On the topology of nearly–integrable Hamiltonians at simple resonances

L. Biasco & L. Chierchia

Dipartimento di Matematica e Fisica

Università degli Studi Roma Tre

Largo San L. Murialdo 1 - 00146 Roma, Italy

biasco@mat.uniroma3.it, luigi@mat.uniroma3.it

February 4, 2019

Abstract

We show that, in general, averaging at simple resonances a real–analytic, nearly–integrable Hamiltonian, one obtains a one–dimensional system with a cosine–like potential; "in general" means for a generic class of holomorphic perturbations and apart from a finite number of simple resonances with small Fourier modes; "cosine–like" means that the potential depends only on the resonant angle, with respect to which it is a Morse function with one maximum and one minimum.
Furthermore, the (full) transformed Hamiltonian is the sum of an effective one–dimensional Hamiltonian (which is, in turn, the sum of the unperturbed Hamiltonian plus the cosine–like potential) and a perturbation, which is exponentially small with respect to the oscillation of the potential.
As a corollary, under the above hypotheses, if the unperturbed Hamiltonian is also strictly convex, the effective Hamiltonian at *any simple resonance* (apart a finite number of low–mode resonances) has the phase portrait of a pendulum.
The results presented in this paper are an essential step in the proof (in the "mechanical" case) of a conjecture by Arnold–Kozlov–Neishdadt ([2, Remark 6.8, p. 285]), claiming that the measure of the "non–torus set" in general nearly–integrable Hamiltonian systems has the same size of the perturbation; compare [4], [3].

Contents

1 Introduction

Consider a real–analytic, nearly–integrable Hamiltonian given, in action–angle variables, by

$$H_\varepsilon(y,x) = h(y) + \varepsilon f(y,x) \ , \quad (y,x) \in \mathcal{M} := D \times \mathbb{T}^n \ , \tag{1}$$

where D is a bounded domain in $\mathbb{R}^n$, $\mathbb{T}^n = \mathbb{R}^n/(2\pi\mathbb{Z}^n)$ is the usual flat n dimensional torus and ε is a small parameter measuring the size of the perturbation εf. The phase space $\mathcal{M}$ is endowed with the standard symplectic form $dy \wedge dx$ so that the Hamiltonian flow $\phi^t_{H_\varepsilon}(y_0,x_0) =: (y(t),x(t))$ governed by H_ε is the solution of the standard Hamiltonian equations

$$\begin{cases} \dot y = -\partial_x H_\varepsilon(y,x) \ , \\ \dot x = \partial_y H_\varepsilon(y,x) \ , \end{cases} \qquad \begin{cases} y(0) = y_0 \ , \\ x(0) = x_0 \ , \end{cases} \tag{2}$$

(where t is time and dot is time derivative).

It is well known that, in general, the $\phi^t_{H_\varepsilon}$–dynamics is strongly influenced by *resonances of the (unperturbed) frequencies* $\omega(y) := h'(y) = \partial_y h(y)$, i.e., by rational relations

$$\omega(y)\cdot k = \sum_{j=1}^{n} \omega_j(y) k_j = 0\,,$$

with $k \in \mathbb{Z}^n \setminus \{0\}$; for general information, compare, e.g., [2]. Indeed, assuming a standard KAM non–degeneracy assumption on h, e.g., that the frequency map $y \in D \to \omega(y)$ is a real–analytic diffeomorphism of D onto the "frequency space" $\Omega := h'(D)$, then the action space D can be covered by three open sets

$$D \subseteq D^0 \cup D^1 \cup D^2 \tag{3}$$

so that the following holds. Roughly speaking, $D^0 \times \mathbb{T}^n$ is a fully non–resonant set which is filled, up to an exponentially small set, by *primary KAM tori*, namely, by homotopically trivial, Lagrangian tori $\phi^t_{H_\varepsilon}$–invariant on which the flow is analytically conjugated to the linear flow

$$\theta \in \mathbb{T}^n \mapsto \theta + \omega t$$

with ω satisfying a Diophantine condition

$$|\omega \cdot k| \geqslant \frac{\gamma}{|k|_1^\tau} \qquad \forall\ k \in \mathbb{Z}^n \setminus \{0\}\ , \tag{4}$$

(for some $\gamma, \tau > 0$); $\omega \cdot k$ denoting the standard inner product $\sum \omega_i k_i$ and $|k|_1 := \sum |k_i|$. Furthermore, such tori are deformation of integrable tori.

D^1 is an open $O(\sqrt{\varepsilon})$–neighbourhood of *simple resonances* (i.e., of regions where *exactly one* independent resonance $\omega(y)\cdot k = 0$ holds) and D^2 is a set of measure $O(\varepsilon)$; compare the Covering Lemma (Proposition 2.1) below[1].

The region D^2 contains double (and higher) resonances and, in general, in $D^2 \times \mathbb{T}^n$ there are $O(\varepsilon)$ regions where the dynamics is *non–perturbative*, being "essentially" governed (after suitable rescalings) by an ε–independent Hamiltonian; compare[2] [2].

[1] This description follows by choosing carefully certain parameters (such as the "small divisor constant" α and "Fourier cut–offs" K) as functions of ε and disregarding logarithmic corrections.

[2] [2, Remark 6.8, p. 285]: "*It is natural to expect that in a generic system with three or more degrees of freedom the measure of the "non–torus" set has order ε. Indeed, the $O(\sqrt{\varepsilon})$–neighbourhoods of two resonant surfaces intersect in a domain of measure $\sim \varepsilon$. In this domain, after the partial averaging taking into account the resonances under consideration, normalizing the deviations of the "actions" from the resonant values by the quantity $\sqrt{\varepsilon}$, normalizing time, and discarding the terms of higher order, we obtain a Hamiltonian of the form $1/2(Ap,p) + V(q_1,q_2)$, which does not involve a small parameter (see the definition of the quantity p above). Generally speaking, for this Hamiltonian there is a set of measure ~ 1 that does not contain points of invariant tori. Returning to the original variables we obtain a "non–torus" set of measure $\sim \varepsilon$.*"

The dynamics in the simple–resonance region $D^1 \times \mathbb{T}^n$ is particularly relevant and interesting. For example, it plays a major rôle in *Arnold diffusion*, as showed by Arnold himself [1], who based his famous instability argument on shadowing partially hyperbolic trajectories arising near simple resonances.
On the other hand, in $D^1 \times \mathbb{T}^n$ there appear *secondary KAM tori*, namely n–dimensional KAM tori with different topologies, which depend upon specific characteristics of the perturbation εf. The appearance of secondary tori is a genuine non–integrable effect, since such tori do not exist in the integrable regime.
In the announcement [4] it is claimed that, in the case of mechanical systems – namely systems governed by Hamiltonians of the form $|y|^2/2 + \varepsilon f(x)$ – and for generic potentials f, primary and secondary tori fill the region $D^1 \times \mathbb{T}^n$ up to a set of measure nearly exponentially small, showing that the "non–torus set" is, at most, $O(\varepsilon)$ as conjectured in[3] [2] and studied in [11]. In fact, Theorem 2.1 below is one of the building block of the proof (in the mechanical case) of the Arnold–Kozlov–Neishdadt conjecture as outlined in [3].

This paper is devoted to the fine topological and quantitative analysis of the behaviour of generic systems in the simple resonant region $D^1 \times \mathbb{T}^n$.

In this introduction, we briefly discuss the main aspects of this analysis in the particular case of *purely positional potentials*; precise statements are given in Theorem 2.1 of § 2 below, and, for the general (but more technical and implicit) case, in Theorem 7.2 of § 7.2).
D^1 is the union of suitable regions $D^{1,k}$, which are $O(\sqrt{\varepsilon})$-close to *exact simple resonances* $\{y \in D |\ \omega(y)\cdot k = 0\}$, and which are labelled by *generators* k of one dimensional, maximal sublattices of $\mathbb{Z}^n$ (see (21) below); "exact" meaning that $\omega(y)$ does not verify double or higher resonant relations.
In averaging (or normal form) theory, one typically considers a *finite* but large (possibly, ε–dependent) number of simple resonances. More precisely, one considers generators k with $|k|_1 \leqslant K$, and K can be chosen according to the application one has in mind. Typically, one chooses $K \sim 1/\varepsilon^a$ for a suitable $a > 0$ (as in Nekhoroshev theorem [15], [7]) or $K \sim |\log \varepsilon|^a$ (as in the KAM theory for secondary tori of [4], [3]).
By averaging theory, in any fixed simple resonant region $D^{1,k}$, one can remove the non–resonant angle dependence, so as to symplectically conjugate H_ε, for ε small enough,

[3] See footnote 2 above. Note also that, as it was proved in [9] (in dimension 2) and [12], [14] (in any dimension), the union of *primary* invariant tori fills the phase space up to a set of measure $O(\sqrt{\varepsilon})$. This result is optimal: the phase region inside the separatrix of the pendulum $\frac{1}{2}y^2 + \varepsilon \cos x$, with $y \in \mathbb{R}$ and $x \in \mathbb{T}^1$, does not contain any primary invariant torus, namely a circle which is a global graph over the angle on $\mathbb{T}^1$, and this region has measure $4\sqrt{2\varepsilon}$. Indeed this region is filled by secondary tori, corresponding to oscillations of the pendulum.

to a Hamiltonian of the form

$$h(y) + \varepsilon G^k(y, k \cdot x) + \varepsilon R^k(y, x) \tag{5}$$

where $\theta \to G^k(y, \theta)$ is a function of one angle and R^k is a "very small" remainder with[4] $R^k_{jk}(y) = 0, \forall j \in \mathbb{Z}$. Thus, up to the remainder R^k, the Hamiltonian depends effectively only on the "resonant angle" $\theta := k \cdot x$ and therefore the "effective Hamiltonian" $h + \varepsilon G^k$ is *integrable*: this is the starting point for ("a priori stable") Arnold diffusion or for the KAM theory for secondary tori of [4]–[3].

Obviously, there are here two main issues:

(a) *What is the actual "generic form" of G^k?*

(b) *How small (and compared to what) is the remainder R^k?*

(a) According to averaging (or normal form) theory[5] G^k is "close" to the projection of the potential f on the Fourier modes of the resonant maximal sublattice $k\mathbb{Z}$:

$$\mathrm{p}_{k\mathbb{Z}} f(x) = \sum_{j \in \mathbb{Z}} f_{jk} e^{\mathrm{i} jk \cdot x} . \tag{6}$$

Now, since f is real-analytic on $\mathbb{T}^n$, it is holomorphic in a complex strip $\mathbb{T}^n_s$ around $\mathbb{T}^n$ of width $s > 0$ and its Fourier coefficients decay exponentially fast as[6] $|f_k| \sim \|f\| e^{-|k|s}$. Hence, *typically* (i.e., if $f_k \neq 0$)

$$\sum_{j \in \mathbb{Z}} f_{jk} e^{\mathrm{i} jk \cdot x} = f_k e^{\mathrm{i} k \cdot x} + f_{-k} e^{-\mathrm{i} k \cdot x} + O\big(\|f\| e^{-2|k|s}\big)$$

which, by the reality condition $f_{-k} = \bar{f}_k$, can be written as

$$\sum_{j \in \mathbb{Z}} f_{jk} e^{\mathrm{i} jk \cdot x} = 2|f_k| \cos\big(k \cdot x + \theta^{(k)}\big) + O\big(\|f\| e^{-2|k|s}\big) \tag{7}$$

for a suitable $\theta^{(k)} \in [0, 2\pi)$. Thus,

$$\mathrm{p}_{k\mathbb{Z}} f(x) = \sum_{j \in \mathbb{Z}} f_{jk} e^{\mathrm{i} jk \cdot x} = 2|f_k| \Big(\cos\big(k \cdot x + \theta^{(k)}\big) + o(1) \Big) \tag{8}$$

[4] $R^k_\ell(y)$ denotes the ℓ^{th} Fourier coefficient of $x \mapsto R^k(y, x)$.

[5] For generalities on Averaging Theory, see, e.g., [2, § 6] and references therein.

[6] Precise norms will be introduced in the next section § 2.

provided

$$|k| \gtrapprox \frac{1}{s} \; . \tag{9}$$

In other words, one expects (8) to hold for *generic real–analytic potentials* and for all generators k's satisfying (9).

Indeed, this is the case: we shall introduce certain classes of periodic holomorphic functions $\mathcal{H}_{s,\tau}$, for which (choosing suitably the "tail" fuction τ) (8) holds for generators k satisfying (9).

The class $\mathcal{H}_{s,\tau}$ turns out to be "generic" in several ways:

(i) it contains an open dense set in the class of real–analytic functions having holomorphic extension on a complex neighbourhood of size s of $\mathbb{T}^n$ (in the topology induced by a suitably weighted Fourier norm);

(ii) its unit ball is of measure 1 (with respect to a natural probability measure);

(iii) it is a "prevalent set".

For precise statements see Definition 2.1 and Proposition 3.1 below.

Next, in order for G^k to be close to $\mathtt{p}_{k\mathbb{Z}} f$ in (8), one needs to have a bound of the type

$$\sup_{D^{1,k}\times\mathbb{T}^n} |G^k - \mathtt{p}_{k\mathbb{Z}} f| \ll |f_k| \sim \|f\| e^{-|k|_1 s} \, . \tag{10}$$

As well known, averaging methods involve an *analyticity loss* in complex domains. In particular, the Hamiltonian in (5) and, therefore, G^k, can be analytically defined only in a *smaller* complex strip $\mathbb{T}^n_{s_\star}$ with $s_\star < s$. Therefore, by analyticity arguments, the best one can hope for is an estimate of the type

$$\sup_{D^{1,k}\times\mathbb{T}^n} |G^k - \mathtt{p}_{k\mathbb{Z}} f| \leqslant c \cdot \|f\| e^{-|k|_1 s_\star} \; , \tag{11}$$

for a suitable constant c that can be taken to be smaller than any prefixed positive number. But then, for (11) and (10) to be compatible one sees that one must "essentially" have $s_\star \sim s$ and that standard averaging theory is not enough[7]. To overcome this problem, we provide (Section 4) a normal form lemma with small analyticity loss, "small" meaning that one can take

$$s_\star = s(1 - 1/K) \tag{12}$$

(compare, in particular, (74)). The value (12) is compatible with (10) for $|k|_1 \leqslant K$, showing that, indeed, generically, one has

$$G^k(y, k\cdot x) = 2|f_k| \Big(\cos\big(k\cdot x + \theta^{(k)}\big) + o(1) \Big) \; .$$

[7]Compare, e.g., [15], where $s_\star = s/6$. For a more detailed comparison with the averaging lemma of [15], see also Remark 4.1–(iv) below. Compare also [5] and [6].

In particular we prove that: *The "effective Hamiltonians" $h + \varepsilon G^k$, as k vary ($c/s \leqslant |k| \leqslant K$), have (up to a phase–shift) the same cosine–like form* and, hence, the same topological feature; compare, also, Remark 2.2–(i) below.

Notice also that on low modes this last property, in general, does not holds, as one immediately sees by considering $k = e_1 = (1, 0, \ldots, 0)$ and a potential f such that

$$\mathtt{p}_{e_1\mathbb{Z}} f(x) := \cos x_1 + \cos 2x_1 ,$$

which is a Morse function with two maxima and two minima in $\mathbb{T}^1$.

(b) What we just discussed gives also an indication for the question "with respect to what R^k has to be small". In fact, if, as expected, (8) is the leading behaviour, one should have

$$\|R^k\| \ll |f_k| \ . \tag{13}$$

But in order to perform averaging procedures, one has, typically, control on small divisors up to the truncation order K, so that the remainder will contain high Fourier modes, $|k| \sim K$, of the potential f. Such terms are bounded by $O(e^{-Ks})$, which are of the *same size* of $\|G^k\|$, at least for $|k| \lesssim K$.

To overcome this problem, we introduce in § 5, at difference with standard geometry of resonances (such as in [13], [15], [7]), *two* Fourier cut offs $K_2 \geqslant 3K_1$ in such a way that on the simple resonant regions $D^{1,k}$ one has non–resonance conditions for double and higher resonances up to order K_2, while K_1 is the maximum value of the size of the generators k (i.e., $|k| \leqslant K_1$). Therefore, we will get an estimate of the remainder R^k of the type

$$\|R^k\| \leqslant C\, K_2^a e^{-K_2 s/2} \leqslant C'\, |f_k|\; e^{-K_2 s/8} \ , \tag{14}$$

for suitable constants $C, C' > 0$. The final upshot is the *complete normal form*

$$H_\varepsilon \circ \Psi_k =: h(y) + 2|f_k|\varepsilon\Big(\cos(k \cdot x + \theta^{(k)}) + \tilde{G}^k(y, k \cdot x) + \tilde{R}^k(y, x)\Big) \tag{15}$$

with $\|\tilde{G}^k\| \ll 1$ and $\|\tilde{R}^k\| \leqslant C'\, e^{-K_2 s/8}$; compare Theorem 2.1 below and, in particular, formula (42).

Summarizing: *for all k large enough, G^k is "cosine–like" i.e. a Morse function with one maximum and one minimum (compare Remark 2.2–(i) below) and R^k is exponentially small with respect to the oscillations of G^k (see (42) and (44) below).*

As a consequence we get that, *if $h(y)$ is strictly convex, the effective Hamiltonian has a phase portrait of a pendulum (compare Remark 2.2–(iv) below).*

2 Statements

Assume that H_ε in (1), for some $r, s > 0$, admits holomorphic extension on the complex domain $D_r \times \mathbb{T}^n_s$, where $D_r \subseteq \mathbb{C}^n$ is the open complex neighbourhood of D formed by points $z \in \mathbb{C}^n$ such that[8] $|z - y| < r$, for some $y \in D$ and $\mathbb{T}^n_s$ denotes the open complex neighbourhood of $\mathbb{T}^n$ given by

$$\mathbb{T}^n_s := \{x = (x_1, ..., x_n) \in \mathbb{C}^n : \ |\operatorname{Im} x_j| < s\}/(2\pi\mathbb{Z}^n) \ . \tag{16}$$

The integrable hamiltonian h is supposed to be "KAM non–degenerate" in the following sense.

Assumption A *Let h be a real-analytic function*

$$h : y \in D \subset \mathbb{R}^n \mapsto h(y) \in \mathbb{R} \ , \quad (n \geqslant 2) \ , \tag{17}$$

where D is a bounded domain of $\mathbb{R}^n$ and such that the frequency map

$$y \in D \mapsto \omega(y) := \partial_y h(y) \in \Omega := \omega(D) \subseteq B_M(0) \subset \mathbb{R}^n \ , \qquad M := \sup_D |\omega(y)| \ , \tag{18}$$

is a global diffeomorphism of D onto Ω with Lipschitz constants given by

$$|y - y_0|\bar{L}^{-1} \leqslant |\omega(y) - \omega(y_0)| \leqslant L|y - y_0| \ , \quad (\forall \ y, y_0 \in D) \ . \tag{19}$$

Now, we describe the covering of frequency/action domain, which allows to apply averaging theory (Proposition 4.1 below) to a perturbation of an integrable system with Hamiltonian h at non–resonant (modulus a lattice) zones[9].

Here, the main point is to find a suitable covering of simple resonances, which are the regions where the averaged Hamiltonian is integrable[10], up to a small remainder. All other higher–order resonances are covered by one set, which is of small measure: how small depending on the choice of the various parameters involved and it will vary according to the applications one has in mind.

Let $\mathbb{Z}^n_\star$ denote the set of integer vectors $k \neq 0$ in $\mathbb{Z}^n$ such that the first non–null component is positive:

$$\mathbb{Z}^n_\star := \big\{k \in \mathbb{Z}^n : \ k \neq 0 \text{ and } k_j > 0 \text{ where } j = \min\{i : k_i \neq 0\}\big\} \ , \tag{20}$$

[8] We denote by $|\cdot|$ the usual Euclidean norm.

[9] We use here the term "zone" in a loose way, *not* in the technical meaning of Nekhoroshev's Theory; compare, e.g., [7].

[10] In the sense that the, up to a small remainder, the averaged Hamiltonian depends on one angle.

and denote by $\mathcal{G}_1^n$ the *generators of 1d maximal lattices*, namely, the set of vectors $k \in \mathbb{Z}^n_\star$ such that the greater common divisor (gcd) of their components is 1:

$$\mathcal{G}_1^n := \{k \in \mathbb{Z}^n_\star : \ \gcd(k_1, \ldots, k_n) = 1\} \ . \tag{21}$$

Then, *the list of one–dimensional maximal lattices is given by the sets $\mathbb{Z}k$ with $k \in \mathcal{G}_1^n$.* Given $K > 0$ we set[11]

$$\mathcal{G}_{1,K}^n := \mathcal{G}_1^n \cap \{|k|_{{}_1} \leqslant K\} \, . \tag{22}$$

Proposition 2.1 (Covering Lemma) *Let h and ω be as in Assumption* **A** *(§ 2) and fix $K_{{}_2} \geqslant K_{{}_1} \geqslant 2$ and $\alpha > 0$. Then, the domain D can be covered by three sets $D^i \subseteq D$,*

$$D = D^0 \cup D^1 \cup D^2 \ , \tag{23}$$

so that the following holds.

(i) *D^0 is $(\alpha/2, K_{{}_1})$ completely non–resonant (i.e., non–resonant modulus $\{0\}$), namely,*

$$y \in D^0 \quad \Longrightarrow \quad |\omega(y) \cdot k| \geqslant \alpha/2 \, , \quad \forall \ \ 0 < |k|_{{}_1} \leqslant K_{{}_1} \, . \tag{24}$$

(ii) $D^1 = \displaystyle\bigcup_{k \in \mathcal{G}^n_{1,K_1}} D^{1,k}$*, where, for each*[12] *$k \in \mathcal{G}^n_{1,K_1}$, $D^{1,k}$ is a neighbourhood of a simple resonance $\{y \in D : \omega(y) \cdot k = 0\}$, which is $(2\alpha K_{{}_2}/|k|, K_{{}_2})$ non–resonant modulo $\mathbb{Z}k$, namely,*

$$y \in D^{1,k} \quad \Longrightarrow \quad |\omega(y) \cdot \ell| \geqslant 2\alpha K_{{}_2}/|k| \, , \quad \forall \quad \ell \in \mathbb{Z}^n \ , \ \ell \notin \mathbb{Z}k \ , \ |\ell|_{{}_1} \leqslant K_{{}_2} \ . \tag{25}$$

(iii) *D^2 contains all the resonances of order two or more and has Lebesgue measure small with α^2: more precisely, there exists a constant $c > 0$ depending only on n such that*

$$\mathrm{meas}\,(D^2) \leqslant c \ \bar{L}^n M^{n-2} \ \alpha^2 \ K_{{}_2}^{n+1} K_{{}_1}^{n-1} \ . \tag{26}$$

Remark 2.1 (i) The neighbourhoods $D^{1,k}$ of simple resonances $\{y : \omega(y) \cdot k = 0\}$ are explicitly defined as follows. Denote by $\mathtt{p}_k^\perp$ the orthogonal projection on the subspace perpendicular to[13] k and, for $k \in \mathcal{G}^n_{1,K_1}$, define

$$\Omega^{1,k} := \Big\{\omega \in \mathbb{R}^n : |\omega \cdot k| < \alpha, \ |\,\mathtt{p}_k^\perp \omega| < M, \text{ and } |\,\mathtt{p}_k^\perp \omega \cdot \ell| > \frac{3\alpha K_{{}_2}}{|k|} \, , \forall \ell \in \mathcal{G}^n_{1,K_2} \setminus \mathbb{Z}k\Big\} \tag{27}$$

[11] $|k|_{{}_1} := \sum_{1 \leqslant i \leqslant n} |k_i|$.

[12] Recall (22).

[13] Explicitly, $\mathtt{p}_k^\perp \omega := \omega - \frac{1}{|k|^2}(\omega \cdot k) k$.

Then,

$$D^{1,k} := \{y \in D : \omega(y) \in \Omega^{1,k}\}\,. \tag{28}$$

(ii) The domains D^0, D^2 are explicitly defined in (146), (130) and (134) below.

(iii) The simply resonant regions $D^{1,k}$ in the above Proposition are labelled by generators of 1–d maximal lattices $k \in \mathcal{G}_1^n$ up to size $|k|_{_1} \leqslant K_{_1}$, however, *the non–resonance condition* (25) *holds for integer vectors* ℓ *with* $|\ell|_{_1}$ *up to a (possibly) larger order* $K_{_2}$. This improvement (with respect to having $K_{_2} = K_{_1}$ as, e.g., in [15]) is technical but important if one wants to have sharp control over the averaged Hamiltonian in a normal form near simple resonances; in particular in order to obtain (159) and (182), which lead to (44).

(iv) The non–resonance relations (24) and (25) allow to apply averaging theory and to remove the dependence upon the "non–resonant angle variables" up to exponential order; for precise statements, see Theorem 6.1 in § 6.

We proceed, now, to describe the generic non–degeneracy assumption on periodic holomorphic functions, which will allow to state the main theorem (for the case of positional potentials).

If $s > 0$, we denote by $\mathbb{B}^n_s$ the Banach space of *real–analytic* functions on $\mathbb{T}^n_s$ having zero average and finite ℓ^∞–Fourier norm:

$$\mathbb{B}^n_s := \Big\{ f = \sum_{\substack{k\in\mathbb{Z}^n\\ k\neq 0}} f_k e^{\mathrm{i}k\cdot x} \,:\, \|f\|_s := \sup_{k\in\mathbb{Z}^n} |f_k| e^{|k|_1 s} < \infty \Big\}\,. \tag{29}$$

Note that $f \in \mathbb{B}^n_s$ can be uniquely written as:

$$f(x) = \sum_{k\in\mathcal{G}_1^n} \sum_{j\in\mathbb{Z}\setminus\{0\}} f_{jk} e^{\mathrm{i}jk\cdot x} \tag{30}$$

For functions[14] $f : D_r \times \mathbb{T}^n_s \to \mathbb{C}$ we will also use the (stronger) norm[15]

$$\lvert f \rvert_{D,r,s} = \lvert f \rvert_{r,s} := \sup_{y\in D_r} \sum_{k\in\mathbb{Z}^n} |f_k(y)| e^{|k|_1 s}\,. \tag{31}$$

Definition 2.1 (Non degenerate potentials) *A tail function* τ *is, by definition, a non–increasing, non–negative continuous function*

$$\tau : \delta \in (0,1] \mapsto \tau(\delta) \geqslant 0\ .$$

[14] Not necessarily holomorphic in y.

[15] See Remark 3.1 for details.

Given $s > 0$ and a (possibly s–dependent) tail function τ, we define, for $\delta \in (0,1]$, $\mathcal{H}_{s,\tau}(\delta)$ as the set of functions in $\mathbb{B}^n_s$ such that, for any generator $k \in \mathcal{G}^n_1$, the following holds[16]

$$\textit{if } |k|_1 > \tau(\delta), \textit{ then } \quad |f_k| \geqslant \delta |k|_1^{-n}\, e^{-|k|_1 s}\,, \tag{32}$$

The class $\mathcal{H}_{s,\tau}$ is the union over δ of the classes $\mathcal{H}_{s,\tau}(\delta)$:

$$\mathcal{H}_{s,\tau} := \bigcup_{0<\delta\leqslant 1} \mathcal{H}_{s,\tau}(\delta)\ . \tag{33}$$

The classes $\mathcal{H}_{s,\tau}$ contain (if the tail is choosen properly) the non degenerate potentials for which Theorem 2.1 below holds and, as mentioned in the Introduction, satisfy three main genericity properties[17], as showed in Proposition 3.1 below (compare, also, Remark 3.2).

Theorem 2.1 *Let $n \geqslant 2$, $s > 0$, $0 < \delta, \gamma \leqslant 1$ such that*

$$\gamma\delta < \frac{2^9}{s^n} e^{-n^2/2}\,. \tag{34}$$

Consider a Hamilonian $H_\varepsilon(y,x) = h(y) + \varepsilon f(x)$ as in (1) where h satisfies the non–degeneracy Assumption **A** *(§ 2) and f is purely positional (i.e., independent of the y–variable) with*

$$\|f\|_s = 1\,. \tag{35}$$

Assume that the potential is non–degenerate in the sense that

$$f \in \mathcal{H}_{s,\tau_{\rm o}}(\delta) \tag{36}$$

with tail function

$$\tau_{\rm o}(\delta;\gamma) := \frac{4}{s}\ \log\Big(e + \frac{2^9}{s^n\gamma\delta}\Big)\,. \tag{37}$$

Let $K_2 \geqslant 3K_1 \geqslant 6$ satisfying

$$K_2^{2\nu-3n-3} \geqslant e^{s+5} 2^{n+11} n^{2n} \frac{L}{s^{2n+1}} \frac{1}{\gamma\delta} \qquad \textit{for some} \quad \nu \geqslant \frac{3}{2}n + 2 \tag{38}$$

[16] One could substitute n with every $\bar n > n/2$ in (32); compare Remark 3.3 below. The "weight" $|k|_1^{-n}$ is necessary in order to show that $\mathcal{H}_{s,\tau}(\delta)$ in (33) has positive measure in a suitable probability space; compare Proposition 3.1–(ii) below.

[17] Such properties hold for *any* tail τ, which can be chosen differently according to the particular problem at hand.

and where L defined in (19). Set

$$r_k := \sqrt{\varepsilon}\frac{K_{_2}^{\nu}}{L|k|} \,. \tag{39}$$

Finally assume that

$$\varepsilon \leqslant \frac{(Lr)^2}{K_{_2}^{2\nu}} \,. \tag{40}$$

Then, for any $k \in \mathcal{G}^n_{1,K_{_1}}$ with $\tau_{\rm o}(\delta;\gamma) \leqslant |k|_{_1} \leqslant K_{_1}$, there exists $\theta^{(k)} \in [0, 2\pi)$ and a symplectic change of variables defined in a neighbourhood of the simple resonance $D^{1,k} \times \mathbb{T}^n$ such that the following holds:

$$\Psi_k : D^{1,k}_{r_k/2} \times \mathbb{T}^n_{s(1-1/K_{_2})^2} \to D^{1,k}_{r_k} \times \mathbb{T}^n_{s(1-1/K_{_2})} \,, \tag{41}$$

and

$$\boxed{H_\varepsilon \circ \Psi_k =: h(y) + 2|f_k|\varepsilon\Big(\cos(k\cdot x + \theta^{(k)}) + \mathtt{G}^k(y, k\cdot x) + \mathtt{f}^k(y,x)\Big)} \tag{42}$$

where $\mathtt{G}^k(y,\cdot) \in \mathbb{B}^1_2$ for every $y \in D^{1,k}_{r_k/2}$ and

$$|\mathtt{G}^k|_{D^{1,k},r_k/2,2} \leqslant \gamma \,. \tag{43}$$

Finally,

$$\mathtt{p}_{k\mathbb{Z}}\mathtt{f}^k = 0 \qquad \text{and} \qquad |\mathtt{f}^k|_{D^{1,k},r_k/2,s(1-1/K_{_2})/2} \leqslant \frac{2^{10n}n^{3n}}{s^{3n}\delta}e^{-K_{_2}s/8} \,. \tag{44}$$

Remark 2.2 (i) Recalling (31), estimate (43) means

$$\sup_{y\in D^{1,k}_{r_k/2}} \sum_{j\in\mathbb{Z}} |\mathtt{G}^k_j(y)|e^{2|j|} \leqslant \gamma \,. \tag{45}$$

This implies that for every $y \in D^{1,k}$ the 2π-periodic real function

$$\theta \mapsto \cos(\theta + \theta^{(k)}) + \mathtt{G}^k(y,\theta) \tag{46}$$

behaves like a cosine in the sense that *it is a Morse function with only one maximum and one minimum and no other critical points.* To prove this, notice that by (45) we have

$$\sup_{y\in D^{1,k},x\in\mathbb{T}^1} |\partial_\theta \mathtt{G}^k(y,\theta)| \leqslant \gamma/e^2 \,, \qquad \sup_{y\in D^{1,k},x\in\mathbb{T}^1} |\partial^2_{\theta\theta}\mathtt{G}^k(y,\theta)| \leqslant \gamma/e^2 \,.$$

Therefore, denoting by $\psi(\theta)$ the derivative of the function in (46), we have that $\psi(\theta) > 0$ for $\theta \in (-\theta^{(k)} + \theta_*, \pi - \theta^{(k)} - \theta_*)$ and $\psi(\theta) < 0$ for $\theta \in (\pi - \theta^{(k)} + \theta_*, 2\pi - \theta^{(k)} - \theta_*)$, where $\theta_* := \arcsin(\gamma/e^2)$. Moreover in the interval $(-\theta^{(k)} - \theta_*, -\theta^{(k)} + \theta_*)$ the function $\psi(\theta)$ has a zero and is strictly increasing since $\psi'(\theta) \geqslant \sqrt{1-\gamma/e^2} - \gamma/e^2 =: c > 0$. Finally in the interval $(\pi - \theta^{(k)} - \theta_*, \pi - \theta^{(k)} + \theta_*)$ it has a zero and is strictly decreasing since $\psi'(\theta) \leqslant -c < 0$.

(ii) As a consequence the phase portrait of the effective Hamiltonian

$$h(y) + 2|f_k|\varepsilon\Big(\cos(k \cdot x + \theta^{(k)}) + \mathtt{G}^k(y, k \cdot x)\Big) \tag{47}$$

and that of the Hamiltonian $h(y) + 2|f_k|\varepsilon \cos(k \cdot x + \theta^{(k)})$ are topologically equivalent.

(iii) As well know the effective Hamiltonian (47) is an integrable system as it depends only on one angle. Indeed, fix $k \in \mathbb{Z}^n \backslash \{0\}$ with $\gcd(k_1, \ldots, k_n) = 1$, then, there exists a matrix $A_k \in \mathrm{Mat}_{n\times n}(\mathbb{Z})$ such that[18]

$$A_k = \begin{pmatrix} \hat{A}_k \\ k \end{pmatrix} \in \mathrm{Mat}_{n\times n}(\mathbb{Z}), \quad \hat{A}_k \in \mathrm{Mat}_{(n-1)\times n}(\mathbb{Z}), \quad \det A_k = 1, \quad |\hat{A}_k|_\infty \leqslant |k|_\infty, \tag{48}$$

where $|\cdot|_\infty$ denotes the sup–norm of the matrix and of the vector, respectively. The existence of such a matrix is guaranteed by an elementary result of linear algebra based on Bezout's Lemma (see Lemma A.1 in Appendix A).
Let us perform the linear symplectic change of variables

$$\Phi_k : (Y, X) \mapsto (y, x) := (A_k^T Y, A_k^{-1} X), \tag{49}$$

which is generated by the generating function $S(Y, x) := Y \cdot A_k x$. Note that Φ_k does not mix actions with angles, its projection on the angles is a diffeomorphism of $\mathbb{T}^n$ onto $\mathbb{T}^n$, and, most relevantly, $X_n = k \cdot x$ is the "secular angle".
In the (Y, X)–variables, the secular Hamiltonian in (47) takes the form

$$\mathtt{h}(Y) + 2|f_k|\varepsilon\Big(\cos(X_n + \theta^{(k)}) + \mathtt{G}^k(A_k^T Y, X_n)\Big), \qquad \text{with} \qquad \mathtt{h}(Y) := h(A_k^T Y). \tag{50}$$

Fix $y_0 \in D^{1,k}$ on the exact resonance, namely $\partial_y h(y_0) \cdot k = 0$. Let Y_0 be such that $y_0 = A_k^T Y_0$. We have

$$\partial_{Y_n} \mathtt{h}(Y_0) \overset{(48)}{=} \partial_y h(A_k^T Y_0) \cdot k = \partial_y h(y_0) \cdot k = 0, \qquad \partial^2_{Y_n Y_n} \mathtt{h}(Y_0) = \partial^2_{yy} h(y_0) k \cdot k,$$

[18] Here, k is a row vector.

where $\partial^2_{yy} h$ is the Hessian matrix of h. By Taylor expansion the secular Hamiltonian in (50) takes the form (up to an addictive constant)

$$\frac{1}{2}\Big(\partial^2_{yy} h(y_0) k\cdot k\Big)(Y_n - Y_{0n})^2 + O\Big((Y_n - Y_{0n})^3\Big) + 2|f_k|\varepsilon\Big(\cos(X_n+\theta^{(k)}) + \mathtt{G}^k(A_k^T Y, X_n)\Big)\,. \tag{51}$$

(iv) In particular if the Hamiltonian h is *convex* the coefficient $\partial^2_{yy} h(y_0)k\cdot k =: m_k$ is bounded away from zero and the phase portrait of the secular Hamiltonian in (51) is topologically equivalent, for $|Y_n - Y_{0n}|$ small[19], to that of the pendulum

$$\frac{1}{2} m_k (Y_n - Y_{0n})^2 + 2|f_k|\varepsilon \cos(X_n + \theta^{(k)})\,.$$

The y-dependent case.

Let us briefly turn to the y-dependent case. First we note that it can happen that, even if the potential $f(y,x)$ satisfies the non-degeneracy condition given in Definition 2.1 at some point y_0, there is no neighborhood of y_0 on which the non-degeneracy condition holds. For example consider the potential

$$f(y,x) = f(y_1, x) := \frac{1}{2}\sum_{k\neq 0}\Big(|k|_1^{-n} - \frac{y_1}{r}\Big) e^{-|k|_1 s} e^{\mathrm{i}k\cdot x}\,.$$

We have that $\|f\|_{D,r,s} = 1$ with $D = \{0\}$ and

$$f(0,\cdot) \in \mathcal{H}_{s,0}(1/2)\,.$$

However, $f_k(rj^{-n}) = 0$ for every $|k| = j$; in particular for every odd number $j = 2h+1$, $h \geqslant 1$ and $k := (h+1, h, 0, \ldots, 0) \in \mathcal{G}^n_1$. Then for every $\delta > 0$, tail function $\tau > 0$ and odd $j \geqslant 3$, we have

$$f(rj^{-n},\cdot) \notin \mathcal{H}_{s,\tau}(\delta)\,.$$

Then, we will prove that the non-degeneracy condition holds in a set of large measure. In particular we fix $\mu > 0$ and prove that, if for a certain point $y_0 \in D$ the potential $f(y_0,\cdot) \in \mathcal{H}_{s,\tau_*}(\delta)$ for a suitable $\tau_* = \tau_*(\mu)$ and $|k| \geqslant \tau_*(\mu)$, then (42)-(44) holds (with $f_k = f_k(y)$ and for a suitable phase $\theta^{(k)} = \theta^{(k)}(y)$) *for every* $y \in B_{r/2e}(y_0)$ *up to a set of relative measure smaller that* μ.
For the precise statement, we refer to Theorem 7.2 below.

[19] Namely in the region $D^{1,k} \times \mathbb{T}^n$ in the original variables.

3 Functional setting and generic holomorphic classes

3.1 Analytic function spaces

(a) Norms and neighbourhoods

In this paper $|\cdot|$ denotes the standard Euclidean norm on $\mathbb{C}^n$ and its subspaces, namely $|x|^2 = \sum_{j=1}^n |x_j|^2$ for every $x \in \mathbb{C}^n$.

$|k|_1$ denotes the 1-norm $\sum |k_j|$.

For linear maps and matrices A (which we shall always identify), $|A|$ denotes the standard "operator norm" $|A| = \sup\limits_{u \neq 0} |Au|/|u|$.

$|M|_\infty$, with M matrix (or vector), denotes the maximum norm $\max_{ij} |M_{ij}|$ (or $\max_i |M_i|$).

We shall use three different (non–equivalent) norms on holomorphic functions with domain $\mathbb{T}^n_s$, $D_r \times \mathbb{T}^n_s$ or D_r (D being a susbset of $\mathbb{R}^n$): given a holomorphic function f with values in $\mathbb{C}^m$ and domain $\mathbb{T}^n_s$, $D_r \times \mathbb{T}^n_s$ or D_r, with $s, r > 0$ we denote by

$$\sum_{k \in \mathbb{Z}^n} f_k e^{\mathrm{i} k \cdot x} \qquad \text{or} \qquad f(y,x) = \sum_{k \in \mathbb{Z}^n} f_k(y) e^{\mathrm{i} k \cdot x}$$

its Fourier expansion and define the following *sup–norm, ℓ^∞–Fourier norm and ℓ^1–Fourier norm*:

$$|f|_s := \sup_{\mathbb{T}^n_s} |f| \ , \quad |f|_r := \sup_{D_r} |f| \ , \quad |f|_{r,s} := \sup_{D_r \times \mathbb{T}^n_s} |f| \ , \tag{52}$$

$$\|f\|_s := \sup_{k \in \mathbb{Z}^n} |f_k| e^{|k|_1 s} \ , \quad \|f\|_{r,s} := \sup_{k \in \mathbb{Z}^n} \big(\sup_{y \in D_r} |f_k(y)| e^{|k|_1 s} \big) \ , \tag{53}$$

$$\lvert\!\lvert\!\lvert f \rvert\!\rvert\!\rvert_s := \sum_{k \in \mathbb{Z}^n} |f_k| e^{|k|_1 s} \ , \quad \lvert\!\lvert\!\lvert f \rvert\!\rvert\!\rvert_{r,s} := \sup_{y \in D_r} \sum_{k \in \mathbb{Z}^n} |f_k(y)| e^{|k|_1 s} \ . \tag{54}$$

If the (real) domain need to be specified, we set, respectively,

$$|f|_{D,r,s} := |f|_{r,s} \ , \quad \|f\|_{D,r,s} := \|f\|_{r,s} \ , \quad \lvert\!\lvert\!\lvert f \rvert\!\rvert\!\rvert_{D,r,s} := \lvert\!\lvert\!\lvert f \rvert\!\rvert\!\rvert_{r,s} \ . \tag{55}$$

Remark 3.1 (i) The space of functions $f : \mathbb{T}^n_s \to \mathbb{C}^m$ endowed with the sup–norm $|\cdot|_s$ or the ℓ^1–Fourier norm $\lvert\!\lvert\!\lvert \cdot \rvert\!\rvert\!\rvert_s$ is a Banach algebra, while $\{f : \mathbb{T}^n_s \to \mathbb{C}^m \text{ s.t. } \|f\|_s < \infty\}$ is just a Banach space (not a Banach algebra). However, the norm $\|\cdot\|_s$ is particularly suited to describe $\{f : \mathbb{T}^n_s \to \mathbb{C}$ s.t. $\|f\|_s < \infty\}$ as a probability space; compare item **(h)** below.

(ii) As already mentioned the three norms in (55) are not equivalent. Indeed, for any $\sigma > 0$, one has[20]

$$\|f\|_{r,s} \leqslant |f|_{r,s} \leqslant \lfloor f \rceil_{r,s} \leqslant (\coth^n(\sigma/2) - 1)\|f\|_{r,s+\sigma} \leqslant (2n/\sigma)^n \|f\|_{r,s+\sigma}\,. \tag{56}$$

The Banach subspace of $\{f : \mathbb{T}^n_s \to \mathbb{C} : \|f\|_s < \infty\}$ of real–analytic functions with zero average ($f_0 = 0$) will be the natural ambient function space. Generic elements of such a space will be the typical potentials to which our uniformaveraging theory applies. We givet it a name:

The following two definitions are needed in order to decompose a holomorphic function on $\mathbb{T}^n$ into a sum over generators of 1-d maximal lattices of holomorphic functions on $\mathbb{T}^1$. Later the Fourier modes $k \in \mathcal{G}^n_1$ will be identified with simple resonances.

(b) Lattice Fourier projectors

Given $f(y,x) = \sum_{k\in\mathbb{Z}^n} f_k(y)e^{\mathrm{i}k\cdot x}$ and a sublattice Λ of $\mathbb{Z}^n$, we denote by $\mathtt{p}_\Lambda$ the projection on the Fourier coefficients in Λ, namely

$$\mathtt{p}_\Lambda f := \sum_{k\in\Lambda} f_k(y)e^{\mathrm{i}k\cdot x}\,. \tag{57}$$

and by $\mathtt{p}^\perp_\Lambda$ its "orthogonal" operator (projection on the Fourier modes in $\mathbb{Z}^n \setminus \Lambda$):

$$\mathtt{p}^\perp_\Lambda f := \sum_{k\notin\Lambda} f_k(y)e^{\mathrm{i}k\cdot x}\,.$$

Obviously

$$\lfloor \mathtt{p}_\Lambda f \rceil_{r,s}\,,\ \lfloor \mathtt{p}^\perp_\Lambda f \rceil_{r,s} \leqslant \lfloor f \rceil_{r,s}\,. \tag{58}$$

[20] We have $\sum_{k\in\mathbb{Z}^n\setminus 0} e^{-|k|_1\sigma} = \coth^n(\sigma/2) - 1$. Moreover $\coth^n x - 1 \leqslant (n/x)^n$. Indeed for $0 < x \leqslant 1$ the estimates follows by $\coth x < 2/\sinh x < 2/x$. In the case $x > 1$ we have

$$\coth^n x - 1 \leqslant (1 + e^{1-2x})^n - 1 \leqslant n(1 + 1/e)^{n-1}e^{1-2x} \leqslant (n/x)^n\,,$$

where in the second inequality we have used that $(1 + y)^n \leqslant 1 + n(1 + 1/e)^{n-1}y$ for $0 \leqslant y \leqslant 1/e$, while in the last one we exploit $\max_{x\geqslant 1} x^n e^{-2x} = (n/2e)^n$.

(c) Fourier Truncation operators

Given $N > 0$, we introduce the following "truncation" and "high–mode" operators T_N and $T_N^\perp$:

$$T_N f(y,x) := \sum_{|k|_1 \leqslant N} f_k(y) e^{\mathrm{i}k\cdot x}\,, \qquad T_N^\perp f(y,x) := \sum_{|k|_1 > N} f_k(y) e^{\mathrm{i}k\cdot x}\,. \tag{59}$$

Note that $\mathtt{p}_\Lambda$ and T_N commute. Note, also, that

$$\| T_N f \|_{r,s}\,,\ \| T_N^\perp f \|_{r,s} \leqslant \| f \|_{r,s}\,, \tag{60}$$

and that

$$\| T_N^\perp f \|_{r,s-\sigma} \leqslant e^{-(N+1)\sigma} \| f \|_{r,s}\,, \qquad 0 < \sigma < s\,. \tag{61}$$

(d) 1d–Fourier projectors

Given $f \in \mathbb{B}^n_s$ and $k \in \mathcal{G}^n_1$ we define the following **1d–Fourier projector**

$$f \in \mathbb{B}^n_s \longmapsto \pi_{k\mathbb{Z}} f =: F^k \in \mathbb{B}^1_{|k|_1 s} \quad \text{where} \quad F^k(\theta) := \sum_{j \in \mathbb{Z}\setminus\{0\}} f_{jk} e^{\mathrm{i}j\theta}\,, \tag{62}$$

f_{jk} being the Fourier coefficient of f with Fourier index $jk \in \mathbb{Z}^n$.

It is immediate to see that:
Any $f \in \mathbb{B}^n_s$ can be uniquely written as:

$$\boxed{f(x) = \sum_{k \in \mathcal{G}^n_1} F^k(k\cdot x)} \tag{63}$$

Notice also that, if $k \in \mathcal{G}^n_1$ and $\| f \|_{r,s} < \infty$, then

$$\| F^k \|_{r,|k|_1 s} \leqslant \| f \|_{r,s}\,. \tag{64}$$

We are now ready to define the main function spaces.

Finally, we introduce a probability measure on the unit ball in $\mathbb{B}^n_s$.

(e) Denote by $\ell^\infty(\mathbb{Z}^n_\star)$ the Banach space of complex sequences (over $\mathbb{Z}^n_\star$) given by

$$\ell^\infty(\mathbb{Z}^n_\star) := \left\{ z \in \mathbb{C}^{\mathbb{Z}^n_\star} \text{ s.t. } z_k \neq 0 \text{ and } |z|_\infty := \sup_{k \in \mathbb{Z}^n_\star} |z_k| < +\infty \right\}\,. \tag{65}$$

Then, the map

$$j : f \in \mathbb{B}^n_s \to \big\{ f_k e^{|k|_1 s} \big\}_{k \in \mathbb{Z}^n_\star} \in \ell^\infty(\mathbb{Z}^n_\star) \tag{66}$$

is an isomorphism of Banach spaces[21], which allows to identify functions in $\mathbb{B}^n_s$ with points in $\ell^\infty(\mathbb{Z}^n_\star)$ and the Borellians of $\mathbb{B}^n_s$ with those of $\ell^\infty(\mathbb{Z}^n_\star)$.
Denote by $\mathbf{B}_1$ the closed ball of radius one in $\mathbb{B}^n_s$ and by $\mathcal{B}$ the Borellians in $\mathbf{B}_1$. On $\mathbf{B}_1$ we can introduce the following natural (product) *probability measure*. Consider, first, the probability measure given by the normalized Lebesgue–product measure on the unit closed ball of $\ell^\infty(\mathbb{Z}^n_\star)$, namely, the unique probability measure μ on the Borellians of $\{z \in \ell^\infty(\mathbb{Z}^n_\star) : |z|_\infty \leqslant 1\}$ such that, given Lebesgue measurable sets A_k in the unit complex disk $\mathrm{A}_k \subseteq D := \{w \in \mathbb{C} : \ |w| \leqslant 1\}$ with $\mathrm{A}_k \neq D$ only for finitely many k, one has

$$\mu\Big(\prod_{k \in \mathbb{Z}^n_\star} \mathrm{A}_k \Big) = \prod_{\{k \in \mathbb{Z}^n_\star :\, \mathrm{A}_k \neq D\}} \frac{1}{\pi} \mathrm{meas}(\mathrm{A}_k)$$

where "meas" denotes the Lebesgue measure on the unit complex disk D.
Then, *the isometry j in (66) naturally induces a* **probability measure**[22] μ_s **on the Borellians $\mathcal{B}$.**

3.2 Generic properties of periodic holomorphic classes

Here we discuss some properties of the classes $\mathcal{H}_{s,\tau}$ of non-degenerate introduced in Definition 2.1.

Remark 3.2 (i) Since $f \in \mathbb{B}^n_s$, one has that $|f_k| \leqslant \|f\|_s e^{-|k|_1 s}$ for all k's and (32) says that, when k is a generator of maximal 1d–lattices (later corresponding to simple resonances), the k–Fourier coefficient does not vanish and is controlled in a quantitive way from below: $|k|_1^{-n}$ is a suitable weight (needed in the proof of Proposition 3.1 below), while δ is any number satisfiying

$$\inf_{|k|_1 > \tau(\delta)} |f_k| |k|_1^n \, e^{|k|_1 s} \geqslant \delta > 0 \; . \tag{67}$$

(ii) It is easy to construct functions in $\mathcal{H}_{s,\tau}(\delta)$. For example let

$$f(x) := 2\delta \sum_{k \in \mathcal{G}^n_1} |k|_1^{-n} e^{-|k|_1 s} \cos(k \cdot x) \; , \tag{68}$$

[21]Recall that since the functions in $\mathbb{B}^n_s$ are *real*–analytic one has the reality condition $f_k = \bar{f}_{-k}$.
[22]I.e, $\mu_s(\mathbf{B}_1) = 1$.

which has Fourier coefficients

$$f_k = \begin{cases} \delta |k|_1^{-n} e^{-|k|_1 s}, & \text{if } \pm k \in \mathcal{G}_1^n \\ 0, & \text{otherwise} \end{cases}$$

and 1d–Fourier projections

$$F^k(\theta) = \delta |k|_1^{-n} e^{-|k|_1 s} \cos\theta \ .$$

Then, $f \in \mathcal{H}_{s,0}(\delta)$ and, also, $f \in \mathcal{H}_{s,\tau}(\delta)$ for any choice of tail function $\tau(\delta)$.

Here, we show that the classes $\mathcal{H}_{s,\tau}$ are "general" in several (topological and measure theoretical) ways.

Proposition 3.1 (Properties of $\mathcal{H}_{s,\tau}$) *Let $s > 0$ and τ be a tail function. Then:*

(i) *The set $\mathcal{H}_{s,\tau} \subseteq \mathbb{B}_s^n$ contains an open dense set.*

(ii) *$\mathcal{H}_{s,\tau} \cap \mathbf{B}_1 \in \mathcal{B}$ and $\mu_s(\mathcal{H}_{s,\tau} \cap \mathbf{B}_1) = 1$.*

(iii) *$\mathcal{H}_{s,\tau}$ is a prevalent set*[23]*.*

Proof

(i) *$\mathcal{H}_{s,\tau}$ contains an open subset $\mathcal{H}'_{s,\tau}$ which is dense in the unit ball of $\mathbb{B}_s^n$.*
Let us define $\mathcal{H}'_{s,\tau}$ as $\mathcal{H}_{s,\tau}$ but with the difference that (32) is replaced by the *stronger condition*[24]

$$\exists\, \delta > 0 \text{ s.t. } |f_k| \geqslant \delta\ e^{-|k|_1 s}\,, \quad \forall\ k \in \mathcal{G}_1^n,\ |k|_1 > \tau(\delta) \tag{69}$$

Let us first prove that $\mathcal{H}'_{s,\tau}$ is open. Let $f \in \mathcal{H}'_{s,\tau}$. We have to show that there exists $\rho > 0$ such that if $\|g\|_s < \rho$, then $f + g \in \mathcal{H}'_{s,\tau}$. Fix $\delta > 0$ such that (69) holds and, by continuity of $\tau(\delta)$, choose $\rho < \delta$ small enough such that $[\tau(\delta)] > \tau(\delta') - 1$, where $\delta' := \delta - \rho$ and $[\cdot]$ denotes integer part. Then, since $\tau(\delta)$ is not increasing, it is immediate to verify that $|k|_1 > \tau(\delta) \iff |k|_1 > \tau(\delta')$. Moreover

$$|f_k + g_k| e^{|k|_1 s} \geqslant |f_k| e^{|k|_1 s} - \|g\|_s \geqslant \delta - \rho = \delta'\,, \qquad \forall\ k \in \mathcal{G}_1^n,\ |k|_1 > \tau(\delta')\,,$$

[23] We recall that a Borel set P of a Banach space X is called *prevalent* if there exists a compactly supported probability measure ν on the Borellians of X such that $\nu(x + P) = 1$ for all $x \in X$; compare, e.g., [8]

[24] Note, however, that $\mu_s(\mathcal{H}'_s) = 0$.

namely $f+g$ satisfies (69) (with δ' instead of δ).

Let us now show that $\mathcal{H}'_{s,\tau}$ is dense in the unit ball of $\mathbb{B}^n_s$. Take f in the unit ball of $\mathbb{B}^n_s$ and $0<\lambda<1$. We have to find $\tilde f \in \mathcal{H}'_{s,\tau}$ with $\|\tilde f - f\|_s \leqslant \lambda$. Let $\delta := \lambda/4$ and denote by f_k and $\tilde f_k$ (to be defined) be the Fourier coefficients of, respectively, f and $\tilde f$. We, then, let $\tilde f_k = f_k$ unless $k \in \mathcal{G}^n_1$, $|k|_{{}_1} > \tau(\delta)$ and $|f_k| e^{|k|_1 s} < \delta$, in which case, $\tilde f_k = \delta e^{-|k|_1 s}$. It is, now, easy to check that $\tilde f \in \mathcal{H}'_s$ and is λ–close to f.

(ii) $\mathcal{H}_{s,\tau} \cap \mathbf{B}_1 \in \mathcal{B}$ *and* $\mu_s(\mathcal{H}_{s,\tau} \cap \mathbf{B}_1) = 1$

We shall prove that, for every $\delta > 0$, the measure of the sets of potentials f that do not satisfy (32) is $O(\delta^2)$, the result will follow letting $\delta \to 0$.

By the identification (66), the measure of the set of potentials f that do not satisfy (32) with a given δ is bounded by

$$\delta^2 \sum_{k\in\mathbb{Z}^n} |k|_{{}_1}^{-2n} . \tag{70}$$

Remark 3.3 *Recalling footnote 16, one could impose the condition* $|f_k| \geqslant \delta |k|_{{}_1}^{-\bar n}\, e^{-|k|_1 s}$ *in* (32)*. Then* (70) *would become* $\delta^2 \sum_{k\in\mathbb{Z}^n} |k|_{{}_1}^{-2\bar n}$*, which is still fine if* $\bar n > n/2$*.*

(iii) $\mathcal{H}_{s,\tau}$ *is prevalent.*

Consider the following compact subset of $\ell^\infty(\mathbb{Z}^n_*)$: let $\mathcal{K} := \{z = \{z_k\}_{k\in\mathbb{Z}^n_*} : z_k \in D_{1/|k|_{{}_1}}\}$, where $D_{1/|k|_{{}_1}} := \{w \in \mathbb{C} :\ |w| \leqslant 1/|k|_{{}_1}\}$, and let ν be the unique probability measure supported on $\mathcal{K}$ such that, given Lebesgue measurable sets $\mathrm{A}_k \subseteq D_{1/|k|_{{}_1}}$, with $\mathrm{A}_k \neq D_{1/|k|_{{}_1}}$ only for finitely many k, one has

$$\nu\Big(\prod_{k\in\mathbb{Z}^n_*} \mathrm{A}_k\Big) := \prod_{\{k\in\mathbb{Z}^n_*:\, \mathrm{A}_k \neq D_{1/|k|_{{}_1}}\}} \frac{|k|_{{}_1}^2}{\pi} \mathrm{meas}(\mathrm{A}_k) .$$

The isometry j_s in (66) naturally induces a probability measure ν_s on $\mathbb{B}^n_s$ with support in the compact set $\mathcal{K}_s := j_s^{-1}\mathcal{K}$. Reasoning as in the proof of $\mu_s(\mathcal{H}_{s,\tau}) = 1$, one can show that $\nu_s(\mathcal{H}_{s,\delta}) \geqslant 1 - \mathrm{const}\,\delta^2$. It is also easy to check that, for every $g \in \mathbb{B}^n_s$, the translated set $\mathcal{H}_{s,\delta} + g$ satisfies $\nu_s(\mathcal{H}_{s,\delta} + g) \geqslant \nu_s(\mathcal{H}_{s,\delta})$. Thus, one gets $\nu_s(\mathcal{H}_{s,\tau} + g) = \nu_s(\mathcal{H}_{s,\tau}) = 1$, $\forall\ g \in \mathbb{B}^n_s$, which means that $\mathcal{H}_{s,\tau}$ is prevalent (recall footnote 23). ∎

4 A normal form lemma with "small" analyticity loss

In this section we describe an analytic normal form lemma for nearly–integrable Hamiltonians $H(y,x) = h(y) + f(y,x)$, which allows to average out non–resonant Fourier modes of the perturbation f on suitable non–resonant regions, and allows for "very small" analyticity loss in the angle variables, a fact, which will be crucial in our applications.

We recall ([13], [15]) that, given an integrable Hamiltonian $h(y)$, positive numbers α, K and a lattice $\Lambda \subset \mathbb{Z}^n$, a (real or complex) domain U is (α, K) non–resonant modulo Λ (with respect to h) if

$$|h'(y) \cdot k| \geqslant \alpha \ , \quad \forall\ y \in U\ , \forall\ k \in \mathbb{Z}^n \setminus \Lambda\ ,\ |k|_{{}_1} \leqslant K\ . \tag{71}$$

The main point of the following "Normal Form Lemma" is that the "new" averaged Hamiltonian is defined, in the fast variable (angle) domain, in a region "almost equal" to the original domain, "almost equal" meaning a complex strip of width $s(1-1/K)$ if s is the width of the initial angle analyticity. More precisely, we have:

Proposition 4.1 (Normal form with "small" analyticity loss)
Let $r, s, \alpha > 0$, $K \in \mathbb{N}$, $K \geqslant 2$, $D \subseteq \mathbb{R}^n$, and let Λ be a lattice of $\mathbb{Z}^n$. Let

$$H(y,x) = h(y) + f(y,x) \tag{72}$$

be real–analytic on $D_r \times \mathbb{T}^n_s$ with $\lVert f \rVert_{r,s} < \infty$. Assume that D_r is $(\alpha$,$K)$ non–resonant modulo Λ and that

$$\vartheta_\star := \frac{2^{11} K^2}{\alpha r s} \lVert f \rVert_{r,s} < 1\,. \tag{73}$$

Then, there exists a real–analytic symplectic change of variables

$$\Psi : (y', x') \in D_{r_\star} \times \mathbb{T}^n_{s_\star} \ \mapsto\ (y,x) \in D_r \times \mathbb{T}^n_s \quad \text{with} \quad r_\star := r/2\,, \quad s_\star := s(1-1/K) \tag{74}$$

satisfying

$$|y - y'|_{{}_1} \leqslant \frac{\vartheta_\star}{2^7 K} r\,, \qquad \max_{1 \leqslant i \leqslant n} |x_i - x_i'| \leqslant \frac{\vartheta_\star}{16 K^2} s\,, \tag{75}$$

and such that

$$H \circ \Psi = h + f^\flat + f_\star\,, \qquad f^\flat := \ \mathtt{p}_\Lambda f + T_K^\perp \mathtt{p}_\Lambda^\perp f \tag{76}$$

with

$$\lVert f_\star \rVert_{r_\star, s_\star} \leqslant \frac{1}{K} \vartheta_\star \lVert f \rVert_{r,s}\,, \qquad \lVert T_K\, \mathtt{p}_\Lambda^\perp f_\star \rVert_{r_\star, s_\star} \leqslant (\vartheta_\star/8)^K \frac{8}{eK} \lVert f \rVert_{r,s}\,. \tag{77}$$

Moreover, re-writing (76) *as*

$$H \circ \Psi = h + g + f_{\star\star} \qquad \textit{where} \quad \mathrm{p}_\Lambda g = g \ , \quad \mathrm{p}_\Lambda f_{\star\star} = 0 \, , \tag{78}$$

one has

$$|g - \mathrm{p}_\Lambda f|_{r_\star, s_\star} \leqslant \frac{1}{K}\vartheta_\star |f|_{r,s} \ , \qquad |f_{\star\star}|_{r_\star, s/2} \leqslant 2e^{-(K-2)\bar{s}} |f|_{r,s} \, , \tag{79}$$

where

$$\bar{s} := \min\left\{ \frac{s}{2}, \log\frac{8}{\vartheta_\star} \right\} \ . \tag{80}$$

Remark 4.1

(i) The "novelty" of this lemma is that the bounds in (77) and the first one in (79) hold on the large angle domain $\mathbb{T}^n_{s_\star}$ with $s_\star = s(1 - 1/K)$. In particular the first estimate in (77) (or, equivalently, in (79)) will be important in our analysis in order to obtain (159), (162) and, therefore, (170), (177) and finally (181), which is the key to prove (43) in Theorem 2.1. The drawback of the gain in angle–analyticity strip is that the power of K in the smallness condition (73) is not optimal: for example in [15] the power of K is one (but $s_\star = s/6$, which would not work in our applications).

(ii) Having information on non–resonant Fourier modes up to order K, the best one can do is to average out the non–resonant Fourier modes up to order K, namely, to "kill" the term $T_K\, \mathrm{p}_\Lambda^\perp f$ of the Fourier expansion of the perturbation. This explains the "flat" term $f^\flat = \mathrm{p}_\Lambda f + T_K^\perp\, \mathrm{p}_\Lambda^\perp f$ surviving in (76) and which cannot be removed in general. Now, think of the remainder term $f_\star$ as

$$f_\star = \mathrm{p}_\Lambda f_\star + \left(T_K^\perp\, \mathrm{p}_\Lambda^\perp f_\star + T_K\, \mathrm{p}_\Lambda^\perp f_\star \right) \ ;$$

then, $\mathrm{p}_\Lambda f_\star$ is a $(\vartheta_\star |f|_{r,s}/K)$–perturbation of the part in normal form (i.e., with Fourier modes in Λ), while $T_K^\perp\, \mathrm{p}_\Lambda^\perp f_\star$ is, by (61), a term exponentially small with K (see also below) and $T_K\, \mathrm{p}_\Lambda^\perp f_\star$ is a very small remainder bounded by $8(\vartheta_\star/8)^K |f|_{r,s}/eK$.

(iii) We note that (78) follows from (76). Indeed we take

$$g = \mathrm{p}_\Lambda f + \mathrm{p}_\Lambda f_\star \, , \qquad f_{\star\star} = T_K^\perp\, \mathrm{p}_\Lambda^\perp f + \mathrm{p}_\Lambda^\perp f_\star = T_K\, \mathrm{p}_\Lambda^\perp f_\star + T_K^\perp\, \mathrm{p}_\Lambda^\perp (f_\star + f) \, .$$

Then the first estimate in (79) follows by the first bound in (77) and (58). Regarding the second estimate in (79), we first note by (77) and (61) (used with

$f \rightsquigarrow f_\star$ $N \rightsquigarrow K$, $r \rightsquigarrow r_\star$, $s \rightsquigarrow s_\star$, and $\sigma \rightsquigarrow \frac{s}{2} - \frac{s}{K}$ so that $s_\star - \sigma = s/2$ and $e^{-(K+1)\sigma} \leqslant e^{-(K-2)s/2}$)

$$\| T_K^\perp f_\star \|_{r_\star, s/2} = \| T_K^\perp f_\star \|_{r_\star, s_\star - \sigma} \leqslant e^{-(K+1)\sigma} \| f_\star \|_{r_\star, s_\star} \leqslant e^{-(K-2)s/2} \vartheta_\star \| f \|_{r,s} / K \,.$$

By (58), (77) and (61) we get

$$\begin{aligned} \| f_{\star\star} \|_{r_\star, s/2} &\leqslant \| T_K \, \mathtt{p}_\Lambda^\perp f_\star \|_{r_\star, s/2} + \| T_K^\perp f_\star \|_{r_\star, s/2} + \| T_K^\perp f \|_{r_\star, s/2} \\ &\leqslant (\vartheta_\star / 8)^K \frac{8}{eK} \| f \|_{r,s} + e^{-(K-2)s/2} (\vartheta_\star / K + e^{-3s/2}) \| f \|_{r,s} \\ &\leqslant 2 e^{-(K-2)\bar{s}} \| f \|_{r,s} \,. \end{aligned}$$

(iv) Let us compare our results with more standard formulations, such as the Normal Form Lemma in § 2 of [15]. In that formulation, imposing the weaker smallness condition $\| f \|_{r,s} \leqslant \mathrm{const}\, \alpha r / K$, the normal form Hamiltonian writes $h + \mathtt{g} + \mathtt{f}$ with $\mathtt{f}$ exponentially small (of order $\| f \|_{r,s} e^{-Ks/6}$) and, regarding $\mathtt{g}$ one knows that

$$\| \mathtt{g} - T_K \, \mathtt{p}_\Lambda f \|_{r/2, s/6} \leqslant \mathrm{const.} \frac{K}{\alpha r} \| f \|_{r,s}^2 \,. \tag{81}$$

For our purposes we need to prove that, when $k \in \mathbb{Z}^n_\sharp$, $|k|_1 \leqslant K_1 \leqslant K$ ($k \in \mathbb{Z}^n_{K_1}$ indexes the simple resonance we want to consider while $l \in \mathbb{Z}^n_K$ indexes the second order resonance beyond k) and $|f_k| / \| f \|_{r,s} \geqslant \delta |k|_1^{-n} e^{-|k|_1 s}$, the quantity

$$\frac{1}{|f_k|} \sup_{y \in D_{r/2}} |g_k(y) - f_k|$$

is small. Indeed by (79) we have

$$\begin{aligned} \frac{1}{|f_k|} \sup_{y \in D_{r/2}} |g_k(y) - f_k| &\leqslant \frac{\vartheta_\star}{K} \| f \|_{r,s} \frac{e^{-|k|_1 s_\star}}{|f_k|} \leqslant \frac{\vartheta_\star}{K} \frac{|k|_1^n e^{(s - s_\star)|k|_1}}{\delta} \\ &= \frac{\vartheta_\star}{K} \frac{|k|_1^n e^{s|k|_1/K}}{\delta} \leqslant \frac{\vartheta_\star}{K} e^s \frac{K_1^n}{\delta} \,, \end{aligned} \tag{82}$$

which is small when

$$K_1 \ll \left(\frac{\alpha r s \delta}{K \| f \|_{r,s}} \right)^{1/n} \,. \tag{83}$$

Consider, for example, the function $f = \varepsilon \hat{f}$ with ε small and $\hat{f}$ defined in (68). We have that $\| f \|_{r,s} = c\delta\varepsilon$, for a suitable constant $c > 0$. In this case (83) writes

$$K_1 \ll \left(\frac{\alpha r s}{K \varepsilon} \right)^{1/n} \,. \tag{84}$$

On the other hand by estimate (81) one only have

$$\frac{1}{|f_k|}\sup_{y\in D_{r/2}}|\mathrm{g}_k(y)-f_k|\leqslant \text{const}\,.\frac{K\varepsilon}{\alpha r}|k|_1^n e^{|k|_1 s}e^{-|k|_1 s/6}\leqslant \text{const}\,.\frac{\varepsilon K}{\alpha r}K_0^n e^{\frac{5}{6}K_0 s}\,,$$

which is small only for

$$K_1 \ll \frac{6}{5s}\log\frac{\alpha r}{K\varepsilon}\,, \tag{85}$$

that is a considerably stronger bound than the one in (84).
Since we are considering simple resonances indexed by $|k|_1 \leqslant K_1$, the non resonant region will be non-resonant only up to order K_1; therefore we have that the perturbation, after normal form in the non-resonant region, will be of magnitude

$$\varepsilon e^{-K_1 s/6} \gg \varepsilon(K\varepsilon/\alpha r)^{1/5}\,,$$

when the bound (85) applies. This estimate is very bad. On the other hand, in our case, the weaker bound (84) applies and we obtain that the perturbation is exponentially small.

Given a function ϕ we denote by X_ϕ^t the hamiltonian flow at time t generated by ϕ and by "ad" the linear operator $u\mapsto \mathrm{ad}_\phi u := \{u,\phi\}$ and ad^ℓ its iterates:

$$\mathrm{ad}_\phi^0 u := u\,,\qquad \mathrm{ad}_\phi^\ell u := \{\mathrm{ad}_\phi^{\ell-1}u,\phi\}\,,\qquad \ell\geqslant 1\,,$$

as standard, $\{\cdot,\cdot\}$ denotes Poisson bracket[25].
Recall the identity ("Lie series expansion")

$$u\circ X_\phi^1 = \sum_{\ell\geqslant 0}\frac{1}{\ell!}\mathrm{ad}_\phi^\ell u = \sum_{\ell=0}^{\infty}\frac{\partial_t^\ell(u\circ X_\phi^t)}{\ell!}\Big|_{t=0}\,, \tag{86}$$

valid for analytic functions and small ϕ. We recall the following technical lemma by [15].

Lemma 4.1 (Lemma B.3 of [15]) *For $0<\rho<r$, $0<\sigma<s$, $D\subseteq\mathbb{R}^n$*

$$\sup_{y\in D_r}\sum_{1\leqslant i\leqslant n}\|\partial_{x_i}\phi(y,\cdot)\|_{s-\sigma}\leqslant\frac{1}{e\sigma}\|\phi\|_{r,s}\,,\qquad \sup_{y\in D_{r-\rho}}\max_{1\leqslant i\leqslant n}\|\partial_{y_i}\phi(y,\cdot)\|_s\leqslant\frac{1}{\rho}\|\phi\|_{r,s}\,,$$

[25] Explicitly, $\{u,v\}=\sum_{i=1}^n(u_{x_i}v_{y_i}-u_{y_i}v_{x_i})$.

By Lemma 4.1 we get (see also Lemma B.4 of [15])

Lemma 4.2 *For* $0<\rho<\bar r:=\{r_0,r\}$, $0<\sigma<\bar s:=\{s_0,s\}$,

$$|\{f,g\}|_{\bar r-\rho,\bar s-\sigma}\leqslant\frac{1}{e}\left(\frac{1}{(r_0-\bar r+\rho)(s-\bar s+\sigma)}+\frac{1}{(r-\bar r+\rho)(s_0-\bar s+\sigma)}\right)|f|_{r_0,s_0}|g|_{r,s}\,. \tag{87}$$

Summing the Lie series in (86) (see Lemma B5 of [15]) we get, also,

Lemma 4.3 *Let* $0<\rho<r_0$ *and* $0<\sigma<s_0$. *Assume that*

$$\hat\vartheta:=\frac{4e|\phi|_{r_0,s_0}}{\rho\sigma}\leqslant 1\,. \tag{88}$$

Then for every $\rho<r'\leqslant r_0$, $\sigma<s'\leqslant s_0$, *the time-1-flow* X^1_ϕ *of vector field* X_ϕ *define a good canonical transformation*

$$X^1_\phi : D_{r'-\rho}\times\mathbb{T}^n_{s'-\sigma}\ \to\ D_{r'-\rho/2}\times\mathbb{T}^n_{s'-\sigma/2} \tag{89}$$

satisfying

$$|y-y'|_1\leqslant\hat\vartheta\frac{\rho}{4e}\,,\qquad \max_{1\leqslant i\leqslant n}|x_i-x_i'|\leqslant\hat\vartheta\frac{\sigma}{4} \tag{90}$$

Moreover let $r>\rho, s>\sigma$ *and set*

$$\bar r:=\min\{r_0,r\}\,,\qquad \bar s:=\min\{s_0,s\}\,.$$

Then for any $j\geqslant 0$

$$\begin{aligned}|u\circ X^1_\phi-\sum_{h\leqslant j}\mathrm{ad}^h_\phi u|_{\bar r-\rho,\bar s-\sigma}&\leqslant\sum_{h>j}\frac{1}{h!}|\mathrm{ad}^h_\phi u|_{\bar r-\rho,\bar s-\sigma}\\&\leqslant 2(\hat\vartheta/2)^j|\{u,\phi\}|_{\bar r-\rho/2,\bar s-\sigma/2}\end{aligned} \tag{91}$$

for every function u *with* $|u|_{r,s}<\infty$.
In particular when $r\leqslant r_0, s\leqslant s_0$

$$|u\circ X^1_\phi-u|_{r-\rho,s-\sigma}\leqslant\sum_{h\geqslant1}\frac{1}{h!}|\mathrm{ad}^h_\phi u|_{r-\rho,s-\sigma}\leqslant 2\hat\vartheta|u|_{r,s}\,, \tag{92}$$

$$|u\circ X^1_\phi-u-\{u,\phi\}|_{r-\rho,s-\sigma}\leqslant\hat\vartheta^2|u|_{r,s}\,, \tag{93}$$

Proof We first note that by Lemma 4.1 (applied with $r_0 \rightsquigarrow r$, $s_0 \rightsquigarrow s$) for every $(y,x) \in D_{r_0-\rho} \times \mathbb{T}^n_{s_0-\sigma}$ we have

$$|\partial_x \phi(y,x)|_1 \leqslant \frac{1}{e\sigma} \| \phi \|_{r_0,s_0} = \frac{\hat{\vartheta}\rho}{4e} \leqslant \frac{\rho}{4e}\,, \qquad \max_{1\leqslant i\leqslant n} |\partial_{y_i}\phi(y,x)| \leqslant \frac{1}{\rho}\|\phi\|_{r_0,s_0} = \frac{\hat{\vartheta}\sigma}{4} \leqslant \frac{\sigma}{4}\,.$$

Then (89) holds.
For $h \geqslant 1$, set for brevity

$$\|\cdot\|_i := \|\cdot\|_{\bar{r}-\frac{\rho}{2}-i\tilde{\rho},\bar{s}-\frac{\sigma}{2}-i\tilde{\sigma}}\,, \qquad 0 \leqslant i \leqslant h\,, \qquad \tilde{\rho} := \frac{\rho}{2h}\,, \quad \tilde{\sigma} := \frac{\sigma}{2h}\,.$$

We get

$$\begin{aligned} &\|\mathrm{ad}^i_\phi\{u,\phi\}\|_i \\ &\overset{(87)}{\leqslant} \frac{1}{e}\left(\frac{1}{\tilde{\rho}(s_0-\bar{s}+i\tilde{\sigma}+\sigma/2)} + \frac{1}{\tilde{\sigma}(r_0-\bar{r}+i\tilde{\rho}+\rho/2)}\right)\|\phi\|_{r_0,s_0}\|\mathrm{ad}^{i-1}_\phi\{u,\phi\}\|_{i-1} \\ &\leqslant \frac{8h^2}{e\rho\sigma}\frac{1}{h+i}\|\phi\|_{r_0,s_0}\|\mathrm{ad}^{i-1}_\phi\{u,\phi\}\|_{i-1}\,, \end{aligned}$$

and, iterating,

$$\|\mathrm{ad}^h_\phi\{u,\phi\}\|_h \leqslant \frac{8h^2}{e\rho\sigma}\frac{h!}{(2h)!}\|\phi\|_{r_0,s_0}\|\{u,\phi\}\|_{r-\rho/2,s-\sigma/2} \leqslant h!(\hat{\vartheta}/2)^h\|\{u,\phi\}\|_{r-\rho/2,s-\sigma/2}$$

by Stirling's formula. Then

$$\sum_{h\geqslant j}\frac{1}{(h+1)!}\|\mathrm{ad}^{h+1}_\phi u\|_{\bar{r}-\rho,\bar{s}-\sigma} \leqslant \sum_{h\geqslant j}\frac{1}{h+1}(\hat{\vartheta}/2)^h\|\{u,\phi\}\|_{r-\rho/2,s-\sigma/2}$$

proving (91) in view of (88).
Finally (92) and (93) follows by (91) and since $\|\{u,\phi\}\|_{\bar{r}-\rho/2,\bar{s}-\sigma/2} \leqslant 2e^{-1}\hat{\vartheta}\|u\|_{r,s}$ by (87).
∎

Given $K \geqslant 2$ and a lattice Λ, recall the definition of $f^\flat$ in (76) and define

$$f^K := f - f^\flat = T_K\, \mathrm{p}^\perp_\Lambda f\ ,$$

so that we have the decomposition (valid for any f):

$$f = f^\flat + f^K\,, \qquad f^\flat := P_\Lambda f + T^\perp_K\, \mathrm{p}^\perp_\Lambda f\,, \qquad f^K := T_K\, \mathrm{p}^\perp_\Lambda f\,. \tag{94}$$

Lemma 4.4 *Let $0 < \rho < r$ and $0 < \sigma < s$. Consider a real–analytic Hamiltonian*

$$H = H(y,x) = h(y) + f(y,x) \qquad \textit{analytic on } \ D_r \times \mathbb{T}^n_s\,. \tag{95}$$

Suppose that D_r is (α,K) non–resonant modulo Λ for h (with $K \geqslant 2$). Assume that

$$\check{\vartheta} := \frac{4e}{\alpha\rho\sigma} |f^K|_{r,s} \leqslant 1\,. \tag{96}$$

Then there exists a real–analytic symplectic change of coordinates

$$\Psi := X^1_\phi \,:\, D_{r_+} \times \mathbb{T}^n_{s_+} \ni (y',x') \ \to \ (y,x) \in D_r \times \mathbb{T}^n_s\,, \quad r_+ := r-\rho\,, \quad s_+ := s-\sigma\,,$$

generated by a function $\phi = \phi^K = T_K\, \mathtt{p}^\perp_\Lambda \phi$ with

$$|\phi|_{r,s} \leqslant |f^K|_{r,s}/\alpha\,, \tag{97}$$

satisfying

$$|y-y'|_{_1} \leqslant \check{\vartheta}\frac{\rho}{4e}\,, \qquad \max_{1\leqslant i\leqslant n} |x_i - x'_i| \leqslant \check{\vartheta}\frac{\sigma}{4}\,, \tag{98}$$

such that

$$H\circ\Psi = h(y') + f_+(y',x')\,, \qquad f_+ := f^\flat + f_\star \tag{99}$$

with

$$|f_\star|_{r_+,s_+} \leqslant \ 4\check{\vartheta}|f|_{r,s}\,. \tag{100}$$

Notice that, by (94) and (100), one has

$$f^K_+ = f^K_\star\ , \quad |f_+|_{r_+,s_+} \leqslant |f_\star|_{r_+,s_+} + |f|_{r,s} \leqslant (1+4\check{\vartheta})|f|_{r,s}\ . \tag{101}$$

Notice also that

$$f^\flat_+ - f^\flat \overset{(99)}{=} f^\flat_\star \quad \Longrightarrow \quad |f^\flat_+ - f^\flat|_{r_+,s_+} \leqslant |f_\star|_{r_+.s_+} \overset{(100)}{\leqslant} 4\check{\vartheta}|f|_{r,s}\ . \tag{102}$$

Proof Let us define

$$\phi = \phi(y,x) := \sum_{|m|\leqslant K, m\notin\Lambda} \frac{f_m(y)}{\mathrm{i}h'(y)\cdot m} e^{\mathrm{i}m\cdot x}\,,$$

and note that ϕ solves the homological equation

$$\{h,\phi\} + f^K = 0\,. \tag{103}$$

Since D_r is (α,K) non–resonant modulo Λ the estimate (97) holds. We now use Lemma 4.3 with parameters $r_0 \rightsquigarrow r, s_0 \rightsquigarrow s$. With these choices it is $\hat{\vartheta} = \check{\vartheta}$, and, by (96) $\check{\vartheta} \leqslant 1$. Thus, (88) holds and Lemma 4.3 applies. (98) follows by (90). We have

$$H \circ \Psi = h + f^\flat + f_\star$$

with

$$f_\star = (h \circ \Psi - h - \{h, \phi\}) + (f \circ \Psi - f)\,.$$

Since

$$h \circ \Psi - h - \{h, \phi\} = \sum_{\ell \geqslant 2} \frac{1}{\ell!} \mathrm{ad}_\phi^\ell h = \sum_{\ell \geqslant 1} \frac{1}{(\ell+1)!} \mathrm{ad}_\phi^\ell \{h, \phi\} \overset{(103)}{=} - \sum_{\ell \geqslant 1} \frac{1}{(\ell+1)!} \mathrm{ad}_\phi^\ell f^K\,,$$

we have

$$\| h \circ \Psi - h - \{h, \phi\} \|_{r_+, s_+} \leqslant \sum_{\ell \geqslant 1} \frac{1}{\ell!} \| \mathrm{ad}_\phi^\ell f^K \|_{r_+, s_+} \overset{(92)}{\leqslant} 2\check{\vartheta} \| f^K \|_{r,s} \leqslant 2\check{\vartheta} \| f \|_{r,s}\ .$$

Finally, applying again Lemma 4.3 with $u = f$, by (92), we get $\| f \circ \Psi - f \|_{r_+, s_+} \leqslant 2\check{\vartheta} \| f \|_{r,s}$, proving (100) and concluding the proof of Lemma 4.4. ∎

As a preliminary step we apply Lemma 4.4 to the Hamiltonian $H = h + f$ in (72) with $\rho = r/4$ and $\sigma = s/2K$. By (58), (60), (94) and (73) hypothesis (96) holds, namely

$$\vartheta_{-1} := \frac{2^5 e K}{\alpha r s} \| f^K \|_{r,s} \leqslant 1\,. \tag{104}$$

Then there exists a real–analytic symplectic change of coordinates

$$\Psi_{-1} : D_{r_0} \times \mathbb{T}^n_{s_0} \ni (y^{(0)}, x^{(0)}) \ \to\ (y, x) \in D_r \times \mathbb{T}^n_s\,, \quad r_0 := \frac{3}{4} r\,, \quad s_0 := \left(1 - \frac{1}{2K}\right) s\,,$$

satisfying

$$|y - y^{(0)}|_1 \leqslant \vartheta_{-1} \frac{r}{16e}\,, \qquad \max_{1 \leqslant i \leqslant n} |x_i - x_i^{(0)}| \leqslant\ \vartheta_{-1} \frac{s}{8K}\,, \tag{105}$$

such that

$$H \circ \Psi_{-1} =: H_0 = h(y^{(0)}) + f_0(y^{(0)}, x^{(0)})\,, \quad f_0 = f^\flat + f_\star\,, \quad f^\flat := P_\Lambda f + T_K^\perp \mathrm{p}_\Lambda^\perp f\,, \tag{106}$$

with

$$\| f_\star \|_{r_0, s_0} \leqslant\ 4\vartheta_{-1} \| f \|_{r,s}\,. \tag{107}$$

Recalling (94) and (106) we get

$$f_0^K = f_\star^K$$

and, by (107) and (104),

$$\| f_0^K \|_{r_0,s_0} \leqslant 4\vartheta_{-1} \| f \|_{r,s} \leqslant \frac{2^7 eK}{\alpha rs} \| f \|_{r,s}^2 \,. \tag{108}$$

Then, setting

$$\vartheta_0 := \delta \| f_0^K \|_{r_0,s_0} \qquad \text{with} \qquad \delta := \frac{2^5 e\, K^3}{\alpha rs} \,, \tag{109}$$

we have

$$\vartheta_0 \leqslant \Big(\frac{2^6 eK^2}{\alpha rs} \| f \|_{r,s} \Big)^2 \overset{(73)}{\leqslant} (\vartheta_\star/8)^2 \leqslant \frac{1}{2^6} \,. \tag{110}$$

Finally, since $f_0^\flat - f^\flat = f_\star^\flat$ by (102) we get

$$\| f_0^\flat - f^\flat \|_{r_0,s_0} \leqslant 4\vartheta_{-1} \| f \|_{r,s} \overset{(104)}{\leqslant} \frac{2^7 eK}{\alpha rs} \| f \|_{r,s}^2 \overset{(73)}{\leqslant} \frac{1}{4K} \vartheta_\star \| f \|_{r,s} \,. \tag{111}$$

The idea is to construct Ψ by applying K times Lemma 4.4.
Let

$$\rho := \frac{r}{4K} \,, \qquad \sigma := \frac{s}{2K^2} \,,$$

$$r_i := \frac{3}{4} r - i\rho \,, \qquad s_i := \Big(1 - \frac{1}{2K} \Big) s - i\sigma \,, \qquad \| \cdot \|_i := \| \cdot \|_{r_i,s_i} \,, \tag{112}$$

Fix $1 \leqslant j \leqslant K$ and make the following **inductive assumptions**:
Assume that there exist, for $1 \leqslant i \leqslant j$, real–analytic symplectic transformations

$$\Psi_{i-1} := X^1_{\phi_{i-1}} \ : \ D_{r_i} \times \mathbb{T}^n_{s_i} \ni (y^{(i)}, x^{(i)}) \ \to \ (y^{(i-1)}, x^{(i-1)}) \in D_{r_{i-1}} \times \mathbb{T}^n_{s_{i-1}} \,,$$

generated by a function $\phi_{i-1} = \phi_{i-1}^K$ with

$$\| \phi_{i-1} \|_{i-1} \leqslant \| f_{i-1}^K \|_{i-1} / \alpha \,, \tag{113}$$

satisfying

$$|y^{(i-1)} - y^{(i)}|_1 \leqslant \vartheta_{i-1} \frac{r}{16eK} \,, \qquad \max_{1 \leqslant \ell \leqslant n} |x_\ell^{(i-1)} - x_\ell^{(i)}| \leqslant \vartheta_{i-1} \frac{s}{8K^2} \,, \tag{114}$$

such that

$$H_i := H_{i-1} \circ \Psi_{i-1} =: h + f_i = h + f_i^K + f_i^\flat \tag{115}$$

satisfies, for $1 \leqslant i \leqslant j$, *the estimates*

$$\vartheta_i \leqslant \left(\frac{2^8 K^2 |f|_{r,s}}{\alpha r s}\right)^{i+1} \overset{(73)}{=} \left(\frac{\vartheta_\star}{8}\right)^{i+1}, \qquad |f_i^\flat - f_{i-1}^\flat|_i \leqslant \frac{1}{\delta}\left(\frac{\vartheta_\star}{8}\right)^{i+1}, \tag{116}$$

where

$$\vartheta_i := \delta |f_i^K|_i \,. \tag{117}$$

Let us first show that the inductive hypothesis is true for $j = 1$ (which implies $i = 1$). Indeed by (110) we see that we can apply Lemma 4.4 with $f \rightsquigarrow f_0^K$ and $\check\vartheta \rightsquigarrow \vartheta_0 = \delta|f_0^K|_0$. Thus, we obtain the existence of $\Psi_0 = X^1_{\phi_0}$, generated by a function $\phi_0 = \phi_0^K$ with

$$|\phi_0|_{r_0,s_0} \leqslant \frac{1}{\alpha}|f_0^K|_{r_0,s_0} \overset{(108)}{\leqslant} \frac{2^7 eK}{\alpha^2 r s}|f|^2_{r,s}\,, \tag{118}$$

satisfying (113) and[26] (114), so that $(h + f_0^K) \circ \Psi_0 =: h + \tilde f_1$ and, by (99) and (100),

$$|\tilde f_1|_1 \leqslant 4\vartheta_0 |f_0^K|_0 \overset{(110)}{\leqslant} \frac{1}{4}|f_0^K|_0 \overset{(108)}{\leqslant} \frac{2^5 eK}{\alpha r s}|f|^2_{r,s}\,. \tag{119}$$

We have that $f_1 = \tilde f_1 + f_0^\flat \circ \Psi_0$. Then[27]

$$f_1^K = \tilde f_1^K + (f_0^\flat \circ \Psi_0 - f_0^\flat)^K\,, \qquad f_1^\flat - f_0^\flat = \tilde f_1^\flat + (f_0^\flat \circ \Psi_0 - f_0^\flat)^\flat\,. \tag{120}$$

Write

$$f_0^\flat \circ \Psi_0 - f_0^\flat = (f_0^\flat - f^\flat)\circ\Psi_0 - (f_0^\flat - f^\flat) + (f^\flat \circ \Psi_0 - f^\flat - \{f^\flat, \phi_0\}) + \{f^\flat, \phi_0\}\,.$$

By (92) (with $u \rightsquigarrow f_0^\flat - f^\flat$, $r \rightsquigarrow r_0$, $s \rightsquigarrow s_0$) we have

$$|(f_0^\flat - f^\flat)\circ\Psi_0 - (f_0^\flat - f^\flat)|_1 \leqslant 2\vartheta_0 |f_0^\flat - f^\flat|_0 \leqslant \frac{2^4 eK}{\alpha r s}|f|^2_{r,s}$$

by (110) and (111). By (91) with $u \rightsquigarrow f^\flat$, $\phi \rightsquigarrow \phi_0$, $j \rightsquigarrow 1$, $\bar r \rightsquigarrow r_0$, $\bar s \rightsquigarrow s_0$,

$$|f^\flat \circ \Psi_0 - f^\flat - \{f^\flat, \phi_0\}|_1 \leqslant 2\vartheta_0 |\{f^\flat, \phi_0\}|_{r_0-\rho/2, s_0-\sigma/2} \leqslant \frac{2^9 K^3}{\alpha^2 r^2 s^2}|f|^3_{r,s} \overset{(73)}{\leqslant} \frac{K}{4\alpha r s}|f|^2_{r,s}\,,$$

---

[26] Note also that $(f_0^K)^\flat = 0$

[27] Note that $(f_0^\flat)^K = 0$ and $(f_0^\flat)^\flat = f_0^\flat$.

by (110), (118) and (87) (with $f \rightsquigarrow \phi_0$, $g \rightsquigarrow f^\flat$). Analaogously by (87) we get

$$\| \{f^\flat, \phi_0\} \|_1 \leqslant \frac{2^4 K^2}{ers} \|\phi_0\|_0 \|f\|_{r,s} \overset{(118)}{\leqslant} \frac{2^{11} K^3}{\alpha^2 r^2 s^2} \|f\|_{r,s}^3 \overset{(73)}{\leqslant} \frac{K}{\alpha r s} \|f\|_{r,s}^2 .$$

Summarizing:

$$\| f_0^\flat \circ \Psi_0 - f_0^\flat \|_1 \leqslant \frac{2^6 K}{\alpha r s} \|f\|_{r,s}^2 .$$

Then, by (119) and (120) we get

$$\| f_1^K \|_1 \, , \ \| f_1^\flat - f_0^\flat \|_1 \ \leqslant \ \frac{2^7 K}{\alpha r s} \|f\|_{r,s}^2 \tag{121}$$

checking (116) in the case $j = i = 1$.

Now take $2 \leqslant j \leqslant K$ and assume that the inductive hypothesis holds true for $1 \leqslant i \leqslant j$ and let us prove that it holds also for $i = j + 1$. By (116) and (73) we can apply Lemma 4.4 with $f \rightsquigarrow f_j^K$ and $\check\vartheta \rightsquigarrow \vartheta_j$. Thus, we obtain the existence of $\Psi_j = X^1_{\phi_j}$, generated by a function $\phi_j = \phi_j^K$ with

$$\|\phi_j\|_j \overset{(113)}{\leqslant} \frac{1}{\alpha} \|f_j^K\|_j \overset{(117)}{=} \frac{\vartheta_j}{\alpha \delta} , \tag{122}$$

so that $(h + f_j^K) \circ \Psi_j =: h + \tilde f_{j+1}$ and, by (99) and (100),

$$\|\tilde f_{j+1}\|_{j+1} \leqslant 4 \vartheta_j \|f_j^K\|_j \overset{(117)}{=} \frac{4}{\delta} \vartheta_j^2 \overset{(116)}{\leqslant} \frac{4}{\delta} (\vartheta_\star/8)^{2j+2} \overset{(73)}{\leqslant} \frac{1}{2^{3j-2}\delta} (\vartheta_\star/8)^{j+2} \leqslant \frac{1}{2^4 \delta} (\vartheta_\star/8)^{j+2} , \tag{123}$$

since $j \geqslant 2$. We have that $f_{j+1} = \tilde f_{j+1} + f_j^\flat \circ \Psi_j$. Then[28]

$$f_{j+1}^K = \tilde f_{j+1}^K + (f_j^\flat \circ \Psi_j - f_j^\flat)^K , \qquad f_{j+1}^\flat - f_j^\flat = \tilde f_{j+1}^\flat + (f_j^\flat \circ \Psi_j - f_j^\flat)^\flat . \tag{124}$$

Writing

$$f_j^\flat = f^\flat + (f_0^\flat - f^\flat) + \sum_{h=1}^{j} f_h^\flat - f_{h-1}^\flat$$

we have

$$\begin{aligned} f_j^\flat \circ \Psi_j - f_j^\flat \ &= \ \{f^\flat, \phi_j\} \\ &\quad + f^\flat \circ \Psi_j - f^\flat - \{f^\flat, \phi_j\} \\ &\quad + (f_0^\flat - f^\flat) \circ \Psi_j - (f_0^\flat - f^\flat) \\ &\quad + \sum_{i=1}^{j} \Big((f_i^\flat - f_{i-1}^\flat) \circ \Psi_j - (f_i^\flat - f_{i-1}^\flat) \Big) \end{aligned} \tag{125}$$

[28]Note that $(f_j^\flat)^K = 0$ and $(f_j^\flat)^\flat = f_j^\flat$.

where $\Psi_j = X^1_{\phi_j}$. By (87) with $f \rightsquigarrow \phi_j$, $g \rightsquigarrow f^\flat$, $r_0 \rightsquigarrow r_j$, $s_0 \rightsquigarrow s_j$, we get, by (113) and (117),

$$\|\{f^\flat, \phi_j\}\|_{j+1} \leqslant \frac{2^4 K^2}{ers} \|\phi_j\|_j \|f\|_{r,s} \leqslant \frac{2^4 K^2 \vartheta_j}{e\alpha rs\delta} \|f\|_{r,s} \stackrel{(73)}{=} \frac{1}{e 2^4 \delta} (\vartheta_\star/8) \vartheta_j \stackrel{(116)}{\leqslant} \frac{1}{e 2^4 \delta} (\vartheta_\star/8)^{j+2} .$$

By (91) with $u \rightsquigarrow f^\flat$, $\phi \rightsquigarrow \phi_j$, $j \rightsquigarrow 1$, $\bar{r} \rightsquigarrow r_j$, $\bar{s} \rightsquigarrow s_j$, reasoning as above we get

$$\|f^\flat \circ \Psi_j - f^\flat - \{f^\flat, \phi_j\}\|_{j+1} \leqslant \vartheta_j \|\{f^\flat, \phi_j\}\|_{r_j-\rho/2, s_j-\sigma/2} \leqslant \frac{\vartheta_j}{4e\delta} (\vartheta_\star/8)^{j+2} \leqslant \frac{1}{2^6 e\delta} (\vartheta_\star/8)^{j+2}$$

by (116) and (73). By (92) (with $u \rightsquigarrow f^\flat_0 - f^\flat$, $r \rightsquigarrow r_j$, $s \rightsquigarrow s_j$) we have

$$\|(f^\flat_0 - f^\flat) \circ \Psi_j - (f^\flat_0 - f^\flat)\|_{j+1} \leqslant 2\vartheta_j \|f^\flat_0 - f^\flat\|_j \leqslant \frac{2^8 eK}{\alpha rs} \|f\|^2_{r,s} \vartheta_j \leqslant \frac{1}{4\delta} (\vartheta_\star/8)^{j+2}$$

by (111), (116), (109) and (73). Analogously, for $1 \leqslant i \leqslant j$, by (92) (now with $u \rightsquigarrow f^\flat_i - f^\flat_{i-1}$)

$$\|(f^\flat_i - f^\flat_{i-1}) \circ \Psi_j - (f^\flat_i - f^\flat_{i-1})\|_{j+1} \leqslant 2\vartheta_j \|f^\flat_i - f^\flat_{i-1}\|_j \leqslant \frac{2}{\delta} (\vartheta_\star/8)^{j+i+2}$$

by (116). Then by (73)

$$\Big\|\sum_{i=1}^{j} \Big((f^\flat_i - f^\flat_{i-1}) \circ \Psi_j - (f^\flat_i - f^\flat_{i-1})\Big)\Big\|_{j+1} \leqslant \frac{2}{7\delta} (\vartheta_\star/8)^{j+2} .$$

Whence:

$$\|f^\flat_j \circ \Psi_j - f^\flat_j\|_{j+1} \leqslant \frac{4}{7\delta} (\vartheta_\star/8)^{j+2} .$$

Then by (123) we get

$$\|\tilde{f}_{j+1}\|_{j+1} + \|f^\flat_j \circ \Psi_j - f^\flat_j\|_{j+1} \leqslant \frac{1}{\delta} (\vartheta_\star/8)^{j+2} .$$

By (124) we get (116) with $i = j + 1$. This completes the proof of the induction.
Now, we can conclude the proof of Proposition 4.1. Set

$$\Psi := \Psi_{-1} \circ \Psi_0 \circ \cdots \circ \Psi_{K-1} \ .$$

Notice that, by (112), $r_K = r/2 = r_\star$ and $s_K = s(1 - 1/K) = s_\star$. By the induction, it is

$$H \circ \Psi = H_{K-1} \circ \Psi_{K-1} \stackrel{(115)_K}{=} h + f_K =: h + f^\flat + f_\star , \tag{126}$$

with $f^\flat = \mathtt{p}_\Lambda f + T_K^\perp \mathtt{p}_\Lambda^\perp f$ (recall (76)). Note that by (116) and (110)

$$\sum_{i=1}^K \vartheta_{i-1} \leqslant \sum_{i=1}^K (\vartheta_\star/8)^i \leqslant \vartheta_\star/7\,. \tag{127}$$

Since $(y', x') = (y^{(K)}, x^{(K)})$ by (105), (114) and triangular inequality we get

$$\begin{aligned} |y'-y|_1 &\leqslant |y-y^{(0)}|_1 + \sum_{i=1}^K |y^{(i)} - y^{(i-1)}|_1 \leqslant \frac{r\vartheta_{-1}}{16e} + \frac{r}{16eK}\sum_{i=1}^K \vartheta_{i-1} \\ &\overset{(127)}{\leqslant} \frac{r}{16e}\left(\vartheta_{-1} + \frac{\vartheta_\star}{7K}\right) \overset{(104)}{\leqslant} \frac{r}{16e}\left(\frac{\vartheta_\star}{8K} + \frac{\vartheta_\star}{7K}\right)\,, \end{aligned}$$

then (75) follows (the estimate on the angle being analogous).
Since $T_K P_\Lambda^\perp f^\flat = (f^\flat)^K = 0$ (for any f, recall (94)) we have

$$|T_K P_\Lambda^\perp f_\star|_{r_\star,s_\star} = |f_K^K|_K \overset{(117)}{=} \delta^{-1}\vartheta_K \overset{(116)}{\leqslant} \delta^{-1}(\vartheta_\star/8)^{K+1} = (\vartheta_\star/8)^K \frac{8}{eK}|f|_{r,s}\,, \tag{128}$$

proving the second estimates in (77).
Finally, (using that $K \geqslant 2$ and that $\vartheta_\star \leqslant 1$)

$$\begin{aligned} |f_\star|_{r_\star,s_\star} &\overset{(126)}{=} |f_K - f^\flat|_K \overset{(94)}{=} |f_K^K + f_K^\flat - f^\flat|_K \leqslant |f_K^K|_K + |f_0^\flat - f^\flat|_0 + \sum_{i=1}^K |f_i^\flat - f_{i-1}^\flat|_i \\ &\overset{(111),(116)}{\leqslant} |f_K^K|_K + \frac{1}{4K}\vartheta_\star |f|_{r,s} + \frac{1}{\delta}\sum_{i=1}^K (\vartheta_\star/8)^{i+1} \\ &\overset{(128),(127)}{\leqslant} (\vartheta_\star/8)^K \frac{8}{eK}|f|_{r,s} + \frac{1}{4K}\vartheta_\star|f|_{r,s} + \frac{\vartheta_\star^2}{56\delta} \leqslant \frac{1}{K}\vartheta_\star|f|_{r,s}\,, \end{aligned}$$

which proves also the first estimate in (77). ∎

5 Geometry of resonances

We, first, discuss the Covering Lemma in frequency space in a ball $B_M(0) \subset \{\omega \in \mathbb{R}^n\}$ and then we shall pull back through ω^{-1} in the action domain.
We define a *covering* $\{\Omega^i\}$ of $B_M(0)$

$$\Omega^0 \cup \Omega^1 \cup \Omega^2 \supset B_M(0)\ , \tag{129}$$

as follows.

Ω^0: The definition of the *completely non–resonant zone* Ω^0 is nearly tautological:

$$\Omega^0 := \{\omega \in B_M(0) : \min_{k\in\mathcal{G}^n_{1,K_1}} |\omega \cdot k| > \alpha/2\} . \tag{130}$$

Ω^1: Recalling the definition of $\Omega^{1,k}$ in (27) we set

$$\Omega^1 := \bigcup_{k\in\mathcal{G}^n_{1,K_1}} \Omega^{1,k} . \tag{131}$$

Ω^2: The set Ω^2 is the union of neighbourhoods of exact double resonances[29]

$$R_{k,\ell} := \{\omega \cdot k = \omega \cdot \ell = 0\} , \qquad k \in \mathcal{G}^n_{1,K_1} , \ell \in \mathcal{G}^n_{1,K_2}, \ \ell \notin \mathbb{Z}k , \tag{132}$$

namely:

$$\Omega^2 = \bigcup_{k\in\mathcal{G}^n_{1,K_1}} \bigcup_{\substack{\ell\in\mathcal{G}^n_{1,K_2} \\ \ell\notin\mathbb{Z}k}} \Omega^2_{k,\ell} \tag{133}$$

where

$$\Omega^2_{k,\ell} := \{|\omega \cdot k| < \alpha\} \cap \{|\,\mathrm{p}^\perp_k\omega| < M\} \cap \{|\,\mathrm{p}^\perp_k\omega \cdot \ell| \leqslant 3\alpha K_2/|k|\} . \tag{134}$$

Indeed, *from these definitions,* (129) *follows immediately.*

Next, let us point out the non–resonance properties satisfied by the frequencies in Ω^i.

(i) If $0 \neq |k|_1 \leqslant K_1$, then there exists $\bar{k} \in \mathcal{G}^n_{1,K_1}$ and a $0 \neq j \in \mathbb{Z}$ such that $k = j\bar{k}$ and, therefore,

$$\omega \in \Omega^0 \quad \Longrightarrow \quad |\omega \cdot k| = |j||\omega \cdot \bar{k}| \geqslant |\omega \cdot \bar{k}| \geqslant \min_{k\in\mathcal{G}^n_{1,K_1}} |\omega \cdot k| > \alpha/2 . \tag{135}$$

(ii) Let $\omega \in \Omega^{1,k}$ with $k \in \mathcal{G}^n_{1,K_1}$ and let $\ell \notin \mathbb{Z}k$, $|\ell|_1 \leqslant K_2$. Then, there exist $j \in \mathbb{Z}\backslash\{0\}$ and $\ell' \in \mathcal{G}^n_{1,K_2}$ such that $\ell = j\ell'$. Hence,

$$\begin{aligned} |\omega \cdot \ell| &= |j|\,|\omega \cdot \ell'| \geqslant |\omega \cdot \ell'| = \Big|\frac{(\omega \cdot k)(k \cdot \ell')}{|k|^2} + \mathrm{p}^\perp_k\omega \cdot \ell'\Big| \\ &\geqslant |\,\mathrm{p}^\perp_k\omega \cdot \ell'| - \frac{\alpha K_2}{|k|} > \frac{3\alpha K_2}{|k|} - \frac{\alpha K_2}{|k|} = \frac{2\alpha K_2}{|k|} . \end{aligned} \tag{136}$$

[29] Recall (22).

(iii) It remains to evaluate the measure of Ω^2. To do this, we first prove the following

Lemma 5.1 *If $\omega \in \Omega^2_{k,\ell}$ with $k \in \mathcal{G}^n_{1,K_1}$, $\ell \in \mathcal{G}^n_{1,K_2}$, $\ell \notin \mathbb{Z}k$, then*

$$\mathrm{dist}(\omega, R_{k,\ell}) \leqslant \sqrt{10}\,\alpha K_2 |k|\,|\ell|\,. \tag{137}$$

Moreover,

$$\mathrm{meas}\,(\Omega^2_{k,\ell}) \leqslant 3 \cdot 2^n\, M^{n-2} \alpha^2 \frac{K_2}{|k|}\ . \tag{138}$$

Proof Let $v \in \mathbb{R}^n$ be the projection of ω onto $R^\perp_{k,\ell}$, which is the plane generated by k and ℓ (recall that, by hypothesis, k and ℓ are not parallel). Then,

$$\mathrm{dist}\,(\omega, R_{k,\ell}) = \mathrm{dist}\,(v, R_{k,\ell}) = |v| \tag{139}$$

and

$$|v \cdot k| = |\omega \cdot k| < \alpha\,, \qquad |\,\mathtt{p}^\perp_k v \cdot \ell| = |\,\mathtt{p}^\perp_k \omega \cdot \ell| \leqslant 3\alpha K_2/|k|\,. \tag{140}$$

Set

$$h := \mathtt{p}^\perp_k \ell = \ell - \frac{\ell \cdot k}{|k|^2} k\,. \tag{141}$$

Then, v decomposes in a unique way as

$$v = ak + bh$$

for suitable $a, b \in \mathbb{R}$. By (140),

$$|a| < \frac{\alpha}{|k|^2}\,, \qquad |\,\mathtt{p}^\perp_k v \cdot \ell| = |bh \cdot \ell| \leqslant 3\alpha K_2/|k|\,, \tag{142}$$

and

$$|h \cdot \ell| \overset{(141)}{=} \frac{|\ell|^2|k|^2 - (\ell \cdot k)^2}{|k|^2} \geqslant \frac{1}{|k|^2}$$

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

$$|b| \leqslant 3\alpha K_2 |k|\,, \tag{143}$$

and (137) follows since $|h| \leqslant |\ell|$ and $|v| = \sqrt{a^2|k|^2 + b^2|h|^2} \leqslant \sqrt{10}\alpha K_2\,|k|\,|\ell|$.
To estimate the measure of $\Omega^2_{k,\ell}$ we write $\omega \in R_{k,\ell}$ as $\omega = v + v^\perp$ with $v^\perp$ in the orthogonal complement of the plane generated by k and ℓ. Since $|v^\perp| \leqslant |\omega| < M$

and v lies in a rectangle of sizes of length $2\alpha/|k|^2$ and $6\alpha K_2|k|$ (compare (142) and (143)) we find

$$\text{meas}\,(\Omega^2_{k,\ell}) \leqslant \frac{2\alpha}{|k|^2}\,(6\alpha K_2|k|)(2M)^{n-2} = 3\cdot 2^n\, M^{n-2}\alpha^2\frac{K_2}{|k|}\ , \tag{144}$$

finishing the proof of Lemma 5.1. ∎

From (133) and (144) it follows immediately (recall that $n \geqslant 2$) that

$$\text{meas}\,(\Omega^2) \leqslant cM^{n-2}\alpha^2 K_2^{n+1}\, K_1^{n-1}\ , \tag{145}$$

for a suitable constant c depending only on n.

Proof (of Proposition 2.1) Recalling (130), (131) and (133), for $i = 0, 1, 2$ set

$$D^i := \{y \in D : \omega(y) \in \Omega^i\}\ . \tag{146}$$

Then (129) implies (23), while (135), (136) and (145) imply immediately (24), (25) and[30] (26) respectively, proving Proposition 2.1. ∎

6 Averaging Theory

Assumption B
Let $r, s > 0$ and let h satisfy Assumption **A** *in § 2.*
Let $f : D_r \times \mathbb{T}^n_s \to \mathbb{C}$ be a holomorphic function with

$$\|f\|_{D,r,s} = 1 \tag{147}$$

and define

$$H_\varepsilon(y,x) := h(y) + \varepsilon f(y,x)\ , \qquad (y,x) \in D_r \times \mathbb{T}^n_s\ ,\ \varepsilon > 0\ . \tag{148}$$

Let K_2, K_1, ν and α be such that

$$K_2 \geqslant 3K_1 \geqslant 6\ , \qquad \nu \geqslant n+2\,, \qquad \alpha := \sqrt{\varepsilon}K_2^\nu\,. \tag{149}$$

[30]Recall the definition of $\bar{L}$ in Assumption **A**, § 2.

For $k \in \mathcal{G}^n_{1,K_1}$, *define*

$$r_0 := \frac{\alpha}{4LK_1} = \sqrt{\varepsilon}\frac{K_2^\nu}{4LK_1} \ ; \qquad\qquad r_k := \frac{\alpha}{L|k|} = \sqrt{\varepsilon}\frac{K_2^\nu}{L|k|} \ , \tag{150}$$

$$\bar\vartheta := 2^{14} n^{2n} \ \frac{L}{s^{2n+1}} \ \frac{1}{K_2^{2\nu-2n-3}} \ ; \qquad\qquad \vartheta := 2^{2n+10} n^{2n} \frac{L}{s^{2n+1}} \frac{1}{K_2^{2\nu-2n-3}} \ . \tag{151}$$

Putting together the Normal Form Lemma (Proposition 4.1) and the Covering Lemma (Proposition 2.1) there follows easily the following *averaging theorem for non–resonant and simply resonant zones*:

Theorem 6.1 *Let Assumption* **B** *hold and assume that* ε *satisfies* (40) *and*

$$K_2^{2\nu-n-4} \geqslant 2^{13+n} n^n \ \frac{Le^{s/2}}{s^{n+1}} \ , \tag{152}$$

then the following holds.
(i) *There exists a symplectic change of variables*

$$\Psi_0 : D^0_{r_0/2} \times \mathbb{T}^n_{s(1-1/K_1)^2} \to D^0_{r_0} \times \mathbb{T}^n_{s(1-1/K_1)} \,, \tag{153}$$

such that

$$H_\varepsilon \circ \Psi_0 = h(y) + \varepsilon g^{\rm o}(y) + \varepsilon f^{\rm o}_{\star\star}(y,x) \ , \qquad\qquad \langle f^{\rm o}_{\star\star}\rangle = 0 \,, \tag{154}$$

where $\langle\cdot\rangle = \mathtt{p}_{\{0\}}$ *denotes the average with respect to the angles* x *and*

$$\sup_{D^0_{r_0/2}} |g^{\rm o} - \langle f\rangle| \leqslant \bar\vartheta \,, \qquad \| f^{\rm o}_{\star\star}\|_{D^0, r_0/2, s(1-1/K_1)/2} \leqslant 2\Big(\frac{2nK_1}{s}\Big)^n e^{-(K_1-3)s/2} \,. \tag{155}$$

(ii) $D^1 = \displaystyle\bigcup_{k\in\mathcal{G}^n_{1,K_1}} D^{1,k}$ *and for any* $k \in \mathcal{G}^n_{1,K_1}$ *there exists a symplectic change of variables*

$$\Psi_k : D^{1,k}_{r_k/2} \times \mathbb{T}^n_{s_\star} \to D^{1,k}_{r_k} \times \mathbb{T}^n_{s(1-1/K_2)} \,, \qquad\qquad s_\star := s(1-1/K_2)^2 \,, \tag{156}$$

such that

$$H_\varepsilon \circ \Psi_k = h(y) + \varepsilon g^k(y,x) + \varepsilon f^k_{\star\star}(y,x) \tag{157}$$

where

$$g^k = \mathtt{p}_{k\mathbb{Z}} g^k \,, \qquad\qquad \mathtt{p}_{k\mathbb{Z}} f^k_{\star\star} = 0 \,, \tag{158}$$

and

$$\| g^k - \mathtt{p}_{k\mathbb{Z}} f\|_{D^{1,k}, r_k/2, s_\star} \leqslant \vartheta \,, \qquad \| f^k_{\star\star}\|_{D^{1,k}, r_k/2, s(1-1/K_2)/2} < 2\Big(\frac{2nK_2}{s}\Big)^n e^{-(K_2-3)s/2} \ . \tag{159}$$

Remark 6.1 (i) The functions g^k and $\mathtt{p}_{_{k\mathbb{Z}}} f$ depend, effectively, only on one angle $\theta \in \mathbb{T}^1$: more precisely, setting

$$\begin{cases} F_j^k(y) := f_{jk}(y) \\ \\ G_j^k(y) := g_{jk}^k(y) \end{cases} \qquad \begin{cases} F^k(y,\theta) := \sum_{j\in\mathbb{Z}} F_j^k(y) e^{\mathrm{i}j\theta} \\ \\ G^k(y,\theta) := \sum_{j\in\mathbb{Z}} G_j^k(y) e^{\mathrm{i}j\theta} \end{cases} \tag{160}$$

we have (recall (63))

$$(\mathtt{p}_{_{k\mathbb{Z}}} f)(y,x) = F^k(y, k\cdot x)\,, \qquad g^k(y,x) = G^k(y, k\cdot x)\,. \tag{161}$$

From (159) and (64) it follows

$$\left|\!\left| G^k - F^k \right|\!\right|_{D^{1,k}, r_k/2, |k|_1 s_*} \leqslant \vartheta\ . \tag{162}$$

The function $\theta \in \mathbb{T}^1_{|k|_1 s_*} \to G^k(y,\theta)$ will be called the **effective potential** since, disregarding the small remainder f^k_{**}, it governs the Hamiltonian evolution at simple resonances.

(ii) We have assumed that $\|f\|_{r,s} = 1$ (see (147)), since this is the natural assumption in term of genericity properties, however the Normal Form Lemma is formulated in term of the *stronger* norm $|\!|\cdot|\!|$. We need therefore to restrict slightly the angle–analyticity domain in order to pass to the norm $|\!|\cdot|\!|$. This can be done through (56), which yields (for $r = r_0$ or $r = r_k$ and $K = K_{_1}$ or $K_{_2}$)

$$|\!| f |\!|_{r, s(1-1/K)} \overset{(56),(147)}{\leqslant} \left(\frac{2nK}{s}\right)^n\ . \tag{163}$$

(iii) The choice of α in (149) is not restrictive (since it is done through the introduction of ν, a new parameter) and it has the effect of making disappear ε from the smallness conditions and from the definition of the smallness parameters $\bar\vartheta$ and ϑ.
According to the choice of $K_{_1}$ and $K_{_2}$ one will get different kind of statements.

Remark 6.2 (i) Observe that $r_0 \leqslant r_k \leqslant \sqrt{\varepsilon} K_{_2}^\nu / L$ so that assumption (40) ensures the necessary condition:

$$r_0 \leqslant r_k \leqslant \frac{\sqrt{\varepsilon} K_{_2}^\nu}{L} \leqslant r\ . \tag{164}$$

(ii) The hypotheses of the Normal Form Lemma (Proposition 4.1) concern *a complex domain* D_r, while the non–resonance properties of the Covering Lemma (Proposition 2.1) hold on *real domains*. The following simple observation allows to use directly the Covering Lemma:

If a set $D \subseteq \mathbb{R}^n$ is (α, K) non–resonant modulo Λ for h, then the complex domain D_r is $(\alpha - LrK, K)$ non–resonant modulo Λ, provided $LrK < \alpha$, where L is the Lipschitz constant of ω on the complex domain D_r.

Indeed, if $y \in D_r$ there exists $y_0 \in D$ such that $|y - y_0| < r$ and $|\omega(y_0) \cdot k| \geqslant \alpha$ for all $k \in \mathbb{Z}^n \setminus \Lambda$, $|k|_1 \leqslant K$. Thus, for such k's, one has

$$|\omega(y) \cdot k| = |\omega(y_0) \cdot k - (\omega(y_0) - \omega(y)) \cdot k| \geqslant |\omega(y_0) \cdot k| - LrK \geqslant \alpha - LrK \ . \quad \blacksquare$$

Proof (of Theorem 6.1) (i): By Remark 6.2–(ii), (24) and the choice of r_0 in (150), the domain $D^0_{r_0}$ is $(\alpha/4, K_1)$ completely non–resonant (or non–resonant modulo the trivial lattice $\{0\}$) and, in view of (163) and (152), one can apply Proposition 4.1 to H_ε in (148) with[31]

$$\begin{aligned} &f \rightsquigarrow \varepsilon f \ , \quad D \rightsquigarrow D^0 \ , \quad r \rightsquigarrow r_0 \ , \quad \Lambda \rightsquigarrow \{0\} \ , \quad \alpha \rightsquigarrow \alpha/4 \ , \\ &K \rightsquigarrow K_1 \ , \quad s \rightsquigarrow s(1 - 1/K_1) \ . \end{aligned} \tag{165}$$

Thus, recalling (73), using (150) and that $K_1 \geqslant 2$, one sees that

$$\begin{aligned} \vartheta_* \rightsquigarrow \vartheta_0 \quad &:= \quad 2^{15} \frac{L K_1^3 \|f\|_{r_0, s(1-1/K_1)}}{K_2^{2\nu} s(1 - 1/K_1)} \\ &\overset{(163),(149)}{<} \quad 2^{16} \frac{L K_1^3}{s K_2^{2\nu}} \Big(\frac{2nK_1}{s}\Big)^n \\ &\overset{(149)}{\leqslant} \quad 2^{13} n^n \frac{L}{s^{n+1}} \frac{1}{K_2^{2\nu - n - 3}} \overset{(152)}{\leqslant} e^{-s/2} \leqslant 1 \ , \end{aligned} \tag{166}$$

showing that (73) holds and also that $\bar{s} \rightsquigarrow s(1 - 1/K_1)/2$ in (80). Then, by (79) and (163), one has:

$$\sup_{D^0_{r_0/2}} |g^{\mathrm{o}} - \langle f \rangle| \leqslant \vartheta_0 \Big(\frac{2nK_1}{s}\Big)^n \overset{(149)}{\leqslant} \Big(\frac{nK_2}{s}\Big)^n \vartheta_0 \overset{(166),(151)}{\leqslant} \bar{\vartheta} \,,$$

$$\|f^{\mathrm{o}}_{**}\|_{D^0, r_0/2, s(1-1/K_1)/2} \leqslant 2 e^{-(K_1 - 2)s(1-1/K_1)/2} \Big(\frac{2nK_1}{s}\Big)^n \leqslant 2 \Big(\frac{2nK_1}{s}\Big)^n e^{-(K_1 - 3)s/2} \,,$$

from which (155) follows.

(ii): By Remark 6.2–(ii), the definition of r_k in (150) and (25), the domain $D^{1,k}_{r_k}$ is

$$(2\alpha K_2/|k| - r_k L K_2, K_2) = (\alpha K_2/|k|, K_2)$$

[31] Recall that the notation "$a \rightsquigarrow b$" means "with a replaced by b".

non–resonant modulo $\mathbb{Z}k$.
Using again (163), we can apply Proposition 4.1 with

$$\begin{array}{llll} f \rightsquigarrow \varepsilon f\ , & D \rightsquigarrow D^{1,k}\ , & r \rightsquigarrow r_k\ , & \alpha \rightsquigarrow \alpha K_2/|k|\ , \\ K \rightsquigarrow K_2\ , & s \rightsquigarrow s(1-1/K_2)\ , & \Lambda \rightsquigarrow \mathbb{Z}k\ , & \end{array} \tag{167}$$

and (recall (73) and that $|k| \leqslant K_1$)

$$\begin{aligned} \vartheta_\star \rightsquigarrow \vartheta_k \quad &:= \quad 2^{11}\, \frac{L K_2^2 |k|^2 \varepsilon \| f \|_{r_k, s(1-1/K_2)}}{\alpha^2 s(1-1/K_2)} \\ &\overset{(149),(163)}{\leqslant} \quad 2^{n+10} n^n \frac{L}{s^{n+1}} \frac{1}{K_2^{2\nu-n-4}} \overset{(152)}{\leqslant} e^{-s/2} \leqslant 1\ , \end{aligned} \tag{168}$$

showing that (73) holds and also that $\bar{s} \rightsquigarrow s(1-1/K_2)/2$ in (80). From (79), (168) and (163) there follows (159); indeed

$$\begin{aligned} \| g^k - \mathtt{p}_{k\mathbb{Z}} f \|_{D^{1,k}, r_k/2, s_\star} &\leqslant \frac{1}{K_2} \Big(\frac{2nK_2}{s} \Big)^n \vartheta_k \overset{(168),(151)}{\leqslant} \vartheta\ , \\ \| f_{\star\star}^k \|_{D^{1,k}, r_k/2, s(1-1/K_2)/2} &\leqslant 2 e^{-(K_2-2)s(1-1/K_2)/2} \Big(\frac{2nK_2}{s} \Big)^n \leqslant 2 \Big(\frac{2nK_2}{s} \Big)^n e^{-(K_2-3)s/2}\ , \end{aligned}$$

∎

7 Proofs of main results

Under the above standing hypotheses, apart from a finite number of simple resonances *the effective potential G^k at simple resonances is close to a (shifted) cosine*:

Proposition 7.1 *Let the assumptions of Theorem 6.1 hold, let $k \in \mathcal{G}^n_{1,K_1}$ and let G^k be as in (160), (158). Then, if*

$$|k|_1 > 3/s\ , \tag{169}$$

one has that

$$\| G^k - T_1 F^k \|_{D^{1,k}, r_k/2, 2} \leqslant \vartheta\, e^{s+5} e^{-|k|_1 s} + 2^8\, e^{-2|k|_1 s}\ . \tag{170}$$

Proof Observe that by definition of[32] T_N and $T_N^\perp$,

$$G^k - T_1 F^k = T_1 G^k - T_1 F^k + T_1^\perp G^k\ . \tag{171}$$

[32] Recall (59).

Now, since $3/s < |k|_1 \leqslant K_1 \leqslant K_2/3$,

$$\sup_{D^{1,k}_{r_k/2}} |G^k_{\pm 1} - F^k_{\pm 1}| \overset{(162)}{\leqslant} \vartheta e^{-|k|_1 s_\star} \overset{(156)}{\leqslant} \vartheta e^{-|k|_1 s(1-2/K_2)} \leqslant \vartheta e^s e^{-|k|_1 s}$$

so that

$$\|T_1 G^k - T_1 F^k\|_{D^{1,k}, r_k/2, 2} = |G^k_1 - F^k_1| e^2 + |G^k_{-1} - F^k_{-1}| e^2 < 2e^2 e^s \vartheta e^{-|k|_1 s} \ . \tag{172}$$

Next, recalling (160), we have that

$$\|T_1^\perp G^k\|_{D^{1,k}, r_k/2, 2} = \sum_{\substack{|j|\geqslant 2\\ j\in\mathbb{Z}}} |g_{jk}| e^{2|j|} \leqslant \sum_{|j|\geqslant 2} |f_{jk}| e^{2|j|} + \sum_{|j|\geqslant 2} |g_{jk} - f_{jk}| e^{2|j|} \ . \tag{173}$$

Let us estimate the two sums separately. Since $\|f\|_s = 1$, $|f_\ell| \leqslant e^{-|\ell|_1 s}$ so that $|f_{jk}| \leqslant e^{-|j||k|_1 s}$ and:

$$\sum_{|j|\geqslant 2} |f_{jk}| e^{2|j|} \leqslant \sum_{|j|\geqslant 2} e^{-|j||k|_1 s}\ e^{2|j|} = 2\ \frac{e^{-2(|k|_1 s - 2)}}{1 - e^{-(|k|_1 s - 2)}} \leqslant 4e^4 e^{-2(|k|_1 s)} \ , \tag{174}$$

where in the last inequality we used the assumption $|k|_1 s > 3 > 2 + \log 2$.
Then (again, because $|k|_1 s > 3$), we see that

$$\begin{aligned}
\sum_{|j|\geqslant 2} |g_{jk} - f_{jk}| e^{2|j|} &= \sum_{|j|\geqslant 2} |g_{jk} - f_{jk}| e^{|j||k|_1 s_\star}\ e^{-|j||k|_1 s_\star + 2|j|} \\
&\overset{(162)}{\leqslant} \sup_{j\geqslant 2} \left(e^{-j(|k|_1 s_\star - 2)} \right) \vartheta \leqslant e^{-2(|k|_1 s_\star - 2)}\ \vartheta \\
&\leqslant e^4 e^{-2|k|_1 s(1-2/K_2)} \vartheta \leqslant \vartheta\ e^{2s+4} e^{-2|k|_1 s} \ .
\end{aligned} \tag{175}$$

Putting (174) and (175) together, by (171) and (169), (170) follows. ∎

Proposition 7.2 *Let the assumptions of Theorem 6.1 hold; let $s > 0$, $0 < \delta \leqslant 1$ and fix any $0 < \gamma \leqslant 1$. Assume (38) and (34). If $k \in \mathcal{G}^n_{1,K_1}$ satisfies*

$$|k|_1 > \tau_o(\delta;\gamma) \ , \tag{176}$$

(with $\tau_o(\delta;\gamma)$ defined in (37)) then,

$$\|G^k - T_1 F^k\|_{D^{1,k}, r_k/2, 2} \leqslant \gamma\, \delta_k \ , \tag{177}$$

where

$$\delta_k := \delta |k|_1^{-n} e^{-|k|_1 s} \ . \tag{178}$$

Remark 7.1 Conditions (38) and (176) are stronger than the ones on ν in (149), (152) and (169). In particular the assumptions of Proposition 7.1 hold.

Proof of Proposition 7.2 As mentioned in the above remark, Proposition 7.1 holds. Let us estimate the two terms in (170) separately. Recalling the definition of ϑ in (151) (and that $|k|_{_1} \leqslant K_{_1} \leqslant K_{_2}/3$), we find:

$$\begin{aligned} \vartheta\, e^{s+5} e^{-|k|_1 s} &\overset{(151)}{=} e^{s+5} e^{-|k|_1 s} 2^{2n+10} n^{2n} \frac{L}{s^{2n+1}} \frac{1}{K_{_2}^{2\nu-2n-3}} \\ &= e^{s+5} 2^{2n+11} n^{2n} \frac{L}{s^{2n+1}} \frac{|k|_{_1}^n}{K_{_2}^{2\nu-2n-3}} \frac{1}{\gamma\delta} \frac{\gamma\delta_k}{2} \\ &\leqslant e^{s+5} 2^{n+11} n^{2n} \frac{L}{s^{2n+1}} \frac{1}{K_{_2}^{2\nu-3n-3}} \frac{1}{\gamma\delta} \frac{\gamma\delta_k}{2} \\ &\overset{(38)}{\leqslant} \frac{\gamma\delta_k}{2}\,. \end{aligned} \tag{179}$$

As for the second term in (170), we use the following calculus lemma, whose elementary check is left to the reader:

Lemma 7.1 *If $a > 2\log 2$, $0 < \varepsilon < e^{-a^2/2}$ and $t > 4\log\varepsilon^{-1}$, then $e^{-t}t^a < \varepsilon$.*

Indeed, by the lemma (with $a \rightsquigarrow n$, $t \rightsquigarrow s|k|_{_1}$ and $\varepsilon \rightsquigarrow s^n\gamma\delta/2^9$) and in view of (34) and (176), one has

$$2^8 e^{-2|k|_1 s} = (s|k|_{_1})^n e^{-|k|_1 s} \frac{2^9}{s^n\gamma\delta} \cdot \frac{\gamma\delta_k}{2} < \frac{\gamma\delta_k}{2}\,. \tag{180}$$

The bounds (179) and (180) prove the claim. ∎

The quantity δ_k defined in (178) is a "Fourier–measure" for the non–degeneracy of analytic potentials f holomorphic on $\mathbb{T}^n_s$, since such potential will have, in general, Fourier coefficients $f_k \sim e^{-s|k|_1}$.

7.1 Positional potentials: proof of Theorem 2.1

In order to conclude the proof of Theorem 2.1 we need the following

Lemma 7.2 *Let $s > 0$, $0 < \delta \leqslant 1$ and fix any $0 < \gamma \leqslant 1$. Let the assumptions of Theorem 6.1 hold; assume (38), (34) and that the positional potential $f \in \mathcal{H}_{s,\tau_o}(\delta)$ with the tail function τ_o defined in (176). Then, for $\tau_o(\delta;\gamma) \leqslant |k|_1 \leqslant K_1 \leqslant K_2/3$, one has*

$$\sup_{y \in D^{1,k}_{r_k/2}} \frac{\| G^k(y,\cdot) - T_1 F^k(\cdot) \|_2}{|f_k|} \leqslant \gamma \tag{181}$$

and

$$\frac{1}{|f_k|} \| f^k_{**} \|_{D^{1,k}, r_k/2, s(1-1/K_2)/2} \leqslant \frac{2^{10n} n^{3n}}{s^{3n}\delta} e^{-K_2 s/8} \,. \tag{182}$$

Proof Since Proposition 7.2 holds, by (177) and since $f \in \mathcal{H}_{s,\tau_o}(\delta)$ we get (181). By (159) we get

$$\begin{aligned}
\frac{1}{|f_k|} \| f^k_{**} \|_{D^{1,k}, r_k/2, s(1-1/K_2)/2} &< \frac{|k|_1^n e^{|k|_1 s}}{\delta} 2 \Big(\frac{2nK_2}{s} \Big)^n e^{-(K_2-3)s/2} \\
&\leqslant \frac{2^{n+1} n^n}{s^n \delta} K_1^n K_2^n e^{-(K_2 - 2K_1)s/2} \overset{(149)}{\leqslant} \frac{n^n}{s^n\delta} K_2^{2n} e^{-K_2 s/6} \,.
\end{aligned}$$

Then using that[33]

$$K_2^{2n} e^{-K_2 s/24} \leqslant \Big(\frac{48n}{s\, e} \Big)^{2n} \leqslant \frac{2^{10n} n^{2n}}{s^n} \,,$$

we prove (182). ∎

Recalling the definition of T_N given in (59), we have that

$$T_1 F^k(\theta) = f_k e^{\mathrm{i}\theta} + f_{-k} e^{-\mathrm{i}\theta} = 2|f_k| \cos(\theta + \theta^{(k)})$$

for a suitable constant $\theta^{(k)}$. Setting

$$\begin{aligned}
\mathtt{G}^k(y,\theta) &:= \frac{G^k(y,\theta) - \cos(\theta + \theta^{(k)})}{2|f_k|} \,, \\
\mathtt{f}^k(y,x) &:= \frac{f^k_{**}(y,\theta)}{2|f_k|} \,,
\end{aligned}$$

we get (42). Finally Theorem 2.1 follows from Lemma 7.2, in particular (43) and (44) follow from (181) and (182), respectively. ∎

[33] Using that for $\alpha > 0$ we have $\max_{x>0} x^\alpha e^{-x} = (\alpha/e)^\alpha$.

7.2 The general case (y-dependent potentials)

For $k \in \mathbb{Z}^n \backslash \{0\}$ let $b_k > 0$ such that

$$\sum_{k \neq 0} b_k < \infty \, .$$

For $\mathcal{Z} \subseteq \mathbb{Z}^n \backslash \{0\}$ we set

$$b_{\mathcal{Z}} := \sum_{k \in \mathcal{Z}} b_k \, .$$

For definiteness we will fix

$$b_k := |k|^{-\frac{n}{2}} \, ,$$

but every other possible choice is fine.

Proposition 7.3 *Let $r, \mu > 0$ and $\mathcal{Z} \subseteq \mathbb{Z}^n \backslash \{0\}$. For any $k \in \mathcal{Z}$ let $\varphi_k(y)$ be holomorphic functions on the complex ball $\{y \in \mathbb{C}^n \ : \ |y| < r\}$ with*

$$\sup_{|y|<r} |\varphi_k(y)| \leqslant 1 \, , \qquad \text{and} \qquad |\varphi_k(0)| \geqslant \hat{\delta}_k > 0 \, .$$

Then, for every $y \in \mathbb{R}^n$ with $|y| < r/2e$, up, at most, to a set of measure[34]

$$\frac{1}{2} \operatorname{meas}_{n-1}(S^{n-1}) b_{\mathcal{Z}} \left(\frac{r}{2e}\right)^n \mu \, ,$$

we have

$$|\varphi_k(y)| \geqslant \hat{\delta}_k \left(\frac{\mu b_k}{30 e^3}\right)^{\log 1/\hat{\delta}_k} \, , \qquad \forall \, k \in \mathcal{Z} \, . \tag{183}$$

The proof relies on the following classical result in function theory (see, e.g., [10]):

Lemma 7.3 (Cartan's Estimate) *Assume that $\mathtt{f} : \mathbb{C} \to \mathbb{C}$ is holomorphic and bounded by $M > 0$ on the complex ball $|z| < 2eR$. If $|\mathtt{f}(0)| = 1$ then, for $0 < \eta < 1$*

$$|\mathtt{f}(z)| \geqslant \left(\frac{\eta}{15 e^3}\right)^{\log M} \tag{184}$$

for any $z \in \mathbb{C}$, $|z| < R$ up to a set of balls of radii r_j satisfying

$$\sum_j r_j \leqslant \eta R \, .$$

[34] As usual $S^{n-1} := \{y \in \mathbb{R}^n | \ |y| = 1\}$.

Remark 7.2 *Note that (184) holds in the complex ball $|z| < R$ up to a set of measure smaller than $\pi\eta^2 R^2$. Moreover it holds on the real interval $(-R, R)$ up to a set of (real) measure $2\eta R$.*

Proof of Proposition 7.3 Fix $k \in \mathcal{Z}$. Fix $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ with $|\xi| = 1$ and $\xi_1 \geqslant 0$. We apply Cartan's estimates simultaneously for every $k \in \mathcal{Z}$ with

$$\mathtt{f}(z) \rightsquigarrow \frac{\varphi_k(z\xi)}{\varphi_k(0)}\,, \qquad R \rightsquigarrow \frac{r}{2e}\,, \qquad M \rightsquigarrow \frac{1}{\hat{\delta}_k}\,, \qquad \eta \rightsquigarrow \frac{\mu}{2}\, b_k.$$

By (184) estimate (183) for the fixed k holds on the segment $\{y\xi \ : \ y \in (-r/2e, r/2e)\}$, up, at most, to a set of measure[35]

$$\frac{\mu}{2e}\, b_k r\ .$$

Integrating on the half-sphere $|\xi| = 1$, $\xi_1 \geqslant 0$, we get that (183) for the fixed k holds on the ball $|y| < r/2e$ up, at most, to a set of measure

$$\frac{1}{2}\,\mathrm{meas}_{\,n-1}(S^{n-1}) \left(\frac{r}{2e}\right)^n \mu b_k\,.$$

Summing on all $k \in \mathcal{Z}$ we get that (183) holds for all $k \in \mathcal{Z}$. ∎

Fix $0 < \mu, \gamma < 1$. Define the following tail function

$$\tau_*(\delta; \gamma, \mu) := \tag{185}$$

$$\frac{2^6 n^2}{\tilde{s}} \max\left\{\log^3 \frac{2^6 n^2}{\tilde{s}}\,,\ \left(\log \frac{30e^3}{\delta\mu}\right) \log^2\left(\frac{4}{\tilde{s}} \log \frac{30e^3}{\delta\mu}\right)\,,\ \log \frac{30e^3}{\mu} \log \frac{1}{\delta}\,,\ \log \frac{2^{10}}{\delta\gamma}\right\}\,,$$

where

$$\tilde{s} := \min\{s, 1\}\,.$$

Fix $y_0 \in D$ and assume that

$$f(y_0, \cdot) \in \mathcal{H}_{s,\tau_*}(\delta)\,. \tag{186}$$

Set

$$\varphi_k(y) := f_k(y) e^{|k|_1 s}\,. \tag{187}$$

[35] Recall Remark 7.2.

We have that

$$\sup_{y\in\mathbb{C}^n,\,|y-y_0|<r} |\varphi_k(y)| \overset{(147)}{\leqslant} 1\,, \qquad |\varphi_k(y_0)| \geqslant \hat\delta_k := \delta/|k|_1^n\,, \qquad \forall\, k\in\mathcal{G}^n_1\,, \quad |k|_1 > \tau_*(\delta)\,. \tag{188}$$

Let $\mu > 0$. Then by Proposition[36] 7.3 there exists a set[37]

$$\mathcal{D}\subseteq B_{r/2e}(y_0) \quad \text{satisfying} \quad \text{meas}\,(B_{r/2e}(y_0)\backslash\mathcal{D}) \leqslant \frac{b}{2}\,\text{meas}_{\,n-1}(S^{n-1})\left(\frac{r}{2e}\right)^n \mu\,, \tag{189}$$

with

$$b := \sum_{|k|_1>\tau_*(\delta;\gamma,\mu)} |k|_1^{-n/2} \leqslant \sum_{k\neq 0} |k|_1^{-n/2}\,,$$

such that

$$\begin{aligned} &|f_k(y)|e^{|k|_1 s} = |\varphi_k(y)| \geqslant \delta_k(\mu) := \hat\delta_k\left(\frac{\mu}{30e^3|k|_1^{n/2}}\right)^{\log 1/\hat\delta_k}\,, \\ &\forall\, y\in\mathcal{D}\,,\ k\in\mathcal{G}^n_1\,, \quad |k|_1 > \tau_*(\delta)\,. \end{aligned} \tag{190}$$

Theorem 7.1 *Let the assumption of Theorem 6.1 hold. Fix $0<\mu,\delta<1/e^8$ and $0<\gamma<1$. Assume that for some $y_0\in D$ we have $f(y_0,\cdot)\in\mathcal{H}_{s,\tau_*}(\delta)$. Set*

$$\tilde\mu := \mu/30e^3\,, \quad \tilde n := 2\nu - 2n - 3\,, \quad \kappa := 2^{2n+10} n^{2n}\frac{L}{s^{2n+1}}\,. \tag{191}$$

Assume that

$$K_2 \geqslant \max\left\{ K_1^{\frac{2n^2}{\tilde n}\log K_1}\,,\ K_1^{\frac{9}{\tilde n}\log\frac{1}{\delta\tilde\mu}}\,,\ e^{\frac{4}{\tilde n}\log\frac{1}{\delta}\log\frac{1}{\tilde\mu}}\,,\ \left(\frac{4e^{s+5}\kappa}{\delta\gamma}\right)^{\frac{4}{\tilde n}}\,,\ \frac{2^5}{s}\log^2\frac{1}{\delta\mu}\,,\ \frac{2^{14}n^4}{s^2}\right\}. \tag{192}$$

If $k\in\mathcal{G}^n_{1,K_1}$ with $|k|_1 > \tau_*(\delta;\gamma,\mu)$ then

$$\sup_{y\in(D^{1,k}\cap\mathcal{D})_{\hat r_k}} \frac{\|G^k(y,\cdot) - T_1F^k(y,\cdot)\|_2}{|f_k(y)|} \leqslant \gamma\,, \tag{193}$$

$$\sup_{y\in(D^{1,k}\cap\mathcal{D})_{\hat r_k}} \frac{\|f^k_{**}(y,\cdot)\|_{s(1-1/K_2)/2}}{|f_k(y)|} < \frac{4e^{3s/2}n^n}{\delta s^n}e^{-K_2 s/8}\,, \tag{194}$$

[36] With $\varphi_k(y) \rightsquigarrow \varphi_k(y+y_0)$.

[37] Both $\mathcal{D}$ and $B_{r/2e}(y_0)$ are real sets.

where $\mathcal{D}$ was defined in (189) *and*[38]

$$\hat{r}_k := \frac{1}{2}\min\{r_k\,,\,\delta_k(\mu)\}\,. \tag{195}$$

Proof First we note that by (190), (188), (195) and Cauchy estimates

$$|\varphi_k(y)| \geqslant \frac{1}{2}\delta_k(\mu)\,, \qquad \forall\, y \in \mathcal{D}_{\hat{r}_k}\,,\ k \in \mathcal{G}^n_1\,, \quad |k|_{_1} > \tau_*(\delta)\,. \tag{196}$$

By (170), (187) and (196) we have that for every $y \in \mathcal{D}\,,\ k \in \mathcal{G}^n_{1,K_1}$

$$\frac{\|G^k(y,\cdot) - T_1 F^k(y,\cdot)\|_2}{|f_k(y)|} \leqslant \frac{2e^{s+5}\,\vartheta\ \ + 2^9\, e^{-|k|_1 s}}{\delta_k(\mu)}\ . \tag{197}$$

Then, in order to prove (193), it is enough to show that

$$\frac{4e^{s+5}\,\vartheta}{\delta_k(\mu)} \leqslant \gamma\,, \qquad \frac{2^{10}\, e^{-|k|_1 s}}{\delta_k(\mu)} \leqslant \gamma\,. \tag{198}$$

Let us consider the first inequality in (198). Since by (191) and recalling (151) we have $\vartheta = \kappa/K_{_2}^{\tilde{n}}$, then, recalling (188) and (190), for $|k|_{_1} \leqslant K_{_1}$

$$\frac{4e^{s+5}\,\vartheta}{\delta_k(\mu)} \leqslant \frac{4e^{s+5}\kappa K_{_1}^n}{\delta K_{_2}^{\tilde{n}}}\left(\frac{K_{_1}^{n/2}}{\tilde{\mu}}\right)^{\log K_{_1}^n\delta^{-1}} = \frac{4e^{s+5}\kappa}{\delta}e^A\,,$$

where

$$A := \left(\frac{n}{2}\log K_{_1} + \log\frac{1}{\tilde{\mu}}\right)\left(n\log K_{_1} + \log\frac{1}{\delta}\right) + n\log K_{_1} - \tilde{n}\log K_{_2}\,.$$

Since

$$A \leqslant -\frac{\tilde{n}}{4}\log K_{_2}$$

by (192), we obtain that

$$\frac{4e^{s+5}\,\vartheta}{\delta_k(\mu)} \leqslant \frac{4e^{s+5}\kappa}{\delta}e^{-\frac{\tilde{n}}{4}\log K_{_2}} \overset{(192)}{\leqslant} \gamma\,,$$

[38] Note that by (164) $\hat{r}_k \leqslant r/2$.

proving the first estimate in (198).
Regarding the second inequality in (198) we have

$$\frac{2^{10}\, e^{-|k|_1 s}}{\delta_k(\mu)} = \frac{2^{10}}{\delta} e^{B}\,,$$

with

$$B := \left(\frac{n}{2}\log|k|_{1} + \log\frac{1}{\tilde\mu}\right)\left(n\log|k|_{1} + \log\frac{1}{\delta}\right) + n\log|k|_{1} - |k|_{1} s\,.$$

We note that

$$B \leqslant -\frac{1}{4}|k|_{1} s$$

by (185), indeed[39]

$$\frac{2}{s}n^2 \leqslant \frac{|k|_{1}}{\log^2|k|_{1}}\,, \qquad \frac{4n}{s}\left(\frac{1}{2}\log\frac{1}{\delta} + \log\frac{1}{\tilde\mu} + 1\right) \leqslant \frac{|k|_{1}}{\log|k|_{1}}\,.$$

Then

$$\frac{2^{10}}{\delta} e^{-\frac{1}{4}|k|_1 s} \overset{(185)}{\leqslant} \gamma\,.$$

This proves (198) and, therefore, completes the proof of (193).
Let us now show (194). By (159) and (196) we get

$$\begin{aligned}
\sup_{y\in(D^{1,k}\cap\mathcal{D})_{\hat r_k}} \frac{\| f^k_{**}(y,\cdot)\|_{s(1-1/K_2)/2}}{|f_k(y)|} &< \frac{2e^{|k|_1 s}}{\delta_k(\mu)} 2\Big(\frac{2nK_{2}}{s}\Big)^n e^{-(K_2-3)s/2}\\
&\overset{(190)}{=} \frac{2e^{|k|_1 s}|k|_{1}^n}{\delta}\left(\frac{30e^3|k|_{1}^{n/2}}{\mu}\right)^{\log(|k|_1^n/\delta)} 2\Big(\frac{2nK_{2}}{s}\Big)^n e^{-(K_2-3)s/2}\\
&\overset{(149)}{\leqslant} \frac{4e^{3s/2}n^n}{\delta s^n}\left(\frac{30e^3K_{2}^{n/2}}{3^{n/2}\mu}\right)^{\log(K_2^n/3^n\delta)} K_{2}^{2n} e^{-K_2 s/6}\\
&= \frac{4e^{3s/2}n^n}{\delta s^n} e^{-K_2 s/8} e^{-Q}
\end{aligned}$$

where

$$Q := \frac{1}{8}K_{2} s - \left(n\log\frac{K_{2}}{3} + \log\frac{1}{\delta}\right)\left(\frac{n}{2}\log\frac{K_{2}}{3} + \log\frac{1}{\mu} + \log 30 + 3\right) - 2n\log K_{2}\,.$$

[39] Note that $x/\log x \geqslant \alpha$ if $x \geqslant \alpha\log^2\alpha$ and $\alpha \geqslant 5$; analogously $x/\log^2 x \geqslant \alpha$ if $x \geqslant \alpha\log^3\alpha$ and $\alpha \geqslant 2^6$.

Then (194) follows if we prove that $Q \geqslant 0$. Recalling (149) we get[40]

$$Q \geqslant \frac{1}{8} K_{_2} s - 8n^2 \log^2 K_{_2} + 2 \log^2 \frac{1}{\delta\mu} \geqslant 0$$

by (192). ∎

We rewrite Theorem 7.1 in the fashion of Theorem 2.1.

Theorem 7.2 *Let $n \geqslant 2$, $0 < s \leqslant 1$. Fix $0 < \mu, \delta < 1/e^8$ and $0 < \gamma < 1$. Consider a Hamilonian $H_\varepsilon(y,x) = h(y) + \varepsilon f y, (x)$ as in (1) such that h satisfies the non–degeneracy Assumption **A** (§ 2) and f has norm one: $\|f\|_{D,r,s} = 1$. Assume that for some $y_0 \in D$ we have $f(y_0, \cdot) \in \mathcal{H}_{s,\tau_*}(\delta)$, with $\tau_* = \tau_*(\delta; \gamma, \mu)$ defined in (185). Let $K_{_2} \geqslant 3K_{_1} \geqslant 6$ with $K_{_1}$ satisfying (192) and (152). Let $\hat{r}_k$ as in (195) and $\mathcal{D}$ as in (189) . Finally assume that ε satisfies (40).*
Then, for any $k \in \mathcal{G}^n_{1,K_1}$ with $\tau_*(\delta; \gamma, \mu) \leqslant |k|_{_1} \leqslant K_{_1}$, there exists a symplectic change of variables Ψ_k as in (41) such that the following holds.
For every $y \in (D^{1,k} \cap \mathcal{D})_{\hat{r}_k}$ there exist a phase $\theta^{(k)}(y)$ and functions $\mathtt{G}^k(y, \cdot) \in \mathbb{B}^1_2$ and $\mathtt{f}^k(y, \cdot) \in \mathbb{B}^n_{s(1-1/K_2)^2}$ satisfying

$$\boxed{H_\varepsilon \circ \Psi_k =: h(y) + 2\varepsilon\, |f_k(y)| \Big(\cos(k \cdot x + \theta^{(k)}(y)) + \mathtt{G}^k(y, k \cdot x) + \mathtt{f}^k(y, x) \Big)} \tag{199}$$

with

$$\sup_{y \in (D^{1,k} \cap \mathcal{D})_{\hat{r}_k}} |\mathtt{G}^k(y, \cdot)|_2 \leqslant \gamma \tag{200}$$

and

$$\sup_{y \in (D^{1,k} \cap \mathcal{D})_{\hat{r}_k}} |\mathtt{f}^k|_{s(1-1/K_2)/2} \leqslant \frac{4e^{3s/2} n^n}{\delta s^n} e^{-K_2 s/8} \ . \tag{201}$$

Proof It directly follows from Theorem 7.1. We only note that $\theta^{(k)}(y)$ is defined such that

$$|f_k(y)| \cos(k \cdot x + \theta^{(k)}(y)) = T_1 F^k(y, x) \,,$$

while

$$\begin{aligned} \mathtt{G}^k(y, x) \ &:= \ \frac{G^k(y, \cdot) - T_1 F^k(y, \cdot)}{|f_k(y)|} \,, \\ \mathtt{f}^k(y, x) \ &:= \ \frac{f^k_{\star\star}(y, x)}{|f_k(y)|} \,, \end{aligned}$$

[40]Using that $\log^2 x \leqslant \sqrt{x}$ for $x \geqslant 2^{13}$.

Note that $\theta^{(k)}(y)$, $\mathtt{G}^k(y,x)$ and $\mathtt{f}^k(y,x)$ are not analytic in y (due to the presence of $|f_k(y)|$), but, obviously, $H_\varepsilon \circ \Psi_k$ is real–analytic in x *and* y. ∎

Acknowledgments We are grateful to V. Kaloshin and A. Sorrentino for useful discussions.

A An elementary result in linear algebra

Lemma A.1 *Given* $k \in \mathbb{Z}^n$, $k \neq 0$ *there exists a matrix* $A = (A_{ij})_{1\leqslant i,j\leqslant n}$ *with integer entries such that* $A_{nj} = k_j \ \forall\ 1 \leqslant j \leqslant n$, $\det A = d := \gcd(k_1, \ldots, k_n)$, *and* $|A|_\infty = |k|_\infty$.

Proof The argument is by induction over n. For $n = 1$ the lemma is obviously true. For $n = 2$, it follows at once from[41]

Bezout's Lemma *Given two integers* a *and* b *not both zero, there exist two integers* x *and* y *such that* $ax+by = d := \gcd(a,b)$, *and such that* $\max\{|x|, |y|\} \leqslant \max\{|a|/d, |b|/d\}$.

Indeed, if x and y are as in Bezout's Lemma with $a = k_1$ and $b = k_2$ one can take $A = \begin{pmatrix} y & -x \\ k_1 & k_2 \end{pmatrix}$. Now, assume, by induction for $n \geqslant 3$ that the claim holds true for $(n-1)$ and let us prove it for n. Let $\bar k = (k_1, ..., k_{n-1})$ and $\bar d = \gcd(k_1, ..., k_{n-1})$ and notice that $\gcd(\bar d, k_n) = d$. By the inductive assumption, there exists a matrix $\bar A = \begin{pmatrix} \tilde A \\ \bar k \end{pmatrix} \in \mathrm{Mat}_{(n-1)\times(n-1)}(\mathbb{Z})$ with $\tilde A \in \mathrm{Mat}_{(n-2)\times(n-1)}(\mathbb{Z})$, such that $\det \bar A = \bar d$ and $|\bar A|_\infty = |\bar k|_\infty$. Now, let x and y be as in Bezout's Lemma with $a = \bar d$, and $b = k_n$. We claim that A can be defined as follows:

$$A = \begin{pmatrix} \tilde k & \tilde x \\ \bar A & \begin{pmatrix} 0 \\ \vdots \\ 0 \\ k_n \end{pmatrix} \end{pmatrix}, \qquad \tilde k = (-1)^n y \, \frac{\bar k}{\bar d}, \qquad \tilde x := (-1)^{n+1} x . \tag{202}$$

[41] The first statement in this formulation of Bezout's Lemma is well known and it can be found in any textbook on elementary number theory; the estimates on x and y are easily deduced from the well known fact that given a solution x_0 and y_0 of the equation $ax + by = d$, all other solutions have the form $x = x_0 + k(b/d)$ and $y = y_0 - k(a/d)$ with $k \in \mathbb{Z}$ and by choosing k so as to minimize $|x|$.

First, observe that since $\bar{d}$ divides k_j for $j \leqslant (n-1)$, $\tilde{k} \in \mathbb{Z}^{n-1}$. Then, expanding the determinant of A from last column, we get

$$
\begin{aligned}
\det A &= (-1)^{n+1}\tilde{x} \det \bar{A} + k_n \det \begin{pmatrix} \tilde{k} \\ \tilde{A} \end{pmatrix} \\
&= (-1)^{n+1}\tilde{x}\, \bar{d} + k_n (-1)^{n-2} \det \begin{pmatrix} \tilde{A} \\ \tilde{k} \end{pmatrix} \\
&= (-1)^{n+1}\tilde{x}\, \bar{d} + k_n (-1)^{n-2}(-1)^n \frac{y}{\bar{d}} \det \bar{A} \\
&= x\bar{d} + k_n y = d \ .
\end{aligned}
$$

Finally, by Bezout's Lemma, we have that $\max\{|x|, |y|\} \leqslant \max\{\bar{d}/d, |k_n|/d\}$, so that

$$
|\tilde{k}|_\infty = |y| \frac{|\bar{k}|_\infty}{\bar{d}} \leqslant \frac{|\bar{k}|_\infty}{d} \leqslant |k|_\infty \ , \quad |\tilde{x}| = |x| \leqslant \frac{|k_n|}{d} \leqslant |k|_\infty \ ,
$$

which, together with $|\bar{A}|_\infty = |\bar{k}|_\infty$, shows that $|A|_\infty = |k|_\infty$. ∎

References

[1] V. I. Arnold. Instability of dynamical systems with several degrees of freedom. Dokl. Akad. Nauk SSSR 156, 9–12 (1964). Engl. transl.: Sov. Math., Dokl. 5, 581–585 (1964)

[2] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. Mathematical aspects of classical and celestial mechanics, *volume 3 of Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, third edition, 2006. [Dynamical systems. III], Translated from the Russian original by E. Khukhro.

[3] L. Biasco, and L. Chierchia. KAM Theory for secondary tori, arXiv:1702.06480v1 [math.DS]

[4] L. Biasco, and L. Chierchia. On the measure of Lagrangian invariant tori in nearly–integrable mechanical systems. Rend. Lincei Mat. Appl. **26** (2015), 1–10

[5] A. Delshams; P. Gutiérrez, Effective stability and KAM theory. J. Differential Equations 128 (1996), no. 2, 415–490

[6] A. Giorgilli; L. Galgani, Rigorous estimates for the series expansions of Hamiltonian perturbation theory. Celestial Mech. 37 (1985), no. 2, 95–112

[7] M. Guzzo, L. Chierchia and G. Benettin. The steep Nekhoroshev Theorem, Commun. Math. Phys. 342, 569-601 (2016)

[8] B. R. Hunt, V. Y. Kaloshin, Prevalence, chapter 2, Handbook in dynamical systems, edited by H. Broer, F. Takens, B. Hasselblatt, Vol. 3, 2010, pg. 43-87

[9] V. F. Lazutkin, Concerning a theorem of Moser on invariant curves (Russian), Vopr. Dinamich. Teor. Rasprostr. Seism. Voln. 14 (1974), 109-120.

[10] B. Y. Levin. Lectures on Entire Functions, American Mathematical Soc., 1996.

[11] A.G. Medvedev, A.I. Neishtadt, D.V. Treschev, Lagrangian tori near resonances of near–integrable Hamiltonian systems, Nonlinearity, **28**:7 (2015), 2105–2130

[12] A. I. Neishtadt, Estimates in the Kolmogorov theorem on conservation of condition- ally periodic motions, J. Appl. Math. Mech. 45 (1981), no. 6, 766-772

[13] N.N. Nekhoroshev. An exponential estimate of the time of stability of nearly- integrable Hamiltonian systems I, Uspekhi Mat. Nauk 32 (1977), 5–66; Russian Math. Surveys 32 (1977), 1–65

[14] J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, Comm. Pure Appl. Math., v. 35 (1982), no. 1, 653-695

[15] J. Pöschel, Nekhoroshev estimates for quasi–convex Hamiltonian systems. Math. Z. **213**, pag. 187 (1993).