Branching of Cantor manifolds of elliptic tori and applications to PDEs

Massimiliano Berti, Luca Biasco

Abstract. We consider infinite dimensional Hamiltonian systems. First we prove the existence of "Cantor manifolds" of elliptic tori -of any finite higher dimension- accumulating on a given elliptic KAM torus. Then, close to an elliptic equilibrium, we show the existence of Cantor manifolds of elliptic tori which are "branching" points of other Cantor manifolds of higher dimensional tori. We also provide a positive answer to a conjecture of Bourgain [8] proving the existence of invariant elliptic KAM tori with tangential frequency constrained to a fixed Diophantine direction. These results are obtained under the natural nonresonance and nondegeneracy conditions. As applications we prove the existence of new kinds of quasi periodic solutions of the one dimensional nonlinear wave equation. The proofs are based on averaging normal forms and a sharp KAM theorem, whose advantages are an explicit characterisation of the Cantor set of parameters, quite convenient for measure estimates, and weaker smallness conditions on the perturbation.

Contents

1 Introduction

A central topic in the theory of Hamiltonian partial differential equations (PDEs) concerns the existence of quasi-periodic solutions. In the last years several existence results have been proved using both KAM theory, see e.g. Wayne [28], Kuksin [22], Pöschel [25], [23], Eliasson-Kuksin [16] (and references therein), or Newton-Nash-Moser implicit function techniques, see e.g. Craig-Wayne [14], Bourgain [8]-[10], Berti-Bolle [5] and with Procesi [6]. We mention also to the recent approach with Lindstedt series by Gentile-Procesi [18]. An advantage of the KAM approach is to provide not only the existence of an invariant torus but also a normal form around it. This would allow, in principle, to study the dynamics of the PDE in its neighbourhood.

All the existing literature considers quasi-periodic solutions of PDEs in a neighbourhood of an elliptic equilibrium, see for a survey Kuksin [22], Craig [13], or perturbations of finite gap solutions of integrable PDEs, see Kuksin [22], Kappeler-Pöschel [20].

In this paper we want to study the dynamics of infinite dimensional Hamiltonian systems near an elliptic torus, developing, in particular, an abstract KAM theory for proving the existence of "Cantor manifolds" of elliptic invariant tori tangent to a given elliptic torus.

For finite dimensional Hamiltonian systems, the dynamics close to a lagrangian KAM torus has been deeply investigated by Giorgilli-Morbidelli [19], proving, in particular, the existence of invariant tori with asymptotic density exponentially close to 1. On the other hand the existence of lower dimensional tori in a neighbourhood of an elliptic torus requires, also in finite dimension, a more refined KAM theorem (it is a corollary of our general results). The difficulty comes from the presence of the elliptic directions.

Our first result states, roughly, the following (see Theorem 2.1 for a precise statement):

Given a finite dimensional torus with an elliptic KAM normal form around it, we prove, under the natural non-resonance and non-degeneracy assumptions, the existence of "Cantor manifolds" of elliptic tori -of any finite higher dimension- accumulating on it.

This result is based on two main steps. We first perform a Birkhoff normalisation (see the "averaging" Proposition 6.1) assuming the weakest, natural, non-resonance conditions on the tangential and normal frequencies of the torus (see (2.12)). These are similar to those used in Bambusi [1], Bambusi-Grébert [3], for an elliptic equilibrium. The next step is to apply some KAM Theorem. Due to the third order monomials on the high mode variables in (6.3)-(6.4), the KAM theorems available in the literature would apply only requiring stronger non-resonance assumptions, see Remark 2.2. Then we use the sharper KAM Theorem 5.1. Note that these refined estimates are required only for small amplitude solutions and not for perturbations of linear PDEs as considered in [21], [22], [28] where the size of the perturbation is an external parameter.

As a second result we prove an abstract theorem describing a branching phenomenon of Cantor manifolds of elliptic tori of increasing dimension (see Theorem 3.1 for a precise statement):

Close to an elliptic equilibrium there exist, under the natural non-resonance and non-degeneracy assumptions, Cantor manifolds of elliptic tori which are "branching points" of other Cantor manifolds of higher dimensional tori.

This result relies on an application of Theorem 2.1. The main difficulty is to check that, after the first application of the KAM theorem close to the equilibrium, the perturbed frequencies of the deformed elliptic torus, fulfil the non-resonance conditions required in Theorem 2.1. This is achieved in section 7, thanks to the explicit form of the Cantor set of non-resonant parameters provided by the basic KAM Theorem 5.1.

Theorem 3.1 can be also seen as a "building block" for constructing small amplitude almost periodic solutions for PDEs without external parameters. Actually, with the present estimates, we can prove the existence of only finitely many branches of finite dimensional elliptic tori. The existence of almost periodic solutions has been proved in Pöschel [26] with a similar scheme, for the nonlinear Schrödinger equation, with potentials as *external* parameters, and adding a regularising nonlinearity.

These abstract results, valid for infinite dimensional Hamiltonian systems, can be applied to Hamiltonian PDEs like Schrödinger, beam and wave equations. For concreteness we focus on the nonlinear wave equation (NLW). Moreover NLW is more difficult for KAM theory than the Schrödinger and beam equation for the weaker asymptotic growth of the frequencies. As an application of Theorem 3.1 we show in Theorem 4.1 *the existence of a new kind of quasi-periodic solutions of*

$$\begin{cases} u_{tt} - u_{xx} + \mathrm{m}u + f(u) = 0 \\ u(t,0) = u(t,\pi) = 0 \end{cases} \tag{1.1}$$

for almost all the masses $\mathrm{m} > 0$ and for real analytic, odd, nonlinearities of the form

$$f(u) = \sum_{k\geq 3,\mathrm{odd}} a_k u^k \,, \qquad a_3 \neq 0 \,, \tag{1.2}$$

These quasi-periodic solutions are different from the ones of [25] since they accumulate to a torus and not to the origin.

As already said, a basic tool for the previous results is an application of the sharp KAM Theorem of section 5. Its main advantages are:

- *(i) the KAM smallness condition are weaker than in [24], see comments after KAM Theorem 5.1.*

This is achieved by a modification of the iterative scheme of [22], [24], as described in section 5.

- *(ii) The final Cantor set of parameters, satisfying the Melnikov non-resonance conditions for all the KAM iterative steps, is completely explicit in terms of the final frequencies only, see (5.13).*

A new aspect of Theorem 5.1 is the *complete* separation between the iterative scheme for the construction of invariant tori and the existence of enough non-resonant frequencies at every step of the iterative process, see [5] for a similar construction in the Nash-Moser setting. In previous KAM theorems the Cantor set of non-resonant parameters is known "a posteriori" ([23]). The key point here is that the final frequencies are always well defined also if the iterative KAM process stops after finitely many steps (and so there are no invariant tori for any value of the parameters). The present formulation simplifies considerably the necessary measure estimates, see, as applications, Theorems 5.2, 5.3, and section 7.1. The characterisation in (5.13) of the Cantor set in terms of the final frequencies only is new also for finite dimensional elliptic tori; for lagrangian tori in finite dimension see [12],[11]. It simplifies also the measure estimates of degenerate KAM theory, see for example [4] for an extension to PDEs. In particular it allows to avoid the notions of "links" and "chains" used in [27]. Actually, thanks to the explicit characterisation of the Cantor set (5.13) we are able to answer positively to a conjecture by Bourgain in [8], proving

- *the existence of elliptic invariant KAM tori with tangential frequency constrained to a fixed Diophantine direction, see Theorem 3.2; for the application to NLW equation (1.1) see Theorem 4.2.*

This kind of results was proved for finite dimensional Hamiltonian systems by Eliasson [15] and Bourgain [8] who raised the question if a similar result can be achieved also for infinite dimensional Hamiltonian systems. For a result for NLW in this direction see [17].

We hope that the results and techniques of this paper will be used to develop a more general description of the dynamics of the PDE close to an elliptic torus, proving, for example, stability results as in Bambusi [1], Bambusi-Grébert [3].

Before presenting precisely our results, we introduce the functional setting and the main notations concerning infinite dimensional Hamiltonian systems.

Acknowledgments: We thank Michela Procesi for useful comments.

Functional setting and notations

Phase space. We consider the Hilbert space of complex-valued sequences

$$\ell^{a,p} := \Big\{ z = (z_1, z_2, \dots) \ : \ \|z\|^2_{a,p} := \sum_{j\geq 1} |z_j|^2 j^{2p} e^{2ja} < +\infty \Big\}$$

with $a > 0$, $p > 1/2$, and the toroidal phase space

$$(x, y, w) \in \mathbb{T}^n_s \times \mathbb{C}^n \times \ell^{a,p}_b \,, \qquad w := (z, \bar z) \in \ell^{a,p}_b := \ell^{a,p} \times \ell^{a,p} \,,$$

where $\mathbb{T}^n_s$ is the complex open s-neighbourhood of the n-torus $\mathbb{T}^n := \mathbb{R}^n/(2\pi\mathbb{Z})^n$.

Hamiltonian system. Given $H : \mathbb{T}^n_s \times \mathbb{C}^n \times \ell^{a,p}_b \to \mathbb{C}$ we consider the Hamiltonian system

$$(\dot{x}, \dot{y}, \dot{w}) = X_H(x, y, w) \tag{1.3}$$

with Hamiltonian vector field

$$X_H := (\partial_y H, -\partial_x H, -\mathrm{i}J\partial_w H) \in \mathbb{C}^n \times \mathbb{C}^n \times \ell^{a,p}_b$$

where

$$J := \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} : \ell^{a,p} \times \ell^{a,p} \to \ell^{a,p} \times \ell^{a,p}\,.$$

Given two functions $H, F : \mathbb{T}^n_s \times \mathbb{C}^n \times \ell^{a,p}_b \to \mathbb{C}$ we define their Poisson bracket

$$\{H, F\} := \partial_x H \cdot \partial_y F - \partial_y H \cdot \partial_x F - \mathrm{i}J\partial_w F \cdot \partial_w H\,. \tag{1.4}$$

Analytic functions. Given a complex Banach space E, we consider analytic functions

$$f : D(s, r) \times \Pi \to E \tag{1.5}$$

possibly depending on parameters $\xi \in \Pi \subset \mathbb{R}^m$ defined on the open neighbourhood of the origin

$$D(s, r) := \Big\{ |\mathrm{Im}\, x| < s\,, |y| < r^2\,, \|w\|_{a,p} < r \Big\} \subset \mathbb{T}^n_s \times \mathbb{C}^n \times \ell^{a,p}_b\,,\ 0 < s, r < 1\,,$$

where $|y| := \sup_{j=1,\dots,n} |y_j|$. We define the sup-norm

$$|f|_{s,r} := |f|_{s,r,\Pi,E} := \sup_{(x,y,w;\xi) \in D(s,r)\times\Pi} \|f(x, y, w; \xi)\|_E\,. \tag{1.6}$$

We denote simply by $|\cdot|_s$ the sup-norm of functions independent of (y, w).

Any analytic function can be developed in a totally convergent power series

$$P(x, y, w; \xi) = \sum_{i,j\geq 0} P_{ij}(x; \xi) y^i w^j$$

where

$$P_{ij}(x) := P_{ij}(x; \xi) \in \mathcal{L}\Big(\overbrace{\mathbb{C}^n \times \ldots \times \mathbb{C}^n}^{i-times} \times \overbrace{\ell^{a,p}_b \times \ldots \times \ell^{a,p}_b}^{j-times}, \mathbb{C} \Big) \tag{1.7}$$

are multilinear, symmetric, bounded maps. For simplicity of notation, we will often omit the explicit dependence on ξ. By the Riesz Representation Theorem, we identify the 1-forms $P_{10}(x) \in (\mathbb{C}^n)^*$, resp. $P_{01}(x) \in (\ell^{a,p})^*$, with vectors $P_{10}(x) \in \mathbb{C}^n$, resp. $P_{01}(x) \in \ell^{a,p}_b$, writing

$$P_{10}(x)y = P_{10}(x) \cdot y\,, \quad \text{resp.} \quad P_{01}(x)w = P_{01}(x) \cdot w\,,$$

where "$\cdot$" denotes the scalar product on $\mathbb{C}^n$, resp. $\ell^{a,p}_b$. Moreover we identify as usual the bilinear symmetric form $P_{02}(x) \in \mathcal{L}(\ell^{a,p}_b \times \ell^{a,p}_b, \mathbb{C})$ with the operator $P_{02}(x) \in \mathcal{L}(\ell^{a,p}_b, \ell^{a,p}_b)$ defined by

$$P_{02}(x)w^2 = P_{02}(x)w \cdot w\,, \quad \forall w \in \ell^{a,p}_b\,.$$

We define

$$P_{\leq 2} := P_{00} + P_{01}w + P_{10}y + P_{02}w \cdot w\,. \tag{1.8}$$

In general we identify the P_{ij} in (1.7) with the vector valued multilinear forms, for $j \geq 1$,

$$P_{ij}(x)\,,\ \partial^i_y \partial^j_w P(x, y, w) \in \mathcal{L}\Big(\overbrace{\mathbb{C}^n \times \ldots \times \mathbb{C}^n}^{i-times} \times \overbrace{\ell^{a,p}_b \times \ldots \times \ell^{a,p}_b}^{(j-1)-times}, \ell^{a,p}_b \Big)\,. \tag{1.9}$$

If the Hamiltonian vector field maps $X_P : D(s,r) \to \mathbb{C}^n \times \mathbb{C}^n \times \ell_b^{a,\bar p}$ with $\bar p \geq p$, then, for $j \geq 1$,

$$|P_{ij}|_s = \sup_{(x;\xi)\in \mathbb{T}_s\times\Pi} \|P_{ij}(x;\xi)\| < \infty$$

where $\|\cdot\|$ denotes the operatorial norm on $\mathcal{L}(\overbrace{\mathbb{C}^n \times \ldots \times \mathbb{C}^n}^{i-times} \times \overbrace{\ell_b^{a,p} \times \ldots \times \ell_b^{a,p}}^{(j-1)-times}, \ell_b^{a,\bar p})$.

The [ ]-operator. We define the operator $[\cdot]$ acting on monomials $Q := q(x) y^i z^a \bar z^{\bar a}$, $i, a, \bar a \in \mathbb{N}^\infty$, by

$$[Q] := \begin{cases} \langle Q \rangle = \langle q \rangle y^i z^a \bar z^{\bar a} & \text{if } \ a = \bar a \\ 0 & \text{otherwise} \end{cases} \tag{1.10}$$

where $\langle q \rangle := (2\pi)^{-n} \int_{\mathbb{T}^n} q(x)dx$ denotes the average with respect to the angles.

Lipschitz norms. Given a function f as in (1.5) we define the Lipschitz semi-norm

$$|f|^{\rm lip}_{s,r} = \sup_{\xi,\zeta\in\Pi,\xi\neq\zeta} \frac{|f(\cdot;\xi) - f(\cdot;\zeta)|_{s,r}}{|\xi - \zeta|} \tag{1.11}$$

and, given $\lambda \geq 0$, the Lipschitz norm

$$|\cdot|^\lambda_{r,s} := |\cdot|_{r,s} + \lambda |\cdot|^{\rm lip}_{r,s}\,. \tag{1.12}$$

We will always use the symbol "λ" in this role, not to be confused with exponentiation. We denote the Lipschitz norm of functions independent of (y,w) more simply by $|\cdot|^\lambda_s$.

Miscellanea. Given $l \in \mathbb{Z}^\infty$ we define

$$|l| := \sum_{j\geq1} |l_j|\,, \quad |l|_p := \sum_{j\geq1} j^p |l_j|\,, \quad \langle l \rangle_d := \max\Big(1, \Big|\sum_{j\geq1} j^d l_j\Big|\Big)$$

and the unit versors $e_j := (0, \ldots, 0, 1, 0, \ldots)$ with zero components except the j-th one. We define the space

$$\ell_\infty^{-\delta} := \Big\{\Omega := (\Omega_1, \Omega_2, \ldots),\, \Omega_j \in \mathbb{R} \ : \ |\Omega|_{-\delta} := \sup_{j\geq1} j^{-\delta} |\Omega_j| < +\infty\Big\}$$

and the Lipschitz norm

$$|\Omega|^\lambda_{-\delta} := \sup_{\xi\in\Pi} |\Omega(\xi)|_{-\delta} + \lambda |\Omega|^{\rm lip}_{-\delta} \quad \text{where} \quad |\Omega|^{\rm lip}_{-\delta} := \sup_{\xi,\zeta\in\Pi,\xi\neq\zeta} \frac{|\Omega(\xi) - \Omega(\zeta)|_{-\delta}}{|\xi - \zeta|}\,. \tag{1.13}$$

Finally, for $\tau > n-1$, $\eta > 0$, we define the set of Diophantine vectors

$$\mathcal{D}_{\eta,\tau} := \Big\{\omega \in \mathbb{R}^n \ : \ |\omega \cdot k| \geq \frac{\eta}{1 + |k|^\tau}\,, \ \forall\, k \in \mathbb{Z}^n \setminus \{0\}\Big\}\,. \tag{1.14}$$

2 Cantor manifolds of tori close to an elliptic torus

The KAM-normal form Hamiltonian

$$H = H(x,y,z,\bar z) = N + P = \omega \cdot y + \Omega \cdot z\bar z + \sum_{2i+j\geq3} P_{ij}(x) y^i w^j \tag{2.1}$$

possesses the elliptic invariant torus

$$\mathcal{T}_0 = \mathbb{T}^n \times \{0\} \times \{0\} \times \{0\} \tag{2.2}$$

with tangential and normal frequencies $\omega := (\omega_1, \ldots, \omega_n) \in \mathbb{R}^n$, $\Omega := (\Omega_{n+1}, \ldots)$ respectively. In (2.1) the variables are $w = (z, \bar z)$ with $z = (z_{n+1}, \ldots)$. We assume

- Frequency asymptotics. The $\Omega_j \in \mathbb{R}$ and there exists $d \geq 1$ such that

$$\Omega_j = j^d + \dots, \qquad j \geq 1, \tag{2.3}$$

where the dots stand for lower order terms in j.

If $d = 1$ we denote by κ the largest positive number such that

$$\frac{\Omega_i - \Omega_j}{i - j} = 1 + O(j^{-\kappa}), \ \forall i > j, \quad \text{and} \quad \mu := \begin{cases} 1 & \text{if } \ d > 1 \\ \kappa/(\kappa+1) & \text{if } \ d = 1. \end{cases} \tag{2.4}$$

- Regularity. The vector field X_P is real analytic and

$$X_P : D(s, r) \to \mathbb{C}^n \times \mathbb{C}^n \times \ell_b^{a,\bar{p}} \quad \text{with} \quad \begin{cases} \bar{p} \geq p & \text{if } \ d > 1 \\ \bar{p} > p & \text{if } \ d = 1. \end{cases} \tag{2.5}$$

We aim to prove the existence of finite dimensional elliptic tori of *any* arbitrary dimension $\hat{n} \geq n$ accumulating onto the elliptic torus $\mathcal{T}_0$. We denote the augmented frequencies

$$\hat{\omega} := (\omega_1, \dots, \omega_n, \Omega_{n+1}, \dots, \Omega_{\hat{n}}) \in \mathbb{R}^{\hat{n}}, \quad \hat{\Omega} := (\Omega_{\hat{n}+1}, \dots), \tag{2.6}$$

the coordinates

$$z = (\tilde{z}, \hat{z}), \ \tilde{z} := (z_{n+1}, \dots, z_{\hat{n}}), \ \hat{z} := (z_{\hat{n}+1}, \dots), \ w = (\tilde{w}, \hat{w}), \ \tilde{w} = (\tilde{z}, \bar{\tilde{z}}), \ \hat{w} = (\hat{z}, \bar{\hat{z}}),$$

and the actions

$$\hat{y} := (y, \tilde{y}), \quad \tilde{y} := \frac{1}{2}(z_{n+1}\bar{z}_{n+1}, \dots, z_{\hat{n}}\bar{z}_{\hat{n}}), \quad \hat{Z} = \frac{1}{2}(z_{\hat{n}+1}\bar{z}_{\hat{n}+1}, \dots).$$

We decompose any $l = (l_{n+1}, \dots) \in \mathbb{Z}^\infty$ as

$$l = (\tilde{l}, \hat{l}) \quad \text{with} \quad \tilde{l} := (l_{n+1}, \dots, l_{\hat{n}}), \quad \hat{l} := (l_{\hat{n}+1}, \dots). \tag{2.7}$$

Given P_{ij} (see (1.7)) we define the coefficients $P_{i\tilde{\jmath}\hat{\jmath}}$, for $\tilde{\jmath}, \hat{\jmath} \in \mathbb{N}$ with $\tilde{\jmath} + \hat{\jmath} = j$, by the relation

$$P_{ij} y^i w^j = \sum_{\tilde{\jmath}+\hat{\jmath}=j} P_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}.$$

We introduce the symmetric $\hat{n}$-dimensional "twist" matrix

$$\hat{A} \in \mathrm{Mat}(\hat{n} \times \hat{n}), \quad \hat{A} := \begin{pmatrix} 2[P_{200}] & [P_{120}] \\ & \\ [P_{120}] & 2[P_{040}] \end{pmatrix} \tag{2.8}$$

where the matrices $[P_{200}]$, $[P_{040}]$, $[P_{120}]$ are defined by[1]

$$[P_{200}]y \cdot y := [P_{200} y^2], \qquad [P_{040}]\tilde{y} \cdot \tilde{y} := [P_{040} \tilde{w}^4], \qquad [P_{120}]y \cdot \tilde{y} := [P_{120} y \tilde{w}^2] \tag{2.9}$$

and the [ ] operator in (1.10). We also define $[P_{102}]$, $[P_{022}]$, by

$$[P_{102}]y \cdot \hat{Z} := [P_{102} y \hat{w}^2], \quad [P_{022}]\tilde{y} \cdot \hat{Z} := [P_{022} \tilde{w}^2 \hat{w}^2]$$

and

$$\hat{B} := \begin{pmatrix} [P_{102}] & [P_{022}] \end{pmatrix} \in \mathcal{L}(\mathbb{C}^{\hat{n}}, \ell_\infty^{\bar{p}-p}), \tag{2.10}$$

the last property being valid thanks to the regularizing property (2.5). We set

$$\tau := \begin{cases} 2(d-1)^{-1} + n + 1 & \text{if } d > 1 \\ (n+2)(\delta_* - 1)\delta_*^{-1} + 1 & \text{if } d = 1 \end{cases} \tag{2.11}$$

with δ_* fixed below.

[1] The matrices $[P_{200}] \in \mathrm{Mat}(n \times n)$, $[P_{040}] \in \mathrm{Mat}\big((\hat{n} - n) \times (\hat{n} - n)\big)$, $[P_{120}] \in \mathrm{Mat}\big((\hat{n} - n) \times n\big)$. Similarly $[P_{102}] \in \mathrm{Mat}(\infty \times n)$, $[P_{022}] \in \mathrm{Mat}\big(\infty \times (\hat{n} - n)\big)$.

Theorem 2.1. (Higher dimensional tori close to an elliptic torus) *Consider an Hamiltonian H as in (2.1) satisfying (2.3), (2.5), and, if $d = 1$, $\mu > 9/14$ (see (2.4)). Fix $\hat{n} \geq n$. There exists a constant $c > 0$ such that, if the following assumptions hold:*

- **(Melnikov conditions)** *For some $\alpha > 0$,*

$$|\omega \cdot k + \Omega \cdot l| \geq \alpha \frac{\langle l \rangle_d}{1 + |k|^\tau}, \quad \forall k \in \mathbb{Z}^n, \ l = (\tilde{l}, \hat{l}) \in \Lambda_{\hat{n},D}, \ (k,l) \neq 0, \tag{2.12}$$

  where τ is defined in (2.11) with $\delta_* = p - \bar{p}$, and

$$\Lambda_{\hat{n},D} := \Big\{ |l| \leq D, |\hat{l}| \leq 2 \Big\} \cup \Big\{ |\tilde{l}| = D, |\hat{l}| = 1 \Big\}, \qquad D := \begin{cases} 4 & \text{if } d > 1 \\ 6 & \text{if } d = 1 . \end{cases}$$

- **(Twist)** *$\hat{A}$ is invertible.*

- **(Non-resonance)** $\forall\, 0 < |\hat{l}| \leq 2$ *there hold*

$$\big(\hat{\Omega} - \hat{B}\hat{A}^{-1}\hat{\omega}\big) \cdot \hat{l} \neq 0 . \tag{2.13}$$

- **(Smallness)** *The third order terms satisfy*

$$(|P_{11}|_s + |P_{03}|_s)^2 \leq c\alpha, \tag{2.14}$$

then there exists an $\hat{n}$-dimensional Cantor manifold of real analytic, elliptic, diophantine $\hat{n}$-dimensional tori accumulating onto the n-dimensional elliptic torus $\mathcal{T}_0$.

The above Cantor manifold has the same geometric structure described in [23]. The constant c depends on $n, \tau, s, d, A, B, \hat{n}, \hat{\omega}, \hat{\Omega}, \hat{A}, \hat{B}$.

Remark 2.1. *By (2.3), (2.4) and the regularizing property (2.10) of $\hat{B}$, (2.13) implies*

$$\inf_{0 < |\hat{l}| \leq 2} |(\hat{\Omega} - \hat{B}\hat{A}^{-1}\hat{\omega}) \cdot \hat{l}| > 0 .$$

Indeed $|(\hat{\Omega} - \hat{B}\hat{A}^{-1}\hat{\omega}) \cdot \hat{l}| \geq 1/2$ up to a finite subset of $\{0 < |\hat{l}| \leq 2\}$.

The proof of Theorem 2.1 is based on two main steps. The former is the "averaging" Proposition 6.1 in which we use the Melnikov conditions (2.12), that are similar to those used in [1]-[3] close to an elliptic equilibrium. The latter is an application of the basic KAM Theorem 5.1, case-(H2).

Remark 2.2. *Condition (H2) of Theorem 5.1 is strictly weaker than the KAM condition in [24] (see comments after Theorem 5.1) and applies under the natural Melnikov conditions (2.12). The KAM Theorem [24] would require the stronger Melnikov conditions (2.12) with $D = 6$ for $d > 1$ and $D = 7$ for $d = 1$ and $\mu = 2/3$ (as for NLW, see (4.5)). See also Remarks 6.1 and 6.3.*

3 Branching of Cantor manifolds of elliptic tori

We consider an Hamiltonian

$$H = \Lambda + Q + R \tag{3.1}$$

where R is a higher order perturbation of an integrable normal form $\Lambda + Q$. In complex coordinates $(\zeta, \bar{\zeta})$ and, setting

$$I := \frac{1}{2}(\zeta_1 \bar{\zeta}_1, \ldots, \zeta_n \bar{\zeta}_n), \qquad Z := \frac{1}{2}(\zeta_{n+1}\bar{\zeta}_{n+1}, \ldots),$$

the normal form consists of the terms

$$\Lambda := \mathrm{a}\cdot I + \mathrm{b}\cdot Z\,, \quad Q := \frac{1}{2}\mathrm{A}I\cdot I + \mathrm{B}I\cdot Z \tag{3.2}$$

where a, b and A, B denote, respectively, vectors and matrices with constant coefficients. Fixed $\hat{n} \geq n$, we assume that:

(A) The normal form $\Lambda + Q$ is nondegenerate in the following sense:

TWIST. (A$_1$) $\det\mathrm{A} \neq 0$

NONRESONANCE.

(A$_2$) $\mathrm{b}\cdot l \neq 0\,, \quad \forall\, 1 \leq |l| \leq 2$

(A$_3$) $\mathrm{a}\cdot k + \mathrm{b}\cdot l \neq 0 \quad or \quad \mathrm{A}k + \mathrm{B}^\intercal l \neq 0\,,\ \forall\, k \in \mathbb{Z}^n\,,\ l \in \Lambda_{\hat{n},D}\,,\ (k,l) \neq 0\,.$

Moreover, if $d = 1$, $\mathrm{a}\cdot k + \mathrm{b}\cdot(\tilde{l},0) \pm h \neq 0 \quad or \quad \mathrm{A}k + \mathrm{B}^\intercal(\tilde{l},0) \neq 0\,,$
$\forall\, 0 < |k| \leq K_0\,,\ |\tilde{l}| \leq D-2\,,\ \ 1 \leq h \leq L_0 + \hat{n}(D-2)\,.$

The constants K_0, L_0 depend only on $d, D, \mathrm{a}, \mathrm{b}, \mathrm{A}, \mathrm{B}$, see (7.34).

(B) FREQUENCY ASYMPTOTICS. There is $d \geq 1$ and $\delta_* < d-1$ such that $\mathrm{b}_j = j^d + \ldots + O(j^{\delta_*})$.

(C) REGULARITY. The vector fields X_Q, X_R are real analytic from some neighbourhood of the origin of $\ell_b^{a,p}$ into $\ell_b^{a,\bar{p}}$ with $\bar{p} \geq p$ defined in (2.5). By increasing δ_*, if necessary, we may also assume

$$p - \bar{p} \leq \delta_* < d-1\,. \tag{3.3}$$

Concerning the higher order perturbation R we assume

$$|R| = O(\|z\|_{a,p}^4) + O(\|\zeta\|_{a,p}^{\mathrm{g}})\,, \quad z := (\zeta_{n+1}, \zeta_{n+2}, \ldots)\,, \quad \mathrm{g} > 1 + 3\mu^{-1}\,, \quad \mu \in (9/14, 1]\,, \tag{3.4}$$

where μ is defined as in (2.4) and, for $d = 1$, κ is the largest positive constant such that

$$\left|\frac{\mathrm{b}_i - \mathrm{b}_j}{i-j} - 1\right| \leq a_* j^{-\kappa}\,, \quad \forall\, i > j \tag{3.5}$$

for some $a_* > 0$. For $d = 1$, by increasing δ_*, if necessary, we can assume $-\delta_* < \kappa$.

Fix $\hat{n} \geq n$. We define the augmented frequency vectors

$$\hat{\mathrm{a}} := (\mathrm{a}, \mathrm{b}_{n+1}, \ldots, \mathrm{b}_{\hat{n}}) \in \mathbb{R}^{\hat{n}}\,, \quad \hat{\mathrm{b}} := (\mathrm{b}_{\hat{n}+1}, \mathrm{b}_{\hat{n}+2} \ldots)\,, \tag{3.6}$$

the symmetric "twist" matrix

$$\hat{\mathrm{A}} \in \mathrm{Mat}(\hat{n}\times\hat{n})\,, \quad \hat{\mathrm{A}}_{ij} := \begin{cases} \mathrm{A}_{ij} & \text{if } i,j \leq n \\ \mathrm{B}_{ij} & \text{if } j \leq n < i \leq \hat{n} \\ \langle \partial^4_{\zeta_i\bar{\zeta}_i\zeta_j\bar{\zeta}_j} R_{|\zeta=\bar{\zeta}=0}\rangle & \text{if } n < i,j \leq \hat{n} \end{cases} \tag{3.7}$$

and

$$\hat{\mathrm{B}} \in \mathrm{Mat}(\hat{n}\times\infty)\,, \quad \hat{\mathrm{B}}_{ij} := \begin{cases} \mathrm{B}_{ij} & \text{if } j \leq n < i \\ \langle \partial^4_{\zeta_i\bar{\zeta}_i\zeta_j\bar{\zeta}_j} R_{|\zeta=\bar{\zeta}=0}\rangle & \text{if } n < j \leq \hat{n} < i. \end{cases} \tag{3.8}$$

($\hat{\mathrm{A}}$) We assume

TWIST. ($\hat{\mathrm{A}}_1$) $\det\hat{\mathrm{A}} \neq 0$

NONRESONANCE.

($\hat{\mathrm{A}}_2$) $\mathrm{b}\cdot l \neq 0\,, \quad \forall\, l \in \Lambda_{\hat{n},D}\,,$ where $\Lambda_{\hat{n},D}$ is defined in (2.12).
Moreover, if $d = 1$, $\inf_{l\in\Lambda_{\hat{n},D}} |\mathrm{b}\cdot l| > 0\,.$

($\hat{\mathrm{A}}_3$) $(\hat{\mathrm{b}} - \hat{\mathrm{B}}\hat{\mathrm{A}}^{-1}\hat{\mathrm{a}})\cdot\hat{l} \neq 0\,, \forall\, \hat{l} = (l_{\hat{n}+1}, l_{\hat{n}+2}, \ldots) \quad$ with $\quad |\hat{l}| = 1, 2\,.$

Clearly $(\hat{\mathrm{A}}_2)$ is stronger than (A_2).

Theorem 3.1. (Branching of Cantor manifolds of elliptic tori) *Fix $\hat{n} \geq n$. Suppose $H = \Lambda + Q + R$ satisfies assumptions* (A),(B),(C), $(\hat{\mathrm{A}})$ *and (3.4). Then*

- **(i)** *there exists an n-dimensional Cantor manifold of real analytic, elliptic, diophantine, invariant n-dimensional tori.*
- **(ii)** *Each of these n-dimensional elliptic tori possesses another Cantor manifold of real analytic, elliptic, diophantine $\hat{n}$-dimensional tori, which is tangent to the torus with asymptotically full density.*

The new result is clearly (ii). Part (i) was proved in Kuksin-Pöschel in [23].

We prove Theorem 3.1 as follows. After a Birkhoff normal form step, we introduce the actions as parameters, and, applying Theorem 5.1-(H3), we find a Cantor manifold of n-dimensional tori close to the origin with asymptotically full density (part (i)). For proving part (i) we only require

$$(\mathrm{A}_1), (\mathrm{A}_2), (\mathrm{B}), (\mathrm{C}), (3.4) \;\text{ and }\; \mathrm{a} \cdot k + \mathrm{b} \cdot l \neq 0 \;\text{ or }\; \mathrm{A}k + \mathrm{B}^\intercal l \neq 0\,, \ \forall\, k \in \mathbb{Z}^n\,,\ |l| \leq 2\,,\ (k,l) \neq 0\,, \tag{3.9}$$

as in [23]. In order to prove part (ii) the crucial point is to show that, thanks to assumptions (A_3) and $(\hat{\mathrm{A}})$, it is possible to take the parameters still in a set of asymptotically full measure, such that the hypotheses of Theorem 2.1 hold. This is verified in subsection 7.1, strongly exploiting the explicit form of the Cantor set Π_∞ in (5.13) proved in the basic KAM Theorem 5.1.

Another minor advantage of the application of the improved KAM Theorem 5.1 is the following. Since condition (H3) is strictly weaker, when $d = 1$, than the KAM condition in [24] (see comments after Theorem 5.1), Theorem 5.1 simultaneously applies to both cases $d > 1$ and $d = 1$.

Actually we can also improve the result of Theorem 3.1-(i) proving the existence of elliptic tori with tangential frequency restricted to a fixed Diophantine direction, extending to infinite dimensional systems the results of Bourgain [8] and Eliasson [15].

Theorem 3.2. *Assume* $(\mathrm{A}_1), (\mathrm{A}_2), (\mathrm{B}), (\mathrm{C})$, *(3.4),* $\mathrm{a} \neq 0$ *and* $(\mathrm{b} - \mathrm{B}\mathrm{A}^{-1}\mathrm{a}) \cdot l \neq 0$, $\forall\, 1 \leq |l| \leq 2$. *Then if $\bar{\omega} \in \mathcal{D}_{\alpha_0,\tau}$ (see (1.14)) with $\alpha_0 := \rho_0^{1+c}$, $\rho_0 := |\bar{\omega} - \mathrm{a}| > 0$, and $c > 0$ is small enough, then*

$$|\mathcal{T}|(2c\rho_0)^{-1} \;\to\; 1 \quad \text{as} \quad \rho_0 \to 0\,, \tag{3.10}$$

where $\mathcal{T} \subset [1 - c\rho_0, 1 + c\rho_0]$ are the t such that $t\bar{\omega}$ is the tangential frequency of a n-dimensional torus found in Theorem 3.1-(i).

Note that the hypotheses of Theorem 3.2 imply (3.9).

4 Application to nonlinear wave equation

Now we apply the results of section 3 to the NLW. We first write the NLW equation (1.1) as an infinite dimensional Hamiltonian system introducing coordinates $\mathrm{q}, \mathrm{p} \in \ell^{a,p}$, $a > 0$, $p > 1/2$, setting

$$u = \sum_{j \geq 1} \frac{\mathrm{q}_j}{\sqrt{\lambda_j}} \phi_j\,,\ v = u_t = \sum_{j \geq 1} \mathrm{p}_j \sqrt{\lambda_j} \phi_j \quad \text{where} \quad \lambda_j := \sqrt{j^2 + \mathrm{m}}\,,\ \phi_j := \sqrt{2/\pi}\, \sin(jx)\,.$$

The Hamiltonian of NLW is

$$H_{NLW} = \int_0^\pi \Big(\frac{v^2}{2} + \frac{u_x^2}{2} + \mathrm{m}\frac{u^2}{2} + g(u) \Big)\, dx = \Lambda + G = \frac{1}{2} \sum_{j \geq 1} \lambda_j (\mathrm{q}_j^2 + \mathrm{p}_j^2) + G(\mathrm{q})\,,$$

where

$$g(s) := \int_0^s f(t) dt\,, \quad G(\mathrm{q}) := \int_0^\pi g\Big(\sum_{j \geq 1} \mathrm{q}_j \lambda_j^{-1/2} \phi_j \Big)\, dx\,.$$

For $1 \le n \le \hat n$ we choose arbitrarily the "tangential sites"

$$\mathcal{I} := \{i_1, \ldots, i_n\} \subseteq \hat{\mathcal{I}} := \{i_1, \ldots, i_n, i_{n+1}, \ldots, i_{\hat n}\} \subset \mathbb{N}^+ . \tag{4.1}$$

By [25] there is a symplectic map transforming H_{NLW} in its partial Birkhoff normal form on the $\hat{\mathcal{I}}$-modes

$$H = \Lambda + \bar G + \check G + K$$

where $X_{\bar G}, X_{\check G}, X_K$ are analytic from some neighborhood of the origin in $\ell^{a,p}$ into $\ell^{a,p+1}$,

$$\bar G = \frac{1}{2} \sum_{i \, or \, j \in \hat{\mathcal{I}}} \bar G_{ij} \mathrm{z}_i \bar{\mathrm{z}}_i \mathrm{z}_j \bar{\mathrm{z}}_j \, , \ \bar G_{ij} := \frac{6}{\pi} \frac{4 - \delta_{ij}}{\lambda_i \lambda_j} \, , \ \mathrm{z}_j = \frac{1}{\sqrt 2}(\mathrm{q}_j + \mathrm{i} \mathrm{p}_j) \, , \ \bar{\mathrm{z}}_j = \frac{1}{\sqrt 2}(\mathrm{q}_j - \mathrm{i} \mathrm{p}_j) \, ,$$

$\check G$ is of order four and depends only on z_i, $i \notin \hat{\mathcal{I}}$, K is of order six and depends on all the variables z_i, $i \in \mathbb{N}$ (for more details we refer to [25] or [7]).

In order to write H in the form (3.1) we renumber the indexes in such a way that the first n modes correspond to the $\mathcal{I}$-modes and the first $\hat n$ modes to the $\hat{\mathcal{I}}$-modes. More precisely we construct a re-ordering $\mathbb{N}^+ \to \mathbb{N}^+$, $j \mapsto i_j$ which is bijective and increasing from $\{1, \ldots, n\}$ onto $\mathcal{I}$, from $\{n+1, \ldots, \hat n\}$ onto $\hat{\mathcal{I}} \setminus \mathcal{I}$ and from $\mathbb{N}^+ \setminus \{1, \ldots, \hat n\}$ onto $\mathbb{N}^+ \setminus \hat{\mathcal{I}}$. Calling the variables

$$\zeta_j := \mathrm{z}_{i_j} \, , \quad \forall \, j \ge 1 \, ,$$

the Hamiltonian H assumes the form (3.1)-(3.2) with

$$\mathrm{a} := (\lambda_{i_1}, \ldots, \lambda_{i_n}) \, , \ \mathrm{b} := (\lambda_{i_{n+1}}, \ldots) \, , \ \mathrm{A} := (\bar G_{i_h i_k})_{1 \le h,k \le n} \, , \ \mathrm{B} := (\bar G_{i_h i_k})_{1 \le k \le n < h} \, , \tag{4.2}$$

and

$$R := \frac{1}{2} \sum_{h \, or \, k \le \hat n, \, h,k > n} \bar G_{i_h i_k} \zeta_h \bar\zeta_h \zeta_k \bar\zeta_k + \check G + K \, . \tag{4.3}$$

Let us verify the hypotheses of Theorem 3.1. By [25] the matrix A in (4.2) is invertible, actually

$$(\mathrm{A}^{-1})_{hk} = \frac{\pi}{6} \left(\frac{4}{4n-1} - \delta_{hk} \right) \mathrm{a}_h \mathrm{a}_k \, , \quad 1 \le h, k \le n \, . \tag{4.4}$$

Then (A_1) holds. Assumption (A_2) holds because the frequencies λ_j are simple and non zero. Still in [25] it is verified that (B), (C) are satisfied with

$$d = 1 \, , \quad \delta_* = -1 \, , \quad \bar p = p + 1 \, ,$$

as well as (3.4) with (see (4.2) and (3.5))

$$\mathrm{g} = 6 \, , \quad \mu = 2/3 > 9/14 \, , \quad \kappa = 2 \, . \tag{4.5}$$

Assumptions (A_3) (which is new with respect to [25]) will be a corollary of the next lemma.

Lemma 4.1. $\forall 0 < |l| < \infty$, *the function* $f_l : (0, \infty) \to \mathbb{R}$, $f_l(\mathrm{m}) := (\mathrm{b} - \mathrm{B}\mathrm{A}^{-1}\mathrm{a}) \cdot l$ *is analytic and non constant.*

Proof. By (4.2) and (4.4) we get $(\mathrm{B}\mathrm{A}^{-1})_{ij} = 4 \mathrm{a}_j \mathrm{b}_i^{-1} / (4n-1)$ and

$$f_l(\mathrm{m}) = \sum_{j > n} l_j \mathrm{b}_j^{-1} (\alpha \mathrm{m} + \beta + i_j^2) \quad \text{with} \quad \alpha := \frac{-1}{4n-1} \, , \quad \beta := -\frac{4 \sum_{1 \le j \le n} i_j^2}{4n-1} \, .$$

Let $j_* := \max\{j > n \, : \, l_j \neq 0\}$ and $i_* := \max\{i_j \, : \, l_j \neq 0\}$. For $m > i_*^2$ we expand the analytic functions $\mathrm{b}_j(\mathrm{m})^{-1}$ in power series

$$\mathrm{b}_j^{-1} = \frac{1}{\sqrt{\mathrm{m}}} \sum_{k \ge 0} c_k \left(i_j^2 / \mathrm{m} \right)^k \quad \text{with} \quad c_0 := 1 \, , \ c_k := -\frac{1}{2} \left(-\frac{1}{2} - 1 \right) \cdots \left(-\frac{1}{2} - k + 1 \right) / k! \neq 0 \, .$$

Then

$$f_l(\mathrm{m}) = \alpha\sqrt{\mathrm{m}} \sum_{n<j\le j_*} l_j + \frac{1}{\sqrt{\mathrm{m}}} \sum_{k\ge 0} c_k p_k \mathrm{m}^{-k} \quad \text{where} \quad p_k := \sum_{n<j\le j_*} l_j i_j^{2k} q_{ki_j}$$

and $q_{ki} := L_i + \dfrac{\alpha i^2}{2(k+1)}$, $L_i := (1-\alpha)i^2 + \beta$. We prove that $f_l(\mathrm{m})$ is not constant showing that $p_k \ne 0$ for k large enough. Note that $|q_{ki_*}| \ge 1/k^2$ for k large enough: if $L_{i_*} \ne 0$ then $q_{ki_*} \to L_{i_*} \ne 0$, otherwise $|q_{ki_*}| = |\alpha i_*^2 (2k+2)^{-1}| \ge 1/k^2$ for k large. Moreover $|q_{ki_j}| \le 2i_*^2$, $\forall\, k$. Hence

$$|p_k| \ge i_*^{2k} |q_{ki_*}| - (i_*-1)^{2k} \sum_{n<j\le j_*} |l_j||q_{ki_j}| \ge i_*^{2k} k^{-2} - |l|(i_*-1)^{2k} 2i_*^2 \to \infty$$

as $k \to \infty$. ■

Corollary 4.1. *Assumption* (A_3) *is satisfied with the exception of a countable set of* m*'s in* $(0,\infty)$.

Proof. If $l \in \Lambda_{\hat n, D}$ and $\mathrm{A}k + \mathrm{B}^\intercal l = 0$, then $\mathrm{a}\cdot k + \mathrm{b}\cdot l = (\mathrm{b} - \mathrm{B}\mathrm{A}^{-1}\mathrm{a})\cdot l \ne 0$ except at most countably many m's. Analogously, if $\mathrm{A}k + \mathrm{B}^\intercal(\tilde l, 0) = 0$, then $\mathrm{a}\cdot k + \mathrm{b}\cdot(\tilde l, 0) \pm h = (\mathrm{b} - \mathrm{B}\mathrm{A}^{-1}\mathrm{a})\cdot(\tilde l, 0) \pm h \not\equiv 0$. ■

The last condition of Theorem 3.1 to verify is $(\hat{\mathrm{A}})$, where $\hat{\mathrm{a}}$, $\hat{\mathrm{b}}$, $\hat{\mathrm{A}}$, $\hat{\mathrm{B}}$, defined in (3.6), (3.7), (3.8), are like a, b, A, B in (4.2) changing $\hat n$ with n. Then $(\hat{\mathrm{A}}_1)$ holds as well as $(\hat{\mathrm{A}}_3)$, except countably many m. Finally, assumption $(\hat{\mathrm{A}}_2)$ holds for almost every $\mathrm{m} \in (0,\infty)$ as a consequence of Theorem 3.12 of [3] (see also Theorem 6.5 of [1]). More precisely $\inf_{l\in\Lambda_{\hat n, D}} |\mathrm{b}(\mathrm{m})\cdot l| > 0$ is a consequence of the nonresonance condition (r-NR) of [3] with $r = D+2$, $N = \hat n$. Then Theorem 3.1 applies.

Theorem 4.1. *Suppose f is real analytic and (1.2) holds. Fix $\hat n \ge n$. For all the choices of indices $\mathcal{I}$, $\hat{\mathcal{I}}$ as in (4.1), for almost all the masses* m *the conclusions (i)-(ii) of Theorem 3.1 apply to the NLW equation (1.1).*

Conclusion (i) was proved in Pöschel [25] for all $\mathrm{m} \in \mathbb{R}$ with the restriction $\min_{1\le j<n} i_{j+1} - i_j \le n-1$. On the other hand, the quasi-periodic solutions obtained in (ii) are new, since they accumulate onto a n-torus and not at the origin. They are not the $\hat n$-dimensional tori bifurcating from the fourth order Birkhoff normal form of (1.1).

As a consequence of Corollary 4.1 we can prove the existence of quasi-periodic solutions with tangential frequency restricted to a fixed direction, see [17] for a similar result.

Theorem 4.2. *Suppose that f is real analytic and (1.2) holds. Then, excluding a countable set of masses* $\mathrm{m} \in (0,\infty)$*, the conclusion of Theorem 3.2 applies to the NLW equation (1.1).*

5 A sharp basic KAM theorem

We consider a family of integrable Hamiltonians

$$N := N(x, y, z, \bar z; \xi) := e(\xi) + \omega(\xi)\cdot y + \Omega(\xi)\cdot z\bar z \tag{5.1}$$

defined on $\mathbb{T}^n_s \times \mathbb{C}^n \times \ell^{a,p} \times \ell^{a,p}$. The frequencies $\omega = (\omega_1, \ldots, \omega_n)$ and $\Omega = (\Omega_{n+1}, \Omega_{n+2}, \ldots)$ depend on m-parameters

$$\xi \in \Pi \subset \mathbb{R}^m\,, \quad m \le n\,, \quad \Pi \text{ compact with positive Lebesgue measure, } \rho := \mathrm{diam}(\Pi)\,.$$

For each ξ there is an invariant n-torus $\mathcal{T}_0 = \mathbb{T}^n \times \{0\} \times \{0\} \times \{0\}$ with frequency $\omega(\xi)$. In its normal space, the origin $(z, \bar z) = 0$ is an elliptic fixed point with proper frequencies $\Omega(\xi)$. The aim is to prove the persistence of a large portion of this family of linearly stable tori under small analytic perturbations $H = N + P$.

We assume

(A*) Parameter dependence. The map $\omega : \Pi \to \mathbb{R}^n$, $\xi \mapsto \omega(\xi)$, is Lipschitz continuous.

(B*) Frequency asymptotics. There exist $d \geq 1$ and $\delta_* < d-1$ such that

$$\Omega_i(\xi) = \bar{\Omega}_i + \Omega_i^*(\xi) \in \mathbb{R}\,, \quad i \geq 1\,,$$

where $\bar{\Omega}_i = i^d + \ldots$ and $\Omega^* : \Pi \to \ell_\infty^{-\delta_*}$ is Lipschitz continuous.

By (A*) and (B*), the Lipschitz semi-norms (defined as in (1.11)) of the frequency maps satisfy

$$|\omega|^{\mathrm{lip}} + |\Omega|^{\mathrm{lip}}_{-\delta_*} \leq M < +\infty\,. \tag{5.2}$$

(C*) Regularity. The perturbation P is real analytic in the space coordinates, Lipschitz in the parameters, and for every $\xi \in \Pi$ the hamiltonian vector field maps $X_P : D(s,r) \to \mathbb{C}^n \times \mathbb{C}^n \times \ell_b^{a,\bar{p}}$ with $\bar{p}$ satisfying (2.5). More precisely, using the notations (1.6), (1.11), we assume

$$|X_P|_{r,s,E,\Pi} + |X_P|^{\mathrm{lip}}_{r,s,E,\Pi} < +\infty \quad \text{where} \quad E := \mathbb{C}^n \times \mathbb{C}^n \times \ell_b^{a,\bar{p}}\,. \tag{5.3}$$

Moreover, we also assume (3.3).

We introduce the *group* (under composition) of maps

$$\begin{aligned}\mathcal{E}_s \;\; := \;\; \Big\{ & \Psi : (x_+, y_+, w_+; \xi) \mapsto (x, y, w) \;\; \text{of the form} \\ & x = x_{00}(x_+;\xi)\,,\;\; w = w_{00}(x_+;\xi) + w_{01}(x_+;\xi) w_+\,, \\ & y = y_{00}(x_+;\xi) + y_{01}(x_+;\xi) w_+ + y_{10}(x_+;\xi) y_+ + y_{02}(x_+;\xi) w_+ \cdot w_+\,, \\ & \text{where } x_{00}, y_{ij}, w_{ij} \text{ are analytic and bounded on } \mathbb{T}^n_s \text{ and Lipschitz on } \Pi \Big\}.\end{aligned} \tag{5.4}$$

In Theorem 5.1 the symplectic map yielding the KAM normal form (5.8) has the form $\Phi = I + \Psi$ with Ψ like in (5.4), as in [24]. It will be the composition of infinitely many time-1-flow maps (each having the form $I + \Psi$, $\Psi \in \mathcal{E}_s$) generated by Hamiltonians in $\mathcal{F}_s$ defined in (8.7).

Theorem 5.1. (Sharp basic KAM theorem) *Suppose that $H = N + P$ satisfies assumptions* (A*), (B*), (C*). *Let $\alpha > 0$ be a parameter and assume that*

$$\Theta := \max\Big\{1, |P_{11}|_s^\lambda, |P_{03}|_s^\lambda, \sum_{2i+j=4} |\partial_y^i \partial_w^j P|_{s,r}^\lambda\Big\} \;\; \text{with} \;\; \lambda = \frac{\alpha}{M} \;\; \text{satisfies} \;\; \Theta \leq \frac{\sqrt{\alpha}}{3r}\,. \tag{5.5}$$

Then there is $\gamma := \gamma(n,\tau,s) > 0$ such that, if one of the following KAM-conditions

- *(H1)* $\varepsilon_1 := \max\Big\{\frac{|P_{00}|_s^\lambda}{r^2\alpha^2}, \frac{|P_{01}|_s^\lambda}{r\alpha^{3/2}}, \frac{|P_{10}|_s^\lambda}{\alpha}, \frac{|P_{02}|_s^\lambda}{\alpha}\Big\} \leq \gamma\,,$
- *(H2)* $\varepsilon_2 := \max\Big\{\frac{|P_{00}|_s^\lambda}{r^2\alpha^{5/4}}, \frac{|P_{01}|_s^\lambda}{r\alpha^{3/2}}, \frac{|P_{10}|_s^\lambda}{\alpha}, \frac{|P_{02}|_s^\lambda}{\alpha}\Big\} \leq \gamma$ *and* $|P_{11}|_s^\lambda \leq \frac{\alpha^{5/4}}{r}\,,$
- *(H3)* $\varepsilon_3 := \max\Big\{\frac{|P_{00}|_s^\lambda}{r^2\alpha^{\mu}}, \frac{|P_{01}|_s^\lambda}{r\alpha}, \frac{|P_{10}|_s^\lambda}{\alpha}, \frac{|P_{02}|_s^\lambda}{\alpha}\Big\} \leq \gamma$ *and* $|P_{11}|_s^\lambda\,,\; |P_{03}|_s^\lambda \leq \frac{\alpha}{r}$
  where $\mu := 1$ if $d > 1$ and $0 < \mu \leq 1$ if $d = 1$,

holds, then there exist:

- **(Frequencies)** *Lipschitz functions $\omega_\infty : \Pi \to \mathbb{R}^n$, $\Omega_\infty : \Pi \to \ell_\infty^{-d}$, satisfying*

$$|\omega_\infty - \omega|^\lambda\,,\; |\Omega_\infty - \Omega|^\lambda_{\bar{p}-p} \leq \gamma^{-1}\alpha\varepsilon_i \tag{5.6}$$

and $|\omega_\infty|^{\mathrm{lip}}$, $|\Omega_\infty|^{\mathrm{lip}}_{-\delta_*} \leq 2M$.

- **(KAM normal form)** *A Lipschitz family of analytic symplectic maps*

$$\Phi : D(s/4, r/4) \times \mathbf{\Pi}_\infty \ni (x_\infty, y_\infty, w_\infty; \xi) \mapsto (x, y, w) \in D(s, r) \tag{5.7}$$

of the form $\Phi = \mathrm{I} + \Psi$ *with* $\Psi \in \mathcal{E}_{s/4}$*, where* $\mathbf{\Pi}_\infty$ *is defined in (5.12), such that,*

$$H^\infty(\cdot;\xi) := H \circ \Phi(\cdot;\xi) = \omega_\infty(\xi) y_\infty + \Omega_\infty(\xi) z_\infty \bar{z}_\infty + P^\infty \quad \textit{has} \quad P^\infty_{\leq 2} = 0 \tag{5.8}$$

see (1.8). Moreover

$$\begin{cases} |P^\infty_{11} - P_{11}|_{s/4} \leq \gamma^{-1} \varepsilon_i (|P_{11}|_s + \alpha^{p_a - 1/2}) \\ |P^\infty_{03} - P_{03}|_{s/4} \leq \gamma^{-1} \varepsilon_i (|P_{03}|_s + |P_{11}|_s + \alpha^{p_a - 1/2})\,. \end{cases} \tag{5.9}$$

- **(Smallness estimates)** *The map* Ψ *satisfies*

$$|x_{00}|^\lambda_{s/4}, |y_{00}|^\lambda_{s/4} \frac{\alpha^{1-p_a}}{r^2}, |y_{01}|^\lambda_{s/4} \frac{\alpha^{1-p_b}}{r}, |y_{10}|^\lambda_{s/4}, |y_{02}|^\lambda_{s/4}, |w_{01}|^\lambda_{s/4}, |w_{00}|^\lambda_{s/4} \frac{\alpha^{1-p_b}}{r} \leq \gamma^{-1} \varepsilon_i \tag{5.10}$$

according $(Hi)_{i=1,2,3}$ *holds, where*

$$p_a := \begin{cases} 2 & \text{if } \textit{(H1)} \\ 5/4 & \text{if } \textit{(H2)} \\ 1 & \text{if } \textit{(H3)} \end{cases} \qquad p_b := \begin{cases} 3/2 & \text{if } \textit{(H1)} \text{ or } \textit{(H2)} \\ 1 & \text{if } \textit{(H3)}. \end{cases} \tag{5.11}$$

- **(Cantor set)** *The Cantor set is explicitely*

$$\mathbf{\Pi}_\infty := \begin{cases} \Pi_\infty & \text{if } (H1) \text{ or } (H2) \text{ or } (H3) - (d > 1) \\ \Pi_\infty \cap \omega^{-1}(\mathcal{D}_{\alpha^\mu, \tau}) & \text{if } (H3) - (d = 1) \end{cases} \tag{5.12}$$

where $\mathcal{D}_{\alpha^\mu,\tau}$ *is defined in (1.14) with* $\eta = \alpha^\mu$*, and*

$$\Pi_\infty := \Big\{ \xi \in \Pi \ : \ |\omega_\infty(\xi) \cdot k + \Omega_\infty(\xi) \cdot l| \geq 2\alpha \frac{\langle l \rangle_d}{1 + |k|^\tau} \ , \ \forall (k, l) \in \mathbb{Z}^n \times \mathbb{Z}^\infty \setminus \{0\}\,, |l| \leq 2 \Big\}\,. \tag{5.13}$$

Then, $\forall \xi \in \mathbf{\Pi}_\infty$*, the map* $x_\infty \mapsto \Phi(x_\infty, 0, 0; \xi)$ *is a real analytic embedding of an elliptic, diophantine,* n*-dimensional torus with frequency* $\omega_\infty(\xi)$ *for the system with Hamiltonian* H*, see (1.3).*

Note that (5.8) is the KAM normal form in an *open* neighborhood of the invariant elliptic torus. Regarding the smallness conditions we note that:
- In (H1) we make assumptions only on P_{00}, P_{01}, P_{10}, P_{02}. This is quite natural because, if they vanish, then the torus $\mathcal{T}_0$ in (2.2) is yet invariant, elliptic, and in normal form.
- In (H2) we relax the smallness assumption on P_{00}, at the expense of a smallness condition on P_{11}. Note that in (H2) we do not require any assumption on P_{03}. We apply (H2) looking for tori in a neighborhood of a fixed torus (where, in general, P_{03} does not vanish), see the proof of Theorem 2.1.
- In (H3) we further relax the smallness assumptions on P_{00} and P_{01}, at the expense of stronger conditions on P_{11} and P_{03}. We apply (H3) looking for tori close to an elliptic equilibrium (where, after a Birkhoff normal form, both P_{11} and P_{03} are small), see the proof of Theorem 3.1.

Comparison with the KAM Theorem [24]. The KAM condition in [24] on X_P in (2.5) is

$$\alpha^{-1} |X_P|^\lambda_{r,s} \leq \textit{const} \quad \text{with} \quad \lambda = \alpha/M\,, \tag{5.14}$$

where $|X_P|^\lambda_{r,s} := |X_P|_{r,s,E,\Pi} + \lambda |X_P|^{\mathrm{lip}}_{r,s,E,\Pi}$ is defined in (1.12) and

$$E := \Big\{ (x, y, w) \in \mathbb{C}^n \times \mathbb{C}^n \times \ell^{a,p}_b \text{ with norm } |(x, y, w)|_r := |x| + r^{-2}|y| + r^{-1}\|w\|_{a,\bar{p}} \Big\}\,.$$

We note that (5.14) implies the KAM condition (H3). For example, since $P_{10} = (\partial_y P)(x,0,0)$, we deduce, by (5.14), that $|P_{10}|_s^\lambda \leq const\, \alpha$. Similarly (5.14) implies all the other conditions in (H3). In the *case $d=1$ condition (H3) is strictly weaker than (5.14)*, since $\mu \leq 1$. This is why, we prove the result of [25] for NLW (where $\mu = 2/3$), avoiding the use of theorem D in [24] (see Theorem 4.1 and the proof of Theorem 3.1-(i)).

On the other hand the KAM conditions (H1)-(H2) are quite different than (5.14). The iterative scheme in [22], [24] would not converge assuming only (H1) or (H2). We discuss below the differences of the KAM iteration process used to prove Theorem 3.1.

Finally note that, if $|P_{03}|_s^\lambda = O(1)$, then (5.14) implies $\alpha \geq const\, r$. This causes difficulties for verifying the measure estimates because, as $r \to 0$, also the size of the parameters domain shrinks to zero, see remark 6.3.

The KAM Theorem 5.1 is completed by the following remarks.

Remark 5.1. (Analytic case) *If the Hamiltonian H is analytic in $\xi \in \Pi$ with $\Pi \subset \mathbb{C}^m$ we can prove the existence of limit-frequency maps $\xi \mapsto (\omega_\infty(\xi), \Omega_\infty(\xi))$ that are of class C^∞ and, $\forall q \geq 1$,*

$$|\omega_\infty - \omega|_{C^q}\,, \ |\Omega_\infty - \Omega|_{\bar{p}-p, C^q} \leq C(q)\varepsilon_0 \alpha^{1-q}\,. \tag{5.15}$$

See remark 8.1. Moreover in the KAM conditions (H1)-(H3) we can substitute $|P_{ij}|_s^\lambda$ with $|P_{ij}|_s$ thanks to Cauchy estimates.

Remark 5.2. (Lipeomorphism) *If $\omega : \Pi \to \omega(\Pi)$ is a homeomorphism which is Lipschitz in both directions (Lipeomorphism), with*

$$|\omega^{-1}|^{\rm lip} \leq L \quad and \quad \varepsilon_i \leq \frac{\gamma}{2LM}\,, \tag{5.16}$$

then $\omega_\infty : \Pi \to \omega_\infty(\Pi)$ is a Lipeomorphism with $|\omega_\infty^{-1}|^{\rm lip} \leq 2L$.

Remark 5.3. (Dependence on n) *The constant γ depends on the dimension n of the torus like, for example, $\gamma = \tilde{\tau}^{-c\tilde{\tau}}$ where $\tilde{\tau} := (\tau+n)\ln\big((\tau+n)/s\big)$ and $c>0$ is an absolute constant, see Remark 8.2. We have not tried to improve such super-exponential estimate to get larger values of γ.*

Let us briefly comment on the assumptions of Theorem 5.1.

Remark 5.4. *The condition $\Theta \geq 1$ in (5.5) is not restrictive because, rescaling the variables*

$$y \to \rho^2 y\,, \ w \to \rho w\,, \ H \to \rho^{-2} H\,, \tag{5.17}$$

we can always verify $\max\{|P_{11}|_s, |P_{03}|_s, \sum_{2i+j=4} |\partial_y^i \partial_w^j P|_{s,r}\} \geq 1$. *On the other hand note that the KAM conditions (H1)-(H3) are invariant under the above rescaling.*

Remark 5.5. *The KAM condition (H3) is obtained, for $d=1$, performing a normal form step before the KAM iteration, see section 8.4. Such condition is used for the wave equation. Note that if $\mu \to 0$ the condition (H3) improves, but, on the contrary, the measure $|\mathcal{D}_{\alpha^\mu,\tau}|$ decreases (see (1.14)-(5.12)).*

The scheme of proof of Theorem 5.1 is different than in [24]. In order to find the symplectic map Φ which transforms the Hamiltonian H into the KAM normal form $H_\infty := H \circ \Phi$ in (5.8), i.e.

$$P_{\leq 2}^\infty := P_{00}^\infty + P_{01}^\infty w + P_{10}^\infty y + P_{02}^\infty w \cdot w \equiv 0\,,$$

we perform infinitely many symplectic maps Φ_ν, $\nu \geq 1$, as in [24]. Each Hamiltonian has the form

$$H^\nu = N^\nu + P^\nu \qquad \text{where} \qquad N^\nu = \omega_\nu(\xi)\cdot y + \Omega_\nu(\xi)\cdot z\bar{z} \tag{5.18}$$

and P^ν is analytic on $D(s_\nu, r_\nu)$ with $r_\nu > r_0/4 > 0$ for all $\nu \geq 0$. It is natural to look at the map

$$(P_{00}^\nu, P_{01}^\nu, P_{10}^\nu, P_{02}^\nu) \quad \mapsto \quad (P_{00}^{\nu+1}, P_{01}^{\nu+1}, P_{10}^{\nu+1}, P_{02}^{\nu+1})$$

after any KAM step. An explicit calculus shows that the new $P^{\nu+1}_{\leq 2}$ is *not* a quadratic function of $P^{\nu}_{\leq 2}$: in the terms $(P^{\nu+1}_{10}, P^{\nu+1}_{02})$ there are linear combinations of $P^{\nu}_{00}, P^{\nu}_{01}$, see Lemma 8.13, with coefficients P^{ν}_{11}, P^{ν}_{03}, P^{ν}_{12}, P^{ν}_{20}. These terms come from the transformation of the cubic and quartic terms of P^{ν} under Φ^{ν}. However, after three iterations, the map

$$(P^{\nu}_{00}, P^{\nu}_{01}, P^{\nu}_{10}, P^{\nu}_{02}) \quad \mapsto \quad (P^{\nu+3}_{00}, P^{\nu+3}_{01}, P^{\nu+3}_{10}, P^{\nu+3}_{02})$$

turns out to be quadratic, see Lemma 8.16. Then the superexponential convergence of the iterative process is guaranteed under the smallness conditions $(H1)$-$(H3)$ on the initial $P_{00}, P_{01}, P_{10}, P_{02}$, where α and r occur with different weights. Note that the exponents of r come from the natural rescaling (5.17), while the different exponents of α by explicit computations. Unlike the usual KAM scheme in [22], [21], [24], the KAM normal form H^{∞} converges directly on an open neighborhood of the torus.

Note that also the KAM iterative scheme in [24] is not quadratic, see, for example formula (13) in [24]. This problem is solved letting the domain of the normal form shrink to zero (see also [21]), so that at the end of the iteration the normal form converges on the KAM torus only. The convergence on an open neighborhood of the torus is then recovered by a posteriori arguments.

The Cantor set Π_{∞}

Note that the Cantor set Π_{∞} in (5.13) depends *only* on the final frequencies $(\omega_{\infty}, \Omega_{\infty})$. It could be empty. In such a case the iterative process stops after finitely many steps and no invariant torus survives for any value of the parameters. However ω_{∞}, Ω_{∞}, and so Π_{∞}, are always well defined.

The idea is as follows. Each KAM step can be performed only for the parameters ξ such that the frequencies $\omega_{\nu}(\xi)$, $\Omega_{\nu}(\xi)$, satisfy the second order Melnikov non-resonance conditions (8.42). Actually this set could be empty. However we can always extend the frequency maps $\omega_{\nu}(\xi)$, $\Omega_{\nu}(\xi)$, to the whole set of parameters $\xi \in \Pi$, see the iterative Lemma 8.17-$(S2)_{\nu}$. This extension is Lipschitz continuous and, if the Hamiltonian is analytic, it is C^{∞}, see remark 8.1. Finally we verify in Lemma 8.19 that if ξ belongs to the Cantor set Π_{∞} then all the Melnikov non-resonance conditions required to perform the previous KAM step are all satisfied. We exploit that $(\omega_{\nu}, \Omega_{\nu})$ converge to $(\omega_{\infty}, \Omega_{\infty})$ superexponentially fast.

Note that we do *not* claim that the frequencies of the final invariant torus satisfy the second order Melnikov non-resonance conditions, fact already proved in [24]. We state a stronger claim, namely that *if* the parameter ξ is in Π_{∞} then the torus is preserved.

The number of parameters m in Theorem 5.1 is arbitrary. It could be strictly less than n (degenerate KAM theory). In the PDE applications of this paper we have $m = n$ and the frequency map is a Lipeomorphism. In such a case the final frequency ω_{∞} is a Lipeomorphism too, see remark 5.2. Then the following measure estimate follows by classical arguments [21], [22], [24], [20] (see also subsection 7.1).

Let κ be the largest number such that (2.4) holds uniformly on Π and set μ as in (2.4).

Theorem 5.2. (Measure estimate I) *Let $\omega : \Pi \to \omega(\Pi)$ be a Lipeomorphism and (5.16) hold. If*

$$\Omega(\xi) \cdot l \neq 0 \,, \quad \forall\, |l| = 1, 2 \,, \quad \forall\, \xi \in \Pi \,, \tag{5.19}$$

$$|\{\xi \in \Pi \ : \ \omega(\xi) \cdot k + \Omega(\xi) \cdot l = 0\}| = 0 \,, \quad \forall k \in \mathbb{Z}^n ,\, l \in \mathbb{Z}^{\infty} \,, |l| \leq 2 \,, \quad (k, l) \neq 0 \,, \tag{5.20}$$

then, taking τ as in (2.11), $|\Pi \setminus \mathbf{\Pi}_{\infty}| \to 0$ as $\alpha \to 0$. If, moreover, $\omega(\xi)$, $\Omega(\xi)$ are affine functions of ξ

$$|\Pi \setminus \mathbf{\Pi}_{\infty}| \leq C \rho^{n-1} \alpha^{\mu} \qquad \text{where} \qquad \rho := \mathrm{diam}(\Pi) \,. \tag{5.21}$$

The following theorem states that, given a Diophantine versor $\bar{\omega}$, there exist many invariant elliptic KAM tori with tangential frequency $t\bar{\omega}$, $t \in \mathbb{R}^+$.

Theorem 5.3. (Measure estimate II) *Assume that $\omega(\xi)$, $\Omega(\xi)$ are affine functions of ξ, $\partial_\xi\omega$ is invertible, and*

$$\Big(\Omega - \partial_\xi\Omega(\partial_\xi\omega)^{-1}\omega\Big)_{|\xi=0} \cdot l \neq 0\,, \quad \forall\, 0 < |l| \leq 2\,. \tag{5.22}$$

Suppose that $0 \notin \omega(\Pi)$. If γ defined in Theorem 5.1 is small enough, there exists $K > 1$ such that for every versor $\bar\omega \in \mathcal{D}_{K\alpha,\tau}$,

$$|\omega_\infty(\Pi \setminus \Pi_\infty) \cap \bar\omega\mathbb{R}^+| \leq K\alpha^\mu \tag{5.23}$$

(here $|\cdot|$ denotes the one dimensional Lebesgue measure).

Condition (5.22) is similar to condition (2) of [15] where it is required for $0 < |l| \leq 3$ (see also (2.13) with $\hat n = n$). By Fubini theorem (5.23) implies (5.21) integrating along the directions $\bar\omega$.

6 Proof of Theorem 2.1

We have

$$\frac12 \hat A\hat y\cdot\hat y = \Big[\sum_{2i+\tilde\jmath=4} P_{i\tilde\jmath 0} y^i \tilde w^{\tilde\jmath}\Big] \quad \text{and} \quad \hat B\hat y\cdot\hat Z = \Big[\sum_{2i+\tilde\jmath=2} P_{i\tilde\jmath 2} y^i \tilde w^{\tilde\jmath}\hat w^2\Big]\,.$$

Proposition 6.1. (Averaging) *Let H be as in (2.1). Suppose that (2.12) holds. Then there exists a constant $C := C(n,\tau,s,d,\hat n) > 1$ large enough, $0 < r_+ < r/4$ small enough and a symplectic map*

$$\Phi : (x_+, y_+, w_+) \in D(s_+, r_+) \to (x,y,w) \in D(s,r)\,, \quad s_+ := s/4\,,$$

close to the identity, such that, defining

$$H^+ := H \circ \Phi =: N + P^+\,,$$

the Hamiltonian vector field X_{P^+} has the same regularity of X_P, $P^+_{ij} = 0$ if $2i + j \leq 2$ and[2]

$$P^+_{i\tilde\jmath\hat\jmath} y^i \tilde w^{\tilde\jmath}\hat w^{\hat\jmath} = \Big[P^+_{i\tilde\jmath\hat\jmath} y^i \tilde w^{\tilde\jmath}\hat w^{\hat\jmath}\Big] \quad \textit{if} \;\; 2i+\tilde\jmath+\hat\jmath \leq D+1 \;\; \textit{and} \;\; \tilde\jmath+\hat\jmath \leq 4\,,\ \hat\jmath \leq 2 \quad \textit{or} \quad \hat\jmath = 1\,. \tag{6.1}$$

Moreover

$$\|[P^+_{i\tilde\jmath\hat\jmath}] - [P_{i\tilde\jmath\hat\jmath}]\| \leq C\kappa_3^2/\alpha\,, \quad \kappa_3 := |P_{11}|_s + |P_{03}|_s\,, \quad \forall\, 2i+\tilde\jmath+\hat\jmath = 4\,, \quad \tilde\jmath = 0,2,4\,, \quad \hat\jmath = 0,2\,. \tag{6.2}$$

In other words, in the case $d > 1$, $D = 4$,

$$\begin{aligned} H^+ \;=\;& \hat\omega\cdot\hat y_+ + \hat\Omega\cdot\hat z_+\bar{\hat z}_+ + P_{003}(x_+)\hat w_+^3 + \frac12\hat A_+\hat y_+\cdot\hat y_+ + \hat B_+\hat y_+\cdot\hat z_+\bar{\hat z}_+ + P^+_{004}(x_+)\hat w_+^4 \\ &+ P^+_{013}(x_+)\tilde w_+^3\hat w_+ + \sum_{2i+\tilde\jmath+\hat\jmath=5,\hat\jmath\neq1} P^+_{i\tilde\jmath\hat\jmath}(x_+) y_+^i \tilde w_+^{\tilde\jmath}\hat w_+^{\hat\jmath} + \sum_{2i+\tilde\jmath+\hat\jmath\geq6} P^+_{i\tilde\jmath\hat\jmath}(x_+) y_+^i \tilde w_+^{\tilde\jmath}\hat w_+^{\hat\jmath}\,, \end{aligned} \tag{6.3}$$

while, in the case $d = 1$, $D = 6$,

$$\begin{aligned} H^+ \;=\;& \hat\omega\cdot\hat y_+ + \hat\Omega\cdot\hat z_+\bar{\hat z}_+ + P_{003}(x_+)\hat w_+^3 + \frac12\hat A_+\hat y_+\cdot\hat y_+ + \hat B_+\hat y_+\cdot\hat z_+\bar{\hat z}_+ + P^+_{004}(x_+)\hat w_+^4 \\ &+ P^+_{013}(x_+)\tilde w_+^3\hat w_+ + \sum_{2i+\tilde\jmath+\hat\jmath=5,6,\ \hat\jmath\leq2} \Big[P^+_{0\tilde\jmath\hat\jmath}(x_+)\tilde w_+^{\tilde\jmath}\hat w_+^{\hat\jmath}\Big] + \sum_{2i+\tilde\jmath+\hat\jmath=5,6\ ,\hat\jmath\geq3} P^+_{i\tilde\jmath\hat\jmath}(x_+) y_+^i \tilde w_+^{\tilde\jmath}\hat w_+^{\hat\jmath}\,, \\ &+ \sum_{2i+\tilde\jmath+\hat\jmath=7,\hat\jmath\neq1} P^+_{i\tilde\jmath\hat\jmath}(x_+) y_+^i \tilde w_+^{\tilde\jmath}\hat w_+^{\hat\jmath} + \sum_{2i+\tilde\jmath+\hat\jmath\geq8} P^+_{i\tilde\jmath\hat\jmath}(x_+) y_+^i \tilde w_+^{\tilde\jmath}\hat w_+^{\hat\jmath}\,, \end{aligned} \tag{6.4}$$

where $\hat A_+ \in \mathrm{Mat}(\hat n \times \hat n)$ and $B_+ \in \mathcal{L}(\mathbb{C}^{\hat n}, \ell_\infty^{\bar p - p})$ satisfy

$$\|\hat A_+ - \hat A\|\,,\ \|\hat B_+ - \hat B\| \leq C(|P_{11}|_s + |P_{03}|_s)^2\alpha^{-1}\,. \tag{6.5}$$

[2] In particular the terms $P^+_{110}, P^+_{101}, P^+_{030}, P^+_{021}, P^+_{012}, P^+_{111}, P^+_{031}, P^+_{041}$ vanish.

Proof. We start with some general considerations. We define the degree of a monomial

$$F = F_{ij} y^i w^j = F_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}} \qquad \text{as} \qquad \deg F := 2i + j = 2i + \tilde{\jmath} + \hat{\jmath}\,.$$

The Poisson brackets of two monomials is a monomial with

$$\deg\{F, G\} = \deg F + \deg G - 2 \quad \text{or} \quad \{F, G\} = 0\,. \tag{6.6}$$

We denote X_F^t the hamiltonian flow generated by F at time t. Then

$$H \circ X_F^1 = \sum_{j \geq 0} L_F^j H / j! \quad \text{where} \quad L_F^j H := \{L_F^{j-1} H, F\} \quad \text{and} \quad L_F^0 H := H\,. \tag{6.7}$$

Let $H = N + P$ be as in (2.1) and suppose that $F = F_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}$ solves the homological equation

$$\{N, F\} + P_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}} = [P_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}]\,. \tag{6.8}$$

By (6.7) and (6.6) we see that the terms of H and $H \circ X_F^1$ with degree less or equal than $\deg F$ of are the same, except for $P_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}$ which is normalised into $[P_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}]$. On the other hand the terms of degree equal to $\deg F + 1$ are changed by a quantity of order $|F|\kappa_3$.
For brevity for the rest of this proof $a \lessdot b$ means that there exists a constant $c = c(n, \tau, s, D, \hat{n}) > 0$ such that $a \leq cb$.
By the Melnikov condition (2.12) there is a solution $F = F_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}$ of the homological equation (6.8) for every $(i, \tilde{\jmath}, \hat{\jmath})$ satisfying the conditions in (6.1). Indeed the existence of F and the estimate

$$|F_{i\tilde{\jmath}\hat{\jmath}}|_{s(1-1/D)} \lessdot |P_{i\tilde{\jmath}\hat{\jmath}}|_s / \alpha \tag{6.9}$$

follows as in Lemmata 1-2 of [24]; we just note that the small divisors involved in the definition of every monomial $f(x) y^m \tilde{z}^{\tilde{a}} \bar{\tilde{z}}^{\tilde{\bar{a}}} \hat{z}^{\hat{a}} \bar{\hat{z}}^{\hat{\bar{a}}}$ of F are $\omega \cdot k + \tilde{\Omega}(\tilde{a} - \tilde{\bar{a}}) + \hat{\Omega}(\hat{a} - \hat{\bar{a}})$, with $\tilde{\Omega} := (\Omega_{n+1}, \ldots, \Omega_{\hat{n}})$, $k \in \mathbb{Z}^n$, $\tilde{a}, \tilde{\bar{a}} \in \mathbb{N}^{\hat{n}-n}$, $\hat{a}, \hat{\bar{a}} \in \mathbb{N}^\infty$ and $|\tilde{a} + \tilde{\bar{a}}| = \tilde{\jmath}$, $|\hat{a} + \hat{\bar{a}}| = \hat{\jmath}$ (then $|\tilde{a} + \tilde{\bar{a}}| \leq \tilde{\jmath}$, $|\hat{a} + \hat{\bar{a}}| \leq \hat{\jmath}$).

We now proceed normalising the terms of degree three with

$$(i, \tilde{\jmath}, \hat{\jmath}) = (1, 1, 0),\ (1, 0, 1),\ (0, 3, 0),\ (0, 2, 1),\ (0, 1, 2)\,. \tag{6.10}$$

Let us define $F^{(3)} := \sum F_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}$ where the sum is taken over the indexes in (6.10). Let $s_3 := s(1 - 1/D)$. For $r_3 > 0$ we have that $|\partial_x F^{(3)}|_{s_3} \lessdot r_3^3$, $|\partial_y F^{(3)}|_{s_3} \lessdot r_3$, $|\partial_w F^{(3)}|_{s_3} \lessdot r_3^2$, since $2i + \tilde{\jmath} + \hat{\jmath} \geq 3$. Therefore we can choose r_3 small enough such that $X^1_{F^{(3)}} : D(s_3, r_3) \to D(s, r)$. Moreover the terms of order three of $H \circ X^1_{F^{(3)}}$ are the same of H except for $P_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}$ with indexes as in (6.10) that are normalised; note that, being of odd degree, they actually annihilate. On the other hand the term of degree four are slightly changed by a quantity of order $|F^{(3)}|_{s_3} \kappa_3 \lessdot \kappa_3^2 / \alpha$ by (6.9).
We now normalise the terms of degree four with

$$(i, \tilde{\jmath}, \hat{\jmath}) = (1, 1, 1)\,, (0, 3, 1)\,, (2, 0, 0)\,, (1, 2, 0)\,, (1, 0, 2)\,, (0, 4, 0)\,, (0, 2, 2)\,. \tag{6.11}$$

Let us define $F^{(4)} := \sum F_{i\tilde{\jmath}\hat{\jmath}} y^i \tilde{w}^{\tilde{\jmath}} \hat{w}^{\hat{\jmath}}$ where the sum is taken over the indexes in (6.11). If $r_4 > 0$ is small enough and $s_3 := s(1 - 2/D)$ we have that $X^1_{F^{(4)}} : D(s_4, r_4) \to D(s_3, r_3)$. The terms of order three and four of $H \circ X^1_{F^{(3)}} \circ X^1_{F^{(4)}}$ are the same of $H \circ X^1_{F^{(3)}}$ except for those with indexes as in (6.10) that are normalised. Note that the terms corresponding to the first two triples in (6.11) annihilate.
The normalisation of all the other terms of degree up to $D + 1$ is analogous. ■

Remark 6.1. *The cubic terms $P_{003}(x_+) \hat{w}_+^3$ on the high modes can not be removed by some averaging procedure because the tangential and normal frequencies satisfy only the second order Melnikov non-resonance conditions (2.12).*

We introduce parameters

$$\xi \in (0, \rho_*]^{\hat n}, \quad \rho_* \in (0, r_+^2/4),$$

and new symplectic variables

$$(x_*, y_*, w_*) = (x_*, \hat y_+ - \xi, \hat w_+) \in D(s_*, r_*) \subset \mathbb{T}^{\hat n}_{s_*} \times \mathbb{C}^{\hat n} \times \ell^{a,p}_b, \ s_* \leq s_+, \ r_* \leq \sqrt{\rho}_*/2$$

where the $\hat n$-dimensional angles are defined by

$$x_{*j} := x_{+j}, \ \forall 1 \leq j \leq n, \quad \sqrt{2(\xi_j + y_{*j})}\big(e^{-\mathrm{i}x_{*j}}, e^{\mathrm{i}x_{*j}}\big) := w_{+j}, \forall \ n < j \leq \hat n.$$

After this symplectic change of coordinates the Hamiltonian H^+ becomes

$$H^* = N^* + P^* = \omega_*(\xi) \cdot y_* + \Omega_*(\xi) \cdot z_* \bar z_* + \sum_{2i+j \geq 0} P^*_{ij}(x_*; \xi) y_*^i w_*^j \tag{6.12}$$

with

$$\omega_*(\xi) := \hat\omega + \hat A_+ \xi, \qquad \Omega_*(\xi) := \hat\Omega + \hat B_+ \xi, \tag{6.13}$$

and, by (6.3), (6.4), denoting for simplicity $|\cdot| := |\cdot|^\lambda_{s_*}$,

$$\text{if } \ d > 1, \ \ |P^*_{00}|, |P^*_{01}| = O(\rho_*^{5/2}), \ \ |P^*_{10}|, |P^*_{02}| = O(\rho_*^2), \ \ |P^*_{11}| = O(\rho_*^{3/2}), \ \ |P^*_{03}| = O(1), \tag{6.14}$$

$$\text{if } \ d = 1, \ \ |P^*_{00}|, |P^*_{01}| = O(\rho_*^{7/2}), \ \ |P^*_{10}|, |P^*_{02}| = O(\rho_*^3), \ \ |P^*_{11}| = O(\rho_*^{5/2}), \ \ |P^*_{03}| = O(1). \tag{6.15}$$

Moreover for $\alpha_* > 0$ and $\lambda := \alpha_*/M$, with $M := \|\hat A_+\| + \|\hat B_+\|$ (recall (5.2)).

We now apply the KAM Theorem 5.1. Take

$$\alpha_* := 9\Theta^2 r_*^2, \ \ \rho_* := r_*^{2\vartheta} \ \text{ where } \ \vartheta \in (9/10, 1) \ \text{ if } \ d > 1, \ \ \vartheta \in (9/14, \mu) \ \text{ if } \ d = 1. \tag{6.16}$$

Remark 6.2. *Other choices of $\alpha_* \geq 9\Theta^2 r_*^2$ are clearly possible, giving different estimates on the Cantor manifold.*

Theorem 2.1 follows applying Theorems 5.1, 5.2 with[3] $H = H^*$, $P = P^*$, $r := r_*$ $\alpha := \alpha_*$, etc. Let us verify the hypotheses of the above theorems. It is immediate to check (A*), (B*), (C*). Let Θ as in (5.5) (with respect to the perturbation P^*); note that $\Theta = O(1)$ with respect to ξ. By (6.14)-(6.16) the KAM condition $(H2)$ of Theorem 5.1 holds.

$$\alpha_*^\mu/\rho_* = \begin{cases} O(r_*^{2(1-\vartheta)}) & \text{for } d > 1 \\ O(r_*^{2(\mu-\vartheta)}) & \text{for } d = 1 \end{cases} \to 0 \quad \text{as} \quad r_* \to 0. \tag{6.17}$$

Since $\hat A_+ = \hat A(Id + \hat A^{-1}(\hat A_+ - \hat A))$, by the twist condition, (6.5) and (2.14) we get that $\hat A_+$ is invertible with

$$\|\hat A_+^{-1} - \hat A^{-1}\| \leq 2\|\hat A^{-1}\|^2 \|\hat A_+ - \hat A\|, \tag{6.18}$$

taking c in (2.14) small enough. Therefore, $\xi \to \omega_*(\xi)$ is a diffeomorphism, see (6.13).

We finally verify that the frequencies ω_*, Ω_* satisfy (5.19) and (5.20). The non-resonance assumption (2.12) implies $|\hat\Omega \cdot l| \geq \alpha$, $\forall 1 \leq |l| \leq 2$, and so[4]

$$|\Omega_*(\xi) \cdot l| \overset{(6.13)}{\geq} |\hat\Omega \cdot l| - |\hat B_+ \xi \cdot l| \geq \alpha - 2\rho_* \|\hat B_+\| \overset{(6.5),(2.14)}{\geq} \alpha - 2\rho_*(\|B_+\| + c) \geq \alpha/2$$

if r_* is small enough. So (5.19) holds.

[3] We apply Theorems 5.1 and 5.2 with $\alpha := \alpha_*$. Here α_* is the parameter defined in (6.16) which is small with r_* and has not to be confused with the *fixed* α appearing in the statement of Theorem 2.1.

[4] Recall that α is fixed and independent of ρ_* and r_* (see also the previous footnote).

Since $\omega_*(\xi) \cdot k + \Omega_*(\xi) \cdot l$ is an affine function of ξ, the condition (5.20) holds if

$$\hat{\omega} \cdot k + \hat{\Omega} \cdot l \neq 0 \quad \text{or} \quad \hat{A}_+ k + \hat{B}_+^{\intercal} l \neq 0 \, .$$

Suppose that $\hat{A}_+ k + \hat{B}_+^{\intercal} l = 0$, then $k = -\hat{A}_+^{-1} \hat{B}_+^{\intercal} l$ and

$$\hat{\omega} \cdot k + \hat{\Omega} \cdot l = (\hat{\Omega} - \hat{B}_+ \hat{A}_+^{-1} \hat{\omega}) \cdot l = (\hat{\Omega} - \hat{B} \hat{A}^{-1} \hat{\omega}) \cdot l + \Big(\hat{B}(\hat{A}^{-1} - \hat{A}_+^{-1}) + (\hat{B} - \hat{B}_+) \hat{A}_+^{-1} \Big) \hat{\omega} \cdot l \neq 0$$

by (2.13) and remark 2.1, (6.5), (6.18) taking c in (2.14) small enough.

Then theorems 5.1 and 5.2 apply and we obtain a family of elliptic $\hat{n}$-dimensional tori parametrized by $\xi \in \mathbf{\Pi}_\infty$, where the set $\mathbf{\Pi}_\infty$ has asymptotically full measure as $r \to 0$ by (5.21) and (6.17).

Remark 6.3. *The KAM theorem in [24] does not apply. Indeed, with only the estimates (6.14)-(6.15) the KAM condition (5.14) implies*

$$\mathit{const} \geq \alpha^{-1} |X_P|_{r,s,E,\Pi} \geq \begin{cases} \mathit{const}\, \alpha^{-1}(\rho^{5/2} r^{-2} + r) = \mathit{const}\, \alpha^{-1}(r^{5\vartheta - 2} + r) & \text{if } d > 1 \\ \mathit{const}\, \alpha^{-1}(\rho^{7/2} r^{-2} + r) = \mathit{const}\, \alpha^{-1}(r^{7\vartheta - 2} + r) & \text{if } d = 1 \end{cases}$$

which is incompatible with the measure estimate $\alpha \ll r^{2\vartheta/\mu}$ (recall (5.21)).

7 Proof of Theorem 3.1

We divide the proof in several steps.

Step 1) Partial Birkhoff Normal Form on $\hat{n} \geq n$ modes

By the non-resonance assumption ($\hat{\mathrm{A}}_2$) where $D \geq 4$, we transform H in partial Birkhoff normal form, up to order 4, on the first $\hat{n} \geq n$ modes, namely

$$H = \hat{\mathrm{a}} \cdot \hat{I} + \hat{\mathrm{b}} \cdot \hat{\zeta} \bar{\hat{\zeta}} + P = \hat{\mathrm{a}} \cdot \hat{I} + \hat{\mathrm{b}} \cdot \hat{\zeta} \bar{\hat{\zeta}} + \frac{1}{2} \hat{\mathrm{A}} \hat{I} \cdot \hat{I} + \hat{\mathrm{B}} \hat{I} \cdot \hat{\zeta} \bar{\hat{\zeta}} + O(|\tilde{\zeta}| \|\hat{\zeta}\|_{a,p}^3) + O(\|\hat{\zeta}\|_{a,p}^4) + O(\|\zeta\|_{a,p}^g) \quad (7.1)$$

where $\hat{\mathrm{a}}$, $\hat{\mathrm{b}}$ are defined in (3.6), the matrices $\hat{A}$, $\hat{B}$ in (3.7), (3.8), $g := \min(\mathrm{g}, 6)$, and

$$\tilde{\zeta} := (\tilde{\zeta}_{n+1}, \ldots, \tilde{\zeta}_{\hat{n}}) \, , \quad \hat{\zeta} := (\hat{\zeta}_{\hat{n}+1}, \hat{\zeta}_{\hat{n}+2}, \ldots) \, , \quad \zeta = (\tilde{\zeta}, \hat{\zeta}) \, , \quad \tilde{I} := \tilde{\zeta} \bar{\tilde{\zeta}} \, , \quad \hat{I} := (I, \tilde{I}) \, .$$

The proof of this statement follows as in [23], [25], [2]. Note that the term $O(|\tilde{\zeta}| \|\hat{\zeta}\|_{a,p}^3)$ can not be removed because ($\hat{\mathrm{A}}_2$) requires only second order Melnikov non-resonance conditions for $n > \hat{n}$.

Step 2) Parameters and action-angle variables on n modes

We introduce parameters

$$\xi \in (0, \rho]^n \, , \quad \rho \in (0,1) \, , \tag{7.2}$$

and angle-action variables (x, y) on the first n modes, setting

$$\zeta_j =: \sqrt{2(\xi_j + y_j)} e^{-\mathrm{i} x_j} \, , \quad 1 \leq j \leq n \, . \tag{7.3}$$

Then $I = \xi + y$ and the Hamiltonian (7.1) assumes the form

$$H = \omega(\xi) \cdot y + \Omega(\xi) \cdot z \bar{z} + \sum_{i,j \geq 0} P_{ij}^*(x; \xi) y^i w^j \quad \text{with} \quad \omega(\xi) := \mathrm{a} + \mathrm{A}\xi \, , \quad \Omega(\xi) := \mathrm{b} + \mathrm{B}\xi \, , \tag{7.4}$$

$z = (\zeta_{n+1}, \ldots)$, $w := (z, \bar{z})$, and

$$|P_{ij}^*|_s^\lambda = O(|\xi|^{\frac{g}{2} - i - \frac{j}{2}}) \, , \ \forall \, 2i + j \leq 3 \, , \quad |P_{ij}^* - P_{ij}|_s^\lambda = O(|\xi|^{\frac{g}{2} - 2}) \, , \ \forall \, 2i + j = 4 \, . \tag{7.5}$$

The Hamiltonian H is real analytic on $D(s,r)$, for some $0<s<1$, $0<r<\rho/2$.

Step 3) Apply the KAM Theorem 5.1 and Theorem 5.2 to H

The assumptions (A^*), (B^*), (C^*) of Theorem 5.1 are implied by (B), (C), as in [23]. We take

$$\alpha := 9\Theta^2 r^2\,, \quad \rho := r^{2\vartheta}\,, \quad \vartheta \in (\bar\mu,\mu) \quad \text{where} \quad \bar\mu := \max\{2(1+\mu)g^{-1}\,,\, 3(g-1)^{-1}\} < \mu \le 1 \tag{7.6}$$

by (3.4).

Remark 7.1. *The parameter domain* Π *can not be the whole* $(0,\rho]^n$ *(see (7.2)) because, by (7.3), the Hamiltonian* H *will be analytic in* $D(s,r)$ *only excluding* $|\xi|\le Cr^2$. *This difficulty can be handled as in [23], section 7, step 5. For simplicity of exposition we skip this technical detail in the following.*

The KAM condition (H3) reduces, by (7.5)-(7.6), to

$$\varepsilon_3 = O(\max\{r^{g\vartheta-2-2\mu}\,, r^{\vartheta(g-1)-3}\}) \le \gamma \qquad \text{and} \qquad O(r^{(g-3)\vartheta-1}) < 1\,, \tag{7.7}$$

which are both verified for r small enough because $(g-3)\vartheta-1>0$ and

$$\varepsilon_3 \to 0 \quad \text{as} \quad r\to 0\,.$$

By Theorem 5.1 there is, $\forall \xi \in \mathbf{\Pi}_\infty$ defined in (5.12), an analytic symplectic map $\Phi(\cdot;\xi) : D(s/4,r/4) \to D(s,r)$ such that

$$H^\infty := H\circ\Phi = \omega_\infty(\xi)\cdot y_\infty + \Omega_\infty(\xi)\cdot z_\infty \bar z_\infty + P^\infty \quad \text{with} \quad P^\infty_{ij} = 0\,, \forall\, 2i+j\le 2\,.$$

Moreover the assumptions (5.19), (5.20) of Theorem 5.2 hold by (7.4) and (A). By Theorem 5.2 the Cantor set of parameters $\mathbf{\Pi}_\infty$ has asymptotically full measure

$$\frac{|\Pi/\mathbf{\Pi}_\infty|}{|\Pi|} = O\Big(\frac{\alpha^\mu\rho^{n-1}}{\rho^n}\Big) = O(r^{2(\mu-\vartheta)}) \to 0 \quad \text{as} \quad r\to 0\,. \tag{7.8}$$

By (5.9), (5.10) with $p_a=1$, and (7.6), we get

$$\begin{cases} |P^\infty_{11} - P^*_{11}| \le C\big(|P^*_{11}| + r\big)\varepsilon_3 \\ |P^\infty_{03} - P^*_{03}| \le C\big(|P^*_{03}| + |P^*_{11}| + r\big)\varepsilon_3 \end{cases} \qquad |P^\infty_{ij} - P^*_{ij}| \le C\varepsilon_3\,, \quad \forall\, 2i+j=4\,, \tag{7.9}$$

where $|\cdot| := |\cdot|^\lambda_{s/4}$ and $C := C(\gamma,\Theta)$. Moreover, (5.6), (7.4), (7.6),

$$\begin{cases} |\omega_\infty(\xi) - \mathrm{a}| \le \gamma^{-1}\alpha\varepsilon_3 + \|\mathrm{A}\||\xi| \le Cr^{2\vartheta} \\ |\Omega_\infty(\xi) - \mathrm{b}|_{\bar p - p} \le \gamma^{-1}\alpha\varepsilon_3 + \|\mathrm{B}\||\xi| \le Cr^{2\vartheta}\,. \end{cases} \tag{7.10}$$

Step 4) Apply Theorem 2.1 to H^∞

Assumptions (2.3), (2.5) of Theorem 2.1 hold by (7.4). The non-resonance assumption (2.12) holds

$$\text{for any } \xi \in \begin{cases} \Pi_0 & \text{if} \quad d>1 \\ \Pi_0 \cap \omega^{-1}(\mathcal{D}_{\alpha^\mu,\tau}) & \text{if} \quad d=1 \end{cases}$$

where

$$\Pi_0 := \Big\{\xi\in\Pi \ : \ |\omega_\infty(\xi)\cdot k + \Omega_\infty(\xi)\cdot l| \ge 2\alpha\frac{\langle l\rangle_d}{1+|k|^\tau}\,, \forall\, k\in\mathbb{Z}^n\,,\ l\in\Lambda_{\hat n,D}\Big\} \subset \Pi_\infty \tag{7.11}$$

and $\Lambda_{\hat n,D}$ is defined in (2.12). In the next section we prove that also Π_0 has asymptotically full measure

$$\frac{|\Pi\setminus\Pi_0|}{|\Pi|} = O\Big(\frac{\rho^{n-1}\alpha^\mu}{\rho^n}\Big) = O(r^{2(\mu-\vartheta)}) \to 0 \quad \text{as} \quad r\to 0\,. \tag{7.12}$$

Step 5) Check the Twist condition

The matrices $\hat{A}$, $\hat{B}$ defined in (2.8), (2.10) (with $P = P^\infty$) satisfy, by (7.9), (7.5), (7.6),

$$\|\hat{A} - \hat{\mathrm{A}}\| \leq C(\varepsilon_3 + r^{\theta(g-4)}), \qquad \|\hat{B} - \hat{\mathrm{B}}\| \leq C(\varepsilon_3 + r^{\theta(g-4)}). \tag{7.13}$$

The matrix $\hat{\mathrm{A}}$ is invertible by $(\hat{\mathrm{A}}_1)$. The twist condition follows for r small enough.

Step 6) Check the non-resonance condition (2.13)

By (7.10), (7.13), for every $0 < |\hat{l}| \leq 2$

$$\left|(\hat{\Omega} - \hat{B}\hat{A}^{-1}\hat{\omega}) \cdot \hat{l} - (\hat{\mathrm{b}} - \hat{\mathrm{B}}\hat{\mathrm{A}}^{-1}\hat{\mathrm{a}}) \cdot \hat{l}\right| \to 0 \quad \text{as} \quad r \to 0. \tag{7.14}$$

Assumption $(\hat{\mathrm{A}}_3)$ and remark 2.1 imply

$$\inf_{0<|\hat{l}|\leq 2} |(\hat{\mathrm{b}} - \hat{\mathrm{B}}\hat{\mathrm{A}}^{-1}\hat{\mathrm{a}}) \cdot \hat{l}| > 0,$$

and (2.13) follows by (7.14) for r small enough.

Step 7) Check the smallness condition (2.14)

By (7.9), we get, for r small enough,

$$|P^\infty_{11}| + |P^\infty_{03}| \leq 2|P^*_{11}| + 2|P^*_{03}| + O(\varepsilon_3 r) \overset{(7.5),(7.6)}{=} O(r^{(g-3)\vartheta} + \varepsilon_3 r). \tag{7.15}$$

Then

$$(|P^\infty_{11}| + |P^\infty_{03}|)^2 \alpha^{-1} \overset{(7.6)}{\leq} C r^{2(g-3)\vartheta - 2} + \varepsilon_3^2 \to 0 \quad \text{as} \quad r \to 0.$$

Proof of Theorem 3.2. We apply Theorem 5.3 to H in (7.4). The hypotheses of Theorem 5.3 hold, in particular condition (5.22) is $(\mathrm{b} - \mathrm{B}\mathrm{A}^{-1}\mathrm{a}) \cdot l \neq 0$, $\forall\, 1 \leq |l| \leq 2$. Moreover $0 \notin \omega(\Pi)$ because $\mathrm{a} \neq 0$ and ρ (namely r) is small enough. We fix $\rho_0 := c\rho$. The segment $[1 - c\rho_0, 1 + c\rho_0]\bar{\omega} \subset \omega_\infty(\Pi)$ for c small enough. Moreover, $\alpha_0 := \rho_0^{1+c} = (c\rho)^{1+c} > K\alpha$ by (7.6), for r and c small enough, where $K > 1$ is the constant defined in Theorem 5.3. Then $\bar{\omega} \in \mathcal{D}_{K\alpha,\tau}$ and (3.10) follows by (5.23) and since $\alpha^\mu/\rho \to 0$ as $r \to 0$ by (7.6).

Remark 7.2. *Actually $\omega_\infty(\Pi)$ is not a neighborhood of the frequency* a*, since $\Pi = (0,\rho]^n$ is not a neighborhood of* 0*. Nevertheless this small technical point is bypassed as follows. For $1 \leq j \leq n$, inverting the signs in the definition (7.3), namely $\zeta_j := \sqrt{2(\xi_j - y_j)}e^{+\mathrm{i}x_j}$, the new tangential frequency in (7.4) becomes $\omega(\xi) = \mathrm{a} + \mathrm{A}(\xi_1, \ldots, -\xi_j, \ldots, \xi_n)$. Taking all the possible choices of $1 \leq j \leq n$ and $\pm$ signs, $\xi \in \Pi$ span a whole neighbourhood of the frequency* a*, except for n hyperplanes passing through* a *(but not through the origin).*

7.1 Measure estimates

The next proposition implies (7.12) concluding the proof of Theorem 3.1.

Proposition 7.1. *$|\Pi \setminus \Pi_0| \leq c\rho^{n-1}\alpha^\mu$ where μ is defined in (3.4) and the constant c depends on* $\mathrm{a}, \mathrm{b}, \mathrm{A}, \mathrm{B}, n, \hat{n}, d, D, a_*, \kappa, \delta_*$.

We have to estimate

$$\Pi \setminus \Pi_0 = \bigcup_{k \in \mathbb{Z}^n,\, l \in \Lambda_{\hat{n},D}} R_{kl}(\alpha) \tag{7.16}$$

where R_{kl} are the "resonant zones"

$$R_{kl}(\alpha) := \left\{ \xi \in \Pi \, : \, |\omega_\infty(\xi) \cdot k + \Omega_\infty(\xi) \cdot l| < \frac{2\alpha\langle l \rangle_d}{1 + |k|^\tau} \right\}.$$

In the case $d > 1$ there are at most finitely many nonempty resonant zones $R_{kl}(\alpha)$. This is a consequence of the next lemmata. The case $d = 1$ is more complex.

Lemma 7.1. *Let $d > 1$. There are $D_* \geq 1$, $\sigma_* > 0$, such that*

$$\langle l \rangle_d \geq D_*^{-1} |l|_{\sigma_*} |l|_{\delta_*} , \quad \forall l \in \Lambda_{\hat{n},D} . \tag{7.17}$$

PROOF. We consider only the more difficult case $l = (\tilde{l}, \hat{l})$, $\hat{l} = e_i - e_j$, $i > j$. We have

$$\langle l \rangle_d \geq i^d - (i-1)^d - D\hat{n}^d \geq i^{d-1} - D\hat{n}^d > i^{d-1}/2 \quad \text{for} \quad i^{d-1} > 2D\hat{n}^d . \tag{7.18}$$

Defining $\delta_0 := \max\{\delta_*, 0\}$, $\sigma_* := d - 1 - \delta_0 > 0$, we have

$$|l|_{\sigma_*} |l|_{\delta_*} \leq D i^{\sigma_*} D i^{\delta_0} = D^2 i^{d-1} . \tag{7.19}$$

Let $D_* := 2D^3 \hat{n}^d$. If $i^{d-1} > 2D\hat{n}^d$ then (7.17) follows by (7.18); if $i^{d-1} \leq 2D\hat{n}^d$, by (7.19). ■

Remark 7.3. *For $d = 1$, $D \geq 3$ (as in this paper) the bound (7.17) is false. Taking for example $l = l^{(j)} := e_{\hat{n}+j} - e_j - e_{\hat{n}}$ with $j > \hat{n}$ we have*

$$\langle l^{(j)} \rangle = 1 , \ |l^{(j)}|_{\delta_*} \geq \hat{n}^{\delta_*} , \ |l^{(j)}|_{\sigma_*} \geq j^{\sigma_*} \to \infty \text{ as } j \to \infty .$$

This motivates assumption (A_3) for $d = 1$. The bound (7.17) is true for $d = 1$, $D = 2$, see [24].

Lemma 7.2. *There exists $\beta_0 > 0$ (depending on $d, \mathrm{b}, \hat{n}, D$) such that*

$$|\mathrm{b} \cdot l| \geq 4\beta_0 \langle l \rangle_d , \quad \forall l \in \Lambda_{\hat{n},D} . \tag{7.20}$$

PROOF. We consider only the subtlest case $l = (\tilde{l}, \hat{l})$, $|\hat{l}| = 2$, $\hat{l} = e_i - e_j$, $i > j$. We have

$$|\mathrm{b} \cdot l| \geq |\mathrm{b}_i - \mathrm{b}_j| - c_1 , \qquad \langle l \rangle_d \leq i^d - j^d + c_2 , \tag{7.21}$$

for some $c_1 := c_1(D, \mathrm{b}_{n+1}, \ldots, \mathrm{b}_{\hat{n}})$, $c_2 := c_2(d, \hat{n}, D) > 0$. By ($\mathrm{A}_2$) and (B) there is $\beta_1 > 0$ such that

$$|\mathrm{b}_i - \mathrm{b}_j| \geq 2\beta_1 (i^d - j^d) , \quad \forall i > j . \tag{7.22}$$

By (7.21), (7.22), for $\beta_0 \leq \beta_1/4$ we have that

$$\beta_1 (i^d - j^d) \geq \beta_1 c_2 + c_1 \quad \Longrightarrow \quad |\mathrm{b} \cdot l| \geq 4\beta_0 \langle l \rangle_d . \tag{7.23}$$

Let $d > 1$. If $i > i_0$ we have $i^d - j^d \geq d i_0^{d-1}$, so (7.23) follows for i_0 large. On the other hand, the set of $|\tilde{l}| \leq D - 2$, $j < i \leq i_0$ is finite and $\langle l \rangle_d \leq D i_0^d$. Hence (7.20) follows by ($\hat{\mathrm{A}}_2$) for β_0 small enough. Let now $d = 1$. Take h large such that $\beta_1 h \geq \beta_1 c_2 + c_1$. Then (7.23) holds for $i - j \geq h$. On the other hand, if $i - j < h$, we have $\langle l \rangle_1 \leq h + \hat{n} D$ and (7.20) follows by ($\hat{\mathrm{A}}_2$) for β_0 small enough. ■

In the following r is small enough.

Lemma 7.3. $|\Omega_\infty(\xi) \cdot l| \geq 3\beta_0 \langle l \rangle_d$, *$\forall \xi \in \Pi$, $l \in \Lambda_{\hat{n},D}$.*

PROOF. By (7.10), $\bar{p} - p \geq -\delta_*$, and Lemma 7.2, we have

$$|\Omega_\infty(\xi) \cdot l| \geq |\mathrm{b} \cdot l| - |l|_{\delta_*} |\Omega_\infty(\xi) - \mathrm{b}|_{-\delta_*} \geq 4\beta_0 \langle l \rangle_d - C |l|_{\delta_*} r^{2\vartheta} .$$

If $d > 1$ Lemma 7.1 implies $|l|_{\delta_*} \leq D_* \langle l \rangle_d$ and the thesis follows for r small enough. If $d = 1$ we have $\delta_* < 0$ (see (3.3)). Therefore $|l|_{\delta_*} \leq D + 1$ and we conclude again for r small. ■

Lemma 7.4. *If $R_{kl}(\alpha) \neq \emptyset$, $\alpha \leq \beta_0$, then*

$$|k| \geq \theta \langle l \rangle_d \quad \text{with} \quad \theta := \beta_0 / (1 + |\mathrm{a}|) . \tag{7.24}$$

PROOF. If there exists $\xi \in R_{kl}(\alpha)$ then $|\omega_\infty(\xi)\cdot k + \Omega_\infty(\xi)\cdot l| < 2\alpha\langle l\rangle_d$ and, using Lemma 7.3,

$$|k||\omega_\infty(\xi)| \geq |k\cdot\omega_\infty(\xi)| \geq |\Omega_\infty(\xi)\cdot l| - 2\alpha\langle l\rangle_d \geq 3\beta_0\langle l\rangle_d - 2\alpha\langle l\rangle_d \geq \beta_0\langle l\rangle_d\,.$$

By (7.10) we have $|\omega_\infty(\xi)| \leq |\mathrm{a}| + 1$ for r small enough, implying (7.24). ■

From now on we always assume $\alpha \leq \beta_0$ taking r small enough. By the previous lemma we shall restrict the union in (7.16) when $|k| \geq \theta\langle l\rangle_d$. In particular we shall always assume $k \neq 0$. In the following $a \lessdot b$ means that there is a constant c, depending on the same quantities as the constant of Proposition 7.1, such that $a \leq cb$. Moreover M, L defined in (5.2), (5.16) respectively, are, here,

$$M = \|\mathrm{A}\| + \|\mathrm{B}\|,\ L = \|\mathrm{A}^{-1}\|\,.$$

Lemma 7.5. *If $|k| \geq 8LM|l|_{\delta_*}$ then $R_{kl}(\alpha) \lessdot \rho^{n-1}\alpha/(1+|k|^\tau)$.*

PROOF. Assume that r is small enough such that $\varepsilon_3 \leq \gamma/(2LM)$. By remark 5.2 the frequency map ω_∞ is invertible from Π to $\tilde\Pi := \omega_\infty(\Pi)$ with $|\omega_\infty^{-1}|^{\mathrm{lip}} \leq 2L$. We introduce the final frequencies $\zeta = \omega_\infty(\xi)$ as parameters over the domain $\tilde\Pi$. Then $\tilde\Omega(\zeta) := \Omega_\infty\big(\omega_\infty^{-1}(\zeta)\big)$ satisfies (see remark 5.2)

$$|\tilde\Omega|_{-\delta_*} \leq |\Omega_\infty|^{\mathrm{lip}}_{-\delta_*}|\omega_\infty^{-1}|^{\mathrm{lip}} \leq 2M2L = 4ML\,. \tag{7.25}$$

Choose a vector $v \in \{-1,1\}^n$ such that $v\cdot k = |k|$ and write $\zeta = sv + w$ with $s \in \mathbb{R}$ and $w \perp v$. Then

$$\zeta\cdot k + \tilde\Omega(\zeta)\cdot l = s|k| + \tilde\Omega(sv+w)\cdot l =: f_{kl}(s) \tag{7.26}$$

and the resonant zones write

$$\tilde R_{kl}(\alpha) := \omega_\infty\big(R_{kl}(\alpha)\big) = \Big\{\zeta = sv + w \in \tilde\Pi\ :\ |f_{kl}(s)| < 2\alpha\frac{\langle l\rangle_d}{1+|k|^\tau}\Big\}\,.$$

By (7.26), (7.25) we have

$$f_{kl}(s_2) - f_{kl}(s_1) \geq (s_2 - s_1)|k| - 4ML|l|_{\delta_*}(s_2 - s_1) \geq |k|(s_2 - s_1)/2$$

because $|k| \geq 8LM|l|_{\delta_*}$. Fubini's theorem implies

$$|\tilde R_{kl}(\alpha)| \leq \frac{2}{|k|}(\operatorname{diam}\tilde\Pi)^{n-1}2\alpha\frac{\langle l\rangle_d}{1+|k|^\tau}\,.$$

Going back to the original parameter domain Π by the inverse map ω_∞^{-1} and noting that $\operatorname{diam}\tilde\Pi \leq 2M\operatorname{diam}\Pi$ (by remark 5.2), $\langle l\rangle_d \leq \theta^{-1}|k|$ (by Lemma 7.4), the final estimate follows. ■

We estimate the other resonant zones $R_{kl}(\alpha)$ using that the unperturbed frequencies in (7.4) are affine functions of ξ and assumption (A_3). We have

$$\omega_\infty(\xi)\cdot k + \Omega(\xi)\cdot l = a_{kl} + b_{kl}\cdot\xi + \mathcal{R}_{kl}(\xi) \tag{7.27}$$

where

$$a_{kl} := \mathrm{a}\cdot k + \mathrm{b}\cdot l \in \mathbb{R}\,,\quad b_{kl} := \mathrm{A}k + \mathrm{B}^\intercal l \in \mathbb{R}^n\,, \tag{7.28}$$

and

$$\mathcal{R}_{kl}(\xi) := (\omega_\infty(\xi) - \omega(\xi))\cdot k + (\Omega_\infty(\xi) - \Omega(\xi))\cdot l\,. \tag{7.29}$$

Assumption (A_3) implies that

$$\delta_{kl} := \min\{|a_{kl}|\,,\,|b_{kl}|\} > 0\,,\quad \forall k \in \mathbb{Z}^n\,,\ l \in \Lambda_{\hat n,D},\ (k,l) \neq 0\,.$$

Moreover (7.29), (5.6), imply

$$|\mathcal{R}_{kl}(\xi)| \lessdot \varepsilon_3\alpha(|k| + |l|_{\delta_*})\,,\quad |\mathcal{R}_{kl}|^{\mathrm{lip}} \lessdot \varepsilon_3(|k| + |l|_{\delta_*})\,. \tag{7.30}$$

Lemma 7.6. *Fix* $K_* > 0$. *For all* $0 < |k| \le K_*$, $l \in \Lambda_{\hat{n},D}$, $(k,l) \neq 0$,

$$\alpha \le \theta\delta_{kl}/4 \quad \Longrightarrow \quad |R_{kl}(\alpha)| \lessdot \rho^{n-1}\alpha/\delta_{kl}\,. \tag{7.31}$$

PROOF. If $d > 1$, by Lemma 7.1, (7.24), and $\delta_* < 0$, we get

$$|l|_{\delta_*} \le \begin{cases} \langle l \rangle_d \le K_*/\theta & \text{if } d > 1 \\ D+1 & \text{if } d = 1\,. \end{cases} \tag{7.32}$$

CASE I: $|a_{kl}| = \delta_{kl}$. By (7.27), (7.30), (7.32) we get, for r small enough,

$$\begin{aligned} |\omega_\infty(\xi)\cdot k + \Omega_\infty(\xi)\cdot l| \quad &\ge \quad |a_{kl}| - (\|A\||k| + \|B\||l|)r^{2\vartheta} - |\mathcal{R}_{kl}| \ge |a_{kl}| - cK_* r^{2\vartheta} \ge \delta_{kl}/2 \\ &\overset{(7.31)}{\ge} \quad 2\alpha\theta^{-1} \overset{(7.24)}{\ge} 2\alpha\langle l \rangle_d |k|^{-1} \ge \frac{2\alpha\langle l \rangle_d}{1+|k|^\tau} \end{aligned}$$

implying that $R_{kl}(\alpha) = \emptyset$.

CASE II: $|b_{kl}| = \delta_{kl}$. Set $\xi = \xi_s = b_{kl}|b_{kl}|^{-1}s + w$ with $s \in \mathbb{R}$, $w \perp b_{kl}$. By (7.27), (7.30), (7.32) the function $f_{kl}(s) := \omega_\infty(\xi_s)\cdot k + \Omega_\infty(\xi_s)\cdot l$ satisfies, taking r small,

$$g_{kl}(s_2) - g_{kl}(s_1) \ge \frac{|b_{kl}|}{2}(s_2 - s_1) = \frac{\delta_{kl}}{2}(s_2 - s_1)\,.$$

Arguing as in Lemma 7.5 by Fubini's theorem we obtain

$$|R_{kl}(\alpha)| \lessdot \frac{\rho^{n-1}\alpha\langle l \rangle_d}{\delta_{kl}(1+|k|^\tau)} \le \frac{\rho^{n-1}\alpha|k|}{\delta_{kl}\theta(1+|k|^\tau)}\,,$$

and the thesis follows (since $\tau \ge 1$). ■

We now distinguish the cases $d > 1$ and $d = 1$.

• **Case** $d > 1$

Let

$$L_* := 8D_*LM\theta^{-1}\,, \qquad K_* := 8LM \max_{|l|_{\sigma_*} \le L_*} |l|_{\delta_*}\,.$$

Lemma 7.7. $|R_{kl}(\alpha)| \lessdot \rho^{n-1}\alpha/(1+|k|^\tau)$, $\forall k \in \mathbb{Z}^n$, $l \in \Lambda_{\hat{n},D}$.

PROOF. If $|k| \le K_*$, $|l|_{\sigma_*} \le L_*$, (7.7) follows by Lemma 7.6. Then we can suppose that $|k| > K_*$ or $|l|_{\sigma_*} > L_*$. If $R_{kl}(\alpha) \neq \emptyset$ and $|l|_{\sigma_*} > L_*$, then

$$|k| \ge \theta\langle l \rangle_d \overset{(7.17)}{\ge} \theta|l|_{\sigma_*}|l|_{\delta_*}/D_* \overset{(7.24)}{\ge} 8LM|l|_{\delta_*}\,.$$

On the other hand, when $|l|_{\sigma_*} \le L_*$ we have $|k| > K_* \ge 8LM|l|_{\delta_*}$. So, in both cases Lemma 7.5 applies proving (7.7). ■

Lemma 7.8. $\mathrm{card}\{\, l : \langle l \rangle_d \le \theta^{-1}|k| \} \lessdot |k|^{\frac{2}{d-1}}$.

PROOF. We claim that

$$c_\flat \langle l \rangle_d \ge |l|_{d-1}\,, \qquad c_\flat := 2D^2\hat{n}^d\,. \tag{7.33}$$

We consider only the case $l = (\tilde{l}, e_i - e_j)$, $i > j$. We have $|l|_{d-1} \le Di^{d-1}$. If $i^{d-1} \le 2Dm^d$, then $c_\flat\langle l \rangle_d \ge c_\flat \ge Di^{d-1} \ge |l|_{d-1}$. Otherwise by (7.18) $\langle l \rangle_d \ge i^{d-1}/2 \ge Di^{d-1}/c_\flat \ge |l|_{d-1}/c_\flat$ and (7.33) follows. Therefore

$$\mathrm{card}\{\, l \;:\; \langle l \rangle_d \le \theta^{-1}|k| \} \le \mathrm{card}\{\, l \;:\; |l|_{d-1} \le c_\flat\theta^{-1}|k| \} \lessdot |k|^{\frac{2}{d-1}}\,.$$

■

By (7.16), (7.24) and Lemmata 7.7, 7.8, we deduce

$$|\Pi \setminus \Pi_0| \le \sum_{|k|\ge\theta\langle l\rangle} |R_{kl}(\alpha)| \lessdot \sum_k \rho^{n-1}\alpha|k|^{\frac{2}{d-1}}/(1+|k|^\tau) \overset{(2.11)}{\lessdot} \rho^{n-1}\alpha$$

namely Proposition 7.1 in the case $d > 1$.

• **Case** $d = 1$

Set

$$K_0 := 8(D+1)ML\,, \qquad L_0 := K_0/\theta\,. \tag{7.34}$$

Lemma 7.9. $\inf\{\delta_{kl}\ :\ 0 < |k| \le K_0\,,\ \langle l\rangle_1 \le L_0\} > 0$.

PROOF. Let $l = (\tilde{l}, \hat{l})$. Since the set $\{\langle l\rangle_1 \le L_0\} \cap \{|\hat{l}| = 0\}$ is finite, we consider $|\hat{l}| = 1$ or 2. If $\hat{l} = \hat{l}^{(j)} = \pm e_j$, $j > \hat{n}$ we have $a_{kl} = \mathrm{a}\cdot k + \mathrm{b}\cdot(\tilde{l}, 0) \pm \mathrm{b}_j \to \pm\infty$ as $j \to \infty$. The same holds for $\hat{l} = \pm(e_i + e_j)$, $i, j > \hat{n}$. It remains only the case $\hat{l} = \pm(e_i - e_j)$, $i > j$. Then $\hat{l} = \hat{l}^{(j)} = \pm(e_{h+j} - e_j)$ for some $1 \le h \le L_0 + \hat{n}(D-2)$ (since $L_0 \ge \langle l\rangle_1 \ge h - |\tilde{l}| \ge h - \hat{n}(D-2)$). As $j \to \infty$ we have

$$\begin{aligned} a_{kl} &= \mathrm{a}\cdot k + \mathrm{b}\cdot(\tilde{l}, 0) \pm (\mathrm{b}_{h+j} - \mathrm{b}_j) \to \mathrm{a}\cdot k + \mathrm{b}\cdot(\tilde{l}, 0) \pm h\,, \\ b_{kl} &= \mathrm{A}k + \mathrm{B}^\intercal(\tilde{l}, 0) + \mathrm{B}^\intercal(0, \pm(e_{h+j} - e_j)) \to \mathrm{A}k + \mathrm{B}^\intercal(\tilde{l}, 0)\,. \end{aligned}$$

We conclude by Assumption (A_3). ■

Lemma 7.10. *For all $k \in \mathbb{Z}^n$, $l \in \Lambda_{\hat{n},D}$, there hold $|R_{kl}(\alpha)| \lessdot \rho^{n-1}\alpha/(1+|k|^\tau)$.*

PROOF. If $|k| \ge K_0 \ge 8LM|l|_{\delta_*}$ because $|l|_{\delta_*} \le (D+1)$ (recall $\delta_* < 0$) the estimate follows by Lemma 7.5. If $|k| < K_0$ we conclude by Lemmata 7.6 and 7.9. ■

We can not estimate $\cup_l R_{kl}(\alpha)$ with $\sum_l |R_{kl}(\alpha)|$ because, even with the constraint $\langle l\rangle_1 \le |k|/\theta$, there exist infinitely many $l = (\tilde{l}, e_{h+j} - e_j)\,,\ j > \hat{n}$, with $\langle l\rangle_1 \le \hat{n}D + h$, $\forall h \ge 1$. We need more refined estimates. We decompose

$$\Lambda_{\hat{n},D} = \Lambda_1 \cup \Lambda_2\,, \quad \Lambda_2 := \Big\{ l = (\tilde{l}, \hat{l})\,,\ \hat{l} = \pm(e_{h+j} - e_j)\,,\ j > \hat{n}\,,\ h \ge 1 \Big\}\,, \quad \Lambda_1 := \Lambda_{\hat{n},D} \setminus \Lambda_2\,.$$

Lemma 7.11. $\mathrm{card}\big(\Lambda_1 \cap \{\langle l\rangle_1 \le |k|/\theta\}\big) \lessdot |k|^2$.

PROOF. We consider only the case $|\hat{l}| = 2$, $\hat{l} = \pm(e_i + e_j)$, $i, j > \hat{n}$ (the cases $|\hat{l}| = 0, 1$ are simpler). We have $|\tilde{l}| \le D - 2$ and $|i + j| \le |k|\theta^{-1} + \hat{n}D \lessdot |k|$, implying the lemma. ■

Lemmata 7.10, 7.11 imply

$$\Big| \bigcup_{l\in\Lambda_1} R_{kl}(\alpha) \Big| \lessdot \frac{|k|^2}{1+|k|^\tau}\rho^{n-1}\alpha\,. \tag{7.35}$$

We now consider the more difficult case $l \in \Lambda_2$. We define

$$Q_{k\tilde{l}hj}(\alpha) := \Big\{ \xi \in \Pi\ :\ |\omega_\infty(\xi)\cdot k + \Omega_\infty(\xi)\cdot(\tilde{l}, 0) + h| \le \delta_{khj} \Big\}$$

where

$$\delta_{khj} := \frac{2\alpha|k|}{\theta(1+|k|^\tau)} + \frac{2(1+\|\mathrm{B}\|)\rho}{j^{-\delta_*}} + \frac{a_* h}{j^\kappa}\,.$$

Lemma 7.12. *Let $1 \le h \le \theta^{-1}|k| + \hat{n}(D-2)$, $j > \hat{n}$. For r small enough*

$$|Q_{k\tilde{l}hj}(\alpha)| \lessdot \rho^{n-1}\left(\frac{\alpha}{1+|k|^\tau} + \frac{\rho}{j^{-\delta_*}} + \frac{1}{j^\kappa}\right) . \tag{7.36}$$

Moreover, if $l^{(j)} = (\tilde{l}, \hat{l}^{(j)}) \in \Lambda_2$, $\hat{l}^{(j)} = e_{h+j} - e_j$, then $R_{kl^{(j)}}(\alpha) \subseteq Q_{k\tilde{l}hj}(\alpha)$.

Proof. If $|k| \ge K_0$, arguing as in the proof of Lemma 7.5, for r small enough we get $|Q_{k\tilde{l}hj}(\alpha)| \lessdot \rho^{n-1}\delta_{khj}/|k|$ and the estimate follows since $h \le \theta^{-1}|k| + \hat{n}(D-2)$. On the other hand, if $|k| < K_0$ we have $h \le L_0 + \hat{n}(D-2)$; by assumption (A$_3$) and arguing as in the proof of Lemmata 7.6 and 7.9, for r small enough we have $|Q_{k\tilde{l}hj}(\alpha)| \lessdot \rho^{n-1}\delta_{khj}$ and the estimate follows as above.

We now prove that $R_{kl^{(j)}}(\alpha) \subseteq Q_{k\tilde{l}hj}(\alpha)$. We have $\Omega_\infty(\xi) \cdot l^{(j)} = \Omega_\infty(\xi) \cdot (\tilde{l}, 0) + \Omega_\infty(\xi) \cdot (0, \hat{l}^{(j)})$. By (5.6) and (3.5) we have

$$\begin{aligned} |\Omega_\infty(\xi) \cdot (0, \hat{l}^{(j)}) - h| &\le |\Omega_\infty(\xi) \cdot (0, \hat{l}^{(j)}) - \mathrm{b} \cdot \hat{l}^{(j)} - \mathrm{B}\xi \cdot \hat{l}^{(j)}| + |\mathrm{B}\xi \cdot \hat{l}^{(j)}| + |\mathrm{b}_{j+h} - \mathrm{b}_j - h| \\ &\le 2\gamma^{-1}\alpha\varepsilon_3 |\hat{l}^{(j)}|_{\delta_*} + \|\mathrm{B}\|\rho|\hat{l}^{(j)}|_{\delta_*} + a_* h j^{-\kappa} \\ &\le 2(\|\mathrm{B}\|+1)\rho j^{\delta_*} + a_* h j^{-\kappa} \end{aligned}$$

(for r small enough $2\alpha \le \rho$); the thesis follows since $\langle l \rangle_1 \le \theta^{-1}|k|$ by Lemma 7.4. ■

We choose

$$j_0 := \left(\frac{1+|k|^\tau}{\alpha}\right)^{\frac{1}{1+\kappa}} . \tag{7.37}$$

Since $R_{kl^{(j)}}(\alpha) \subset Q_{k\tilde{l}hj}(\alpha) \subseteq Q_{k\tilde{l}hj_0}(\alpha)$ for $j \ge j_0$, we have

$$\left|\bigcup_{j>\hat{n}} R_{kl^{(j)}}(\alpha)\right| \le \sum_{\hat{n}<j<j_0} |R_{kl^{(j)}}(\alpha)| + |Q_{k\tilde{l}hj_0}| \lessdot \rho^{n-1}\left(\frac{\alpha j_0}{1+|k|^\tau} + \frac{\rho}{j_0^{-\delta_*}} + \frac{1}{j_0^\kappa}\right) \tag{7.38}$$

by Lemma 7.10 and (7.36). By (7.38), (7.37), (7.6) choosing $\vartheta \in (\max\{\bar{\mu}, \mu + \delta_*(1+\kappa)^{-1}\}, \mu)$ (note $\delta_* < 0$) we get, for r small enough (recall that $-\delta_* \le \kappa$)

$$\left|\bigcup_{j>\hat{n}} R_{kl^{(j)}}(\alpha)\right| \lessdot \rho^{n-1}\frac{\alpha^\mu}{(1+|k|^\tau)^{\frac{\delta_*}{\delta_*-1}}} .$$

Since $\langle l \rangle_1 \le |k|/\theta$ implies $h \le \hat{n}(D-2) + |k|/\theta$, and $\mathrm{card}\{\tilde{l} : |\tilde{l}| \le D-2\} \lessdot 1$ we have

$$\left|\bigcup_{l\in\Lambda_2} R_{kl}(\alpha)\right| \lessdot \rho^{n-1}\frac{\alpha^\mu}{(1+|k|^\tau)^{\frac{\delta_*}{\delta_*-1}}} . \tag{7.39}$$

By (7.39) and (7.35) we get

$$\left|\bigcup_{l\in\Lambda_{\hat{n},D}} R_{kl}(\alpha)\right| \lessdot \rho^{n-1}\frac{\alpha^\mu |k|^2}{(1+|k|^\tau)^{\frac{\delta_*}{\delta_*-1}}} .$$

Summing over k and by the choice of τ in (2.11) we get Proposition 7.1 also when $d = 1$.

8 Proof of the basic KAM Theorem 5.1

8.1 Technical lemmata

We first give some lemmata on composition of families of analytic functions depending in a Lipschitz way on parameters. We recall that the Lipschitz norms defined in (1.12) satisfy the algebra property

$$|fg|^\lambda_{s,r} \le |f|^\lambda_{s,r}|g|^\lambda_{s,r} .$$

Lemma 8.1. *If $h(\cdot;\xi)$ is analytic in $\mathbb{T}^n_s$ and $|\psi|^\lambda_{s-\sigma} \le \sigma/2$ then*

$$g(x;\xi) := h(x+\psi(x;\xi);\xi) \quad \textit{satisfies} \quad |g|^\lambda_{s-\sigma} \le |h|^\lambda_s + \frac{2}{\sigma}|h|_s|\psi|^\lambda_{s-\sigma} \le 2|h|^\lambda_s \,. \tag{8.1}$$

If $\Psi \in \mathcal{E}_{s-\sigma}$ (see (5.4)) satisfies

$$\frac{|x_{00}|^\lambda_{s-\sigma}}{\sigma}\,,\ \frac{|y_{00}|^\lambda_{s-\sigma}}{r^2}\,,\ \frac{|y_{01}|^\lambda_{s-\sigma}}{r}\,,\ |y_{10}|^\lambda_{s-\sigma}\,,\ |y_{02}|^\lambda_{s-\sigma}\,,\ \frac{|w_{00}|^\lambda_{s-\sigma}}{\sigma r}\,,\ \frac{|w_{01}|^\lambda_{s-\sigma}}{\sigma} \le \frac{\delta}{16}\,, \tag{8.2}$$

with $0 \le \delta \le 1$, then, for all $H(\cdot;\xi)$ analytic in $D(s,r)$,

$$\tilde H(x,y,w;\xi) := H\big((x,y,w)+\Psi(x,y,w;\xi);\xi\big) \quad \textit{satisfies} \quad |\tilde H|^\lambda_{s-\sigma,r-\delta r} \le 2|H|^\lambda_{s,r}\,. \tag{8.3}$$

Proof. Since $h(\cdot;\xi)$ is analytic in $\mathbb{T}^n_s$, by Cauchy estimates,

$$|\psi|_{s-\sigma} \le \frac{\sigma}{2} \quad \Longrightarrow \quad |g|^{\rm lip}_{s-\sigma} \le |\partial_x h|_{s-\frac{\sigma}{2}}|\psi|^{\rm lip}_{s-\sigma} + |h|^{\rm lip}_s \le \frac{2}{\sigma}|h|_s|\psi|^{\rm lip}_{s-\sigma} + |h|^{\rm lip}_s$$

and (8.1) follows. The proof of (8.3) is similar. ■

We now estimate derivatives of the composed functions.

Lemma 8.2. *Given $H : D(s,r)\times\Pi \to \mathbb{C}$. There exists $c_0>0$ such that, if*

$$\Phi : D(\tilde s,\tilde r) \ni (x_+,y_+,w_+) \mapsto (x,y,w) \in D(s,r) \quad \textit{with } 0<\tilde r \le \frac{r}{2}\,,\ 0<\tilde s \le \frac{s}{2}\,,$$

and $\Phi = I + \Psi$ with $\Psi \in \mathcal{E}_{\tilde s}$ satisfies

$$\frac{|x_{00}|^\lambda_{\tilde s}}{s}\,,\ \frac{|y_{00}|^\lambda_{\tilde s}}{r^2}\,,\ \frac{|y_{01}|^\lambda_{\tilde s}}{r}\,,\ |y_{10}|^\lambda_{\tilde s}\,,\ |y_{02}|^\lambda_{\tilde s}\,,\ \frac{|w_{00}|^\lambda_{\tilde s}}{sr}\,,\ \frac{|w_{01}|^\lambda_{\tilde s}}{s} \le c_0\,, \tag{8.4}$$

then $\tilde H := H\circ\Phi$ is analytic on $D(\tilde s,\tilde r)$, $\forall \xi\in\Pi$, and

$$|\partial_{y_+^i w_+^j}\tilde H|^\lambda_{\tilde s,\tilde r} \le 3\Theta\,,\ \forall\, 2i+j=4\,, \quad \textit{where} \quad \Theta := \max\Big\{1, \sum_{2i+j=4}|\partial_{y^i w^j}H|^\lambda_{s,r}\Big\} \tag{8.5}$$

(we use the short notation $H\circ\Phi$ to mean $H(\cdot,\xi)\circ\Phi$, $\forall\xi\in\Pi$).

Proof. For c_0 small enough, conditions (8.4) imply (8.2) with

$$s \to \frac{3s}{4}\,,\ r \to \frac{3r}{4}\,,\ \sigma := \frac{3s}{4} - \tilde s \ge \frac{s}{4}\,,\ \delta := \frac{3r-4\tilde r}{3r} \ge \frac{1}{3}\,.$$

Then (8.3) implies, for c_0 small enough,

$$\begin{aligned} |\partial_{y_+ w_+^2}\tilde H|^\lambda_{\tilde s,\tilde r} \ \le\ & 2\Big[|\partial_{y^3}H|^\lambda_{\frac{3s}{4},\frac{3r}{4}}(|y_{01}|^\lambda_{\tilde s} + |y_{02}|^\lambda_{\tilde s} r)^2 + 2|\partial_{y^2w}H|^\lambda_{\frac{3s}{4},\frac{3r}{4}}(1+|w_{01}|^\lambda_{\tilde s})(|y_{01}|^\lambda_{\tilde s} + |y_{02}|^\lambda_{\tilde s} r) \\ & + 2|\partial_{y^2}H|^\lambda_{s,r}|y_{02}|^\lambda_{\tilde s} + |\partial_{yw^2}H|^\lambda_{s,r}(1+|w_{01}|^\lambda_{\tilde s})^2\Big](1+|y_{10}|^\lambda_{\tilde s}) \le 3\Theta\,, \end{aligned}$$

using that, by Cauchy estimates,

$$|\partial_{y^3}H|^\lambda_{\frac{3s}{4},\frac{3r}{4}} \le 16r^{-2}|\partial_{y^2}H|^\lambda_{s,r} \le 16r^{-2}\Theta\,, \quad |\partial_{y^2w}H|^\lambda_{\frac{3s}{4},\frac{3r}{4}} \le 4r^{-1}|\partial_{y^2}H|^\lambda_{s,r} \le 4r^{-1}\Theta\,.$$

The other estimates are analogous. ■

We conclude with a lemma on Fourier series. Fixed an integer $K>0$, we denote

$$T_K f(x;\xi) := \sum_{k\in\mathbb{Z}^n, |k|\le K} f_k(\xi)e^{{\rm i}k\cdot x} \quad \text{and} \quad T_K^\perp := I - T_K\,.$$

Lemma 8.3. *Let $f(\cdot;\xi)$ be analytic on $\mathbb{T}^n_s$. There is $C := C(n)$ such that, $\forall 0 \le \sigma \le s$, $K\sigma \ge 1$,*

$$K^{-n}e^{K\sigma}|T_K^\perp f|^\lambda_{s-\sigma}\,,\ \sigma K^{-n}e^{K\sigma}|T_K^\perp f'|^\lambda_{s-\sigma}\,,\ \sigma^n|T_K f|^\lambda_{s-\sigma}\,,\ \sigma^{n+1}|T_K f'|^\lambda_{s-\sigma} \le C|f|^\lambda_s\,. \tag{8.6}$$

PROOF. We have

$$|T_K^\perp f'|_{s-\sigma} \le \sum_{|k|>K}|k||f_k|e^{|k|(s-\sigma)} \le |f|_s\sum_{|k|>K}|k|e^{-|k|\sigma} \le |f|_s\sum_{l>K}4^n l^n e^{-l\sigma}$$

and the last sum is bounded by $C(n)\sigma^{-1}K^n e^{-K\sigma}$ if $K\sigma \ge 1$. The other estimates are analogous. ■
In the following we will always assume $K\sigma \ge 1$.

8.2 A class of symplectic transformations

We introduce the space of Hamiltonians

$$\begin{aligned}\mathcal{F}_s \ := \ \Big\{ & F(x;\xi) = F_{00}(x;\xi) + F_{01}(x;\xi)\cdot w + F_{10}(x;\xi)\cdot y + F_{02}(x;\xi)w\cdot w \\ & \text{where } F_{ij} \text{ are analytic and bounded on } \mathbb{T}^n_s \text{ and Lipschitz in } \xi\in\Pi\Big\}\,.\end{aligned} \tag{8.7}$$

Note that the terms that we want to eliminate from the perturbation through the KAM iteration have such a form. We shall also take "auxiliary" Hamiltonians in $\mathcal{F}_s$ whose time one flow generates the KAM symplectic transformations, see Lemma 8.9.

The next lemmata will be used to estimate the perturbation after the KAM step, see Lemma 8.11. The time one flow map generated by Hamiltonians in $\mathcal{F}_s$ has the form $I+\Psi$ with Ψ as in (5.4), see Lemma 8.6. Lemma 8.4 shows that $\mathcal{F}_s$ is closed under composition with such maps. We estimate the transformed map in a slightly smaller analytic strip for the convergence of the KAM iteration.

Lemma 8.4. (Composition) *If $F\in\mathcal{F}_s$, $\Psi\in\mathcal{E}_{s-\sigma}$, $0<\sigma\le s$, with $|x_{00}|^\lambda_{s-\sigma}\le\sigma/2$, then $S := F\circ(I+\Psi)\in\mathcal{F}_{s-\sigma}$ and*

$$\begin{aligned}S_{00} &= \tilde F_{00} + \tilde F_{10}\cdot y_{00} + \tilde F_{01}\cdot w_{00} + \tilde F_{02}w_{00}\cdot w_{00}\\ S_{01} &= (I+w_{01}^\intercal)\tilde F_{01} + y_{01}^\intercal\tilde F_{10} + 2(I+w_{01}^\intercal)\tilde F_{02}w_{00}\\ S_{10} &= (I+y_{10}^\intercal)\tilde F_{10}\\ S_{02} &= \tilde F_{10}\cdot y_{02} + (I+w_{01}^\intercal)\tilde F_{02}(I+w_{01})\end{aligned}$$

where $\tilde F_{ij} = \tilde F_{ij}(x_+) := F_{ij}\big(x_+ + x_{00}(x_+)\big)$. By (8.1), $|\tilde F_{ij}|^\lambda_{s-\sigma}\le 2|F_{ij}|^\lambda_s$.

It is a merely algebraic calculus that the space $\mathcal{F}_s$ is closed under the Poisson brackets (see (1.4)).

Lemma 8.5. (Poisson bracket) *Let $R,F\in\mathcal{F}_s$ then $G := \{R,F\}\in\mathcal{F}_{s'}$, $\forall 0<s'<s$, and*

$$\begin{aligned}G_{00} &= F_{10}\cdot R'_{00} - R_{10}\cdot F'_{00} - \mathrm{i}R_{01}\cdot JF_{01}\\ G_{01} &= F_{10}\cdot R'_{01} - R_{10}\cdot F'_{01} + 2\mathrm{i}F_{02}JR_{01} - 2\mathrm{i}R_{02}JF_{01}\\ G_{10} &= F_{10}\cdot R'_{10} - R_{10}\cdot F'_{10}\\ G_{02} &= F_{10}\cdot R'_{02} - R_{10}\cdot F'_{02} - 4\mathrm{i}R_{02}JF_{02}\,.\end{aligned}$$

Given $F\in\mathcal{F}_s$, we consider the associated Hamiltonian system (see (1.3))

$$\begin{cases}\dot x = F_{10}(x)\\ \dot y = -F'_{00}(x) - F'_{01}(x)w - F'_{10}(x)y - F'_{02}(x)w\cdot w\\ \dot w = -\mathrm{i}JF_{01}(x) - 2\mathrm{i}JF_{02}(x)w\end{cases} \tag{8.8}$$

with initial condition $(x^0, y^0, w^0) = (x_+, y_+, w_+)$. For all $\xi \in \Pi$, the hamiltonian flow at time t

$$X_F^t(\cdot;\xi) : (x_+, y_+, w_+) \mapsto (x^t, y^t, w^t)(x_+, y_+, w_+)$$

defines a symplectic diffeomorphism which is close to the identity for $0 \le t \le 1$ and F small. In the next lemma we estimate each component of these symplectic diffeomorphisms separately. These finer estimates are required by our approach. This is a difference with respect to [24].

Lemma 8.6. (Hamiltonian flow) *Let $0 < \sigma < s \le 1$ and $F \in \mathcal{F}_s$ satisfy, for some $\lambda \ge 0$,*

$$|F_{10}|_s^\lambda \le \sigma/12\,, \ |F_{02}|_s^\lambda \le 1/12\,. \tag{8.9}$$

Then, for all $t \in [0,1]$, $X_F^t = I + \Psi^t$ with $\Psi^t \in \mathcal{E}_{s-\sigma}$ satisfying

$$|x_{00}^t|_{s-\sigma}^\lambda \le 2|F_{10}|_s^\lambda\ ,\ |y_{00}^t|_{s-\sigma}^\lambda \le \frac{12}{\sigma}\Big(|F_{00}|_s^\lambda + 9(|F_{01}|_s^\lambda)^2\Big)\ ,\ |y_{10}^t|_{s-\sigma}^\lambda \le \frac{6}{\sigma}|F_{10}|_s^\lambda\,, \tag{8.10}$$

$$|y_{01}^t|_{s-\sigma}^\lambda \le \frac{36}{\sigma}|F_{01}|_s^\lambda\,,\ |y_{02}^t|_{s-\sigma}^\lambda \le \frac{27}{\sigma}|F_{02}|_s^\lambda\,,\ |w_{00}^t|_{s-\sigma}^\lambda \le 6|F_{01}|_s^\lambda\,,\ |w_{01}^t|_{s-\sigma}^\lambda \le 6|F_{02}|_s^\lambda\,.$$

Moreover, if, for $0 < \delta < 1$,

$$|F_{00}|_s \le \frac{\delta r^2 \sigma}{72}\,, \quad |F_{01}|_s \le \frac{\delta r \sigma}{216}\,, \quad |F_{10}|_s \le \frac{\delta\sigma}{24}\,, \quad |F_{02}|_s \le \frac{\delta\sigma}{108}\,, \tag{8.11}$$

then $X_F^t(\cdot;\xi) : D(s-\sigma, r-\delta r) \subseteq D(s,r)$, $\forall 0 \le t \le 1$, $\forall \xi \in \Pi$.

Proof. In the Appendix. ■

Finally we study the composition of two symplectic maps of the form $I + \Psi$ with $\Psi \in \mathcal{E}_s$. The symplectic transformation (5.7) of Theorem 5.1 is the composition of infinitely many maps of this form, see the iterative Lemma 8.17-$(S6)_\nu$.

Lemma 8.7. (Composition of diffeomorphisms) *Let $0 < s < \tilde{s}$, $\tilde{\Phi} = I + \tilde{\Psi}$ with $\tilde{\Psi} \in \mathcal{E}_{\tilde{s}}$, and $\Phi = I + \Psi$ with $\Psi \in \mathcal{E}_s$ satisfy $2|x_{00}|_{\tilde{s}}^\lambda/(\tilde{s} - s) \le \eta \le 1$. Then the composite map has the form*

$$\tilde{\Phi} \circ \Phi = I + \hat{\Psi} \quad \textit{with} \quad \hat{\Psi} \in \mathcal{E}_s \quad \textit{and}$$

$$\begin{aligned}
&|\hat{x}_{00} - x_{00}|_s \le (1+\eta)|\tilde{x}_{00}|_{\tilde{s}}\,, \quad |\hat{w}_{00} - w_{00}|_s \le (1+\eta)|\tilde{w}_{00}|_{\tilde{s}} + 2|\tilde{w}_{01}|_{\tilde{s}}|w_{00}|_s \\
&|\hat{w}_{01} - w_{01}|_s \le (1+\eta)|\tilde{w}_{01}|_{\tilde{s}}(1 + |w_{01}|_s) \\
&|\hat{y}_{00} - y_{00}|_s \le (1+\eta)|\tilde{y}_{00}|_{\tilde{s}} + 2|\tilde{y}_{01}|_{\tilde{s}}|w_{00}|_s + 2|\tilde{y}_{10}|_{\tilde{s}}|y_{00}|_s + 2|\tilde{y}_{02}|_{\tilde{s}}|w_{00}|_s^2 \\
&|\hat{y}_{01} - y_{01}|_s \le (1+\eta)|\tilde{y}_{01}|_{\tilde{s}}(1 + |w_{01}|_s) + 2|\tilde{y}_{10}|_{\tilde{s}}|y_{01}|_s + 4|\tilde{y}_{02}|_{\tilde{s}}|w_{00}|_s(1 + |w_{01}|_s) \\
&|\hat{y}_{10} - y_{10}|_s \le (1+\eta)|\tilde{y}_{10}|_{\tilde{s}}(1 + |y_{10}|_s) \\
&|\hat{y}_{02} - y_{02}|_s \le (1+\eta)|\tilde{y}_{02}|_{\tilde{s}}(1 + |w_{01}|_s)^2 + 2|\tilde{y}_{10}|_{\tilde{s}}|y_{02}|_s
\end{aligned} \tag{8.12}$$

where for brevity $|\cdot|_{\tilde{s}} := |\cdot|_{\tilde{s}}^\lambda$, $|\cdot|_s := |\cdot|_s^\lambda$.

Proof. We have $\hat{\Psi} - \Psi = \tilde{\Psi} \circ (I + \Psi)$. The estimate on $\hat{x}_{00}$ follows by $\hat{x}_{00}(x_+) - x_{00}(x_+) = \tilde{x}_{00}\big(x_+ + x_{00}(x_+)\big)$ and (8.1). All the other estimates follow analogously. ■

8.3 The KAM step

At the generic ν-th step we have an Hamiltonian $H^\nu = N^\nu + P^\nu$ like in (5.18). Both ω_ν, Ω_ν are Lipschitz in Π_ν with $|\omega_\nu|^{\rm lip} + |\Omega_\nu|_{-\delta_*}^{\rm lip} \le M_\nu$. We set

$$\Theta_\nu := \max\Big\{1, |P_{11}^\nu|_{s_\nu}^{\lambda_\nu}, |P_{03}^\nu|_{s_\nu}^{\lambda_\nu}, \sum_{2i+j=4} |\partial_y^i \partial_w^j P^\nu|_{s_\nu, r_\nu}^{\lambda_\nu}\Big\} \quad \text{with} \quad \lambda_\nu := \frac{\alpha_\nu}{M_\nu}\,. \tag{8.13}$$

We simplify notations in the next section dropping the index ν and writing "+" for $\nu+1$. So $P = P^\nu$, $P^+ = P^{\nu+1}$, etc.

The symplectic change of coordinates

We write

$$H = N + P = N + R + (P - R) \qquad \text{where} \qquad R := T_K P_{\leq 2} \tag{8.14}$$

and $P_{\leq 2}$ is defined in (1.8). Then we consider the homological equation

$$\{N, F\} + R = [R] \tag{8.15}$$

where

$$[R] := \hat{e} + \hat{\omega} \cdot y + \hat{\Omega} z \cdot \bar{z}\,, \quad \hat{e} := \langle P_{00} \rangle\,, \quad \hat{\omega} := \langle P_{10} \rangle\,, \quad \hat{\Omega} := \mathrm{diag}_{j \geq 1} \langle \partial^2_{z_j \bar{z}_j} P_{|y=0, w=0} \rangle \tag{8.16}$$

and $\langle \cdot \rangle$ denotes the average with respect to the angles.

Lemma 8.8. (homological equation) *Suppose that, uniformly on* Π,

$$|\omega(\xi) \cdot k + \Omega(\xi) \cdot l| \geq \alpha \frac{\langle l \rangle_d}{1 + |k|^\tau}\,, \quad \forall\, (k, l) \neq 0,\ |k| \leq K,\ |l| \leq 2\,. \tag{8.17}$$

Let $0 < \sigma < s$. *Then,* $\forall R \in \mathcal{F}_s$, *the equation (8.15) has a solution* $F \in \mathcal{F}_{s-\sigma}$ *satisfying* $[F] = 0$ *and*

$$|F_{ij}|^\lambda_{s-\sigma} \leq \frac{\mathtt{K} |P_{ij}|^\lambda_s}{\alpha \sigma^{2\tau + n + 1}}\,, \quad 0 \leq 2i + j \leq 2\,, \quad 0 \leq \lambda \leq \frac{\alpha}{M}\,, \tag{8.18}$$

with $\mathtt{K} := \mathtt{K}(n, \tau) \geq 1$. *We can take* $\mathtt{K} = (\tau + n)^{c(\tau + n)}$ *for some absolute constant* $c > 0$.

Proof. The proof is given in [24], Lemmata 1-2 with the only difference that (8.17) holds for every k. The truncation $|k| \leq K$ does not affect the estimates, since $T_K P_{ij}$ and, therefore, F_{ij} are Fourier polynomials of order K. ■

By Lemma 8.8 and 8.6 we deduce:

Lemma 8.9. (symplectic map) *There exist* $C_0 := C_0(n, \tau) > 1$ *large enough -we can take* $C_0 := \mathtt{K}^c$ *for some absolute constant* $c > 0$ *with* $\mathtt{K}$ *defined in Lemma 8.8- such that, if*

$$\frac{|P_{00}|^\lambda_s}{r^2}\,, \quad \frac{|P_{01}|^\lambda_s}{r}\,, \quad |P_{10}|^\lambda_s\,, \quad |P_{02}|^\lambda_s \leq \frac{\delta \alpha \sigma^\beta}{16 C_0}\,, \tag{8.19}$$

where

$$\beta := 2\tau + n + 2\,, \tag{8.20}$$

$0 < 2\sigma < s < 1$, $0 < \delta < 1$, $0 \leq \lambda \leq \alpha/M$, *the symplectic maps*

$$\Phi^t = I + \Psi^t := X^t_F \ :\ D(s - 2\sigma, r - \delta r) \to D(s - \sigma, r - \delta r/2) \tag{8.21}$$

are well defined $\forall t \in [0, 1]$, *and* $\Psi^t \in \mathcal{E}_{s-2\sigma}$ *satisfy*

$$\begin{aligned}
&|x^t_{00}|^\lambda_{s-2\sigma} \leq C_0 \frac{|P_{10}|^\lambda_s}{\alpha \sigma^{\beta-1}}\,, \quad |y^t_{00}|^\lambda_{s-2\sigma} \leq C_0 \frac{|P_{00}|^\lambda_s}{2\alpha\sigma^\beta} + C_0 \frac{(|P_{01}|^\lambda_s)^2}{2\alpha^2 \sigma^{2\beta-1}}\,, \\
&|y^t_{10}|^\lambda_{s-2\sigma} \leq C_0 \frac{|P_{10}|^\lambda_s}{\alpha \sigma^\beta}\,, \quad |y^t_{01}|^\lambda_{s-2\sigma} \leq C_0 \frac{|P_{01}|^\lambda_s}{\alpha \sigma^\beta}\,, \quad |y^t_{02}|^\lambda_{s-2\sigma} \leq C_0 \frac{|P_{02}|^\lambda_s}{\alpha \sigma^\beta}\,, \\
&|w^t_{00}|^\lambda_{s-2\sigma} \leq C_0 \frac{|P_{01}|^\lambda_s}{\alpha \sigma^{\beta-1}}\,, \quad |w^t_{01}|^\lambda_{s-2\sigma} \leq C_0 \frac{|P_{02}|^\lambda_s}{\alpha \sigma^{\beta-1}}\,.
\end{aligned} \tag{8.22}$$

Note that (8.19)-(8.22) imply (8.2) (with $| \cdot |^\lambda_{s-2\sigma}$ instead of $| \cdot |^\lambda_{s-\sigma}$).

The Hamiltonian transformed under the symplectic map $\Phi^+ := X^1_F$ defined in (8.21) is

$$H^+ := H \circ \Phi^+ = N + \hat{N} + \int_0^1 \{(1 - t)\hat{N} + tR, F\} \circ X^t_F \, dt \ + (P - R) \circ \Phi^+ =: N^+ + P^+ \tag{8.23}$$

where $N^+ := N + \hat{N}$ and $\hat{N} := [R]$ is defined in (8.16).

The new normal form N^+

We now estimate $N^+ := N + \hat{N}$ where $\hat{N} := \hat{e} + \hat{\omega} \cdot y + \hat{\Omega} z \cdot \bar{z}$. We identify $\hat{\Omega}$ with the vector

$$\hat{\Omega} = (\hat{\Omega}_i)_{i \geq n+1}\,, \quad \hat{\Omega}_i := \langle \partial^2_{z_i \bar{z}_i} P_{|y=0,w=0} \rangle\,.$$

Lemma 8.10. $|\hat{\omega}| \leq |P_{10}|_s$, $|\hat{\omega}|^{\rm lip} \leq |P_{10}|^{\rm lip}_s$, $|\hat{\Omega}|_{\bar{p}-p} \leq |P_{02}|_s$, $|\hat{\Omega}|^{\rm lip}_{\bar{p}-p} \leq |P_{02}|^{\rm lip}_s$ *and*

$$|\hat{\omega} \cdot k + \hat{\Omega} \cdot l| \leq |P_{10}|_s |k| + 2|P_{02}|_s \langle l \rangle_d\,, \quad \forall (k,l) \in \mathbb{Z}^n \times \mathbb{Z}^\infty\,. \tag{8.24}$$

Proof. We have $\hat{\Omega}_j = (\langle P_{02} \rangle e_j, e_j)_p$ where $(\cdot,\cdot)_p$ and e_j (respectively $(\cdot,\cdot)_{\bar{p}}$ and $\bar{e}_j$) denote the scalar product and the j-th element of the basis in $\ell^{a,p}_b$ (respectively $\ell^{a,\bar{p}}_b$). We have $\bar{e}_i = i^{p-\bar{p}} e_j$ and, if $u \in \ell^{a,\bar{p}}_b$, $u = \sum_i \bar{u}_i \bar{e}_i = \sum_i u_i e_i$, then $u_i = i^{p-\bar{p}} \bar{u}_i$. Denoting $u := \langle P_{02} \rangle e_i$, we get

$$i^{\bar{p}-p} |\hat{\Omega}_i| = i^{\bar{p}-p} |(u, e_i)_p| = i^{\bar{p}-p} |u_i| = |\bar{u}_i| = |(u, \bar{e}_i)_{\bar{p}}| \leq \|u\|_{a,\bar{p}} \leq |P_{02}|_s$$

(recall that $|P_{02}|_s = \sup_{x \in \mathbb{T}_s} \|P_{02}(x)\|_{\mathcal{L}(\ell^{a,p}_b, \ell^{a,\bar{p}}_b)}$) implying $|\hat{\Omega}|_{\bar{p}-p} \leq |P_{02}|_s$. Similarly $|\hat{\omega}| \leq |P_{10}|_s$. Then

$$|\hat{\omega} \cdot k + \hat{\Omega} \cdot l| \leq |\hat{\omega}||k| + |\hat{\Omega}|_{-\delta_*} |l|_{\delta_*} \leq |\hat{\omega}||k| + |\hat{\Omega}|_{\bar{p}-p} 2\langle l \rangle_d \leq |P_{10}|_s |k| + 2|P_{02}|_s \langle l \rangle_d$$

using (3.3) and $|l|_{\delta_*} \leq |l|_{d-1} \leq 2\langle l \rangle_d$, $\forall |l| \leq 2$. The same estimates holds for $|\cdot|^{\rm lip}$. ■

The new perturbation P^+

Notation. For the rest of this section, $A \lessdot B$ means that $A \leq \mathtt{K}^c B$ where $\mathtt{K}$ is defined in Lemma 8.8 and $c > 0$ is some absolute constant.

By (8.23), and since $\hat{N} = [R]$, we have to estimate $P^+ = P^* + \tilde{P}$ where

$$P^* := \int_0^1 \{(1-t)[R] + tR, F\} \circ X^t_F \, dt\,, \qquad \tilde{P} := (P - R) \circ \Phi^+\,.$$

We estimate P^* in Lemma 8.11 and $\tilde{P}$ in Lemma 8.13.

We introduce the rescaled quantities

$$a := \frac{|P_{00}|^\lambda_s}{r^2 \alpha^{p_a}}\,, \quad b := \frac{|P_{01}|^\lambda_s}{r \alpha^{p_b}}\,, \quad c := \frac{|P_{10}|^\lambda_s}{\alpha}\,, \quad d := \frac{|P_{02}|^\lambda_s}{\alpha} \tag{8.25}$$

where p_a, p_b are defined in (5.11). Since $p_a, p_b \geq 1$, if

$$a, b, c, d \leq \frac{\delta \sigma^\beta}{16 C_0} \tag{8.26}$$

(the constant C_0 is defined in Lemma 8.9), then (8.19) and, so, (8.22) hold.

Note that the P^*_{ij} in (8.27), $0 \leq 2i + j \leq 2$, are "quadratic" in the variables a, b, c, d (i.e. P_{ij}).

Lemma 8.11. $P^* := \int_0^1 \{(1-t)[R] + tR, F\} \circ X^t_F \, dt \in \mathcal{F}_{s-2\sigma}$ *and*

$$\begin{aligned} |P^*_{00}|^\lambda_{s-2\sigma} &\lessdot \sigma^{2-6\beta} r^2 \alpha^{p_a} (ac + b^2)\,, \quad |P^*_{01}|^\lambda_{s-2\sigma} \lessdot \sigma^{2-6\beta} r \alpha^{p_b} b(c+d)\,, \\ |P^*_{10}|^\lambda_{s-2\sigma} &\lessdot \sigma^{2-6\beta} \alpha c^2\,, \qquad\qquad\quad |P^*_{02}|^\lambda_{s-2\sigma} \lessdot \sigma^{2-6\beta} \alpha d(c+d)\,, \end{aligned} \tag{8.27}$$

where β is defined in (8.20).

Proof. We estimate $\int_0^1 t\{R,F\}\circ X_F^t\,dt$. The term $\int_0^1 (1-t)\{[R],F\}\circ X_F^t\,dt$ is analogous. The statement follows by Lemma 8.4 (with $s\to s-\frac{3\sigma}{2}$, $s-\sigma\to s-2\sigma$), Lemma 8.5 (with $G=\{R,F\}$), Lemma 8.3, and (8.1), (8.6), (8.18), (8.19), (8.25) (8.26). Indeed, using $r,\alpha<1$ and $2p_b\geq p_a+1$, we get

$$
\begin{aligned}
|P^*_{00}|^\lambda_{s-2\sigma} &\lessdot |G_{00}|^\lambda_{s-\frac{3\sigma}{2}} + |G_{10}|^\lambda_{s-\frac{3\sigma}{2}}|y_{00}|^\lambda_{s-2\sigma} + |G_{01}|^\lambda_{s-\frac{3\sigma}{2}}|w_{00}|^\lambda_{s-2\sigma} + |G_{02}|^\lambda_{s-\frac{3\sigma}{2}}(|w_{00}|^\lambda_{s-2\sigma})^2\\
&\lessdot |F_{10}|^\lambda_{s-\sigma}|T_KP'_{00}| + |T_KP_{10}||F'_{00}| + |F_{01}||T_KP_{01}|\\
&\quad+\Big(|T_KP'_{10}||F_{10}| + |T_KP_{10}||F'_{10}|\Big)\sigma^{1-2\beta}r^2\alpha^{p_a-1}(a+b^2)\\
&\quad+\Big(|F_{10}||T_KP'_{01}| + |T_KP_{10}||F'_{01}| + |T_KP_{01}||F_{02}| + |T_KP_{02}||F_{01}|\Big)\sigma^{1-2\beta}r\alpha^{p_b-1}b\\
&\quad+\Big(|F_{10}||T_KP'_{02}| + |T_KP_{10}||F'_{02}| + |T_KP_{02}||F_{02}|\Big)\sigma^{2-4\beta}r^2\alpha^{2p_b-2}b^2\\
&\lessdot \alpha^{-1}\sigma^{2-6\beta}\Big[|P_{00}|^\lambda_s|P_{10}| + |P_{01}|^2 + |P_{10}|^2r^2\alpha^{p_a-1}(a+b^2)\\
&\quad+(|P_{10}|+|P_{02}|)|P_{01}|(1+r\alpha^{p_b-1}b) + |P_{02}|^2r^2\alpha^{2p_b-2}b^2\Big]\\
&\lessdot \alpha^{-1}\sigma^{2-6\beta}\Big[r^2\alpha^{p_a+1}ac + r^2\alpha^{p_b}b^2 + r^2\alpha^{p_a+1}(a+b^2)c^2\\
&\quad+r^2\alpha^{2p_b}(c+d)b^2 + r^2\alpha^{2p_b-2}b^2d^2\Big] \lessdot \sigma^{2-6\beta}r^2\alpha^{p_a}ac
\end{aligned}
$$

where in the second term of the chain of inequalities all the norms are $|\cdot|^\lambda_{s-\sigma}$, in the third term all the norms are $|\cdot|^\lambda_s$, and we used Cauchy inequalities. Next

$$
\begin{aligned}
|P^*_{01}|^\lambda_{s-2\sigma} &\lessdot |G_{01}|^\lambda_{s-\frac{3\sigma}{2}} + |G_{10}|^\lambda_{s-\frac{3\sigma}{2}}|y_{01}|^\lambda_{s-2\sigma} + |G_{02}|^\lambda_{s-\frac{3\sigma}{2}}|w_{00}|^\lambda_{s-2\sigma}\\
&\lessdot |F_{10}|^\lambda_{s-\sigma}|T_KP'_{01}| + |T_KP_{10}||F'_{01}| + |F_{02}||T_KP_{01}| + |T_KP_{02}||F_{01}|\\
&\quad+\big(|T_KP'_{10}||F_{10}| + |T_KP_{10}||F'_{10}| + |F_{10}||T_KP'_{02}| + |T_KP_{10}||F'_{02}| + |T_KP_{02}||F_{02}|\big)\times\\
&\quad\times\sigma^{1-2\beta}r\alpha^{p_b-1}b\\
&\lessdot \sigma^{1-4\beta}r\alpha^{p_b}[b(c+d) + bc^2 + bd(c+d)] \lessdot \sigma^{1-4\beta}r\alpha^{p_b}b(c+d)
\end{aligned}
$$

where in the second line all the norms are $|\cdot|^\lambda_{s-\sigma}$. Moreover

$$
|P^*_{10}|^\lambda_{s-2\sigma} \lessdot |G_{10}|^\lambda_{s-\frac{3\sigma}{2}} \lessdot |F_{10}|^\lambda_{s-\sigma}|T_KP'_{10}|^\lambda_{s-\sigma} + |F'_{10}|^\lambda_{s-\sigma}|T_KP_{10}|^\lambda_{s-\sigma} \lessdot \sigma^{-2\beta}\alpha c^2\,.
$$

Finally

$$
\begin{aligned}
|P^*_{02}|^\lambda_{s-2\sigma} &\lessdot |G_{10}|^\lambda_{s-\frac{3\sigma}{2}}|y_{02}|^\lambda_{s-2\sigma} + |G_{02}|^\lambda_{s-\frac{3\sigma}{2}}\\
&\lessdot \big(|F_{10}|^\lambda_{s-\sigma}|T_KP'_{10}| + |T_KP_{10}||F'_{10}|\big)\sigma^{1-2\beta}d\\
&\quad+|T_KP'_{02}||F_{10}| + |F'_{02}||T_KP_{10}| + |T_KP_{02}||F_{02}|\\
&\lessdot \sigma^{1-4\beta}\alpha(c^2d + cd + d^2) \lessdot \sigma^{1-4\beta}\alpha(c+d)d
\end{aligned}
$$

where in the second line all the norms are $|\cdot|^\lambda_{s-\sigma}$. ■

We define the higher order terms of the perturbation

$$
P_4 := \sum_{2i+j\geq 4} P_{ij}(x)y^iw^j \qquad \text{so that} \qquad P = P_{\leq 2} + P_{11}yw + P_{03}w^3 + P_4 \tag{8.28}
$$

($P_{\leq 2}$ was defined in (1.8)). Note that $\partial_y^i\partial_w^jP = \partial_y^i\partial_w^jP_4$ if $2i+j=4$. We also define

$$
\Phi_{00} := \Phi^+_{|\{y_+=0,w_+=0\}} = \big(x_+ + x_{00}(x_+;\xi), y_{00}(x_+;\xi), w_{00}(x_+;\xi)\big)\,.
$$

By Lemma 8.9, $\Phi_{00}: D(s-2\sigma)\to D(s-\sigma, r-\delta r/2)$, $\forall\xi\in\Pi$.

Lemma 8.12. *We have*

$$|P_4 \circ \Phi_{00}| \lessdot \Theta\big(\delta^{-1}|y_{00}|^2 + \delta^{-1}|y_{00}||w_{00}|^2 + |w_{00}|^4\big)$$
$$|(\partial_y P_4) \circ \Phi_{00}| \lessdot \Theta\big(\delta^{-1}|y_{00}| + |w_{00}|^2\big)\,,\ |(\partial^2_{yw} P_4) \circ \Phi_{00}| \lessdot \Theta\big((\delta r)^{-1}|y_{00}| + |w_{00}|\big)$$
$$|(\partial_w P_4) \circ \Phi_{00}| \lessdot \Theta\big((\delta r)^{-1}|y_{00}|^2 + \delta^{-1}|y_{00}||w_{00}| + |w_{00}|^3\big)$$
$$|(\partial^2_{ww} P_4) \circ \Phi_{00}| \lessdot \Theta\big(\delta^{-1}|y_{00}| + |w_{00}|^2\big)\,,\ |(\partial^3_{www} P_4) \circ \Phi_{00}| \lessdot \Theta\big((\delta r)^{-1}|y_{00}| + |w_{00}|\big)$$
$$|(\partial^3_{yyw} P_4) \circ \Phi_{00}| \lessdot \Theta(\delta r)^{-1}\,,\quad |(\partial^3_{yyy} P_4) \circ \Phi_{00}| \lessdot \Theta(\delta r)^{-2}$$

where all the norms $|\ | := |\ |^\lambda_{s-2\sigma}$ *and* Θ *is defined in (5.5).*

PROOF. We only prove the estimate for $\partial^3_w P_4 \circ \Phi_{00}$ where, for brevity, $\partial^3_w := \partial_{www}$. For all $(x, y, w; \xi) \in D(s, r - \delta r/2) \times \Pi$, since $\partial^3_w P_4(x, 0, 0; \xi) = 0$ (by definition of P_4), we have

$$\|\partial^3_w P_4(x, y, w; \xi)\| = \|\partial^3_w P_4(x, y, w; \xi) - \partial^3_w P_4(x, 0, 0; \xi)\|$$
$$\le \sup_{0\le t\le 1} \|\partial^3_w \partial_y P_4(x, ty, tw; \xi)\||y| + \sup_{0\le t\le 1} \|\partial^4_w P_4(x, ty, tw; \xi)\|\|w\|_{a,p} \le \Theta((\delta r)^{-1}|y| + \|w\|_{a,p})$$

($\|\cdot\|$ denote the operatorial norm) because, by Cauchy estimates, and the definition of Θ,

$$|\partial^3_w \partial_y P_4|_{s,(1-\frac{\delta}{2})r} \lessdot (\delta r)^{-1}|\partial^2_w \partial_y P_4|_{s,r} \lessdot \Theta(\delta r)^{-1}\,. \tag{8.29}$$

Then $\forall |y| < (r - \delta r/2)^2$, $\|w\|_{a,p} < r - \delta r/2$,

$$|\partial^3_w P_4(\cdot, y, w; \cdot)|_s\,, \sigma|\partial^3_w \partial_x P_4(\cdot, y, w; \cdot)|_{s-\sigma} \lessdot \Theta((\delta r)^{-1}|y| + \|w\|_{a,p})\,. \tag{8.30}$$

Then, since Lemma 8.9 implies $|x_{00}|^\lambda_{s-2\sigma} \le \sigma/16$, $|y_{00}| < (r - \delta r/2)^2$, $|w_{00}|_{s-2\sigma} < r - \delta r/2$,

$$|\partial^3_w P_4 \circ \Phi_{00}|_{s-2\sigma} \le \sup_{x\in\mathbb{T}^n_s,\, \zeta\in\Pi} |\partial^3_w P_4(x, y_{00}(x_+; \xi), w_{00}(x_+; \xi); \zeta)| \lessdot \Theta\Big(\frac{|y_{00}|_{s-2\sigma}}{\delta r} + |w_{00}|_{s-2\sigma}\Big)\,.$$

With similar estimates $|\partial^3_w P_4 \circ \Phi_{00}|^{\text{lip}}_{s,r} \lessdot \Theta\lambda^{-1}(|y_{00}|^\lambda_{s-2\sigma}(\delta r)^{-1} + |w_{00}|^\lambda_{s-2\sigma})$. ■

We now estimate $\tilde P := (P - R) \circ \Phi^+$. Note the "linear" term in the variables a, b, c, d.

Lemma 8.13. $\tilde P := (P - R) \circ \Phi^+ = (P_{11}yw + P_{03}w^3 + P_4 + T^\perp_K P_{\le 2}) \circ \Phi^+ \in \mathcal{F}_{s-2\sigma}$ *and*

$$\begin{aligned}
\sigma^{8\beta-4}|\tilde P_{00}|^\lambda_{s-2\sigma} &\lessdot |P_{11}|^\lambda_s r^3\alpha^{p_a+p_b-2}(ab + b^3) + |P_{03}|^\lambda_s r^3\alpha^{3p_b-3}b^3 + \Theta\delta^{-1}r^4\alpha^{2p_a-2}(a^2 + b^4)\\
&\quad + K^n e^{-K\sigma} r^2\alpha^{p_a}(a + b^2)\\
\sigma^{6\beta-3}|\tilde P_{01}|^\lambda_{s-2\sigma} &\lessdot |P_{11}|^\lambda_s r^2\alpha^{p_a-1}(a + b^2) + |P_{03}|^\lambda_s r^2\alpha^{2p_b-2}b^2 + \Theta\delta^{-1}r^3\alpha^{p_a+p_b-2}(a + b^2)b\\
&\quad + K^n e^{-K\sigma} r\alpha^{p_b} b\\
\sigma^{4\beta-2}|\tilde P_{10}|^\lambda_{s-2\sigma} &\lessdot |P_{11}|^\lambda_s r\alpha^{p_b-1}b + \Theta\delta^{-1}r^2\alpha^{p_a-1}(a + b^2) + K^n e^{-K\sigma}\alpha c\\
\sigma^{4\beta-2}|\tilde P_{02}|^\lambda_{s-2\sigma} &\lessdot (|P_{11}|^\lambda_s + |P_{03}|^\lambda_s) r\alpha^{p_b-1}b + \Theta\delta^{-1}r^2\alpha^{p_a-1}(a + b^2) + K^n e^{-K\sigma}\alpha d\\
\sigma^{2\beta-1}|\tilde P_{11} - P_{11}|^\lambda_{s-2\sigma} &\lessdot |P_{11}|^\lambda_s(c + d) + \Theta\delta^{-1}r\alpha^{p_a-1}(a + b)\\
\sigma^{2\beta-1}|\tilde P_{03} - P_{03}|^\lambda_{s-2\sigma} &\lessdot (|P_{11}|^\lambda_s + |P_{03}|^\lambda_s)d + \Theta\delta^{-1}r\alpha^{p_a-1}(a + b)\,,
\end{aligned}$$

where β *is defined in (8.20).*

PROOF. Let for simplicity $\Phi^+ := \Phi$. We have

$$\tilde P_{00} = \big((P - R) \circ \Phi\big)_{|y_+=0,w_+=0}\,,\quad \tilde P_{01} = \partial_{w_+}\big((P - R) \circ \Phi\big)_{|y_+=0,w_+=0}\,, \tag{8.31}$$
$$\tilde P_{10} = \partial_{y_+}\big((P - R) \circ \Phi\big)_{|y_+=0,w_+=0}\,,\quad \tilde P_{02} = \frac{1}{2}\partial^2_{w_+w_+}\big((P - R) \circ \Phi\big)_{|y_+=0,w_+=0}\,,$$
$$\tilde P_{11} = \partial^2_{y_+w_+}\big((P - R) \circ \Phi\big)_{|y_+=0,w_+=0}\,,\quad \tilde P_{03} = \frac{1}{6}\partial^3_{w_+w_+w_+}\big((P - R) \circ \Phi\big)_{|y_+=0,w_+=0}\,.$$

For brevity we set $|\cdot| := |\cdot|_s^\lambda$, $|\cdot|_* := |\cdot|_{s-2\sigma}^\lambda$. The $P_{ij}^\perp(x^+) := T_K^\perp P_{ij}(x^+ + x_{00}(x_+))$, $0 \le 2i+j \le 2$, satisfy, since $|x_{00}|_{s-2\sigma}^\lambda \le \delta\sigma/16$ (by Lemma 8.9),

$$|P_{ij}^\perp|_* \overset{(8.1)}{\le} |T_K^\perp P_{ij}|_{s-\sigma} \overset{(8.6)}{\lessdot} K^n e^{-K\sigma}|P_{ij}|\,. \tag{8.32}$$

All the following estimates are a consequence of (8.31), the definition of P_4 in (8.28), Lemmata 8.12 and 8.9, (8.25), (8.26), (8.32) and $2p_b \ge p_a + 1$. Setting $Q := P_4 + T_K^\perp P_{\le 2}$ we have

$$\begin{aligned}
|\tilde P_{00}|_* &\lessdot |P_{11}||y_{00}|_*|w_{00}|_* + |P_{03}||w_{00}|_*^3 + |Q\circ\Phi_{00}|_* \\
&\lessdot |P_{11}|\sigma^{2-4\beta}r^3\alpha^{p_a+p_b-2}(ab+b^3) + |P_{03}|\sigma^{3-3\beta}r^3\alpha^{3p_b-3}b^3 \\
&\quad +\Theta\left(\delta^{-1}|y_{00}|_*^2 + \delta^{-1}|y_{00}|_*|w_{00}|_*^2 + |w_{00}|_*^4\right) \\
&\quad +|P_{00}^\perp|_* + |P_{01}^\perp|_*|w_{00}|_* + |P_{10}^\perp|_*|y_{00}|_* + |P_{02}^\perp|_*|w_{00}|_*^2 \\
&\lessdot |P_{11}|\sigma^{2-4\beta}r^3\alpha^{p_a+p_b-2}(ab+b^3) + |P_{03}|\sigma^{3-3\beta}r^3\alpha^{3p_b-3}b^3 \\
&\quad +\Theta\delta^{-1}\sigma^{4-8\beta}r^4\big(\alpha^{2p_a-2}(a+b^2)^2 + \alpha^{4p_b-4}b^4\big) \\
&\quad +K^n e^{-K\sigma}\sigma^{2-4\beta}r^2\alpha^{p_a}\big(a+b^2+c(a+b^2)+db^2\big)\,.
\end{aligned}$$

Next

$$\begin{aligned}
|\tilde P_{01}|_* &\lessdot |P_{11}|\big(|y_{01}|_*|w_{00}|_* + |I+w_{01}|_*|y_{00}|_*\big) + |P_{03}||w_{00}|_*^2|I+w_{01}|_* + |\partial_{w_+}(Q\circ\Phi)_{|y_+=0,w_+=0}|_* \\
&\lessdot |P_{11}|\sigma^{2-4\beta}r^2\alpha^{p_a-1}(a+b^2) + |P_{03}|\sigma^{2-4\beta}r^2\alpha^{2p_b-2}b^2 + |(\partial_y Q)\circ\Phi_{00}|_*|y_{01}|_* \\
&\quad +|(\partial_w Q)\circ\Phi_{00}|_*|I+w_{01}|_* \\
&\lessdot |P_{11}|\sigma^{2-4\beta}r^2\alpha^{p_a-1}(a+b^2) + |P_{03}|\sigma^{2-4\beta}r^2\alpha^{2p_b-2}b^2 + \Theta\big(\delta^{-1}|y_{00}|_* + |w_{00}|_*^2\big)|y_{01}|_* \\
&\quad +\Theta\big((\delta r)^{-1}|y_{00}|_*^2 + \delta^{-1}|y_{00}|_*|w_{00}|_* + |w_{00}|_*^3\big) \\
&\quad +|P_{01}^\perp|_*|I+w_{01}|_* + |P_{10}^\perp|_*|y_{01}|_* + |P_{02}^\perp|_*|w_{00}|_*|I+w_{01}|_* \\
&\lessdot |P_{11}|\sigma^{2-4\beta}r^2\alpha^{p_a-1}(a+b^2) + |P_{03}|r^2\sigma^{2-4\beta}\alpha^{2p_b-2}b^2 + \Theta\delta^{-1}\sigma^{3-6\beta}r^3\alpha^{p_a+p_b-2}(a+b^2)b \\
&\quad +K^n e^{-K\sigma}\sigma^{1-2\beta}r\alpha^{p_b}b\,.
\end{aligned}$$

Moreover

$$\begin{aligned}
|\tilde P_{10}|_* &\lessdot |P_{11}||w_{00}|_*|I+y_{10}|_* + |\partial_{y_+}(Q\circ\Phi)_{|y_+=0,w_+=0} \\
&\lessdot |P_{11}|\sigma^{1-2\beta}r\alpha^{p_b-1}b + \Theta\big(\delta^{-1}|y_{00}|_* + |w_{00}|_*^2\big) + |P_{10}^\perp|_*|I+y_{10}|_* \\
&\lessdot |P_{11}|\sigma^{1-2\beta}r\alpha^{p_b-1}b + \Theta\delta^{-1}\sigma^{2-4\beta}r^2\alpha^{p_a-1}(a+b^2) + K^n e^{-K\sigma}\alpha c\,.
\end{aligned}$$

By (8.22) and (8.26) we have $|y_{01}|_* \lessdot \delta r$ and then

$$\begin{aligned}
|\tilde P_{02}|_* &\lessdot |P_{11}|\big(|y_{02}|_*|w_{00}|_* + |I+w_{01}|_*|y_{01}|_*\big) + |P_{03}||w_{00}|_*|I+w_{01}|_*^2 \\
&\quad +|\partial^2_{w_+w_+}(Q\circ\Phi)_{|y_+=0,w_+=0}|_* \\
&\lessdot (|P_{11}|+|P_{03}|)\sigma^{2-4\beta}r\alpha^{p_b-1}b + |(\partial^2_{yy}Q)\circ\Phi_{00}|_*|y_{01}|_*^2 + |(\partial^2_{yw}Q)\circ\Phi_{00}|_*|I+w_{01}|_*|y_{01}|_* \\
&\quad +|(\partial_y Q)\circ\Phi_{00}|_*|y_{02}|_* + |(\partial^2_{ww}Q)\circ\Phi_{00}|_*|I+w_{01}|_*^2 \\
&\lessdot (|P_{11}|+|P_{03}|)\sigma^{2-4\beta}r\alpha^{p_b-1}b + \Theta\big((\delta r)^{-1}|y_{00}|_* + |w_{00}|_*\big)|y_{01}|_* \\
&\quad +\Theta\big(\delta^{-1}|y_{00}|_* + |w_{00}|_*^2\big)|y_{02}|_* + \Theta\big(\delta^{-1}|y_{00}|_* + |w_{00}|_*^2\big) + |P_{10}^\perp|_*|y_{02}|_* + |P_{02}^\perp|_*|I+w_{01}|_*^2 \\
&\lessdot (|P_{11}|+|P_{03}|)\sigma^{2-4\beta}r\alpha^{p_b-1}b + \Theta\delta^{-1}\big(|y_{01}|_*^2 + |y_{00}|_* + |w_{00}|_*|y_{01}|_* + |w_{00}|_*^2\big) \\
&\quad +|P_{10}^\perp|_*|y_{02}|_* + |P_{02}^\perp|_* \\
&\lessdot (|P_{11}|+|P_{03}|)\sigma^{2-4\beta}r\alpha^{p_b-1}b + \Theta\delta^{-1}\sigma^{2-4\beta}r^2\alpha^{p_a-1}(a+b^2) + K^n e^{-K\sigma}\sigma^{1-2\beta}\alpha d\,.
\end{aligned}$$

The estimates of $|\tilde P_{11} - P_{11}|_*$ and $|\tilde P_{03} - P_{03}|_*$ follow in the same way. ■

Recollecting the previous informations we state the following key lemma of the KAM step.

Lemma 8.14. (KAM step) *Assume (8.26). Then, $\forall \xi \in \Pi$ satisfying (8.17), there is a symplectic map*

$$\Phi^+(\cdot;\xi) : D(s-2\sigma, r-\delta r) \to D(s-\sigma, r) \quad \text{with} \quad 0 < 2\sigma < s\,,\ 0<\delta<1\,,$$

satisfying (8.22), such that

$$H^+ := H \circ \Phi^+ = N^+ + P^+ = (N+\hat{N}) + P^+ = (N+[P]) + P^+$$

and $P^+ = P^* + \tilde{P}$ satisfies the estimates of Lemmata 8.11 and 8.13.

We define a_+, b_+, c_+, d_+ like a, b, c, d in (8.25), with $P^+_{ij}, s_+ := s-2\sigma, \alpha_+, r_+$ instead of P_{ij}, s, α, r.

Lemma 8.15. *Assume (8.26), $\Theta r^2 \le 18\alpha$ and $|P_{11}|^\lambda_s \le 9\alpha^{p_e}/r$, $|P_{03}|^\lambda_s \le 9\alpha^{p_f}/r$, where*

$$p_e := 3 - p_a - p_f = \begin{cases} 1/2 & \text{if (H1)} \\ 5/4 & \text{if (H2)} \\ 1 & \text{if (H3)} \end{cases} \quad \text{and} \quad p_f := \begin{cases} 1/2 & \text{if (H1) or (H2)} \\ 1 & \text{if (H3)}\,. \end{cases} \tag{8.33}$$

We have that

$$\begin{aligned} a_+ &\le C_1(ac + b^2 + a^2 + K^n e^{-K\sigma} a)/\delta\sigma^{\tilde\beta} \\ b_+ &\le C_1(a + b^2 + bc + bd + K^n e^{-K\sigma} b)/\delta\sigma^{\tilde\beta} \\ c_+ &\le C_1(b + c^2 + a + K^n e^{-K\sigma} c)/\delta\sigma^{\tilde\beta} \\ d_+ &\le C_1(b + cd + d^2 + a + K^n e^{-K\sigma} d)/\delta\sigma^{\tilde\beta} \end{aligned} \tag{8.34}$$

where $\tilde\beta := 16\tau + 8n + 12$ and $C_1 = \mathtt{K}^c$ for some absolute constant $c > 0$ ($\mathtt{K}$ defined in Lemma 8.8).

Proof. By Lemma 8.14 (see the estimates of Lemmata 8.11 and 8.13), $\tilde\beta = 8\beta - 4$, we get

$$\begin{aligned} \sigma^{\tilde\beta} a_+ &\lessdot ac + b^2 + (ab+b^3)\alpha^{p_b+p_e-2} + b^3\alpha^{3p_b+p_f-p_a-3} + \Theta\delta^{-1}(a^2+b^4)r^2\alpha^{p_a-2} + K^n e^{-K\sigma} a \\ \sigma^{\tilde\beta} b_+ &\lessdot bc + bd + (a+b^2)\alpha^{p_a+p_e-p_b-1} + b^2\alpha^{p_b+p_f-2} + \Theta\delta^{-1}(ab+b^3)r^2\alpha^{p_a-2} + K^n e^{-K\sigma} b \\ \sigma^{\tilde\beta} c_+ &\lessdot c^2 + b\alpha^{p_b+p_e-2} + \Theta\delta^{-1}(a+b^2)r^2\alpha^{p_a-2} + K^n e^{-K\sigma} c \\ \sigma^{\tilde\beta} d_+ &\lessdot cd + d^2 + b\alpha^{p_b+p_e-2} + b\alpha^{p_b+p_f-2} + \Theta\delta^{-1}(a+b^2)r^2\alpha^{p_a-2} + K^n e^{-K\sigma} d \end{aligned}$$

which imply (8.34) thanks to $\Theta r^2 \le 18\alpha$, (5.11), (8.33) and (8.26). ■

8.4 KAM iteration

We *fix* χ such that

$$1 < \chi < 2^{1/3}\,, \qquad \chi^4 + 1 > \chi^5\,. \tag{8.35}$$

Below an "absolute constant " (denoted by $c, c_i, c', \ldots$) is a constant depending (possibly) on χ only.

Lemma 8.16. *Let $\{(a_j, b_j, c_j, d_j)\}_{0\le j\le\nu}$ a sequence of positive numbers satisfying*

$$\begin{aligned} a_{j+1} &\le \kappa^{j+1}(a_j c_j + b_j^2 + a_j^2 + K_*^n e^{-K_* 2^j} a_j) \\ b_{j+1} &\le \kappa^{j+1}(a_j + b_j^2 + b_j c_j + b_j d_j + K_*^n e^{-K_* 2^j} b_j) \\ c_{j+1} &\le \kappa^{j+1}(b_j + c_j^2 + a_j + K_*^n e^{-K_* 2^j} c_j) \\ d_{j+1} &\le \kappa^{j+1}(b_j + c_j d_j + d_j^2 + a_j + K_*^n e^{-K_* 2^j} d_j)\,, \qquad \forall\, 0 \le j \le \nu - 1\,, \end{aligned} \tag{8.36}$$

where $\kappa > e^e$ and $K_* \ge 2^6 + 6\ln\kappa + 16n^2$. There exist $0 < \gamma_0 := \gamma_0(\kappa,\chi) \le 1/3$ such that

$$a_0,\, b_0,\, c_0,\, d_0 \le \varepsilon_0 \le \gamma_0 \implies a_j,\, b_j,\, c_j,\, d_j \le \gamma_0^{-1}\,\varepsilon_0\, e^{-\chi^j}\,,\ \forall\, 0 \le j \le \nu\,. \tag{8.37}$$

In particular one can take $\gamma_0 = \kappa^{-c\ln(\ln\kappa)}$ for some $c = c(\chi) \ge 1$.

PROOF. In the last three inequalities in (8.36) appear the linear terms a_j, b_j. This seems in contrast with a superconvergent iterative scheme, i.e. (8.37). However we recover a quadratic scheme iterating three times, i.e. the estimate of $(a_{j+3}, b_{j+3}, c_{j+3}, d_{j+3})$ in terms of (a_j, b_j, c_j, d_j) is quadratic. The detailed computations are given in the Appendix. ■

For $\nu \in \mathbb{N}$ we define

- $\sigma_\nu := \sigma_0 2^{-\nu}\,, \quad \sigma_0 := \frac{s_0}{8}\,, \quad s_{\nu+1} := s_\nu - 2\sigma_\nu \searrow \frac{s_0}{2}\,,$
- $\delta_\nu := 2^{-\nu-3}\,, \quad r_{\nu+1} := (1-\delta_\nu) r_\nu \searrow r_0 \prod_{\nu=0}^{\infty}(1-\delta_\nu) > \frac{r_0}{2}\,, \quad D_\nu := D(s_\nu, r_\nu)\,,$
- $1 > \alpha_0 \geq \alpha_\nu := \frac{\alpha_0}{2}(1+2^{-\nu}) \searrow \frac{\alpha_0}{2}\,, \quad M_\nu := M_0(2-2^{-\nu}) \nearrow 2M_0\,, \quad \lambda_\nu := \frac{\alpha_\nu}{M_\nu} \searrow \frac{\alpha_0}{4M_0},$
- $K_\nu := K_0 4^\nu\,, \quad K_0 := \frac{8K_*}{s_0}\,, \quad K_{-1} := 0\,, \quad K_* := 2^6 + 6\ln\kappa + 16n^2.$

Note that $K_\nu \sigma_\nu = K_* 2^\nu \geq 1$. Let us define

$$\kappa := 4C_1(4/s_0)^{\tilde\beta} \tag{8.38}$$

where $C_1 = \mathtt{K}^c$, $\tilde\beta = 16\tau + 8n + 12$ are introduced in Lemma 8.15 and $\mathtt{K} = (n+\tau)^{c(n+\tau)}$ in Lemma 8.8 (here c denotes absolute, possibly different, costants). We set

$$\gamma_0 := \gamma_0(\kappa, \chi) \text{ as in Lemma 8.16 with } \kappa \text{ in (8.38)}\,. \tag{8.39}$$

Note that, for some $1 < c_1 < c_2$,

$$e^{c_1\tau_0} \leq \kappa \leq e^{c_2\tau_0}\,, \quad \tau_0^{-c_2\tau_0} \leq \gamma_0 \leq \tau_0^{-c_1\tau_0}\,, \quad \text{with} \quad \tau_0 := (\tau+n)\ln\big((\tau+n)/s_0\big)\,. \tag{8.40}$$

In the following lemma we set $|\cdot|_\nu := |\cdot|_{s_\nu}^{\lambda_\nu}$ for brevity.

Lemma 8.17. (Iterative Lemma) *Let $H^0 = N^0 + P^0 : D_0 \times \Pi_{-1} \to \mathbb{C}$ be analytic in D_0 with $\Pi_{-1} \subset \mathbb{R}^m$, $N^0 := e_0 + \omega_0(\xi)\cdot y + \Omega_0(\xi)\cdot z\bar z$ in normal form and $|\omega_0|^{\rm lip} + |\!|\Omega_0|\!|^{\rm lip}_{-\delta_*} \leq M_0$. Define*

$$a_0 := \frac{|P^0_{00}|_0}{r_0^2\alpha_0^{p_a}}\,,\ b_0 := \frac{|P^0_{01}|_0}{r_0\alpha_0^{p_b}}\,,\ c_0 := \frac{|P^0_{10}|_0}{\alpha_0}\,,\ d_0 := \frac{|P^0_{02}|_0}{\alpha_0}\,.$$

There exist $C_\star = \gamma_0^{-c^*} > 1$, $\gamma_\star = \gamma_0^{c_\star} < 1$ (for some absolute constants $c_\star > c^* > 1$), such that, if the smallness conditions

$$\max\{a_0\,, b_0\,, c_0\,, d_0\} =: \varepsilon_0 \leq \gamma_\star\,, \quad r_0|P^0_{11}|_0 \leq \alpha_0^{p_e}, \quad r_0|P^0_{03}|_0 \leq \alpha_0^{p_f}\,, \quad 2\Theta_0 r_0 \leq \sqrt{\alpha_0}\,, \tag{8.41}$$

are satisfied (the constant Θ_0 is defined as in (5.5) for P^0), then:

(S1)$_\nu$ *$\forall 0 \leq j \leq \nu$ there exist $H^j = N^j + P^j : D_j \times \Pi_{j-1} \to \mathbb{C}$, analytic in D_j, with $N^j := e_j + \omega_i(\xi)\cdot y + \Omega_i(\xi)\cdot z\bar z$ in normal form and*

$$\Pi_j := \Big\{\, \xi \in \Pi_{j-1} \,:\, |\omega_j(\xi)\cdot k + \Omega_j(\xi)\cdot l| \geq \alpha_j \frac{\langle l\rangle_d}{1+|k|^\tau}\,,\ \forall (k,l)\neq 0\,,\ |k| \leq K_j\,,\ |l| \leq 2 \Big\}\,. \tag{8.42}$$

Moreover, $\forall 1 \leq j \leq \nu$, $H^j = H^{j-1}\circ\Phi^j$ where $\Phi^j : D_j \times \Pi_{j-1} \to D_{j-1}$ are a Lipschitz family of real analytic symplectic maps of the form $\Phi^j = I + \Psi^j$ with $\Psi^j \in \mathcal{E}_{s_j}$ satisfying

$$\begin{aligned} &|x^j_{00}|_j\,, |y^j_{10}|_j \leq C_\star 2^{(2\beta-1)(j-1)} c_{j-1}\,, \quad |y^j_{00}|_j \leq C_\star 2^{(2\beta-1)(j-1)} r_0^2 \alpha_0^{p_a-1}(a_{j-1} + b^2_{j-1})\,,\\ &|y^j_{01}|_j\,, |w^j_{00}|_j \leq C_\star 2^{(2\beta-1)(j-1)} r_0\alpha_0^{p_b-1} b_{j-1}\,, \quad |y^j_{02}|_j\,, |w^j_{01}|_j \leq C_\star 2^{(2\beta-1)(j-1)} d_{j-1}\,, \end{aligned} \tag{8.43}$$

where

$$a_j := \frac{|P^j_{00}|_j}{r_j^2 \alpha_j^{p_a}}\,, \quad b_j := \frac{|P^j_{01}|_j}{r_j \alpha_j^{p_b}}\,, \quad c_j := \frac{|P^j_{10}|_j}{\alpha_j}\,, \quad d_j := \frac{|P^j_{02}|_j}{\alpha_j}\,. \tag{8.44}$$

(S2)$_\nu$ *$\forall 0 \le j \le \nu$ there exist Lipschitz extensions $\tilde\omega_j$, $\tilde\Omega_j$ of ω_j, Ω_j defined on Π_{-1} and, for $j \ge 1$,*

$$|\tilde\omega_j - \tilde\omega_{j-1}|\,, |\tilde\omega_j - \tilde\omega_{j-1}|^{\rm lip} \le |P^{j-1}_{10}|_{s_{j-1}}\,, \ |\tilde\Omega_j - \tilde\Omega_{j-1}|_{\bar p - p}\,, |\tilde\Omega_j - \tilde\Omega_{j-1}|^{\rm lip}_{\bar p - p} \le |P^{j-1}_{02}|_{s_{j-1}}\,, \tag{8.45}$$

$$|\tilde\omega_j|^{\rm lip} + |\tilde\Omega_j|^{\rm lip}_{-\delta_*} \le M_j\,. \tag{8.46}$$

(S3)$_\nu$ *$\{(a_j, b_j, c_j, d_j)\}_{0\le j\le\nu}$ satisfy (8.36) with κ defined in (8.38).*

(S4)$_\nu$ *$\forall 0 \le j \le \nu - 1$, the $a_j\,, b_j\,, c_j\,, d_j \le \gamma_0^{-1}\varepsilon_0\, e^{-\chi^j}$ with γ_0 defined in (8.39).*

(S5)$_\nu$ *$\forall\, 1 \le j \le \nu - 1$ we have $\Theta_j \le 9\Theta_0$ (see (8.13)), and*

$$|P^j_{11} - P^{j-1}_{11}|_j \le 2^{-j-1} C_\star \varepsilon_0 (|P^0_{11}|_0 + \alpha_0^{p_a - 1/2})\,, \tag{8.47}$$

$$|P^j_{03} - P^{j-1}_{03}|_j \le 2^{-j-1} C_\star \varepsilon_0 (|P^0_{03}|_0 + |P^0_{11}|_0 + \alpha_0^{p_a - 1/2})\,. \tag{8.48}$$

(S6)$_\nu$ *$\forall 1 \le j \le \nu$, the composed map $\tilde\Phi^j := \Phi^1 \circ \Phi^2 \circ \cdots \circ \Phi^j = I + \tilde\Psi^j$ with $\tilde\Psi^j \in \mathcal{E}_{s_j}$ satisfies*

$$\begin{aligned} &|\tilde x^j_{00}|_j\,, |\tilde y^j_{10}|_j\,, |\tilde y^j_{02}|_j\,, |\tilde w^j_{01}|_j \le C_\star^2 (1 - 2^{-j})\varepsilon_0\,, \\ &|\tilde y^j_{00}|_j \le C_\star^2 (1 - 2^{-j}) r_0^2 \alpha_0^{p_a - 1}\varepsilon_0\,, \quad |\tilde y^j_{01}|_j\,, |\tilde w^j_{00}|_j \le C_\star^2 (1 - 2^{-j}) r_0 \alpha_0^{p_b - 1}\varepsilon_0\,. \end{aligned} \tag{8.49}$$

Proof. The statements $(S1)_0$, $(S2)_0$, $(S5)_0$, follow by the hypothesis of the lemma, (8.41) and setting $\tilde\omega_0 := \omega_0$, $\tilde\Omega_0 = \Omega_0$. The $(S4)_0$ holds by (8.41) because $\gamma_0 \le 1/3$ (see Lemma 8.16). The $(S6)_1$ follows by $(S1)_0$. Note that $(S3)_0$ trivially holds since there is nothing to verify in (8.36) for $\nu = 0$.

Then, by induction, we prove the statements $(Si)_{\nu+1}$, $i = 1, \ldots, 6$.

(S4)$_{\nu+1}$ follows by (8.41), $(S3)_\nu$ and Lemma 8.16.

(S1)$_{\nu+1}$. By $(S4)_{\nu+1}$ we have, since $\varepsilon_0 \le \gamma_\star = \gamma_0^{c_\star}$,

$$a_\nu, b_\nu, c_\nu, d_\nu \le \gamma_0^{-1}\varepsilon_0 e^{-\chi^\nu} \le \gamma_0^{c_\star - 1} e^{-\chi^\nu} \le \frac{\delta_\nu \sigma_\nu^\beta}{16 C_0} \tag{8.50}$$

for $c_\star$ large enough. Indeed, since $\sigma_\nu := s_0 2^{-\nu}/8$, $\delta_\nu := 2^{-\nu-3}$, $\beta := 2\tau + n + 2$, we get

$$\sup_{\nu\ge 0} \frac{e^{-\chi^\nu}}{\delta_\nu \sigma_\nu^\beta} = \sup_{\nu \ge 0} s_0^{-\beta} e^{-\chi^\nu} 2^{(\beta+1)(\nu+3)} \le \Big(\frac{\beta}{s_0}\Big)^{c\beta} \le \Big(\frac{\tau + n}{s_0}\Big)^{c(\tau+n)}\,.$$

Then (8.50) follows, for $c_\star$ large enough, by (8.40) and $C_0 = K^c = (\tau + n)^{c'(\tau+n)}$, see Lemma 8.9.

Then, by (8.50), $\forall \xi \in \Pi_\nu$, Lemma 8.14 applies with $N = N^\nu$, $P = P_\nu$, $s = s_\nu$, $\sigma = \sigma_\nu$, $r = r_\nu$, $\alpha = \alpha_\nu$, $\delta = \delta_\nu$, $M = M_\nu$. There exists a real analytic symplectic map $\Phi^{\nu+1} : D_{\nu+1} \times \Pi_\nu \to D_\nu$, Lipschitz in Π_ν, such that,

$$H^{\nu+1} = H^\nu \circ \Phi^{\nu+1} =: N^{\nu+1} + P^{\nu+1}\,, \quad N^{\nu+1} := N^\nu + [P^\nu]\,.$$

The estimates (8.43) follow by (8.22) and (8.44), taking $C_\star$ large enough (namely c^* large enough).

(S2)$_{\nu+1}$. The frequency maps $\omega_{\nu+1}$ $\Omega_{\nu+1}$ are defined on Π_ν and, by Lemma 8.10, satisfy the estimates

$$|\omega_{\nu+1} - \omega_\nu| \le |P^\nu_{10}|_{s_\nu}\,, |\omega_{\nu+1} - \omega_\nu|^{\rm lip} \le |P^\nu_{10}|^{\rm lip}_{s_\nu} \tag{8.51}$$

$$|\Omega_{\nu+1} - \Omega_\nu|_{\bar p - p} \le |P^\nu_{02}|_{s_\nu}\,, |\Omega_{\nu+1} - \Omega_\nu|^{\rm lip}_{\bar p - p} \le |P^\nu_{02}|^{\rm lip}_{s_\nu}\,. \tag{8.52}$$

By the Kirszbraun theorem (see e.g. [23]), used componentwise, they can be extended to maps $\tilde{\omega}_{\nu+1}$, $\tilde{\Omega}_{\nu+1}$ defined on the whole Π_{-1} preserving the same sup-norm and Lipschitz seminorms (8.51)-(8.52). As a consequence, and since $|\ |_{-\delta_*} \leq |\ |_{\bar{p}-p}$ (recall (3.3)), we get

$$\begin{aligned} |\tilde{\omega}_{\nu+1}|^{\rm lip} + |\tilde{\Omega}_{\nu+1}|^{\rm lip}_{-\delta_*} &\leq M_\nu + |P^\nu_{10}|^{\rm lip}_\nu + |P^\nu_{02}|^{\rm lip}_\nu \leq M_\nu + \lambda_\nu^{-1}\alpha_\nu(c_\nu + d_\nu) \\ &= M_\nu(1 + c_\nu + d_\nu) \leq M_{\nu+1} \end{aligned}$$

by $(S4)_\nu$ and for $c_\star$ large enough.

$(\mathbf{S3})_{\nu+1}$ follows by (8.34) and the definition of κ. The assumptions of Lemma 8.15 hold by (8.50), by

$$\Theta_\nu r_\nu^2 \overset{(S5)_\nu}{\leq} 9\Theta_0 r_\nu^2 \leq 9\Theta_0 r_0^2 \overset{(8.41)}{\leq} 9\alpha_0/2 \leq 18\alpha_\nu$$

and $|P^\nu_{11}|_\nu \leq 9\alpha_\nu^{p_e}/r_\nu$, $|P^\nu_{03}|_\nu \leq 9\alpha_\nu^{p_f}/r_\nu$, that follow by $(S5)_\nu$. Indeed, by (8.48) with $j = \nu$, and, since $p_a \geq p_e \geq p_f$, we get by (8.41)

$$\begin{aligned} |P^\nu_{03}|_\nu &\leq |P^0_{03}|_0 + C_\star\varepsilon_0(|P^0_{03}|_0 + |P^0_{11}|_0 + \alpha_0^{p_a - 1/2}) \leq 2|P^0_{03}|_0 + |P^0_{11}|_0 + \alpha_0^{p_a-1/2} \\ &\leq 3r_0^{-1}\alpha_0^{p_f} + \alpha_0^{p_f - 1/2} \leq 4r_0^{-1}\alpha_0^{p_f} \leq 9r_\nu^{-1}\alpha_\nu^{p_f}\,, \end{aligned} \tag{8.53}$$

for $c_\star$ large enough (with respect to c^*). The estimate $|P^\nu_{11}|_\nu \leq 9\alpha_\nu^{p_e}/r_\nu$ follows as well.

$(\mathbf{S5})_{\nu+1}$. By the last inequality of Lemma 8.13, $(S4)_{\nu+1}$, (8.41) and $\Theta_\nu \leq 9\Theta_0$ we deduce

$$\begin{aligned} |P^{\nu+1}_{03} - P^\nu_{03}|_{\nu+1} &\leq \mathtt{K}^c\gamma_0^{-1}\varepsilon_0 2^{3\beta\nu}e^{-\chi^\nu}(|P^\nu_{11}|_\nu + |P^\nu_{03}|_\nu + \Theta_\nu r_\nu \alpha_\nu^{p_a-1}) \\ &\leq 2^{-\nu-2}C_\star\varepsilon_0(|P^0_{11}|_0 + |P^0_{03}|_0 + \alpha_0^{p_a-1/2}) \end{aligned}$$

with c^* large enough. The proof of (8.47) for $j = \nu + 1$ is analogous.

Finally, by $(S6)_\nu$ and $c_\star$ large enough, we apply Lemma 8.2 with $\Phi = \tilde{\Phi}^\nu = I + \tilde{\Psi}^{\nu+1}$. Then (8.5) yields $\Theta_{\nu+1} \leq 9\Theta_0$ because $\partial_{y^i w^j}H^{\nu+1} = \partial_{y^i w^j}P^{\nu+1}$ for $2i + j = 4$.

$(\mathbf{S6})_{\nu+1}$ By $(S1)_\nu$ we can apply Lemma 8.7 with $\tilde{\Phi} = \tilde{\Phi}^\nu$, $\Phi = \Phi^{\nu+1}$, $\hat{\Psi} = \tilde{\Psi}^{\nu+1}$. Then $\tilde{\Psi}^{\nu+1} \in \mathcal{E}_{s_{\nu+1}}$ and $(S6)_{\nu+1}$ follows. The estimate for $\tilde{y}^{\nu+1}_{00}$ follows by the bound in $(S6)_\nu$ for $|\tilde{y}^\nu_{00}|_\nu$ and the inequalities

$$\begin{aligned} |\tilde{y}^{\nu+1}_{00} - \tilde{y}^\nu_{00}|_{\nu+1} &\overset{(8.12)}{\leq} |y^{\nu+1}_{00}|_{\nu+1} + 2^{\nu+3}s_0^{-1}|x^{\nu+1}_{00}|_{\nu+1}|\tilde{y}^\nu_{00}|_\nu \\ &\quad +2(|\tilde{y}^\nu_{01}|_\nu|w^{\nu+1}_{00}|_{\nu+1} + |\tilde{y}^\nu_{10}|_\nu|y^{\nu+1}_{00}|_{\nu+1} + |\tilde{y}^\nu_{02}|_\nu|w^{\nu+1}_{00}|^2_{\nu+1}) \\ &\overset{(S1)_{\nu+1}}{\leq} C_\star^2 2^{-\nu-1}r_0^2\alpha_0^{p_a-1}\varepsilon_0 \end{aligned}$$

with c^* large enough and, then, $c_\star$ large enough (w.r.t. c^*). All the other estimates are analogous. ■

Corollary 8.1. *For all $\xi \in \Pi_{\alpha_0} := \cap_{\nu\geq 0}\Pi_\nu$ the sequence $\tilde{\Phi}^\nu = I + \tilde{\Psi}^\nu$ converges uniformly on $D(s_0/2, r_0/2)$ to an analytic symplectic map $\Phi = I + \Psi$ where $\Psi \in \mathcal{E}_{s_0/2}$ satisfies*

$$|x_{00}|^{\lambda_0}_{s_0/2}, |y_{00}|^{\lambda_0}_{s_0/2}\frac{\alpha_0^{1-p_a}}{r_0^2}, |y_{01}|^{\lambda_0}_{s_0/2}\frac{\alpha_0^{1-p_b}}{r_0}, |y_{10}|^{\lambda_0}_{s_0/2}, |y_{02}|^{\lambda_0}_{s_0/2}, |w_{01}|^{\lambda_0}_{s_0/2}, |w_{00}|^{\lambda_0}_{s_0/2}\frac{\alpha_0^{1-p_b}}{r_0} \leq \gamma_0^{-c}\varepsilon_0 \tag{8.54}$$

and the perturbation $P^\infty_{\leq 2}(\cdot, \xi) = 0$.

Proof. The $\tilde{\Phi}^{\nu+1} - \tilde{\Phi}^\nu = \Psi^{\nu+1} \circ \tilde{\Phi}^\nu$ is a Cauchy sequence by (8.43), $(S4)_{\nu+1}$ and $(S6)_\nu$. Estimates (8.54) follow by (8.49) and since $|\cdot|^{\lambda_0/4}_{s_0/2} \leq 4|\cdot|^{\lambda_0}_{s_0/2}$. Finally $P^\infty_{\leq 2}(\cdot, \xi) = 0$, $\forall \xi \in \Pi_{\alpha_0}$, follows by (8.44) and $(S4)_\nu$. ■

Let us define

$$\omega_\infty := \lim_{\nu\to\infty}\tilde{\omega}_\nu\,, \quad \Omega_\infty := \lim_{\nu\to\infty}\tilde{\Omega}_\nu\,.$$

It could happen that $\Pi_{\nu_0} = \emptyset$ for some ν_0. In such a case $\Pi_{\alpha_0} = \emptyset$ and the iterative process stops after finitely many steps. However, we can always set $\tilde{\omega}_\nu := \tilde{\omega}_{\nu_0}$, $\tilde{\Omega}_\nu := \tilde{\Omega}_{\nu_0}$, $\forall \nu \geq \nu_0$, and ω_∞, Ω_∞ are always well defined.

Lemma 8.18. $|\tilde{\omega}_\nu - \omega_\infty|$, $|\tilde{\Omega}_\nu - \Omega_\infty|_{\bar{p}-p}$, $|\tilde{\omega}_\nu - \omega_\infty|^{\rm lip}$, $|\tilde{\Omega}_\nu - \Omega_\infty|^{\rm lip}_{\bar{p}-p} \leq \gamma_0^{-c}\alpha_0\varepsilon_0 e^{-\chi^\nu}$.

PROOF. By (8.45), (8.44), $(S4)_\nu$, we have

$$|\tilde{\omega}_\nu - \omega_\infty| \leq \Big| \sum_{j=\nu}^{\infty} \tilde{\omega}_{j+1} - \tilde{\omega}_j \Big| \leq \gamma_0^{-1}\alpha_0\varepsilon_0 \sum_{j=\nu}^{\infty} e^{-\chi^j} \leq \gamma_0^{-c}\alpha_0\varepsilon_0 e^{-\chi^\nu} .$$

The other estimates are analogous. ■

End of the proof of Theorem 5.1

CASE 1: HYPOTHESES (H1), (H2), OR (H3)-($d > 1$). We apply the iterative Lemma with

$$s_0 := s,\ r_0 := \frac{r}{2},\ \alpha_0 := \alpha,\ N^0 := N,\ P^0 := P,\ \Theta_0 := \Theta,\ M_0 := M,\ \Pi_{-1} := \Pi .$$

The smallness assumption (8.41) follows by (5.5), (H1), (H2), (H3), (8.33), taking $\gamma \leq \gamma_\star$. Theorem 5.1 follows by the conclusions of Lemma 8.17, Corollary 8.1 and Lemma 8.18. Finally we prove the characterisation of the Cantor set in terms of the limit frequencies $(\omega_\infty, \Omega_\infty)$.

Lemma 8.19. $\Pi_\infty \subseteq \Pi_\alpha := \cap_{\nu \geq 0}\Pi_\nu$.

PROOF. By (3.3) we get $|l|_{p-\bar{p}} \leq |l|_{d-1} \leq 2\langle l \rangle_d$. If $\xi \in \Pi_\infty$, we have, $\forall \nu \geq 0$, $\forall |k| \leq K_\nu$, $|l| \leq 2$,

$$|\omega_\nu(\xi)\cdot k + \Omega_\nu(\xi)\cdot l| \geq 2\alpha \frac{\langle l \rangle_d}{1+|k|^\tau} - |\omega_\nu(\xi) - \omega_\infty(\xi)||k| - 2|\Omega_\nu - \Omega_\infty|_{\bar{p}-p}\langle l \rangle_d \geq \alpha \frac{\langle l \rangle_d}{1+|k|^\tau} \qquad (8.55)$$

because, by Lemma 8.18, for γ small enough,

$$|\omega_\nu(\xi) - \omega_\infty(\xi)| \leq \frac{\alpha}{2(1+K_\nu^\tau)K_\nu} ,\ |\Omega_\nu - \Omega_\infty|_{\bar{p}-p} \leq \frac{\alpha}{4(1+K_\nu^\tau)} .$$

Since $\alpha \geq \alpha_\nu$, by (8.55) we deduce $\Pi_\infty \subset \Pi_\nu$, $\forall \nu \geq 0$. ■

CASE 2: HYPOTHESIS (H3)-($d = 1$). We first perform one step of averaging. The homological equation

$$\{N, F\} + P_{00} = \langle P_{00} \rangle$$

has a solution $\hat{F} := \hat{F}_{00}$, for all $\xi \in \Pi$ such that[5] $\omega(\xi) \in \mathcal{D}_{\alpha^\mu,\tau}$ (see (1.14)). The symplectic map $\hat{\Phi} := X^1_{\hat{F}} : D(s/2, r/2) \to D(s,r)$ has the form

$$\hat{\Phi}(x_+, y_+, w_+) = (x_+, y_+ + \hat{y}_{00}(x_+), w_+)$$

and $|\hat{y}_{00}|_{s/2} \lessdot \alpha^{-\mu}|P_{00}|_s$, where, here and in the following, $|\cdot|_s$ and $|\cdot|_{s/2}$ are short for $|\cdot|_s^\lambda$ and $|\cdot|_{s/2}^\lambda$ respectively. Then $\hat{H} := H \circ \hat{\Phi} = N + \hat{P}$ satisfies

$$\begin{aligned}
|\hat{P}_{00}|_{s/2} &\lessdot \alpha^{-\mu}|P_{00}|_s|P_{10}|_s + \alpha^{-2\mu}|P_{00}|_s^2 \lessdot \varepsilon_3^2 r^2\alpha + \varepsilon_3^2 r^4 \leq 2\varepsilon_3^2 r^2 \alpha \\
|\hat{P}_{01}|_{s/2} &\lessdot |P_{01}|_s + \alpha^{-\mu}|P_{11}|_s|P_{00}|_s \lessdot |P_{01}|_s + \alpha^{1/2}\varepsilon_3 r^2 \leq \alpha\varepsilon_3 r \\
|\hat{P}_{10}|_{s/2} &\lessdot |P_{10}|_s + \alpha^{-\mu}|P_{00}|_s \lessdot |P_{10}|_s + \varepsilon_3 r^2 \leq \varepsilon_3 \alpha \\
|\hat{P}_{02}|_{s/2} &\lessdot |P_{02}|_s + \alpha^{-\mu}|P_{00}|_s \leq \varepsilon_3\alpha
\end{aligned}$$

and so

$$\tilde{\varepsilon} := \max\Big\{ r^{-2}\alpha^{-1}|\hat{P}_{00}|_{s/2} ,\ \alpha^{-1}r^{-1}|\hat{P}_{01}|_{s/2} ,\ \alpha^{-1}|\hat{P}_{10}|_{s/2} ,\ \alpha^{-1}|\hat{P}_{02}|_{s/2} \Big\} \lessdot \varepsilon_3 .$$

[5] Actually it is sufficient to require in (1.14) only finitely many non-resonance conditions, i.e. for $|k| \leq \bar{K}$.

Moreover

$$|\hat{P}_{11}-P_{11}|_{s/2}\,,|\hat{P}_{03}-P_{03}|_{s/2}\lessdot|\hat{y}_{00}|_{s/2}\lessdot\alpha^{-\mu}|P_{00}|_s\leq\varepsilon_3 r^2\leq\varepsilon_3\alpha\,,$$

whence $|\hat{P}_{11}|_{s/2}\,,|\hat{P}_{03}|_{s/2}\leq 2\alpha/r$, if γ is small enough. By Lemma 8.2 we get $\hat{\Theta}\leq 3\Theta$. We apply the iterative Lemma with

$$H^0:=\hat{H},\ N^0:=N,\ P^0:=\hat{P}\ ,\ s_0:=\frac{s}{2},\ r_0:=\frac{r}{2},\ \alpha_0:=\alpha,\ \Theta_0:=3\Theta\,,\ M_0:=M\,,\ \varepsilon_0:=\tilde{\varepsilon}\,,$$
$$\Pi_{-1}:=\Pi\setminus\omega^{-1}(\mathcal{D}_{\alpha^\mu,\tau})\,.$$

Then (8.41) follows since $\tilde{\varepsilon}\lessdot\varepsilon_3\leq\gamma$, taking γ small enough (with respect to $\gamma_\star$).

We now prove remark 5.1 for analytic Hamiltonians.

Remark 8.1. *We only modify the statement $(S2)_\nu$ stating the existence of C^∞-extensions of the frequency maps ω_∞, Ω_∞. We follow the cut-off procedure of [5]. The small divisor condition (8.42) holds with $\alpha_j/2$ instead of α_j in the neighborhood*

$$\mathcal{N}(\Pi_j):=\Big\{\xi\in\Pi_{j-1}\ :\ \mathrm{dist}(\xi,\Pi_j)\leq c\alpha_j K_j^{-(\tau+1)}\Big\}\tag{8.56}$$

where c is a small constant. Then H^{j+1} exists for all $\xi\in\mathcal{N}(\Pi_j)$ and, the KAM iteration implies

$$|\omega_{j+1}-\omega_j|\,,\quad |\Omega_{j+1}-\Omega_j|_{\bar{p}-p}\leq C\alpha_j\varepsilon_0 e^{-\chi^j}\,.$$

By a cut-off procedure we define C^∞-functions $\tilde{\Omega}_{j+1}-\tilde{\Omega}_j$ for all the parameters $\xi\in\Pi_{-1}$ coinciding with $\Omega_{j+1}-\Omega_j$ on Π_j and equal to zero outside $\mathcal{N}(\Pi_j)$. Moreover, by (8.56), the derivatives of such extended frequency maps satisfy

$$|D^q(\tilde{\Omega}_{j+1}-\tilde{\Omega}_j)|_{\bar{p}-p}\leq C\alpha_j\varepsilon_0 e^{-\chi^j}/(\alpha_j K_j^{-(\tau+1)})^q\leq C(q)\frac{\varepsilon_0}{\alpha^{q-1}}e^{-\chi^j}K_j^{(\tau+1)q}\,,\quad\forall q\geq 1\,.$$

An analogous estimate hold for $\tilde{\omega}_{j+1}-\tilde{\omega}_j$. Summing in $j\geq 1$ we get (5.15).

We now discuss the estimates of remark 5.3.

Remark 8.2. *By Lemma 8.17 the small constant $\gamma:=\gamma(n,\tau,s)$ of Theorem 5.1 can be taken $\gamma:=\gamma_0^c$ where γ_0 is defined in (8.39). Then (8.40) implies the estimate for γ given in Remark 5.3.*

Proof of remark 5.2. By (5.6), (1.13), $\lambda=\alpha/M$, we get

$$|\omega_\infty-\omega|^{\mathrm{lip}},|\Omega_\infty-\Omega|^{\mathrm{lip}}_{-\delta_*}\leq M\varepsilon_i/\gamma\tag{8.57}$$

By (5.2), (3.3) we have $|\omega_\infty|^{\mathrm{lip}}$, $|\Omega_\infty|^{\mathrm{lip}}_{-\delta_*}\leq M+M\varepsilon_i/\gamma\leq 2M$. Let $\xi_1,\xi_2\in\Pi$ and $\omega_j:=\omega_\infty(\xi_j)$, $j=1,2$. We have $|\xi_1-\xi_2|=|\omega_\infty^{-1}(\omega_1)-\omega_\infty^{-1}(\omega_2)|\leq L|\omega_1-\omega_2|$ and

$$\begin{aligned}|\omega_\infty(\xi_1)-\omega_\infty(\xi_2)| &\geq |\omega_1-\omega_2|-|(\omega_\infty-\omega)(\xi_1)-(\omega_\infty-\omega)(\xi_2)|\\ &\geq \big(L^{-1}-|\omega_\infty-\omega|^{\mathrm{lip}}\big)|\xi_1-\xi_2|\\ &\overset{(8.57)}{\geq} (L^{-1}-\gamma^{-1}M\varepsilon_i)|\xi_1-\xi_2|\geq(2L)^{-1}|\xi_1-\xi_2|\,.\end{aligned}$$

Therefore ω_∞ is injective and $|\omega_\infty^{-1}|^{\mathrm{lip}}\leq 2L$.

Proof of Theorem 5.3. We have $\omega(\xi)=\mathrm{a}+\mathrm{A}\xi$, $\det A\neq 0$, $\Omega(\xi)=\mathrm{b}+\mathrm{B}\xi$ and (B^*) implies

$$b_i=i^d+\text{lower order terms}\,,\ i>n\,,\quad B\in\mathcal{L}(\mathbb{C}^n,\ell_\infty^{-\delta_*})\,,\quad\delta_*<d-1\,.\tag{8.58}$$

Since Π is compact and $0\notin\omega(\Pi)$ there exist $0<t_-<t_+$ such that

$$\omega_\infty(\Pi)\cap\bar{\omega}\mathbb{R}^+\subset[t_-,t_+]\bar{\omega}\,.$$

By remark 5.2, for ε_i small enough, the perturbed frequency map ω_∞ is invertible. Then, for all $t \in [t_-, t_+]$ such that $t\bar\omega \in \omega_\infty(\Pi)$ we define

$$\bar\Omega_\infty(t) := \Omega_\infty\big(\omega_\infty^{-1}(t\bar\omega)\big) = \mathrm{b} + \mathrm{BA}^{-1}(t\bar\omega - \mathrm{a}) + r(t)$$

where $r(t)$ is a Lipschitz map satisfying, by (5.6) and (8.58),

$$|r|_{-\delta_*}\,\alpha^{-1}, |r|^{\mathrm{lip}}_{-\delta_*} \le c\varepsilon_i \le c\gamma\,. \tag{8.59}$$

The map $r(t)$ can be extended to a Lipschitz map on the whole $\mathbb{R}$ preserving the bounds (8.59) by the Kirszbraun theorem applied componentwise. Defining

$$f_{kl}(t) := t\bar\omega \cdot k + \bar\Omega_\infty(t)\cdot l = (\mathrm{b} - \mathrm{BA}^{-1}\mathrm{a})\cdot l + t(k + \mathrm{A}^{-1}\mathrm{B}^\intercal l)\cdot\bar\omega + r(t)\cdot l \tag{8.60}$$

we have to estimate the resonant set

$$\omega_\infty(\Pi\setminus\Pi_\infty)\cap\bar\omega\mathbb{R}^+ \subseteq \bigcup_{k\in\mathbb{Z}^n, |l|\le 2, (k,l)\neq 0} R_{kl} \quad \text{where} \quad R_{kl} := \left\{ t\in[t_-,t_+] \ : \ |f_{kl}(t)| < \frac{2\alpha\langle l\rangle_d}{1+|k|^\tau}\right\}.$$

Let $\Lambda_{i_0} := \{|l|\le 2 \ : \ l_i = 0\,,\ \forall\, i > i_0\}$. Note that Λ_{i_0} is a finite set.

Lemma 8.20. *There exists $\beta_1 > 0$ (small enough) and i_0 (large enough) such that*

$$\alpha\le\beta_1\,,\quad l\notin\Lambda_{i_0}\,,\quad |k|\le\langle l\rangle_d/8t_+ \quad\Longrightarrow\quad R_{kl}=\emptyset\,. \tag{8.61}$$

Proof. We first prove that if i_0 is large enough then

$$|(\mathrm{b}-\mathrm{BA}^{-1}\mathrm{a} + t\mathrm{BA}^{-1}\bar\omega)\cdot l| \ge \langle l\rangle_d/4\,,\quad \forall\, t\in[t_-,t_+]\,,\ 0<|l|\le 2\,,\ l\notin\Lambda_{i_0}\,. \tag{8.62}$$

We consider only the subtlest case $l = e_i - e_j$, $i > j$. Since $l \notin \Lambda_{i_0}$, we have $i > i_0$. By (8.58) we get $|\mathrm{b}\cdot l| \ge \langle l\rangle_d/2$ for i_0 large enough. If $d > 1$ then $\langle l\rangle_d = i^d - j^d \ge d i^{d-1}$. Then (8.62) follows for i_0 large enough since, by (8.58), $|(\mathrm{BA}^{-1}\mathrm{a} + t\mathrm{BA}^{-1}\bar\omega)\cdot l| \le C i^{\delta_*}$ and $\delta_* < d-1$. If $d = 1$, $\delta_* < 0$ and it is enough to prove that $i - j \ge C j^{\delta_*}$ for some $C > 1$. For all $j > j_0$ such that $C j_0^{\delta_*} \le 1$ the thesis follows because $i - j \ge 1$. For all $j \le j_0$ the thesis follows taking $i_0 \ge j_0 + C$. By (8.60), (8.62), (8.59), if $t_+|k| \le \langle l\rangle_d/8$ and $\alpha\le\beta_1$ is small enough, then

$$|f_{kl}(t)| \ge \frac14\langle l\rangle_d - const\,\alpha - t_+|k| \ge \frac19\langle l\rangle_d > \frac{2\alpha\langle l\rangle_d}{1+|k|^\tau}$$

implying that $R_{kl} = \emptyset$. ■

Lemma 8.21. *For $\bar\omega\in\mathcal{D}_{K\alpha,\tau}$ with $K > 2/t_-$ then $R_{k0} = \emptyset$. Moreover for α small*

$$|R_{kl}| \le const\frac{\alpha\langle l\rangle_d}{1+|k|^\tau}\,,\quad \forall\, k\in\mathbb{Z}^n\,,\ |l|\le 2\,,\ (k,l)\neq 0\,. \tag{8.63}$$

Proof. Since $\bar\omega\in\mathcal{D}_{K\alpha,\tau}$ with $K > 2/t_-$ then, for $t\in[t_-,t_+]$,

$$|f_{k0}(t)| = |t\bar\omega\cdot k| \ge t_-|\bar\omega\cdot k| \ge 2\alpha/(1+|k|^\tau) \quad\Longrightarrow\quad R_{k0}=\emptyset\,.$$

We then discuss $l \neq 0$. Moreover, by Lemma 8.20, we consider only $l\in\Lambda_{i_0}$ or $|k| > \langle l\rangle_d/8t_+$. By the hypotheses (5.22) and (8.58), arguing as in Remark 2.1,

$$c_l := (\mathrm{b}-\mathrm{BA}^{-1}\mathrm{a})\cdot l \qquad \text{satisfies} \qquad |c_l| \ge \bar\delta > 0\,,\ \forall\, 0<|l|\le 2\,. \tag{8.64}$$

Now set $m_{kl} := (k + \mathrm{A}^{-1}\mathrm{B}^\intercal l)\cdot\bar\omega$. If $|m_{kl}| < \bar\delta/(3t_+)$, by (8.60), (8.64), (8.59), for α small enough,

$$|f_{kl}(t)| \ge |c_l| - \frac{\bar\delta}{3} - 2c\gamma\alpha \ge \frac{\bar\delta}{2} \overset{(8.61)}{\ge} \frac{2\alpha\langle l\rangle_d}{1+|k|^\tau} \quad\Longrightarrow\quad R_{kl}=\emptyset\,.$$

If $|m_{kl}| \ge \bar\delta/(3t_+)$ we have $|f_{kl}(t_2) - f_{kl}(t_1)| \ge |t_2-t_1|(|m_{kl}| - 2c\gamma) \ge |t_2-t_1|\bar\delta/(4t_+)$ for γ small enough and (8.63) follows with $const = 8t_+/\bar\delta$. ■

Now the proof of (5.23) proceeds as in [24] or in subsection 7.1 above (recalling Remark 7.3, now (7.17) holds also for $d = 1$ since $\hat n = n$, $D = 2$). Note that (8.61) and (8.63) are the analogue of Lemma 7.4 and Lemmata 7.7 (case $d > 1$), 7.10 (case $d = 1$) respectively.

9 Appendix

PROOF OF LEMMA 8.6. We take $0 \le t \le 1$. For brevity we write $|\cdot|$ instead of $|\cdot|^\lambda$.
STEP 1. The solution of the first equation in (8.8) with $x^0 = x_+$ has the form

$$x^t = x_+ + x^t_{00}(x_+) \quad \text{where} \quad x^t_{00}(x_+) = \int_0^t F_{10}\big(x_+ + x^\tau_{00}(x_+)\big)\, d\tau\,.$$

By (8.9) and (8.1) we get $|x^t_{00}|_{s-\sigma} \le \sigma/2$ and the estimate (8.10) for x^t_{00} follows.
STEP 2. Substituting x^t in the third equation in (8.8) we get

$$\dot w^t = -\mathrm{i}J\tilde F^t_{01} - 2\mathrm{i}J\tilde F^t_{02} w^t =: b^t + A^t w^t \quad \text{where} \quad \tilde F^t_{ij} := F_{ij}\big(x_+ + x^t_{00}(x_+)\big)\,. \tag{9.65}$$

By (8.1) we have $|\tilde F^t_{ij}|_{s-\sigma} \le 2|F_{ij}|_s$ and so

$$|b^t|_{s-\sigma} \le 2|F_{01}|_s\,, \quad |A^t|_{s-\sigma} \le 4|F_{02}|_s \overset{(8.9)}{\le} 1/3\,. \tag{9.66}$$

Let M^t be the solution of the homogeneous system $\dot M^t = A^t M^t$ with $M^0 = I$. We have

$$|M^t - I|_{s-\sigma} \le \int_0^t |A^\tau|_{s-\sigma}|M^\tau|_{s-\sigma}\, d\tau \overset{(9.66)}{\le} \frac13 \sup_{0\le t\le 1}|M^t|_{s-\sigma} \le \frac13 + \frac13 \sup_{0\le t\le 1}|M^t - I|_{s-\sigma}$$

whence

$$|M^t|_{s-\sigma} \le \frac32 \quad \text{and} \quad |M^t - I|_{s-\sigma} \le \frac32 \sup_{0\le t\le 1}|A^t|_{s-\sigma} \overset{(9.66)}{\le} 6|F_{02}|_s \overset{(8.9)}{\le} \frac12\,. \tag{9.67}$$

Then, by Neumann series,

$$|(M^t)^{-1}|_{s-\sigma} \le \sum_{j\ge 0}|M^t - I|^j_{s-\sigma} \le 2\,. \tag{9.68}$$

The solution of the non-homogeneous problem (9.65) with $w^0 = w_+$ is

$$w^t = w_+ + (M^t - I)w_+ + M^t \int_0^t (M^\tau)^{-1} b^\tau\, d\tau =: w_+ + w^t_{01}(x_+)w_+ + w^t_{00}(x_+)\,. \tag{9.69}$$

The estimates (8.10) on w^t_{00} and w^t_{01} follow by (9.69), (9.67), (9.68), (9.66).
STEP 3. Finally, substituting x^t and w^t in the second equation (8.8), we get

$$\dot y^t = -\hat F^t_{00} - \hat F^t_{01} w^t - \hat F^t_{02} w^t \cdot w^t - \hat F^t_{10} y^t =: \hat b^t + \hat A^t y^t \tag{9.70}$$

where $\hat F^t_{ij} := F'_{ij}\big(x_+ + x^t_{00}(x_+)\big)$, $\hat A^t = -\hat F^t_{10}$, and, using (9.69),

$$\begin{aligned} \hat b^t \;\; = \;\; & -\Big(\hat F^t_{00} + \hat F^t_{01} w^t_{00} + \hat F^t_{02} w^t_{00}\cdot w^t_{00}\Big) - \Big(\hat F^t_{01}(I + w^t_{01}) + 2(w^t_{00})^\intercal \hat F^t_{02}(I + w^t_{01})\Big) w_+ \\ & -\Big((I + w^t_{01})^\intercal \hat F^t_{02}(I + w^t_{01})\Big) w_+ \cdot w_+\,. \end{aligned} \tag{9.71}$$

Since $|x^t_{00}|_{s-\sigma} \le \sigma/2$, by Cauchy estimates and (8.1) we get

$$|\hat F^t_{ij}|_{s-\sigma} \le 2|F'_{ij}|_{s-\frac{\sigma}{2}} \le \frac{4}{\sigma}|F_{ij}|_s \quad \Longrightarrow \quad |\hat A^t|_{s-\sigma} \le \frac{4}{\sigma}|F_{10}|_s \overset{(8.9)}{\le} \frac13\,. \tag{9.72}$$

Let $\hat M^t$ be the solution of $\dot{\hat M}^t = A^t \hat M^t$ with $\hat M^0 = I$. Reasoning as in Step 2 we get

$$|\hat M^t|_{s-\sigma} \le \frac32\,, \quad |\hat M^t - I|_{s-\sigma} \le \frac32 |\hat A^t|_{s-\sigma} \le \frac{6}{\sigma}|F_{10}|_s \overset{(8.9)}{\le} \frac12 \quad \text{and} \quad |(\hat M^t)^{-1}|_{s-\sigma} \le 2\,. \tag{9.73}$$

The solution of the non-homogeneous system (9.70) with $y^0 = y_+$ is

$$\begin{aligned} y^t &= y_+ + (\hat{M}^t - I)y_+ + \hat{M}^t \int_0^t (\hat{M}^\tau)^{-1} \hat{b}^\tau \, d\tau \\ &= y_+ + y^t_{00}(x_+) + y^t_{01}(x_+)w_+ + y^t_{10}(x_+)y_+ + y^t_{02}(x_+)w_+ \cdot w_+ \end{aligned}$$

where, by (9.71),

$$\begin{aligned} y^t_{00} &= -\hat{M}^t \int_0^t (\hat{M}^\tau)^{-1} \Big(\hat{F}^\tau_{00} + \hat{F}^\tau_{01} w^\tau_{00} + \hat{F}^\tau_{02} w^\tau_{00} \cdot w^\tau_{00} \Big) \, d\tau \\ y^t_{01} &= -\hat{M}^t \int_0^t (\hat{M}^\tau)^{-1} \Big(\hat{F}^\tau_{01}(I + w^\tau_{01}) + 2(w^\tau_{00})^\intercal \hat{F}^\tau_{02}(I + w^\tau_{01}) \Big) \, d\tau \\ y^t_{10} &= \hat{M}^t - I \\ y^t_{02} &= -\hat{M}^t \int_0^t (\hat{M}^\tau)^{-1} \Big((I + w^\tau_{01})^\intercal \hat{F}^\tau_{02} (I + w^\tau_{01}) \Big) d\tau \, . \end{aligned}$$

The estimates (8.10) on y^t_{ij} follow by (9.73), (9.72) and the previous estimates for w_{00}, w_{01}.
We finally prove that $X^t_F : D(s - \sigma, r - \delta r) \to D(s, r)$. If $(x_+, y_+, w_+) \in D(s - \sigma, r - \delta r)$ then

$$|\mathrm{Im}\, x^t(x_+)| = |\mathrm{Im}\, x_+ + \mathrm{Im}\, x^t_{00}(x_+)| \le s - \sigma + |x^t_{00}|_{s-\sigma} \overset{(8.10)}{\le} s - \sigma + 2|F_{10}|_s \overset{(8.9)}{<} s \, .$$

The estimates $|y^t(x_+, y_+, w_+)| < r^2$, $\|w^t(x_+, w_+)\|_{a,p} < r$, follow as well by (8.10), (8.11). ■

Proof of Lemma 8.16. Let $\gamma_0 := \tilde{\gamma}_0^3 e^{-\chi^4}$ where

$$\tilde{\gamma}_0 := \frac{1}{8} \inf_{j \ge 0} \Big\{ \kappa^{-j-1} e^{(\chi-1)\chi^{j+1}} \, , \quad \kappa^{-j-1} e^{(2-\chi)\chi^j} \, , \quad \kappa^{-j-1} e^{(\chi^4+1-\chi^5)\chi^j} \, , \quad \kappa^{-j-1} e^{(2-\chi^3)\chi^{j+2}} \Big\} .$$

Note that $\tilde{\gamma}_0 \ge \kappa^{-\tilde{c} \ln(\ln \kappa)}$ for some $\tilde{c} = \tilde{c}(\chi) \ge 1$, since $\inf_{j \ge 1} \kappa^{-j} e^{\alpha \chi^j} \ge \kappa^{-\bar{c} \ln(\ln \kappa)}$ for some $\bar{c} = \bar{c}(\chi, \alpha) \ge 1$, (recall $\kappa > e^e$). By the choice of χ we have $0 < \tilde{\gamma}_0 < 1$.
We claim that

$$a_j \le \varepsilon_0 e^{\chi^4 - \chi^{j+4}} \, , \quad b_j \le \tilde{\gamma}_0^{-1} \varepsilon_0 e^{\chi^4 - \chi^{j+2}} \, , \quad c_j \, , \ d_j \le \tilde{\gamma}_0^{-2} \varepsilon_0 e^{\chi^4 - \chi^j} \, , \qquad \forall \, 0 \le j \le \nu \, . \tag{9.74}$$

Note that (8.37) follows by (9.74) since $\tilde{\gamma}_0^{-2} e^{\chi^4} \le \gamma_0^{-1}$. We prove (9.74) by induction over j. The case $j = 0$ follows by a_0, b_0, c_0, $d_0 \le \gamma_0$. Then we prove that (9.74) holds for $j+1$. We have

$$\begin{aligned} a_{j+1} &\le \kappa^{j+1}(a_j c_j + b_j^2 + a_j^2 + K_*^n e^{-K_* 2^j} a_j) \\ &\le e^{2\chi^4} \varepsilon_0^2 \kappa^{j+1} (\tilde{\gamma}_0^{-2} e^{-\chi^{j+4} - \chi^j} + \tilde{\gamma}_0^{-2} e^{-2\chi^{j+2}} + e^{-2\chi^{j+4}}) + \varepsilon_0 \kappa^{j+1} K_*^n e^{\chi^4 - \chi^{j+4} - K_* 2^j} \\ &\le \varepsilon_0 e^{\chi^4 - \chi^{j+5}} \end{aligned}$$

since, $\forall \, j \ge 0$,

$$\begin{aligned} &\varepsilon_0 \tilde{\gamma}_0^{-2} e^{\chi^4} \le \tilde{\gamma}_0 \le \frac{1}{8} \kappa^{-j-1} e^{(\chi^4+1-\chi^5)\chi^j} \, , \quad \varepsilon_0 \tilde{\gamma}_0^{-2} e^{\chi^4} \le \tilde{\gamma}_0 \le \frac{1}{8} \kappa^{-j-1} e^{(2-\chi^3)\chi^{j+2}} \, , \\ &\varepsilon_0 e^{\chi^4} \le \tilde{\gamma}_0 \le \frac{1}{8} \kappa^{-j-1} e^{(2-\chi)\chi^{j+4}} \, , \quad \kappa^{j+1} K_*^n e^{1 + \chi^{j+5} - \chi^{j+4} - K_* 2^j} \le 1 \, . \end{aligned}$$

The first three estimates directly follow by the definition of $\tilde{\gamma}_0$. The last one holds since, by

$$K_* \ge 2^6 + 6 \ln \kappa + 16 n^2 \, , \ 1 + \chi^{j+5} - \chi^{j+4} - K_* 2^j \le \chi^{j+5} - K_* 2^j \le -K_* 2^{j-1}$$

and[6] $(j+1)\ln\kappa + n\ln K_* - K_* 2^{j-1} \le 0$. We have

$$\begin{aligned} b_{j+1} &\le \kappa^{j+1}(a_j + b_j^2 + b_j(c_j+d_j) + K_*^n e^{-K_* 2^j} b_j) \\ &\le e^{\chi^4}\varepsilon_0\kappa^{j+1}(e^{-\chi^{j+4}} + \tilde{\gamma}_0^{-2}\varepsilon_0 e^{\chi^4 - 2\chi^{j+2}} + 2\tilde{\gamma}_0^{-3}\varepsilon_0 e^{\chi^4-\chi^{j+2}-\chi^j}) + \tilde{\gamma}_0^{-1}\varepsilon_0\kappa^{j+1}K_*^n e^{\chi^4-\chi^{j+2}-K_*2^j} \\ &\le \tilde{\gamma}_0^{-1}\varepsilon_0 e^{\chi^4-\chi^{j+3}} \end{aligned}$$

since, $\forall\, j \ge 0$, $\kappa^{j+1}K_*^n e^{1+\chi^{j+3}-\chi^{j+2}-K_*2^j} \le 1$ and

$$\tilde{\gamma}_0 \le \frac{1}{8}\kappa^{-j-1}e^{(\chi-1)\chi^{j+3}}, \quad \tilde{\gamma}_0^{-1}e^{\chi^4}\varepsilon_0 \le \tilde{\gamma}_0 \le \frac{1}{8}\kappa^{-j-1}e^{(2-\chi)\chi^{j+2}},$$
$$\tilde{\gamma}_0^{-2}e^{\chi^4}\varepsilon_0 \le \tilde{\gamma}_0 \le \frac{1}{8}\kappa^{-j-1}e^{(\chi^2+1-\chi^3)\chi^j}.$$

reasoning as above (note that $\chi^2+1 > \chi^3$). Finally

$$\begin{aligned} c_{j+1} &\le \kappa^{j+1}(a_j + b_j + c_j^2 + K_*^n e^{-K_*2^j}c_j) \\ &\le e^{\chi^4}\varepsilon_0\kappa^{j+1}(e^{-\chi^{j+4}} + \tilde{\gamma}_0^{-1}e^{-\chi^{j+2}} + \tilde{\gamma}_0^{-4}e^{\chi^4}\varepsilon_0 e^{-2\chi^j}) + \tilde{\gamma}_0^{-2}\varepsilon_0\kappa^{j+1}K_*^n e^{\chi^4-\chi^j-K_*2^j} \\ &\le \tilde{\gamma}_0^{-2}\varepsilon_0 e^{\chi^4-\chi^{j+1}} \end{aligned}$$

since, $\forall\, j \ge 0$, $\kappa^{j+1}K_*^n e^{1+\chi^{j+1}-\chi^j-K_*2^j} \le 1$, and

$$\tilde{\gamma}_0^2 \le \tilde{\gamma}_0 \le \frac{1}{8}\kappa^{-j-1}e^{(\chi^3-1)\chi^{j+1}}, \quad \tilde{\gamma}_0 \le \frac{1}{8}\kappa^{-j-1}e^{(\chi-1)\chi^{j+1}}, \quad \tilde{\gamma}_0^{-2}e^{\chi^4}\varepsilon_0 \le \tilde{\gamma}_0 \le \frac{1}{8}\kappa^{-j-1}e^{(2-\chi)\chi^j}.$$

The estimate $d_{j+1} \le \tilde{\gamma}_0^{-2}\varepsilon_0 e^{\chi^4-\chi^{j+1}}$ follows as well. ■

References

[1] Bambusi D., *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Phys. 234, no. 2, 253-285, 2003.

[2] Bambusi D., Berti M., *A Birkhoof-Lewis type theorem for some Hamiltonian PDE's*, SIAM Journal on Mathematical Analysis, 37, 1, 83-102, 2005.

[3] Bambusi D., Grébert B.,*Birkhoff normal form for partial differential equations with tame modulus*, Duke Math. J. 135, no. 3, 507-567, 2005.

[4] Bambusi D., Magistrelli E., *Degenerate KAM theory for PDEs*, preprint 2010.

[5] Berti M., Bolle P., *Cantor families of periodic solutions for completely resonant nonlinear wave equations*, Duke Mathematical Journal , 134, issue 2, 359-419, 2006.

[6] Berti M., Bolle P., Procesi M., *An abstract Nash-Moser Theorem with parameters and applications to PDEs*, Annales de l'Institute H. Poincaré, Analyse nonlineaire, 27, 377-399, 2010.

[7] Biasco L., Di Gregorio L., *A Birkhoff-Lewis type Theorem for the nonlinear wave equation*, Arch. Rat. Mech. 196, no. 1, 303–362, 2010.

[8] Bourgain J., *On Melnikov's persistency problem*, Math. Res. Lett., no. 4, 445–458, 1997.

[9] Bourgain J., *Quasi-periodic solutions of Hamiltonian perturbations of* 2D *linear Schrödinger equations*, Ann. of Math. 148, 363-439, 1998.

[6] This inequality holds for $j = 0, 1$, by $K_* \ge 2^6 + 6\ln\kappa + 16n^2$, while, for $j \ge 2$, $(j+1)\ln\kappa + n\ln K_* - K_*2^{j-1} \le (j+1)\ln\kappa + n\ln K_* - K_*(j-1) \le 3\ln\kappa + n\ln K_* - K_* \le 0$.

[10] Bourgain J., *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton, 2005.

[11] Chierchia L., *A direct method for the construction of solutions of the Hamilton-Jacobi equation*, Meccanica, 25, 246-252, 1990.

[12] Chierchia L., Gallavotti G., *Smooth prime integrals for quasi-integrable Hamiltonian systems*, Il Nuovo Cimento, Vol. 67 B, n. 277-295, 1982.

[13] Craig W., *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthèses, 9, Société Mathématique de France, Paris, 2000.

[14] Craig W., Wayne C. E., *Newton's method and periodic solutions of nonlinear wave equation*, Comm. Pure Appl. Math. 46, 1409-1498, 1993.

[15] Eliasson L. H., *Perturbations of stable invariant tori for Hamiltonian systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 15, no. 1, 115-147, 1988.

[16] Eliasson L.H., Kuksin S., *KAM for non-linear Schrödinger equation*, Annals of Math., to appear.

[17] Geng J., Ren X.: *Lower dimensional invariant tori with prescribed frequency for nonlinear wave equation*, to appear in J. Diff. Eq..

[18] Gentile G., Procesi M., *Periodic solutions for a class of nonlinear partial differential equations in higher dimension*, Comm. Math. Phys., v. 289, n.3, 863-906, 2009.

[19] Giorgilli A., Morbidelli A., *Superexponential Stability of KAM Tori*, J. Stat. Phys., 78, 5/6, 1995.

[20] Kappeler T., Pöschel J., *KAM and KdV*, Springer, 2003.

[21] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional. Anal. i Prilozhen. 21, no. 3, 22–37, 95, 1987.

[22] Kuksin S., *Analysis of Hamiltonian PDEs*, Oxford Lecture Series in Mathematics and its Applications, 19. Oxford University Press, 2000.

[23] Kuksin S., Pöschel J., *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Ann. of Math, 2, 143, no. 1, 149-179, 1996.

[24] Pöschel J., *A KAM-Theorem for some nonlinear PDEs*, Ann. Scuola Norm. Sup. Pisa, Cl. Sci., 23, 119-148, 1996.

[25] Pöschel J., *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv., 71, no. 2, 269-296, 1996.

[26] Pöschel J., *On the construction of almost periodic solutions for a nonlinear Schrödinger equation*, Ergod. Th. and Dynam. Sys., 22, 15371549, 2002.

[27] Rüssmann H., *Invariant tori in non-degenerate nearly integrable Hamiltonian systems*, Regular and Chaotic Dynamics, 6, 2, 119-204, 2001.

[28] Wayne E., *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. 127, 479-528, 1990.

2000AMS SUBJECT CLASSIFICATION: 37K55, 35L05, 37K50,
Massimiliano Berti, Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università degli Studi Napoli Federico II, Via Cintia, Monte S. Angelo, I-80126, Napoli, Italy, `m.berti@unina.it`.

Luca Biasco, Dipartimento di Matematica, Università di Roma 3, Largo San Leonardo Murialdo, I-00146, Roma, Italy, `biasco@mat.uniroma3.it`.

Supported by the European Research Council under FP7 "New connections between dynamical systems and Hamiltonian PDEs with small divisors phenomena".