Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 385 (2024) 325–361

Journal of Differential Equations

www.elsevier.com/locate/jde

# Global properties of generic real–analytic nearly–integrable Hamiltonian systems

L. Biasco, L. Chierchia *

*Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, Largo San Leonardo Murialdo 1 - 00146 Roma, Italy*

Received 14 July 2023; accepted 17 December 2023
Available online 3 January 2024

**Abstract**

We introduce a new class $\mathbb{G}^n_s$ of generic real analytic potentials on $\mathbb{T}^n$ and study global analytic properties of natural nearly–integrable Hamiltonians $\frac{1}{2}|y|^2 + \varepsilon f(x)$, with potential $f \in \mathbb{G}^n_s$, on the phase space $\mathcal{M} = B \times \mathbb{T}^n$ with $B$ a given ball in $\mathbb{R}^n$. The phase space $\mathcal{M}$ can be covered by three sets: a 'non–resonant' set, which is filled up to an exponentially small set of measure $e^{-c\mathtt{K}}$ (where $\mathtt{K}$ is the maximal size of resonances considered) by primary maximal KAM tori; a 'simply resonant set' of measure $\sqrt{\varepsilon}\mathtt{K}^a$ and a third set of measure $\varepsilon\mathtt{K}^b$ which is 'non perturbative', in the sense that the $\mathtt{H}$–dynamics on it can be described by a natural system which is *not* nearly–integrable. We then focus on the simply resonant set – the dynamics of which is particularly interesting (e.g., for Arnol'd diffusion, or the existence of secondary tori) – and show that on such a set the secular (averaged) 1 degree–of–freedom Hamiltonians (labeled by the resonance index $k \in \mathbb{Z}^n$) can be put into a universal form (which we call 'Generic Standard Form'), whose main analytic properties are controlled by *only one parameter, which is uniform in the resonance label $k$*.


*Keywords:* KAM theory; Hamiltonian perturbation theory; Averaging theory; Nearly-integrable systems; Geometry of resonances

## 0. Introduction

The paper is divided in three parts.

* Corresponding author.
*E-mail addresses:* luca.biasco@uniroma3.it (L. Biasco), luigi.chierchia@uniroma3.it (L. Chierchia).

https://doi.org/10.1016/j.jde.2023.12.018

**1.** In the first part we discuss generic properties of (multi–periodic) analytic functions introducing a new class $\mathbb{G}^n_s$ of functions real analytic on the complex neighborhood of $\mathbb{T}^n$ given by

$$\mathbb{T}^n_s := \{x = (x_1, ..., x_1) \in \mathbb{C}^n : \ |\operatorname{Im} x_j| < s\}/(2\pi\mathbb{Z}^n) .$$

Such a class – related but smaller than the sets $\mathcal{H}_{s,\tau}$ introduced in [6] – is generic, as it contains an open and dense set in the norm $\|f\|_s = \sup_k |f_k| e^{s|k|_1}$, and has full probability measure with respect to a natural weighted product probability measure on Fourier coefficients.

The class $\mathbb{G}^n_s$ may be described as follows. Consider a real–analytic zero average function $f$ and consider its projection $\pi_{\mathbb{Z}k} f$ onto the Fourier modes proportional to a given $k \in \mathbb{Z}^n \setminus \{0\}$ (with components with no common divisors), which is given by

$$\theta \in \mathbb{T} \mapsto \pi_{\mathbb{Z}k} f(\theta) := \sum_{j\in\mathbb{Z}} f_{jk} e^{\mathrm{i} j\theta} .$$

These one dimensional projections arise naturally, e.g., in averaging theory, where they are the leading terms of the averaged ('secular') Hamiltonians around simple resonances $\{y | y \cdot k = \sum y_j k_j = 0\}$. Denoting by $\mathcal{G}^n$ the set of generators of 1–dimensional maximal lattices in $\mathbb{Z}^n$, the class $\mathbb{G}^n_s$ is formed by real–analytic zero average functions $f$ with $\|f\|_s \leqslant 1$, which satisfy

$$\delta_f = \varliminf_{\substack{|k|_1 \to +\infty \\ k\in\mathcal{G}^n}} |f_k| e^{|k|_1 s} |k|_1^n > 0 ,$$

and such that the Fourier–projection $\pi_{\mathbb{Z}k} f$ is a Morse function with distinct critical values for all $k \in \mathcal{G}^n$ with $|k|_1 \leqslant \mathtt{N}$, where $\mathtt{N}$ is a *a–priori* Fourier cut–off depending only on $n, s$ and $\delta_f$.

A remarkable feature of this class of functions is that the Fourier projection $\pi_{\mathbb{Z}k} f$ is close (in a large analytic norm) to a shifted rescaled cosine,

$$\pi_{\mathbb{Z}k} f(\theta) \sim |f_k| \cos(\theta + \theta_k) ; \qquad \forall k \in \mathcal{G}^n , |k|_1 \geqslant \mathtt{N} ,$$

(Proposition 1.1 below), allowing to have uniform control of the analytic properties of secular Hamiltonians as $|k|_1 \to +\infty$.

We believe that the class $\mathbb{G}^n_s$ is a good candidate to address analytic problems in dynamical systems whenever generic results – such as generic existence of secondary tori in nearly–integrable Hamiltonian systems[1] or Arnol'd diffusion[2] – are considered.

**2.** In the rest of the paper, we consider natural nearly–integrable Hamiltonian systems with $n \geqslant 2$ degrees of freedom with Hamiltonian $\mathtt{H} = \frac{1}{2}|y|^2 + \varepsilon f(x)$ $(n \geqslant 2)$, with potential $f$ in the class $\mathbb{G}^n_s$ with a fixed $s > 0$, on a bounded phase space $\mathcal{M} = \mathrm{B} \times \mathbb{T}^n \subset \mathbb{R}^n \times \mathbb{T}^n$; in fact, in view of the model, it is not restrictive to simply consider $\mathrm{B} = \{y \in \mathbb{R}^n \text{ s.t. } |y| < 1\}$.

We, then, introduce a covering of the action space $\mathrm{B} = \mathcal{R}^0 \cup \mathcal{R}^1 \cup \mathcal{R}^2$, depending on two 'Fourier cut–offs' $\mathtt{K} > \mathtt{K}_{\mathrm{o}} > \mathtt{N}$ ($\mathtt{N}$ as above), so that: $\mathcal{R}^0$ is a non–resonant set up to order $\mathtt{K}_{\mathrm{o}}$; $\mathcal{R}^1$ is union of neighborhoods $\mathcal{R}^{1,k}$ of simple resonances $\{y \in \mathrm{B} : y \cdot k = 0\}$ of maximal order $\mathtt{K}_{\mathrm{o}}$, which are

[1] I.e., tori which are not a continuous deformation of integrable ($\varepsilon = 0$) flat tori; for references, see [25], [23], [31], [1], [6], [19], [7].

[2] See [2]; compare also, e.g., [11], [14], [33], [22], [32], [15], [9], [16], [21], [17], [10], [12].

non resonant modulo $\mathbb{Z}k$ up to order $\mathtt{K}$, and $\mathcal{R}^2$ is a set of measure proportional to $\varepsilon \mathtt{K}^b$ for a suitable $b > 1$ (which depends only on $n$); similar 'geometry–of–resonances' analysis is typical of Nekhoroshev's theory.[3]

The set $\mathcal{R}^2$ is *a non perturbative set*, namely, it is a set where the $\mathtt{H}$–dynamics is equivalent to the dynamics of a Hamiltonian, which is *not nearly–integrable*: indeed, in the simplest non trivial case $n = 2$ such a Hamiltonian is given by $|y|^2/2 + f(x)$.

On the other hand the set $(\mathcal{R}^0 \cup \mathcal{R}^1) \times \mathbb{T}^n$ is suitable for high order perturbation theory, and, following the averaging theory developed in [6], we construct high order normal forms (Theorem 2.1) so that on $\mathcal{R}^0 \times \mathbb{T}^n$ the above Hamiltonian $\mathtt{H}$ is conjugated, up to an exponentially small term of $O(e^{-\mathtt{K}_o s/3})$, to an integrable Hamiltonian, which depends only on action variables and it is close to $|y|^2/2$. By classical KAM theory, it then follows that this set is filled by primary[4] KAM tori up to a set of measure of order $O(e^{-\mathtt{K}_o s/7})$. Actually, here there is a delicate point: the symplectic map realizing the above mentioned conjugation moves the boundary of the phase space $\mathrm{B} \times \mathbb{T}^n$ by a quantity much larger than $O(e^{-\mathtt{K}_o s/7})$, therefore, in order to get the exponentially small measure estimate on the 'non–torus set' one needs to introduce a second covering which takes care of the dynamics close to the boundary: this is done in Lemma 2.3 below.

The analysis on the dynamics in $\mathcal{R}^1 \times \mathbb{T}^n$ is *much* more complicate. In each of the neighborhoods $\mathcal{R}^{1,k}$, which cover the set $\mathcal{R}^1$ as $|k|_1 \leqslant \mathtt{K}_o$, one can perform resonant averaging theory so as to conjugate $\mathtt{H}$ to still an integrable system, which however depends on the resonant angle $\mathtt{x}_1 = k \cdot x$. The averaged systems with secular Hamiltonians $\overline{\mathtt{H}}_k(\mathtt{y}, \mathtt{x}_1)$ are therefore 1D–Hamiltonian systems (one degree–of–freedom systems in the symplectic variables $(\mathtt{y}_1, \mathtt{x}_1)$ depending also on adiabatic actions $\mathtt{y}_2, ..., \mathtt{y}_n$), which are close to natural systems with potentials $\pi_{\mathbb{Z}k} f$. Such potentials, for low $k$'s, are rather general: for instance, they may have an arbitrary large number of separatrices depending on the particular structure of $f$. The global analytic properties of the Hamiltonians $\overline{\mathtt{H}}_k(\mathtt{y}, \mathtt{x}_1)$ is the argument of the third (and main) part of this paper.

**3.** In the third part we prove that the secular Hamiltonians $\overline{\mathtt{H}}_k(\mathtt{y}, \mathtt{x}_1)$ described in the previous item *can be symplectically conjugated, for all $|k|_1 \leqslant \mathtt{K}_o$, to 1D–Hamiltonians in the standard form introduced in* [5] (see, also, Definition 3.1 below). In few words, a standard 1D–Hamiltonian (which depends on $(n - 1)$ external parameters) is a one degree–of–freedom Hamiltonian system close to a natural system with a generic potential, which may be controlled essentially by *only one parameter*, namely, the parameter $\kappa$ appearing in Eq. (81) below; here, 'essentially' means, roughly speaking, that $\kappa$ governs the main scaling properties of the Hamiltonian $\overline{\mathtt{H}}_k$. What is particularly relevant is that the $\kappa$ parameter of the secular Hamiltonians $\overline{\mathtt{H}}_k$ is shown to be *independent of $k$*, as it depends only on $n$, $s$, the above parameter $\delta_f$, and on a fourth parameter $\beta$, which measures the Morse properties of the potentials $\pi_{\mathbb{Z}k} f$ with $|k|_1 \leqslant \mathtt{N}$; compare Eq. (87) and Remark 3.3–(i).

This uniformity allows to analyze global analytic properties: for example, the action–angle map for standard Hamiltonians, as discussed in [5], depends only on the parameter $\kappa$ of the standard Hamiltonian and therefore can be used *simultaneously* for all the secular Hamiltonians $\overline{\mathtt{H}}_k$, allowing for a nearly–integrable description of $\mathtt{H}$ on $\mathcal{R}^{1,k} \times \mathbb{T}^n$ with uniformly exponentially small perturbations.

The results presented in this paper may be useful in attacking some of the fundamental open problems in the analytic theory of nearly–integrable Hamiltonian systems such as Arnol'd diffusion

[3] Compare [26], [4], [30], [13], [27], [14], [18], [8], [28].

[4] Primary tori are smooth deformation of the flat Lagrangian integrable ($\varepsilon = 0$) tori.

for generic real analytic systems, and provide indispensable tools to develop a ‘singular KAM Theory’, namely a KAM theory dealing simultaneously with primary and secondary persistent Lagrangian tori in the full phase space, except for the non–perturbative set $\mathcal{R}^2$. In particular, Theorem 3.1 below is the starting point for, e.g., the following result, which (up to the logarithmic correction and in the case of natural systems) proves a conjecture by Arnold, Kozlov and Neishtadt.[5]

**Theorem.** ([7]) *Fix $n \geqslant 2$, $s > 0$, $f \in \mathbb{G}^n_s$, $B$ an open ball in $\mathbb{R}^n$, and let $\mathtt{H}(y, x; \varepsilon) := \frac{1}{2}|y|^2 + \varepsilon f(x)$. Then, there exists a constant $c > 1$ such that, for all $0 < \varepsilon < 1$, all points in $B \times \mathbb{T}^n$ lie on a maximal KAM torus for $\mathtt{H}$, except for a subset of measure bounded by $c\,\varepsilon|\log\varepsilon|^\gamma$ with $\gamma := 11n + 4$.*

Let us remark that, since it is well known that the asymptotic (as $\varepsilon \to 0$) density of non–integrable *primary* tori is $1 - c\sqrt{\varepsilon}$ (see [24], [29]), the difference of order of the density of invariant maximal tori in the above theorem must come from secondary tori, i.e., the tori in $\mathcal{R}^1 \times \mathbb{T}^n$ whose leading dynamics is governed by the secular Hamiltonians $\overline{\mathtt{H}}_k(\mathtt{y}, \mathtt{x}_1)$ discussed in this paper.

## 1. Generic real analytic periodic functions

We begin with a few definitions.

**Definition 1.1. (Norms on real analytic periodic functions)**
For $s > 0$ and $n \in \mathbb{N} = \{1, 2, 3...\}$, consider the Banach space of zero average real analytic periodic functions $f : x \in \mathbb{T}^n := \mathbb{R}^n/(2\pi\mathbb{Z}^n) \mapsto \sum_{k\in\mathbb{Z}} f_k e^{\mathrm{i}k\cdot x}$, $f_0 = 0$, with finite norm[6]

$$\|f\|_s := \sup_{k\in\mathbb{Z}^n} |f_k| e^{|k|_1 s}\,,$$

and denote by $\mathbb{B}^n_s$ the closed unit ball of functions $f$ with $\|f\|_s \leqslant 1$.

Besides the norm $\|\cdot\|_s$, we shall also use the following two (non equivalent) norms

$$|f|_s := \sup_{\mathbb{T}^n_s} |f|\,, \qquad \boldsymbol{|}\, f \,\boldsymbol{|}_s := \sum_{k\in\mathbb{Z}^n} |f_k| e^{|k|_1 s}\,.$$

Such norms satisfy the relations

$$\|f\|_s \leqslant |f|_s \leqslant \boldsymbol{|}\, f \,\boldsymbol{|}_s\,.$$

Notice also the following ‘smoothing property’ of the norm $\boldsymbol{|}\cdot\boldsymbol{|}_s$: *if $s' \leqslant s$, then for any $N \geqslant 1$, one has*

[5] See, [3, Remark 6.18, § 6.3–C].
[6] As usual $|k|_1 := \sum |k_j|$.

$$f(y,x)=\sum_{|k|_1\geqslant N} f_k(y)e^{\mathrm{i}k\cdot x} \qquad \Longrightarrow \qquad |\, f\,|_{s'}\leqslant e^{-(s-s')N}\,|\, f\,|_s\,. \tag{1}$$

**Definition 1.2. (Generators and Fourier projectors)**
(i) Let $\mathbb{Z}^n_\star$ be the set of integer vectors $k\neq 0$ in $\mathbb{Z}^n$ such that the first non–null component is positive:

$$\mathbb{Z}^n_\star := \big\{k\in\mathbb{Z}^n:\ k\neq 0 \text{ and } k_j>0 \text{ where } j=\min\{i: k_i\neq 0\}\big\}\,, \tag{2}$$

and denote by $\mathcal{G}^n$ the set of *generators of 1d maximal lattices* in $\mathbb{Z}^n$, namely, the set of vectors $k\in\mathbb{Z}^n_\star$ such that the greater common divisor (gcd) of their components is 1:

$$\mathcal{G}^n := \{k\in\mathbb{Z}^n_\star:\ \gcd(k_1,\dots,k_n)=1\}\,.$$

Let us also denote by $\mathcal{G}^n_K$ the generators of size not exceeding $K\geqslant 1$,

$$\mathcal{G}^n_K := \mathcal{G}^n\cap\{|k|_1\leqslant K\}\,,$$

(ii) Given a zero average real analytic periodic function and $k\in\mathcal{G}^n$, we define

$$\theta\in\mathbb{T}\mapsto \pi_{\mathbb{Z}k}f(\theta):=\sum_{j\in\mathbb{Z}} f_{jk}e^{\mathrm{i}j\theta}\,. \tag{3}$$

Notice that $f$ can be uniquely written as

$$f(x)=\sum_{k\in\mathcal{G}^n}\pi_{\mathbb{Z}k}f(k\cdot x)\,.$$

**Definition 1.3.** Let $\beta>0$. A function $F\in C^2(\mathbb{T},\mathbb{R})$ is called a $\beta$**–Morse function** if

$$\min_{\theta\in\mathbb{T}}\big(|F'(\theta)|+|F''(\theta)|\big)\geqslant\beta\,,\quad \min_{i\neq j}|F(\theta_i)-F(\theta_j)|\geqslant\beta\,,$$

where $\theta_i\in\mathbb{T}$ are the critical points of $F$.

**Definition 1.4. (Cosine–like functions)** Let $0<\mathtt{g}<1/4$. We say that a real analytic function $G:\mathbb{T}_1\to\mathbb{C}$ is $\mathtt{g}$–cosine–like if, for some $\eta>0$ and $\theta_0\in\mathbb{R}$, one has

$$|G(\theta)-\eta\cos(\theta+\theta_0)|_1 := \sup_{|\,\mathrm{Im}\,\theta|<1}|G(\theta)-\eta\cos(\theta+\theta_0)|\leqslant\eta\mathtt{g}\,.$$

Notice that this notion is invariant by rescalings: $G$ is $\mathtt{g}$–cosine–like if and only if $\lambda G$ is $\mathtt{g}$–cosine–like for any $\lambda>0$. Beware of the usage of $|\cdot|_1$ as sup norm on $\mathbb{T}_1$, the complex strip of width 2.

Now, the main definition.

**Definition 1.5. (The analytic class $\mathbb{G}^n_s$)** We denote by $\mathbb{G}^n_s$ the subset of functions $f \in \mathbb{B}^n_s$ such that the following two properties hold:

$$\varliminf_{\substack{|k|_1 \to +\infty \\ k \in \mathcal{G}^n}} |f_k| e^{|k|_1 s} |k|_1^n > 0 \,, \tag{4}$$

$$\forall\, k \in \mathcal{G}^n \,, \ \pi_{\mathbb{Z}k} f \ \text{is a Morse function with distinct critical values}\,.$$

**Remark 1.1.** (i) If $f \in \mathbb{B}^n_s$, then the function $\pi_{\mathbb{Z}k} f$ belongs to $\mathbb{B}^1_{|k|_1 s}$ and therefore has a domain of analyticity which increases with the norm of $k$.

(ii) A simple example of function in $\mathbb{G}^n_s$ is given by

$$f(x) := 2 \sum_{k \in \mathcal{G}^n} e^{-s|k|_1} \cos k \cdot x \,.$$

Indeed, one checks immediately that

$$\|f\|_s = 1 \,, \qquad \varliminf_{\substack{|k|_1 \to +\infty \\ k \in \mathcal{G}^n}} |f_k| e^{|k|_1 s} |k|_1^n = +\infty \,, \qquad \pi_{\mathbb{Z}k} f(\theta) = 2e^{-s|k|_1} \cos\theta \,.$$

(iii) The critical points of an analytic Morse function on $\mathbb{T}$, by compactness, cannot accumulate, hence, there are a finite, even number of them, which are, alternately, a relative strict maximum and a relative strict minimum. In particular, if $G$ is $\beta$–Morse, then the number of its critical points can be bounded by $\pi\sqrt{2\max|G''|/\beta}$. Indeed, if $\theta \neq \theta'$ are critical points of $G$, then, by (79) one has

$$\beta \leqslant |G(\theta) - G(\theta')| \leqslant \tfrac{1}{2}(\max|G''|)\, \mathrm{dist}(\theta, \theta')^2 \,,$$

which implies that the minimal distance between two critical points is $\sqrt{2\beta/\max|G''|}$ and the claim follows.

(iv) The requirement of having different critical values in the definition of $\beta$–Morse functions is not strictly necessary and it has been made only for technical convenience.

### *1.1. Uniform behavior of large-mode Fourier projections*

If a function $f \in \mathbb{B}^n_s$ satisfies (4), then, *apart from a finite number of Fourier modes, its Fourier projections $\pi_{\mathbb{Z}k} f$ are close to a shifted rescaled cosine*, a fact that allows, e.g., to have a uniform analytic theory of high order perturbation theory.

To discuss this matter, let us first point out that for any sequence of real numbers $\{a_k\}$ and for any function $N(\delta)$ such that $\lim_{\delta\downarrow 0} N(\delta) = +\infty$ one has

$$\varliminf a_k > 0 \quad \iff \quad \exists\, \delta > 0 \ \text{ s.t. } \ a_k \geqslant \delta \,, \ \forall\, k \geqslant N(\delta) \,. \tag{5}$$

We shall apply this remark to the minimum limit in (4) with a particular choice of the function $N(\delta)$, namely, we define $\mathtt{N}(\delta) = \mathtt{N}(\delta; n, s)$ as

$$\mathtt{N}(\delta) := 2\max\left\{1\,,\ \frac{1}{s}\log\frac{c_{\rm o}}{s^n\,\delta}\right\}, \qquad c_{\rm o} := 2^{44}\,(2n/e)^n\,. \tag{6}$$

For later use, we point out that[7]

$$\mathtt{N} \geqslant 2\mathtt{c}_s\,, \quad \text{where} \quad \mathtt{c}_s := \max\left\{1, 1/s\right\}. \tag{7}$$

From (5) it follows that if $f$ satisfies (4), one can find $0 < \delta \leqslant 1$ such that

$$|f_k| \geqslant \delta |k|_{_1}^{-n}\, e^{-|k|_1 s}\,, \qquad \forall\, k \in \mathcal{G}^n\,,\ |k|_{_1} \geqslant \mathtt{N}\,. \tag{8}$$

The main feature of the above choice of $\mathtt{N}$ is that, for $|k|_{_1} \geqslant \mathtt{N}$, $\pi_{\mathbb{Z}k} f$ is very close to a shifted rescaled cosine function:

**Proposition 1.1.** *Let $\delta > 0$, $f \in \mathbb{B}^n_s$ and assume (8). Then, for any $k \in \mathcal{G}^n$ with $|k|_{_1} \geqslant \mathtt{N}$, $\pi_{\mathbb{Z}k} f$ is $2^{-40}$–cosine–like (Definition 1.4).*

**Proof.** We shall prove something slightly stronger, namely, that there exists $\theta_k \in [0, 2\pi)$ so that

$$\pi_{\mathbb{Z}k} f(\theta) = 2|f_k|\big(\cos(\theta+\theta_k) + F^k_\star(\theta)\big)\,, \quad F^k_\star(\theta) := \frac{1}{2|f_k|}\sum_{|j|\geqslant 2} f_{jk} e^{\mathrm{i}j\theta}\,, \tag{9}$$

with $F^k_\star \in \mathbb{B}^1_1$ and (recall the definition of the norms in (24))

$$|F^k_\star|_1 \leqslant \|F^k_\star\|_1 \leqslant 2^{-40}\,. \tag{10}$$

Indeed, by definition of $\pi_{\mathbb{Z}k} f$,

$$\pi_{\mathbb{Z}k} f(\theta) := \sum_{j\in\mathbb{Z}\setminus\{0\}} f_{jk} e^{\mathrm{i}j\theta} = \sum_{|j|=1} f_{jk} e^{\mathrm{i}j\theta} + \sum_{|j|\geqslant 2} f_{jk} e^{\mathrm{i}j\theta}\,,$$

and, defining $\theta_k \in [0, 2\pi)$ so that $e^{\mathrm{i}\theta_k} = f_k/|f_k|$, one has

$$\frac{1}{2|f_k|}\sum_{|j|=1} f_{jk} e^{\mathrm{i}j\theta} = \mathrm{Re}\Big(\frac{f_k}{|f_k|} e^{\mathrm{i}\theta}\Big) = \mathrm{Re}\, e^{\mathrm{i}(\theta+\theta_k)} = \cos(\theta+\theta_k)\,,$$

which yields (9). Now, since $f \in \mathbb{B}^n_s$ it is $|f_k| \leqslant e^{-|k|_1 s}$ and, by (8), $|f_k| \geqslant \delta|k|_{_1}^{-n}\, e^{-|k|_1 s}$. Therefore, for $|k|_{_1} \geqslant \mathtt{N}$, one has

[7] In fact, if $s \geqslant 1$ then $\mathtt{N} \geqslant 2 \geqslant 2/s$, while if $s < 1$ then the logarithm in (6) is larger than one, so that $\mathtt{N} \geqslant 2/s$ also in this case.

$$
\begin{aligned}
|F_\star^k|_1 &\overset{(9)}{=} \frac{1}{2|f_k|}\sum_{|j|\geqslant 2}|f_{jk}|e^{|j|} \leqslant \frac{|k|_1^n e^{|k|_1 s}}{2\delta}\sum_{|j|\geqslant 2}|f_{jk}|e^{|j|}\\
&\leqslant \frac{|k|_1^n e^{|k|_1 s}}{2\delta}\sum_{|j|\geqslant 2} e^{-|j|(|k|_1 s-1)}\\
&\leqslant \frac{2e^2|k|_1^n}{\delta}\, e^{-|k|_1 s} = \frac{2^{n+1}e^2}{s^n\delta} e^{-\frac{|k|_1 s}{2}}\ \Big(\frac{|k|_1 s}{2}\Big)^n e^{-\frac{|k|_1 s}{2}}\\
&\leqslant \Big(\frac{2n}{es}\Big)^n \frac{2e^2}{\delta}\, e^{-\frac{\mathtt{N}s}{2}} \leqslant 2^{-40}, \qquad (11)
\end{aligned}
$$

where the geometric series converges since $|k|_1 s \geqslant \mathtt{N}s \geqslant 2$ (by (7)) and last inequality follows by definition of $\mathtt{N}$ in (6). $\square$

**Remark 1.2.** In fact, the particular form of $\mathtt{N}$ is used *only* in the last inequality in (11).

Next, we need an elementary calculus lemma:

**Lemma 1.1.** *Assume that $F \in C^2(\mathbb{T}, \mathbb{R})$, $\bar\theta$ and $0 < \mathtt{c} < 1/2$ are such that*

$$\|F - \cos(\theta + \bar\theta)\|_{C^2} \leqslant \mathtt{c}\,,$$

*where $\|F\|_{C^2} := \max_{0\leqslant k\leqslant 2}\sup|F^{(k)}|$. Then, $F$ has only two critical points and it is $(1-2\mathtt{c})$–Morse (Definition 1.3).*

**Proof.** By considering the translated function $\theta \to F(\theta - \bar\theta)$, one can reduce oneself to the case $\bar\theta = 0$ ($F$ is $\beta$–Morse, if and only if $\theta \to F(\theta - \bar\theta)$ is $\beta$–Morse).
Thus, we set $\bar\theta = 0$, and note that, by assumption $|F'| = |F' + \sin\theta - \sin\theta| \geqslant |\sin\theta| - \mathtt{c}$, and, analogously, $|F''| \geqslant |\cos\theta| - \mathtt{c}$. Hence, $|F'| + |F''| \geqslant |\sin\theta| + |\cos\theta| - 2\mathtt{c} \geqslant 1 - 2\mathtt{c}$. Next, let us show that $F$ has a unique strict maximum $\theta_0 \in I := (-\pi/6, \pi/6)$ (mod $2\pi$). Writing $F = \cos\theta + g$, with $g := F - \cos\theta$, one has that $F'(-\pi/6) = 1/2 + g'(\pi/6) \geqslant 1/2 - \mathtt{c} > 0$, and, similarly $F'(\pi/6) \leqslant -1/2 + \mathtt{c}$, thus $F$ has a critical point in $I$, and, since $-F'' = \cos\theta - g'' \geqslant \cos\theta - \mathtt{c} \geqslant \sqrt{3}/2 - \mathtt{c} > 0$, $F$ is strictly concave in $I$, showing that such critical point is unique and it is a strict local minimum. In fact, similarly one shows that $F$ has a second critical point $\theta_1 \in (\pi - \pi/6, \pi + \pi/6)$ where $F$ is strictly convex, so that $\theta_1$ is a strict local minimum; but, since in the complementary of these intervals $F$ is strictly monotone (as it is easy to check), it follows that $F$ has a unique global strict maximum and a unique global strict minimum. Finally, $F(\theta_0) - F(\theta_1) \geqslant \sqrt{3} - 2\mathtt{c} > 1 - 2\mathtt{c}$ and the claim follows. $\square$

From Proposition 1.1 and Lemma 1.1 one gets immediately:

**Proposition 1.2.** *Let $\delta > 0$, $f \in \mathbb{B}^n_s$ and assume* (8)*. Then, for every $k \in \mathcal{G}^n$ with $|k|_1 \geqslant \mathtt{N}$, the function $\pi_{\mathbb{Z}k} f$ is $|f_k|$–Morse.*

**Proof.** As in the proof of Proposition 1.1, we get

$$\left|\frac{\pi_{\mathbb{Z}k} f}{2 f_k} - \cos(\theta+\theta^k)\right|_1 \overset{(9)}{=} |F_\star^k|_1 \leqslant \| F_\star^k \|_1 \overset{(10)}{\leqslant} 2^{-40}, \tag{12}$$

which implies that the function $F := \pi_{\mathbb{Z}k} f/(2f_k)$ is $C^2$–close to a (shifted) cosine: Indeed, by Cauchy estimates $\|\cdot\|_{C^2} \leqslant 2|\cdot|_1$, so that

$$\|F - \cos(\theta+\theta^k)\|_{C^2} = \max_{0\leqslant j\leqslant 2} \max_{\mathbb{T}} |\partial_\theta^j (F - \cos(\theta+\theta^k))| \leqslant 2|F_\star^k|_1 \overset{(12)}{\leqslant} 2^{-39}.$$

By Lemma 1.1 we see that $F$ is $(1-2^{-38})$–Morse, and the claim follows by rescaling. □

*1.2. Genericity*

In this section we prove that $\mathbb{G}_s^n$ is a generic set in $\mathbb{B}_s^n$.

**Definition 1.6.** Given $n, s > 0$, $0 < \delta \leqslant 1$ and $\beta > 0$ and $\mathtt{N} = \mathtt{N}(\delta)$ as in (6) we call $\mathbb{G}_s^n(\delta,\beta)$ the set of functions in $\mathbb{B}_s^n$ which satisfy (8) together with:

$$\pi_{\mathbb{Z}k} f \text{ is } \beta\text{−Morse}, \qquad \forall\, k \in \mathcal{G}^n,\ |k|_1 \leqslant \mathtt{N}. \tag{13}$$

Then, the following lemma holds:

**Lemma 1.2.** *Let $n, s > 0$. Then,* $\mathbb{G}_s^n = \bigcup_{\substack{\delta\in(0,1]\\ \beta>0}} \mathbb{G}_s^n(\delta,\beta)$.

**Proof.** Assume $f \in \mathbb{G}_s^n$ and let $0 < \delta_0 \leqslant 1$ be smaller than limit inferior in (4). Then, there exists $N_0$ such that $|f_k| > \delta_0 |k|_1^{-n} e^{-|k|_1 s}$, for any $|k|_1 \geqslant N_0$, $k \in \mathcal{G}^n$. Since $\lim_{\delta\to 0} \mathtt{N} = +\infty$, there exists $0 < \delta < \delta_0$ such that $\mathtt{N} > N_0$. Hence, if $|k|_1 \geqslant \mathtt{N}$ and $k \in \mathcal{G}^n$, (8) holds.
Since $\pi_{\mathbb{Z}k} f$ is, for any $|k|_1 \leqslant \mathtt{N}$, a Morse function with distinct critical values one can, obviously, find a $\beta > 0$ for which (13) holds. Hence $f \in \mathbb{G}_s^n(\delta,\beta)$.
Now, let $f \in \bigcup \mathbb{G}_s^n(\delta,\beta)$. Then, there exist $\delta \in (0,1]$ and $\beta > 0$ such that (8) and (13) hold. Then, (4) follows immediately from (8). By Proposition 1.1, for any $k \in \mathcal{G}^n$ with $|k|_1 > \mathtt{N}$, $\pi_{\mathbb{Z}k} f$ is $2^{-40}$–cosine–like, showing (Lemma 1.1) that $\pi_{\mathbb{Z}k} f$ is Morse with distinct critical values also for $|k|_1 \geqslant \mathtt{N}$. The proof is complete. □

**Proposition 1.3.** $\mathbb{G}_s^n$ *contains an open and dense set in* $\mathbb{B}_s^n$.

To prove this result we need a preliminary elementary result on real analytic periodic functions:

**Lemma 1.3.** *Let $F = \sum F_j e^{\mathrm{i} j\theta}$ be a real analytic function on $\mathbb{T}$. There exists a compact set $\Gamma \subseteq \mathbb{C}$ (depending on $F_j$ for $|j| \geqslant 2$) of zero Lebesgue measure such that if the Fourier coefficient $F_1$ does not belong to $\Gamma$, then $F$ is a Morse function with distinct critical values.*

**Proof.** Without loss of generality we may assume that $F$ has zero average. Then, letting $z := F_1 \in \mathbb{C}$, we write $F$ as

$$F(\theta) = ze^{\mathrm{i}\theta} + \bar{z}e^{-\mathrm{i}\theta} + G(\theta) := ze^{\mathrm{i}\theta} + \bar{z}e^{-\mathrm{i}\theta} + \sum_{|j|\geqslant 2} F_j e^{\mathrm{i}j\theta}\,. \tag{14}$$

When $G \equiv 0$ the claim is true with $\Gamma = \{0\}$.
Assume that $G \not\equiv 0$. Observe that, since $G$ is real–analytic, the equations $F'(\theta) = 0 = F''(\theta)$ are equivalent to the single equation $z = \frac{1}{2}e^{-\mathrm{i}\theta}\big(\mathrm{i}G'(\theta) + G''(\theta)\big)$, which, as $\theta \in \mathbb{T}$, describes a smooth closed 'critical' curve $\Gamma_1$ in $\mathbb{C}$.
Observe also that $F$ has distinct critical points $\theta_1, \theta_2 \in \mathbb{T}$ with the same critical values if and only if the following three real equations are satisfied:

$$F'(\theta_1) = 0\,, \qquad F'(\theta_2) = 0\,, \qquad F(\theta_1) - F(\theta_2) = 0\,. \tag{15}$$

We claim that if $z, \theta_1, \theta_2$ satisfy (15) then

$$z = \zeta(\theta_1, \theta_2)\,, \qquad g(\theta_1, \theta_2) = 0\,, \tag{16}$$

with $\zeta$ and $g$ real analytic on $\mathbb{T}^2$ given by

$$\zeta(\theta_1,\theta_2) := \begin{cases} \frac{\mathrm{i}}{2(e^{\mathrm{i}\theta_1} - e^{\mathrm{i}\theta_2})}\big(G'(\theta_1) - G'(\theta_2) + \mathrm{i}G(\theta_1) - \mathrm{i}G(\theta_2)\big)\,, & \text{for } \theta_1 \neq \theta_2\,; \\ \frac{1}{2e^{\mathrm{i}\theta_1}}\big(G''(\theta_1) + \mathrm{i}G'(\theta_1)\big)\,, & \text{for } \theta_1 = \theta_2\,, \end{cases}$$
$$g(\theta_1,\theta_2) := \big(1 - \cos(\theta_1 - \theta_2)\big)\big(G'(\theta_1) + G'(\theta_2)\big) - \sin(\theta_1 - \theta_2)\big(G(\theta_1) - G(\theta_2)\big)\,.$$

Indeed, summing up the third equation in (15) with the difference of the first two equations multiplied by $-\mathrm{i}$, we get

$$2(e^{\mathrm{i}\theta_1} - e^{\mathrm{i}\theta_2})z - \mathrm{i}\big(G'(\theta_1) - G'(\theta_2) + \mathrm{i}G(\theta_1) - \mathrm{i}G(\theta_2)\big) = 0\,,$$

which is equivalent to $z = \zeta(\theta_1, \theta_2)$. Then, by definition $g(\theta_1, \theta_1) = 0$, while if $\theta_1 \neq \theta_2$, substituting $z = \zeta(\theta_1, \theta_2)$ in the first equation in (15) and multiplying by $1 - \cos(\theta_1 - \theta_2)$ we get $g(\theta_1, \theta_2) = 0$ also for $\theta_1 \neq \theta_2$. Thus, (16) holds.
Next, we claim that the real analytic function $g(\theta_1, \theta_2)$ is not identically zero. Assume by contradiction that $g$ is identically zero. Then $g(\theta_2 + t, \theta_2) \equiv 0$ for every $\theta_2$ and $t$, and taking the fourth derivative with respect to $t$ evaluated at $t = 0$, we see that $G'''(\theta_2) + G'(\theta_2) = 0$, for all $\theta_2$. The general (real) solution of the such differential equation is given by $G(\theta_2) = ce^{\mathrm{i}\theta_2} + \bar{c}e^{-\mathrm{i}\theta_2} + c_0$, with $c \in \mathbb{C}$, $c_0 \in \mathbb{R}$, which contradicts the fact that, by definition, $G_j = 0$ for $|j| \leqslant 1$. Thus, $g(\theta_1, \theta_2)$ is not identically zero and, therefore, the set $\mathcal{Z} \subseteq \mathbb{T}^2$ of its zeros is compact and has zero Lebesgue measure.[8] Clearly, also the set $\Gamma_2 := \zeta(\mathcal{Z}) \subseteq \mathbb{C}$ is compact and has zero measure, and, therefore, if we define $\Gamma = \Gamma_1 \cup \Gamma_2$, we see that the lemma holds also in the case $G \not\equiv 0$. $\square$

[8] Compare, e.g., Corollary 10, p. 9 of [20].

**Proof of Proposition 1.3.** Let $\tilde{\mathbb{G}}^n_s(\delta,\beta)$ denote the subset of functions in $\mathbb{G}^n_s(\delta,\beta)$ satisfying the (stronger) condition[9]

$$|f_k| > \delta\, e^{-|k|_1 s}\,, \qquad \forall\, k \in \mathcal{G}^n\,,\ |k|_1 \geqslant \mathtt{N} = \mathtt{N}(\delta)\,, \tag{17}$$

and let $\tilde{\mathbb{G}}^n_s = \bigcup_{\substack{0<\delta\leqslant 1\\ \beta>0}} \tilde{\mathbb{G}}^n_s(\delta,\beta)$. We claim that $\tilde{\mathbb{G}}^n_s$ is an open subset of $\mathbb{B}^n_s$. Let $f \in \tilde{\mathbb{G}}^n_s(\delta,\beta)$ for some $0<\delta\leqslant 1, \beta>0$ and let us show that there exists $0<\delta'\leqslant \delta/2$ such that if $g\in\mathbb{B}^n_s$ with $\|g-f\|_s<\delta'\leqslant\delta/2$, then $g\in\tilde{\mathbb{G}}^n_s(\delta',\beta')$ with $\beta':=\min\{\beta,\delta e^{-s\mathtt{N}(\delta/2)}\}/2$. Indeed

$$|g_k|e^{|k|_1 s} \geqslant |f_k|e^{|k|_1 s} - \|g-f\|_s > \delta-\delta' \geqslant \delta/2\,, \qquad \forall\, k\in\mathcal{G}^n\,,\ |k|_1 \geqslant \mathtt{N}(\delta)\,,$$

namely $g$ satisfies (17) with $\delta/2$ instead of $\delta$. We already know that $\pi_{\mathbb{Z}k} f$ is $\beta$–Morse $\forall\, k\in\mathcal{G}^n$, $|k|_1<\mathtt{N}(\delta)$. Moreover, by Proposition 1.2, we know that $\pi_{\mathbb{Z}k} f$ is $|f_k|$–Morse for $k\in\mathcal{G}^n$ with $|k|_1\geqslant\mathtt{N}(\delta)$. In conclusion, by (17), we get that $\pi_{\mathbb{Z}k} f$ is $2\beta'$–Morse $\forall\, k\in\mathcal{G}^n$, $|k|_1<\mathtt{N}(\delta/2)$. Since the $\|\cdot\|_s$–norm is stronger than the $C^2$–one, taking $\delta'$ small enough we get that $\pi_{\mathbb{Z}k} g$ is $\beta'$–Morse $\forall\, k\in\mathcal{G}^n$, $|k|_1<\mathtt{N}(\delta/2)$.

Let us now show that $\tilde{\mathbb{G}}^n_s$ is dense in $\mathbb{B}^n_s$. Fix $f$ in $\mathbb{B}^n_s$ and $0<\lambda<1$. We have to find $g\in\tilde{\mathbb{G}}^n_s$ with $\|g-f\|_s\leqslant\lambda$. Let $\delta:=\lambda/4$ and denote by $f_k$ and $g_k$ (to be defined) be the Fourier coefficients of, respectively, $f$ and $g$. It is enough to define $g_k$ only for $k\in\mathbb{Z}^n_\star$ since, for $k\in-\mathbb{Z}^n_\star$ we set $g_k:=\bar{g}_{-k}$, since $g$ must be real analytic. Set $g_k:=f_k$ for $k\in\mathbb{Z}^n_\star\setminus\mathcal{G}^n$. For $k\in\mathcal{G}^n$, $|k|_1\geqslant\mathtt{N}(\delta)$, we set $g_k:=f_k$ if $|f_k|e^{|k|_1 s}>\delta$ and $g_k:=2\delta e^{-|k|_1 s}$ otherwise. Consider now $k\in\mathcal{G}^n$, $|k|_1<\mathtt{N}(\delta)$. We make use of Lemma 1.3 with $F=\pi_{\mathbb{Z}k} g$, $z=F_1=g_k$. Thus, by Lemma 1.3, there exists a compact set $\Gamma_k\subseteq\mathbb{C}$ (depending on $F_k$ for $|k|\geqslant 2$) of zero measure such that if $g_k\notin\Gamma_k$ the function $\pi_{\mathbb{Z}k} g$ is a Morse function with distinct critical values. We conclude the proof of the density choosing $|g_k|<e^{-|k|_1 s}$, $|f_k-g_k|\leqslant\lambda e^{-|k|_1 s}$ with $g_k\notin\Gamma_k$. $\square$

### *1.3. Full measure*

Here we show that $\mathbb{G}^n_s$ *is a set of probability 1 with respect to the standard product probability measure on* $\mathbb{B}^n_s$. More precisely, consider the space[10] $\mathrm{D}^{\mathbb{Z}^n_\star}$, where $\mathrm{D}:=\{w\in\mathbb{C}:\ |w|\leqslant 1\}$, endowed with the product topology.[11] The product $\sigma$-algebra of the Borel sets of $\mathrm{D}^{\mathbb{Z}^n_\star}$ is the $\sigma$–algebra generated by the cylinders $\bigotimes_{k\in\mathbb{Z}^n_\star} A_k$, where $A_k$ are Borel sets of D, which differs from D only for a finite number of $k$. The probability product measure $\mu_\otimes$ on $\mathrm{D}^{\mathbb{Z}^n_\star}$ is then defined by letting

$$\mu_\otimes\Big(\bigotimes_{k\in\mathbb{Z}^n_\star} A_k\Big) := \prod_{k\in\mathbb{Z}^n_\star} |A_k|\,,$$

where $|\cdot|$ denotes the normalized ($|\mathrm{D}|=1$) Lebesgue measure on D. The (weighted) Fourier bijection[12]

[9] Here, we explicitly indicate the dependence on $\delta$, while $n$ and $s$ are fixed. Recall that $\mathtt{N}(\delta)$ is decreasing.

[10] $\mathbb{Z}^n_\star$ was defined in (2).

[11] By Tychonoff's Theorem, $\mathrm{D}^{\mathbb{Z}^n_\star}$ with the product topology is a compact Hausdorff space.

[12] $f$ is real analytic so that $f_{-k}=\bar{f}_k$.

$$\mathcal{F} : f(x) = \sum_{k \in \mathbb{Z}^n_\star} f_k e^{\mathrm{i}k\cdot x} + \bar{f}_k e^{-\mathrm{i}k\cdot x} \in \mathbb{B}^n_s \to \big\{ f_k e^{|k|_1 s} \big\}_{k \in \mathbb{Z}^n_\star} \in \ell^\infty(\mathbb{Z}^n_\star) \tag{18}$$

induces a product topology on $\mathbb{B}^n_s$ and a *probability product measure* $\mu$ on the product $\sigma$-algebra $\mathcal{B}$ of the Borellians in $\mathbb{B}^n_s = \mathcal{F}^{-1}\big(\mathrm{D}^{\mathbb{Z}^n_\star}\big)$ (with respect to the induced product topology), i.e., given $B \in \mathcal{B}$, we set $\mu(B) := \mu_\otimes(\mathcal{F}(B))$. Then one has:

**Proposition 1.4.** $\mathbb{G}^n_s \in \mathcal{B}$ *and* $\mu(\mathbb{G}^n_s) = 1$.

**Proof.** First note that, for every $\delta, \beta > 0$ the set $\mathbb{G}^n_s(\delta, \beta)$ is closed with respect to the product topology. Indeed for every $k \in \mathcal{G}^n$ the set $\{f \in \mathbb{B}^n_s \text{ s.t. } |f_k| \geqslant \delta |k|_1^{-n} e^{-|k|_1 s}\}$ is a closed cylinder. Moreover also the set $\{f \in \mathbb{B}^n_s \text{ s.t. } \pi_{\mathbb{Z}k} f \text{ is } \beta\text{–Morse}\}$ is closed w.r.t. the product topology. In fact we prove that the complementary $E := \{f \in \mathbb{B}^n_s \text{ s.t. } \pi_{\mathbb{Z}k} f \text{ is not } \beta\text{–Morse}\}$ is open w.r.t. the product topology. Indeed if $f^* \in E$ there exists a $r > 0$ small enough such that $E_r := \{f \in \mathbb{B}^n_s \text{ s.t. } \|\pi_{\mathbb{Z}k} f - \pi_{\mathbb{Z}k} f^*\|_{C^2} < r\} \subseteq E$. Define the open cylinder

$$E_{\rho,J} := \{f \in \mathbb{B}^n_s \text{ s.t. } |f_{jk} - f^*_{jk}| < \frac{\rho}{|j|_1^2} e^{-|jk|_1 s} \text{ for } j \in \mathbb{Z},\ 0 < |j|_1 \leqslant J\}.$$

We claim that $E_{\rho,J} \subseteq E_r$ for suitably small $\rho$ and large $J$ (depending on $r$ and $s$). Indeed, when $f \in E_{\rho,J}$

$$\|\pi_{\mathbb{Z}k} f - \pi_{\mathbb{Z}k} f^*\|_{C^2} \leqslant 3 \sum_{j \neq 0} |j|_1^2 |f_{jk} - f^*_{jk}| \leqslant 3\rho \sum_{0 < |j|_1 \leqslant J} e^{-|jk|_1 s} + 6 \sum_{|j|_1 > J} |j|_1^2 e^{-|jk|_1 s} < r$$

for suitably small $\rho$ and large $J$. Therefore $E_{\rho,J} \subseteq E_r \subseteq E$ and $E$ is open in the product topology. In conclusion, taking the intersection over $k \in \mathcal{G}^n$, we get that $\mathbb{G}^n_s(\delta, \beta)$ is closed with respect to the product topology.

Recalling Lemma 1.2, we note that $\mathbb{G}^n_s$ can be written as $\mathbb{G}^n_s = \bigcup_{h \in \mathbb{N}} \mathbb{G}^n_s(1/h, 1/h)$. Thus $\mathbb{G}^n_s$ is Borellian.

Let us now prove that $\mu(\mathbb{G}^n_s) = 1$. Fix $0 < \delta \leqslant 1$ and denote by $\mathbb{G}^n_s(\delta)$ the subset of functions in $\mathbb{B}^n_s$ satisfying (8) and such that $\pi_{\mathbb{Z}k} f$ is a Morse function with distinct critical values for every $k \in \mathcal{G}^n$. Recall (18) and define

$$\mathbb{P}_\delta := \mathcal{F}(\mathbb{G}^n_s(\delta)) \subseteq \ell^\infty(\mathbb{Z}^n_\star).$$

Fix $\hat{g} = (g_k)_{k \in \mathbb{Z}^n_\star \setminus \mathcal{G}^n} \in \ell^\infty(\mathbb{Z}^n_\star \setminus \mathcal{G}^n)$ with $|g_k| \leqslant 1$ for every $k \in \mathbb{Z}^n_\star \setminus \mathcal{G}^n$. Consider the section

$$\mathbb{P}^{\hat{g}}_\delta := \big\{ \check{g} = (g_k)_{k \in \mathcal{G}^n},\ |g_k| \leqslant 1 \ \text{ s.t } \ |g_k| \geqslant \delta |k|_1^{-n} \text{ if } |k|_1 \geqslant \mathtt{N},\ \ g_k e^{-|k|_1 s} \notin \Gamma_k,\ \text{if } |k|_1 < \mathtt{N} \big\},$$

where the sets $\Gamma_k$ (depending on $\hat{g}$) were defined in the proof of Proposition 1.3 so that, for every $k \in \mathcal{G}^n$, $|k|_1 < \mathtt{N}$, if $g_k e^{-|k|_1 s} \notin \Gamma_k$ then the function[13]

[13] Recall (14).

$$g_k e^{-|k|_1 s} e^{\mathrm{i}\theta} + \bar{g}_k e^{-|k|_1 s} e^{-\mathrm{i}\theta} + \sum_{|j|\geqslant 2} \hat{g}_{jk} e^{-|jk|_1 s} e^{\mathrm{i}j\theta} = \pi_{\mathbb{Z}k} f\,, \text{ with } f := \mathcal{F}^{-1}(g)\,, \ \ g = (\check{g}, \hat{g})\,,$$

is a Morse function with distinct critical values. Then, since every $\Gamma_k$ has zero measure

$$\mu_\otimes|_{\ell^\infty(\mathcal{G}^n)}(\mathbb{P}^{\hat{g}}_\delta) = \prod_{k\in\mathcal{G}^n, |k|_1 \geqslant \mathtt{N}} (1 - \delta^2 \, |k|_1^{-2n}) \geqslant 1 - c\delta^2\,,$$

for a suitable constant $c = c(n)$. Since the above estimate holds for every $\hat{g} \in \ell^\infty(\mathbb{Z}^n_\star \setminus \mathcal{G}^n)$, by Fubini's Theorem we get

$$\mu_\otimes|_{\ell^\infty(\mathcal{G}^n)}(\mathbb{P}^{\hat{g}}_\delta) = \mu_\otimes(\mathbb{P}_\delta) = \mu(\mathbb{G}^n_s(\delta)) \geqslant 1 - c\delta^2\,.$$

Then,

$$\mu(\mathbb{G}^n_s) = \lim_{\delta\to 0^+} \mu(\mathbb{G}^n_s(\delta)) = 1\,. \quad \square$$

## 2. Averaging, coverings and normal forms

In the rest of the paper we consider *real–analytic, nearly–integrable natural Hamiltonian systems*

$$\begin{cases} \dot{y} = -\mathtt{H}_x(y,x) \\ \dot{x} = \mathtt{H}_y(y,x) \end{cases}, \qquad (y,x) \in \mathbb{R}^n \times \mathbb{T}^n\,,$$

$$\mathtt{H}(y,x;\varepsilon) := \tfrac{1}{2}|y|^2 + \varepsilon f(x)\,, \qquad n \geqslant 2\,, \ 0 < \varepsilon < 1\,. \tag{19}$$

As usual, 'dot' denotes derivative with respect to 'time' $t \in \mathbb{R}$; $\mathtt{H}_y$ and $\mathtt{H}_x$ denote the gradients with respect to $y$ and $x$; $|y|^2 := y \cdot y := \sum_j |y_j|^2$; $\mathbb{T}^n$ denotes the standard flat torus $\mathbb{R}^n/(2\pi\mathbb{Z}^n)$, and the phase space $\mathbb{R}^n \times \mathbb{T}^n$ is endowed with the standard symplectic form $dy \wedge dx = \sum_j dy_j \wedge dx_j$.

In this section, we discuss the high order normal forms of generic natural systems, especially in neighborhoods of simple resonances.

As standard in perturbation theory, we consider a bounded phase space $\mathcal{M} \subseteq \mathbb{R}^n \times \mathbb{T}^n$. By translating actions and rescaling the parameter $\varepsilon$, it is not restrictive to take

$$\mathcal{M} := \mathrm{B} \times \mathbb{T}^n\,, \qquad \text{with} \qquad \mathrm{B} := B_1(0) := \{y \in \mathbb{R}^n \text{ s.t. } |y| < 1\}\,. \tag{20}$$

The first step in averaging theory is to construct suitable coverings so as to control resonances where small divisors appear. Let us recall that *a resonance* $\mathcal{R}_k$ (with respect to the free Hamiltonian $\frac{1}{2}|y|^2$) is the hyperplane $\{y \in \mathbb{R}^n : y \cdot k = 0\}$, where $k \in \mathcal{G}^n$, and its order is given by $|k|_1$; a *double resonance* $\mathcal{R}_{k,\ell}$ is the intersection of two resonances: $\mathcal{R}_{k,\ell} = \mathcal{R}_k \cap \mathcal{R}_\ell$ with $k \neq \ell$ in $\mathcal{G}^n$; the order of $\mathcal{R}_{k,\ell}$ is given by $\max\{|k|_1, |\ell|_1\}$.

**Notations.** The real or complex (open) balls of radius $r > 0$ and center $y_0 \in \mathbb{R}$ or $z_0 \in \mathbb{C}^n$ are denoted by

$$B_r(y_0) := \{y \in \mathbb{R}^n : |y - y_0| < r\}, \qquad D_r(z_0) := \{z \in \mathbb{C}^n : |z - z_0| < r\}; \tag{21}$$

if $V \subset \mathbb{R}^n$ and $r > 0$, $V_r$ denotes the complex neighborhood of $V$ given by[14]

$$V_r := \bigcup_{y \in D} D_r(y). \tag{22}$$

We shall also use the notation $\mathrm{Re}\,(V_r)$ to denote the *real* $r$–neighborhood of $V \subset \mathbb{R}^n$, namely,

$$\mathrm{Re}\,(V_r) := V_r \cap \mathbb{R}^n = \bigcup_{y \in V} B_r(y). \tag{23}$$

For a set $V \subseteq \mathbb{R}^n$ and for $r, s > 0$, given a function $f : (y, x) \in V_r \times \mathbb{T}^n_s \to f(y, x)$, we denote

$$|f|_{V,r,s} = |f|_{r,s} := \sup_{V_r \times \mathbb{T}^n_s} |f|, \qquad \| f \|_{V,r,s} = \| f \|_{r,s} := \sup_{y \in V_r} \sum_{k \in \mathbb{Z}^n} |f_k(y)| e^{|k|_1 s}, \tag{24}$$

where $f_k(y)$ denotes the $k$–th Fourier coefficient of $x \in \mathbb{T}^n \mapsto f(y, x)$; for a function depending only on $y \in V_r$ we set $|f|_{V,r} = |f|_r := \sup_{V_r} |f|$.

### 2.1. Non–resonant and simply–resonant sets

Denote by $\mathrm{p}_k$ and $\mathrm{p}_k^{\perp}$ the orhogonal projections

$$\mathrm{p}_k y := (y \cdot e_k)\, e_k, \qquad \mathrm{p}_k^{\perp} y := y - \mathrm{p}_k y, \qquad e_k := k/|k|, \tag{25}$$

and, for any $\mathtt{K} \geqslant \mathtt{K}_{\rm o} \geqslant 2$ and $\alpha > 0$, define the following sets:

$$\mathcal{R}^0 := \{y \in \mathrm{B} : \min_{k \in \mathcal{G}^n_{\mathtt{K}_{\rm o}}} |y \cdot k| > \tfrac{\alpha}{2}\}, \tag{26}$$

$$\begin{cases} \mathcal{R}^{1,k} := \left\{ y \in \mathrm{B} : |y \cdot k| < \alpha;\ |\mathrm{p}_k^{\perp} y \cdot \ell| > \frac{3\alpha \mathtt{K}}{|k|}, \forall \ell \in \mathcal{G}^n_{\mathtt{K}} \setminus \mathbb{Z}k \right\}, \quad (k \in \mathcal{G}^n_{\mathtt{K}_{\rm o}}); \\ \mathcal{R}^1 := \bigcup_{k \in \mathcal{G}^n_{\mathtt{K}_{\rm o}}} \mathcal{R}^{1,k}; \end{cases} \tag{27}$$

where, as above, $\mathrm{B} = B_1(0)$.
Eq. (26) implies that $\mathcal{R}^0$ is a $(\alpha/2)$–non–resonant set up to order $\mathtt{K}_{\rm o}$, i.e.,

$$|y \cdot k| > \frac{\alpha}{2}, \qquad \forall\, y \in \mathcal{R}^0,\ \forall\, 0 < |k| \leqslant \mathtt{K}_{\rm o}. \tag{28}$$

[14] $|u| := \sqrt{u \cdot \bar{u}}$ denotes the standard Euclidean norm on vectors $u \in \mathbb{C}^n$ (and its subspaces); 'bar', as usual, denotes complex–conjugated.

Indeed, fix $y \in \mathcal{R}^0$ and $k \in \mathbb{Z}^n$ with $0 < |k| \leqslant \mathtt{K}_{\mathtt{o}}$. Then, there exists $\bar{k} \in \mathcal{G}^n_{\mathtt{K}_{\mathtt{o}}}$ and $j \in \mathbb{Z} \setminus \{0\}$ such that $k = j\bar{k}$, so that

$$|y \cdot k| = |j|\, |\bar{k} \cdot y| \geqslant |\bar{k} \cdot y| > \alpha/2\,.$$

From (27) it follows that $\mathcal{R}^{1,k}$ is $(2\alpha\mathtt{K}/|k|)$–non resonant modulo $\mathbb{Z}k$ up to order $\mathtt{K}$, namely:

$$|y \cdot \ell| \geqslant 2\alpha\mathtt{K}/|k|\,, \quad \forall k \in \mathcal{G}^n_{\mathtt{K}_{\mathtt{o}}}\,,\ \forall y \in \mathcal{R}^{1,k}\,,\ \forall \ell \notin \mathbb{Z}k\,,\ |\ell| \leqslant \mathtt{K}\,. \tag{29}$$

Indeed, fix $y \in \mathcal{R}^{1,k}$, $k \in \mathcal{G}^n_{\mathtt{K}_{\mathtt{o}}}$, $\ell \notin \mathbb{Z}k$ with $|\ell| \leqslant \mathtt{K}$. Then, there exist $j \in \mathbb{Z} \setminus \{0\}$ and $\bar{\ell} \in \mathcal{G}^n_{\mathtt{K}}$ such that $\ell = j\bar{\ell}$. Hence,

$$\begin{aligned}|y \cdot \ell| &= |j|\,|y \cdot \bar{\ell}| \geqslant |y \cdot \bar{\ell}| = |\,\mathrm{p}_k^{\perp} y \cdot \bar{\ell} + \mathrm{p}_k y \cdot \bar{\ell}| \geqslant |\,\mathrm{p}_k^{\perp} y \cdot \bar{\ell}| - \frac{\alpha\mathtt{K}}{|k|}\\ &> \frac{3\alpha\mathtt{K}}{|k|} - \frac{\alpha\mathtt{K}}{|k|} = \frac{2\alpha\mathtt{K}}{|k|}\,.\end{aligned}$$

Relations (28) and (29) yield quantitative control on the small divisors that appear in perturbation theory allowing for *high* order averaging theory as we now proceed to show.

*Averaging*

To perform averaging, we need to introduce a few parameters (Fourier cut–offs, a small divisor threshold, radii of analyticity) and some notation.

Let

$$\begin{cases} \mathtt{K} \geqslant 6\mathtt{K}_{\mathtt{o}} \geqslant 12\,, \quad \nu := \frac{9}{2}n + 2\,, \quad \alpha := \sqrt{\varepsilon}\mathtt{K}^{\nu}\,, \quad r_{\mathtt{o}} := \frac{\alpha}{16\mathtt{K}_{\mathtt{o}}}\,, \quad r'_{\mathtt{o}} := \frac{r_{\mathtt{o}}}{2}\,,\\ s_{\mathtt{o}} := s\big(1 - \frac{1}{\mathtt{K}_{\mathtt{o}}}\big)\,,\ s'_{\mathtt{o}} := s_{\mathtt{o}}\big(1 - \frac{1}{\mathtt{K}_{\mathtt{o}}}\big)\,,\ s_\star := s\big(1 - \frac{1}{\mathtt{K}}\big)\,,\ s'_\star := s_\star\big(1 - \frac{1}{\mathtt{K}}\big)\,,\\ r_k := \alpha/|k| = \sqrt{\varepsilon}\mathtt{K}^{\nu}/|k|\,, \quad r'_k := \frac{r_k}{2}\,,\ s'_k := |k|_1 s'_\star\,, \qquad (\forall\, k \in \mathcal{G}^n_{\mathtt{K}_{\mathtt{o}}})\,. \end{cases} \tag{30}$$

**Remark 2.1.** (i) The action space B can be trivially covered by three sets as follows

$$\mathrm{B} = \mathcal{R}^0 \cup \mathcal{R}^1 \cup \mathcal{R}^2\,, \qquad \mathcal{R}^2 := \mathrm{B} \setminus (\mathcal{R}^0 \cup \mathcal{R}^1)\,.$$

As just pointed out, on the set $(\mathcal{R}^0 \cup \mathcal{R}^1) \times \mathbb{T}^n$ one can construct detailed, high order normal forms, while $\mathcal{R}^2$ is a *small set of measure of order* $\varepsilon^2\mathtt{K}^{\gamma}$ (compare (59) below).

(ii) It is important to notice that $\mathcal{R}^2$, which is a neighborhood of double resonances of order $\mathtt{K}$, is a *non perturbative set*, as pointed out in [3]. Indeed, consider for simplicity the case $n = 2$, where the only double resonance is the origin $y = 0$. Rescaling variables and time by setting $y = \sqrt{\varepsilon}\mathrm{y}$, $\mathrm{x} = x$, $\mathrm{t} = \sqrt{\varepsilon}t$, the Hamiltonian $t$–flow of $\frac{1}{2}|y|^2 + \varepsilon f(x)$ on $\{y : |y| < \varepsilon\} \times \mathbb{T}^2 \subseteq \mathcal{R}^2 \times \mathbb{T}^2$ is equivalent to the t–flow on $\{|\mathrm{y}| < 1\} \times \mathbb{T}^2$ of the Hamiltonian $\frac{1}{2}\mathrm{y}^2 + f(\mathrm{x})$, which *does not depend upon* $\varepsilon$.

Next result is based on 'refined Averaging Theory' as presented in [6]. The main technical point in this approach is the minimal loss of regularity in the angle analyticity domain and the usage of two Fourier cut–offs; for a discussion on these fine points, we refer to the Introduction in [6].

**Lemma 2.1** *(Averaging Lemma). Let* $\mathtt{H}$ *be as in* (19) *with* $f \in \mathbb{B}^n_s$ *and let* (30) *hold. There exists a constant* $\mathtt{b}_0 = \mathtt{b}_0(n,s) > 1$ *such that if* $\mathtt{K}_{\rm o} \geqslant \mathtt{b}_0$ *the following holds.*

(a) *There exists a real analytic symplectic map*

$$\Psi_{\rm o} : \mathcal{R}^0_{r'_{\rm o}} \times \mathbb{T}^n_{s'_{\rm o}} \to \mathcal{R}^0_{r_{\rm o}} \times \mathbb{T}^n_{s_{\rm o}}\,, \tag{31}$$

*such that, denoting by* $\langle\cdot\rangle$ *the average over angles* $x$*,*

$$\mathtt{H}_{\rm o}(y,x) := \big(\mathtt{H}\circ\Psi_{\rm o}\big)(y,x) = \frac{|y|^2}{2} + \varepsilon\big(g^{\rm o}(y) + f^{\rm o}(y,x)\big)\,, \quad \langle f^{\rm o}\rangle = 0\,, \tag{32}$$

*where* $g^{\rm o}$ *and* $f^{\rm o}$ *are real analytic on* $\mathcal{R}^0_{r'_{\rm o}} \times \mathbb{T}^n_{s'_{\rm o}}$ *and satisfy*

$$|g^{\rm o}|_{r'_{\rm o}} \leqslant \vartheta_{\rm o} := \frac{1}{\mathtt{K}^{6n+1}}\,, \qquad \| f^{\rm o} \|_{r'_{\rm o},s'_{\rm o}} \leqslant e^{-\mathtt{K}_{\rm o}s/3}\,. \tag{33}$$

(b) *For each* $k \in \mathcal{G}^n_{\mathtt{K}_{\rm o}}$*, there exists a real analytic symplectic map*

$$\Psi_k : \mathcal{R}^{1,k}_{r'_k} \times \mathbb{T}^n_{s'_\star} \to \mathcal{R}^{1,k}_{r_k} \times \mathbb{T}^n_{s_\star}\,, \tag{34}$$

*such that*

$$\begin{aligned} \mathtt{H}_k(y,x) &:= \big(\mathtt{H}\circ\Psi_k\big)(y,x) \\ &= \frac{|y|^2}{2} + \varepsilon\big(g^k_{\rm o}(y) + g^k(y,k\cdot x) + f^k(y,x)\big)\,, \qquad \pi_{\mathbb{Z}k} f^k = 0\,, \end{aligned} \tag{35}$$

*where* $g^k_{\rm o}$ *is real analytic on* $\mathcal{R}^{1,k}_{r'_k}$*;* $g^k(y,\cdot) \in \mathbb{B}^1_{s'_k}$ *for every* $y \in \mathcal{R}^{1,k}_{r'_k}$ *(in particular,* $\langle g^k(y,\cdot)\rangle = 0$*);* $f^k$ *is real analytic on* $\mathcal{R}^{1,k}_{r'_k} \times \mathbb{T}^n_{s'_\star}$*, and:*

$$|g^k_{\rm o}|_{r'_k} \leqslant \vartheta_{\rm o}\,, \qquad \| g^k - \pi_{\mathbb{Z}k} f \|_{r'_k,s'_k} \leqslant \vartheta_{\rm o}\,, \qquad \| f^k \|_{r'_k,\frac{s_\star}{2}} \leqslant e^{-\mathtt{K}s/3}\,. \tag{36}$$

(c) *Finally, denoting by* $\pi_y$ *and* $\pi_x$ *the projections onto, respectively, the action variable* $y$ *and the angle variable* $x$*, one has*

$$|\pi_y\Psi_{\rm o} - y|_{r'_{\rm o},s'_{\rm o}} \leqslant \tfrac{r_{\rm o}}{2^7\mathtt{K}_{\rm o}}\,, \quad |\pi_y\Psi_k - y|_{r'_k,s'_\star} \leqslant \tfrac{r_k}{2^7\mathtt{K}}\,, \tag{37}$$

*and, for every fixed* $y$*,* $\pi_x\Psi_{\rm o}(y,\cdot)$*, and* $\pi_x\Psi_k(y,\cdot)$ *are diffeomorphisms on* $\mathbb{T}^n$*.*

**Proof.** The statements follow from Theorem 6.1 in [6, p. 3553] with obvious notational changes, which we proceed to spell out. The correspondence of symbols between this paper and [6] is the following[15]:

[15] In these identities, the first symbol is the one used here, the second one is that used in [6].

$$\mathcal{R}^0=\Omega^0\,;\ \ \mathcal{R}^{1,k}=\Omega^{1,k}\,,\ \ \tfrac{|y|^2}{2}=h(y)\,;\ \ \mathtt{K}_{\rm o}=\mathtt{K}_1\,,\ \ \mathtt{K}=\mathtt{K}_2\,,$$
$$g^{\rm o}=\mathrm{g}^{\rm o}\,;\ \ f^{\rm o}=f^{\rm o}_{\star\star}\,;\ \ g^k_{\rm o}(y)+g^k(y,k\cdot x)=\mathrm{g}^k(y,x)\,;\ \ f^k=f^k_{\star\star}\,;$$

the constants $\bar{L}$ and $L$ in Definition 2.1 in [6, p. 3532] in the present case are $\bar{L}=L=1$ (since the frequency map here is the identity map); the projection $\mathrm{p}_{\mathbb{Z}k}$ introduced in [6, p. 3529] is different from the projection $\pi_{\mathbb{Z}k}$ defined here, the relation between the two being: $\pi_{\mathbb{Z}k}f(k\cdot x)=\mathrm{p}_{\mathbb{Z}k}f(x)$; finally, the norm $|\cdot|_{D,r,s}$ in [6, p. 3534] corresponds here to the norm $\|\cdot\|_{D,r,s}$, hence:

$$|g^k_{\rm o}|_{r'_k}+\|g^k-\pi_{\mathbb{Z}k}f\|_{r'_k,s'_k}=\|\mathrm{g}^k-\mathrm{p}_{\mathbb{Z}k}f\|_{D^{1,k},r_k/2,s_\star}\,.$$

Now, Assumption A in [6, p. 3533] holds. Indeed:

- the action–analyticity radii are the same as in [6] (compare (30) with Eq. (140) in [6]);
- the angle–analyticity radii defined here are the same as in Eq.s (144) and (147) in [6] (with different names);
- In [6] it is assumed that $\mathtt{K}\geqslant 3\mathtt{K}_{\rm o}\geqslant 6$ (see Eq. (139) in [6]), which in view of (30), is satisfied. Also $\nu$ in [6] is assumed to satisfy $\nu\geqslant n+2$, which in (30) is defined as $\nu=\frac{9}{2}n+2$.
- By taking $\mathtt{b}_0$ big enough condition (143) is satisfied.
- finally, to meet the smallness condition (40) in [6], namely $\varepsilon\leqslant r^2/\mathtt{K}^{2\nu}$ (where $r$ is the analyticity radius of the unperturbed Hamiltonian, which here is a free parameter), one can take $r=\mathtt{K}^\nu$ so that condition (40) in [6] becomes simply $\varepsilon\leqslant 1$.

Thus, Theorem 6.1 of [6] can be applied, and (32) and (35) are immediately recognized as, respectively, Eq.'s (145) and (148) in [6]. Since $\bar{\vartheta}$ and $\vartheta$ in Eq. 141 of [6] are of the form $c(n,s)/\mathtt{K}^{7n+1}$, we see that, by taking $\mathtt{b}_0$ big enough, they can be bounded by $\vartheta_{\rm o}=1/\mathtt{K}^{6n+1}$ in (33). Analogously, the exponential estimates on the perturbation functions in (146) and (150) of [6] are, respectively, of the form $c(n,s)\,\mathtt{K}_{\rm o}^n e^{-\mathtt{K}_{\rm o}s/2}$ and $c(n,s)\,\mathtt{K}^n e^{-\mathtt{K}s/2}$, which, by taking $\mathtt{b}_0$ big enough, can be bounded, respectively, by $e^{-\mathtt{K}_{\rm o}s/3}$ and $e^{-\mathtt{K}s/3}$ as claimed. Thus (a) and (b) are proven. Finally, from (71) and (69) in [6, p. 3541] it follows at once (37) and the injectivity of the angle maps. $\square$

For high Fourier modes, a more precise and uniform normal form can be achieved[16]:

**Lemma 2.2** *(Cosine–like Normal forms). Let* $\mathtt{H}$ *be as in* (19) *with* $f\in\mathbb{B}^n_s$ *satisfying* (8) *and let* (30) *hold. There exists a constant* $\mathbf{c_0}=\mathbf{c_0}(n,s,\delta)\geqslant\max\{\mathtt{N},\,\mathtt{b}_0\}$ *such that if* $\mathtt{K}_{\rm o}\geqslant\mathbf{c_0}$ *then the following holds. For any* $k\in\mathcal{G}^n_{\mathtt{K}_{\rm o}}$ *such that* $|k|_1\geqslant\mathtt{N}$*, then the Hamiltonian* $\mathtt{H}_k$ *in* (35) *takes the form:*

$$\mathtt{H}_k=\frac{|y|^2}{2}+\varepsilon g^k_{\rm o}(y)+2|f_k|\varepsilon\left[\cos(k\cdot x+\theta_k)+F^k_\star(k\cdot x)+g^k_\star(y,k\cdot x)+f^k_\star(y,x)\right],\qquad(38)$$

*where* $\theta_k$ *and* $F^k_\star$ *are as in Proposition 1.1 and:*

[16] This lemma should be compared with Theorem 2.1 in [6].

$$g_\star^k := \frac{1}{2|f_k|}\left(g^k - \pi_{\mathbb{Z}k} f\right), \qquad f_\star^k := \frac{1}{2|f_k|} f^k . \tag{39}$$

*Furthermore,* $g_\star^k(y,\cdot) \in \mathbb{B}_1^1$ *(for every* $y \in \mathcal{R}^{1,k}_{r'_k}$*),* $\pi_{\mathbb{Z}k} f_\star^k = 0$*, and one has:*

$$|\, g_\star^k \,|_{r'_k,1} \leqslant \vartheta := \frac{1}{\mathtt{K}^{5n}}, \qquad |\, f_\star^k \,|_{r'_k, \frac{s_\star}{2}} \leqslant e^{-\mathtt{K}s/7} . \tag{40}$$

Observe that, under the assumptions of Lemma 2.2, by (30) and (7) it is

$$\mathtt{K} \geqslant 6\mathtt{K}_{\rm o} \geqslant 6\mathtt{N} \geqslant 12\mathtt{c}_s \geqslant 12 . \tag{41}$$

**Proof.** First of all observe that the hypotheses of Lemma 2.2 imply those of Lemma 2.1 so that the results of Lemma 2.1 hold.
From (39) it follows that $g^k(y,\theta) = \pi_{\mathbb{Z}k} f(\theta) + 2|f_k| g_\star^k(y,\theta)$, which together with (9) and (35) of Lemma 2.1, implies immediately the relations (38). To prove the first estimate in (40), we observe that, since $|k|_1 \geqslant \mathtt{N}$, recalling (30) and (41) one has

$$s'_k = |k|_1 s \left(1 - \tfrac{1}{\mathtt{K}}\right)^2 > \mathtt{N}s\, \tfrac{4}{5} > 1 . \tag{42}$$

Thus, $g_\star^k(y,\cdot)$ is bounded on a 'large' angle–domain of size larger than 1 and has zero average (since $g_\star^k(y,\cdot) \in \mathbb{B}^1_{|k|_1 s'_\star}$). Now, recall the smoothing property (1) (with $N = 1$), recall that $\mathtt{K}_{\rm o} \leqslant \mathtt{K}/6$, and take $\mathbf{c_0}$ large enough. Then,

$$\begin{aligned}
|\, g_\star^k \,|_{r'_k,1} &\overset{(39)}{:=} \frac{1}{2|f_k|} \,|\, g^k - \pi_{\mathbb{Z}k} f \,|_{r'_k,1} \overset{(8)}{\leqslant} \frac{|k|_1^n e^{|k|_1 s}}{2\delta} \,|\, g^k - \pi_{\mathbb{Z}k} f \,|_{r'_k,1} \\
&\overset{(1),\,(42)}{\leqslant} \frac{|k|_1^n e^{|k|_1 s}}{2\delta} \,|\, g^k - \pi_{\mathbb{Z}k} f \,|_{r'_k, s'_k} \cdot e^{-(s'_k - 1)} \overset{(36)}{\leqslant} \frac{|k|_1^n e}{2\delta}\, \vartheta_{\rm o}\, e^{|k|_1 (s - s'_\star)} \\
&\overset{(30)}{=} \frac{|k|_1^n e}{2\delta}\, \vartheta_{\rm o}\, e^{\frac{|k|_1}{\mathtt{K}} s (2 - \frac{1}{\mathtt{K}})} \overset{(33)}{\leqslant} \frac{\mathtt{K}_{\rm o}^n e}{2\delta}\, \frac{1}{\mathtt{K}^{6n+1}}\, e^{2s \frac{\mathtt{K}_{\rm o}}{\mathtt{K}}} \leqslant \frac{1}{\mathtt{K}^{5n}} \overset{(40)}{=} \vartheta .
\end{aligned}$$

Furthermore, possibly increasing $\mathbf{c_0}$, one also has

$$\begin{aligned}
|\, f_\star^k \,|_{r'_k, \frac{s_\star}{2}} &\overset{(39)}{=} \frac{1}{2|f_k|} \,|\, f^k \,|_{r'_k, \frac{s_\star}{2}} \overset{(8)}{\leqslant} \frac{|k|_1^n e^{|k|_1 s}}{2\delta} \,|\, f^k \,|_{r'_k, \frac{s_\star}{2}} \overset{(36)}{\leqslant} \frac{|k|_1^n e^{|k|_1 s}}{2\delta}\, e^{-\frac{\mathtt{K}s}{3}} \\
&\leqslant \frac{\mathtt{K}_{\rm o}^n}{2\delta}\, e^{-\mathtt{K}s(\frac{1}{3} - \frac{\mathtt{K}_{\rm o}}{\mathtt{K}})} \leqslant \frac{\mathtt{K}^n}{2\delta \cdot 6^n}\, e^{-\mathtt{K}s/6} \leqslant e^{-\mathtt{K}s/7} . \qquad \square
\end{aligned}$$

### 2.2. *Coverings*

As mentioned in the Introduction, the averaging symplectic maps $\Psi_{\rm o}$ and $\Psi_k$ of Lemma 2.1 may displace boundaries by $\sqrt{\varepsilon}\mathtt{K}^\nu$ (compare (30) and (37)) so one cannot use the secular Hamiltonians to describe the dynamics all the way up to the boundary of $\mathrm{B} \times \mathbb{T}^n$. Such a problem – which is essential, for example, in achieving the results described at the end of the Introduction – may be overcome by introduce *a second covering*, as we proceed now to explain.

Recall the definitions of $\mathcal{R}^0$ and $\mathcal{R}^{1,k}$ in (26) and (27); recall (30), the notation in (23) and define

$$\widetilde{\mathcal{R}}^0 := \mathrm{Re}\,(\mathcal{R}^0_{r_o'/2})\,, \qquad \widetilde{\mathcal{R}}^{1,k} := \mathrm{Re}\,(\mathcal{R}^{1,k}_{r_k'/2})\,, \qquad (k \in \mathcal{G}^n_{\mathrm{K_o}})\,. \tag{43}$$

Then, the following result holds:

**Lemma 2.3.** ***(Covering Lemma)***

$$\mathcal{R}^0 \times \mathbb{T}^n \subseteq \Psi_o\big(\widetilde{\mathcal{R}}^0 \times \mathbb{T}^n\big)\,, \tag{44}$$

$$\mathcal{R}^{1,k} \times \mathbb{T}^n \subseteq \Psi_k\big(\widetilde{\mathcal{R}}^{1,k} \times \mathbb{T}^n\big)\,, \qquad \forall k \in \mathcal{G}^n_{\mathrm{K_o}}\,, \tag{45}$$

$$\mathcal{R}^2 := \mathrm{B} \setminus (\mathcal{R}^0 \cup \mathcal{R}^1) \subseteq \bigcup_{k \in \mathcal{G}^n_{\mathrm{K_o}}} \bigcup_{\substack{\ell \in \mathcal{G}^n_{\mathrm{K}} \\ \ell \notin \mathbb{Z}k}} \mathcal{R}^2_{k,\ell}\,, \tag{46}$$

*where*

$$\mathcal{R}^2_{k\ell} := \Big\{y \in \mathrm{B} : |y \cdot k| < \alpha;\ |\mathrm{p}_k^\perp y \cdot \ell| \leqslant \tfrac{3\alpha \mathrm{K}}{|k|}\Big\}\,, \qquad (k \in \mathcal{G}^n_{\mathrm{K_o}}\,,\ \ell \in \mathcal{G}^n_{\mathrm{K}} \setminus \mathbb{Z}k)\,. \tag{47}$$

**Remark 2.2.** (i) From the definition of $\mathcal{R}^2$ in (46) it follows trivially that $\{\mathcal{R}^i\}$ is a covering of B so that $\mathrm{B} = \mathcal{R}^0 \cup \mathcal{R}^1 \cup \mathcal{R}^2$.
(ii) Notice that from the definition of $\widetilde{\mathcal{R}}^{1,k}$ in (43), one has that

$$\widetilde{\mathcal{R}}^{1,k}_{r_k'/2} \subseteq \mathcal{R}^{1,k}_{r_k'}\,. \tag{48}$$

(iii) Relations (44) and (45) allow to map back the dynamics of the averaged Hamiltonians (32) and (35) so as to describe the dynamics also *arbitrarily close to the boundary of the starting phase space*.

For the proof of the Covering Lemma we shall use the following immediate consequence of the Contraction Lemma[17]:

**Lemma 2.4.** *Fix $y_0 \in \mathbb{R}^n$, $r > 0$ and let $\phi : D_{2r}(y_0) \to \mathbb{C}^n$ be a real analytic map satisfying*

$$\sup_{D_{2r}(y_0)} |\phi(y) - y| \leqslant M \tag{49}$$

*for some $0 < M < r$. Then, $y_0 \in \phi(\overline{B_r(y_0)})$.*

**Proof.** Let $V_0 := \overline{B_r(0)}$. Solving the equation $\phi(y) = y_0$ for some $y \in \overline{B_r(y_0)}$ is equivalent to solve the fixed point equation $w = \psi_0(w) := -\psi(y_0 + w)$ for $w \in V_0$ having set $\psi(y) := \phi(y) - y$. By (49) it follows that $\psi_0 : V_0 \to V_0$ and by the mean value theorem and Cauchy estimates we get that, for every $w, w' \in V_0$,

[17] Recall the definitions in (21); as usual $\overline{A}$ denotes the closure of the set $A$.

$$|\psi_0(w) - \psi_0(w')| = |\psi(y_0 + w) - \psi(y_0 + w')| \leqslant \frac{M}{r}|w - w'|\,,$$

showing that $\psi_0$ is a contraction on $V_0$ (since $M/r < 1$) and the claim follows by the standard Contraction Lemma. $\square$

**Proof of (44).** We start by proving that

$$\forall\ (y_0, x) \in \mathcal{R}^0 \times \mathbb{T}^n\,,\ \exists!\ (y, x_0) \in \widetilde{\mathcal{R}}^0 \times \mathbb{T}^n:\ \ \Psi_{\rm o}(y, x) = (y_0, x_0)\,. \tag{50}$$

Define

$$M := \frac{r_{\rm o}}{2^7 {\rm K}_{\rm o}} \overset{(30)}{=} \frac{\alpha}{2^{11}{\rm K}_{\rm o}^2} < \frac{\alpha}{2^{10}{\rm K}_{\rm o}^2} =: r < \frac{\alpha}{2^7 {\rm K}_{\rm o}} \overset{(30)}{=} \frac{r'_{\rm o}}{4}\,. \tag{51}$$

Fix $(y_0, x) \in \mathcal{R}^0 \times \mathbb{T}^n$ and let $\phi(y) := \pi_y \Psi_{\rm o}(y, x)$. Then, by (51),

$$\sup_{D_{2r}(y_0)} |\phi(y) - y| \leqslant \sup_{D_{r'_{\rm o}}(y_0)} |\phi(y) - y| \leqslant |\pi_y \Psi_{\rm o} - y|_{r'_{\rm o}, s'_{\rm o}} \overset{(37)}{\leqslant} M\,.$$

Thus, by Lemma 2.4, since by (51) $2r < r'_{\rm o}/2$, by definition of $\widetilde{\mathcal{R}}^0$, we have that

$$y_0 \in \pi_y \Psi_{\rm o}\big(\overline{B_r(y_0)} \times \{x\}\big) \subseteq \pi_y \Psi_{\rm o}\big(\widetilde{\mathcal{R}}^0 \times \{x\}\big)\,,$$

which implies that $\Psi_{\rm o}(y, x) = (y_0, x_0)$ with $x_0 \in \mathbb{T}^n$ proving (50). Now, observe that the map $(y_0, x) \in \mathcal{R}^0 \times \mathbb{T}^n \mapsto (y, x_0) \in \widetilde{\mathcal{R}}^0 \times \mathbb{T}^n$ in (50) is nothing else than the diffeomorphism associated to the near–to–identity generating function $y_0 \cdot x + \psi_0(y_0, x)$ of the near–to–identity symplectomorphism $\Psi_{\rm o}$. Thus, for each $y_0 \in \mathcal{R}^0$, the map $x \in \mathbb{T}^n \mapsto x_0 = x + \partial_{y_0}\psi_0(y_0, x)$ is a diffeomorphism of $\mathbb{T}^n$ with inverse given by $x_0 \in \mathbb{T}^n \mapsto x = x_0 + \chi(y_0, x_0)$ for a suitable (small) real analytic map $\chi$. Therefore, given $(y_0, x_0) \in \mathcal{R}^0 \times \mathbb{T}^n$, if we take $x = x_0 + \chi(y_0, x_0)$ in (50) we obtain that there exist $(y, x) \in \widetilde{\mathcal{R}}^0 \times \mathbb{T}^n$ such that $(y_0, x_0) = \Psi_0(y, x)$, proving (44). $\square$

**Proof of (45).** The argument is completely analogous: Again, we start by proving that

$$\forall\ k \in \mathcal{G}^n_{{\rm K}_{\rm o}}\,,\ \forall\ (y_0, x) \in \mathcal{R}^{1,k} \times \mathbb{T}^n\,,\ \exists!\ (y, x_0) \in \widetilde{\mathcal{R}}^{1,k} \times \mathbb{T}^n:\ \ \Psi_k(y, x) = (y_0, x_0)\,. \tag{52}$$

Fix $k \in \mathcal{G}^n_{{\rm K}_{\rm o}}$ and define

$$M := \frac{r_k}{2^7 {\rm K}_{\rm o}} \overset{(30)}{=} \frac{\alpha}{2^7 |k|\, {\rm K}} < \frac{\alpha}{2^6 |k| {\rm K}} =: r < \frac{r'_k}{4} \overset{(30)}{=} \frac{\alpha}{8|k|}\,. \tag{53}$$

Fix $(y_0, x) \in \mathcal{R}^{1,k} \times \mathbb{T}^n$, and let $\phi(y) := \pi_y \Psi_k(y, x)$. By (53),

$$\sup_{D_{2r}(y_0)} |\phi(y) - y| \leqslant \sup_{D_{r'_k}(y_0)} |\phi(y) - y| \leqslant |\pi_y \Psi_k - y|_{r'_k, s_\star} \overset{(37)}{\leqslant} M\,.$$

Thus, by Lemma 2.4 we have

$$y_0 \in \pi_y \Psi_k \big(\overline{B_r(y_0)} \times \{x\}\big) \subseteq \pi_y \Psi_k \big(\widetilde{\mathcal{R}}^{1,k} \times \{x\}\big)\,,$$

which implies that $\Psi_k(y, x) = (y_0, x_0)$ for some $x_0 \in \mathbb{T}^n$ proving (52). Now, the argument given in the non–resonant case applies also in this case. □

**Proof of (46).** If $y \in \mathcal{R}^2$ then, since $y \notin \mathcal{R}^0$, there exists $k \in \mathcal{G}^n_{\mathrm{K}_{\mathrm{o}}}$ such that $|y \cdot k| < \alpha$, in which case, since $y \notin \mathcal{R}^1$, there exists $\ell \in \mathcal{G}^n_{\mathrm{K}} \setminus \mathbb{Z}k$ such that $|\,\mathrm{p}^\perp_k y \cdot \ell| \leqslant \frac{3\alpha \mathrm{K}}{|k|}$, hence $y \in \mathcal{R}^2_{k,\ell}$ for some $k \in \mathcal{G}^n_{\mathrm{K}_{\mathrm{o}}}$ and $\ell \in \mathcal{G}^n_{\mathrm{K}} \setminus \mathbb{Z}k$. □

Next, we show that the measure of $\mathcal{R}^2$ is proportional to[18] $\alpha^2$:

**Lemma 2.5.** *There exists a constant $c_\star = c_\star(n) > 1$ such that:*

$$\mathrm{meas}\,(\mathcal{R}^2) \leqslant c_\star\, \alpha^2\, \mathrm{K}^{2n}\,. \tag{54}$$

**Proof.** Let us estimate the measure of $\mathcal{R}^2_{k,\ell}$ in (47). Denote by $v \in \mathbb{R}^n$ the projection of $y$ onto the plane generated by $k$ and $\ell$ (recall that, by hypothesis, $k$ and $\ell$ are not parallel). Then,

$$|v \cdot k| = |y \cdot k| < \alpha\,, \qquad |\,\mathrm{p}^\perp_k v \cdot \ell| = |\,\mathrm{p}^\perp_k y \cdot \ell| \leqslant 3\alpha \mathrm{K}/|k|\,. \tag{55}$$

Set

$$h := \mathrm{p}_k \ell = \ell - \frac{\ell \cdot k}{|k|^2} k\,. \tag{56}$$

Then, $v$ decomposes in a unique way as $v = ak + bh$ for suitable $a, b \in \mathbb{R}$. By (55),

$$|a| < \frac{\alpha}{|k|^2}\,, \qquad |\,\mathrm{p}_k v \cdot \ell| = |bh \cdot \ell| \leqslant 3\alpha \mathrm{K}/|k|\,, \tag{57}$$

and, since $|\ell|^2|k|^2 - (\ell \cdot k)^2$ is a positive integer (recall, that $k$ and $\ell$ are integer vectors not parallel),

$$|h \cdot \ell| \overset{(56)}{=} \frac{|\ell|^2|k|^2 - (\ell \cdot k)^2}{|k|^2} \geqslant \frac{1}{|k|^2}\,.$$

Hence,

$$|b| \leqslant 3\alpha \mathrm{K}|k|\,. \tag{58}$$

Then, write $y \in \mathcal{R}^2_{k,\ell}$ as $y = v + v^\perp$ with $v^\perp$ in the orthogonal complement of the plane generated by $k$ and $\ell$. Since $|v^\perp| \leqslant |y| < 1$ and $v$ lies in the plane spanned by $k$ and $\ell$ inside a rectangle of sizes of length $2\alpha/|k|^2$ and $6\alpha \mathrm{K}|k|$ (compare (57) and (58)) we find

[18] A similar result can be found in [6, p. 3533].

$$\text{meas}\,(\mathcal{R}^2_{k,\ell}) \leqslant \tfrac{2\alpha}{|k|^2}\,(6\alpha \mathrm{K}|k|)\,2^{n-2} = 3\cdot 2^n\,\alpha^2\,\tfrac{\mathrm{K}}{|k|}\,, \qquad \forall \begin{cases} k \in \mathcal{G}^n_{\mathrm{K}_o}\,, \\ \ell \in \mathcal{G}^n_{\mathrm{K}} \setminus \mathbb{Z}k\,. \end{cases}$$

Since $\sum_{k\in\mathcal{G}^n_{\mathrm{K}_o}} |k|^{-1} \leqslant c\,\mathrm{K}_o^{n-1}$ for a suitable $c = c(n)$, and $\mathrm{K}_o \leqslant \mathrm{K}/6$, (54) follows. □

**Remark 2.3.** In view of (54) and (30), we have

$$\text{meas}\,(\mathcal{R}^2) \leqslant c_\star\,\varepsilon\,\mathrm{K}^\gamma\,, \qquad \gamma := 11n + 4\,. \tag{59}$$

Thus, if $\mathrm{V}_n = \pi^{\frac{n}{2}}/\Gamma(1+\frac{n}{2})$ denotes the volume of the Euclidean unit ball B in $\mathbb{R}^n$, we have that

$$\varepsilon < \frac{\mathrm{V}_n}{c_\star\,\mathrm{K}^\gamma} \qquad \Longrightarrow \qquad \text{meas}\,(\mathcal{R}^2) < \text{meas}\,\mathrm{B}\,. \tag{60}$$

### *2.3. Normal form theorem*

In the normal form around simple resonances the 'averaged Hamiltonian' in (35) (i.e., the Hamiltonian obtained disregarding the exponentially small term $f^k$) depends on angles through the linear combination $k\cdot x$, which, since $k \in \mathcal{G}^n$ defines *a new well–defined angle* $\mathrm{x}_1 \in \mathbb{T}$. This fact calls for a linear symplectic change of variables:

**Lemma 2.6.** *Let the hypotheses of Lemma 2.2 hold.*
(i) *For any $k \in \mathcal{G}^n_{\mathrm{K}_o}$ there exists a matrix $\hat{\mathrm{A}} \in \mathbb{Z}^{(n-1)\times n}$ such that*[19]

$$\begin{gathered} \mathrm{A} := \begin{pmatrix} k \\ \hat{\mathrm{A}} \end{pmatrix} = \begin{pmatrix} k_1 \cdots k_n \\ \hat{\mathrm{A}} \end{pmatrix} \in \ \mathrm{SL}(n,\mathbb{Z})\,, \\ |\hat{\mathrm{A}}|_\infty \leqslant |k|_\infty\,,\ \ |\mathrm{A}|_\infty = |k|_\infty\,,\ \ |\mathrm{A}^{-1}|_\infty \leqslant (n-1)^{\frac{n-1}{2}}\,|k|_\infty^{n-1}\,. \end{gathered} \tag{61}$$

(ii) *Let $\Phi_0$ be the linear, symplectic map on $\mathbb{R}^n \times \mathbb{T}^n$ onto itself defined by*

$$\Phi_0 : (\mathrm{y}, \mathrm{x}) \mapsto (y, x) = (\mathrm{A}^T \mathrm{y}, \mathrm{A}^{-1}\mathrm{x})\,. \tag{62}$$

*Then,*

$$\mathrm{x}_1 = k\cdot x\,, \qquad y = \mathrm{y}_1 k + \hat{\mathrm{A}}^T \hat{\mathrm{y}}\,, \qquad \big[\hat{\mathrm{y}} := (\mathrm{y}_2, \ldots, \mathrm{y}_n)\big]\,. \tag{63}$$

*Furthermore, letting*[20]

$$\mathscr{D}^k := \mathrm{A}^{-T}\widetilde{\mathcal{R}}^{1,k}\,, \qquad \begin{cases} \tilde{r}_k := \frac{r_k}{c_1 |k|} \\ \tilde{s}_k := \frac{s}{c_1 |k|^{n-1}} \end{cases}\,, \qquad c_1 := 5n(n-1)^{\frac{n-1}{2}}\,, \tag{64}$$

[19] $\mathrm{SL}(n,\mathbb{Z})$ denotes the group of $n\times n$ matrices with entries in $\mathbb{Z}$ and determinant 1; $|M|_\infty$, with $M$ matrix (or vector), denotes the maximum norm $\max_{ij} |M_{ij}|$ (or $\max_i |M_i|$).
[20] $\widetilde{\mathcal{R}}^{1,k}$ is defined in (43); recall, also, (30).

*with* A *as in* (i), *we find*

$$\Phi_0 : \mathscr{D}^k_{\tilde r_k} \times \mathbb{T}^n_{\tilde s_k} \to \widetilde{\mathcal{R}}^{1,k}_{r'_k/2} \times \mathbb{T}^n_{s_\star/2}\,, \qquad \Phi_0(\mathscr{D}^k \times \mathbb{T}^n) = \widetilde{\mathcal{R}}^{1,k} \times \mathbb{T}^n\,. \tag{65}$$

(iii) $\mathrm{H}_k$ *in* (35), *in the symplectic variables* $(\mathrm{y}, \mathrm{x}) = \big((\mathrm{y}_1, \hat{\mathrm{y}}), \mathrm{x}\big)$, *takes the form:*

$$\mathcal{H}_k(\mathrm{y}, \mathrm{x}) := \mathrm{H}_k \circ \Phi_0(\mathrm{y}, \mathrm{x}) = \overline{\mathrm{H}}_k(\mathrm{y}, \mathrm{x}_1) + \varepsilon \bar f^k(\mathrm{y}, \mathrm{x})\,, \quad (\mathrm{y}, \mathrm{x}) \in \mathscr{D}^k_{\tilde r_k} \times \mathbb{T}^n_{\tilde s_k}\,, \tag{66}$$

*where the 'secular Hamiltonian'*

$$\overline{\mathrm{H}}_k(\mathrm{y}, \mathrm{x}_1) := \frac{1}{2}|\mathrm{A}^T\mathrm{y}|^2 + \varepsilon g^k_0(\mathrm{A}^T\mathrm{y}) + \varepsilon g^k(\mathrm{A}^T\mathrm{y}, \mathrm{x}_1)\,, \quad \bar f^k(\mathrm{y}, \mathrm{x}) := f^k(\mathrm{A}^T\mathrm{y}, \mathrm{A}^{-1}\mathrm{x}) \tag{67}$$

*is a real analytic function for* $\mathrm{y} \in \mathscr{D}^k_{\tilde r_k}$ *and*[21] $\mathrm{x}_1 \in \mathbb{T}_{s'_k}$.

**Remark 2.4.** In the above Lemma 2.6 (and also often in what follows), to simplify symbols, we may omit the dependence upon $k$ in the notation, but of course A, $\hat{\mathrm{A}}$ and $\Phi_0$ *do depend upon the simple resonance label* $k \in \mathcal{G}^n_{\mathrm{K}_o}$.

**Proof of Lemma 2.6.** (i) From Bézout's lemma it follows that[22]:
*given* $k \in \mathbb{Z}^n$, $k \neq 0$ *there exists a matrix* $\mathrm{A} = (\mathrm{A}_{ij})_{1\leqslant i,j\leqslant n}$ *with integer entries such that* $A_{nj} = k_j$ $\forall\, 1 \leqslant j \leqslant n$, $\det \mathrm{A} = \gcd(k_1, ..., k_1)$*, and* $|\mathrm{A}|_\infty = |k|_\infty$.
Hence, since $k \in \mathcal{G}^n$, $\gcd(k_1, ..., k_1) = 1$, and (61) follows.[23]
(ii) $\Phi_0$ is symplectic since it is generated by the generating function $\mathrm{y} \cdot \mathrm{A}x$.
The relations in (63) follow at once from the definition of $\Phi_0$.
Let us prove (65): $\mathrm{y} \in \mathscr{D}^k_{\tilde r_k}$ if and only if $\mathrm{y} = \mathrm{y}_0 + z$ with $\mathrm{y}_0 \in \mathscr{D}^k$ and $|z| < \tilde r_k$. Thus,

$$|\mathrm{A}^T z| \overset{(61)}{\leqslant} n|k||z| < n|k|\tilde r_k \overset{(64)}{<} \frac{r_k}{4} \overset{(30)}{=} \frac{r'_k}{2}\,.$$

Since, by definition of $\mathscr{D}^k$, $\mathrm{A}^T\mathrm{y}_0 \in \widetilde{\mathcal{R}}^{1,k}$, we have that $\mathrm{A}^T\mathrm{y} \in \widetilde{\mathcal{R}}^{1,k}_{r'_k/2}$.
Let, now, x belong to $\mathbb{T}^n_{\tilde s_k}$. Then, for any $1 \leqslant j \leqslant n$, recalling the definitions of $s_\star$ and $s'_\star$ in (30), we find

$$\Big|\mathrm{Im}\,(\mathrm{A}^{-1}\mathrm{x})_j\Big| = \Big|\sum_{i=1}^n (\mathrm{A}^{-1})_{ij}\,\mathrm{Im}\,\mathrm{x}_j\Big| \overset{(61)}{<} n(n-1)^{\frac{n-1}{2}}|k|^{n-1}\tilde s_k \overset{(64)}{\leqslant} \frac{s_\star}{2} < s'_\star\,.$$

Thus, $\mathrm{A}^{-1}\mathrm{x}$ belong to $\mathbb{T}^n_{s'_\star}$, and (65) follows.
(iii) Eq.'s (66)–(67) follow immediately from the definition of the symplectic map $\Phi_0$ in (62) and (63). The statement on the angle–analyticity domain of $\overline{\mathrm{H}}_k$ follows from part (b) of Lemma 2.1. □

[21] Recall (30).
[22] See Appendix A of [6, p. 3564] for a detailed proof.
[23] Notice that the bound on $|\mathrm{A}^{-1}|_\infty$ follows from D'Alembert expansion of determinants, observing that for any $m \times m$ matrix M, one has $|\det \mathrm{M}| \leqslant m^{m/2}|\mathrm{M}|^m_\infty$.

We summarize the above lemmata in the following

**Theorem 2.1** *(Normal Form Theorem). Let* $\mathtt{H}$ *be as in* (19) *with* $f \in \mathbb{B}^n_s$ *satisfying* (8) *with* $\mathtt{N}$ *as in* (6)*, and let* (30) *hold. There exists a constant*[24] $\mathbf{c_0} = \mathbf{c_0}(n, s, \delta) \geqslant \max\{\mathtt{N}, \mathtt{b}_0\}$ *such that, if* $\mathtt{K}_{\mathrm{o}} \geqslant \mathbf{c_0}$, $k \in \mathcal{G}^n_{\mathtt{K}_{\mathrm{o}}}$*, and* $\mathscr{D}^k$, $\tilde{r}_k$, $\tilde{s}_k$ *are as in* (64)*, then there exist real analytic symplectic maps*

$$\Psi_{\mathrm{o}} : \mathcal{R}^0_{r'_{\mathrm{o}}} \times \mathbb{T}^n_{s'_{\mathrm{o}}} \to \mathcal{R}^0_{r_{\mathrm{o}}} \times \mathbb{T}^n_{s_{\mathrm{o}}}\,, \qquad \Psi^k : \mathscr{D}^k_{\tilde{r}_k} \times \mathbb{T}^n_{\tilde{s}_k} \to \mathcal{R}^{1,k}_{r_k} \times \mathbb{T}^n_{s_\star} \tag{68}$$

*having the following properties.*

(i) $\mathtt{H}_{\mathrm{o}}(y,x) := \big(\mathtt{H} \circ \Psi_{\mathrm{o}}\big)(y,x) = \frac{|y|^2}{2} + \varepsilon\big(g^{\mathrm{o}}(y) + f^{\mathrm{o}}(y,x)\big)$*, with* $g^{\mathrm{o}}$ *and* $f^{\mathrm{o}}$ *satisfying* (33) *and* $\langle f^{\mathrm{o}}\rangle = 0$.

(ii)

$$\mathcal{H}_k(\mathrm{y},\mathrm{x}) := \mathtt{H} \circ \Psi^k(\mathrm{y},\mathrm{x}) = \overline{\mathtt{H}}_k(\mathrm{y},\mathrm{x}_1) + \varepsilon \bar{f}^k(\mathrm{y},\mathrm{x})\,, \quad (\mathrm{y},\mathrm{x}) \in \mathscr{D}^k_{\tilde{r}_k} \times \mathbb{T}^n_{\tilde{s}_k}\,, \tag{69}$$

*where*

$$\overline{\mathtt{H}}_k(\mathrm{y},\mathrm{x}_1) := \frac{1}{2}|\mathrm{A}^T\mathrm{y}|^2 + \varepsilon \mathtt{g}^k_{\mathrm{o}}(\mathrm{y}) + \varepsilon \mathtt{g}^k(\mathrm{y},\mathrm{x}_1) \tag{70}$$

*is a real analytic function for* $\mathrm{y} \in \mathscr{D}^k_{\tilde{r}_k}$ *and* $\mathrm{x}_1 \in \mathbb{T}_{s'_k}$*. In particular* $\mathtt{g}^k(y,\cdot) \in \mathbb{B}^1_{s'_k}$ *for every* $y \in \mathscr{D}^k_{\tilde{r}_k}$*. Furthermore, the following estimates hold:*

$$|\mathtt{g}^k_{\mathrm{o}}|_{\tilde{r}_k} \leqslant \vartheta_{\mathrm{o}} = \frac{1}{\mathtt{K}^{6n+1}}\,, \qquad |\mathtt{g}^k - \pi_{\mathbb{Z}k} f\,|_{\tilde{r}_k, s'_k} \leqslant \vartheta_{\mathrm{o}}\,, \qquad |\,\bar{f}^k\,|_{\tilde{r}_k,\tilde{s}_k} \leqslant e^{-\mathtt{K}s/3}\,. \tag{71}$$

(iii) *If* $|k|_1 \geqslant \mathtt{N}$*, there exists* $\theta_k \in [0, 2\pi)$ *such that*

$$\mathcal{H}_k = \frac{1}{2}|\mathrm{A}^T\mathrm{y}|^2 + \varepsilon \mathtt{g}^k_{\mathrm{o}}(\mathrm{y}) + 2|f_k|\varepsilon\left[\cos(\mathrm{x}_1 + \theta_k) + F^k_\star(\mathrm{x}_1) + \mathtt{g}^k_\star(\mathrm{y},\mathrm{x}_1) + \mathtt{f}^k_\star(\mathrm{y},\mathrm{x})\right], \tag{72}$$

*where* $F^k_\star$ *is as in Proposition* 1.1 *and satisfies* $F^k_\star \in \mathbb{B}^1_1$ *and* $|F^k_\star|_1 \leqslant 2^{-40}$.
*Moreover,* $\mathtt{g}^k_\star(y,\cdot) \in \mathbb{B}^1_1$ *(for every* $y \in \mathscr{D}^k_{\tilde{r}_k}$*),* $\pi_{\mathbb{Z}k}\mathtt{f}^k_\star = 0$*, and one has*

$$|\,\mathtt{g}^k_\star\,|_{\tilde{r}_k,1} \leqslant \vartheta = \tfrac{1}{\mathtt{K}^{5n}}\,, \qquad |\,\mathtt{f}^k_\star\,|_{\tilde{r}_k,\tilde{s}_k} \leqslant e^{-\mathtt{K}s/7}\,. \tag{73}$$

**Proof.** The first relation in (68) is (31). Define

$$\Psi^k := \Psi_k \circ \Phi_0\,. \tag{74}$$

Then, since $s_\star/2 < s'_\star$ (compare (30)), by (65), (48) we get the second relation in (68).
(i) follows from point (a) of Lemma 2.1.

[24] $\mathtt{b}_0$ is defined in Lemma 2.1.

(ii) (69), (70) and (71) follow from, respectively, (66), (67), (36) and point (ii) of Lemma 2.6 setting

$$\mathtt{g}_{\mathrm{o}}^k(\mathtt{y}) := g_{\mathrm{o}}^k(\mathrm{A}^T\mathtt{y}), \qquad \mathtt{g}^k(\mathtt{y},\mathtt{x}_1) := g^k(\mathrm{A}^T\mathtt{y},\mathtt{x}_1). \tag{75}$$

(iii) follows by Proposition 1.1 and Lemma 2.2. In particular (72) follows from (38). Furthermore,

$$\mathtt{g}_\star^k := \frac{1}{2|f_k|}\big(\mathtt{g}^k - \pi_{\mathbb{Z}k} f\big), \qquad \mathtt{f}_\star^k := \frac{1}{2|f_k|}\bar{f}^k \tag{76}$$

and noting that $\mathtt{g}_\star^k(\mathtt{y},\mathtt{x}_1) = g_\star^k(\mathrm{A}^T\mathtt{y},\mathtt{x}_1)$ and that, by (67), $\mathtt{f}_\star^k(\mathtt{y},\mathtt{x}) = f_\star^k(\mathrm{A}^T\mathtt{y},\mathrm{A}^{-1}\mathtt{x})$, we see that (73) follows from (40) and (65). $\square$

## 3. Generic standard form at simple resonances

In this final section we show that *the secular Hamiltonians* $\overline{\mathrm{H}}_k$ (67) *in Theorem 2.1 can be symplectically put into a suitable standard form, uniformly in* $k \in \mathcal{G}^n_{\mathrm{K}_\mathrm{o}}$

The precise definition of 'standard form' is taken from [5], where the analytic properties of action–angle variables of such Hamiltonian systems are discussed.

**Definition 3.1.** Let $\hat{D} \subseteq \mathbb{R}^{n-1}$ be a bounded domain, $\mathtt{R} > 0$ and $D := (-\mathtt{R},\mathtt{R}) \times \hat{D}$. We say that the real analytic Hamiltonian $\mathrm{H}_\flat$ is in Generic Standard Form with respect to the symplectic variables $(p_1, q_1) \in (-\mathtt{R},\mathtt{R}) \times \mathbb{T}$ and 'external actions'

$$\hat{p} = (p_2, ..., p_n) \in \hat{D}$$

if $\mathrm{H}_\flat$ has the form

$$\mathrm{H}_\flat(p, q_1) = \big(1 + \nu(p, q_1)\big)p_1^2 + \mathtt{G}(\hat{p}, q_1), \tag{77}$$

where:

- $\nu$ and $\mathtt{G}$ are real analytic functions defined on, respectively, $D_\mathtt{r} \times \mathbb{T}_\mathtt{s}$ and $\hat{D}_\mathtt{r} \times \mathbb{T}_\mathtt{s}$ for some $0 < \mathtt{r} \leqslant \mathtt{R}$ and $\mathtt{s} > 0$;
- $\mathtt{G}$ has zero average and there exists a function $\bar{\mathtt{G}}$ (the 'reference potential') depending only on $q_1$ such that, for some[25] $\beta > 0$,

$$\bar{\mathtt{G}} \text{ is } \beta\text{–Morse}, \qquad \langle \bar{\mathtt{G}} \rangle = 0; \tag{78}$$

[25] Recall Definition 1.3.

- the following estimates hold:

$$\begin{cases} \sup\limits_{\mathbb{T}^1_{\mathtt{s}}} |\bar{\mathtt{G}}| \leqslant \epsilon\,, \\ \sup\limits_{\hat{D}_{\mathtt{r}} \times \mathbb{T}^1_{\mathtt{s}}} |\mathtt{G} - \bar{\mathtt{G}}| \leqslant \epsilon\mu\,, \quad \text{for some} \quad 0 < \epsilon \leqslant \mathtt{r}^2/2^{16}\,, \ \ 0 \leqslant \mu < 1\,, \\ \sup\limits_{D_{\mathtt{r}} \times \mathbb{T}^1_{\mathtt{s}}} |\nu| \leqslant \mu\,. \end{cases} \tag{79}$$

We shall call $(\hat{D}, \mathtt{R}, \mathtt{r}, \mathtt{s}, \beta, \epsilon, \mu)$ *the analyticity characteristics of* $\mathtt{H}_\flat$ *with respect to the unperturbed potential* $\bar{\mathtt{G}}$.

**Remark 3.1.** If $\mathtt{H}_\flat$ is in Generic Standard Form, then the parameters $\beta$ and $\epsilon$ satisfy the relation[26]

$$\tfrac{\epsilon}{\beta} \geqslant \tfrac{1}{2}\,. \tag{80}$$

Furthermore, one can always fix $\kappa \geqslant 4$ such that:

$$\tfrac{1}{\kappa} \leqslant \mathtt{s} \leqslant 1\,, \qquad 1 \leqslant \tfrac{\mathtt{R}}{\mathtt{r}} \leqslant \kappa\,, \qquad \tfrac{1}{2} \leqslant \tfrac{\epsilon}{\beta} \leqslant \kappa\,. \tag{81}$$

Such a parameter $\kappa$ rules the main scaling properties of these Hamiltonians.

### *3.1. Main theorem*

In the following we shall often use the following notation: If $w$ is a vector with $n$ or $2n$ components, $\hat{w} = (w)\widehat{\ }$ denotes the last $(n-1)$ components; if $w$ is vector with $2n$ components, $\check{w} = (w)\check{\ }$ denotes the first $n+1$ components. Explicitly:

$$w = (y, x) = \big((y_1, ..., y_n), (x_1, ... x_n)\big) \quad \Longrightarrow \quad \begin{cases} \hat{w} = (w)\widehat{\ } = (x_2, ..., x_n) = \hat{x}\,, \\ \hat{y} = (y)\widehat{\ } = (y_2, ..., y_n)\,, \\ \check{w} = (w)\check{\ } = (y, x_1)\,, \\ w = (\check{w}, \hat{w})\,. \end{cases} \tag{82}$$

**Definition 3.2.** Given a domain $\hat{\mathrm{D}} \subseteq \mathbb{R}^{n-1}$, we denote by $\mathfrak{G}_\dagger$ the abelian group of symplectic diffeomorphisms $\Psi_{\mathtt{g}}$ of $(\mathbb{R} \times \hat{\mathrm{D}}) \times \mathbb{R}^n$ given by

$$(p, q) \in (\mathbb{R} \times \hat{\mathrm{D}}) \times \mathbb{R}^n \overset{\Psi_{\mathtt{g}}}{\mapsto} (P, Q) = (p_1 + \mathtt{g}(\hat{p}), \hat{q}, q_1, \hat{q} - q_1 \partial_{\hat{p}} \mathtt{g}(\hat{p})) \in \mathbb{R}^{2n}\,,$$

with $\mathtt{g} : \hat{\mathrm{D}} \to \mathbb{R}$ smooth.

**Remark 3.2.** The group properties of $\mathfrak{G}_\dagger$ are trivial:

$$\mathrm{id}_{\mathfrak{G}_\dagger} = \Psi_0\,, \qquad \Psi_{\mathtt{g}}^{-1} = \Psi_{-\mathtt{g}}\,, \qquad \Psi_{\mathtt{g}} \circ \Psi_{\mathtt{g}'} = \Psi_{\mathtt{g}+\mathtt{g}'}\,.$$

[26] By (79), $\beta \leqslant |\bar{\mathtt{G}}(\theta_i) - \bar{\mathtt{G}}(\theta_i)| \leqslant 2 \max_{\mathbb{T}} |\bar{\mathtt{G}}| \leqslant 2\epsilon$.

Notice, however, that, unless $\partial_{\hat p}\mathtt{g} \in \mathbb{Z}^{n-1}$, maps in $\Psi_{\mathtt{g}} \in \mathfrak{G}_{\dagger}$ *do not induce well defined angle maps* $q \in \mathbb{T}^n \mapsto (q_1, \hat q - q_1\partial_{\hat p}\mathtt{g}(\hat p)) \in \mathbb{T}^n$.

Now, let $f \in \mathbb{G}^n_s$ satisfy[27] (8) and (13) for some $0 < \delta \leqslant 1$ and $\beta > 0$ with $\mathtt{N}$ defined in (6), let $k \in \mathcal{G}^n_{\mathtt{K}_{\rm o}}$, recall (30) and define the following parameters[28]

$$
\begin{aligned}
&\mathtt{R} = \alpha/|k|^2 = \sqrt{\varepsilon}\mathtt{K}^\nu/|k|^2\,, \quad c_2 = 4n^{\frac32}c_1\,, \quad \mathtt{r} = \mathtt{R}/c_2\,, \quad \varepsilon_k = \tfrac{2\varepsilon}{|k|^2}\,,\\
&\hat D = \Big\{\hat I \in \mathbb{R}^{n-1} : |\mathrm{p}_k^\perp \hat{\mathrm{A}}^T\hat I| < 1\,, \min_{\substack{\ell\in\mathcal{G}^n_{\mathtt{K}}\\ \ell\notin\mathbb{Z}k}} \big|\big(\mathrm{p}_k^\perp \hat{\mathrm{A}}^T\hat I\big)\cdot\ell\big| \geqslant \tfrac{3\alpha\mathtt{K}}{|k|}\Big\}\,,\ D = (-\mathtt{R},\mathtt{R})\times\hat D\,,\\
&\upbeta = \begin{cases}\varepsilon_k\beta, & \text{if } |k|_1 < \mathtt{N}\\ \varepsilon_k|f_k|, & \text{if } |k|_1 \geqslant \mathtt{N}\end{cases}, \qquad \chi_k = \begin{cases}1\,, & \text{if } |k|_1 < \mathtt{N}\\ |f_k|\,, & \text{if } |k|_1 \geqslant \mathtt{N}\end{cases}, \qquad \epsilon = \mathtt{c}_s\varepsilon_k\,\chi_k\,,\\
&\mathtt{s} = \begin{cases}\min\{\frac{s}{2},1\}, & \text{if } |k|_1 < \mathtt{N}\\ 1\,, & \text{if } |k|_1 \geqslant \mathtt{N}\end{cases}, \qquad \check{\mathtt{s}} := \begin{cases}s'_k\,, & \text{if } |k|_1 < \mathtt{N}\,,\\ 1, & \text{if } |k|_1 \geqslant \mathtt{N}\end{cases}, \qquad \upmu = \vartheta = \frac{1}{\mathtt{K}^{5n}}\,.
\end{aligned}
\tag{83}
$$

**Theorem 3.1** *(Generic Standard Form at simple resonances). Let* $\mathtt{H}$ *be as in* (19) *with* $f \in \mathbb{G}^n_s$ *satisfying* (8) *and* (13) *for some* $0 < \delta \leqslant 1$ *and* $\beta > 0$ *with* $\mathtt{N}$ *defined in* (6)*. Assume that*[29] $\mathtt{K}_{\rm o} \geqslant \max\{c_2, \mathbf{c_0}\}$*. Then, with the definitions given in* (83)*, the following holds for all* $k \in \mathcal{G}^n_{\mathtt{K}_{\rm o}}$.

(i) *There exists a real analytic symplectic transformation*

$$
\Phi_\star : (\mathrm{p},\mathrm{q}) \in D\times\mathbb{R}^n \to (\mathrm{y},\mathrm{x}) = \Phi_\star(\mathrm{p},\mathrm{q}) \in \mathbb{R}^{2n}\,,
\tag{84}
$$

*such that:* $\Phi_\star$ *fixes* $\hat{\mathrm{p}}$ *and*[30] $\mathrm{q}_1$*; for every* $\hat{\mathrm{p}} \in \hat D$ *the map* $(\mathrm{p}_1,\mathrm{q}_1) \mapsto (\mathrm{y}_1,\mathrm{x}_1)$ *is symplectic; the* $(n+1)$*–dimensional map*[31] $\check\Phi_\star$ *depends only on the first* $n+1$ *coordinates* $(\mathrm{p},\mathrm{q}_1)$*, is* $2\pi$*–periodic in* $\mathrm{q}_1$ *and, if* $\mathscr{D}^k = \mathrm{A}^{-T}\mathcal{R}^{1,k}$ *and* $\overline{\mathtt{H}}_k$ *are as in Theorem 2.1, one has*[32]

$$
\begin{aligned}
&\check\Phi_\star : D_{\mathtt{r}}\times\mathbb{T}_{\check{\mathtt{s}}} \to \mathscr{D}^k_{\tilde r_k}\times\mathbb{T}_{\check{\mathtt{s}}}\,,\\
&\overline{\mathtt{H}}_k\circ\check\Phi_\star(\mathrm{p},\mathrm{q}) =: \tfrac{|k|^2}{2}\big(\mathtt{H}_k(\mathrm{p},\mathrm{q}_1) + \hat h_k(\hat{\mathrm{p}})\big)\,,\\
&\sup_{\hat{\mathrm{p}}\in\hat D_{2\mathtt{r}}}\big|\hat h_k(\hat{\mathrm{p}}) - \hat{\mathcal{Q}}_k(\hat{\mathrm{p}})\big| \leqslant 6\,\varepsilon_k\upmu\,, \qquad \hat{\mathcal{Q}}_k(\hat{\mathrm{p}}) := \tfrac{1}{|k|^2}|\mathrm{p}_k^\perp\hat{\mathrm{A}}^T\hat{\mathrm{p}}|^2\,.
\end{aligned}
\tag{85}
$$

(ii) $\mathtt{H}_k$ *in* (85) *is in Generic Universal Form according to Definition 3.1:*

$$
\mathtt{H}_k(\mathrm{p},\mathrm{q}_1) = \big(1+\upnu_k(\mathrm{p},\mathrm{q}_1)\big)\mathrm{p}_1^2 + \mathtt{G}_k(\hat{\mathrm{p}},\mathrm{q}_1)\,,
$$

*having reference potential*

$$
\bar{\mathtt{G}} = \bar{\mathtt{G}}_k := \varepsilon_k\,\pi_{\mathbb{Z}k}f\,,
\tag{86}
$$

[27] Recall that by Lemma 1.2 such $\delta$ and $\beta$ always exist.
[28] Here and in what follows we shall not always indicate explicitly the dependence upon $k$. Recall the definitions of $c_1$, $\hat{\mathrm{A}}$ and $\mathtt{c}_s$ in, respectively, (64) Lemma 2.6 and (7).
[29] $\mathbf{c_0}$ is defined in Theorem 2.1.
[30] I.e., in (84) it is $\mathrm{y} = \hat{\mathrm{p}}$, $\mathrm{x}_1 = \mathrm{q}_1$.
[31] Recall the notation in (82).
[32] $r_k$ and $s'_k$ are defined in (30), $\tilde r_k$ in (64).

*analyticity characteristics given in* (83)*, and* $\mathtt{\kappa}$ *verifying* (81) *with*

$$\mathtt{\kappa} = \mathtt{\kappa}(n, s, \beta) := \max\left\{c_2\,, 4\mathtt{c}_s\,, \mathtt{c}_s/\beta\right\}. \tag{87}$$

(iii) *Finally,* $\Phi_\star = \Phi_1 \circ \Phi_2 \circ \Phi_3$*, where*[33]*:* $\Phi_1 := \Psi_{\mathtt{g}_1} \in \mathfrak{G}_\dagger$ *with* $\mathtt{g}_1(\hat{\mathtt{p}}) := -\frac{1}{|k|^2}(\hat{\mathtt{A}}k)\cdot\hat{\mathtt{p}}$*;* $\Phi_3 := \Psi_{\mathtt{g}_3} \in \mathfrak{G}_\dagger$ *for a suitable real analytic function* $\mathtt{g}_3(\hat{\mathtt{p}})$ *satisfying*

$$|\mathtt{g}_3|_{4\mathtt{r}} < \tfrac{\varepsilon_k}{\mathtt{r}}\mu\,,$$

*and* $\Phi_2(\mathtt{p},\mathtt{q}) = (\mathtt{p}_1 + \eta_2, \hat{\mathtt{p}}, \mathtt{q}_1, \hat{\mathtt{q}} + \chi_2)$ *for suitable real analytic functions* $\eta_2 = \eta_2(\hat{\mathtt{p}}, \mathtt{q}_1)$ *and* $\chi_2 = \chi_2(\hat{\mathtt{p}}, \mathtt{q}_1)$ *satisfying*

$$|\eta_2|_{4\mathtt{r},\check{\mathtt{s}}} < \frac{\varepsilon_k\chi_k}{\mathtt{r}}\mu\,,\qquad |\chi_2|_{2\mathtt{r},\check{\mathtt{s}}} < \frac{4\varepsilon_k\chi_k}{\mathtt{r}^2}\mu\,.$$

**Remark 3.3.** (i) One of the main points of the above theorem is that the parameter $\kappa$ in (87) *does not depend on* $k$. Incidentally, we point out that $\kappa$ depends (indirectly) also on $\delta$, since $\delta$ appears in the definition of $\mathtt{N}$ and $\beta$ is the uniform Morse constant of the first $\mathtt{N}$ reference potentials.

(ii) Note that by (83), (85), (30) and (7)

$$\min\left\{\tfrac{s}{2}, 1\right\} \leqslant \mathtt{s} \leqslant \check{\mathtt{s}} \leqslant s'_k\,. \tag{88}$$

In particular, the composition $\overline{\mathtt{H}}_k \circ \check{\Phi}_\star$ is well defined; compare Theorem 2.1–(ii).
As for the action analyticity radii, notice that, by the definitions in (30), (64) and (83), one has

$$r_k = \mathtt{R}\,|k|\,,\qquad \tilde{r}_k = \frac{\mathtt{R}}{c_1}\,. \tag{89}$$

(iii) The three maps which define $\Phi_\star$ have the following purposes: The first one is needed to decouple the 'kinetic energy' of the 1–d.o.f. secular system; the second one is introduced so as to get a purely positional 1–dimensional potential; finally, the third one puts the momentum coordinate of the equilibria in 0.

(iv) The proof is fully constructive and the explicit definition of $\mathtt{H}_k$ is given in (121), (99), (101), (113), (108) and (92) below.

### 3.2. *Proof of the main theorem*

The proof is articulated in three lemmata.
The first lemma shows how to 'block–diagonalize' the kinetic energy. For $k \in \mathcal{G}^n_{\mathtt{K}_{\mathtt{o}}}$, recall the definition of the matrices A and $\hat{\mathrm{A}}$ in (61), and define[34]

$$\mathtt{y} = \mathrm{U}\mathtt{Y} := \begin{pmatrix} 1 & -\frac{1}{|k|^2}(\hat{\mathrm{A}}k)^T \\ 0 & \mathrm{I}_{n-1}\end{pmatrix}\mathtt{Y}\,,\qquad \text{i.e.}\qquad \begin{cases}\mathtt{y}_1 = \mathtt{Y}_1 - \frac{1}{|k|^2}\hat{\mathrm{A}}k\cdot\hat{\mathtt{Y}}\,,\\ \hat{\mathtt{y}} = \hat{\mathtt{Y}}\,.\end{cases} \tag{90}$$

33 Recall Definition 3.2.
34 $\mathrm{I}_m$ denotes the $(m\times m)$–identity matrix and recall the notation in (82).

Then, one has

**Lemma 3.1.** (i) *Let* $\Phi_1$ *be the map* $\Phi_1(\mathtt{Y},\mathtt{X}) = (\mathrm{U}\mathtt{Y}, \mathrm{U}^{-T}\mathtt{X})$. *Then,* $\Phi_1$ *is symplectic and*

$$\mathscr{D}^k = \mathrm{U}D\,, \qquad \check{\Phi}_1 : D_{4\mathtt{r}} \times \mathbb{T}_{\breve{\mathtt{s}}} \to \mathscr{D}^k_{\tilde{r}_k} \times \mathbb{T}_{\breve{\mathtt{s}}}\,. \tag{91}$$

(ii) *Let*

$$\begin{cases} \mathtt{G}^{(1)}_{\mathrm{o}} := \varepsilon_k g^k_{\mathrm{o}}(\mathrm{A}^T\mathrm{U}\mathtt{Y})\,, \quad \mathtt{G}^{(1)}(\mathtt{Y},\mathtt{X}_1) := \varepsilon_k g^k(\mathrm{A}^T\mathrm{U}\mathtt{Y},\mathtt{X}_1)\,, \\ \overline{\mathtt{H}}^{(1)}(\mathtt{Y},\mathtt{X}_1) := \mathtt{Y}_1^2 + \mathtt{G}^{(1)}_{\mathrm{o}}(\mathtt{Y}) + \mathtt{G}^{(1)}(\mathtt{Y},\mathtt{X}_1)\,, \qquad \langle \mathtt{G}^{(1)}(\mathtt{Y},\cdot)\rangle = 0\,. \end{cases} \tag{92}$$

*Then, if* $\overline{\mathtt{H}}_k$ *is as in* (67)*, one has*

$$\overline{\mathtt{H}}_k \circ \check{\Phi}_1(\mathtt{Y},\mathtt{X}_1) = \tfrac{|k|^2}{2}\,\overline{\mathtt{H}}^{(1)}(\mathtt{Y},\mathtt{X}_1) + \tfrac12 |\mathrm{p}_k^\perp \hat{\mathrm{A}}^T\hat{\mathtt{Y}}|^2\,, \tag{93}$$

*with* $\overline{\mathtt{H}}^{(1)}$ *real analytic on* $D_{4\mathtt{r}} \times \mathbb{T}_{\breve{\mathtt{s}}}$ *and* $\langle \mathtt{G}^{(1)}(\mathtt{Y},\cdot)\rangle = 0$*, and the following estimates hold*[35]*:*

$$|\mathtt{G}^{(1)}_{\mathrm{o}}|_{4\mathtt{r}} \leqslant \eta_{\mathrm{o}} := 2\varepsilon_k\vartheta = \frac{2\varepsilon_k}{\mathrm{K}^{5n}}\,, \qquad |\mathtt{G}^{(1)} - \bar{\mathtt{G}}_k|_{4\mathtt{r},\breve{\mathtt{s}}} \leqslant \eta := \chi_k\eta_{\mathrm{o}} \leqslant \eta_{\mathrm{o}}\,. \tag{94}$$

**Proof.** (i) $\Phi_1$ is symplectic since it is generated by the generating function $\mathrm{U}\mathtt{Y}\cdot\mathtt{x}$.
From the definitions of A and U in, respectively, (61) and (90), it follows

$$(\mathrm{A}^T\mathrm{U})\mathtt{Y} = \mathtt{Y}_1 k + \hat{\mathrm{A}}^T\hat{\mathtt{Y}} - \tfrac{(\hat{\mathrm{A}}k)\cdot\hat{\mathtt{Y}}}{|k|^2}k = \mathtt{Y}_1 k + \hat{\mathrm{A}}^T\hat{\mathtt{Y}} - \tfrac{\hat{\mathrm{A}}^T\hat{\mathtt{Y}}\cdot k}{|k|^2}k = \mathtt{Y}_1 k + \mathrm{p}_k^\perp\hat{\mathrm{A}}^T\hat{\mathtt{Y}}\,. \tag{95}$$

Thus, $\mathtt{y} = (\mathrm{A}^T\mathrm{U})\mathtt{Y}$ if and only if $\mathtt{y}\cdot k = \mathtt{Y}_1|k|^2$ and $\mathrm{p}_k^\perp\mathtt{y} = \mathrm{p}_k^\perp\hat{\mathrm{A}}^T\hat{\mathtt{Y}}$, which is equivalent to say $(\mathrm{A}^T\mathrm{U})D = \mathcal{R}^{1,k}$, which in view of (64), is equivalent to $\mathscr{D}^k = \mathrm{U}D$. Now, by (61),

$$|\mathrm{U}|,\, |\mathrm{U}^{-1}| \leqslant n^{\frac32}\,, \tag{96}$$

where, as usual, for a matrix $M$ we denote by $|M| = \sup_{u\neq 0}|Mu|/|u|$ the standard operator norm.
Thus, by (83) and (89) we have (for complex $z$)

$$|z| < 4\mathtt{r} \implies |\mathrm{U}z| < n^{\frac32}4\mathtt{r} = 4n^{\frac32}\frac{\mathrm{R}}{c_2} = \frac{\mathrm{R}}{c_1} = \tilde{r}_k\,, \tag{97}$$

which, since $\mathtt{X}_1 = \mathtt{x}_1$, implies that $\check{\Phi}_1 : D_{4\mathtt{r}} \times \mathbb{T}_{\breve{\mathtt{s}}} \to \mathscr{D}^k_{\tilde{r}_k} \times \mathbb{T}_{\breve{\mathtt{s}}}$, proving (91).

(ii) By the previous item, the composition $\overline{\mathtt{H}}_k \circ \check{\Phi}_1$ is well defined and analytic on $D_{4\mathtt{r}} \times \mathbb{T}_{\breve{\mathtt{s}}}$. From (95) it follows that $|\mathrm{A}^T\mathrm{U}\mathtt{Y}|^2 = |k|^2\mathtt{Y}_1^2 + |\mathrm{p}_k^\perp\hat{\mathrm{A}}^T\hat{\mathtt{Y}}|^2$, and (93) follows. Notice that since $g^k(y,\cdot) \in \mathbb{B}^1_{s_k'}$ (compare Lemma 2.1), $\mathtt{G}^{(1)}$ has zero average over $\mathbb{T}$.
By the definition of $\mathtt{G}^{(1)}_{\mathrm{o}}$ and $\mathtt{G}^{(1)}$ in (92), by (97), (36) in[36] Lemma 2.1, the estimates on $|\mathtt{G}^{(1)}_{\mathrm{o}}|_{4\mathtt{r}}$

35 $\vartheta$ is defined in (40); $\bar{\mathtt{G}}_k$ is the reference potential in (86). Notice that, by (83), $\chi_k \leqslant 1$ for all $k$.
36 Recall that, by (30), (64), $\tilde{r}_k < r_k' = r_k/2$. Recall also the definitions of $\vartheta_{\mathrm{o}}$ and $\vartheta$ in (33) and (40).

and on $|\mathtt{G}^{(1)} - \bar{\mathtt{G}}_k|_{4\mathtt{r},\check{\mathtt{s}}}$ for $|k|_1 < \mathtt{N}$ in (94) follow. The estimate for $|k|_1 \geqslant \mathtt{N}$ in (94) follows from Lemma 2.2: see in particular (39), (40) and (9). □

Next lemma shows how one can remove the dependence on $\mathtt{Y}_1$ in the potential $\mathtt{G}^{(1)}$.

**Lemma 3.2.** *If let* $\mathtt{K} \geqslant c_2$ *then,*

$$\frac{\eta_o}{\mathtt{r}^2} < \frac{1}{2^{10}}\, \frac{\check{\mathtt{s}}}{\pi+\check{\mathtt{s}}} < 1\,, \tag{98}$$

*and the following statements hold.*
(i) *The fixed point equation*

$$\mathfrak{p} = -\tfrac{1}{2}\, \partial_{\mathtt{Y}_1} \mathtt{G}^{(1)}_{\mathtt{o}}(\mathfrak{p}, \hat{\mathtt{P}}) - \tfrac{1}{2}\, \partial_{\mathtt{Y}_1} \mathtt{G}^{(1)}(\mathfrak{p}, \hat{\mathtt{P}}, \mathtt{Q}_1) \tag{99}$$

*has a unique solution* $\mathfrak{p} : (\hat{\mathtt{P}}, \mathtt{Q}_1) \in \hat{D} \times \mathbb{T} \mapsto \mathfrak{p}(\hat{\mathtt{P}}, \mathtt{Q}_1) \in \mathbb{R}$ *real analytic on* $\hat{D}_{4\mathtt{r}} \times \mathbb{T}_{\check{\mathtt{s}}}$, *satisfying*

$$|\mathfrak{p}|_{4\mathtt{r},\check{\mathtt{s}}} < \frac{\eta_o}{3\mathtt{r}}\,. \tag{100}$$

*Furthermore, if we define*

$$\begin{cases} \mathfrak{p}_o(\hat{\mathtt{P}}) := \langle \mathfrak{p}(\hat{\mathtt{P}}, \cdot) \rangle \\ \tilde{\mathfrak{p}} := \mathfrak{p} - \mathfrak{p}_o \end{cases}, \qquad \begin{cases} \phi(\hat{\mathtt{P}}, \mathtt{X}_1) := \displaystyle\int_0^{\mathtt{X}_1} \tilde{\mathfrak{p}}(\hat{\mathtt{P}}, \theta)d\theta \\ \hat{\mathfrak{q}}(\hat{\mathtt{P}}, \mathtt{Q}_1) := -\partial_{\hat{\mathtt{P}}}\, \phi(\hat{\mathtt{P}}, \mathtt{Q}_1) \end{cases} \tag{101}$$

*then,* $\mathtt{Q}_1 \to \hat{\mathfrak{q}}(\hat{\mathtt{P}}, \mathtt{Q}_1)$ *is a real analytic periodic function, and one has*

$$|\mathfrak{p}_o|_{4\mathtt{r}} < \frac{1}{3}\, \frac{\eta_o}{\mathtt{r}}\,, \qquad |\tilde{\mathfrak{p}}|_{4\mathtt{r},\check{\mathtt{s}}} < \frac{1}{3}\, \frac{\eta}{\mathtt{r}}\,, \qquad |\hat{\mathfrak{q}}|_{2\mathtt{r},\check{\mathtt{s}}} < \frac{\eta}{6\mathtt{r}^2}\,(\pi + \check{\mathtt{s}})\,. \tag{102}$$

(ii) *The real analytic symplectic map* $\Phi_2$ *generated by* $\mathtt{P}\cdot\mathtt{X} + \phi(\hat{\mathtt{P}}, \mathtt{X}_1)$, *namely,*

$$\Phi_2 : (\mathtt{P}, \mathtt{Q}) \mapsto (\mathtt{Y}, \mathtt{X}) \quad \text{with} \quad \begin{cases} \mathtt{Y}_1 = \mathtt{P}_1 + \tilde{\mathfrak{p}}(\hat{\mathtt{P}}, \mathtt{Q}_1) \\ \hat{\mathtt{Y}} = \hat{\mathtt{P}} \end{cases}, \quad \begin{cases} \mathtt{X}_1 = \mathtt{Q}_1 \\ \hat{\mathtt{X}} = \hat{\mathtt{Q}} + \hat{\mathfrak{q}}(\hat{\mathtt{P}}, \mathtt{Q}_1) \end{cases}, \tag{103}$$

*satisfies:*

$$\check{\Phi}_2 : D_{2\mathtt{r}} \times \mathbb{T}_{\check{\mathtt{s}}} \to D_{3\mathtt{r}} \times \mathbb{T}_{\check{\mathtt{s}}}\,, \tag{104}$$

*and*

$$\begin{aligned} \overline{\mathtt{H}}^{(2)}(\mathtt{P}, \mathtt{Q}_1) &:= \overline{\mathtt{H}}^{(1)} \circ \check{\Phi}_2(\mathtt{P}, \mathtt{Q}_1) \\ &= \big(1 + \tilde{\nu}(\mathtt{P}, \mathtt{Q}_1)\big)\big(\mathtt{P}_1 - \mathfrak{p}_o(\hat{\mathtt{P}})\big)^2 + \mathtt{G}^{\mathtt{o}}_\sharp(\hat{\mathtt{P}}) + \mathtt{G}_k(\hat{\mathtt{P}}, \mathtt{Q}_1)\,, \end{aligned} \tag{105}$$

*for suitable functions* $\tilde{\nu}$, $\mathtt{G}^{\mathtt{o}}_\sharp$ *and* $\mathtt{G}_k$ *(explicitly defined in* (113) *below, with* $\langle \mathtt{G}_k \rangle = 0$*) real analytic on, respectively,* $D_{2\mathtt{r}} \times \mathbb{T}_{\check{\mathtt{s}}}$, $\hat{D}_{2\mathtt{r}}$ *and* $\hat{D}_{2\mathtt{r}} \times \mathbb{T}_{\check{\mathtt{s}}}$, *which satisfy the bounds:*

$$|\tilde{\nu}|_{2\mathtt{r},\check{\mathtt{s}}} \leqslant \frac{\eta_o}{\mathtt{r}^2}\,, \qquad |\mathtt{G}^o_\sharp|_{2\mathtt{r}} \leqslant 3\eta_o \qquad |\mathtt{G}_k - \bar{\mathtt{G}}_k|_{2\mathtt{r},\check{\mathtt{s}}} \leqslant 2\eta\,. \tag{106}$$

**Proof.** We start by proving (98). Recalling (88), (7) and (41), we have

$$\frac{\pi+\check{\mathtt{s}}}{\check{\mathtt{s}}} \leqslant 1 + 2\pi\,\mathtt{c}_s < 8\mathtt{c}_s < \mathtt{K}. \tag{107}$$

Now, by the definitions in (94), (93), (83), (30), we find

$$\frac{\eta_o}{\mathtt{r}^2} = 4c_2^2\,\frac{|k|^2}{\mathtt{K}^{14n+4}} \overset{(107)}{\leqslant} 4c_2^2\,\frac{1}{\mathtt{K}^{14n+1}}\,\frac{\check{\mathtt{s}}}{\pi+\check{\mathtt{s}}}\,,$$

which yields (98) since, by assumption, $\mathtt{K} > \mathtt{K}_o \geqslant c_2$.

(i) Let us denote by $\mathbf{X} := \hat{D}_{4\mathtt{r},\check{\mathtt{s}}} \times \mathbb{T}_{\check{\mathtt{s}}}$ and by $\mathcal{X}$ the complete metric space formed by the real analytic complex–valued functions $u : \mathbf{X} \to \{z \in \mathbb{C} : |z| \leqslant \mathtt{r}/2\}$, equipped with the metric given by the distance in sup–norm on $\mathbf{X}$. Let us also denote:

$$\mathtt{G}^\sharp := \mathtt{G}^{(1)}_o + \mathtt{G}^{(1)}\,, \qquad \tilde{\mathtt{G}}^\sharp := \mathtt{G}^{(1)} - \bar{\mathtt{G}}_k\,. \tag{108}$$

Note that $\mathtt{G}^{(1)}$ and $\tilde{\mathtt{G}}^\sharp$ have zero average. Consider the operator $F : u \in \mathcal{X} \mapsto F(u)$, where $F(u)(\hat{\mathtt{P}},\mathtt{Q}_1) := -\frac{1}{2}\partial_{\mathtt{Y}_1}\mathtt{G}^\sharp(u(\hat{\mathtt{P}},\mathtt{Q}_1),\hat{\mathtt{P}},\mathtt{Q}_1)$. If $u \in \mathcal{X}$, then, by Cauchy estimate we get

$$\begin{aligned}
\sup_{\mathbf{X}} |F(u)| &= \tfrac{1}{2}\sup_{\mathbf{X}} \big|\partial_{\mathtt{Y}_1}\mathtt{G}^\sharp(u(\hat{\mathtt{P}},\mathtt{Q}_1),\hat{\mathtt{P}},\mathtt{Q}_1)\big| \\
&= \tfrac{1}{2}\sup_{\mathbf{X}} \big|\partial_{\mathtt{Y}_1}\big[\mathtt{G}^\sharp(u(\hat{\mathtt{P}},\mathtt{Q}_1),\hat{\mathtt{P}},\mathtt{Q}_1) - \bar{\mathtt{G}}_k(\mathtt{Q}_1)\big]\big| \\
&\leqslant \frac{1}{2}\,\frac{\big|\mathtt{G}^\sharp - \bar{\mathtt{G}}_k\big|_{4\mathtt{r},\check{\mathtt{s}}}}{4\mathtt{r} - \frac{\mathtt{r}}{2}} \\
&\overset{(94)}{\leqslant} \frac{1}{2}\,\frac{\eta_o + \eta}{4\mathtt{r} - \frac{\mathtt{r}}{2}} \leqslant \frac{2}{7}\,\frac{\eta_o}{\mathtt{r}} \overset{(98)}{<} \frac{2}{7}\,\mathtt{r} < \frac{\mathtt{r}}{2}\,.
\end{aligned} \tag{109}$$

Thus, $F : \mathcal{X} \to \mathcal{X}$. Let us check that $F$ is, in fact, a contraction on $\mathcal{X}$. If $u, v \in \mathcal{X}$, then, again, by Cauchy estimate, (94) and (98), we get[37]

$$\begin{aligned}
\sup_{\mathbf{X}} |F(u) - F(v)| &\leqslant \tfrac{1}{2}\sup_{\mathbf{X}} \big|\partial_{\mathtt{Y}_1}\big(\mathtt{G}^\sharp(u,\hat{\mathtt{P}},\mathtt{Q}_1) - \mathtt{G}^\sharp(v,\hat{\mathtt{P}},\mathtt{Q}_1)\big)\big| \\
&\leqslant \frac{1}{2}\,\big|\partial^2_{\mathtt{Y}_1}(\mathtt{G}^\sharp - \bar{\mathtt{G}}_k)\big|_{\frac{\mathtt{r}}{2},\check{\mathtt{s}}} \cdot \sup_{\mathbf{X}} |u - v| \\
&\leqslant \frac{1}{2}\,\frac{\big|\mathtt{G}^\sharp - \bar{\mathtt{G}}_k\big|_{4\mathtt{r},\check{\mathtt{s}}}}{\big(4\mathtt{r} - \frac{\mathtt{r}}{2}\big)^2} \cdot \sup_{\mathbf{X}} |u - v| \\
&\overset{(94)}{\leqslant} \frac{4}{49}\,\frac{\eta_o}{\mathtt{r}^2} \cdot \sup_{\mathbf{X}} |u - v| \overset{(98)}{<} \frac{1}{8} \cdot \sup_{\mathbf{X}} |u - v|\,,
\end{aligned} \tag{110}$$

[37] $u$ and $v$, in the r.h.s. of the first inequality, are evaluated at $(\mathtt{Q}_1, \hat{\mathtt{P}})$.

showing that $F$ is a contraction on $\mathcal{X}$. Thus, by the standard Contraction Lemma, it follows that there exists a unique $\mathfrak{p} \in \mathcal{X}$ solving (99).

Since $\mathfrak{p} = F(\mathfrak{p})$, one sees that (100) follows from (109).
The first bound in (102) follows immediately from (100).
To prove the second estimate in (102), write[38]

$$\partial_{\mathtt{Y}_1} \mathtt{G}^\sharp(\mathfrak{p}) = \partial_{\mathtt{Y}_1} \mathtt{G}^\sharp(\mathfrak{p}_{\mathrm{o}} + \tilde{\mathfrak{p}}) = \partial_{\mathtt{Y}_1} \mathtt{G}^\sharp(\mathfrak{p}_{\mathrm{o}}) + w\tilde{\mathfrak{p}}\,, \quad \text{with} \quad w := \int_0^1 \partial^2_{\mathtt{Y}_1} \mathtt{G}^\sharp(\mathfrak{p}_{\mathrm{o}} + t\tilde{\mathfrak{p}})dt\,. \tag{111}$$

As above, by Cauchy estimates,

$$|w|_{4\mathtt{r},\check{\mathtt{s}}} \leqslant \frac{2}{49} \frac{\eta_{\mathrm{o}} + \eta}{\mathtt{r}^2} < \frac{1}{8}\,. \tag{112}$$

Thus, by (111), Cauchy estimates, and (112), observing that[39] $\langle \partial_{\mathtt{Y}_1} \mathtt{G}^{(1)}(\mathfrak{p}_{\mathrm{o}})\rangle = 0$, one finds

$$\begin{aligned}
|\tilde{\mathfrak{p}}| = |\mathfrak{p} - \mathfrak{p}_{\mathrm{o}}| &\overset{(111)}{=} \tfrac{1}{2}\big|\partial_{\mathtt{Y}_1}\mathtt{G}^\sharp(\mathfrak{p}_{\mathrm{o}}) - \langle \partial_{\mathtt{Y}_1}\mathtt{G}^\sharp(\mathfrak{p}_{\mathrm{o}})\rangle + w\tilde{\mathfrak{p}} - \langle w\tilde{\mathfrak{p}}\rangle\big| \\
&= \tfrac{1}{2}\big|\partial_{\mathtt{Y}_1}\mathtt{G}^{(1)}(\mathfrak{p}_{\mathrm{o}}) - \langle \partial_{\mathtt{Y}_1}\mathtt{G}^{(1)}(\mathfrak{p}_{\mathrm{o}})\rangle + w\tilde{\mathfrak{p}} - \langle w\tilde{\mathfrak{p}}\rangle\big| \\
&= \tfrac{1}{2}\big|\partial_{\mathtt{Y}_1}\big(\mathtt{G}^{(1)}(\mathfrak{p}_{\mathrm{o}}) - \bar{\mathtt{G}}_k\big) + w\tilde{\mathfrak{p}} - \langle w\tilde{\mathfrak{p}}\rangle\big| \\
&\overset{(94),(112)}{\leqslant} \frac{1}{2}\Big(\frac{2}{7}\frac{\eta}{\mathtt{r}}\Big) + \frac{1}{2}|\tilde{\mathfrak{p}}|\,,
\end{aligned}$$

which yields immediately the second bound in (102).
Next, since $\tilde{\mathfrak{p}}$ has zero average over the torus, the function $\phi$ defined in (101) defines a (real analytic) periodic function such that $\partial_{\mathtt{X}_1}\phi = \tilde{\mathfrak{p}}$. Furthermore, by the second estimate in (102), one has[40] $|\phi|_{4\mathtt{r},\check{\mathtt{s}}} < \frac{\eta}{3\mathtt{r}}\,(\pi + \check{\mathtt{s}})$, so that, by Cauchy estimates, also last bounds in (102) follow.

(ii) By the definition of $\Phi_2$ in (103), by (102) and (98), the relations in (104) follow at once.
Now, define[41]

$$\begin{cases}
\tilde{\mathtt{v}}(\mathtt{P},\mathtt{Q}_1) := \displaystyle\int_0^1 (1-t)\partial^2_{\mathtt{Y}_1}\mathtt{G}^\sharp\big(\mathfrak{p}_{\mathrm{o}} + t(\mathtt{P}_1 - \mathfrak{p}_{\mathrm{o}}), \hat{\mathtt{P}}, \mathtt{Q}_1\big)dt\,, \\
\mathtt{G}^{\mathrm{o}}_\sharp(\hat{\mathtt{P}}) := \langle \mathfrak{p}(\hat{\mathtt{P}},\cdot)^2\rangle + \langle \mathtt{G}^\sharp\big(\mathfrak{p}(\cdot,\hat{\mathtt{P}}),\hat{\mathtt{P}},\cdot\big)\rangle\,, \\
\mathtt{G}_k(\hat{\mathtt{P}},\mathtt{Q}_1) := \mathfrak{p}(\hat{\mathtt{P}},\mathtt{Q}_1)^2 + \mathtt{G}^\sharp\big(\mathfrak{p}(\mathtt{Q}_1,\hat{\mathtt{P}}),\hat{\mathtt{P}},\mathtt{Q}_1\big) - \mathtt{G}^{\mathrm{o}}_\sharp(\hat{\mathtt{P}})\,,
\end{cases} \tag{113}$$

38 To simplify notation, we drop, here, from the notation the explicit dependence on $\hat{\mathtt{P}}$ and $\mathtt{Q}_1$ of $\mathtt{G}^\sharp$.
39 Recall that $\langle \mathtt{G}^{(1)}(\mathtt{Y},\cdot)\rangle = 0$ as stated in Lemma 3.1.
40 $\pi + \check{\mathtt{s}}$ is an estimate of the length of the integration path in (101), as the real part of $\mathtt{Q}_1$ can be taken in $[-\pi,\pi)$.
41 Here, $\mathfrak{p}_{\mathrm{o}} = \mathfrak{p}_{\mathrm{o}}(\hat{\mathtt{P}})$.

then, by Taylor's formula, (92), (103), (108) and (99), one finds[42]

$$
\begin{aligned}
\overline{\mathtt{H}}^{(2)}(\mathtt{P}_1,\mathtt{Q}_1) \;:=\; & \overline{\mathtt{H}}^{(1)}\circ\check{\Phi}_2(\mathtt{P}_1,\mathtt{Q}_1)=(\mathtt{P}_1+\tilde{\mathfrak{p}})^2+\mathtt{G}^\sharp(\mathtt{P}_1+\tilde{\mathfrak{p}},\mathtt{Q}_1)\\
\overset{(101)}{=}\; & \big(\mathfrak{p}+(\mathtt{P}_1-\mathfrak{p}_{\rm o})\big)^2+\mathtt{G}^\sharp\big(\mathfrak{p}+(\mathtt{P}_1-\mathfrak{p}_{\rm o}),\mathtt{Q}_1\big)\\
\overset{(113)}{=}\; & (\mathtt{P}_1-\mathfrak{p}_{\rm o})^2+2(\mathtt{P}_1-\mathfrak{p}_{\rm o})\mathfrak{p}+\mathfrak{p}^2+\mathtt{G}^\sharp(\mathfrak{p},\mathtt{Q}_1)+\partial_{\mathtt{Y}_1}\mathtt{G}^\sharp(\mathfrak{p},\mathtt{Q}_1)(\mathtt{P}_1-\mathfrak{p}_{\rm o})\\
& +(\mathtt{P}_1-\mathfrak{p}_{\rm o})^2\tilde{\mathtt{v}}\\
\overset{(99)}{=}\; & (1+\tilde{\mathtt{v}})(\mathtt{P}_1-\mathfrak{p}_{\rm o})^2+\mathfrak{p}^2+\mathtt{G}^\sharp(\mathfrak{p},\mathtt{Q}_1)\\
\overset{(113)}{=}\; & (1+\tilde{\mathtt{v}})(\mathtt{P}_1-\mathfrak{p}_{\rm o})^2+\mathtt{G}^{\rm o}_\sharp+\mathtt{G}_k(\mathtt{Q}_1)\,,
\end{aligned}
\tag{114}
$$

proving (105).

Let us now prove (106). Note that for $\mathtt{P}\in D_{2\mathtt{r}}$ by (98) and (102) the segment $\big(\mathfrak{p}_{\rm o}+t(\mathtt{P}_1-\mathfrak{p}_{\rm o}),\hat{\mathtt{P}}\big)$, $t\in[0,1]$, still belongs to $D_{\frac{5}{2}\mathtt{r}}$, hence, by definition of $\tilde{\mathtt{v}}$ in (113), by Cauchy estimate[43] and (94) one obtains the first estimate in (106).

By the definition of $\mathtt{G}^{\rm o}_\sharp$ in (113), by (100), (94) and (98) one gets immediately the second estimate in (106).

As for the third estimate in (106), by the definitions given, one has that[44]

$$
\mathtt{G}_k-\bar{\mathtt{G}}_k=\big(\mathfrak{p}^2-\langle\mathfrak{p}^2\rangle\big)+\big(\mathtt{G}^\sharp(\mathfrak{p},\cdot)-\langle\mathtt{G}^\sharp(\mathfrak{p},\cdot)\rangle-\bar{\mathtt{G}}_k\big)\,. \tag{115}
$$

Let us estimate the terms in brackets separately. For $\hat{\mathtt{P}}\in\hat{D}_{2\mathtt{r}}$ and $\mathtt{Q}_1\in\mathbb{T}_{\check{\mathtt{s}}}$, one finds

$$
|\mathfrak{p}^2-\langle\mathfrak{p}^2\rangle|=|2\tilde{\mathfrak{p}}\mathfrak{p}_{\rm o}+\tilde{\mathfrak{p}}^2-\langle\tilde{\mathfrak{p}}^2\rangle|\leqslant(2|\mathfrak{p}_{\rm o}|+2|\tilde{\mathfrak{p}}|)\,|\tilde{\mathfrak{p}}|\overset{(102)}{\leqslant}\frac{4}{9}\eta_{\rm o}\frac{\eta}{\mathtt{r}^2}\overset{(98)}{<}\frac{\eta}{2}\,. \tag{116}
$$

To estimate the second term in (115), we define

$$
z(t):=\mathtt{G}^\sharp(\mathfrak{p}_{\rm o}+t\tilde{\mathfrak{p}},\mathtt{Q}_1)-\langle\mathtt{G}^\sharp(\mathfrak{p}_{\rm o}+t\tilde{\mathfrak{p}},\cdot)\rangle\,,
$$

and observe that (recall (108)) $z(0)=\mathtt{G}^{(1)}(\mathfrak{p}_{\rm o},\mathtt{Q}_1)$ and that, by Cauchy estimates, we get[45]

$$
\begin{aligned}
|z'(s)|&\leqslant|\tilde{\mathfrak{p}}|\int_0^1\Big|\partial_{\mathtt{Y}_1}\big(\mathtt{G}^\sharp(\mathfrak{p}_{\rm o}+t\tilde{\mathfrak{p}},\mathtt{Q}_1)-\langle\mathtt{G}^\sharp(\mathfrak{p}_{\rm o}+t\tilde{\mathfrak{p}},\cdot)\rangle\big)\Big|dt\\
&\leqslant|\tilde{\mathfrak{p}}|\,\frac{2|\mathtt{G}^\sharp|_{4\mathtt{r},\check{\mathtt{s}}}}{4\mathtt{r}-\frac{\mathtt{r}}{2}}\leqslant|\tilde{\mathfrak{p}}|\,\frac{4\eta_{\rm o}}{4\mathtt{r}-\frac{\mathtt{r}}{2}}\overset{(102)}{<}\frac{8}{21}\frac{\eta_{\rm o}}{\mathtt{r}^2}\,\eta\overset{(98)}{<}\frac{1}{2}\eta\,.
\end{aligned}
\tag{117}
$$

---

42 $g(t_0+\tau)=g(t_0)+g'(t_0)\tau+\Big(\int_0^1(1-t)g''(t_0+t\tau)dt\Big)\tau^2$ with $g=\mathtt{G}^{(1)}(\cdot,\mathtt{Q}_1)$, $t_0=\mathfrak{p}$ and $\tau=\mathtt{P}_1-\mathfrak{p}_{\rm o}$. For ease of notation we drop the (dumb) dependence upon $\hat{\mathtt{P}}$ in these formulae.

43 Compare, also, the estimates done in (110).

44 Dropping, again, in the notation the dumb variable $\hat{\mathtt{P}}$.

45 Reasoning as in (109).

Thus,

$$\left|\mathtt{G}^\sharp(\mathfrak{p},\cdot)-\langle\mathtt{G}^\sharp(\mathfrak{p},\cdot)\rangle-\bar{\mathtt{G}}_k\right|\leqslant|z(0)-\bar{\mathtt{G}}_k|+\int_0^1|z'(t)|dt\leqslant|\mathtt{G}^{(1)}(\mathfrak{p}_o,\cdot)-\bar{\mathtt{G}}_k|+\frac{1}{2}\eta\overset{(94)}{\leqslant}\frac{3}{2}\eta\,.$$

Putting together this estimate and (116) one gets also the third estimate in (106). □

The final transformation is again just a translation, which is done so that *all equilibria of the secular system will lie on the angle–axis in its 2–dimensional phase space*.

**Lemma 3.3.** *The real analytic symplectic map $\Phi_3\in\mathfrak{G}_\dagger$ defined as*

$$\Phi_3:(\mathtt{p},\mathtt{q})\mapsto(\mathtt{P},\mathtt{Q})\quad\text{with}\quad\begin{cases}\mathtt{P}_1=\mathtt{p}_1+\mathfrak{p}_o(\hat{\mathtt{p}})\\ \hat{\mathtt{P}}=\hat{\mathtt{p}}\end{cases},\qquad\begin{cases}\mathtt{Q}_1=\mathtt{q}_1\\ \hat{\mathtt{Q}}=\hat{\mathtt{q}}-\mathtt{q}_1\partial_{\hat{\mathtt{p}}}\mathfrak{p}_o(\hat{\mathtt{p}})\end{cases},\tag{118}$$

*satisfies*

$$\check{\Phi}_3:D_{\mathtt{r}}\times\mathbb{T}_{\check{\mathtt{s}}}\to D_{2\mathtt{r}}\times\mathbb{T}_{\check{\mathtt{s}}}\,.\tag{119}$$

*Furthermore, one has:*

$$\overline{\mathtt{H}}^{(2)}\circ\check{\Phi}_3(\mathtt{p},\mathtt{q}_1)=\big(1+\mathtt{v}_k(\mathtt{p},\mathtt{q}_1)\big)\,\mathtt{p}_1^2+\mathtt{G}^{\mathrm{o}}_\sharp(\hat{\mathtt{p}})+\mathtt{G}_k(\hat{\mathtt{p}},\mathtt{q}_1)\,,\tag{120}$$

*where*

$$\mathtt{v}_k(\mathtt{p},\mathtt{q}_1):=\tilde{\mathtt{v}}\big(\mathfrak{p}_o(\hat{\mathtt{p}})+\mathtt{p}_1,\hat{\mathtt{p}},\mathtt{q}_1\big)\,,\tag{121}$$

*and the following bounds hold:*

$$|\mathtt{v}_k|_{\mathtt{r},\check{\mathtt{s}}}\leqslant\frac{\eta_o}{\mathtt{r}^2}\,,\qquad|\mathtt{G}^{\mathrm{o}}_\sharp|_{2\mathtt{r}}\leqslant3\eta_o\qquad|\mathtt{G}_k-\bar{\mathtt{G}}_k|_{2\mathtt{r},\check{\mathtt{s}}}\leqslant2\eta\,.\tag{122}$$

**Proof.** Just observe that, if $|\mathtt{p}_1|<\mathtt{r}$, then, by (102) and (98), it follows that, for all $\mathtt{p}\in D_{\mathtt{r}}$,

$$|\mathfrak{p}_o(\hat{\mathtt{p}})+\mathtt{p}_1|<\tfrac{\eta_o}{3\mathtt{r}}+\mathtt{r}\leqslant\tfrac{\mathtt{r}}{3}+\mathtt{r}=\tfrac{4}{3}\mathtt{r}<2\mathtt{r}\,,$$

so that (119) holds. Finally, by (106), we get[46] (122). □

We are ready for the

**Proof of Theorem 3.1.** Recall the definitions of $\Phi_j$, $1\leqslant j\leqslant3$, in, respectively, Lemma 3.1, (103) and (118) and define $\Phi_\star:=\Phi_1\circ\Phi_2\circ\Phi_3$, and $\hat{h}_k(\hat{I}):=\frac{1}{|k|^2}|\mathtt{p}_k^\perp\hat{\mathrm{A}}^T\hat{I}|^2+\mathtt{G}^{\mathrm{o}}_\sharp(\hat{I})$. Then, the expression for $\mathtt{H}_k$ in (85) follows by (93), (105) and (120).

[46] $\mathtt{G}_k$ *and* $\bar{\mathtt{G}}_k$ *are the same* as in (106) of Lemma 3.2.

By (119) and Lemma 3.3, the Hamiltonian function $\mathtt{H}_k$ is real analytic on $D_{\mathtt{r}} \times \mathbb{T}_{\breve{\mathtt{s}}}$, where $D = (-\mathtt{R}, \mathtt{R}) \times \hat{D}$ (compare (83)).
By (13) and Proposition 1.2 we have that $\bar{\mathtt{G}}_k$ in (86) is $\beta$–Morse with $\beta$ as in (83).
Let us, now, estimate $|\bar{\mathtt{G}}_k|_{\mathtt{s}}$. Consider, first, $|k|_1 < \mathtt{N}$. Then, estimating $|f_{jk}|$ by $e^{-|j||k|_1 s}$, by the definition of $\mathtt{s}$, we get

$$|\bar{\mathtt{G}}_k|_{\mathtt{s}} \overset{(86)}{=} \varepsilon_k |\pi_{\mathbb{Z}k} f|_{\mathtt{s}} \leqslant \varepsilon_k |\pi_{\mathbb{Z}k} f|_{s/2} = \varepsilon_k \sum_{j\neq 0} |f_{jk}| e^{\frac{|j||k|_1 s}{2}}$$
$$\leqslant \frac{8\varepsilon}{|k|^2} \frac{e^{-s/2}}{2(1-e^{-s/2})} < \frac{8\varepsilon}{|k|^2} \frac{1}{s} \,.$$

If $|k|_1 \geqslant \mathtt{N}$ one has

$$|\bar{\mathtt{G}}_k|_{\mathtt{s}} = |\bar{\mathtt{G}}_k|_1 \overset{(9)}{=} \frac{4\varepsilon}{|k|^2} |f_k| |\cos(\theta+\theta_k) + F^k_\star(\theta)|_1 \overset{(10)}{\leqslant} \frac{4\varepsilon}{|k|^2} |f_k| (\cosh 1 + 2^{-40}) < \frac{8\varepsilon}{|k|^2} |f_k| \,.$$

Thus, by definitions of $\chi_k$ in (83) and $\mathtt{c}_s$, one gets

$$|\bar{\mathtt{G}}_k|_{\mathtt{s}} \leqslant \epsilon \,, \tag{123}$$

with $\epsilon$ as in (83). Next, since $\chi_k \leqslant 1 \leqslant \mathtt{c}_s$,

$$|\mathtt{G}_k - \bar{\mathtt{G}}_k|_{\mathtt{r},\mathtt{s}} \overset{(106)}{\leqslant} 2\eta \overset{(94)}{=} \frac{8\varepsilon}{|k|^2} \chi_k \vartheta \overset{(83)}{\leqslant} \epsilon\mu \,. \tag{124}$$

By (122), (83), (94), using the inequalities $|k| \leqslant \mathtt{K}_{\rm o} \leqslant \mathtt{K}/6$, recalling (85), (30), and the hypothesis $\mathtt{K}_{\rm o} \geqslant c_2$ (in the last inequality), one sees that

$$|\nu_k|_{\mathtt{r},\mathtt{s}} \leqslant c_2^2 \frac{|k|^2}{\mathtt{K}^{2\nu}} \vartheta \leqslant \frac{c_2^2}{36} \frac{1}{\mathtt{K}^{2(\nu-1)}} \vartheta < \vartheta = \mu \,. \tag{125}$$

Then (79) follows by (123), (124) and (125).
Finally, observe that, by the definitions in (83) and (94) one has

$$\epsilon/\beta = \begin{cases} \frac{4\mathtt{c}_s}{\beta} \,, & \text{if } |k|_1 < \mathtt{N} \,, \\ 4\mathtt{c}_s \,, & \text{if } |k|_1 \geqslant \mathtt{N} \,. \end{cases} \tag{126}$$

Then, (81) with $\kappa$ as in (87), follows immediately by the definitions in (83), (80) and (126).

It remains to prove claim (iii): From (118), it follows that $\mathtt{g}_3 = \mathfrak{p}_{\rm o}$, and from (103), it follows that $\eta_2 = \tilde{\mathfrak{p}}$ and $\chi_2 = \hat{\mathfrak{q}}$. The relative estimates follow immediately by (102), (94), (40) and (83). □

**Data availability**

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

## Acknowledgments

This work has been partly supported by the Italian PRIN grant PNRR-M4C2-I1.1-PRIN 2022-PE1-Stability in Hamiltonian dynamics and beyond-F53D23002730006-Finanziato dall'U.E.-NextGenerationEU.

## References

[1] A.A. Arabanov, A.D. Morozov, On resonances in Hamiltonian systems with three degrees of freedom, Regul. Chaotic Dyn. 24 (6) (2019) 628–648.
[2] V.I. Arnol'd, Instability of dynamical systems with many degrees of freedom, Dokl. Akad. Nauk SSSR 156 (1964) 9–12.
[3] V.I. Arnold, V.V. Kozlov, A.I. Neishtadt, Mathematical Aspects of Classical and Celestial Mechanics, third edition, Encyclopaedia of Mathematical Sciences, vol. 3, Springer-Verlag, Berlin, 2006, Dynamical systems. III, Translated from the Russian original by E. Khukhro.
[4] G. Benettin, L. Galgani, A. Giorgilli, A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems, Celest. Mech. 37 (1) (1985) 1–25.
[5] L. Biasco, L. Chierchia, Complex Arnol'd–Liouville maps, preprint, arXiv:2306.00875v1, 2023.
[6] L. Biasco, L. Chierchia, On the topology of nearly-integrable Hamiltonians at simple resonances, Nonlinearity 33 (2020) 3526–3567.
[7] L. Biasco, L. Chierchia, Quasi–periodic motions in generic nearly–integrable mechanical systems, Atti Accad. Naz. Lincei, Cl. Sci. Fis. Mat. Nat. 33 (3) (2022) 575–580.
[8] A. Bounemoura, B. Fayad, L. Niederman, Superexponential stability of quasi–periodic motion in Hamiltonian systems, Commun. Math. Phys. 350 (1) (2017) 361–386.
[9] P. Bernard, V. Kaloshin, K. Zhang, Arnol'd diffusion in arbitrary degrees of freedom and 3–dimensional normally hyperbolic invariant cylinders, Acta Math. 217 (1) (2016) 1–79.
[10] Q. Chen, R. de la Llave, Analytic genericity of diffusing orbits in a priori unstable Hamiltonian systems, Nonlinearity 35 (4) (2022) 1986–2019.
[11] L. Chierchia, G. Gallavotti, Drift and diffusion in phase space, Ann. Inst. Henri Poincaré, Phys. Théor. 60 (1994) 1–144, Erratum: Ann. Inst. Henri Poincaré, Phys. Théor. 68 (1) (1998) 135.
[12] A. Clarke, J. Fejoz, M. Guardia, Topological shadowing methods in Arnold diffusion: weak torsion and multiple time scales, Nonlinearity 36 (1) (2023) 426–457.
[13] A. Delshams, P. Gutiérrez, Effective stability and KAM theory, J. Differ. Equ. 128 (1996) 415–490.
[14] A. Delshams, G. Huguet, Geography of resonances and Arnol'd diffusion in a priori unstable Hamiltonian systems, Nonlinearity 22 (8) (2009) 1997–2077.
[15] A. Delshams, R. de la Llave, T.M. Seara, Instability of high dimensional Hamiltonian systems: multiple resonances do not impede diffusion, Adv. Math. 294 (2016) 689–755.
[16] A. Delshams, R.G. Schaefer, Arnold diffusion for a complete family of perturbations, Regul. Chaotic Dyn. 22 (1) (2017) 78–108.
[17] M. Guardia, J. Paradela, T.M. Seara, C. Vidal, Symbolic dynamics in the restricted elliptic isosceles three body problem, J. Differ. Equ. 294 (2021) 143–177.
[18] M. Guzzo, L. Chierchia, Benettin, The steep Nekhoroshev's theorem, Commun. Math. Phys. 342 (2) (2016) 569–601.
[19] J.-L. Figueras, A. Haro, A. Luque, Effective bounds for the measure of rotations, Nonlinearity 33 (2) (2020) 700–741.
[20] R.C. Gunning, H. Rossi, Analytic Functions of Several Complex Variables, American Mathematical Soc., 2009, 317 pp.
[21] V. Kaloshin, K. Zhang, Arnol'd Diffusion for Smooth Systems of Two and a Half Degrees of Freedom, Annals of Mathematics Studies, vol. 208, Princeton University Press, Princeton, NJ, 2020.
[22] J.N. Mather, Arnol'd diffusion by variational methods, in: Essays in Mathematics and Its Applications, Springer, Heidelberg, 2012, pp. 271–285.
[23] A.G. Medvedev, A.I. Neishtadt, D.V. Treschev, Lagrangian tori near resonances of near–integrable Hamiltonian systems, Nonlinearity 28 (7) (2015) 2105–2130.
[24] A.I. Neishtadt, Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions, J. Appl. Math. Mech. 45 (6) (1981) 766–772.

[25] N.N. Nehoroshev, A. Neishtadt, Problems of Perturbation Theory for Non–Linear Resonant Systems, Doktor, Diss., Moscow Univ., Moscow, 1989, 342 pp. (Russian).
[26] N.N. Nehoroshev, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems (in Russian) Usp. Mat. Nauk 32 (6(198)) (1977) 5–66, 287.
[27] L. Niederman, Dynamics around simple resonant tori in nearly integrable Hamiltonian systems, J. Differ. Equ. 161 (1) (2000) 1–41.
[28] L. Niederman, A. Pousse, P. Robutel, On the co-orbital motion in the three–body problem: existence of quasi–periodic horseshoe–shaped orbits, Commun. Math. Phys. 377 (1) (2020) 551–612.
[29] J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, Commun. Pure Appl. Math. 35 (5) (1982) 653–696.
[30] J. Pöschel, Nekhoroshev estimates for quasi–convex Hamiltonian systems, Math. Z. 213 (1993) 187.
[31] C. Simó, Some questions looking for answers in dynamical systems, Discrete Contin. Dyn. Syst. 38 (12) (2018) 6215–6239.
[32] D. Treschev, Arnol'd diffusion far from strong resonances in multidimensional a priori unstable Hamiltonian systems, Nonlinearity 25 (9) (2012) 2717–2757.
[33] Ke Zhang, Speed of Arnol'd diffusion for analytic Hamiltonian systems, Invent. Math. 186 (2) (2011) 255–290.