{DK}^\top c_\Omega + 2n c_\Omega c_{DF} \sigma_{DK} \right) \frac{\|E\|_\rho}{\delta} =: \frac{C_1}{\delta} \|E\|_\rho.$$

Then, using the Rüssmann estimates (see Equation (22) or Lemma 4.3) we end up with

$$\|\Omega_K\|_{\rho-2\delta} \leq \frac{c_R C_1}{\gamma \delta^{\tau+1}} \|E\|_\rho =: \frac{C_2}{\gamma \delta^{\tau+1}} \|E\|_\rho.$$

Next, we introduce the error in the symplectic character of the frame as follows

$$E_{\text{sym}}(\theta) = P(\theta)^\top \Omega(K(\theta)) P(\theta) - \Omega_0 \quad (39)$$

and a straightforward computation shows that

$$E_{\text{sym}}(\theta) = \begin{pmatrix} \Omega_K(\theta) & \Omega_K(\theta)A(\theta) \\ A(\theta)^\top \Omega_K(\theta) & A(\theta)^\top \Omega_K(\theta)A(\theta) \end{pmatrix}, \quad (40)$$

which is controlled by

$$\|E_{\text{sym}}\|_{\rho-2\delta} \leq \frac{(1 + c_A) \max\{1, c_A\} C_2}{\gamma \delta^{\tau+1}} \|E\|_\rho =: \frac{C_3}{\gamma \delta^{\tau+1}} \|E\|_\rho. \quad (41)$$

Next, we show that the tangent map DF is approximately reducible in the frame $P(\theta)$. To this end, we introduce

$$E_{\text{red}}(\theta) = -\Omega_0 P(\theta + \omega)^\top \Omega(K(\theta + \omega)) DF(K(\theta)) P(\theta) - \Lambda(\theta), \quad (42)$$

where $\Lambda(\theta)$ is given by Equation (12). We decompose $E_{\text{red}}(\theta)$ into four $(n \times n)$ -block components given by:

$$E_{\text{red}}^{1,1}(\theta) = N(\theta + \omega)^\top \Omega(K(\theta + \omega)) DE(\theta) + A(\theta + \omega)^\top \Omega_K(\theta + \omega), \quad (43)$$

$$E_{\text{red}}^{1,2}(\theta) = N(\theta + \omega)^\top \Omega(K(\theta + \omega)) DF(K(\theta)) N(\theta) - T(\theta) = O_n$$

$$E_{\text{red}}^{2,1}(\theta) = -\Omega_K(\theta + \omega) - DK(\theta + \omega)^\top \Omega(K(\theta + \omega)) DE(\theta), \quad (44)$$

$$\begin{aligned} E_{\text{red}}^{2,2}(\theta) = & -\Omega_K(\theta)A(\theta) + DK(\theta + \omega)^\top \Delta\Omega(\theta) DF(K(\theta)) N(\theta) \\ & + DE(\theta)^\top \Omega(F(K(\theta))) DF(K(\theta)) N(\theta). \end{aligned} \quad (45)$$

Then, we conclude that the error of reducibility satisfies

$$\|E_{\text{red}}\|_{\rho-2\delta} \leq \frac{\max\{C_4, C_5 + C_6\}}{\gamma\delta^{\tau+1}} \|E\|_{\rho} =: \frac{C_7}{\gamma\delta^{\tau+1}} \|E\|_{\rho}$$

where

$$\begin{aligned} C_4 &= nc_{N^\top}c_{\Omega}\gamma\delta^{\tau} + c_A C_2, \\ C_5 &= C_2 + n\sigma_{\text{DK}^\top}c_{\Omega}\gamma\delta^{\tau}, \\ C_6 &= c_A C_2 + \sigma_{\text{DK}^\top}c_{\text{D}\Omega}c_{\text{D}F}c_N\gamma\delta^{\tau+1} + 2nc_{\Omega}c_{\text{D}F}c_N\gamma\delta^{\tau}. \end{aligned} \quad (46)$$

Now, we study Equation (25) using the symplectic frame in (8): we introduce $\Delta K(\theta) = P(\theta)\xi(\theta)$ thus obtaining

$$DF(K(\theta))P(\theta)\xi(\theta) - P(\theta + \omega)\xi(\theta + \omega) = -E(\theta).$$

We multiply both sides by $-\Omega_0 P(\theta + \omega)^\top \Omega(K(\theta + \omega))$ and we get

$$\begin{aligned} \Lambda(\theta)\xi(\theta) + E_{\text{red}}(\theta)\xi(\theta) - (I - \Omega_0 E_{\text{sym}}(\theta + \omega))\xi(\theta + \omega) = \\ \Omega_0 P(\theta + \omega)^\top \Omega(K(\theta + \omega))E(\theta), \end{aligned} \quad (47)$$

where we used Equations (39) and (42). In order to obtain an approximate solution of Equation (47), we consider Lemma 2.10 taking

$$\eta(\theta) = \Omega_0 P(\theta + \omega)^\top \Omega(K(\theta + \omega))E(\theta), \quad (48)$$

and $T(\theta)$ given by Equation (13). We choose the solution satisfying $\xi_0^{\text{DK}} = 0$. To control the resulting Equations (32), (33), and (34), we first compute

$$\|\eta^{\text{DK}}\|_{\rho} = \|N(\theta + \omega)^\top \Omega(K(\theta + \omega))E(\theta)\|_{\rho} \leq c_{N^\top}c_{\Omega}\|E\|_{\rho}, \quad (49)$$

$$\|\eta^N\|_{\rho} = \|DK(\theta + \omega)^\top \Omega(K(\theta + \omega))E(\theta)\|_{\rho} \leq \sigma_{\text{DK}^\top}c_{\Omega}\|E\|_{\rho}. \quad (50)$$

On the one hand, using Equations (22) and (50), we obtain

$$\|\mathcal{R}(\eta^N)\|_{\rho-\delta} \leq \frac{c_R\sigma_{\text{DK}^\top}c_{\Omega}}{\gamma\delta^{\tau}} \|E\|_{\rho} =: \frac{C_8}{\gamma\delta^{\tau}} \|E\|_{\rho},$$

and on the other hand, using Hypothesis H_4 and Equations (34), (22) and (49), we have

$$\begin{aligned} \|\xi^N\|_{\rho-\delta} &\leq \frac{C_8 + \sigma_T(c_{N^\top}c_{\Omega}\gamma\delta^{\tau} + c_T C_8)}{\gamma\delta^{\tau}} \|E\|_{\rho} =: \frac{C_9}{\gamma\delta^{\tau}} \|E\|_{\rho}, \\ \|\xi^{\text{DK}}\|_{\rho-2\delta} &\leq \frac{c_R(c_{N^\top}c_{\Omega}\gamma\delta^{\tau} + c_T C_9)}{\gamma^2\delta^{2\tau}} \|E\|_{\rho} =: \frac{C_{10}}{\gamma^2\delta^{2\tau}} \|E\|_{\rho}. \end{aligned} \quad (51)$$

The new parameterization $\bar{K} = K + \Delta K$ and the related objects are controlled using standard computations. The first estimate in (30) follows directly from $\Delta K = DK\xi^{\text{DK}} + N\xi^N$ and estimates in (51):

$$\|\bar{K} - K\|_{\rho-2\delta} = \|\Delta K\|_{\rho-2\delta} \leq \frac{\sigma_{\text{DK}}C_{10} + c_N C_9\gamma\delta^{\tau}}{\gamma^2\delta^{2\tau}} \|E\|_{\rho} =: \frac{\hat{C}_2}{\gamma^2\delta^{2\tau}} \|E\|_{\rho}.$$

Combining this expression with Cauchy estimates we obtain the first estimate in (28):

$$\|\text{D}\bar{K}\|_{\rho-3\delta} \leq \|\text{D}K\|_{\rho} + \|\text{D}\Delta K\|_{\rho-3\delta} \leq \|\text{D}K\|_{\rho} + \frac{n\hat{C}_2}{\gamma^2\delta^{2\tau+1}} \|E\|_{\rho} < \sigma_{\text{DK}}. \quad (52)$$

The last inequality in the previous computation is obtained by including this condition in Hypothesis (26). The control of the transposed object in (28) is analogous:

$$\|D\bar{K}^\top\|_{\rho-3\delta} \leq \|DK^\top\|_\rho + \frac{2n\hat{C}_2}{\gamma^2\delta^{2\tau+1}}\|E\|_\rho < \sigma_{DK^\top}. \quad (53)$$

To control \bar{B} and $\langle \bar{T} \rangle^{-1}$ we use that for every pair of matrices X and Y

$$Y^{-1} = (I + X^{-1}(Y - X))^{-1}X^{-1}. \quad (54)$$

If $\|X^{-1}\|\|Y - X\| < 1$, the Neumann series implies

$$\|Y^{-1} - X^{-1}\| \leq \frac{\|X^{-1}\|^2\|Y - X\|}{1 - \|X^{-1}\|\|Y - X\|}. \quad (55)$$

First, we use Equation (54) taking $X = DK^\top\Omega(K)N_0$ and $Y = D\bar{K}^\top\Omega(\bar{K})N_0$. We obtain the second estimate in (30) with

$$\hat{C}_3 := 2\sigma_B^2 C_{11}, \quad C_{11} := c_{N_0}\hat{C}_2(\sigma_{DK^\top}c_{D\Omega}\delta + 2nc_\Omega),$$

where we assumed that (to be included in (26))

$$\frac{2\sigma_B C_{11}}{\gamma^2\delta^{2\tau+1}}\|E\|_\rho < 1. \quad (56)$$

This computation allows us to set that in order to satisfy the first estimate in (29) we have to include

$$\|\bar{B}\|_{\rho-3\delta} \leq \|B\|_{\rho-3\delta} + \|\bar{B} - B\|_{\rho-3\delta} \leq \|B\|_{\rho-3\delta} + \frac{\hat{C}_3}{\gamma^2\delta^{2\tau+1}}\|E\|_\rho < \sigma_B, \quad (57)$$

into Condition (26). The third estimate in (29) follows from

$$\begin{aligned} \|N_0^\top(\Omega \circ \bar{K})N_0\|_{\rho-2\delta} &\leq \|N_0^\top(\Omega \circ K)N_0\|_{\rho-2\delta} + c_{N_0^\top}c_{D\Omega}c_{N_0}\|\bar{K} - K\|_{\rho-2\delta} \\ &< \|N_0^\top(\Omega \circ K)N_0\|_\rho + \frac{c_{N_0^\top}c_{D\Omega}c_{N_0}\hat{C}_2}{\gamma^2\delta^{2\tau}}\|E\|_\rho < \sigma_{N_0^\top(\Omega \circ K)N_0}, \end{aligned}$$

and including this last inequality into Condition (26).

The third expression in (30) also follows using Equation (54) with $X = T$ and $Y = \bar{T}$. Now we have to control the new matrices $\bar{N}(\theta)$ and $\bar{A}(\theta)$, given by Equations (9) and (11) replacing $K(\theta)$ by $\bar{K}(\theta)$. Specifically, we obtain

$$\|\bar{A} - A\|_{\rho-3\delta} \leq \left(\frac{n}{2}(\sigma_B)^2 c_{D\Omega}\hat{C}_2\delta + \frac{n+1}{2}\sigma_{N_0^\top(\Omega \circ K)N_0}\hat{C}_3 \right) \frac{\|E\|_\rho}{\gamma^2\delta^{2\tau+1}} =: \frac{C_{12}}{\gamma^2\delta^{2\tau+1}}\|E\|_\rho,$$

and observe that $\|\bar{A} - A\|_{\rho-3\delta} = \|\bar{A}^\top - A^\top\|_{\rho-3\delta}$. Moreover

$$\|\bar{N} - N\|_{\rho-3\delta} \leq \frac{(\sigma_{DK}C_{12} + n\hat{C}_2c_A + c_{N_0}\hat{C}_3)\|E\|_\rho}{\gamma^2\delta^{2\tau+1}} =: \frac{C_{13}}{\gamma^2\delta^{2\tau+1}}\|E\|_\rho$$

and

$$\|\bar{N}^\top - N^\top\|_{\rho-3\delta} \leq \frac{(\sigma_{DK^\top}C_{12} + 2n\hat{C}_2c_A + nc_{N_0^\top}\hat{C}_3)\|E\|_\rho}{\gamma^2\delta^{2\tau+1}} =: \frac{C_{13}^*}{\gamma^2\delta^{2\tau+1}}\|E\|_\rho,$$

that allow us to compute

$$\|\bar{T} - T\|_{\rho-3\delta} \leq \frac{C_{14}}{\gamma^2\delta^{2\tau+1}}\|E\|_\rho, \quad (58)$$

with

$$C_{14} := c_N^\top c_N \hat{C}_2 (c_\Omega c_{D^2 F} + c_{D\Omega} c_{DF}) \delta + c_\Omega c_{DF} (c_N^\top C_{13} + c_N C_{13}^*).$$

Introducing Equation (58) into Equation (55) we obtain the third estimate in (30) by defining the constant $\hat{C}_4 := 2\sigma_T^2 C_{14}$ and also the second estimate in (29). Computations are analogous to those performed to control the object \bar{B} . Hence we have to include the condition

$$\frac{\hat{C}_4}{\gamma^2 \delta^{2\tau+1}} \|E\|_\rho < \sigma_T - |\langle T \rangle^{-1}| \quad (59)$$

in (26). Note that the closure of $\bar{K}(\mathbb{T}_{\rho-2\delta}^n)$ lies in \mathcal{B} , since

$$\text{dist}(\bar{K}(\mathbb{T}_{\rho-2\delta}^n), \partial\mathcal{B}) \geq \text{dist}(K(\mathbb{T}_\rho^n), \partial\mathcal{B}) - \|\Delta K\|_{\rho-2\delta} \geq \text{dist}(K(\mathbb{T}_\rho^n), \partial\mathcal{B}) - \frac{\hat{C}_2}{\gamma^2 \delta^{2\tau}} \|E\|_\rho > 0. \quad (60)$$

The last inequality is also included in (26).

Hence, the terms $E_{\text{red}}(\theta)\xi(\theta)$ and $\Omega_0 E_{\text{sym}}(\theta + \omega)\xi(\theta + \omega)$ in Equation (47) are quadratic in $E(\theta)$. Then, using $\Delta K(\theta) = P(\theta)\xi(\theta)$, Equation (47), the definition of $\xi(\theta)$, and also that

$$(-\Omega_0 P(\theta + \omega)^\top \Omega(K(\theta + \omega)))^{-1} = P(\theta + \omega)(I - \Omega_0 E_{\text{sym}}(\theta + \omega))^{-1},$$

it turns out that

$$\begin{aligned} DF(K(\theta))\Delta K(\theta) - \Delta K(\theta + \omega) + E(\theta) = \\ P(\theta + \omega)(I - \Omega_0 E_{\text{sym}}(\theta + \omega))^{-1} E_{\text{lin}}(\theta), \end{aligned} \quad (61)$$

where

$$E_{\text{lin}}(\theta) = E_{\text{red}}(\theta)\xi(\theta) + \Omega_0 E_{\text{sym}}(\theta + \omega)\xi(\theta + \omega) - \left(\left\langle L(\theta + \omega)^\top \Omega(K(\theta + \omega)) E(\theta) \right\rangle \right). \quad (62)$$

After performing one step of the Newton method, the error of invariance associated to the parameterization $\bar{K} = K + \Delta K$ is given by

$$\begin{aligned} \bar{E}(\theta) &= F(K(\theta) + \Delta K(\theta)) - K(\theta) - \Delta K(\theta + \omega) \\ &= P(\theta + \omega)(I - \Omega_0 E_{\text{sym}}(\theta + \omega))^{-1} E_{\text{lin}}(\theta) + \Delta^2 F(\theta), \end{aligned} \quad (63)$$

where we used Equation (61), and

$$\begin{aligned} \Delta^2 F(\theta) &= F(K(\theta) + \Delta K(\theta)) - F(K(\theta)) - DF(K(\theta))\Delta K(\theta) \\ &= \int_0^1 (1-t) D^2 F(K(\theta) + t\Delta K(\theta)) (\Delta K(\theta), \Delta K(\theta)) dt. \end{aligned}$$

The last step of the proof is to see, using the previously computed expressions, that the new error $\bar{E}(\theta)$ is quadratic in $E(\theta)$.

We use Lemma 2.11 to control the modulus of the average:

$$\left| \left\langle L(\theta + \omega)^\top \Omega(K(\theta + \omega)) E(\theta) \right\rangle \right| \leq \left(\frac{2nc_{D^2 a}}{\delta} + \frac{c_{D^2 a}}{2} \right) \|E\|_\rho^2$$

and from the expression of $E_{\text{lin}}(\theta)$ in Equation (62) we obtain

$$\|E_{\text{lin}}\|_{\rho-2\delta} \leq \left(\frac{(C_3 + C_7) \max\{C_9 \gamma \delta^\tau, C_{10}\}}{\gamma^3 \delta^{3\tau+1}} + \frac{2nc_{D^2 a}}{\delta} + \frac{c_{D^2 a}}{2} \right) \|E\|_\rho^2 =: \frac{C_{15}}{\gamma^3 \delta^{3\tau+1}} \|E\|_\rho^2.$$

Using a Neumann series argument we obtain

$$\|(I - \Omega_0 E_{\text{sym}})^{-1}\| \leq \frac{1}{1 - \|\Omega_0 E_{\text{sym}}\|}. \quad (64)$$

Let us consider, as a hypothesis that we include in (26), that

$$\frac{2C_3}{\gamma\delta^{\tau+1}} \|E\|_\rho =: \frac{\hat{C}_1}{\gamma\delta^{\tau+1}} \|E\|_\rho < 1. \quad (65)$$

Using Equations (41), (64) and (65), we obtain $\|(I - \Omega_0 E_{\text{sym}})^{-1}\| < 2$. Then, the new error of invariance, given by Equation (63), satisfies Condition (31):

$$\|\bar{E}\|_{\rho-2\delta} < \left(2c_P C_{15} \gamma \delta^{\tau-1} + \frac{1}{2} c_{D^2F} \hat{C}_2^2\right) \frac{\|E\|_\rho^2}{\gamma^4 \delta^{4\tau}} =: \frac{\hat{C}_5 \|E\|_\rho^2}{\gamma^4 \delta^{4\tau}}. \quad (66)$$

We complete the proof by merging Equations (52), (53), (57), (59), (60) and (65), thus obtaining the expression in (27) that appears in the statement. \square

Proof of Theorem 2.6. Let us consider the approximate F -invariant 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 2.9 recursively we obtain new objects $K_s = K_{s-1}$, $E_s = E_{s-1}$, $B_s = B_{s-1}$, and $T_s = T_{s-1}$. The domain of analyticity of these objects is reduced at every step. To characterize this fact, we introduce parameters $a_1 > 1$, $a_2 > 1$, $a_3 = 3 \frac{a_1 - a_2}{a_1 - 1 a_2 - 1}$ 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 to optimize the convergence of the KAM process for a particular problem. This has been used for example in [23]. Due to the quadratic convergence of the scheme, a good strategy is to optimize the first numbers $\delta_0, \delta_1, \dots, \delta_m$.

We denote the objects at the s -step as K_s, E_s, B_s and T_s , respectively. We observe that Condition (26) is required at every step but the construction has been performed in such a way that we can control $\|DK_s\|_{\rho_s}, \|DK_s^\top\|_{\rho_s}, \|B_s\|_{\rho_s}, \text{dist}(K_s(\mathbb{T}_{\rho_s}^m), \partial\mathcal{B})$, and $|\langle T_s \rangle^{-1}|$ uniformly with respect to s , so the constants that appear in Lemma 2.9 are taken to be the same for all steps by considering the worst value of δ_s , that is, $\delta_0 = \rho_0/a_3$.

Now we proceed by induction. We suppose that we have applied s times Lemma 2.9, for certain $s \geq 0$, so we have to verify that we can apply it again. To this end, we first compute the error E_s in terms of E_0 as follows

$$\|E_s\|_{\rho_s} < \frac{\hat{C}_5}{\gamma^4 \delta_{s-1}^{4\tau}} \|E_{s-1}\|_{\rho_{s-1}}^2 = \frac{\hat{C}_5 a_1^{4\tau(s-1)}}{\gamma^4 \delta_0^{4\tau}} \|E_{s-1}\|_{\rho_{s-1}}^2$$

and iterating this sequence backwards (we use that $1 + 2 + \dots + 2^{s-1} = 2^s - 1$ and $1(s-1) + 2(s-2) + 2^2(s-3) \dots + 2^{s-2} = 2^s - s - 1$) we obtain

$$\|E_s\|_{\rho_s} < \left(\frac{a_1^{4\tau} \hat{C}_5 \|E_0\|_{\rho_0}}{\gamma^4 \delta_0^{4\tau}} \right)^{2^s - 1} a_1^{-4\tau s} \|E_0\|_{\rho_0}. \quad (67)$$

We use this expression in order to verify Condition (26) so we can perform the step $s+1$. Before that, in order to produce a decreasing sequence of errors, we assume that

$$\frac{a_1^{4\tau} \hat{C}_5 \|E_0\|_{\rho_0}}{\gamma^4 \delta_0^{4\tau}} < 1 \quad (68)$$

thus including this condition in Hypothesis (23). Now, to verify the inequality in (26) we observe that in the expression for $\hat{\mathfrak{C}}_1$, given by Equation (27), there are two types of conditions. On the one hand, we have conditions like (65), depending only on the error E_s and δ_s . On the other hand, we have conditions like (52) depending also on other objects at the s -step. For example, Condition (65) is direct using Equation (67) and $\tau \geq n$

$$\frac{2C_3\|E_s\|_{\rho_s}}{\gamma\delta_s^{\tau+1}} < \frac{2C_3a_1^{(\tau+1)s}}{\gamma\delta_0^{\tau+1}} \left(\frac{a_1^{4\tau}\hat{C}_5\|E_0\|_{\rho_0}}{\gamma^4\delta_0^{4\tau}} \right)^{2^s-1} a_1^{-4\tau s}\|E_0\|_{\rho_0} < \frac{2C_3}{\gamma\delta_0^{\tau+1}}\|E_0\|_{\rho_0} < 1,$$

where the last inequality is included in (23). If the condition depends also on other objects, we have to relate it to the initial one. For example, Condition (52) requires to compute, using Equation (68), the following

$$\begin{aligned} \|DK_s\|_{\rho_s} + \frac{n\hat{C}_2\|E_s\|_{\rho_s}}{\gamma^2\delta_s^{2\tau+1}} &< \|DK_0\|_{\rho_0} + \sum_{j=0}^s \frac{n\hat{C}_2\|E_j\|_{\rho_j}}{\gamma^2\delta_j^{2\tau+1}} \\ &< \|DK_0\|_{\rho_0} + \sum_{j=0}^{\infty} \frac{n\hat{C}_2a_1^{(2\tau+1)j}}{\gamma^2\delta_0^{2\tau+1}} \left(\frac{a_1^{4\tau}\hat{C}_5\|E_0\|_{\rho_0}}{\gamma^4\delta_0^{4\tau}} \right)^{2^j-1} a_1^{-4\tau j}\|E_0\|_{\rho_0} \\ &< \|DK_0\|_{\rho_0} + \frac{n\hat{C}_2}{\gamma^2\delta_0^{2\tau+1}} \left(\frac{1}{1-a_1^{1-2\tau}} \right) \|E_0\|_{\rho_0} < \sigma_{DK}. \end{aligned}$$

As usual, the last inequality is included in (23). Then, we perform analogous computations to guarantee the conditions in (26) and we obtain the sufficient condition

$$\frac{\hat{C}_8\|E_0\|_{\rho}}{\gamma^2\delta_0^{2\tau+1}} < 1, \quad (69)$$

where \hat{C}_8 is given by

$$\hat{C}_8 := \max \left\{ 2C_3\gamma\delta_0^{\tau}, \frac{\hat{C}_6}{1-a_1^{1-2\tau}}, \frac{\hat{C}_7}{1-a_1^{-2\tau}} \right\} \quad (70)$$

with

$$\hat{C}_6 := \max \left\{ \frac{n\hat{C}_2}{\sigma_{DK} - \|DK_0\|_{\rho_0}}, \frac{2n\hat{C}_2}{\sigma_{DK^{\top}} - \|DK_0^{\top}\|_{\rho_0}}, \frac{c_{N_0^{\top}}c_{D\Omega}c_{N_0}\delta\hat{C}_2}{\sigma_{N_0^{\top}(\Omega \circ K)N_0} - \|N_0^{\top}(\Omega \circ K)N_0\|_{\rho-2\delta}}, \frac{\hat{C}_3}{\sigma_B - \|B_0\|_{\rho_0}}, \frac{\hat{C}_4}{\sigma_T - |\langle T_0 \rangle^{-1}|} \right\} \quad (71)$$

and

$$\hat{C}_7 := \frac{\hat{C}_2\delta_0}{\text{dist}(K_0(\mathbb{T}_{\rho_0}^n), \partial\mathcal{B})}.$$

Since Hypotheses H_1 to H_4 and Condition (26) are satisfied, we can apply Lemma 2.9 again. Note that the sequence of errors satisfies $\|E_s\|_{\rho_s} \rightarrow 0$ when $s \rightarrow \infty$, so the iterative scheme converges to a true quasi-periodic torus K_{∞} . Condition (23) of the smallness of $\|E_0\|_{\rho_0}$ is obtained by merging Conditions (68) and (69). Indeed, we have

$$\mathfrak{C}_1 := \max \left\{ (a_1a_3)^{4\tau}\hat{C}_5, (a_3)^{2\tau+1}\hat{C}_8\gamma^2\rho_0^{2\tau-1} \right\}, \quad (72)$$

where \hat{C}_5 is given in (31), \hat{C}_8 is given in (70) and we used that $\delta_0 = \rho_0/a_3$. Finally, we obtain the constant

$$\mathfrak{C}_2 := a_3^{2\tau}\hat{C}_2/(1-a_1^{1-2\tau}) \quad (73)$$

that appears in (24), controlling that the torus is close to the initial approximation. \square

3 On the approximation of periodic functions using discrete Fourier transform

In the core of the computer assisted methodology presented in this work, we have to bound the error produced when approximating a periodic function by its discrete Fourier transform. This is a very natural problem that has been considered in the approximation theory literature [66]. It is well known that error estimates improve commensurately as the functions become smoother [27]. We refer the reader to [57, 71] for problems where similar ideas have been used. Motivated by the setting of the present paper, we address the problem for analytic functions. The estimates presented in this section improve the ones given in [27] for this specific case (see Section 3.3).

3.1 Notation regarding discretization of the torus and Fourier transforms

Given a function $f : \mathbb{T}^n \rightarrow \mathbb{C}$, we consider its Fourier series

$$f(\theta) = \sum_{k \in \mathbb{Z}^n} f_k e^{2\pi i k \cdot \theta},$$

where the Fourier coefficients are given by the *Fourier transform* (FT)

$$f_k = \int_{[0,1]^n} f(\theta) e^{-2\pi i k \cdot \theta} d\theta. \quad (74)$$

We consider a sample of points on the regular grid of size $N_F = (N_{F,1}, \dots, N_{F,n}) \in \mathbb{N}^n$

$$\theta_j := (\theta_{j_1}, \dots, \theta_{j_n}) = \left(\frac{j_1}{N_{F,1}}, \dots, \frac{j_n}{N_{F,n}} \right), \quad (75)$$

where $j = (j_1, \dots, j_n)$, with $0 \leq j_\ell < N_{F,\ell}$ and $1 \leq \ell \leq n$. This defines an n -dimensional sampling $\{f_j\}$, with $f_j = f(\theta_j)$. The total number of points is $N_D = N_{F,1} \cdots N_{F,n}$. The integrals in Equation (74) are approximated using the trapezoidal rule on the regular grid, obtaining the discrete Fourier transform (DFT)

$$\tilde{f}_k = \frac{1}{N_D} \sum_{0 \leq j < N_F} f_j e^{-2\pi i k \cdot \theta_j},$$

where the sum runs over integer subindices $j \in \mathbb{Z}^n$ such that $0 \leq j_\ell < N_{F,\ell}$ for $\ell = 1, \dots, n$. Notice that \tilde{f}_k is periodic with respect to the components k_1, \dots, k_n of k , with periods $N_{F,1}, \dots, N_{F,n}$, respectively. The periodic function f is approximated by the discrete Fourier approximation

$$\tilde{f}(\theta) = \sum_{k \in \mathcal{I}_{N_F}} \tilde{f}_k e^{2\pi i k \cdot \theta}, \quad (76)$$

where \mathcal{I}_{N_F} is the finite set of multi-indices given by

$$\mathcal{I}_{N_F} = \left\{ k \in \mathbb{Z}^n \mid -\frac{N_{F,\ell}}{2} \leq k_\ell < \frac{N_{F,\ell}}{2}, 1 \leq \ell \leq n \right\}. \quad (77)$$

Along this section we will use the standard notation $[x]$ for the integer part of x : $[x] = \min \{j \in \mathbb{Z} : x \leq j\}$.

3.2 Error estimates on the approximation of analytic periodic functions

Motivated by the setting of the present paper, we will work in spaces of analytic functions on a complex strip of the torus (see Section 2.2), but most of the arguments can be adapted to other spaces. The main goal is to control the error between \tilde{f} and f , using suitable norms. As a previous step, we establish estimates of the approximation \tilde{f}_k of f_k .

Lemma 3.1. *The coefficients of the DFT are obtained from the coefficients of the FT by*

$$\tilde{f}_k = \sum_{m \in \mathbb{Z}^n} f_{k+N_{\mathbb{F}}(m)},$$

where $N_{\mathbb{F}}(m) = (N_{\mathbb{F},1}m_1, \dots, N_{\mathbb{F},n}m_n)$.

The proof of Lemma 3.1 is direct. Using this result we obtain a bound for the difference between \tilde{f}_k and f_k as follows:

Proposition 3.2. *Let $f : \mathbb{T}_{\hat{\rho}}^n \rightarrow \mathbb{C}$ be an analytic and bounded function in the complex strip $\mathbb{T}_{\hat{\rho}}^n$ of size $\hat{\rho} > 0$. Let \tilde{f} be the discrete Fourier approximation of f in the regular grid of size $N_{\mathbb{F}} = (N_{\mathbb{F},1}, \dots, N_{\mathbb{F},n}) \in \mathbb{N}^n$. Then, for $-\frac{N_{\mathbb{F}}}{2} \leq k < \frac{N_{\mathbb{F}}}{2}$:*

$$|\tilde{f}_k - f_k| \leq s_{N_{\mathbb{F}}}^*(k, \hat{\rho}) \|f\|_{\hat{\rho}}$$

where

$$s_{N_{\mathbb{F}}}^*(k, \hat{\rho}) = \prod_{\ell=1}^n \left(e^{-\pi \hat{\rho} N_{\mathbb{F},\ell}} \frac{e^{2\pi \hat{\rho} (|k_{\ell}| - N_{\mathbb{F},\ell}/2)} + e^{-2\pi \hat{\rho} (|k_{\ell}| - N_{\mathbb{F},\ell}/2)}}{1 - e^{-2\pi \hat{\rho} N_{\mathbb{F},\ell}}} \right) - e^{-2\pi \hat{\rho} |k|_1}.$$

Proof: Let $k \in \mathbb{Z}^n$ be a multi-index. From Lemma 3.1 and standard bounds of the Fourier coefficients of analytic functions, we obtain

$$|\tilde{f}_k - f_k| \leq \sum_{m \in \mathbb{Z}^n \setminus \{0\}} |f_{k+N_{\mathbb{F}}(m)}| \leq \sum_{m \in \mathbb{Z}^n \setminus \{0\}} e^{-2\pi \hat{\rho} |k+N_{\mathbb{F}}(m)|_1} \|f\|_{\hat{\rho}}.$$

Then we define

$$s_{N_{\mathbb{F}}}(k, \hat{\rho}) = \sum_{m \in \mathbb{Z}^n} e^{-2\pi \hat{\rho} |k+N_{\mathbb{F}}(m)|_1}, \quad s_{N_{\mathbb{F}}}^*(k, \hat{\rho}) = \sum_{m \in \mathbb{Z}^n \setminus \{0\}} e^{-2\pi \hat{\rho} |k+N_{\mathbb{F}}(m)|_1}.$$

Notice that $s_{N_{\mathbb{F}}}(k, \hat{\rho}) = s_{N_{\mathbb{F}}}(k', \hat{\rho})$ for every $k' \in \mathbb{Z}^n$ such that $|k_i| = |k'_i|$ for all $i = 1, \dots, n$. Then, we write

$$s_{N_{\mathbb{F}}}(k, \hat{\rho}) = \prod_{\ell=1}^n s_{N_{\mathbb{F},\ell}}(k_{\ell}, \hat{\rho}),$$

where

$$s_{N_{\mathbb{F},\ell}}(k_{\ell}, \hat{\rho}) = \sum_{m \in \mathbb{Z}} e^{-2\pi \hat{\rho} |k_{\ell} + N_{\mathbb{F},\ell} m|}.$$

Then, by defining $r_{\ell} \equiv k_{\ell} \pmod{N_{\mathbb{F},\ell}}$ for $\ell = 1, \dots, n$, we obtain

$$\begin{aligned} s_{N_{\mathbb{F},\ell}}(k_{\ell}, \hat{\rho}) &= s_{N_{\mathbb{F},\ell}}(r_{\ell}, \hat{\rho}) = \sum_{m_{\ell} \geq 0} e^{-2\pi \hat{\rho} (r_{\ell} + N_{\mathbb{F},\ell} m_{\ell})} + \sum_{m_{\ell} < 0} e^{-2\pi \hat{\rho} (-r_{\ell} - N_{\mathbb{F},\ell} m_{\ell})} \\ &= \frac{e^{2\pi \hat{\rho} (r_{\ell} - N_{\mathbb{F},\ell})} + e^{-2\pi \hat{\rho} r_{\ell}}}{1 - e^{-2\pi \hat{\rho} N_{\mathbb{F},\ell}}}. \end{aligned}$$

The result follows directly from $s_{N_{\mathbb{F}}}^*(k, \hat{\rho}) = s_{N_{\mathbb{F}}}(k, \hat{\rho}) - e^{-2\pi \hat{\rho} |k|_1}$. \square

Next, we state the main result of this section, that allows us to control the error between \tilde{f} and f .

Theorem 3.3. *Let $f : \mathbb{T}_{\hat{\rho}}^n \rightarrow \mathbb{C}$ be an analytic and bounded function in the complex strip $\mathbb{T}_{\hat{\rho}}^n$ of size $\hat{\rho} > 0$. Let \tilde{f} be the discrete Fourier approximation of f in the regular grid of size $N_{\mathbb{F}} = (N_{\mathbb{F},1}, \dots, N_{\mathbb{F},n}) \in \mathbb{N}^n$. Then*

$$\|\tilde{f} - f\|_{\rho} \leq C_{N_{\mathbb{F}}}(\rho, \hat{\rho}) \|f\|_{\hat{\rho}},$$

for $0 \leq \rho < \hat{\rho}$, where $C_{N_F}(\rho, \hat{\rho}) = S_{N_F}^{*1}(\rho, \hat{\rho}) + S_{N_F}^{*2}(\rho, \hat{\rho}) + T_{N_F}(\rho, \hat{\rho})$ is given by

$$S_{N_F}^{*1}(\rho, \hat{\rho}) = \prod_{\ell=1}^n \frac{1}{1 - e^{-2\pi\hat{\rho}N_{F,\ell}}} \sum_{\substack{\sigma \in \{-1, 1\}^n \\ \sigma \neq (1, \dots, 1)}} \prod_{\ell=1}^n e^{(\sigma_\ell - 1)\pi\hat{\rho}N_{F,\ell}} \nu_\ell(\sigma_\ell \hat{\rho} - \rho),$$

$$S_{N_F}^{*2}(\rho, \hat{\rho}) = \prod_{\ell=1}^n \frac{1}{1 - e^{-2\pi\hat{\rho}N_{F,\ell}}} \left(1 - \prod_{\ell=1}^n (1 - e^{-2\pi\hat{\rho}N_{F,\ell}}) \right) \prod_{\ell=1}^n \nu_\ell(\hat{\rho} - \rho)$$

and

$$T_{N_F}(\rho, \hat{\rho}) = \left(\frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} \right)^n \left(1 - \prod_{\ell=1}^n (1 - \mu_\ell(\hat{\rho} - \rho) e^{-\pi(\hat{\rho}-\rho)N_{F,\ell}}) \right),$$

with

$$\nu_\ell(\delta) = \frac{e^{2\pi\delta} + 1}{e^{2\pi\delta} - 1} \left(1 - \mu_\ell(\delta) e^{-\pi\delta N_{F,\ell}} \right) \quad \text{and} \quad \mu_\ell(\delta) = \begin{cases} 1 & \text{if } N_{F,\ell} \text{ is even} \\ \frac{2e^{\pi\delta}}{e^{2\pi\delta} + 1} & \text{if } N_{F,\ell} \text{ is odd} \end{cases}.$$

Proof: From the definition of the discrete Fourier approximation \tilde{f} of f , we have

$$\|\tilde{f} - f\|_\rho \leq \sum_{k \in \mathcal{I}_{N_F}} |\tilde{f}_k - f_k| e^{2\pi\rho|k|_1} + \sum_{k \notin \mathcal{I}_{N_F}} |f_k| e^{2\pi\rho|k|_1},$$

where \mathcal{I}_{N_F} is the finite set of multi-indices given by Equation (77). From Proposition 3.2 and the growth rate properties of the Fourier coefficients of an analytic function, we get

$$\|\tilde{f} - f\|_\rho \leq (S_{N_F}^*(\rho, \hat{\rho}) + T_{N_F}(\rho, \hat{\rho})) \|f\|_{\hat{\rho}},$$

where

$$S_{N_F}^*(\rho, \hat{\rho}) = \sum_{k \in \mathcal{I}_{N_F}} s_N^*(k, \hat{\rho}) e^{2\pi\rho|k|_1},$$

and

$$T_{N_F}(\rho, \hat{\rho}) = \sum_{k \notin \mathcal{I}_{N_F}} e^{2\pi(\rho-\hat{\rho})|k|_1}.$$

Next, we obtain a computable expression for $T_{N_F}(\rho, \hat{\rho})$. Notice that

$$T_{N_F}(\rho, \hat{\rho}) = \sum_{k \in \mathbb{Z}^n} e^{2\pi(\rho-\hat{\rho})|k|_1} - \sum_{k \in \mathcal{I}_{N_F}} e^{2\pi(\rho-\hat{\rho})|k|_1} = \left(\frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} \right)^n - \prod_{\ell=1}^n \nu_\ell(\hat{\rho} - \rho),$$

where

$$\nu_\ell(\delta) = \sum_{k_\ell = -\lceil \frac{N_{F,\ell}}{2} \rceil}^{\lceil \frac{N_{F,\ell}-1}{2} \rceil} e^{-2\pi\delta|k_\ell|}.$$

Then, the formula stated in the proposition follows by distinguishing the cases where $N_{F,\ell}$ is odd and even.

To obtain a suitable expression for $S_{N_F}^*(\rho, \hat{\rho})$, we compute

$$\begin{aligned}
S_{N_F}(\rho, \hat{\rho}) &= \sum_{k \in \mathcal{I}_{N_F}} s_{N_F}(k, \hat{\rho}) e^{2\pi\rho|k|_1} \\
&= \sum_{k \in \mathcal{I}_{N_F}} \prod_{\ell=1}^n \left(e^{-\pi\hat{\rho}N_{F,\ell}} \frac{e^{2\pi\hat{\rho}(|k_\ell| - N_{F,\ell}/2)} + e^{-2\pi\hat{\rho}(|k_\ell| - N_{F,\ell}/2)}}{1 - e^{-2\pi\hat{\rho}N_{F,\ell}}} e^{2\pi\rho|k_\ell|} \right) \\
&= \prod_{\ell=1}^n \frac{e^{-\pi\hat{\rho}N_{F,\ell}}}{1 - e^{-2\pi\hat{\rho}N_{F,\ell}}} \sum_{\sigma \in \{-1,1\}^n} \sum_{k \in \mathcal{I}_{N_F}} \prod_{\ell=1}^n e^{-2\pi(\sigma_\ell\hat{\rho}-\rho)|k_\ell|} e^{\pi\sigma_\ell\hat{\rho}N_{F,\ell}} \\
&= \prod_{\ell=1}^n \frac{e^{-\pi\hat{\rho}N_{F,\ell}}}{1 - e^{-2\pi\hat{\rho}N_{F,\ell}}} \sum_{\sigma \in \{-1,1\}^n} \prod_{\ell=1}^n \sum_{-\frac{N_{F,\ell}}{2} \leq k_\ell < \frac{N_{F,\ell}}{2}} e^{-2\pi(\sigma_\ell\hat{\rho}-\rho)|k_\ell|} e^{\pi\sigma_\ell\hat{\rho}N_{F,\ell}} \\
&= \prod_{\ell=1}^n \frac{e^{-\pi\hat{\rho}N_{F,\ell}}}{1 - e^{-2\pi\hat{\rho}N_{F,\ell}}} \sum_{\sigma \in \{-1,1\}^n} \prod_{\ell=1}^n e^{\pi\sigma_\ell\hat{\rho}N_{F,\ell}} \nu_\ell(\sigma_\ell\hat{\rho} - \rho).
\end{aligned}$$

Finally, we use that

$$S_{N_F}^*(\rho, \hat{\rho}) = S_{N_F}(\rho, \hat{\rho}) - \prod_{\ell=1}^n \nu_\ell(\hat{\rho} - \rho),$$

and we decompose the resulting expression in the two functions $S_{N_F}^{*1}$ and $S_{N_F}^{*2}$. \square

Remark 3.4. In the above formulae, there are expressions of the form $1 - \prod_{\ell=1}^n (1 - x_\ell)$, where $0 < x_\ell < 1$ for $\ell = 1 \dots n$. In our applications, it turns out that $0 < x_\ell \ll 1$, so we have to be aware of the propagation of the error when enclosing this expression using interval arithmetics. To this end, we will use the formulae

$$1 - \prod_{\ell=1}^n (1 - x_\ell) = \sum_{j=1}^n (-1)^{j-1} \sum_{\substack{\ell_1 < \dots < \ell_j \\ 1 \leq \ell_i \leq n}} x_{\ell_1} \dots x_{\ell_j}.$$

Notice that the dominant term of the expression is $\sum_{\ell=1}^n x_\ell$.

Remark 3.5. It is interesting to characterize the dominant terms in the expression $C_{N_F}(\rho, \hat{\rho}) = S_{N_F}^{*1}(\rho, \hat{\rho}) + S_{N_F}^{*2}(\rho, \hat{\rho}) + T_{N_F}(\rho, \hat{\rho})$. The dominant term of $S_{N_F}^{*1}(\rho, \hat{\rho})$ corresponds to the multi-indices $\sigma \in \{-1, 1\}^n$ for which only one component is -1 . Hence, we have

$$S_{N_F}^{*1}(\rho, \hat{\rho}) \simeq \left(\frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} \right)^{n-1} \left(\frac{e^{2\pi(\hat{\rho}+\rho)} + 1}{e^{2\pi(\hat{\rho}+\rho)} - 1} \right) \sum_{\ell=1}^n \mu_\ell(\hat{\rho} - \rho) e^{-\pi(\hat{\rho}-\rho)N_{F,\ell}}.$$

Then, we observe that the dominant term of $S_{N_F}^{*2}(\rho, \hat{\rho})$,

$$S_{N_F}^{*2}(\rho, \hat{\rho}) \simeq \left(\frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} \right)^n \sum_{\ell=1}^n e^{-2\pi\hat{\rho}N_{F,\ell}},$$

is much smaller than $S_{N_F}^{*1}(\rho, \hat{\rho})$. Finally,

$$T_{N_F}(\rho, \hat{\rho}) \simeq \left(\frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} \right)^n \sum_{\ell=1}^n \mu_\ell(\hat{\rho} - \rho) e^{-\pi(\hat{\rho}-\rho)N_{F,\ell}},$$

which is of the same order as $S_{N_F}^{*1}(\rho, \hat{\rho})$. Putting this together we obtain

$$C_{N_F}(\rho, \hat{\rho}) \simeq \left(\frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} \right)^{n-1} \left(\frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} + \frac{e^{2\pi(\hat{\rho}+\rho)} + 1}{e^{2\pi(\hat{\rho}+\rho)} - 1} \right) \sum_{\ell=1}^n \mu_\ell(\hat{\rho} - \rho) e^{-\pi(\hat{\rho}-\rho)N_{F,\ell}}$$

and $\mu_\ell(\hat{\rho} - \rho) \leq 1$ implies

$$C_{N_F}(\rho, \hat{\rho}) \simeq O(e^{-\pi(\hat{\rho}-\rho)\min_{\ell}\{N_{F,\ell}\}}).$$

3.3 Comments on the 1-dimensional case

The simplest case $n = 1$ deserves especial attention, as it is a common situation in the literature. Let us formulate Proposition 3.2 and Theorem 3.3 in this case.

Corollary 3.6. *Let $f : \mathbb{T}_{\hat{\rho}} \rightarrow \mathbb{C}$ be an analytic and bounded function in the complex strip $\mathbb{T}_{\hat{\rho}}$ of size $\hat{\rho} > 0$. Let \tilde{f} be the discrete Fourier approximation of f in the regular grid of size $N_F \in \mathbb{N}$. Then, for $k = -\left\lfloor \frac{N_F}{2} \right\rfloor, \dots, \left\lfloor \frac{N_F-1}{2} \right\rfloor$,*

$$|\tilde{f}_k - f_k| \leq s_{N_F}^*(k, \hat{\rho}) \|f\|_{\hat{\rho}}$$

where

$$s_{N_F}^*(k, \hat{\rho}) = \frac{e^{-2\pi\hat{\rho}N_F}}{1 - e^{-2\pi\hat{\rho}N_F}} \left(e^{2\pi\hat{\rho}k} + e^{-2\pi\hat{\rho}k} \right).$$

Corollary 3.7. *Let $f : \mathbb{T}_{\hat{\rho}} \rightarrow \mathbb{C}$ be an analytic and bounded function in the complex strip $\mathbb{T}_{\hat{\rho}}$ of size $\hat{\rho} > 0$. Let \tilde{f} be the discrete Fourier approximation of f in the regular grid of size N_F . Then, for $0 \leq \rho < \hat{\rho}$, we have*

$$\|\tilde{f} - f\|_{\rho} \leq C_{N_F}(\rho, \hat{\rho}) \|f\|_{\hat{\rho}},$$

where $C_{N_F}(\rho, \hat{\rho}) = S_{N_F}^{*1}(\rho, \hat{\rho}) + S_{N_F}^{*2}(\rho, \hat{\rho}) + T_{N_F}(\rho, \hat{\rho})$, with

$$S_{N_F}^{*1}(\rho, \hat{\rho}) = \frac{e^{-2\pi\hat{\rho}N_F}}{1 - e^{-2\pi\hat{\rho}N_F}} \frac{e^{-2\pi(\hat{\rho}+\rho)} + 1}{e^{-2\pi(\hat{\rho}+\rho)} - 1} \left(1 - \mu_1(-\hat{\rho} - \rho) e^{\pi(\hat{\rho}+\rho)N_F} \right),$$

$$S_{N_F}^{*2}(\rho, \hat{\rho}) = \frac{e^{-2\pi\hat{\rho}N_F}}{1 - e^{-2\pi\hat{\rho}N_F}} \frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} \left(1 - \mu_1(\hat{\rho} - \rho) e^{-\pi(\hat{\rho}-\rho)N_F} \right)$$

and

$$T_{N_F}(\rho, \hat{\rho}) = \frac{e^{2\pi(\hat{\rho}-\rho)} + 1}{e^{2\pi(\hat{\rho}-\rho)} - 1} \mu_1(\hat{\rho} - \rho) e^{-\pi(\hat{\rho}-\rho)N_F}.$$

In order to compare with [27] we consider the odd case, for $N_F = 2M + 1$. In this reference, the following uniform bound was obtained

$$|\tilde{f}_k - f_k| \leq \tilde{s}_{N_F}^*(\hat{\rho}) \|f\|_{\hat{\rho}},$$

for $k = -M, \dots, M$, where

$$\tilde{s}_{N_F}^*(\hat{\rho}) = \frac{4e^{-2\pi\hat{\rho}M}}{e^{2\pi\hat{\rho}} - 1}.$$

Let us now compare $s_{N_F}^*(k, \hat{\rho})$ with $\tilde{s}_{N_F}^*(\hat{\rho})$, for $k = -M, \dots, M$:

$$\frac{s_{N_F}^*(k, \hat{\rho})}{\tilde{s}_{N_F}^*(\hat{\rho})} \leq \frac{s_{N_F}^*(M, \hat{\rho})}{\tilde{s}_{N_F}^*(\hat{\rho})} = \frac{1}{4} \left(1 - e^{-2\pi\hat{\rho}} \right) \frac{1 + e^{-2\pi\hat{\rho}(N_F-1)}}{1 - e^{-2\pi\hat{\rho}N_F}} \leq \frac{1}{4}.$$

Notice that the estimates produced in Corollary 3.6 are at least four times better than the estimates produced in [27].

Now we consider Corollary 3.7 in the case $\rho = 0$. First, we observe that $\mu_1(-\delta) = \mu_1(\delta)$, so we get

$$\|\tilde{f} - f\|_0 \leq C_{N_F}(0, \hat{\rho}) \|f\|_{\hat{\rho}},$$

with

$$C_{N_F}(0, \hat{\rho}) = 2 \frac{e^{2\pi\hat{\rho}} + 1}{e^{2\pi\hat{\rho}} - 1} \mu_1(\hat{\rho}) e^{-\pi\hat{\rho}N_F}.$$

In the odd case, with $N_F = 2M + 1$, this expression reads as follows:

$$C_{N_F}(0, \hat{\rho}) = \frac{4 e^{-2\pi\hat{\rho}M}}{(e^{2\pi\hat{\rho}} - 1)(1 - e^{-2\pi\hat{\rho}N_F})}.$$

It is worth pointing out that the best uniform approximation p^* of the form $p(\theta) = \sum_{-M \leq k \leq M} \tilde{p}_k e^{2\pi i k \cdot \theta}$ satisfies (c.f. [27])

$$\|p^* - f\|_0 \leq \frac{2e^{-2\pi\hat{\rho}M}}{(e^{2\pi\hat{\rho}} - 1)} \|f\|_{\hat{\rho}}.$$

Hence, the discrete Fourier approximation is very close to optimal, since the corresponding error (approximately) doubles the less possible error.

3.4 Matrices of periodic functions

In this section we consider some extensions of Theorem 3.3 to deal with matrix functions $A : \mathbb{T}^n \rightarrow \mathbb{C}^{m_1 \times m_2}$. Our goal is to control the propagation of the error when we perform matrix operations. Specifically, we are interested in the study of products and inverses, but the ideas given below can be adapted to control other operations if necessary.

The first result is obtained directly from Theorem 3.3:

Corollary 3.8. *Let us consider two matrix functions $A : \mathbb{T}^n \rightarrow \mathbb{C}^{m_1 \times m_2}$, and $B : \mathbb{T}^n \rightarrow \mathbb{C}^{m_2 \times m_3}$, such that their entries are analytic and bounded functions in the complex strip $\mathbb{T}_{\hat{\rho}}^n$ of size $\hat{\rho} > 0$. We denote by AB the product matrix and \widetilde{AB} the corresponding approximation given by DFT. Given a grid of size $N_F = (N_{F,1}, \dots, N_{F,n})$, we evaluate A and B in the grid, and we interpolate the points $AB(\theta_j) = A(\theta_j)B(\theta_j)$. Then, we have*

$$\|AB - \widetilde{AB}\|_{\rho} \leq C_{N_F}(\rho, \hat{\rho}) \|A\|_{\hat{\rho}} \|B\|_{\hat{\rho}} \quad (78)$$

for every $0 \leq \rho < \hat{\rho}$, where $C_{N_F}(\rho, \hat{\rho})$ is given in Theorem 3.3.

Notice that Corollary 3.8 is useful to control the product of approximated objects. If \tilde{A} and \tilde{B} are the corresponding approximations of A and B given by DFT, then

$$\|\tilde{A}\tilde{B} - \widetilde{\tilde{A}\tilde{B}}\|_{\rho} \leq C_{N_F}(\rho, \hat{\rho}) \|\tilde{A}\|_{\hat{\rho}} \|\tilde{B}\|_{\hat{\rho}} \leq C_{N_F}(\rho, \hat{\rho}) \|\tilde{A}\|_{F, \hat{\rho}} \|\tilde{B}\|_{F, \hat{\rho}} \quad (79)$$

for every $0 \leq \rho < \hat{\rho}$. Notice that since \tilde{A} and \tilde{B} are Fourier series with finite support, then it is interesting to control Equation (79) using Fourier norms.

The second result allows us to control the inverse of a matrix using the discrete Fourier approximation:

Corollary 3.9. *Let us consider a matrix function $A : \mathbb{T}^n \rightarrow \mathbb{C}^{m \times m}$ whose entries are analytic and bounded functions in the complex strip $\mathbb{T}_{\hat{\rho}}^n$ of size $\hat{\rho} > 0$. Given a grid of size $N_F = (N_{F,1}, \dots, N_{F,n})$, we evaluate A in the grid and compute the inverses $X(\theta_j) = A(\theta_j)^{-1}$. Then, if \tilde{X} is the corresponding discrete Fourier approximation associated to the sample $X(\theta_j)$, the error $E(\theta) = I_m - A(\theta)\tilde{X}(\theta)$ satisfies*

$$\|E\|_{\rho} \leq C_{N_F}(\rho, \hat{\rho}) \|A\|_{\hat{\rho}} \|\tilde{X}\|_{\hat{\rho}}, \quad (80)$$

for $0 \leq \rho < \hat{\rho}$. Moreover, if $\|E\|_{\rho} < 1$, there exists an analytic inverse $A^{-1} : \mathbb{T}^n \rightarrow \mathbb{C}^{m \times m}$ satisfying

$$\|A^{-1} - \tilde{X}\|_{\rho} \leq \frac{\|\tilde{X}\|_{\hat{\rho}} \|E\|_{\rho}}{1 - \|E\|_{\rho}}. \quad (81)$$

Proof. To obtain Equation (80) we observe that if $\widetilde{A\tilde{X}}$ is the discrete Fourier approximation of $A\tilde{X}$, then it turns out that

$$(A\tilde{X})(\theta_j) = A(\theta_j)\tilde{X}(\theta_j) = I_m$$

for all points in the grid. This implies that $\widetilde{A\tilde{X}} = I_m$ and we end up with

$$\|E\|_\rho = \|I_m - A\tilde{X}\|_\rho = \|\widetilde{A\tilde{X}} - A\tilde{X}\|_\rho$$

and Inequality (80) follows applying Corollary 3.8. Inequality (81) follows from the expression $E = I_m - A\tilde{X}$, simply writing $A^{-1} = \tilde{X}(I_m - E)^{-1}$ and using a Neumann series argument. \square

4 Dealing with the small divisors

In this Section we discuss two technical auxiliary results that play a fundamental role in KAM theory: the characterization of the Diophantine constants (γ, τ) and the computation of the Rüssmann constant c_R . In Section 4.1 we propose a general method to assign Diophantine constants to a given interval vector of frequencies, such that the corresponding set of Diophantine vectors has positive measure. In Section 4.2 we revisit the classic Rüssmann estimates. To take into account the effect of small divisors, we compute the first elements explicitly and then we control the remaining tail analytically. In this way, with the help of the computer, we obtain sharper estimates than in the classic literature.

4.1 On the characterization of Diophantine constants

A fundamental hypothesis of Theorem 2.6 is the fact that $\omega \in \mathbb{R}^n$ satisfies Diophantine conditions. To ensure it, we enclose ω with an interval vector ϖ and we look for constants (γ, τ) such that ϖ contains (γ, τ) -Diophantine vectors. The estimates presented in this section are based on two elementary observations. First, that we only need to give a lower bound of the measure of vectors $\omega \in \varpi$ satisfying $|k \cdot \omega - m| \geq \gamma|k|_1^{-\tau}$, for every $k \in \mathbb{Z}^n \setminus \{0\}$ and $m \in \mathbb{Z}$. Second, that this lower bound is obtained by splitting the computations in two parts: the low resonances are checked rigorously with the help of the computer, while the measure of the high resonances are bounded analytically.

Consider an interval vector of the form $\varpi = \prod_{i=1}^n [a_i, b_i]$, and Diophantine constants (γ, τ) . For each index $k \in \mathbb{Z}^n \setminus \{0\}$ we define the k -resonant set of type (γ, τ) as

$$\text{Res}_k(\varpi, \gamma, \tau) = \bigcup_{m \in \mathbb{Z}} \left\{ \omega \in \varpi : |k \cdot \omega - m| < \frac{\gamma}{|k|_1^\tau} \right\},$$

so that the resonant set of type (γ, τ) is

$$\text{Res}(\varpi, \gamma, \tau) = \bigcup_{k \in \mathbb{Z}^n \setminus \{0\}} \text{Res}_k(\varpi, \gamma, \tau).$$

The relative measure of the set of (γ, τ) -Diophantine vectors in ϖ is

$$p(\omega, \gamma, \tau) = 1 - \frac{\text{meas}(\text{Res}(\varpi, \gamma, \tau))}{\text{meas}(\varpi)}, \quad (82)$$

where $\text{meas}(A)$ stands for the Lebesgue measure of a Borel set A . Our goal is to obtain positive lower bounds of $p(\omega, \gamma, \tau)$. To do so, we control the resonant set by fixing M sufficiently big, and using the decomposition

$$\text{Res}(\varpi, \gamma, \tau) = \text{Res}_{\leq M}(\varpi, \gamma, \tau) \cup \text{Res}_{> M}(\varpi, \gamma, \tau),$$

where $\text{Res}_{\leq M}(\varpi, \gamma, \tau)$ and $\text{Res}_{> M}(\varpi, \gamma, \tau)$ are, respectively, the sets of resonances with index k satisfying $|k|_1 \leq M$ and $|k|_1 > M$. By choosing γ sufficiently small, we get $\text{Res}_{\leq M}(\varpi, \gamma, \tau) = \emptyset$, and then we have to get an upper bound of the measure of $\text{Res}_{> M}(\varpi, \gamma, \tau)$. These arguments are the core of the proof of the following proposition.

Proposition 4.1. Let $\varpi = \prod_{i=1}^n [a_i, b_i]$ be an interval vector, whose diameter is $\text{diam}(\varpi) = \sqrt{\sum_{i=1}^n (b_i - a_i)^2}$. Given $M \geq n$, we assume that for any $\omega \in \varpi$, $k \in \mathbb{Z}^n$ such that $0 < |k|_1 \leq M$, and $m \in \mathbb{Z}$, we have $k \cdot \omega - m \neq 0$. For any $\tau > n$, we define

$$\gamma_M(\varpi, \tau) = \min\{|k \cdot \omega - m| |k|_1^\tau : \omega \in \varpi, 0 < |k|_1 \leq M, m \in \mathbb{Z}\}.$$

Then, for any positive $\gamma \leq \gamma_M(\varpi, \tau)$, we have

$$p(\varpi, \gamma, \tau) > 1 - \frac{C(\varpi, n)\gamma}{(\tau - n)M^{\tau-n}} \geq 1 - \frac{C(\varpi, n)\gamma_M(\varpi, \tau)}{(\tau - n)M^{\tau-n}}, \quad (83)$$

where

$$C(\varpi, n) = \frac{2^{2n}}{(n-1)!} \frac{(\text{diam}(\varpi))^n}{\text{meas}(\varpi)}.$$

Moreover, the equation for τ

$$1 - \frac{C(\varpi, n)\gamma_M(\varpi, \tau)}{(\tau - n)M^{\tau-n}} = 0 \quad (84)$$

has a unique solution $\tau_M(\varpi)$, for which we define $\gamma_M(\varpi) = \gamma_M(\varpi, \tau_M(\varpi))$. As a consequence, for any pair (γ, τ) with $\tau \geq \tau_M(\varpi)$ and $\gamma \leq \gamma_M(\varpi)$, we have $p(\varpi, \gamma, \tau) > 0$.

Proof. Notice that, for $\gamma \leq \gamma_M(\varpi, \tau)$, we obtain $\text{Res}(\varpi, \gamma, \tau) = \text{Res}_{>M}(\varpi, \gamma, \tau)$, since $\text{Res}_{\leq M}(\varpi, \gamma, \tau) = \emptyset$. In order to get an upper bound of $\text{meas}(\text{Res}_{>M}(\varpi, \gamma, \tau))$, we use the elementary estimate

$$\text{meas}(\text{Res}_k(\varpi, \gamma, \tau)) \leq 2\gamma |k|_1^{-\tau} (\text{diam}(\varpi))^n,$$

for any $k \in \mathbb{Z}^n \setminus \{0\}$ (see the proof of Lemma 2.11 of [19]). Consequently,

$$\text{meas}(\text{Res}_{>M}(\varpi, \gamma, \tau)) \leq 2(\text{diam}(\varpi))^n \gamma \sum_{|k|_1 > M} |k|_1^{-\tau} < \frac{2^{2n}(\text{diam}(\varpi))^n}{(n-1)!(\tau-n)M^{\tau-n}}. \quad (85)$$

The lower bound (83) follows immediately.

Let us consider the function $\hat{p} : (n, \infty) \rightarrow \mathbb{R}$ defined by

$$\hat{p}(\tau) = 1 - \frac{C(\varpi, n)\gamma_M(\varpi, \tau)}{(\tau - n)M^{\tau-n}}. \quad (86)$$

Then, from the definition of $\gamma_M(\varpi, \tau)$,

$$\hat{p}(\tau) = 1 - \frac{C(\varpi, n)}{(\tau - n)M^{\tau-n}} \min \left\{ |k \cdot \omega - m| \left(\frac{|k|_1}{M} \right)^\tau : \omega \in \varpi, 0 < |k|_1 \leq M, m \in \mathbb{Z} \right\},$$

from where we deduce that \hat{p} is a strictly increasing function of τ , and $\lim_{\tau \rightarrow n^+} \hat{p}(\tau) = -\infty$. Moreover, since

$$\hat{p}(\tau) \geq 1 - \frac{C(\varpi, n)}{2(\tau - n)M^{\tau-n}},$$

then $\lim_{\tau \rightarrow \infty} \hat{p}(\tau) = 1$. Hence, there exists a unique $\tau = \tau_M(\varpi) > n$ such that $\hat{p}(\tau) = 0$. The rest of the proof follows immediately. \square

Remark 4.2. The constant $C(\varpi, n)$ in Proposition 4.1 is specially simple if the length of the edges of the interval vector are equal:

$$C(\varpi, n) = \frac{2^{2n} n^{\frac{n}{2}}}{(n-1)!} =: C(n).$$

In particular, we have $C(1) = 4$ and $C(2) = 32$.

In this paper we are not interested in maximizing the measure of Diophantine numbers but simply in guaranteeing that it is positive. Proposition 4.1 provides a simple algorithm to associate pairs (γ, τ) to an interval vector ϖ . Specifically, we take $M \geq n$ and solve Equation (84) to find the pair $(\gamma_M(\varpi), \tau_M(\varpi))$. Notice that $\tau_M(\varpi)$ is the minimum value of τ that guarantees the existence of Diophantine vectors in ϖ for $\gamma \leq \gamma_M(\varpi)$. To produce larger sets of Diophantine vectors, take a pair $(\gamma_M(\varpi, \tau), \tau)$, with $\tau \geq \tau_M(\varpi)$ (since the function \hat{p} defined in the proof is strictly increasing). Giving measure estimates of invariant tori (both in frequency space and in phase space) is a very interesting problem that deserves full attention and will be considered in future work.

To illustrate the previous construction we consider (tight) interval frequencies ϖ enclosing

$$[\omega_{a,b} - 2^{-50}, \omega_{a,b} + 2^{-50}], \quad \text{with} \quad \omega_{a,b} = \frac{\sqrt{b^2 + 4b/a} - b}{2}. \quad (87)$$

In Table 1 we provide Diophantine constants (γ, τ) , for some of these interval frequencies, such that the Lebesgue measure of (γ, τ) -Diophantine frequencies in ϖ is positive. The computations have been performed using the interval arithmetics library MPFI (see [65]) taking a precision of 64 bits. The values of Table 1 are obtained using $M = 1000$. If we take a larger value of M , then we can obtain a smaller lower value of τ . For example, for the golden mean (with $a = b = 1$), taking $M = 10^3$ we obtain $\tau \geq 1.26$, taking $M = 10^4$ we obtain $\tau \geq 1.22$, taking $M = 10^5$ we obtain $\tau \geq 1.19$, and taking $M = 10^6$ we obtain $\tau \geq 1.17$,

| a | b | $\gamma \leq$ | $\tau \geq$ | a | b | $\gamma \leq$ | $\tau \geq$ | a | b | $\gamma \leq$ | $\tau \geq$ |
|-----|-----|-------------------|-------------|-----|-----|-------------------|-------------|-----|-----|-------------------|-------------|
| 1 | 1 | 0.381966011250104 | 1.26 | 3 | 1 | 0.263762615825972 | 1.23 | 5 | 1 | 0.170820393249935 | 1.19 |
| 1 | 2 | 0.267949192431121 | 1.23 | 3 | 2 | 0.290994448735804 | 1.24 | 5 | 2 | 0.183215956619922 | 1.20 |
| 1 | 3 | 0.208712152522079 | 1.21 | 3 | 3 | 0.302775637731993 | 1.24 | 5 | 3 | 0.188194301613412 | 1.20 |
| 1 | 4 | 0.171572875253808 | 1.19 | 3 | 4 | 0.277309053319640 | 1.23 | 5 | 4 | 0.190890230020663 | 1.20 |
| 1 | 5 | 0.145898033750314 | 1.18 | 3 | 5 | 0.223037765858308 | 1.21 | 5 | 5 | 0.192582403567251 | 1.20 |
| 1 | 6 | 0.127016653792582 | 1.17 | 3 | 6 | 0.187329140491556 | 1.20 | 5 | 6 | 0.193743884534261 | 1.20 |
| 2 | 1 | 0.366025403784437 | 1.26 | 4 | 1 | 0.207106781186546 | 1.21 | 6 | 1 | 0.145497224367901 | 1.18 |
| 2 | 2 | 0.413767832000904 | 1.27 | 4 | 2 | 0.224744871391588 | 1.22 | 6 | 2 | 0.154700538379250 | 1.18 |
| 2 | 3 | 0.300011472016747 | 1.24 | 4 | 3 | 0.232050807568876 | 1.22 | 6 | 3 | 0.158312395177698 | 1.19 |
| 2 | 4 | 0.235323972166368 | 1.22 | 4 | 4 | 0.236067977499788 | 1.22 | 6 | 4 | 0.160246899469285 | 1.19 |
| 2 | 5 | 0.192798030208926 | 1.20 | 4 | 5 | 0.238612787525829 | 1.22 | 6 | 5 | 0.161453237111884 | 1.19 |
| 2 | 6 | 0.163806299636515 | 1.19 | 4 | 6 | 0.206140402288459 | 1.21 | 6 | 6 | 0.162277660168378 | 1.19 |

Table 1: Rigorously computed Diophantine constants (γ, τ) ensuring positive measure for several 1-dimensional (tight) interval frequencies ϖ enclosing intervals given by Equation (87). We use the methodology derived from Proposition 4.1 with $M = 1000$.

As an illustration for $n = 2$, we consider (tight) interval frequency vectors ϖ enclosing

$$[\omega_p - 2^{-50}, \omega_p + 2^{-50}] \times [\omega_q - 2^{-50}, \omega_q + 2^{-50}] \subset \mathbb{R}^2, \quad (88)$$

with $\omega_p = \sqrt{p} - [\sqrt{p}]$ and $\omega_q = \sqrt{q} - [\sqrt{q}]$, where $[\cdot]$ stands for the integer part. The computations have been performed using the interval arithmetics library MPFI (see [65]) taking a precision of 64 bits. Diophantine constants associated to these intervals are given in Table 2, using the methodology derived from Proposition 4.1 with $M = 1000$.

| | $p = 2$ | $p = 3$ | $p = 5$ | $p = 7$ | $p = 11$ | $p = 13$ |
|----------|---------|--------------------|--------------------|--------------------|--------------------|--------------------|
| $q = 2$ | | 0.1421950391579065 | 0.0100027079360758 | 0.0216249622296120 | 0.1108017984300913 | 0.0642768464253883 |
| $q = 3$ | 2.40 | | 0.1706190237467459 | 0.3712175538099537 | 0.0587500580246455 | 0.2621757106909654 |
| $q = 5$ | 2.14 | 2.42 | | 0.0090434339906178 | 0.0939733872704087 | 0.1120569499083642 |
| $q = 7$ | 2.20 | 2.51 | 2.13 | | 0.0139696161322762 | 0.1940680771272715 |
| $q = 11$ | 2.37 | 2.30 | 2.35 | 2.17 | | 0.0044815206623438 |
| $q = 13$ | 2.31 | 2.47 | 2.37 | 2.43 | 2.09 | |

Table 2: Rigorously computed Diophantine constants (γ, τ) ensuring positive measure for several 2-dimensional (tight) interval frequencies ϖ enclosing intervals given by Equation (88). We use the methodology derived from Proposition 4.1 with $M = 1000$. An upper value of γ is given above the diagonal and a lower value of τ is given below the diagonal.

4.2 On the Rüssmann estimates

In this section we present a version of the Rüssmann estimates that is tailored to be evaluated with the help of the computer. We use the notation introduced in Section 2.2.

Lemma 4.3. *Let $\omega \in \mathbb{R}^n$ be a (γ, τ) -Diophantine frequency vector, for certain $\gamma > 0$ and $\tau \geq n$ (see Definition 2.5). Then, for any analytic function $v : \mathbb{T}_\rho^n \rightarrow \mathbb{C}$, with $\|v\|_\rho < \infty$ and $\rho > 0$, there exists a unique zero-average analytic solution $u : \mathbb{T}_\rho^n \rightarrow \mathbb{C}$ of $\mathcal{L}u = v - \langle v \rangle$, denoted by $u = \mathcal{R}v$. Moreover, given $L \in \mathbb{N}$, for any $0 < \delta < \rho$ we have*

$$\|u\|_{\rho-\delta} \leq \frac{c_R(\delta)}{\gamma \delta^\tau} \|v\|_\rho, \quad (89)$$

where

$$c_R(\delta) = \sqrt{\gamma^2 \delta^{2\tau} 2^n \sum_{0 < |k|_1 \leq L} \frac{e^{-4\pi|k|_1 \delta}}{4|\sin(\pi k \cdot \omega)|^2} + 2^{n-3} \zeta(2, 2\tau) (2\pi)^{-2\tau} \int_{4\pi\delta(L+1)}^\infty u^{2\tau} e^{-u} du} \quad (90)$$

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

Proof. We follow standard arguments (see [19, 37, 51, 67, 69]), with an eye in the feasibility of computing rigorous upper bounds of finite sums (up to order L).

We control the divisors in the expansion of the function $u(\theta) = \mathcal{R}v(\theta)$, formally given by Equation (20), as

$$|1 - e^{2\pi i k \cdot \omega}| = 2|\sin(\pi k \cdot \omega)| \geq 2^2 \min_{m \in \mathbb{Z}} |k \cdot \omega - m|,$$

where we used that $\sin x \geq 2x/\pi$ if $x \leq \pi/2$ and that $\min_{m \in \mathbb{Z}} |k \cdot \omega - m| < 1/2$. Then, it is natural to introduce the notation $d_k = k \cdot \omega - m_k$, such that $|d_k| = \min_{m \in \mathbb{Z}} |k \cdot \omega - m|$. Notice that the divisors d_k satisfy $d_{k_1} \neq d_{k_2}$ if $k_1 \neq k_2$, and $d_{-k} = -d_k$. The Diophantine condition in (21) reads $|d_k| \geq \gamma |k|_1^{-\tau}$.

We then control the norm of u as

$$\begin{aligned} \|u\|_{\rho-\delta} &\leq \|u\|_{F, \rho-\delta} \leq \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{|\hat{v}_k|}{2|\sin(\pi k \cdot \omega)|} e^{2\pi|k|_1(\rho-\delta)} \\ &\leq \left(\sum_{k \in \mathbb{Z}^n \setminus \{0\}} |\hat{v}_k|^2 e^{4\pi|k|_1 \rho} \right)^{1/2} \left(\sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{e^{-4\pi|k|_1 \delta}}{2^2 \sin^2(\pi k \cdot \omega)} \right)^{1/2}, \end{aligned} \quad (91)$$

where we used Cauchy-Schwarz inequality. On the one hand, the first term is bounded by

$$\sum_{k \in \mathbb{Z}^n \setminus \{0\}} |\hat{v}_k|^2 e^{4\pi|k|_1 \rho} \leq \sum_{k \in \mathbb{Z}^n} |\hat{v}_k|^2 e^{4\pi|k|_1 \rho} \leq 2^n \|v\|_\rho^2 \quad (92)$$

(see [67]), and on the other hand, the second term is bounded by computing the sum up to order L and controlling the tail,

$$\sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{e^{-4\pi|k|_1 \delta}}{4|\sin(\pi k \cdot \omega)|^2} = \sum_{0 < |k|_1 \leq L} \frac{e^{-4\pi|k|_1 \delta}}{4|\sin(\pi k \cdot \omega)|^2} + \sum_{|k|_1 > L} \frac{e^{-4\pi|k|_1 \delta}}{4|\sin(\pi k \cdot \omega)|^2}. \quad (93)$$

By using that $2|\sin(\pi k \cdot \omega)| \geq 2^2 |d_k|$, and the Abel summation formula, we bound the tail by

$$\sum_{|k|_1 > L} \frac{e^{-4\pi|k|_1 \delta}}{4|\sin(\pi k \cdot \omega)|^2} \leq \sum_{\ell=L+1}^\infty \left(\sum_{\substack{k \in \mathbb{Z}^n \setminus \{0\} \\ |k|_1 \leq \ell}} \frac{1}{2^4 |d_k|^2} \right) (e^{-4\pi\ell\delta} - e^{-4\pi(\ell+1)\delta}). \quad (94)$$

Then, given $\ell \in \mathbb{N}$, we define the set of positive divisors up to order ℓ as

$$\mathcal{D}_\ell = \{k \in \mathbb{Z}^n \setminus \{0\} : |k|_1 \leq \ell \text{ and } d_k > 0\},$$

and we sort the divisors according to $0 < d_{k_1} < \dots < d_{k_{\#\mathcal{D}_\ell}}$ with $k_j \in \mathcal{D}_\ell$, for $j = 1, \dots, \#\mathcal{D}_\ell$. We obtain recursively that

$$d_{k_j} = (d_{k_j} - d_{k_{j-1}}) + \dots + (d_{k_2} - d_{k_1}) + d_{k_1} \geq (j-1)\gamma(2\ell)^{-\tau} + \gamma\ell^{-\tau},$$

where we used that $|k_1|_1 \leq \ell$ and $|k_i - k_{i-1}|_1 \leq 2\ell$. Then, we have

$$\sum_{j=1}^{\#\mathcal{D}_\ell} \frac{1}{(d_{k_j})^2} \leq \sum_{j=1}^{\infty} \frac{\ell^{2\tau}}{\gamma^2(1+(j-1)2^{-\tau})^2} = 2^{2\tau} \sum_{j=0}^{\infty} \frac{1}{(2^\tau + j)^2} \frac{\ell^{2\tau}}{\gamma^2} = 2^{2\tau} \zeta(2, 2^\tau) \frac{\ell^{2\tau}}{\gamma^2}.$$

The same result is obtained for the sum corresponding to the negative divisors up to order ℓ .

Finally, we control the sum of Equation (94) as follows

$$\begin{aligned} \sum_{|k|_1 > L} \frac{e^{-4\pi|k|_1\delta}}{4|\sin(\pi k \cdot \omega)|^2} &\leq \frac{2^{2\tau} \zeta(2, 2^\tau)}{2^3 \gamma^2} \sum_{\ell > L} \ell^{2\tau} \int_{\ell}^{\ell+1} 4\pi\delta e^{-4\pi\delta x} dx \\ &\leq \frac{4\pi\delta 2^{2\tau} \zeta(2, 2^\tau)}{2^3 \gamma^2} \int_{L+1}^{\infty} x^{2\tau} e^{-4\pi\delta x} dx = \frac{2^{-3} \zeta(2, 2^\tau)}{\gamma^2 (2\pi\delta)^{2\tau}} \int_{4\pi\delta(L+1)}^{\infty} u^{2\tau} e^{-u} du. \end{aligned} \quad (95)$$

Combining Equations (91), (92), and (95), we end up with the stated estimate. \square

Remark 4.4. Taking $L = 0$ we obtain

$$c_R(\delta) = \sqrt{2^{n-3} \zeta(2, 2^\tau) (2\pi)^{-2\tau} \int_{4\pi\delta}^{\infty} u^{2\tau} e^{-u} du} \leq \sqrt{2^{n-3} \zeta(2, 2^\tau) (2\pi)^{-2\tau} \Gamma(2\tau + 1)}, \quad (96)$$

which gives us the classic (uniform) Rüssmann estimate.

If we use a computer to control the first divisors explicitly, then it turns out that the expression $c_R(\delta)$ in Equation (90) improves the classic estimate in Equation (96). To this end, we enclose ω with an interval vector ϖ , as described in Section 4.1, and we rigorously enclose the finite sum for $0 < |k|_1 \leq L$ using interval arithmetics. We consider upper bounds of the integral in the tail using that, if $y > x$,

$$\int_y^{\infty} u^x e^{-u} du \leq \frac{y}{y-x} y^x e^{-y}.$$

Applying this last estimate requires to take L such that $4\pi\delta(L+1) > 2\tau$ in (95). In practice, we take a value of L large enough in such a way that the contribution of the tail is smaller than the contribution of the finite sum.

In Tables 3 and 4 we present some rigorous bounds of $c_R(\delta)$, given by Equation (90), for several values of δ and we compare them with the classic constant c_R , given by Equation (96). Specifically, in Table 3 we consider the same 1-dimensional interval frequencies characterized in Table 1. In Table 4 we consider the same 2-dimensional interval frequencies characterized in Table 2. The computations have been performed using the interval arithmetics library MPFI (see [65]) taking a precision of 64 bits. We observe that the improvement of Equation (90) is remarkable and, in some cases, we improve the classic constant by several orders of magnitude. As we will see, this represents an important improvement in order to apply the KAM theorem.

5 Validation algorithm to apply the KAM theorem

In this section we present a methodology to perform computer assisted validations of Lagrangian quasi-periodic invariant tori in exact symplectic maps. Given an approximately F -invariant torus (e.g. obtained numerically) we have to rigorously bound the error of such an approximation, and to rigorously verify the *a priori* hypotheses of Theorem 2.6. One of the applications of the proposed methodology falls into one of the main strains of the field of *validated numerics* and, for this reason, we use the term *validation algorithm* in what follows.

| $\omega_{a,b}$ | | $c_R(\delta) \leq$ | | | | | |
|----------------|-----|--------------------|----------------|-----------------|------------------|-------------------|--------------------|
| a | b | $L = 0$ | $\delta = 0.1$ | $\delta = 0.01$ | $\delta = 0.001$ | $\delta = 0.0001$ | $\delta = 0.00001$ |
| 1 | 1 | 6.53700395e-02 | 1.70002315e-02 | 1.01408017e-02 | 5.57856565e-03 | 3.06566441e-03 | 1.68472062e-03 |
| 1 | 2 | 6.76832915e-02 | 1.40967097e-02 | 8.84204788e-03 | 5.26914629e-03 | 3.08884783e-03 | 1.79851814e-03 |
| 1 | 3 | 6.92956514e-02 | 1.29193920e-02 | 8.71895511e-03 | 5.18084015e-03 | 3.29070636e-03 | 1.98865837e-03 |
| 1 | 4 | 7.09671322e-02 | 1.25818941e-02 | 8.36835924e-03 | 5.84462956e-03 | 3.60202865e-03 | 2.21507905e-03 |
| 1 | 5 | 7.18258809e-02 | 1.23983728e-02 | 7.88617196e-03 | 5.47041537e-03 | 3.93080619e-03 | 2.62830129e-03 |
| 1 | 6 | 7.27004339e-02 | 1.24386826e-02 | 8.01767004e-03 | 5.19907101e-03 | 3.59023334e-03 | 2.59913106e-03 |
| 2 | 1 | 6.53700395e-02 | 1.74727821e-02 | 1.04651045e-02 | 5.78939864e-03 | 3.21601209e-03 | 1.75767438e-03 |
| 2 | 2 | 6.46258261e-02 | 1.87353432e-02 | 1.09024839e-02 | 5.85645816e-03 | 3.14665921e-03 | 1.68979306e-03 |
| 2 | 3 | 6.68984767e-02 | 1.63024413e-02 | 9.88023684e-03 | 5.56751156e-03 | 3.12759392e-03 | 1.77052228e-03 |
| 2 | 4 | 6.84822445e-02 | 1.53872642e-02 | 9.71562658e-03 | 5.85967409e-03 | 3.53201756e-03 | 2.12887318e-03 |
| 2 | 5 | 7.01238363e-02 | 1.50961843e-02 | 9.60632258e-03 | 5.78570281e-03 | 3.43840456e-03 | 2.04301934e-03 |
| 2 | 6 | 7.09671322e-02 | 1.48540346e-02 | 9.04822601e-03 | 5.05304849e-03 | 2.95729166e-03 | 1.98349009e-03 |
| 3 | 1 | 6.76832915e-02 | 1.46809408e-02 | 1.00993505e-02 | 5.63999620e-03 | 3.50708184e-03 | 1.95894122e-03 |
| 3 | 2 | 6.68984767e-02 | 1.43496537e-02 | 8.98100848e-03 | 5.22239107e-03 | 3.08135263e-03 | 1.82028295e-03 |
| 3 | 3 | 6.68984767e-02 | 1.58579091e-02 | 9.83769747e-03 | 5.65400813e-03 | 3.24897579e-03 | 1.86887029e-03 |
| 3 | 4 | 6.76832915e-02 | 1.62566583e-02 | 9.69074106e-03 | 5.59759584e-03 | 3.32500705e-03 | 1.99628047e-03 |
| 3 | 5 | 6.92956514e-02 | 1.50813939e-02 | 8.58023012e-03 | 5.31739360e-03 | 3.58299882e-03 | 2.46021534e-03 |
| 3 | 6 | 7.01238363e-02 | 1.42772183e-02 | 7.56136730e-03 | 5.20508341e-03 | 4.02265302e-03 | 2.54652970e-03 |
| 4 | 1 | 6.92956514e-02 | 1.31882573e-02 | 9.51364826e-03 | 6.14222557e-03 | 3.50022063e-03 | 2.17640510e-03 |
| 4 | 2 | 6.84822445e-02 | 1.26639499e-02 | 8.35820776e-03 | 5.03133457e-03 | 3.02541383e-03 | 1.81923509e-03 |
| 4 | 3 | 6.84822445e-02 | 1.32535840e-02 | 9.03724309e-03 | 5.61174299e-03 | 3.33248407e-03 | 1.90879693e-03 |
| 4 | 4 | 6.84822445e-02 | 1.39802574e-02 | 9.34078901e-03 | 5.55619944e-03 | 3.32472818e-03 | 2.05369884e-03 |
| 4 | 5 | 6.84822445e-02 | 1.48184252e-02 | 9.49849042e-03 | 5.85472306e-03 | 3.88991086e-03 | 2.07085456e-03 |
| 4 | 6 | 6.92956514e-02 | 1.38269028e-02 | 8.66584042e-03 | 6.22603715e-03 | 3.90134199e-03 | 1.99958702e-03 |
| 5 | 1 | 7.09671322e-02 | 1.26864212e-02 | 8.86482905e-03 | 6.21678963e-03 | 4.26563195e-03 | 2.68508905e-03 |
| 5 | 2 | 7.01238363e-02 | 1.22725538e-02 | 8.83966465e-03 | 5.80010426e-03 | 3.74378915e-03 | 2.37051014e-03 |
| 5 | 3 | 7.01238363e-02 | 1.24917622e-02 | 9.34589411e-03 | 5.62325597e-03 | 3.07349358e-03 | 1.89609801e-03 |
| 5 | 4 | 7.01238363e-02 | 1.27481886e-02 | 9.35639250e-03 | 5.19209738e-03 | 3.32315143e-03 | 2.18847977e-03 |
| 5 | 5 | 7.01238363e-02 | 1.30462113e-02 | 9.23511743e-03 | 5.79164682e-03 | 3.51817789e-03 | 2.36970971e-03 |
| 5 | 6 | 7.01238363e-02 | 1.33875654e-02 | 9.20387961e-03 | 6.71312444e-03 | 3.57979848e-03 | 2.30908333e-03 |
| 6 | 1 | 7.18258809e-02 | 1.24345830e-02 | 8.50000514e-03 | 5.57984051e-03 | 3.88857456e-03 | 2.76605873e-03 |
| 6 | 2 | 7.18258809e-02 | 1.23988684e-02 | 9.62126814e-03 | 6.51124354e-03 | 4.05219896e-03 | 2.34111803e-03 |
| 6 | 3 | 7.09671322e-02 | 1.22004393e-02 | 9.29930218e-03 | 4.89324773e-03 | 2.80122249e-03 | 2.10296918e-03 |
| 6 | 4 | 7.09671322e-02 | 1.22880935e-02 | 8.91873609e-03 | 4.68287840e-03 | 3.19881083e-03 | 2.50645650e-03 |
| 6 | 5 | 7.09671322e-02 | 1.23862489e-02 | 8.53842453e-03 | 5.33661343e-03 | 3.55154114e-03 | 2.28605325e-03 |
| 6 | 6 | 7.09671322e-02 | 1.24977263e-02 | 8.40455003e-03 | 5.93603968e-03 | 3.81558500e-03 | 2.25377574e-03 |

Table 3: We give a rigorous upper bound of $c_R(\delta)$, given by Equation (90) for several 1-dimensional tight interval frequencies ϖ enclosing intervals given by Equation (87). We use the values (γ, τ) provided in Table 1. The column $L = 0$ corresponds to the classic Rüssmann constant in Equation (96).

| ϖ | | $c_R(\delta) \leq$ | | | | | |
|----------|-----|--------------------|----------------|-----------------|-----------------|------------------|------------------|
| p | q | $L = 0$ | $\delta = 0.1$ | $\delta = 0.05$ | $\delta = 0.01$ | $\delta = 0.005$ | $\delta = 0.001$ |
| 2 | 3 | 3.62859961e-02 | 3.10060284e-03 | 6.00977402e-03 | 1.50395887e-03 | 1.23391239e-03 | 8.15687337e-04 |
| 2 | 5 | 4.24412098e-02 | 6.88136970e-04 | 8.38062576e-04 | 6.85210209e-03 | 3.48793730e-03 | 1.39435290e-03 |
| 2 | 7 | 4.08296165e-02 | 1.51913516e-03 | 1.64056165e-03 | 5.96490692e-03 | 2.86266455e-03 | 1.08117640e-03 |
| 2 | 11 | 3.68950309e-02 | 3.40361447e-03 | 5.37454913e-03 | 2.42226866e-03 | 1.16692138e-03 | 5.72003609e-04 |
| 2 | 13 | 3.81857929e-02 | 3.34411162e-03 | 1.88186836e-03 | 5.70267600e-03 | 6.35005202e-03 | 2.59988091e-03 |
| 3 | 5 | 3.58927516e-02 | 6.12958788e-03 | 3.95221883e-03 | 2.42682994e-03 | 1.30875574e-03 | 1.17588322e-03 |
| 3 | 7 | 3.42417879e-02 | 6.62761410e-03 | 4.15150247e-03 | 2.08654672e-03 | 1.29692066e-03 | 1.82961175e-03 |
| 3 | 11 | 3.84108018e-02 | 5.79145725e-03 | 5.55693971e-03 | 8.86690139e-04 | 7.03975422e-04 | 1.57560298e-03 |
| 3 | 13 | 3.49523429e-02 | 5.16292254e-03 | 6.09178577e-03 | 2.29447926e-03 | 1.84095693e-03 | 1.36098576e-03 |
| 5 | 7 | 4.27225331e-02 | 6.29450376e-03 | 6.95330075e-03 | 8.63897764e-04 | 3.99507510e-04 | 1.95384727e-03 |
| 5 | 11 | 3.73142594e-02 | 2.16195541e-03 | 3.88935752e-03 | 1.44076003e-03 | 7.74345647e-04 | 6.57036231e-03 |
| 5 | 13 | 3.68950309e-02 | 2.73635212e-03 | 4.34566648e-03 | 2.63259686e-03 | 1.00779405e-03 | 1.73212787e-03 |
| 7 | 11 | 4.16193859e-02 | 2.03472877e-03 | 6.68639919e-03 | 2.04760389e-03 | 7.38986151e-04 | 2.71601280e-04 |
| 7 | 13 | 3.56998570e-02 | 4.23255275e-03 | 4.16600489e-03 | 2.44375364e-03 | 1.91977769e-03 | 4.23864594e-03 |
| 11 | 13 | 4.38862946e-02 | 2.62790710e-03 | 7.60559938e-03 | 2.07028161e-03 | 6.59592023e-04 | 2.09531497e-04 |

Table 4: We give a rigorous upper bound of $c_R(\delta)$, given by Equation (90) for several 2-dimensional (tight) interval frequencies ϖ enclosing intervals given by Equation (88). We use the values (γ, τ) provided in Table 2. The column $L = 0$ corresponds to the classic Rüssmann constant in Equation (96).

We use fast Fourier transform to manipulate approximations of periodic functions and then we use the error estimates discussed in Section 3. Hence, the computational cost of the proposed methodology is of order $O(N_D \log N_D)$, where N_D is the total number of used Fourier coefficients and the asymptotic constant depends on the problem. Observe that we do not require to perform any symbolic manipulation of Fourier series (as in [29]) which has cost $O(N_D^2)$. An important aspect of our approach is that the asymptotic cost of the validation coincides with the asymptotic cost of computing invariant tori using the parameterization method (see [8, 31, 37, 41]).

5.1 A validation algorithm

In this section we propose an algorithm to apply Theorem 2.6 in particular problems. The algorithm is stated at a formal level but, when implementing it using a computer, operations must be performed using interval arithmetics. In Algorithm 5.1 we overview the procedure. A detailed discussion of each step is given in separate subsections.

Let us first present some useful notation. We recall that $F : \mathcal{A} \rightarrow \mathcal{A}$ is homotopic to the identity and $K : \mathbb{T}^n \rightarrow \mathcal{A}$ is homotopic to the zero section. For this reason, it is interesting to introduce the notation

$$F(x, y) = (x, 0) + F_p(x, y),$$

and

$$K(\theta) = (\theta, 0) + K_p(\theta).$$

To handle periodic functions, we approximate them using discrete Fourier transform and we control such approximation using the estimates presented in Section 3. Given a periodic function $f : \mathbb{T}^n \rightarrow \mathbb{C}$ and a regular grid of size $N_F = (N_{F,1}, \dots, N_{F,n}) \in \mathbb{N}^n$, satisfying that $N_{F,\ell} = 2^{q_\ell}$, with $q_\ell \in \mathbb{N}$, for $\ell = 1, \dots, n$, we consider a sample of points $\theta_j \in \mathbb{T}^n$, given by (75), that defines an n -dimensional array $\{f_j\}$, $f_j = f(\theta_j)$, where $j = (j_1, \dots, j_n)$, with $0 \leq j_\ell < N_{F,\ell}$ and $1 \leq \ell \leq n$.

We recall that the total number of points is given by $N_D = N_{F,1} \cdots N_{F,n}$ and that we denote by \tilde{f} the discrete Fourier approximation given by Equation (76). For every sample $\{f_j\}$ we introduce the forward discrete Fourier transform

$$\{\tilde{f}_k\} = \text{DFT}_F(\{f_j\}), \quad \text{with} \quad \tilde{f}_k = \frac{1}{N_D} \sum_{0 \leq j < N_F} f_j e^{-2\pi i k \cdot \theta_j}, \quad (97)$$

with $k = (k_1, \dots, k_n)$ and $-\frac{N_{F,\ell}}{2} \leq k_\ell < \frac{N_{F,\ell}}{2}$. Similarly, we introduce the backward discrete Fourier transform

$$\{f_j\} = \text{DFT}_B(\{\tilde{f}_k\}), \quad \text{with} \quad f_j = \sum_{-\frac{N_F}{2} \leq k < \frac{N_F}{2}} \tilde{f}_k e^{2\pi i k \cdot \theta_j}. \quad (98)$$

Notice that Equations (97) and (98) are just formal definitions. When implementing the validation algorithm we will resort to FFT to evaluate these expressions. Such FFT algorithms carry operations in a particular ordering, which may not be optimal in terms of rounding operations (when using interval arithmetics), but allows performing fast computations. From now on we use the notation FFT_F and FFT_B to refer to Equations (97) and (98) evaluated according to the FFT algorithms.

Finally, we give some details on the manipulation of functions discretized in the Fourier space. Given a periodic function f , discretized as $\{\tilde{f}_k\}$, we compute the Fourier discretization of a partial derivative $\partial_{\theta_i} f$ as $\{(\partial_{\theta_i} f)_k\} = \{2\pi k_i i \tilde{f}_k\}$, which corresponds to a diagonal operator in Fourier space. Similarly, given an interval vector ϖ , the composition $f \circ R_\varpi$ is approximated as $\{(\widetilde{f \circ R_\varpi})_k\} = \{e^{2\pi i k \cdot \varpi} \tilde{f}_k\}$.

Algorithm 5.1. *Given an annulus \mathcal{A} endowed with an exact symplectic structure $\omega = d\alpha$ represented by $\Omega(z)$ and $a(z)$, let us consider an exact symplectic map $F : \mathcal{A} \rightarrow \mathcal{A}$ and the following input:*

- **Input 1.** *Sampling of an approximately invariant torus: a sampling $\{K_{p,j}\}$, with $K_{p,j} = K_0(\theta_j) - (\theta_j, 0)$ on a regular grid on \mathbb{T}^n of size $N_F = (N_{F,1}, \dots, N_{F,n})$, with $N_{F,\ell} = 2^{q_\ell}$ and $q_\ell \in \mathbb{N}$, for $\ell = 1, \dots, n$.*

- **Input 2.** Transversal vectors: a map $N_0 : \mathbb{T}^n \rightarrow \mathbb{R}^{2n \times n}$.
- **Input 3.** Frequency vector: a pair of constants $\tau > 0$ and $\gamma > 0$, and an interval vector $\varpi = (\varpi_1, \dots, \varpi_n)$ containing a (γ, τ) -Diophantine vector of frequencies.
- **Input 4.** KAM parameters: constants $\delta, \sigma, \rho, \hat{\rho}, \rho_\infty$, and d_B satisfying $0 < \delta < \rho/3$, $0 < \rho < \hat{\rho}$, $1 < \sigma$, $0 < \rho_\infty < \rho$, and $d_B > 0$.

Then, proceed as follows:

- **Step 0.** Parameterization of the approximately invariant torus: compute the parameterization $K(\theta) = (\theta, 0) + \tilde{K}_p(\theta)$, and characterize the global complex domain \mathcal{B} . Details are provided in Section 5.2.
- **Step 1.** Error of invariance: compute a constant b_E such that $\|E\|_\rho \leq b_E$. Details are provided in Section 5.3.
- **Step 2.** Symplectic frame: compute constants $b_{DK}, b_{DK^\top}, b_{N_0^\top(\Omega \circ K)N_0}, b_B, b_A, b_N$, and b_{N^\top} such that $\|DK\|_\rho \leq b_{DK}$, $\|DK^\top\|_\rho \leq b_{DK^\top}$, $\|N_0^\top(\Omega \circ K)N_0\|_\rho \leq b_{N_0^\top(\Omega \circ K)N_0}$, $\|B\|_\rho \leq b_B$, $\|A\|_\rho \leq b_A$, $\|N\|_\rho \leq b_N$, and $\|N^\top\|_\rho \leq b_{N^\top}$. Details are provided in Section 5.4.
- **Step 3.** Torsion matrix: compute the discrete Fourier approximation \tilde{T} of T and compute a constant b_T such that $|\langle T \rangle| \leq b_T$. Details are provided in Section 5.5.
- **Step 4.** Hypotheses of the theorem: introduce the constants

$$\sigma_{DK} = b_{DK} \sigma, \quad \sigma_{DK^\top} = b_{DK^\top} \sigma, \quad \sigma_{N_0^\top(\Omega \circ K)N_0} = b_{N_0^\top(\Omega \circ K)N_0} \sigma, \quad \sigma_B = b_B \sigma, \quad \sigma_T = b_T \sigma, \quad (99)$$

so that (using $\sigma > 1$) Hypotheses H_1, H_2, H_3, H_4 , and H_5 in Theorem 2.6 are satisfied. Then, compute constants \mathfrak{C}_1 and \mathfrak{C}_2 . Details are provided in Section 5.6.

If the condition

$$\frac{\mathfrak{C}_1 b_E}{\gamma^4 \rho^{4\tau}} < 1$$

holds, then, for every (γ, τ) -Diophantine frequency $\omega \in \varpi$ there exists an F -invariant torus $K_{\infty, \omega}(\mathbb{T}^n)$, analytic in the strip of width $\rho_\infty = \rho/a_2$. Moreover, these invariant tori satisfy

$$\|DK_{\infty, \omega}\|_{\rho_\infty} < \sigma_{DK}, \quad \|DK_{\infty, \omega}^\top\|_{\rho_\infty} < \sigma_{DK^\top}, \quad \|K_{\infty, \omega} - K\|_{\rho_\infty} < \frac{\mathfrak{C}_2 b_E}{\gamma^2 \rho^{2\tau}}.$$

Remark 5.2. Incidentally, the above process gives a lower bound of the relative measure of the vectors ω in ϖ for which the KAM theorem applies. Assume that the pair (γ, τ) has been assigned to the interval vector ϖ by the algorithm derived from Proposition 4.1 (using a given value of M). Notice that for $\gamma_0 < \gamma$ the constant $\mathfrak{C}_1 = \mathfrak{C}_1(\gamma)$ is an upper bound of $\mathfrak{C}_1(\gamma_0)$. Then, if we take γ_0 such that

$$\frac{\mathfrak{C}_1(\gamma) b_E}{\gamma_0^4 \rho^{4\tau}} = 1,$$

it turns out that the theorem also applies for (γ_0, τ) -Diophantine vectors. A straightforward computation gives that

$$p(\varpi, \gamma_0, \tau) > 1 - \frac{C(\varpi, n) \gamma_0}{(\tau - n) M^{\tau - n}} = 1 - \frac{C(\varpi, n) \sqrt[4]{\mathfrak{C}_1(\gamma) b_E}}{(\tau - n) M^{\tau - n} \rho^\tau}. \quad (100)$$

5.2 Implementation details of Step 0

This preliminary step is performed in order to simplify the implementation of the algorithm. We observe that Theorem 2.6 deals with real analytic objects with real analytic derivatives. Notice that this property is not preserved by DFT. A simple way to avoid this problem is to consider the validation of a suitable parameterization. To this end, we perform the following computations:

- We compute $\tilde{K}_p = \{\tilde{K}_{p,k}\} = \text{FFT}_F(\{K_{p,j}\})$, where $\{K_{p,j}\}$ is the sample given in Input 1.
- We set to zero the coefficients $\tilde{K}_{p,k}$ with $k = (k_1, \dots, k_n)$ such that $-N_{F,\ell}/2 \leq k_\ell \leq -N_{F,\ell}/4$ or $N_{F,\ell}/4 \leq k_\ell < N_{F,\ell}/2$ for some index k_ℓ .
- We set the parameterization $K(\theta) = (\theta, 0) + \tilde{K}_p(\theta)$ and redefine the sampling $\{K_{p,j}\} = \text{FFT}_B(\{\tilde{K}_{p,k}\})$, replacing the original one.

Remark 5.3. *An alternative formulation of Algorithm 5.1 would consist in considering the constructed discrete Fourier approximation $\tilde{K} = \{\tilde{K}_{p,k}\}$ in Input 1.*

Now we introduce the domain \mathcal{B} given by

$$\mathcal{B} = \{(x, y) \in \mathbb{C}^n / \mathbb{Z}^n \times \mathbb{C}^n : |\text{Im } x_i| \leq d_{\mathcal{B}} + \rho + \|K_p^{x_i}\|_{F,\rho}, |y_i| \leq d_{\mathcal{B}} + \|K_p^{y_i}\|_{F,\rho}\}, \quad (101)$$

and the domain $\hat{\mathcal{B}}$ given by

$$\hat{\mathcal{B}} = \{(x, y) \in \mathbb{C}^n / \mathbb{Z}^n \times \mathbb{C}^n : |\text{Im } x_i| \leq \hat{\rho} + \|K_p^{x_i}\|_{F,\hat{\rho}}, |y_i| \leq \|K_p^{y_i}\|_{F,\hat{\rho}}\}.$$

Finally, we compute the following upper estimates $\|Da\|_{\mathcal{B}} \leq c_{D_a}$, $\|D^2a\|_{\mathcal{B}} \leq c_{D^2a}$, $\|\Omega\|_{\mathcal{B}} \leq c_{\Omega}$, $\|D\Omega\|_{\mathcal{B}} \leq c_{D\Omega}$, $\|DF\|_{\mathcal{B}} \leq c_{DF}$, and $\|D^2F\|_{\mathcal{B}} \leq c_{D^2F}$, which appear in the statement of Theorem 2.6. We also compute the upper estimates $\|\Omega\|_{\hat{\mathcal{B}}} \leq \hat{c}_{\Omega}$, and $\|F_p\|_{\hat{\mathcal{B}}} \leq \hat{c}_{F_p}$.

5.3 Implementation details of Step 1

To evaluate the error $E(\theta)$ we use the expression

$$E(\theta) = \begin{pmatrix} K_p^x(\theta) + F_p^x(K(\theta)) - K_p^x(\theta + \varpi) - \varpi \\ F_p^y(K(\theta)) - K_p^y(\theta + \varpi) \end{pmatrix}. \quad (102)$$

To evaluate this formula at the grid points, we first compute $F_p \circ K$ and $K_p \circ R_{\varpi}$: the first term is computed directly from the grid, thus obtaining

$$\{(F_p \circ K)_j\} = \{F_p(\theta_j + K_{p,j}^x, K_{p,j}^y)\},$$

and the second term is computed in Fourier space, thus obtaining $\{\widetilde{(K_p \circ R_{\varpi})_k}\}$. Then we compute

$$\{(K_p \circ R_{\varpi})_j\} = \text{FFT}_B(\{\widetilde{(K_p \circ R_{\varpi})_k}\}).$$

From these expressions, the computation of Equation (102) at the grid, $\{E_j\}$, is straightforward. Then, we compute $\{\tilde{E}_k\} = \text{FFT}_F(\{E_j\})$.

Finally, using component-wise Theorem 3.3, we have

$$\begin{aligned} \|\tilde{E} - E\|_{\rho} &= \max\{\|\tilde{E}^x - E^x\|_{\rho}, \|\tilde{E}^y - E^y\|_{\rho}\} \leq C_{N_F}(\rho, \hat{\rho}) \max\{\|E^x\|_{\hat{\rho}}, \|E^y\|_{\hat{\rho}}\} \\ &\leq C_{N_F}(\rho, \hat{\rho}) \max\{\|F_p^x \circ K\|_{\hat{\rho}} + 2\|K_p^x\|_{\hat{\rho}} + |\varpi|, \|F_p^y \circ K\|_{\hat{\rho}} + \|K_p^y\|_{\hat{\rho}}\}, \\ &\leq C_{N_F}(\rho, \hat{\rho}) \max\{\hat{c}_{F_p} + 2\|K_p^x\|_{F,\hat{\rho}} + |\varpi|, \hat{c}_{F_p} + \|K_p^y\|_{F,\hat{\rho}}\}, \end{aligned}$$

thus obtaining

$$\|E\|_{\rho} \leq \|\tilde{E}\|_{F,\rho} + C_{N_F}(\rho, \hat{\rho}) \max\{\hat{c}_{F_p} + 2\|K_p^x\|_{F,\hat{\rho}} + |\varpi|, \hat{c}_{F_p} + \|K_p^y\|_{F,\hat{\rho}}\} =: b_E.$$

5.4 Implementation details of Step 2

To construct the frame $P(\theta)$ we represent the tangent vectors $DK(\theta)$ as

$$\{\widetilde{DK}_k\} = \left\{ \begin{pmatrix} I_n \\ O_n \end{pmatrix} \delta_{k,0} + \widetilde{DK}_{p,k} \right\}, \quad \{DK_j\} = \left\{ \begin{pmatrix} I_n \\ O_n \end{pmatrix} + DK_{p,j} \right\},$$

where $\delta_{k,0}$ is Kronecker's delta, and the computation of $\{\widetilde{DK}_{p,k}\}$ is performed in Fourier space. Finally compute $\{DK_{p,j}\} = \text{FFT}_B(\{\widetilde{DK}_{p,k}\})$.

We compute an approximation of the matrix $G(\theta) = -DK(\theta)^\top \Omega(K(\theta)) N_0(\theta)$ in a grid as

$$\{G_j\} = \{-DK_j^\top \Omega_j N_{0,j}\},$$

where $\Omega_j = \Omega((\theta_j, 0) + K_{p,j})$ and $N_{0,j} = N_0(\theta_j)$. We complement $DK(\theta)$ by computing $N(\theta)$ as

$$\{N_j\} = \{DK_j A_j + N_j^0 B_j\},$$

where

$$\{B_j\} = \{G_j^{-1}\} \quad \text{and} \quad \{A_j\} = \left\{ -\frac{1}{2} (B_j^\top N_{0,j}^\top \Omega_j N_{0,j} B_j) \right\}. \quad (103)$$

Then, we obtain $\{\tilde{N}_k\}$ using DFT, thus ending up with

$$\{\tilde{P}_k\} = \{(\widetilde{DK}_k \quad \tilde{N}_k)\}, \quad \{P_j\} = \{(DK_j \quad N_j)\}.$$

Let us observe that, since the parameterization is a truncated series, we have $DK(\theta) = \widetilde{DK}(\theta)$ and we set

$$\|DK\|_\rho \leq \|\widetilde{DK}\|_{F,\rho} =: b_{DK}, \quad \|DK^\top\|_\rho \leq \|\widetilde{DK}^\top\|_{F,\rho} =: b_{DK^\top}.$$

We also need to control $\|N_0\|_\rho \leq c_{N_0}$, $\|N_0^\top\|_\rho \leq c_{N_0^\top}$, $\|N_0^\top(\Omega \circ K)N_0\|_\rho \leq b_{N_0^\top(\Omega \circ K)N_0}$, $\|N_0\|_{\hat{\rho}} \leq \hat{c}_{N_0}$, and $\|N_0^\top\|_{\hat{\rho}} \leq \hat{c}_{N_0^\top}$. Since the selection of N_0 depends on the particular application at hand, we do not give here explicit details for the estimation of these objects. They follow using the same ideas that we discuss next to control the error of discrete Fourier approximations of the remaining objects. To use Corollary 3.8 we control the norm of $I_n - G(\theta)\tilde{B}(\theta)$ as follows

$$\begin{aligned} \|I_n - G\tilde{B}\|_\rho &\leq C_{N_F}(\rho, \hat{\rho}) \|G\|_{\hat{\rho}} \|\tilde{B}\|_{\hat{\rho}} \leq C_{N_F}(\rho, \hat{\rho}) \|\widetilde{DK}^\top\|_{\hat{\rho}} \|\Omega\|_{\hat{B}} \|N_0\|_{\hat{\rho}} \|\tilde{B}\|_{\hat{\rho}} \\ &\leq C_{N_F}(\rho, \hat{\rho}) \hat{c}_\Omega \hat{c}_{N_0} \|\widetilde{DK}^\top\|_{F,\hat{\rho}} \|\tilde{B}\|_{\hat{\rho}} =: t_B. \end{aligned}$$

Then, if $t_B < 1$ we use Corollary 3.9, thus obtaining

$$\|B\|_\rho \leq \|\tilde{B}\|_{F,\rho} + \frac{t_B \|\tilde{B}\|_{F,\hat{\rho}}}{1 - t_B} =: b_B. \quad (104)$$

Finally, we obtain direct estimates for $\|A\|_\rho$ and $\|N\|_\rho$. On the one hand, using Equation (104), we have

$$\|A\|_\rho \leq \frac{1}{2} \|B^\top\|_\rho \|N_0^\top(\Omega \circ K)N_0\|_\rho \|B\|_\rho \leq \frac{n}{2} b_{N_0^\top(\Omega \circ K)N_0} (b_B)^2 =: b_A, \quad (105)$$

and on the other hand, using Equations (104) and (105), we have

$$\|N\|_\rho \leq \|DK\|_\rho \|A\|_\rho + \|N_0\|_\rho \|B\|_\rho \leq b_{DK} b_A + c_{N_0} b_B =: b_N, \quad (106)$$

$$\|N^\top\|_\rho \leq \|A\|_\rho \|DK^\top\|_\rho + \|B^\top\|_\rho \|N_0^\top\|_\rho \leq b_A b_{DK^\top} + n b_B c_{N_0^\top} =: b_{N^\top}. \quad (107)$$

5.5 Implementation details of Step 3

To compute the torsion matrix $T(\theta)$, we first obtain the shifted normal frame $\{(\tilde{N} \circ R_{\varpi})_k\}$ and $\{(\tilde{N} \circ R_{\varpi})_j\} = \text{FFT}_B(\{(\tilde{N} \circ R_{\varpi})_k\})$. Then, we compute

$$\{T_j\} = \{(N \circ R_{\varpi})_j^\top (\Omega \circ K \circ R_{\varpi})_j (DF \circ K)_j N_j\},$$

and approximate the average $\langle T \rangle$ as follows

$$\tilde{T}_0 = \frac{1}{N_D} \sum_j T_j.$$

Then we compute

$$|\tilde{T}_0 - \langle T \rangle| \leq s_{N_F}^*(0, \rho) \|T\|_\rho \leq s_{N_F}^*(0, \rho) \|N^\top\|_\rho \|\Omega\|_B \|DF\|_B \|N\|_\rho \leq s_{N_F}^*(0, \rho) c_{\Omega} c_{DF} b_N b_{N^\top} := t_T$$

using the upper estimates in (106) and (107) for $\|N\|_\rho$ and $\|N^\top\|_\rho$, respectively. Finally, we check that

$$|\tilde{T}_0^{-1}| t_T < 1,$$

and we obtain (using a Neumann series argument)

$$|\langle T \rangle^{-1}| \leq \frac{|\tilde{T}_0^{-1}|}{1 - |\tilde{T}_0^{-1}| |\tilde{T}_0 - \langle T \rangle|} \leq \frac{|\tilde{T}_0^{-1}|}{1 - |\tilde{T}_0^{-1}| t_T} =: b_T.$$

5.6 Implementation details of Step 4

We notice that the choice in (99) introduces some suitable simplifications in the expression of Constants \mathfrak{C}_1 and \mathfrak{C}_2 . Indeed, after simple manipulations of the expressions described in Section 2.4 and using the Equations (99) and (101), we obtain

$$\mathfrak{C}_1 = \max \{2(a_3)^{\tau+1} \gamma^3 \rho^{3\tau-1} C_3, \mathfrak{C}_3, \mathfrak{C}_4, \mathfrak{C}_5\}, \quad \mathfrak{C}_2 := \frac{a_3^{2\tau} \hat{C}_2}{1 - a_1^{1-2\tau}},$$

where

$$\mathfrak{C}_3 = (a_1 a_3)^{4\tau} \hat{C}_5, \quad \mathfrak{C}_4 = \frac{\sigma_*(a_3)^{2\tau+1} \gamma^2 \rho^{2\tau-1} \hat{C}_2}{(\sigma-1)(1-a_1^{1-2\tau})}, \quad \mathfrak{C}_5 = \frac{(a_3)^{2\tau} \gamma^2 \rho^{2\tau} \hat{C}_2}{d_B(1-a_1^{-2\tau})}. \quad (108)$$

Finally, we provide expressions for the constants σ_* , \hat{C}_2 , and \hat{C}_5 in terms of the initial data. The first one is given by

$$\sigma_* = \max \left\{ \frac{n}{b_{DK}}, \frac{2n}{b_{DK^\top}}, \frac{c_{N_0^\top} c_{D\Omega} c_{N_0} \delta}{b_{N_0^\top(\Omega \circ K)N_0}}, \frac{\beta_1}{b_B}, \frac{\beta_5}{b_T} \right\} \quad (109)$$

where

$$\begin{aligned} \beta_1 &= 2\sigma_B^2 c_{N_0} (\sigma_{DK^\top} c_{D\Omega} \delta + 2nc_\Omega), \\ \beta_2 &= \frac{n}{2} (\sigma_B)^2 c_{D\Omega} \delta + \frac{n+1}{2} \sigma_{N_0^\top(\Omega \circ K)N_0} \beta_1, \\ \beta_3 &= \sigma_{DK} \beta_2 + nc_A + c_{N_0} \beta_1, \\ \beta_4 &= \sigma_{DK^\top} \beta_2 + 2nc_A + nc_{N_0^\top} \beta_1, \\ \beta_5 &= 2(\sigma_T)^2 (c_{N^\top} c_N (c_\Omega c_{D^2F} + c_{D\Omega} c_{DF}) \delta + c_\Omega c_{DF} (c_{N^\top} \beta_3 + c_N \beta_4)) \end{aligned}$$

and

$$c_A = \frac{1}{2} n \sigma_{N_0^\top(\Omega \circ K)N_0} (\sigma_B)^2, \quad c_N = \sigma_{DK} c_A + c_{N_0} \sigma_B, \quad c_{N^\top} = c_A \sigma_{DK^\top} + n \sigma_B c_{N_0^\top}.$$

The constants \hat{C}_2 and \hat{C}_5 follow from the next sequence of computations:

$$\begin{aligned}
c_P &= \sigma_{DK} + c_N, \\
c_T &= c_N^\top c_\Omega c_{DF} c_N, \\
C_1 &= \sigma_{DK}^\top \sigma_{DK} c_{D\Omega} \delta + n \sigma_{DK}^\top c_\Omega + 2n c_\Omega c_{DF} \sigma_{DK}, \\
C_2 &= c_R C_1, \\
C_3 &= (1 + c_A) \max(1, c_A) C_2, \\
C_4 &= n c_N^\top c_\Omega \gamma \delta^\tau + c_A C_2, \\
C_5 &= C_2 + n \sigma_{DK}^\top c_\Omega \gamma \delta^\tau, \\
C_6 &= c_A C_2 + \sigma_{DK}^\top c_{D\Omega} c_{DF} c_N \gamma \delta^{\tau+1} + 2n c_\Omega c_{DF} c_N \gamma \delta^\tau, \\
C_7 &= \max(C_4, C_5 + C_6), \\
C_8 &= 2c_R \sigma_{DK}^\top c_\Omega, \\
C_9 &= C_8 + \sigma_T (c_N^\top c_\Omega \gamma \delta^\tau + c_T C_8), \\
C_{10} &= c_R (c_N^\top c_\Omega \gamma \delta^\tau + c_T C_9), \\
\hat{C}_2 &= \sigma_{DK} C_{10} + c_N C_9 \gamma \delta^\tau, \\
C_{15} &= (C_3 + C_7) \max(C_9 \gamma \delta^\tau, C_{10}) + 2n c_{D_a} \gamma^3 \delta^{3\tau} + \frac{1}{2} c_{D^2_a} \gamma^3 \delta^{3\tau+1}, \\
\hat{C}_5 &= 2c_P C_{15} \gamma \delta^{\tau-1} + \frac{1}{2} c_{D^2_F} \hat{C}_2^2,
\end{aligned}$$

where the value c_R is computed using Equation (90). We recall that if we take the value $L = 0$ in (90) we obtain the classic expression in (96).

Remark 5.4. *It is worth mentioning that the above expressions are very general. Using specific information from a particular problems, it is possible to improve some estimates. For example, if $n = 1$ then the phase space is 2-dimensional and every 1-dimensional subspace is Lagrangian. This has the immediate consequence that we can take $C_1 = 0$, $b_A = c_A = 0$ and $b_{N_0^\top(\Omega \circ K)N_0} = 0$, thus simplifying the computations and reducing the size of the subsequent constants. Indeed, in this case the third element in (109) is removed. For the same reason, if $n = 1$ we do not have to control $\|\bar{A} - A\|_{\rho-3\delta}$ so we can take $C_{12} = 0$, and so, $\beta_2 = 0$. Moreover, we can take advantage of the specific expression of the map F and the matrix N_0 in order to control the norm of the twist matrix T . More precisely, we can replace the estimate $\|T\|_\rho \leq c_T = c_N^\top c_\Omega c_{DF} c_N$ by an ad hoc estimation for the considered problem.*

6 Application of the KAM theorem in some examples

In this section we apply the techniques described in the paper to prove existence of invariant tori in different scenarios. A common feature in the selected examples is that the objects F , Ω , and a are explicit. As a consequence we directly obtain global estimates for $\|DF\|_B$, $\|D^2F\|_B$, $\|\Omega\|_B$, $\|D\Omega\|_B$, $\|Da\|_B$, and $\|D^2a\|_B$. This allows us to focus in the fundamental steps involved in the computer assisted proof (CAP). If the symplectic map is defined by means of an implicit equation or it is given by the discretization of the flow of a Hamiltonian vector field, then the control of global estimates may deserve a particular attention and can become a very difficult problem (see the discussion in Section 7).

In order to avoid repeated explanations in the description of the different examples, we summarize next some general details regarding the input of Algorithm 5.1 that we use to perform the CAPs.

- **Input 1.** The numerical values of the sampling $\{K_{p,j}\}$, with $K_{p,j} = K_0(\theta_j) - (\theta_j, 0)$ in a regular grid of size $N_F = (N_{F,1}, \dots, N_{F,n})$, are provided by means of a data file. These numbers are read rounding to the nearest representable number. These numbers are used to construct the parameterization that will be validated (Step 0 of the Algorithm 5.1).

- **Input 2.** The transversal map N_0 is selected according to the particular problem at hand. For example, if we look for an invariant curve and we know that it is a graph, then we can introduce a constant vector complementing the tangent vector of the curve at every point. We also can use the geometric structure of the problem to obtain the transversal map. The evaluation rule for the map N_0 is provided by means of a separate subroutine. The reader is referred to Chapter 4 of [37] for a detailed discussion that summarizes different approaches in the literature.
- **Input 3.** When the frequency vector is defined by means of an algebraic equation then ϖ is obtained by enclosing the solution of this equation. When the frequency vector has been obtained in a numerical computation, then the enclosing interval vector ϖ is selected according with the precision of the computation. If the Diophantine constants of a target frequency are known, they can be provided (e.g. for the golden mean). Otherwise, we obtain a pair of constants $\tau > n$ and $\gamma > 0$ using the methodology derived from Proposition 4.1.
- **Input 4.** For every parameterization given in Input 1, we need suitable constants $\delta, \sigma, \rho, \hat{\rho}, \rho_\infty$, and d_B satisfying $0 < \delta < \rho/3$, $0 < \rho < \hat{\rho}$, $1 < \sigma$, $0 \leq \rho_\infty < \rho$, and $d_B > 0$. These constants are obtained from the heuristic (non-rigorous) methodology described in Appendix A, and are provided by a data file. They are read rounding to the nearest representable number.

The sampling $\{K_{p,j}\}$ is obtained numerically using the implementation of the parameterization method proposed in Chapter 4 of [37]. Analogous implementations of the method in different contexts have been previously presented in the literature adapted to several problems [8, 9, 10, 21, 31, 38, 41]. We want to remark that the validation algorithm admits complete freedom in using a different method to obtain the approximate parameterization. Indeed, the numerical computation of invariant tori has been a fruitful area of research in the last years and there is a wide set of numerical methods available. The parameterization method is a suitable choice for several reasons: it has cost $O(N_D)$ in storage and only $O(N_D \log(N_D))$ in time, where $N_D = N_{F,1} \cdots N_{F,n}$; it does not use a perturbative setting of the problem so we can consider any invariant torus in phase space, without performing any perturbative analysis of the parameters of the problem; the numerical algorithm of the parameterization method has a similar structure than Algorithm 5.1 so we can take advantage of the same codes using a suitable overloading of arithmetics. Finally, we observe that the bottleneck of Algorithm 5.1 is the error of invariance of the parameterization defined by the sampling $\{K_{p,j}\}$ (see details in Section 5.2). To this end, the fact that the convergence of the parameterization method is quadratic and that the iterations are fast allows us to obtain approximations of invariant tori with very high accuracy.

6.1 Standard map

We first consider the study of quasi-periodic invariant curves of the so-called Chirikov standard map [18]

$$\begin{aligned} F : \mathbb{T} \times \mathbb{R} &\longrightarrow \mathbb{T} \times \mathbb{R} \\ (x, y) &\longmapsto (\bar{x}, \bar{y}) = \left(x + \bar{y}, y - \frac{\varepsilon}{2\pi} \sin(2\pi x)\right) \end{aligned} \quad (110)$$

For $\varepsilon = 0$ the dynamics is very simple: the orbit of any point $(x, y) \in \mathbb{T} \times \mathbb{R}$ is given by the rigid rotation $F^n(x, y) = (x + ny, y)$. Note that if $y = p/q \in \mathbb{Q}$ the corresponding orbit is periodic, i.e., $F^q(x, y) = (x + p, y) = (x, y)$. On the contrary, if $y \in \mathbb{R} \setminus \mathbb{Q}$, the orbit is dense in the invariant curve $\mathbb{T} \times \{y\}$. In any case, the orbit of a given point (x, y) has rotation number $\omega = y$ for every $x \in \mathbb{T}$.

For $\varepsilon > 0$, sufficiently small, KAM theory concludes that “most” of the previous invariant curves persist, although they are slightly deformed. The deformation preserves the homotopy class of these rotational invariant curves, also called *primary tori*. These curves are successively destroyed as ε is increased. The value of ε for which an invariant curve is destroyed is called critical value $\varepsilon_c = \varepsilon_c(\omega)$. A particularly interesting case is the golden rotation $\omega = \frac{\sqrt{5}-1}{2}$.

For the standard map, given by Equation (110), we have $\mathcal{A} = \mathbb{T} \times \mathbb{R}$, $\alpha = \alpha_0 = ydx$ and $\omega = \omega_0 = dy \wedge dx$. Hence, we take $c_\Omega = 1$, $c_{D\Omega} = 0$, $c_{D\alpha} = 1$ and $c_{D^2\alpha} = 0$. We select the transversal vector $N_0 : \mathbb{T} \rightarrow \mathbb{R}^{2 \times 1}$ as

$$N_0(\theta) = N_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

| ε | N_F | ρ | δ | $\sigma - 1$ | d_B | $\hat{\rho}$ | $\frac{\mathfrak{C}_1 b_E}{\gamma^4 \rho^{4\tau}}$ | $\frac{\mathfrak{C}_2 b_E}{\gamma^2 \rho^{2\tau}}$ |
|---------------|-------|--------------|--------------|--------------|--------------|--------------|--|--|
| 0.06 | 128 | 1.606160e-02 | 3.212319e-03 | 1.670325e-01 | 5.064098e-06 | 2.569855e-01 | 1.35e-28 | 9.47e-34 |
| 0.16 | 256 | 1.369960e-02 | 2.739919e-03 | 9.673976e-02 | 2.937365e-06 | 1.369960e-01 | 9.24e-28 | 3.77e-33 |
| 0.26 | 256 | 1.369960e-02 | 2.739919e-03 | 6.974093e-02 | 2.044422e-06 | 1.301462e-01 | 1.74e-26 | 4.94e-32 |
| 0.36 | 512 | 1.369960e-02 | 2.739919e-03 | 5.229422e-02 | 1.400906e-06 | 7.534778e-02 | 4.24e-25 | 8.26e-31 |
| 0.46 | 512 | 1.369960e-02 | 2.739919e-03 | 3.941981e-02 | 9.278480e-07 | 7.534778e-02 | 1.76e-23 | 2.27e-29 |
| 0.56 | 512 | 4.520867e-03 | 8.908112e-04 | 1.268703e-02 | 9.401294e-08 | 6.329214e-02 | 9.39e-24 | 1.24e-30 |
| 0.66 | 1024 | 3.300233e-03 | 5.973272e-04 | 1.047736e-02 | 4.061043e-08 | 3.300233e-02 | 1.88e-23 | 1.11e-30 |
| 0.76 | 1024 | 2.310163e-03 | 4.017675e-04 | 5.924431e-03 | 1.166394e-08 | 3.003212e-02 | 2.32e-18 | 3.98e-26 |
| 0.86 | 2048 | 1.178183e-03 | 1.996921e-04 | 1.921375e-03 | 1.234843e-09 | 1.531638e-02 | 1.74e-17 | 3.19e-26 |
| 0.96 | 32768 | 1.178183e-04 | 1.971855e-05 | 3.648874e-05 | 5.996316e-13 | 1.060365e-03 | 2.34e-12 | 2.09e-24 |

Table 5: Application of Theorem 2.6 using Algorithm 5.1 for the golden invariant curve of the standard map (110) for different values of ε . We use $\gamma = \frac{3-\sqrt{5}}{2}$, $\tau = 1$ and the *ad hoc* Rüssmann estimates in (90). The result (last two columns) is given with 2 significant digits.

In Step 0 of Algorithm 5.1, we introduce the domain \mathcal{B} as

$$\mathcal{B} = \{(x, y) \in \mathbb{C}/\mathbb{Z} \times \mathbb{C} : |\operatorname{Im} x| \leq r_1, |y| \leq r_2\},$$

and the domain $\hat{\mathcal{B}}$ as

$$\hat{\mathcal{B}} = \{(x, y) \in \mathbb{C}/\mathbb{Z} \times \mathbb{C} : |\operatorname{Im} x| \leq \hat{r}_1, |y| \leq \hat{r}_2\},$$

where

$$r_1 = d_B + \rho + \|K_p^x\|_{F,\rho}, \quad r_2 = d_B + \|K_p^y\|_{F,\rho}, \quad \hat{r}_1 = \hat{\rho} + \|K_p^x\|_{F,\hat{\rho}}, \quad \hat{r}_2 = \|K_p^y\|_{F,\hat{\rho}}.$$

Then, global estimates of the symplectic map (110) in these domains are characterized as follows

$$\begin{aligned} \|DF\|_{\mathcal{B}} &\leq c_{DF} := 2 + \varepsilon \cosh(2\pi r_1), \\ \|D^2F\|_{\mathcal{B}} &\leq c_{D^2F} := 2\pi\varepsilon \cosh(2\pi r_1), \\ \|F_p\|_{\hat{\mathcal{B}}} &\leq \hat{c}_{F_p} := \hat{r}_2 + \frac{\varepsilon}{2\pi} \cosh(2\pi \hat{r}_1). \end{aligned}$$

Let us characterize the constants that appear in Step 2 of Algorithm 5.1. Since N_0 is constant, we take $c_{N_0} = 1$, $c_{N_0^\top} = 1$, $\hat{c}_{N_0} = 1$, and $\hat{c}_{N_0^\top} = 1$. Moreover, we observe that $N_0^\top(\Omega \circ K)N_0$ vanishes identically, so we take $c_{N_0^\top(\Omega \circ K)N_0} = 0$ (see Remark 5.4).

In Table 5 we show the fundamental information in the computer assisted proof for the existence of the golden invariant curve for different values of ε . The second column is the number of points N_F in the regular grid where the sampling $\{K_{p,j}\}$ is defined. As it is mentioned in the introduction of this section, this sampling is obtained via the parameterization method asking for a tolerance of 10^{-33} in the error of invariance (using the norm $\|E\|_{F,0}$). For each value of ε and the corresponding sampling, we take the values ρ , δ , σ , d_B and $\hat{\rho}$ that are given in columns 3 to 7. These values have been obtained using the heuristic methodology described in Appendix A. In all computations we take $a_2 = 1000$ so that $\rho_\infty = \rho/1000$. We use the specific Diophantine constants $\gamma = \frac{3-\sqrt{5}}{2}$ and $\tau = 1$ of the golden mean. After applying the rigorous computations described in Algorithm 5.1, using 267 bits, in the last two columns we provide the values of the left-hand side of the KAM condition and the bound of the correction of the true invariant tori, which are given by

$$\frac{\mathfrak{C}_1 b_E}{\gamma^4 \rho^{4\tau}}, \quad \|K_\infty - K\|_{\rho_\infty} < \frac{\mathfrak{C}_2 b_E}{\gamma^2 \rho^{2\tau}},$$

respectively. It is worth mentioning that the results presented in Table 5 are non-perturbative, in the sense that the computations are performed independently of each value of ε , without using any information related to smaller values of the parameter. The computational time of the CAP for the case $\varepsilon = 0.96$ is 117 seconds in a single processor Intel(R) Xeon(R) CPU at 2.40 GHz.

The application of the KAM theorem becomes computationally more demanding as we approach the critical value ε_c . Indeed, from $\varepsilon = 0.02$ to $\varepsilon = 0.96$ the KAM condition has worsened by 16 orders of magnitude. From

this point, the number of Fourier coefficients required to apply the theorem increase dramatically (exponentially). In this situation, the Rüssmann estimates proposed in Section 4.2 play a significant role in improving the applicability of the KAM theorem. For example, if we repeat the CAP for $\varepsilon = 0.96$ using the classic Rüssmann estimates in (96) we obtain

$$\frac{\mathfrak{C}_1 b_E}{\gamma^4 \rho^{4\tau}} \leq 5.42 \cdot 10^{-6}, \quad \frac{\mathfrak{C}_2 b_E}{\gamma^2 \rho^{2\tau}} \leq 1.71 \cdot 10^{-21}.$$

In this case, expression $\mathfrak{C}_1 b_E \gamma^{-4} \rho^{-4\tau}$ is 6 orders of greater than the value obtained using the *ad hoc* estimates in (90). The difference increases when we approach to the critical value. The last value of ε for which we have applied the KAM theorem is the following:

Theorem 6.1. *For $\varepsilon = 0.9716$ the standard map has a rotational invariant curve with golden rotation number.*

Proof. We consider a parameterization K obtained using the parameterization method with $N_F = 8388608$ Fourier coefficients (we show the significant ones in Figure 1). This parameterization satisfies $\|E\|_{F,0} \leq 2.74 \cdot 10^{-41}$. Again, we take the Diophantine constants $\gamma = \frac{3-\sqrt{5}}{2}$ and $\tau = 1$, and we use the improved Rüssmann constant in (90). Setting the parameters

$$\begin{aligned} \rho &= 3.748290 \cdot 10^{-7}, \\ \delta &= 6.273289 \cdot 10^{-8}, \\ \sigma &= 1 + 1.610158 \cdot 10^{-9}, \\ d_B &= 3.159428 \cdot 10^{-21}, \\ \hat{\rho} &= 4.872777 \cdot 10^{-6}, \end{aligned}$$

and applying Algorithm 5.1 with precision of 367 bits we obtain the following rigorous bound:

$$\frac{\mathfrak{C}_1 b_E}{\gamma^4 \rho^{4\tau}} \leq 0.0823.$$

This allows us to apply the KAM theorem. Moreover, the golden curve satisfies

$$\|K_\infty - K\|_{\rho_\infty} < \frac{\mathfrak{C}_2 b_E}{\gamma^2 \rho^{2\tau}} \leq 3.89 \cdot 10^{-22}.$$

The computational time of this CAP is 11404 seconds in a single processor Intel(R) Core(R) CPU at 3.50 GHz. It required the use of almost 32 GB of RAM. \square

Now we illustrate the methodology proposed in Section 4.1 to obtain a pair of constants (γ, τ) for a given interval vector. For example, we consider the interval ϖ obtained by computing $\frac{\sqrt{5}-1}{2}$ using MPFI with 267 bits and taking $M = 1000$. In Table 6 we repeat the CAPs given in Table 5 using the obtained constants $\gamma = 0.38196601125010$ and $\tau = 1.26$. Using the argument in Remark 5.2 we obtain a rigorous upper bound of the relative measure of the set of frequencies in ϖ for which the KAM theorem does not apply.

In Table 7 we present the application of the KAM theorem for other invariant curves. Specifically, we consider the rotation numbers that have been characterized in Table 1. We fix $N_F = 32768$ and we show the maximum number of ε for which we have been able to apply the KAM theorem (taking jumps of length 0.01 in ε). Computations are performed using interval arithmetics with 267 bits. We perform the CAP using both the classic Rüssmann estimates in (96) and the *ad hoc* estimates in (90). Numerical approximations of the critical values of some of these curves have been reported in [31], for example, $\varepsilon_c \simeq 0.957447$ for $\omega_{2,2}$ (we prove existence for $\varepsilon = 0.95$), $\varepsilon_c \simeq 0.87608$ for $\omega_{1,2}$ (we prove existence for $\varepsilon = 0.87$), $\varepsilon_c \simeq 0.89086$ for $\omega_{3,3}$ (we prove existence for $\varepsilon = 0.88$), and $\varepsilon_c \simeq 0.77242$ for $\omega_{1,3}$ (we prove existence for 0.76). As in Theorem 6.1, we can obtain sharper rigorous lower bounds of these critical values by increasing the number of Fourier coefficients. We remark again that the use of *ad hoc* Rüssmann estimates represents a significant advantage in order to apply the KAM theorem. In some cases we improve up to 17 orders of magnitude the size of the smallness condition. The computational time of these CAPs ranges between 75 and 123 seconds in a single processor Intel(R) Xeon(R) CPU at 2.40 GHz.

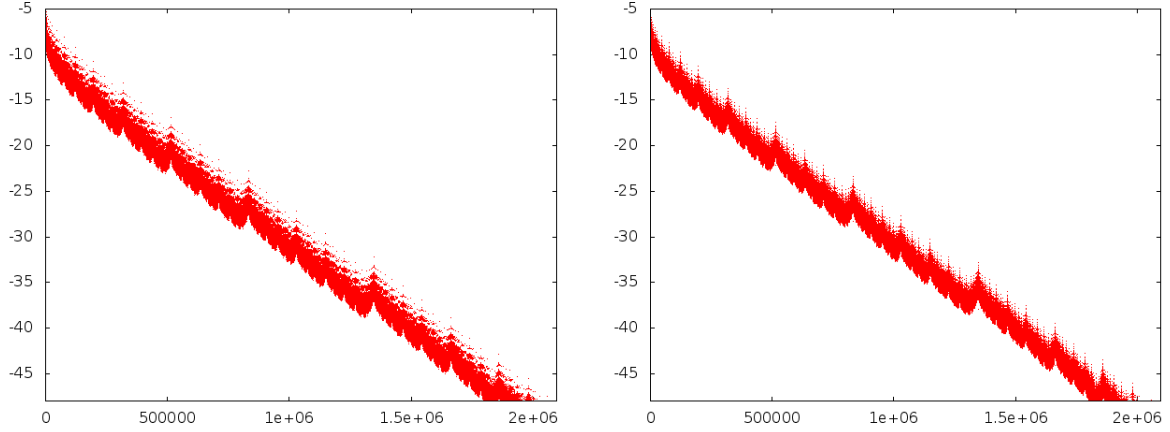


Figure 1: Fourier coefficients (for positive k) of the validated parameterization of the golden invariant curve of the standard map (110) for $\varepsilon = 0.9716$. In the left plot we show $k \mapsto \log_{10}(|K_{p,k}^x|)$ and in the right plot we show $k \mapsto \log_{10}(|K_{p,k}^y|)$.

| ε | $\frac{c_1 b_E}{\gamma^4 \rho^{4\tau}}$ | $\frac{c_2 b_E}{\gamma^2 \rho^{2\tau}}$ | $1 - p(\varpi, \gamma_0, \tau)$ |
|---------------|---|---|---------------------------------|
| 0.06 | 3.50e-28 | 7.56e-34 | 1.33e-07 |
| 0.16 | 2.40e-27 | 3.01e-33 | 2.16e-07 |
| 0.26 | 4.51e-26 | 3.94e-32 | 4.50e-07 |
| 0.36 | 1.10e-24 | 6.60e-31 | 9.99e-07 |
| 0.46 | 4.56e-23 | 1.81e-29 | 2.53e-06 |
| 0.56 | 2.39e-23 | 9.83e-31 | 2.16e-06 |
| 0.66 | 4.25e-23 | 8.64e-31 | 2.49e-06 |
| 0.76 | 5.00e-18 | 3.08e-26 | 4.61e-05 |
| 0.86 | 3.64e-17 | 2.46e-26 | 7.58e-05 |
| 0.96 | 4.84e-12 | 1.61e-24 | 1.45e-03 |

Table 6: Application of Theorem 2.6 using Algorithm 5.1 around the golden invariant curve of the standard map (110) for different values of ε . We use the same implementation parameters as in Table 5. We take the constants $\gamma = 0.38196601125010$, $\tau = 1.26$ and use the *ad hoc* Rüssmann estimates in (90). A rigorous upper bound of the relative measure of the set of frequencies in ϖ for which the KAM theorem does not apply (we use expression (100)). The results are given with 2 significant digits.

6.2 Non-twist standard map

The second application falls in the context of the so-called non-twist maps. It is well known that there is an analogue of KAM theory in the non-twist scenario (see for example [25, 34, 72]). The loss of the twist condition introduces different properties with respect to the twist case, for example the fact that the Birkhoff Graph Theorem does not apply and folded invariant curves are observed. A classic mechanism that creates such folded invariant curves is called *reconnection* (see [24, 72]). Reconnection is a global bifurcation of the invariant manifolds of two or more distinct hyperbolic periodic orbits having the same winding number. This creates a meandering region having folded quasi periodic curves. Among these orbits, of special interest is the one that corresponds to an invariant curve having a local extremum in the rotation number, called shearless invariant curve. In this section we are not interested in applying singular KAM theory to study the shearless invariant curves (in the spirit of [34]) but we will consider invariant curves inside the meandering region that are non-degenerate (in the spirit of [72]). The study of shearless invariant curves and their bifurcations can be performed in combination with the tools in [34].

Let us consider the non-twist standard map

$$\begin{aligned}
 F : \mathbb{T} \times \mathbb{R} &\longrightarrow \mathbb{T} \times \mathbb{R} \\
 (x, y) &\longmapsto (\bar{x}, \bar{y}) = \left(x + (\bar{y} + \lambda_1)(\bar{y} + \lambda_2), y - \frac{\varepsilon}{2\pi} \sin(2\pi x) \right)
 \end{aligned} \tag{111}$$

where $(x, y) \in \mathbb{T} \times \mathbb{R}$ are phase space coordinates and $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\varepsilon > 0$ are parameters. Although this family is not generic (it contains just one harmonic), it describes the essential features of non-twist systems with a local

| a | b | ε | ρ | δ | $\sigma - 1$ | $d_{\mathcal{B}}$ | $\hat{\rho}$ | $\frac{c_1 b_F}{\gamma^4 \rho^{4\tau}}$ | $\frac{c_2 b_F}{\gamma^2 \rho^{2\tau}}$ |
|-----|-----|---------------|--------------|--------------|--------------|-------------------|--------------|---|---|
| 1 | 2 | 0.87 | 9.226561e-05 | 1.544194e-05 | 5.101229e-06 | 4.377668e-14 | 9.687889e-04 | 1.71e-07 | 9.11e-21 |
| 1 | 3 | 0.76 | 1.919125e-04 | 3.211924e-05 | 1.149569e-05 | 1.945721e-13 | 2.111037e-03 | 1.57e-11 | 3.85e-24 |
| 1 | 4 | 0.67 | 2.140562e-04 | 3.582531e-05 | 7.808704e-06 | 1.020386e-13 | 2.140562e-03 | 7.08e-11 | 9.33e-24 |
| 1 | 5 | 0.60 | 3.321562e-04 | 5.559099e-05 | 5.647400e-06 | 6.322830e-14 | 2.159015e-03 | 3.38e-08 | 2.75e-21 |
| 1 | 6 | 0.54 | 1.992937e-04 | 3.335460e-05 | 2.297353e-06 | 1.249607e-14 | 2.092584e-03 | 7.49e-09 | 1.20e-22 |
| 2 | 1 | 0.93 | 1.845312e-04 | 3.088388e-05 | 1.340147e-05 | 3.284838e-13 | 1.107187e-03 | 2.90e-10 | 1.17e-22 |
| 2 | 2 | 0.95 | 1.033375e-04 | 1.729498e-05 | 7.353570e-06 | 1.004165e-13 | 1.033375e-03 | 1.70e-10 | 2.14e-23 |
| 2 | 3 | 0.91 | 1.919125e-04 | 3.211924e-05 | 1.297609e-05 | 3.508463e-13 | 1.151475e-03 | 4.39e-10 | 1.97e-22 |
| 2 | 4 | 0.86 | 1.771500e-04 | 2.964853e-05 | 1.090351e-05 | 2.725768e-13 | 2.037225e-03 | 6.87e-11 | 2.42e-23 |
| 2 | 5 | 0.82 | 1.107187e-04 | 1.853033e-05 | 2.959595e-06 | 2.577350e-14 | 1.051828e-03 | 1.85e-09 | 6.21e-23 |
| 2 | 6 | 0.78 | 9.595623e-05 | 1.605962e-05 | 1.243173e-06 | 6.667071e-15 | 1.055519e-03 | 2.51e-08 | 2.19e-22 |
| 3 | 1 | 0.83 | 2.066750e-04 | 3.458995e-05 | 1.428538e-05 | 3.464827e-13 | 2.066750e-03 | 6.95e-12 | 3.02e-24 |
| 3 | 2 | 0.89 | 1.144094e-04 | 1.914801e-05 | 8.411203e-06 | 1.146242e-13 | 1.029684e-03 | 1.45e-10 | 2.05e-23 |
| 3 | 3 | 0.88 | 1.771500e-04 | 2.964853e-05 | 1.476158e-05 | 3.809012e-13 | 1.151475e-03 | 3.09e-11 | 1.47e-23 |
| 3 | 4 | 0.86 | 1.144094e-04 | 1.914801e-05 | 5.127246e-06 | 6.122894e-14 | 1.029684e-03 | 4.37e-10 | 3.43e-23 |
| 3 | 5 | 0.83 | 1.771500e-04 | 2.964853e-05 | 1.026994e-05 | 1.831959e-13 | 1.860075e-03 | 1.66e-07 | 3.95e-20 |
| 3 | 6 | 0.80 | 1.992937e-04 | 3.335460e-05 | 1.067795e-05 | 2.109941e-13 | 2.092584e-03 | 1.14e-11 | 3.14e-24 |
| 4 | 1 | 0.74 | 1.771500e-04 | 2.964853e-05 | 4.904504e-06 | 4.722540e-14 | 1.948650e-03 | 4.97e-07 | 2.88e-20 |
| 4 | 2 | 0.80 | 1.771500e-04 | 2.964853e-05 | 1.002141e-05 | 1.894322e-13 | 1.860075e-03 | 2.32e-07 | 5.61e-20 |
| 4 | 3 | 0.81 | 9.226561e-05 | 1.544194e-05 | 4.429583e-06 | 3.382568e-14 | 1.014922e-03 | 1.74e-09 | 7.33e-23 |
| 4 | 4 | 0.79 | 2.140562e-04 | 3.582531e-05 | 1.653331e-05 | 4.949205e-13 | 2.140562e-03 | 2.19e-12 | 1.37e-24 |
| 4 | 5 | 0.78 | 1.291718e-04 | 2.161872e-05 | 5.383696e-06 | 5.697885e-14 | 1.097961e-03 | 5.25e-10 | 3.79e-23 |
| 4 | 6 | 0.76 | 1.771500e-04 | 2.964853e-05 | 6.388444e-06 | 9.570476e-14 | 1.948650e-03 | 6.39e-08 | 7.93e-21 |
| 5 | 1 | 0.66 | 1.328625e-04 | 2.223640e-05 | 2.056577e-06 | 8.775943e-15 | 1.062900e-03 | 2.64e-08 | 3.03e-22 |
| 5 | 2 | 0.72 | 1.144094e-04 | 1.914801e-05 | 3.830449e-06 | 2.609694e-14 | 1.029684e-03 | 1.88e-09 | 6.41e-23 |
| 5 | 3 | 0.73 | 1.033375e-04 | 1.729498e-05 | 4.269049e-06 | 3.062480e-14 | 1.033375e-03 | 2.50e-09 | 9.36e-23 |
| 5 | 4 | 0.72 | 2.066750e-04 | 3.458995e-05 | 1.094161e-05 | 2.103895e-13 | 2.066750e-03 | 8.81e-11 | 2.40e-23 |
| 5 | 5 | 0.71 | 1.919125e-04 | 3.211924e-05 | 8.565614e-06 | 1.320616e-13 | 1.919125e-03 | 1.51e-08 | 2.52e-21 |
| 5 | 6 | 0.70 | 1.439343e-04 | 2.408943e-05 | 3.888326e-06 | 3.451748e-14 | 1.079508e-03 | 1.38e-09 | 5.90e-23 |
| 6 | 1 | 0.59 | 1.771500e-04 | 2.964853e-05 | 2.193874e-06 | 1.019455e-14 | 2.037225e-03 | 2.24e-08 | 2.75e-22 |
| 6 | 2 | 0.65 | 8.488436e-05 | 1.420659e-05 | 1.660222e-06 | 4.346395e-15 | 9.337279e-04 | 1.24e-04 | 7.09e-19 |
| 6 | 3 | 0.66 | 8.857498e-05 | 1.482426e-05 | 1.695908e-06 | 6.855800e-15 | 9.743248e-04 | 2.46e-06 | 2.03e-20 |
| 6 | 4 | 0.65 | 2.140562e-04 | 3.582531e-05 | 1.192707e-05 | 2.027028e-13 | 2.140562e-03 | 2.16e-11 | 5.59e-24 |
| 6 | 5 | 0.65 | 9.226561e-05 | 1.544194e-05 | 1.994275e-06 | 7.521929e-15 | 1.014922e-03 | 3.06e-08 | 2.97e-22 |
| 6 | 6 | 0.64 | 1.181000e-04 | 1.976569e-05 | 2.045468e-06 | 1.035221e-14 | 1.062900e-03 | 1.33e-08 | 1.76e-22 |

Table 7: Application of Theorem 2.6 using Algorithm 5.1 for the standard map (110). We consider several interval frequencies ϖ enclosing intervals given by Equation (87). For each curve, we present the larger value of ε (second column) for which we can apply the KAM theorem with *ad hoc* Rüssmann estimates (90) using a grid of size $N_F = 32768$.

quadratic extremum in the rotation number.

For the non-twist standard map, given by Equations (111), we have $\mathcal{A} = \mathbb{T} \times \mathbb{R}$, $\alpha = \alpha_0 = ydx$ and $\omega = \omega_0 = dy \wedge dx$. Hence, we take $c_\Omega = 1$, $c_{D\Omega} = 0$, $c_{D\alpha} = 1$ and $c_{D^2\alpha} = 0$.

Contrary to the example discussed in Section 6.1, we notice that the tangent vectors to the invariant curves are not in general in the horizontal direction, since invariant curves are not expected to be graphs. Using the ambient structure, we select the transversal vectors $N_0 : \mathbb{T} \rightarrow \mathbb{R}^{2 \times 1}$ as

$$N_0(\theta) = \Omega_0 DK(\theta), \quad (112)$$

and it is clear that $DK(\theta)$ and $N_0(\theta)$ form a basis of \mathbb{R}^2 for every $\theta \in \mathbb{T}$.

In Step 0 of Algorithm 5.1, we introduce the domain \mathcal{B} as

$$\mathcal{B} = \{(x, y) \in \mathbb{C}/\mathbb{Z} \times \mathbb{C} : |\operatorname{Im} x| \leq r_1, |y| \leq r_2\},$$

and the domain $\hat{\mathcal{B}}$ as

$$\hat{\mathcal{B}} = \{(x, y) \in \mathbb{C}/\mathbb{Z} \times \mathbb{C} : |\operatorname{Im} x| \leq \hat{r}_1, |y| \leq \hat{r}_2\}.$$

where

$$r_1 = d_{\mathcal{B}} + \rho + \|K_p^x\|_{F,\rho}, \quad r_2 = d_{\mathcal{B}} + \|K_p^y\|_{F,\rho}, \quad \hat{r}_1 = \hat{\rho} + \|K_p^x\|_{F,\hat{\rho}}, \quad \hat{r}_2 = \|K_p^y\|_{F,\hat{\rho}}.$$

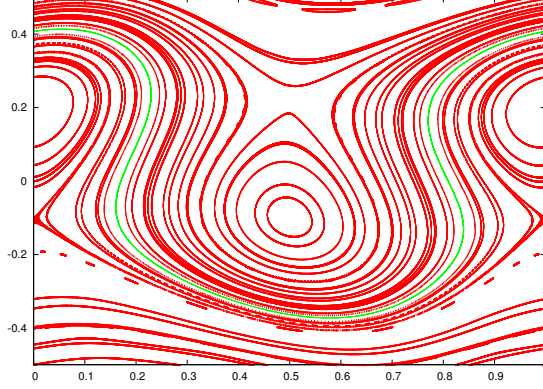


Figure 2: Meandering invariant curves for the non-twist standard map (111) for parameters $\lambda_1 = 0.1$, $\lambda_2 = -0.2$ and $\varepsilon = 0.45$.

Then, global estimates of the symplectic map (111) in these domains are characterized as follows

$$\begin{aligned} \|DF\|_{\mathcal{B}} &\leq c_{DF} := \max \left\{ 1 + |\partial_x \bar{y}|, 1 + (2|\bar{y}| + |\lambda_1 + \lambda_2|)(1 + |\partial_x \bar{y}|) \right\}, \\ \|D^2F\|_{\mathcal{B}} &\leq c_{D^2F} := \max \left\{ |\partial_{xx} \bar{y}|, 2 + 2|\partial_x \bar{y}| + 2|\partial_x \bar{y}|^2 + (2|\bar{y}| + |\lambda_1 + \lambda_2|)|\partial_{xx} \bar{y}| \right\}, \\ \|F_p\|_{\hat{\mathcal{B}}} &\leq \hat{c}_{F_p} := \max \left\{ \hat{r}_2 + \varepsilon \cosh(2\pi \hat{r}_1), (\hat{r}_2 + \varepsilon \cosh(2\pi \hat{r}_1) + |\lambda_1|)(\hat{r}_2 + \varepsilon \cosh(2\pi \hat{r}_1) + |\lambda_2|) \right\}, \end{aligned}$$

where

$$|\bar{y}| = r_2 + \frac{\varepsilon}{2\pi} \cosh(2\pi r_1), \quad |\partial_x \bar{y}| = \varepsilon \cosh(2\pi r_1), \quad |\partial_{xx} \bar{y}| = 2\pi \varepsilon \cosh(2\pi r_1).$$

Let us characterize the constants that appear in Step 2 of Algorithm 5.1. Using Equation (112) and the fact that Ω_0 is constant, we control the analytic norms related to N_0 using

$$c_{N_0} = \|N_0\|_{F,\rho} = \|DK\|_{F,\rho}, \quad c_{N_0^\top} = \|DK^\top\|_{F,\rho}, \quad \hat{c}_{N_0} = \|DK\|_{F,\hat{\rho}}, \quad \hat{c}_{N_0^\top} = \|DK^\top\|_{F,\hat{\rho}}.$$

Notice that we can evaluate the above norms directly since K_p is a trigonometric polynomial. We also observe that $N_0^\top(\Omega \circ K)N_0$ vanishes identically, so we take $c_{N_0^\top(\Omega \circ K)N_0} = 0$.

Let us describe a simple general approach to obtain candidates of invariant curves using a global (non-perturbative) analysis of the problem. Given fixed values of λ_1 , λ_2 and ε , we take an initial condition (x_0, y_0) . To this initial condition we associate a rotation number (frequency) using the extrapolation method of [50]. Notice that we can define a (non-equispaced) sampling of a parameterization as follows $K(n\omega) = (x_n, y_n)$, where (x_n, y_n) is the orbit of the previous initial condition and ω is the computed frequency. Then, we obtain a first equispaced sampling $\{K_{p,j}\}$ by means of polynomial interpolation (as used for example in [64]). If the error of invariance of this approximation is not good enough, we apply the parameterization method.

From now on, we fix $\lambda_1 = 0.1$, $\lambda_2 = -0.2$ and $\varepsilon = 0.45$. In Figure 2 we show the iteration of several initial conditions and we observe meandering (folded) invariant curves. We consider one of the two invariant curves having rotation number $\omega = \frac{\sqrt{5}-1}{32}$, which corresponds to the green curve in Figure 2. Next we show that, for the above values of the parameters, there exists a true invariant curve nearby. To this end, we obtain an approximation of the corresponding parameterization using $N_F = 2048$ and with a numerically estimated error of invariance of 10^{-42} (using the norm $\|\cdot\|_{F,0}$). Computations are performed using interval arithmetics with 267 bits. As input of Algorithm 5.1

we enclose ω with a tight interval of length 10^{-80} , and we use the parameters

$$\begin{aligned}\rho &= 1.223945 \cdot 10^{-3}, \\ \delta &= 2.048444 \cdot 10^{-4}, \\ \sigma &= 1 + 1.601973 \cdot 10^{-11}, \\ d_{\mathcal{B}} &= 8.333835 \cdot 10^{-22}, \\ \hat{\rho} &= 1.835918 \cdot 10^{-2}.\end{aligned}$$

We use also the improved Rüssmann estimates in (90). The obtained result is

$$\frac{\mathfrak{C}_1 b_E}{\gamma^4 \rho^{4\tau}} \leq 0.0343, \quad \frac{\mathfrak{C}_2 b_E}{\gamma^2 \rho^{2\tau}} \leq 3.78 \cdot 10^{-23}.$$

Then, we can ensure the existence of a meandering invariant curve close to the green curve in Figure 2. Notice that, even though the curve is quite regular, we need a very good approximation of the invariant curve in order to apply the KAM theorem. The reason is that the twist condition around the curve is quite weak. Indeed, we obtain $c_T \leq 2388.12$ and $\sigma_T \leq 33.11$ that propagate significantly along the computation of \mathfrak{C}_1 and \mathfrak{C}_2 . Still, we are able to conclude that the distance, measure by the analytic norm, between the true invariant curve and the initial approximation is controlled by $3.78 \cdot 10^{-23}$. The computational time of this CAP is 65 seconds in a single processor Intel(R) Xeon(R) CPU at 2.40 GHz.

6.3 Froeschlé map

Next we illustrate the use of Algorithm 5.1 to prove existence of Lagrangian invariant tori in a higher dimensional case. We consider the so-called Froeschlé map (see [32]), which consists in two coupled standard maps, given by

$$F : \mathbb{T}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{T}^2 \times \mathbb{R}^2$$

$$\begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{pmatrix} \longrightarrow \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} = \begin{pmatrix} x_1 + \bar{y}_1 \\ x_2 + \bar{y}_2 \\ y_1 - \frac{\lambda_1}{2\pi} \sin(2\pi x_1) - \frac{\varepsilon}{2\pi} \sin(2\pi(x_1 + x_2)) \\ y_2 - \frac{\lambda_2}{2\pi} \sin(2\pi x_2) - \frac{\varepsilon}{2\pi} \sin(2\pi(x_1 + x_2)) \end{pmatrix}, \quad (113)$$

where $(x, y) \in \mathbb{T}^2 \times \mathbb{R}^2$ are phase space coordinates and $\lambda_1, \lambda_2, \varepsilon$ are parameters. This family has been extensively studied in the literature as a model to understand instability channels and the destruction of invariant tori [36, 44, 54, 73].

For the Froeschlé map, given by Equation (113), we have $\mathcal{A} = \mathbb{T}^2 \times \mathbb{R}^2$, $\alpha = \alpha_0 = y_1 dx_1 + y_2 dx_2$ and $\omega = \omega_0 = dy_1 \wedge dx_1 + dy_2 \wedge dx_2$. Hence, we take $c_\Omega = 1$, $c_{D\Omega} = 0$, $c_{D\alpha} = 1$ and $c_{D^2\alpha} = 0$.

Using the ambient structure, we select the transversal vectors $N_0 : \mathbb{T} \rightarrow \mathbb{R}^{4 \times 2}$ as

$$N_0(\theta) = \Omega_0 DK(\theta), \quad (114)$$

and it is clear that $DK(\theta)$ and N_0 form a basis of \mathbb{R}^4 for every $\theta \in \mathbb{T}^2$.

In Step 0 of Algorithm 5.1, we introduce the domain \mathcal{B} as

$$\mathcal{B} = \{(x, y) \in \mathbb{C}^2/\mathbb{Z}^2 \times \mathbb{C}^2 : |\operatorname{Im} x_1| \leq r_1, |\operatorname{Im} x_2| \leq r_2, |y_1| \leq r_3, |y_2| \leq r_4\},$$

and the domain $\hat{\mathcal{B}}$ as

$$\hat{\mathcal{B}} = \{(x, y) \in \mathbb{C}^2/\mathbb{Z}^2 \times \mathbb{C}^2 : |\operatorname{Im} x_1| \leq \hat{r}_1, |\operatorname{Im} x_2| \leq \hat{r}_2, |y_1| \leq \hat{r}_3, |y_2| \leq \hat{r}_4\},$$

where

$$r_1 = d_{\mathcal{B}} + \rho + \|K_p^{x_1}\|_{F,\rho}, \quad r_2 = d_{\mathcal{B}} + \rho + \|K_p^{x_2}\|_{F,\rho}, \quad r_3 = d_{\mathcal{B}} + \|K_p^{y_1}\|_{F,\rho}, \quad r_4 = d_{\mathcal{B}} + \|K_p^{y_2}\|_{F,\rho},$$

| ε | $N_{F,1} \times N_{F,2}$ | ρ | δ | $\sigma - 1$ | d_B | $\hat{\rho}$ | $\frac{c_1 b_F}{\gamma^4 \rho^4 \tau}$ | $\frac{c_2 b_F}{\gamma^2 \rho^2 \tau}$ |
|---------------|--------------------------|--------------|--------------|--------------|--------------|--------------|--|--|
| 0.005 | 128 × 128 | 5.037106e-02 | 8.325794e-03 | 9.971678e-09 | 1.304797e-13 | 2.518553e-01 | 1.22e-11 | 1.65e-24 |
| 0.010 | 128 × 128 | 4.149000e-02 | 6.943934e-03 | 2.991465e-09 | 2.984918e-14 | 2.489400e-01 | 8.18e-11 | 2.53e-24 |
| 0.015 | 128 × 128 | 3.381873e-02 | 5.660038e-03 | 1.007737e-09 | 7.594011e-15 | 2.198217e-01 | 3.17e-09 | 2.49e-23 |
| 0.020 | 128 × 128 | 2.640907e-02 | 4.419929e-03 | 3.038407e-10 | 1.678977e-15 | 1.980681e-01 | 1.51e-05 | 2.62e-20 |
| 0.025 | 256 × 256 | 1.480798e-02 | 2.447599e-03 | 2.798552e-09 | 8.363941e-15 | 1.184638e-01 | 4.82e-11 | 4.18e-25 |
| 0.030 | 512 × 512 | 6.403442e-03 | 1.071706e-03 | 1.635121e-10 | 2.072109e-16 | 1.088585e-01 | 3.38e-41 | 7.25e-57 |
| 0.035 | 512 × 512 | 6.339408e-03 | 1.060989e-03 | 1.271988e-10 | 1.484302e-16 | 1.014305e-01 | 8.06e-38 | 1.24e-53 |
| 0.040 | 512 × 512 | 6.339408e-03 | 1.060989e-03 | 1.042153e-10 | 1.129674e-16 | 1.014305e-01 | 1.13e-37 | 1.32e-53 |
| 0.045 | 512 × 512 | 6.339408e-03 | 1.060989e-03 | 8.603873e-11 | 8.652147e-17 | 1.077699e-01 | 9.77e-42 | 8.75e-58 |
| 0.050 | 512 × 512 | 6.339408e-03 | 1.060989e-03 | 7.143414e-11 | 6.649623e-17 | 1.077699e-01 | 1.71e-41 | 1.17e-57 |
| 0.055 | 512 × 512 | 6.339408e-03 | 1.060989e-03 | 5.955329e-11 | 5.127525e-17 | 1.077699e-01 | 2.98e-41 | 1.58e-57 |
| 0.060 | 512 × 512 | 6.339408e-03 | 1.060989e-03 | 5.050749e-11 | 4.016518e-17 | 1.077699e-01 | 1.34e-40 | 5.57e-57 |
| 0.065 | 512 × 512 | 6.339408e-03 | 1.060989e-03 | 4.239981e-11 | 3.109553e-17 | 1.014305e-01 | 1.96e-36 | 6.31e-53 |
| 0.070 | 512 × 512 | 6.339408e-03 | 1.060989e-03 | 3.566570e-11 | 2.408118e-17 | 9.509112e-02 | 7.89e-32 | 1.97e-48 |
| 0.075 | 1024 × 512 | 6.085832e-03 | 1.018549e-03 | 1.128252e-14 | 6.748840e-21 | 7.911581e-02 | 1.45e-16 | 1.01e-36 |

Table 8: Application of Theorem 2.6 using Algorithm 5.1 around the cubic invariant torus of the Froeschlé map (113) for different values of ε . We use the Rüssmann estimates given in (90).

and

$$\hat{r}_1 = \hat{\rho} + \|K_p^{x_1}\|_{F,\hat{\rho}}, \quad \hat{r}_2 = \hat{\rho} + \|K_p^{x_2}\|_{F,\hat{\rho}}, \quad \hat{r}_3 = \|K_p^{y_1}\|_{F,\hat{\rho}}, \quad \hat{r}_4 = \|K_p^{y_2}\|_{F,\hat{\rho}}.$$

Then, global estimates of the the symplectic map (113) in these domains are characterized as follows

$$\begin{aligned} \|DF\|_B &\leq c_{DF} := \max \left\{ 2 + \lambda_1 c_1 + 2\varepsilon c_3, 2 + \lambda_2 c_2 + 2\varepsilon c_3 \right\}, \\ \|D^2F\|_B &\leq c_{D^2F} := \max \left\{ 2\pi\lambda_1 c_1 + 4\pi\varepsilon c_3, 2\pi\lambda_2 c_2 + 4\pi\varepsilon c_3 \right\}, \\ \|F_p\|_{\hat{B}} &\leq \hat{c}_{F_p} := \max \left\{ \hat{r}_3 + \frac{\lambda_1}{2\pi} \hat{c}_1 + \frac{\varepsilon}{2\pi} \hat{c}_3, \hat{r}_4 + \frac{\lambda_3}{2\pi} \hat{c}_3 + \frac{\varepsilon}{2\pi} \hat{c}_3 \right\}, \end{aligned}$$

where

$$c_1 = \cosh(2\pi r_1), \quad c_2 = \cosh(2\pi r_2), \quad c_3 = \cosh(2\pi(r_1 + r_2)),$$

and

$$\hat{c}_1 = \cosh(2\pi \hat{r}_1), \quad \hat{c}_2 = \cosh(2\pi \hat{r}_2).$$

In Step 2 of Algorithm 5.1, we observe that

$$N_0(\theta)^\top \Omega_0 N_0(\theta) = DK(\theta)^\top \Omega_0 DK(\theta),$$

where we used $\Omega_0^\top \Omega_0 = I_4$. Using Theorem 3.3 we take

$$c_{N_0^\top(\Omega_0 K)N_0} = \|DK^\top \Omega_0 DK\|_{F,\rho} + C_{N_F}(\rho, \hat{\rho}) \|DK^\top\|_{F,\hat{\rho}} \|DK\|_{F,\hat{\rho}}.$$

Moreover, using Equation (114), we control the remaining analytic norms related to N_0 with the constants

$$c_{N_0} = \|DK\|_{F,\rho}, \quad c_{N_0^\top} = \|DK^\top\|_{F,\rho}, \quad \hat{c}_{N_0} = \|DK\|_{F,\hat{\rho}}, \quad \hat{c}_{N_0^\top} = \|DK^\top\|_{F,\hat{\rho}}.$$

We consider the frequency vector $\omega = (\nu, \nu^2)$ where ν is the so-called *cubic golden number* (the only real root of $x^3 + x - 1 = 0$). We fix the parameters $\lambda_1 = 0.01$ and $\lambda_2 = 0.02$ and we validate the invariant torus with frequency ω for different values of ε . Several validations are summarized in Table 8 whose format is analogous as in previous examples. In Table 9 we provide some complementary information about the execution of the computer program depending on the number of Fourier coefficients. We enclose the components of ω with intervals of radius μ^2 , where μ is the tolerance requested in the numerical computation of the parameterization (see Table 9). The last validated invariant torus, corresponding to $\varepsilon = 0.075$, is shown in Figure 3.

| $N_{F,1} \times N_{F,2}$ | Precision (bits) | Tolerance | CPU time (sec) |
|--------------------------|------------------|-----------|----------------|
| 128×128 | 267 | 1e-33 | 246 |
| 256×256 | 347 | 1e-38 | 751 |
| 512×512 | 533 | 1e-66 | 3628 |
| 1024×512 | 533 | 1e-66 | 6992 |

Table 9: Auxiliary implementation parameters of the CAPs corresponding to Table 8. The CAPs are performed in a single processor Intel(R) Xeon(R) CPU at 2.40 GHz.

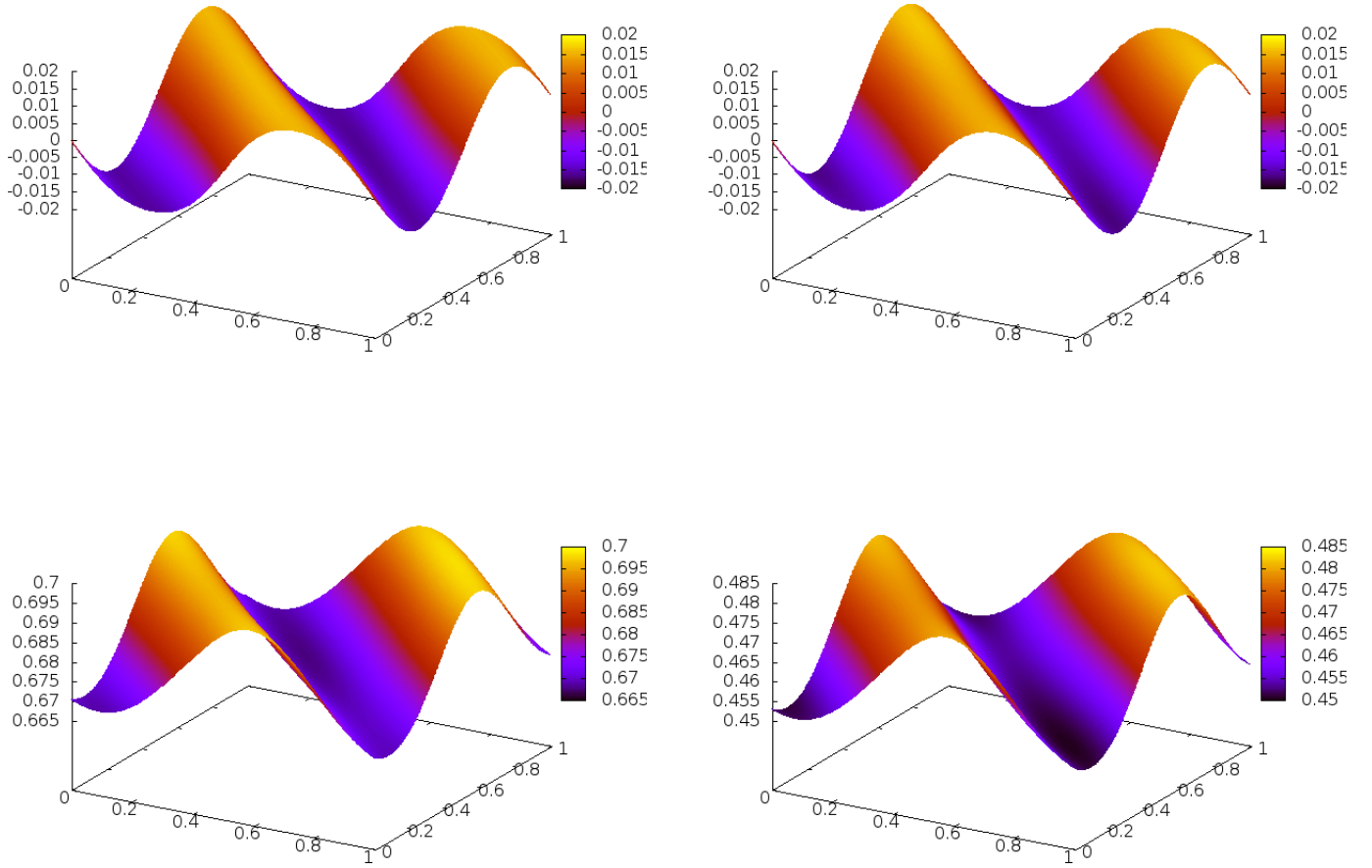


Figure 3: Last validated invariant torus ($\varepsilon = 0.075$) in Table 8. Left-Top plot: $(\theta_1, \theta_2) \mapsto K_p^{x_1}(\theta_1, \theta_2)$. Right-Top plot: $(\theta_1, \theta_2) \mapsto K_p^{x_2}(\theta_1, \theta_2)$. Left-Bottom plot: $(\theta_1, \theta_2) \mapsto K_p^{y_1}(\theta_1, \theta_2)$. Right-Bottom plot: $(\theta_1, \theta_2) \mapsto K_p^{y_2}(\theta_1, \theta_2)$.

The (numerical) breakdown value is $\varepsilon_c \simeq 0.232$, obtained using the methods described in [36, 52] (we also refer the reader to [17] for other possible methods). Notice that using 1024×512 Fourier modes we have proved existence of the invariant torus up to values of ε around 32% of the breakdown. The use of more modes would exceed our threshold of RAM memory (32 Gb).

7 Final remarks and perspectives

We conclude by presenting several remarks and comments on different aspects of the methodology of the paper. Our aim is to point out possible extensions of the methodology and future research.

It is well known that KAM theory has strong interest in the context of Hamiltonian vector fields (see e.g. [2, 4, 5, 19, 46, 78]) and, in particular, in the context of celestial mechanics (see e.g. [3, 16, 48, 49]). Given a Hamiltonian system, the methodology presented in this paper can be applied by defining a discrete dynamical system either by considering its flow or a Poincaré section. This requires to obtain rigorous enclosures for the integrations of the vector field and the variational equations up to second order (we refer the reader to the library [1]) and it may become a difficult and demanding task. Another interesting alternative to study Hamiltonian systems is to characterize the invariance equation on the vector field, thus avoiding to integrate the solutions. In this context, given a Hamiltonian vector field X_H and a frequency vector $\omega \in \mathbb{R}^n$, a Lagrangian torus is characterized by a parameterization K satisfying

$$\partial_\omega K(\theta) = X_H(K(\theta)), \quad \partial_\omega = \sum_{i=1}^n \omega_i \frac{\partial}{\partial \theta_i}.$$

Indeed, a posteriori KAM theorems for vector fields have been presented in [20, 51] (see also [37, Section 4.8.2]). The drawback of this approach, compared with performing a reduced Poincaré section, is that the dimension of phase space is not reduced, so we can be limited by the number of Fourier coefficients required to approximate the torus. Rigorous application of the a posteriori approach for Hamiltonian system is being explored, and it is still not clear which of the above strategies is the best one in general, since the computational cost and the difficulty of the implementation strongly depends on the particular problem and the analyticity of the invariant torus.

It is clear that the a posteriori approach can be used to study close-to-integrable situations (the examples in Section 6 are close-to-integrable when ε is small). In this situation, the classic KAM result (*à la Arnold* [2]) establishes that the Lebesgue measure of the gaps between surviving invariant tori is $O(\varepsilon^{1/2})$. This estimate cannot be improved in general (see e.g. [63]) and it is related to the fact that the threshold value of ε for which the KAM theorem can be applied is such that $\varepsilon = O(\gamma^2)$. However, the KAM theorem used in the present paper is a result *à la Kolmogorov* [46], that is, it is oriented to obtain the existence of a single invariant torus with prescribed frequency. This typically produces a threshold value of ε such that $\varepsilon = O(\gamma^4)$ and hence is translated to a natural lower bound for γ that is of $O(\varepsilon^{1/4})$ (see [4, 46]). The source of this discrepancy is that the Arnold approach requires to solve only one cohomological equation at every iterative step, while the Kolmogorov approach requires to solve two equations. Remarkably, the Kolmogorov approach was adapted to obtain a lower bound for γ that is of $O(\varepsilon^{1/2})$ in [76] (using classic normal forms) and, more recently, in [77] (using the a-posteriori approach). These bounds require carrying out very elaborate estimates and replacing the condition of invariance by a system of three functional equations. It is worth noticing that these works are concerned with the case of Hamiltonian systems and may find additional difficulties to deal with the case of symplectic maps. Of course, it would be interesting to adapt these techniques to obtain an analogous of Theorem 2.6 with estimates of the form $\varepsilon = O(\gamma^2)$. This would be particularly relevant to obtain computer-assisted lower bounds of the measure of invariant tori in phase space (not in frequency space). Nevertheless, in the goal pursued in this paper, the refinement in the above estimates does not represent a significant improvement since we are interested in the existence of a torus with prefixed frequency. For example, we observe that for the values of γ in Tables 1 and 2, the difference between $O(\gamma^2)$ and $O(\gamma^4)$ can be easily overcome by slightly reducing the error of invariance of the approximately invariant torus. In a typical situation, this is feasible and more advantageous than increasing the complexity of the estimates.

Having said that Theorem 2.6 is local (it asserts “just” the existence of an invariant torus), it is a natural question to connect this approach with the existence of local normal forms (*à la Kolmogorov*). A possibility to approach the problem is the following: assume that K_0 is a parameterization of an approximate invariant torus (in the annulus \mathcal{A} or in an open set of \mathbb{R}^{2n}); construct a normal frame N using Equations (9), (10), and (11); and then introduce the change of variables

$$(\theta, I) \rightarrow K(\theta) + N(\theta)I. \tag{115}$$

Then, it is not difficult to see that the obtained system is expressed as a perturbation of a Kolmogorov-like normal form, where I plays the role of an “action” variable. Indeed, Hypothesis H_4 in Theorem 2.6 is translated into the classic Kolmogorov non-degeneracy conditions. Notice that the change of variables (115) is not symplectic so we require to change the symplectic form. In order to implement this approach effectively, it may be interesting to introduce the change (115) using automatic differentiation techniques. Let us highlight that the coordinates in (115) are used only to define the system, and that the idea would be to proceed without performing any sequence of canonical transformations. This may be an interesting idea to explore in order to perform a further analysis of the local dynamics. For instance, if K_0 is an actual invariant torus, one could try to obtain classic estimates on the effective stability with respect to I (see e.g. [58]) or connect the approach with recent results on the local dynamics around KAM tori (see [26]).

Finally, it is worth mentioning that the methodology described in Section 3 extends beyond the analytical setting. It is of independent interest that similar estimates can be deduced in the case of functions with finite regularity. In this case, one uses that the decay of Fourier coefficients of a C^r -function is of the form $|f_k| \leq (2\pi k)^{-r} \|f\|_{C^r}$. The role of increasing the strip of analyticity is replaced by controlling higher order derivatives. Other spaces of interest are Sobolev spaces, which are particularly useful to the study of partial differential equations.

Acknowledgements

We are very grateful to R. Calleja, R. de la Llave, and J. Villanueva for useful and fruitful discussions along the last years. We are also grateful to J. B. van den Berg, M. Breden, R. Castelli, A. Celletti, J. Cyranka, J. Gómez-Serrano, J.-P. Lessard, J.D. Mireles-James, K. Mischaikow, C. Simó, and P. Zgliczynski for their interest and comments. We would like to acknowledge financial support from the Spanish grants MTM2012-32541, MTM2015-67724-P (MINECO/FEDER, UE) and the Catalan grant 2014-SGR-1145. J.-L. F. acknowledges the partial support from Essen, L. and C.-G., for mathematical studies. Moreover, A.L. acknowledges support from a postdoctoral position in the ERC Starting grant 335079. We acknowledge A. Granados and the use of the UPC Dynamical Systems group’s cluster for research computing (see <https://dynamicalsystems.upc.edu/en/computing/>). Finally, we would like to thank the anonymous referees for helpful comments which led us to write Section 7.

References

- [1] <http://capd.ii.uj.edu.pl/download.php>.
- [2] V.I. Arnold. Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspehi Mat. Nauk*, 18(5 (113)):13–40, 1963.
- [3] V.I. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russ. Math. Surveys*, 18:85–192, 1963.
- [4] 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.
- [5] J.B. Bost. Tores invariants des systèmes dynamiques hamiltoniens (d’après Kolmogorov, Arnold, Moser, Rüssmann, Zehnder, Herman, Pöschel, . . .). *Astérisque*, (133-134):113–157, 1986. Seminar Bourbaki, Vol. 1984/85.
- [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] 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.

- [8] 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.
- [9] R. Calleja and J.-Ll. Figueras. Collision of invariant bundles of quasi-periodic attractors in the dissipative standard map. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(3):033114, 2012.
- [10] M. Canadell and A. Haro. Parameterization method for computing quasi-periodic reducible normally hyperbolic invariant tori. In *F. Casas, V. Martínez (eds.), Advances in Differential Equations and Applications*, volume 4 of *SEMA SIMAI Springer Series*. Springer, 2014.
- [11] M. Canadell and A. Haro. A KAM-like theorem for quasi-periodic normally hyperbolic invariant tori. Preprint, 2015.
- [12] A. Castro, D. Córdoba, C. Fefferman, F. Gancedo, and J. Gómez-Serrano. Finite time singularities for the free boundary incompressible Euler equations. *Ann. of Math. (2)*, 178(3):1061–1134, 2013.
- [13] A. Celletti and L. Chierchia. Construction of Analytic KAM Surfaces and Effective Stability Bounds. *Comm. Math. Phys.*, 118(1):199–161, 1988.
- [14] A. Celletti and L. Chierchia. A constructive theory of Lagrangian tori and computer-assisted applications. In *Dynamics Reported*, pages 60–129. Springer, Berlin, 1995.
- [15] A. Celletti and L. Chierchia. On the stability of realistic three-body problems. *Comm. Math. Phys.*, 186(2):413–449, 1997.
- [16] A. Celletti and L. Chierchia. KAM stability and celestial mechanics. *Mem. Amer. Math. Soc.*, 187(878):viii+134, 2007.
- [17] A. Celletti, C. Falcolini, and U. Locatelli. On the break-down threshold of invariant tori in four dimensional maps. *Regul. Chaotic Dyn.*, 9(3):227–253, 2004.
- [18] B.V. Chirikov. A universal instability of many-dimensional oscillator systems. *Phys. Rep.*, 52(5):264–379, 1979.
- [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 A. Luque. Differentiability at the tip of Arnold tongues for Diophantine rotations: numerical studies and renormalization group explanations. *J. Stat. Phys.*, 143(6):1154–1188, 2011.
- [22] R. de la Llave and D. Rana. Accurate strategies for small divisor problems. *Bull. Amer. Math. Soc. (N.S.)*, 22(1):85–90, 1990.
- [23] R. de la Llave and D. Rana. Accurate strategies for K.A.M. bounds and their implementation. In *Computer aided proofs in analysis (Cincinnati, OH, 1989)*, volume 28 of *IMA Vol. Math. Appl.*, pages 127–146. Springer, New York, 1991.
- [24] D. del Castillo-Negrete, J.M. Greene, and P.J. Morrison. Area preserving nontwist maps: periodic orbits and transition to chaos. *Phys. D*, 91(1-2):1–23, 1996.
- [25] A. Delshams and R. de la Llave. KAM theory and a partial justification of Greene’s criterion for nontwist maps. *SIAM J. Math. Anal.*, 31(6):1235–1269 (electronic), 2000.

- [26] L.H. Eliasson, B. Fayad, and R. Krikorian. Around the stability of KAM tori. *Duke Math. J.*, 164(9):1733–1775, 2015.
- [27] C.L. Epstein. How well does the finite Fourier transform approximate the Fourier transform? *Comm. Pure Appl. Math.*, 58(10):1421–1435, 2005.
- [28] C.L. Fefferman and L.A. Seco. Interval arithmetic in quantum mechanics. In *Applications of interval computations (El Paso, TX, 1995)*, volume 3 of *Appl. Optim.*, pages 145–167. Kluwer Acad. Publ., Dordrecht, 1996.
- [29] J.-Ll. Figueras and A. Haro. Reliable computation of robust response tori on the verge of breakdown. *SIAM J. Appl. Dyn. Syst.*, 11:597–628, 2012.
- [30] 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.
- [31] 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.
- [32] C. Froesché. Numerical study of a four-dimensional mapping. *Astron. Astrophys.*, 16:172–189, 1972.
- [33] G. Gallavotti. Perturbation theory for classical Hamiltonian systems. In *Scaling and self-similarity in physics (Bures-sur-Yvette, 1981/1982)*, volume 7 of *Progr. Phys.*, pages 359–426. Birkhäuser Boston, Boston, MA, 1983.
- [34] 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.
- [35] J.M. Greene. A method for determining a stochastic transition. *J. Math. Phys.*, 20(6):1183–1201, 1975.
- [36] A. Haro. *The Primitive Function of an Exact Symplectomorphism*. PhD thesis, Universitat de Barcelona, 1998. Available at <http://hdl.handle.net/2445/42094>.
- [37] A. Haro, M. Canadell, J.-Ll. Figueras, A. Luque, and J.M. Mondelo. The parameterization method for invariant manifolds: from rigorous results to effective computations. To appear in *Applied Mathematical Sciences*, Springer. In press, 2015.
- [38] A. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: numerical algorithms. *Discrete Contin. Dyn. Syst. Ser. B*, 6(6):1261–1300, 2006.
- [39] J. Hass and R. Schlafly. Double bubbles minimize. *Ann. of Math. (2)*, 151(2):459–515, 2000.
- [40] M.-R. Herman. Sur les courbes invariantes par les difféomorphismes de l’anneau. Vol. 2. *Astérisque*, (144):248, 1986. With a correction to: *On the curves invariant under diffeomorphisms of the annulus, Vol. 1* (French) [Astérisque No. 103-104, Soc. Math. France, Paris, 1983; MR 85m:58062].
- [41] 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.
- [42] À. Jorba, R. de la Llave, and M. Zou. Lindstedt series for lower-dimensional tori. In *Hamiltonian systems with three or more degrees of freedom (S’Agaró, 1995)*, volume 533 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 151–167. Kluwer Acad. Publ., Dordrecht, 1999.
- [43] I. Jungreis. A method for proving that monotone twist maps have no invariant circles. *Ergodic Theory Dynam. Systems*, 11(1):79–84, 1991.
- [44] K. Kaneko and R. Bagley. Arnold diffusion, ergodicity and intermittency in a coupled standard mapping. *Physics Letters A*, 110(9):435–440, 1985.

- [45] H. Koch, A. Schenkel, and P. Wittwer. Computer-assisted proofs in analysis and programming in logic: a case study. *SIAM Rev.*, 38(4):565–604, 1996.
- [46] 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.
- [47] O.E. Lanford, III. A computer-assisted proof of the Feigenbaum conjectures. *Bull. Amer. Math. Soc. (N.S.)*, 6(3):427–434, 1982.
- [48] U. Locatelli. Three-body planetary problem: study of KAM stability for the secular part of the Hamiltonian. *Planetary and Space Science*, 46(11):1453–1464, 1998.
- [49] U. Locatelli and A. Giorgilli. Invariant tori in the secular motions of the three-body planetary systems. *Cel. Mech.*, 78(1):47–74, 2000.
- [50] A. Luque and J. Villanueva. Numerical computation of rotation numbers for quasi-periodic planar curves. *Phys. D*, 238(20):2025–2044, 2009.
- [51] A. Luque and J. Villanueva. A KAM theorem without action-angle variables for elliptic lower dimensional tori. *Nonlinearity*, 24(4):1033–1080, 2011.
- [52] A. Luque and J. Villanueva. A numerical method for computing initial conditions of Lagrangian invariant tori using the frequency map. *Phys. D*, 325:63–73, 2016.
- [53] 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.
- [54] R.S. MacKay, J.D. Meiss, and J. Stark. Converse KAM theory for symplectic twist maps. *Nonlinearity*, 2(4):555–570, 1989.
- [55] R.S. MacKay and I.C. Percival. Converse KAM: theory and practice. *Comm. Math. Phys.*, 98(4):469–512, 1985.
- [56] J.N. Mather. Non-existence of invariant circles. *Ergodic Theory Dyn. Syst.*, 4:301–309, 1984.
- [57] G.T. Minton. *Computer-assisted proofs in geometry and physics*. PhD thesis, Massachusetts Institute of Technology, 2013.
- [58] A. Morbidelli and A. Giorgilli. Superexponential stability of KAM tori. *J. Statist. Phys.*, 78(5-6):1607–1617, 1995.
- [59] J. Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*, 1962:1–20, 1962.
- [60] J. Moser. On the theory of quasiperiodic motions. *SIAM Rev.*, 8(2):145–172, 1966.
- [61] J. Moser. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:499–535, 1966.
- [62] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169:136–176, 1967.
- [63] A. I. Neishtadt. Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions. *J. Appl. Math. Mech.*, 45(6):1016–1025, 1981.

- [64] A. Olvera and N.P. Petrov. Regularity properties of critical invariant circles of twist maps, and their universality. *SIAM J. Appl. Dyn. Syst.*, 7(3):962–987, 2008.
- [65] N. Revol and F. Rouillier. Motivations for an arbitrary precision interval arithmetic and the mpfi library. *Reliable computing*, 11(5):275–290, 2005.
- [66] T. J. Rivlin. *An introduction to the approximation of functions*. Blaisdell Publishing Co. Ginn and Co., Waltham, Mass.-Toronto, Ont.-London, 1969.
- [67] 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.
- [68] 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.
- [69] 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.
- [70] D. Salamon and E. Zehnder. KAM theory in configuration space. *Comment. Math. Helv.*, 64(1):84–132, 1989.
- [71] A. Schenkel, J. Wehr, and P. Wittwer. Computer-assisted proofs for fixed point problems in Sobolev spaces. *Math. Phys. Electron. J.*, 6:Paper 3, 67, 2000.
- [72] C. Simó. Invariant curves of analytic perturbed nontwist area preserving maps. *Regul. Chaotic Dyn.*, 3(3):180–195, 1998. J. Moser at 70.
- [73] S. Tompaids. Approximation of invariant surfaces by periodic orbits in high-dimensional maps: some rigorous results. *Experiment. Math.*, 5(3):197–209, 1996.
- [74] W. Tucker. The Lorenz attractor exists. *C. R. Acad. Sci. Paris Sér. I Math.*, 328(12):1197–1202, 1999.
- [75] W. Tucker. *Validated Numerics: A Short Introduction to Rigorous Computations*. Princeton University Press, Princeton, NJ, 2011.
- [76] J. Villanueva. Kolmogorov theorem revisited. *J. Differential Equations*, 244(9):2251–2276, 2008.
- [77] J. Villanueva. A new Approach to the Parameterization Method for Lagrangian Tori of Hamiltonian Systems. Preprint, 2016.
- [78] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. II. *Comm. Pure Appl. Math.*, 29(1):49–111, 1976.

A A heuristic selection of parameters to validate invariant tori

In this appendix we describe a direct method to select the implementation parameters ρ , δ , σ , ρ_∞ , d_B , and $\hat{\rho}$, required in Input 4 of Algorithm 5.1 to rigorously validate an invariant torus. The idea consists in using the structure of the constants that appear in Step 4 (c.f. Section 5.6). We do not claim that the procedure described below is optimal, but allows us to obtain suitable values of the parameters at a moderate computational cost.

We assume that the term $2(a_3)^{3\tau+1}\gamma^3\rho^{3\tau-1}C_3$ does not contribute to the constant \mathfrak{C}_1 and we look for parameters such that the constants in (108) satisfy $\mathfrak{C}_3 = \mathfrak{C}_4 = \mathfrak{C}_5$. The neglected term is in general very small (it stands for the control of the approximate Lagrangian character of the approximated torus) and this assumption allows us to simplify the heuristic analysis of the constants.

We proceed in analogy with the procedure described in [37]. We observe that the dependence on a_2 is very simple: it appears only in the expression $a_3 = 3 \frac{a_1}{a_1-1} \frac{a_2}{a_2-1}$ and in the final strip of analyticity $\rho_\infty = \rho/a_2$. In order to look for the limit condition, we take $a_2 = \infty$ (so $\rho_\infty = 0$) in all subsequent computations.

Now we can describe a very simple algorithm to obtain suitable values of the parameters ρ , δ , σ , d_B and $\hat{\rho}$ for a given parameterization K in a grid of size $N_F = (N_{F,1}, \dots, N_{F,n})$:

- I. We take $\rho_0 = -\log(\|E\|_{F,0})/(2\pi N)$, where E is the error of invariance and $N = \max_i \{N_{F,i}\}$. This will be the initial value of ρ .
- II. For a given value ρ , we consider values $\delta \in [\frac{\rho}{6.5}, \frac{\rho}{4.5}]$ (recall that $a_3 = \frac{\rho}{\delta}$ and $a_1 = \frac{a_3-3}{a_3}$). For any of these values (ρ, δ) we compute σ and d_B solving the equations $\mathfrak{C}_4 = \mathfrak{C}_5$ and $\mathfrak{C}_3 = \mathfrak{C}_4$, which are respectively written as follows:

$$\sigma_*(1 - a_1^{-2\tau})d_B - (\sigma - 1)\delta(1 - a_1^{1-2\tau}) = 0, \quad (116)$$

$$\sigma_*(a_3)^{2\tau+1}\gamma^2\rho^{2\tau-1}\hat{C}_2 - (\sigma - 1)(1 - a_1^{1-2\tau})(a_1a_3)^{4\tau}\hat{C}_5 = 0. \quad (117)$$

Let us recall that σ_* , \hat{C}_2 and \hat{C}_5 depend on ρ , δ , σ , d_B , and $\hat{\rho}$. In order to avoid the dependence on $\hat{\rho}$ in the above expression, we take $C_{N_F}(\rho, \hat{\rho}) = 0$ when computing these constants. A suitable value of $\hat{\rho}$ is fixed later. Then, we select the value of δ that minimizes the expression $\mathfrak{C}_1\gamma^{-4}\rho^{-4\tau}\|E\|_\rho$.

- III. If $\mathfrak{C}_1\gamma^{-4}\rho^{-4\tau}\|E\|_\rho \geq 1$, we decrease the value of ρ and repeat step II, thus obtaining a new value of \mathfrak{C}_1 . We proceed until we find that $\mathfrak{C}_1\gamma^{-4}\rho^{-4\tau}\|E\|_\rho < 1$. If at any point we reach a minimum of the function $\mathfrak{C}_1\gamma^{-4}\rho^{-4\tau}\|E\|_\rho$ then we stop the computations. If this condition is not satisfied at the minimum, then we need a better approximation of the invariant torus.
- IV. Assume that we have obtained values of ρ , δ , σ , and d_B as above. Then, we take a sequence of increasing values of $\hat{\rho}$ (starting at a value slightly greater than ρ) and compute the constant $C_{N_F}(\rho, \hat{\rho})$. Then we compute again the constant \mathfrak{C}_1 and the bound b_E (see Section 5.3). We select a value of $\hat{\rho}$ that minimizes the expression $\mathfrak{C}_1\gamma^{-4}\rho^{-4\tau}b_E$.