A KAM theorem for the defocusing NLS equation ☆

DATASET *in* JOURNAL OF DIFFERENTIAL EQUATIONS · DECEMBER 2011

Impact Factor: 1.68 · DOI: 10.1016/j.jde.2011.11.028

Zhenguo Liang
Fudan University

J. Differential Equations 252 (2012) 4068–4113

Contents lists available at SciVerse ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

A KAM theorem for the defocusing NLS equation ☆

Thomas Kappeler[a,∗], Zhenguo Liang[b]

[a] University of Zürich, Winterthurestrasse 190, CH-8057 Zürich, Switzerland
[b] School of Mathematical Sciences and Key Lab of Mathematics for Nonlinear Science, Fudan University, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 20 October 2011
Available online 19 December 2011

ABSTRACT

In this paper we prove a KAM theorem for the defocusing NLS equation in one space dimension with periodic boundary conditions. The novelty of our result is that it is valid not only near the zero solution, but on the entire Sobolev space $H^N(\mathbb{T},\mathbb{C})$ with $N \in \mathbb{Z}_{\geqslant 1}$. In particular, the invariant tori which persist under small Hamiltonian perturbations might be far away from the zero potential.

1. Introduction

Consider the defocusing nonlinear Schrödinger equation (dNLS) in one space dimension

$$i\partial_t u = -\partial_x^2 u + 2|u|^2 u \tag{1.1}$$

on the Sobolev space $H^N_{\mathbb{C}} \equiv H^N(\mathbb{T},\mathbb{C})$ of complex valued functions on $\mathbb{R}$ of period one,

$$H^N_{\mathbb{C}} = \left\{ u(x) = \sum_{j\in\mathbb{Z}} \hat{u}_j e^{2\pi ijx}\colon\ \|u\|_N < \infty \right\},$$

where

☆ Both authors were partially supported by the Swiss National Science Foundation. In addition, the second author was partially supported by NSFC grant 10701027.

∗ Corresponding author.
E-mail addresses: tk@math.uzh.ch (T. Kappeler), zgliang@fudan.edu.cn (Z. Liang).

0022-0396/$ – see front matter
doi:10.1016/j.jde.2011.11.028

$$\|u\|_N = \left(|\hat{u}_0|^2 + \sum_{j \neq 0} j^{2N} |\hat{u}_j|^2 \right)^{\frac{1}{2}},$$

and $\hat{u}_j$, $j \in \mathbb{Z}$, denote the Fourier coefficients of u. It is an integrable PDE and according to [12], admits global Birkhoff coordinates. Indeed, recall from [12] that the (complex) NLS equation can be viewed as a Hamiltonian system with phase space $H_c^N = H_{\mathbb{C}}^N \times H_{\mathbb{C}}^N$ and Poisson bracket

$$\{F, G\}(\phi_1, \phi_2) = -i \int_0^1 (\partial_{\phi_1} F \partial_{\phi_2} G - \partial_{\phi_2} F \partial_{\phi_1} G)\, dx$$

where $\partial_{\phi_i} F$ denotes the L^2-gradient of F with respect to ϕ_i $(i = 1, 2)$. The Hamiltonian equations of motion are given by

$$\partial_t \phi_1 = -i \partial_{\phi_2} H_{NLS}, \qquad \partial_t \phi_2 = i \partial_{\phi_1} H_{NLS}$$

where

$$H_{NLS}(\phi_1, \phi_2) = \int_0^1 \left(\partial_x \phi_1 \partial_x \phi_2 + \phi_1^2 \phi_2^2 \right) dx.$$

The defocusing NLS equation (1.1) is then obtained by restricting the complex NLS equation to the invariant subspace $H_r^N = \{\phi \in H_c^N \mid \phi_2 = \bar{\phi}_1\}$. Note that H_r^N is a real subspace of H_c^N. To describe the Birkhoff coordinates introduce the model space

$$\mathfrak{h}_r^N = \left\{ (q, p) = (q_j, p_j)_{j \in \mathbb{Z}} \colon q_j, p_j \in \mathbb{R};\ \|q\|_N + \|p\|_N < \infty \right\}$$

where

$$\|p\|_N = \left(p_0^2 + \sum_{j \neq 0} j^{2N} |p_j|^2 \right)^{\frac{1}{2}}.$$

The corresponding complex Hilbert space is denoted by $\mathfrak{h}_c^N$. The space $\mathfrak{h}_r^N$ is endowed with the Poisson structure induced by the standard symplectic form $\sum_{j \in \mathbb{Z}} dq_j \wedge dp_j$. In [12] one finds a detailed proof of the following result on Birkhoff coordinates for (1.1).

Theorem 1.1. *There exists a real analytic map $\Phi : H_r^0 \to \mathfrak{h}_r^0$ with the following properties*

(B1) *Φ is canonical, i.e. for any C^1-functions F, G on $\mathfrak{h}_r^0$, $\{F \circ \Phi, G \circ \Phi\} = \{F, G\} \circ \Phi$.*
(B2) *For any $N \in \mathbb{Z}_{\geqslant 0}$, the restriction of Φ, $\Phi|_{H_r^N} : H_r^N \to \mathfrak{h}_r^N$, is a real analytic diffeomorphism.*
(B3) *Φ defines global Birkhoff coordinates for NLS on H_r^1. That is, on $\mathfrak{h}_r^1$, the transformed NLS Hamiltonian $H_{NLS} \circ \Phi^{-1}$ is a real analytic function of the actions $I_j = \frac{1}{2}(p_j^2 + q_j^2)$, $j \in \mathbb{Z}$.*
(B4) *The differential of Φ at $\phi = 0$, $d_0\Phi$, is the Fourier transform.*

To state our KAM theorem, we need first to introduce some more notations. Let us denote by T_τ, $\tau \in \mathbb{R}$, the flow of translation on L_c^2, i.e. for any $\phi \in L_c^2$, $T_\tau \phi(x) = \phi(x + \tau)$. Note that $\tau \to T_\tau(\phi)$ solves the linear PDE $\partial_\tau \phi = \partial_x \phi$. Actually the latter is a Hamiltonian PDE

$$\partial_\tau \phi_1 = -i \partial_{\phi_2}(i H_2), \qquad \partial_\tau \phi_2 = i \partial_{\phi_1}(i H_2)$$

where H_2 is the Hamiltonian

$$H_2(\phi_1, \phi_2) = \int_0^1 \phi_2 \partial_x \phi_1 \, dx$$

which is the second Hamiltonian in the NLS-hierarchy – see e.g. [12, Section 4]. In particular, H_2 Poisson commutes with H_{NLS},

$$\{H_2, H_{NLS}\} = 0.$$

Actually, a large class of Hamiltonians Poisson commutes with H_2. Indeed, consider a Hamiltonian of the form

$$P(\phi) = \int_0^1 F\big(x, \phi_1(x), \phi_2(x)\big) \, dx$$

where $F = F(x, \zeta, \eta)$ is a polynomial in two complex variables ζ, η, $F(x, \zeta, \eta) = \sum_{\text{finite}} a_{ij}(x)\zeta^i \eta^j$, with coefficients a_{ij} in $C^\infty(\mathbb{T}, \mathbb{C})$. As $H_c^1 \hookrightarrow C^0(\mathbb{T}, \mathbb{C}^2)$ by the Sobolev embedding theorem, for any $N \geqslant 1$, the functional P is defined on H_c^N. Note that for $i = 1, 2$, $\partial_{\phi_i} P = f_i(x, \phi_1(x), \phi_2(x))$ with $f_1 = \partial_\zeta F$, $f_2 = \partial_\eta F$ and that $(\partial_{\phi_1} P, \partial_{\phi_2} P) \in H_c^N$ for any ϕ in H_c^N. By a straightforward computation,

$$\begin{aligned}
\{P, H_2\} &= -i \int_0^1 (\partial_{\phi_1} P \cdot \partial_{\phi_2} H_2 - \partial_{\phi_2} P \cdot \partial_{\phi_1} H_2) \, dx \\
&= -i \int_0^1 (\partial_{\phi_1} P \cdot \partial_x \phi_1 + \partial_{\phi_2} P \cdot \partial_x \phi_2) \, dx \\
&= -i \int_0^1 \frac{d}{dx} F\big(x, \phi(x)\big) \, dx + i \int_0^1 (\partial_x F)\big(x, \phi(x)\big) \, dx.
\end{aligned}$$

As $F(x, \phi(x))$ is 1-periodic, $\int_0^1 \frac{d}{dx} F(x, \phi(x)) \, dx$ vanishes. Furthermore,

$$\int_0^1 (\partial_x F)\big(x, \phi(x)\big) \, dx = \sum_{\text{finite}} \int_0^1 (\partial_x a_{ij})(x) \phi_1(x)^i \phi_2(x)^j \, dx.$$

Hence $\{P, H_2\}$ vanishes identically iff all the coefficients a_{ij} of the polynomial F are constant. More generally, $\{P, H_2\}$ vanishes identically for any Hamiltonian P on its domain of definition if it is of the form

$$P(\phi) = \int_0^1 F\big(\phi_1(x), \phi_2(x)\big) \, dx \tag{1.2}$$

where $F(\zeta, \eta)$ is an arbitrary analytic function on some domain of $\mathbb{C}^2$.

Next we need to introduce notation to parametrize finite-dimensional tori invariant under the defocusing NLS. For $I_A = (I_j)_{j\in A} \in \mathbb{R}^A_{>0}$ with $A \subseteq \mathbb{Z}$ finite and $\mathbb{R}^A_{>0} = (\mathbb{R}_{>0})^A$, denote by $\mathbb{T}_{I_A}$ the torus in $\mathfrak{h}^0_r$ given by

$$\mathbb{T}_{I_A} = \left\{(q_j, p_j)_{j\in\mathbb{Z}}\colon q_j^2 + p_j^2 = 2I_j \ \forall j \in A;\ p_j = q_j = 0\ \forall j \in \mathbb{Z}\setminus A\right\}$$

and by $\mathcal{T}_{I_A}$ its image by Φ^{-1}, $\mathcal{T}_{I_A} = \Phi^{-1}(\mathbb{T}_{I_A})$. For $\Pi \subseteq \mathbb{R}^A_{>0}$ a compact subset of positive Lebesgue measure, denote by $\mathbb{T}_\Pi$ and $\mathcal{T}_\Pi$ the sets

$$\mathbb{T}_\Pi = \bigcup_{I_A\in\Pi} \mathbb{T}_{I_A} \quad \text{and} \quad \mathcal{T}_\Pi = \Phi^{-1}(\mathbb{T}_\Pi).$$

We will consider Hamiltonian perturbations $H_{NLS} + \epsilon K$ on H^N_r, $N \in \mathbb{Z}_{\geqslant 1}$, with the following assumptions on K:

(P1) K is analytic on some open neighborhood $U \equiv U_\Pi$ of $\mathcal{T}_\Pi$ in H^N_c and real valued on $U \cap H^N_r$;
(P2) the L_2-gradients $\partial_{\phi_1} K$, $\partial_{\phi_2} K$ are bounded as functions from U to $H^N_\mathbb{C}$ and verify the normalization condition

$$\sup\left\{\|\partial_{\phi_1} K\|_N + \|\partial_{\phi_2} K\|_N\colon \phi \in U\right\} \leqslant 1;$$

(P3) $\{K, H_2\} \equiv 0$.

Examples of Hamiltonians satisfying conditions (P1)–(P3) are polynomials in ϕ_1, ϕ_2 of the form

$$\sum_{\text{finite}} \int_0^1 a_{ij}\phi_1(x)^i \phi_2(x)^j\, dx$$

where the complex coefficients a_{ij} are constant and satisfy $a_{ij} = \bar{a}_{ji}$.

Our KAM theorem states that for any $A \in \mathbb{Z}$ finite and for any $\epsilon > 0$ sufficiently small, many of the NLS-invariant tori $\mathcal{T}_{I_A}$ persist under perturbation of the NLS Hamiltonian by ϵK with K satisfying (P1)–(P3). Moreover, these tori and their linear flows are only slightly deformed. Let us now state our KAM theorem in a more formal way. Denote by $\mathbb{T}^A$ the $|A|$-dimensional torus $(\mathbb{R}/2\pi\mathbb{Z})^A$ and by $\text{meas}(W)$ the Lebesgue measure of a Lebesgue measurable subset $W \subseteq \mathbb{R}^A$.

Theorem 1.2. *Let $N \in \mathbb{Z}_{\geqslant 1}$ and let $A \subseteq \mathbb{Z}$ be a finite index set. Furthermore let $\Pi \subseteq \mathbb{R}^A_{>0}$ be a compact subset of positive Lebesgue measure. Then for any Hamiltonian K satisfying* (P1)–(P3)*, there exists $\epsilon_0 > 0$ so that the following holds*:

(KAM1) *there exists a family of closed subsets $\Pi_\epsilon \subseteq \Pi$, $|\epsilon| \leqslant \epsilon_0$, with $\lim_{\epsilon\to 0} \text{meas}(\Pi \setminus \Pi_\epsilon) = 0$;*
(KAM2) *for any $|\epsilon| \leqslant \epsilon_0$, there exists a Lipschitz family of real analytic torus embeddings*

$$\Xi_\epsilon : \mathbb{T}^A \times \Pi_\epsilon \to U \cap H^N_r;$$

(KAM3) *for any $|\epsilon| \leqslant \epsilon_0$, there exists a Lipschitz map*

$$f_\epsilon : \Pi_\epsilon \to \mathbb{R}^A$$

such that for any $|\epsilon| \leqslant \epsilon_0$, $I_A \in \Pi_\epsilon$, and $\theta_A \in \mathbb{T}^A$, the curve $t \mapsto \Xi_\epsilon(\theta_A + tf_\epsilon(I_A), I_A)$ is a quasi-periodic solution of

$$\partial_t \phi_1 = -i\partial_{\phi_2} H_{NLS} - i\epsilon\partial_{\phi_2} K, \qquad \partial_t \phi_2 = i\partial_{\phi_1} H_{NLS} + i\epsilon\partial_{\phi_1} K.$$

Related work. Theorem 1.2 confirms that the KAM type theorem of [7], when applied to dNLS, does not only hold near $\phi = 0$, but is actually valid on the entire phase space. In [8], Geng and You prove an abstract KAM result in spaces with exponential weights near an equilibrium solution of certain linear integrable PDEs for a special class of perturbations. They then apply their theorem, among other equations, to the beam equation and to a class of nonlinear Schrödinger equations in arbitrary space dimension. We note that the existence of quasi-periodic solutions of such equations was proved earlier in [2], by the C-W-B method. At the same time, Theorem 1.2 complements the KAM type theorem proved in [10] where instead of imposing condition (P3), dNLS is studied on various invariant subspaces of H_r^N, including the subspace of odd functions and the one of even functions of H_r^N. The perturbations considered in [10] are assumed to induce Hamiltonian vector fields which are tangent to the subspaces considered so that the perturbed equation evolves on these subspaces. For further results on Hamiltonian perturbations of nonlinear Schrödinger equations, see [1–9] and [13–17].

To prove Theorem 1.2 one has to overcome the difficulties caused by the asymptotics of the NLS frequencies $(\omega_j)_{j\in\mathbb{Z}}$. In fact, for $j \in \mathbb{Z}$ large, $\omega_j \sim \omega_{-j}$, i.e., ω_j and ω_{-j} are in 'near resonance'. In earlier work (see [14,10]), NLS-invariant subspaces of H_r^N were considered so that the near resonances mentioned above are no longer relevant when dNLS is restricted to these subspaces. In [8], Geng and You overcome the difficulties caused by these near resonances by imposing a symmetry condition on the perturbations – cf. [8], condition (A4). Condition (P3), introduced above, is a coordinate-free way of formulating their condition (A4). In Section 2 we express condition (P3) in Birkhoff coordinates. It allows to apply a KAM theorem with symmetries, a version of a by now standard abstract KAM theorem of the type obtained in [18] (cf. also [7]), which we state in Section 4. Taking into account the properties of the frequencies of dNLS, discussed in Section 3, Theorem 1.2 is then proved in Section 5. In subsequent work we plan to apply the arguments used in the proof of Theorem 1.2 to other equations as well. In Section 6 we prove the KAM theorem with symmetries stated in Section 4.

2. H_2-symmetry

Let us consider a real analytic Hamiltonian P, defined on an open neighborhood $U \subseteq H_{\mathbb{C}}^N$ of the form introduced in Section 1 with $\Pi \subseteq \mathbb{R}_{>0}^A$ where $A \subseteq \mathbb{Z}$ is finite. We want to compute the Poisson bracket $\{P, iH_2\}$ in Birkhoff coordinates $(q, p) = (q_j, p_j)_{j\in\mathbb{Z}}$. For this purpose it is convenient to introduce action–angle coordinates $I_A = (I_j)_{j\in A}$, $\theta \equiv \theta_A = (\theta_j)_{j\in A}$ and complex coordinates $w = (w_j)_{j\in B}$, $z = (z_j)_{j\in B}$ where $B = \mathbb{Z} \setminus A$. Note that for $j \in A$, one has $I_j > 0$ and hence the angle variable θ_j is well defined $mod\, 2\pi$. The coordinates q, p are related to I_A, θ_A, w, and z as follows: for $j \in A$

$$(q_j, p_j) = \sqrt{2I_j}(\cos\theta_j, -\sin\theta_j),$$

where $I_j = (p_j^2 + q_j^2)/2$ whereas for $j \in B$,

$$w_j = \frac{1}{\sqrt{2}}(q_j - ip_j), \qquad z_j = \frac{1}{\sqrt{2}}(q_j + ip_j).$$

Note that for any $j \in B$, $dw_j \wedge dz_j = i\, dq_j \wedge dp_j$ and $w_j z_j = I_j$ whereas for $j \in A$ one has $d\theta_j \wedge dI_j = dq_j \wedge dp_j$. Assume that $P : U \to \mathbb{C}$ is a real analytic Hamiltonian. Then the Taylor expansion of $P \circ \Phi^{-1}$ at $I_A = \xi \in \Pi$, $w = 0$, $z = 0$ is of the form

$$\sum P_{k\ell mn} e^{ik\cdot\theta} y^\ell w^m z^n \tag{2.1}$$

where $y = I_A - \xi$ and where k, ℓ, m, n are integer vectors, $k \in \mathbb{Z}^A$, $\ell \in \mathbb{Z}^A_{\geqslant 0}$, $m, n \in \mathbb{Z}^B_{\geqslant 0}$ with $|m|, |n| < \infty$. Here $|m| = \sum_{j \in B} m_j$ and in (2.1) we have used the multi-index notation

$$y^\ell = \prod_{j \in A} y_j^{\ell_j}, \qquad k \cdot \theta = \sum_{j \in A} k_j \theta_j, \qquad w^m = \prod_{j \in B} w_j^{m_j}.$$

Further introduce the sequence $\nu = (\nu_j)_{j \in \mathbb{Z}}$ where $\nu_j = j$, for any $j \in \mathbb{Z}$. With the notation $\nu_A = (\nu_j)_{j \in A}$ and $\nu_B = (\nu_j)_{j \in B}$ one then has

$$k \cdot \nu_A = \sum_{j \in A} j k_j \quad \text{and} \quad m \cdot \nu_B = \sum_{j \in B} j m_j.$$

By Theorem 1.1, there exists a neighborhood W of H^0_r in H^0_c so that the Birkhoff map Φ is defined on W and has range $V := \Phi(W) \subseteq \mathfrak{h}^0_c$ so that for any $N \geqslant 0$,

$$\Phi : W \cap H^N_c \to V \cap \mathfrak{h}^N_c \tag{2.2}$$

is a bi-analytic diffeomorphism.

Proposition 2.1.

(i) *On* $\mathfrak{h}^1_c \cap V$, $iH_2 \circ \Phi^{-1}(q, p) = \sum_{j \in \mathbb{Z}} 2\pi j I_j$. *In particular, for* $I_A = \xi + y$ *one has* $iH_2 \circ \Phi^{-1}(q, p) = 2\pi (c + \sum_{j \in A} j y_j + \sum_{j \in B} j w_j z_j)$, *where* $c = \sum_{j \in A} j \xi_j$.

(ii) *Let* $P : U \to \mathbb{C}$ *be given as above. Then, at any point* $I_A = \xi \in \Pi$, $w = 0$, $z = 0$, *the function* $\{P \circ \Phi^{-1}, iH_2 \circ \Phi^{-1}\}$ *admits a Taylor expansion in* $y = I_A - \xi$, w, z *of the form*

$$\{P \circ \Phi^{-1}, iH_2 \circ \Phi^{-1}\} = 2\pi i \sum_{k, \ell, m, n} \big(k \cdot \nu_A + (n - m) \cdot \nu_B\big) P_{k\ell mn} e^{ik\cdot\theta} y^\ell w^m z^n.$$

Proof. (i) follows from [11], Proposition 3.4 and the remark following it and (ii) results from a straightforward computation, taking into account that the Birkhoff coordinates are canonical. □

As an immediate consequence of Proposition 2.1 one has the following

Corollary 2.1. *For* $P : U \to \mathbb{C}$ *with* $\{P, H_2\} \equiv 0$, *the coefficients of the Taylor expansion* (2.1) *of* $P \circ \Phi^{-1}$ *at* $I_A = \xi \in \Pi$, $w = 0$, $z = 0$ *satisfy for any* $k \in \mathbb{Z}^A$, $\ell \in \mathbb{Z}^A_{\geqslant 0}$, $m, n \in \mathbb{Z}^B_{\geqslant 0}$

$$\textit{if } P_{k\ell mn} \neq 0 \quad \textit{then } k \cdot \nu_A + (n - m) \cdot \nu_B = 0. \tag{2.3}$$

Proof. As Φ and hence Φ^{-1} are canonical one has $0 = \{P, H_2\} \circ \Phi^{-1} = \{P \circ \Phi^{-1}, H_2 \circ \Phi^{-1}\}$. The claimed statement then follows from item (ii) of Proposition 2.1. □

As an illustration of implications of (2.3), consider P in Corollary 2.1 with the property that $P \circ \Phi^{-1}$ admits an expansion of the form

$$\sum_{|k| \leqslant K,\ j \in B} p_{kj} e^{ik\cdot\theta} w_j z_{-j}. \tag{2.4}$$

It then follows from (2.3) that $p_{kj} = 0$ for any $j \in B$ with $2|j| > K \max_{i \in A} |i|$. In particular, the sum in (2.4) is finite.

3. NLS frequencies

Let W and V be the open neighborhoods introduced in Section 2 – see (2.2). Note that $H_{NLS} \circ \Phi^{-1}$ is well-defined on $V \cap \mathfrak{h}_c^1$ and analytic there. By Theorem 1.1, $H_{NLS} \circ \Phi^{-1}$ only depends on the action variables I_j, $j \in \mathbb{Z}$, and it then follows that $H_{NLS} \circ \Phi^{-1}$ is a real analytic function of I_j, $j \in \mathbb{Z}$. For any $j \in \mathbb{Z}$,

$$\omega_j := \partial_{I_j} H_{NLS} \circ \Phi^{-1}$$

is called the jth NLS frequency of the (defocusing) NLS. We note that due to Theorem 1.1, the frequencies are analytic functions on V_I^N for any $N \in \mathbb{Z}_{\geqslant 1}$ where $V_I^N \subseteq \ell^{1,2N}(\mathbb{Z}, \mathbb{C})$ denotes the open neighborhood of $\ell^{1,2N}(\mathbb{Z}, \mathbb{R})$ given by

$$V_I^N := \left\{ I = \left(\frac{p_j^2 + q_j^2}{2} \right)_{j \in \mathbb{Z}} : (q_j, p_j)_{j \in \mathbb{Z}} \in V \cap \mathfrak{h}_c^N \right\}. \tag{3.1}$$

Here $\ell_{\mathbb{C}}^{1,\alpha} \equiv \ell^{1,\alpha}(\mathbb{Z}, \mathbb{C})$ denotes the Banach space consisting of all complex sequences $v = (v_j)_{j \in \mathbb{Z}}$ with

$$\|v\|_{\ell^{1,\alpha}} = |v_0| + \sum_{j \neq 0} |j|^\alpha |v_j| < \infty.$$

The expansion of $H_{NLS} \circ \Phi^{-1}$ at $I = 0$ is calculated in [14]. It leads to the following asymptotic expansion of the frequencies in a neighborhood of $I = 0$ in $\ell^{1,2}(\mathbb{Z}, \mathbb{C})$ (see [10, Corollary 3.2])

$$\omega_j = 4\pi^2 j^2 + 4 \sum_{i \neq j} I_i + 2I_j + \mathcal{O}\big(I^2\big)$$

and of their partial derivatives

$$\partial_{I_i} \omega_j = 4 - 2\delta_{ij} + \mathcal{O}(I). \tag{3.2}$$

As an application one obtains the following results (cf. [10]).

Proposition 3.1. *For any $\emptyset \neq A \subseteq \mathbb{Z}$ with $|A| < \infty$, the following functions, when restricted to $\mathbb{R}_{\geqslant 0}^A$, satisfy*

(i) $$\det\big((\partial_{I_i} \omega_j)_{i,j \in A}\big)\big|_{I=0} \neq 0; \quad \textit{in particular } \det\big((\partial_{I_i} \omega_j)_{i,j \in A}\big) \not\equiv 0;$$

(ii) *for any $k \in \mathbb{Z}^A$ and $a, b \in B$,*
 (M1) $k \cdot \omega_A \pm \omega_a \not\equiv 0$;
 (M2) $k \cdot \omega_A \pm (\omega_a + \omega_b) \not\equiv 0$;
 (M3) *if in addition $a \neq b$ then $k \cdot \omega_A + \omega_a - \omega_b \not\equiv 0$.*

Proof. (i) It follows from (3.2) that

$$\det\big((\partial_{I_i} \omega_j)_{i,j \in A}\big)\big|_{I=0} = -(-2)^{|A|}\big(2|A| - 1\big) \neq 0.$$

(ii) Let $A' := A \cup \{a\}$ and $k^{\pm} \in \mathbb{Z}^{A'}$ with $k_j^{\pm} = k_j$ for $j \in A$ and $k_a^{\pm} = \pm 1$. In particular, $k^{\pm} \neq 0$. As by (i), $\det((\partial_{I_i} \omega_j)_{i,j \in A'})$ doesn't vanish identically on $\mathbb{R}_{\geqslant 0}^{A'}$, it follows that there exists $j \in A'$ so that $\partial_{I_j}(\sum_{i \in A} k_i \omega_i \pm \omega_a)$ doesn't vanish identically. This proves (M1). The statements (M2) and (M3) are proved in a similar way. $\square$

Remark 3.1. Consider the case (M2) with $a = b$. Let $A' = A \cup \{a\}$, $k^{\pm} \in A'$ with $k_j^{\pm} = k_j$ for any $j \in A$ and $k_a^{\pm} = \pm 2$. Then $k^{\pm} \neq 0$. Hence again, $\partial_{I_j}(\sum_{i \in A} k_i \omega_i \pm 2\omega_a)$ cannot vanish identically for all $j \in A'$ at the same time.

Proposition 3.1 will allow us to prove Kolmogorov's and Melnikov's conditions for NLS on the entire phase space – see Section 5 for details. Finally we state the asymptotics of the frequencies derived in [10]. There, they are stated for potentials of real type, $\phi \in H_r^1$. The proof of Theorem 5.10 in [10] shows that the asymptotics actually hold on $W \cap H_c^1$.

Proposition 3.2. *For $\phi \in W \cap H_c^1$ or equivalently, for I in V_I^1,*

$$\omega_j = 4\pi^2 j^2 + \mathcal{O}(1) \tag{3.3}$$

locally uniformly on $W \cap H_c^1$. Hence by [13], *Theorem* A.3, *and by Theorem* 1.1,

$$V_I^1 \to \ell^\infty(\mathbb{Z}, \mathbb{C}), \qquad I \mapsto \left(\omega_j - 4j^2\pi^2\right)_{j \in \mathbb{Z}}$$

is real analytic.

Note that the asymptotics (3.3) imply that

$$\omega_j - \omega_{-j} = \mathcal{O}(1). \tag{3.4}$$

It means that the frequencies ω_j and ω_{-j} are not well separated as $|j| \to \infty$. This causes the additional difficulties, alluded to in the introduction, when estimating the measure of the set of good parameters in the proof of Theorem 1.2.

4. An infinite-dimensional KAM theorem with symmetries

Theorem 1.2 is derived from an abstract KAM Theorem with parameters in infinite dimension, first obtained by Kuksin [15] and then further developed by Pöschel [18], cf. also [13]. We need a version of this result taking into account the occurrence of near resonance (3.4). Following the exposition in [13] and [18], consider small perturbations of a family of infinite-dimensional integrable Hamiltonians $H \equiv H(y, u, v; \xi)$ with parameter ξ in the normal form

$$H = \sum_{j \in A} \omega_j(\xi) y_j + \frac{1}{2} \sum_{j \in B} \Omega_j(\xi)\left(u_j^2 + v_j^2\right), \tag{4.1}$$

on the phase space

$$\mathcal{M}^N := \mathbb{T}^A \times \mathbb{R}^A \times \ell^{2,N} \times \ell^{2,N}$$

with coordinates (x, y, u, v) where $A \subseteq \mathbb{Z}$ with $|A| < \infty$, $B = \mathbb{Z} \setminus A$, $N \in \mathbb{Z}_{\geqslant 1}$ and where $\mathbb{T}^A = \mathbb{R}^A / 2\pi\mathbb{Z}^A$ denotes the $|A|$-dimensional torus, conveniently indexed by the set A. Here $\ell^{2,N} \equiv \ell^{2,N}(B, \mathbb{R})$ denotes the Hilbert space of all *real* sequences $u = (u_j)_{j \in B}$ with

$$\|u\|_N^2 = \sum_{j \in B} \langle j \rangle^{2N} |u_j|^2 < \infty,$$

where $\langle j \rangle = 1 \vee |j|$. The 'internal' frequencies, $\omega = (\omega_j)_{j \in A}$, as well as the 'external' ones, $\Omega = (\Omega_j)_{j \in B}$, are real valued and depend on the parameter $\xi \in \Pi \subset \mathbb{R}^A$ and Π is a compact subset of $\mathbb{R}^A$ of positive Lebesgue measure. The symplectic form on $\mathcal{M}^N$ is the standard one given by $\sum_{j\in A} dx_j \wedge dy_j + \sum_{j\in B} du_j \wedge dv_j$. The Hamiltonian equations of motion of H are therefore

$$\dot{x} = \omega(\xi), \qquad \dot{y} = 0, \qquad \dot{u} = \Omega(\xi)v, \qquad \dot{v} = -\Omega(\xi)u,$$

where for any $j \in B$, $(\Omega(\xi)u)_j = \Omega_j(\xi)u_j$. Hence, for any parameter $\xi \in \Pi$, on the $|A|$-dimensional invariant torus,

$$\mathbb{T}_0 = \mathbb{T}^A \times \{0\} \times \{0\} \times \{0\},$$

the flow is rotational with internal frequencies $\omega(\xi) = (\omega_j(\xi))_{j\in A}$. In the normal space, described by the (u, v) coordinates, we have an elliptic equilibrium at the origin, whose frequencies are $\Omega(\xi) = (\Omega_j(\xi))_{j\in B}$. Hence, for any $\xi \in \Pi$, $\mathbb{T}_0$ is an invariant, rotational, linearly stable torus for the Hamiltonian H. Our aim is to prove the persistence of this torus under small perturbations $H + P$ of the integrable Hamiltonian H for a large Cantor set of parameter values ξ. To this end we make assumptions on the frequencies of the unperturbed Hamiltonian H and on the perturbation P.

Assumption $\mathcal{A}$ *(Frequencies).*

(A1) The map $\xi \mapsto \omega(\xi)$ between Π and its image $\omega(\Pi)$ is a homeomorphism which, together with its inverse, is Lipschitz continuous.

(A2) There exists a real sequence $(\overline{\Omega}_j)_{j\in B}$, independent of $\xi \in \Pi$, of the form

$$\overline{\Omega}_j = |j|^d + a_1|j|^{d_1} + \cdots + a_D|j|^{d_D} \tag{4.2}$$

where $d = d_0 > d_1 > \cdots > d_D \geqslant 0$ with $D \in \mathbb{Z}_{\geqslant 0}, d > 1$, and $a_1, \ldots, a_D \in \mathbb{R}$, so that $\xi \mapsto (\Omega_j - \overline{\Omega}_j)_{j\in B}$ is a Lipschitz continuous map on Π with values in $\ell^{\infty,-\delta} \equiv \ell^{\infty,-\delta}(B, \mathbb{R})$ for some $0 \leqslant \delta < 1 \wedge (d-1)$.

(A3) For any (k, e) in $\mathcal{Z} := \{(k, e) \in \mathbb{Z}^A \times \mathbb{Z}^B \setminus (0, 0)\colon\ |e| \leqslant 2;\ k \cdot \nu_A + e \cdot \nu_B = 0\}$ with $e \neq 0$

$$\operatorname{meas}\bigl\{\xi \in \Pi\colon k \cdot \omega(\xi) + e \cdot \Omega(\xi) = 0\bigr\} = 0. \tag{4.3}$$

Recall that for integer vectors such as e, the norm $|e|$ is given by $|e| = \sum_{j\in B} |e_j|$. Furthermore, we note that Assumption (A1) implies that (4.3) holds for $e = 0$.

The second set of assumptions concerns the perturbing Hamiltonian P and its vector field, $X_P = (\partial_y P, -\partial_x P, \partial_v P, -\partial_u P)$. We use the notation $i_\xi X_P$ for X_P evaluated at ξ. Finally, we denote by $\mathcal{M}^N_{\mathbb{C}}$ the complexification of the phase space $\mathcal{M}^N$, $\mathcal{M}^N_{\mathbb{C}} = (\mathbb{C}/2\pi\mathbb{Z})^A \times \mathbb{C}^A \times \ell^{2,N}_{\mathbb{C}} \times \ell^{2,N}_{\mathbb{C}}$. Note that at each point of $\mathcal{M}^N_{\mathbb{C}}$, the tangent space is given by

$$\mathcal{P}^N_{\mathbb{C}} := \mathbb{C}^A \times \mathbb{C}^A \times \ell^{2,N}_{\mathbb{C}} \times \ell^{2,N}_{\mathbb{C}}.$$

Assumption $\mathcal{B}$ *(Perturbation).*

(B1) There exists a neighborhood V of $\mathbb{T}_0$ in $\mathcal{M}^N_{\mathbb{C}}$ such that P is a function on $V \times \Pi$ and its Hamiltonian vector field defines a map

$$X_P : V \times \Pi \to \mathcal{P}^N_{\mathbb{C}}. \tag{4.4}$$

Moreover, $i_\xi X_P$ is real analytic on V for each $\xi \in \Pi$, and $i_w X_P$ is uniformly Lipschitz on Π for each $w \in V$. (Here $i_\xi X_P$ denotes the vector field X_P, evaluated at the parameter value ξ; $i_w X_P$ is defined similarly.)

(B2) $\{P, S\} = 0$ where

$$S = a + b\sum_{j\in A} jy_j + c\sum_{j\in B} j\big(u_j^2 + v_j^2\big)/2 \tag{4.5}$$

with $a \in \mathbb{R}$ and $b, c \in \mathbb{R}\setminus\{0\}$.

To state the KAM theorem we need to introduce some domains and norms. For $s > 0$ and $r > 0$ we introduce the complex $\mathbb{T}_0$-neighborhoods

$$D(s,r) = \big\{|\Im x| < s\big\} \times \big\{|y| < r^2\big\} \times \big\{\|u\|_N + \|v\|_N < r\big\} \subset \mathcal{M}_{\mathbb{C}}^N.$$

Here, for z in $\mathbb{R}^A$ or $\mathbb{C}^A$, $|z| = \max_{j\in A}|z_j|$. For a vector Y in $\mathcal{P}_{\mathbb{C}}^N$ with components (Y_x, Y_y, Y_u, Y_v) introduce the weighted norm

$$\|Y\|_{r,N} = |Y_x| + \frac{1}{r^2}|Y_y| + \frac{1}{r}\|Y_u\|_N + \frac{1}{r}\|Y_v\|_N.$$

Such weights are convenient when estimating the components of a Hamiltonian vector field $X_P = (\partial_y P, -\partial_x P, \partial_v P, -\partial_u P)$ on $D(s,r)$ in terms of r. For a vector field $Y: V\times\Pi \to \mathcal{P}_{\mathbb{C}}^N$ we then define the norms

$$\|Y\|_{r,N;V\times\Pi}^{\sup} = \sup_{(w,\xi)\in V\times\Pi}\big\|Y(w,\xi)\big\|_{r,N},$$

$$\|Y\|_{r,N;V\times\Pi}^{lip} = \sup_{\substack{\xi,\zeta\in\Pi\\ \xi\neq\zeta}} \frac{\|\Delta_{\xi\zeta}Y\|_{r,N;V}^{\sup}}{|\xi-\zeta|},$$

where $\Delta_{\xi\zeta}Y = i_\xi Y - i_\zeta Y$, and

$$\|i_\xi Y\|_{r,N;V}^{\sup} = \sup_{w\in V}\big\|Y(w,\xi)\big\|_{r,N}.$$

In a completely analogous way, the Lipschitz semi-norm of the map $F: \Pi \to \ell^{\infty,-\delta}$ is defined as

$$|F|_{\Pi,\ell^{\infty,-\delta}}^{lip} = \sup_{\substack{\xi,\zeta\in\Pi\\ \xi\neq\zeta}} \frac{\|\Delta_{\xi\zeta}F\|_{\ell^{\infty,-\delta}}}{|\xi-\zeta|}.$$

Finally, let $1 \leqslant M < \infty$ be a constant satisfying

$$|\omega|_{\Pi}^{lip} + |\Omega|_{\Pi,\ell^{\infty,-\delta}}^{lip} \leqslant M. \tag{4.6}$$

Note that if Assumption $\mathcal{A}$ and Assumption $\mathcal{B}$ hold such an M exists.

Theorem 4.1. *Suppose H is a family of Hamiltonians of the form* (4.1) *defined on the phase space $\mathcal{M}^N$, $N \in \mathbb{Z}_{\geqslant 1}$, and depending on parameters in Π so that Assumption $\mathcal{A}$ is satisfied with d and δ. Furthermore, assume that $s > 0$. Then there exist a positive constant γ depending on the finite subset $A \subset \mathbb{Z}$ of* (4.1)*, d, δ, the frequencies ω and Ω of H, and s such that for any perturbed Hamiltonian $H + P$ with P satisfying Assumption $\mathcal{B}$ on a neighborhood V of $\mathbb{T}_0$ in $\mathcal{M}_{\mathbb{C}}^N$, with $D(s,r) \subseteq V$ for some $r > 0$, and the smallness condition*

$$\varepsilon := \|X_P\|^{\sup}_{r,N;D(s,r)\times\Pi} + \frac{\alpha}{M}\|X_P\|^{lip}_{r,N;D(s,r)\times\Pi} \leqslant \alpha\gamma \tag{4.7}$$

for some $0 < \alpha < 1$, the following holds. There exist

(i) *a closed subset $\Pi_* \subset \Pi$, depending on the perturbation P, with* $\operatorname{meas}(\Pi \setminus \Pi_*) \to 0$ *as $\alpha \to 0$,*
(ii) *a Lipschitz family of real analytic torus embeddings $\Psi : \mathbb{T}^A \times \Pi_* \to \mathcal{M}^N$,*
(iii) *a Lipschitz map $f : \Pi_* \to \mathbb{R}^A$,*

such that for any $\xi \in \Pi_*$, $\Psi(\mathbb{T}^A \times \{\xi\})$ is an invariant torus of the perturbed Hamiltonian $H + P$ at ξ and the flow of $H + P$ on this torus is given by

$$\mathbb{T}^A \times \mathbb{R} \to \mathcal{M}^N, \qquad (x,t) \mapsto \Psi\big(x + tf(\xi), \xi\big).$$

Thus for any $x \in \mathbb{T}^A$ and $\xi \in \Pi_*$, the curve $t \mapsto \Psi(x + tf(\xi), \xi)$ is a quasi-periodic solution for the Hamiltonian $i_\xi(H + P)$. Moreover, for any $\xi \in \Pi_*$, the embedding $\Psi(\cdot, \xi) : \mathbb{T}^A \to \mathcal{M}^N$ is real analytic on $D(s/2) = \{|\Im x| < s/2\}$, and

$$\|\Psi - \Psi_0\|^{\sup}_{r,N;\,D(s/2)\times\Pi_*} + \frac{\alpha}{M}\|\Psi - \Psi_0\|^{lip}_{r,N;\,D(s/2)\times\Pi_*} \leqslant \frac{c\varepsilon}{\alpha},$$
$$|f - \omega|^{\sup}_{\Pi_*} + \frac{\alpha}{M}|f - \omega|^{lip}_{\Pi_*} \leqslant c\varepsilon,$$

where

$$\Psi_0 : \mathbb{T}^A \times \Pi \to \mathbb{T}_0, \qquad (x,\xi) \mapsto (x,0,0,0)$$

is the trivial embedding, and c is a positive constant which depends on the same parameters as γ.

Remark 4.1.

(i) Note that (4.2) implies that for any $j \in B$ with $-j \in B$, one has $\overline{\Omega}_{-j} = \overline{\Omega}_j$. Theorem 4.1 continues to hold under a weaker version of (4.2) where the coefficients for $j > 0$ and $j < 0$ might take different values, $a_1^{\pm}, \dots, a_D^{\pm}$. However for the applications we have in mind, condition (A2) as stated suffices. Furthermore, it is straightforward to verify that Theorem 4.1 also continues to hold if δ and/or some of the exponents in (4.2) are negative. We add the condition $\delta \geqslant 0$ and $d_D \geqslant 0$ for convenience.
(ii) Theorem 4.1 remains true if S in Assumption (B2) is replaced by $\sum_{j\in A}\rho(j)y_j + \sum_{j\in B}\rho(j)(u_j^2 + v_j^2)/2$ where $(\rho(j))_{j\in\mathbb{Z}}$ is a real sequence, satisfying for some constants $\kappa_0 > 0$, $\kappa_1 > 0$ and $C_\rho > 0$,

$$\big|\rho(j) - \rho(-j)\big| \geqslant C_\rho |j|^{\kappa_1} > 0, \quad \forall |j| \geqslant \kappa_0 > 0.$$

It turns out that Theorem 4.1 can be shown by adapting the proofs of Theorem A and Corollary C in [18], taking into account the symmetry condition (B2). The latter condition is used in an essential way to obtain the claimed measure estimate of Theorem 4.1 – see Section 6.4.

We conclude this section with a brief outline of the KAM proof in the presence of symmetries. As in the case without symmetries, it employs the rapidly converging iteration scheme of Newton type, involving an infinite sequence of coordinate transformations. At the νth step of the scheme, a Hamiltonian $H_\nu + P_\nu$ is considered where H_ν is a Hamiltonian of the form (4.1), and P_ν is a small perturbation satisfying the symmetry condition $\{P_\nu, S\} = 0$. In the case considered, the Hamiltonian S is in normal form, given by the expression (4.5). One then constructs a canonical transformation Ψ_ν with the property that $(H_\nu + P_\nu) \circ \Psi_\nu$ takes the form $H_{\nu+1} + P_{\nu+1}$ where $H_{\nu+1}$ is again of the form (4.1) and $P_{\nu+1}$ is a much smaller error term than P_ν, satisfying in addition $\{P_{\nu+1}, S\} = 0$. The composition of the infinite sequence of coordinate changes $\Psi_0, \Psi_1, \ldots$ transforms the initial Hamiltonian $H + P$ – at least formally – into a normal form H_∞. For the construction of these coordinate transformations a set of parameters ξ has to be excluded. The measure of this set is then estimated, using that $\{P_\nu, S\} = 0$ for any ν. Let us now describe the construction of the transformation Ψ_ν in more detail. For brevity, we drop the index ν in H_ν, P_ν, R_ν and write

$$H + P = H + R + (P - R),$$

where R is obtained from P by truncating its Fourier and Taylor series expansion. From $\{P, S\} = 0$ one deduces that $\{R, S\} = 0$ as well. The canonical transformation Ψ_ν is constructed as the time-1-map of the flow X_F^t of a Hamiltonian vector field X_F, $\Psi_\nu = X_F^t|_{t=1}$, where the Hamiltonian F satisfies $\{F, S\} = 0$. To find such a Hamiltonian F, one expands $(H + P) \circ X_F^t$ with respect to t at $t = 0$. Recall that for any Hamiltonian G,

$$\frac{d}{dt} G \circ X_F^t = \{G, F\} \circ X_F^t.$$

Hence

$$R \circ X_F^1 = R + \int_0^1 \{R, F\} \circ X_F^t \, dt$$

and

$$H \circ X_F^1 = H + \{H, F\} + \int_0^1 (1 - t) \big\{ \{H, F\}, F \big\} \circ X_F^t \, dt.$$

Altogether, one thus has

$$(H + R) \circ \Psi_\nu = H + R + \{H, F\} + \int_0^1 \big\{ (1 - t)\{H, F\} + R, F \big\} \circ X_F^t \, dt.$$

The latter integral is of quadratic order in R and F and will be part of the new error term. The aim is to determine F in such a way that $H_+ := H + R + \{H, F\}$ is again of the form (4.1) and $\{F, S\} = 0$. Setting $\hat{H} := H_+ - H$, this amounts to solve the system of linear equations

$$\{F, H\} + \hat{H} = R \quad \text{and} \quad \{F, S\} = 0 \tag{4.8}$$

for F and $\hat{H}$ with $\hat{H}$ being of the form (4.1), and R given as above. We will explicitly construct a solution F, $\hat{H}$ of (4.8). It then follows that

$$(1 - t)\{H, F\} + R = (1 - t)\hat{H} + tR,$$

and hence

$$(H + P) \circ \Psi_\nu = H_+ + Q + (P - R) \circ \Psi_\nu \tag{4.9}$$

with

$$Q = \int_0^1 \big\{(1-t)\hat{H} + tR, F\big\} \circ X_F^t \, dt. \tag{4.10}$$

Then H_+ is the new normal form $H_{\nu+1}$ and $Q + (P - R) \circ \Psi_\nu$ the new perturbative term $P_{\nu+1}$. Note that the term Q is of quadratic order in R, F, $\hat{H}$. Furthermore, one has $S \circ X_F^t = S$ as

$$\frac{d}{dt}\big(S \circ X_F^t\big) = \{S, F\} \circ X_F^t = 0 \quad \text{and} \quad S \circ X_F^t\big|_{t=0} = S. \tag{4.11}$$

Hence

$$\big\{(P-R) \circ \Psi_\nu, S\big\} = \big\{(P-R) \circ \Psi_\nu, S \circ \Psi_\nu\big\} = \{P - R, S\} \circ \Psi_\nu = 0$$

and, with $G(t) := (1-t)\hat{H} + tR$,

$$\begin{aligned} \{Q, S\} &= \int_0^1 \big\{\big\{G(t), F\big\} \circ X_F^t, S \circ X_F^t\big\}\, dt \\ &= \int_0^1 \big\{\big\{G(t), F\big\}, S\big\} \circ X_F^t \, dt. \end{aligned}$$

As $\hat{H}$ and S are both in normal form one has $\{\hat{H}, S\} = 0$. Together with the already established identities $\{F, S\} = 0$ and $\{R, S\} = 0$, one then concludes by the Jacobi identity that $\{Q, S\} = 0$. Altogether it follows that $\{P_{\nu+1}, S\} = 0$. In Section 6, we complete the proof of Theorem 4.1.

5. Proof of Theorem 1.2

In this section we show how Theorem 1.2 can be deduced from Theorem 4.1, using similar arguments as in [13] – see also [10]. Recall the set-up of Theorem 1.2. The subset $A \subseteq \mathbb{Z}$ is of finite cardinality, $\Pi \subseteq \mathbb{R}^A_{>0}$ is compact and of positive Lebesgue measure, $\mathbb{T}_\Pi$ is a union of A-tori in h_r^0 indexed by $\xi \in \Pi$, and $\mathcal{T}_\Pi = \Phi^{-1}(\mathbb{T}_\Pi) \subseteq \bigcap_{N \geqslant 0} H_r^N$. Consider the perturbed NLS Hamiltonian $H_\epsilon = H_{NLS} + \epsilon K$, where K is a real analytic map, $K : U \to \mathbb{C}$, with $U \equiv U_\Pi$ a complex neighborhood of $\mathcal{T}_\Pi$ in H_c^N for some $N \in \mathbb{Z}_{\geqslant 1}$, so that properties (P1)–(P3) of Theorem 1.2 hold.

As a first step we apply the Birkhoff map Φ^{-1} of Theorem 1.1,

$$\Phi^{-1} : \mathfrak{h}_r^N \to H_r^N.$$

Since Φ^{-1} is real analytic, there is a complex neighborhood V of $\mathbb{T}_\Pi$ in the complexification of $\mathfrak{h}_r^N$, which is mapped bi-analytically onto the neighborhood U of $\mathcal{T}_\Pi$. If necessary, we choose U and/or V smaller. Hence we have the following diagram where each arrow represents a bi-analytic diffeomorphism given by an approximate restriction of Φ^{-1}:

$$\begin{array}{ccccccc} \mathbb{T}_I & \subset & \mathbb{T}_\Pi & \subset & V & \subset & \mathfrak{h}_c^N \\ \downarrow & & \downarrow & & \downarrow & & \\ \mathcal{T}_I & \subset & \mathcal{T}_\Pi & \subset & U & \subset & H_c^N. \end{array}$$

Now we consider the transformed Hamiltonian $H_\epsilon \circ \Phi^{-1}$. Define $\tilde{H}_{NLS} := H_{NLS} \circ \Phi^{-1}$ and $\tilde{K} := K \circ \Phi^{-1}|_V$ so that

$$H_\epsilon \circ \Phi^{-1} = \tilde{H}_{NLS} + \epsilon \tilde{K}.$$

Then $H_\epsilon \circ \Phi^{-1}$ is real analytic on $V \supset \mathbb{T}_\Pi$. Let us first look at the integrable Hamiltonian $\tilde{H}_{NLS}$. By Theorem 1.1, $\tilde{H}_{NLS}$ depends on (q, p) only through the actions, $I_j = (q_j^2 + p_j^2)/2$, $j \in \mathbb{Z}$. As in Section 3, we view $\tilde{H}_{NLS}$ as a real analytic function of the $I = (I_j)_{j\in\mathbb{Z}}$ defined on V_I^N where V_I^N has been introduced in (3.1). Using Taylor's formula and the definition of the frequencies, $\omega_j(I) := \partial_{I_j}\tilde{H}_{NLS}(I)$, we obtain

$$\tilde{H}_{NLS}\big(I^0 + J\big) = \tilde{H}_{NLS}\big(I^0\big) + \sum_{j\in\mathbb{Z}} \omega_j\big(I^0\big) J_j + Q \tag{5.1}$$

where $Q := \sum_{i,j\in\mathbb{Z}} Q_{ij}(I^0, J) J_i J_j$ and

$$Q_{ij}\big(I^0, J\big) := \int_0^1 (1-t)\partial_{I_i}\omega_j\big(I^0 + tJ\big)\,dt.$$

Note that $\partial_{I_i}\omega_j = \partial_{I_i}\partial_{I_j}\tilde{H}_{NLS}(I)$ and hence the Q_{ij} are symmetric in i and j. Using the asymptotics of ω_j and the analyticity properties of $(\omega_j - 4\pi^2 j^2)_{j\in\mathbb{Z}}$ of Proposition 3.2 it follows from Cauchy's estimate (see e.g. [13, Lemma A.2]) that

$$\sup_{j\in\mathbb{Z}} \left| \sum_{i\in\mathbb{Z}} Q_{ij}\big(I^0, J\big) J_i \right| \leqslant C \| J \|_{\ell^{1,2N}}$$

and hence

$$|Q| = \left| \sum_{i,j} Q_{ij}\big(I^0, J\big) J_i J_j \right| \leqslant C \| J \|^2_{\ell^{1,2N}} \tag{5.2}$$

uniformly in I^0 on some complex neighborhood of $\mathbb{T}_\Pi$ and $\| J \|_{\ell^{1,2N}}$ sufficiently small. Furthermore, by assumption (P2), the Hamiltonian vector field X_K of K, given by $X_K = -i(\partial_{\phi_2} K, -\partial_{\phi_1} K)$, is defined on U and of order 1, $\|X_K\|_N = O(1)$. The Hamiltonian vector field of the transformed Hamiltonian $\tilde{K} = K \circ \Phi^{-1}$,

$$X_{\tilde{K}} = \big(\Phi^{-1}\big)^* X_K = d\Phi \cdot X_K \circ \Phi^{-1},$$

is then defined on V. In view of Theorem 1.1, we may shrink V, if necessary, so that $d\Phi \circ \Phi^{-1}$ is uniformly bounded on V. Hence

$$\|X_{\tilde{K}}\|_N = O(1) \tag{5.3}$$

uniformly on V.

As a second step we introduce symplectic polar coordinates near the tori in the family $\mathbb{T}_\Pi$. For each $\xi = (\xi_j)_{j\in A} \in \Pi$ we then introduce new coordinates by setting for $j \in A$

$$\sqrt{2(\xi_j + y_j)}e^{-ix_j} := q_j + ip_j, \qquad \sqrt{2(\xi_j + y_j)}e^{ix_j} := q_j - ip_j$$

whereas for $j \in B$, the Birkhoff coordinates q_j, p_j play the role of u_j, v_j of Section 4,

$$u_j := q_j, \qquad v_j := p_j.$$

For each $\xi \in \Pi$, this transformation is real analytic and symplectic on $D(s,r) \subseteq V$ for all $s > 0$ and $r > 0$ sufficiently small. In the following we fix such an s, while we keep the freedom of choosing r smaller later in the proof. Using the expansion of $\tilde{H}_{NLS}$ in (5.1) and setting I^0 to be element with components ξ_j for $j \in A$ and 0 for $j \in B$, the integrable Hamiltonian $\tilde{H}_{NLS}$ in the new coordinates is, up to a constant depending only on ξ, given by $H + Q$ with

$$H = H(y,u,v;\xi) = \sum_{j \in A} \omega_j(\xi) y_j + \sum_{j \in B} \Omega_j(\xi)\big(u_j^2 + v_j^2\big)/2, \tag{5.4}$$

where $\Omega_j(\xi) := \omega_j(\xi)$ for $j \in B$, and, according to (5.1), $Q \equiv Q(y,u,v;\xi)$ is given by

$$Q = \sum_{i,j} Q_{ij}(\xi, J) J_i J_j \quad \text{with } J_j = y_j \ (j \in A) \text{ and } J_j = \big(u_j^2 + v_j^2\big)/2 \ (j \in B), \tag{5.5}$$

where we have identified I^0 with ξ. We want to apply Theorem 4.1 for H, defined by (5.4), $P := Q + \epsilon \tilde{K}$, and $S := iH_2 \circ \Phi^{-1}$. We now verify Assumptions (A1)–(A3) and (B1)–(B2). Concerning (A1), recall that by Proposition 3.1, $\det((\frac{\partial \omega_i}{\partial \xi_j})_{i,j \in A}) \not\equiv 0$ on Π. Since this determinant is a real analytic function, it is nonzero almost everywhere on Π. In particular, for any given $\eta > 0$ we may excise from Π a relatively open subset Π_η with $\operatorname{meas}(\Pi_\eta) < \eta$ such that on $\Pi \setminus \Pi_\eta$ the above determinant is uniformly bounded away from zero. Moreover, we may cover $\Pi \setminus \Pi_\eta$ by finitely many closed subsets Π_ι, so that on each subset the map $\xi \to \omega(\xi)$ is a bianalytic homeomorphism onto its image in $\mathbb{R}^A$. Henceforth it suffices to consider each such parameter set Π_ι one at a time.

Next let us verify (A2). The external frequencies Ω_j, $j \in B$, may be written as $\Omega_j(\xi) = \overline{\Omega}_j + \tilde{\Omega}_j(\xi)$ with $\overline{\Omega}_j = 4\pi^2 j^2$ and

$$\tilde{\Omega}_j(\xi) := \Omega_j - \overline{\Omega}_j = \partial_{I_j} \tilde{H}_{NLS}(\xi) - 4\pi^2 j^2.$$

By Proposition 3.2, $\tilde{\Omega} : \xi \mapsto (\tilde{\Omega}_j(\xi))_{j \in B}$ maps Π into $\ell^\infty(B;\mathbb{R})$ and is analytic on a complex neighborhood of Π with values in $\ell^\infty(B,\mathbb{C})$. Hence $\tilde{\Omega}$ is also Lipschitz by Cauchy's estimate. In summary, Assumption (A2) is satisfied with $d = 2$ and $\delta = 0$.

To see that Assumption (A3) holds note that by Proposition 3.1, $k \cdot \omega(\xi) + e \cdot \Omega(\xi) \not\equiv 0$ for every $k \in \mathbb{Z}^A$ and $e \in \mathbb{Z}^B$ with $1 \leqslant |e| \leqslant 2$. Since each such expression is real analytic in ξ, its zero set is a set of measure zero and (A3) follows.

Toward Assumption (B2), first note that by Proposition 2.1(i), $iH_2 \circ \Phi^{-1}$ is of the form S, described in (B2). As Φ^{-1} is canonical and Q, given by (5.5), is in normal form, it follows that $\{Q, iH_2 \circ \Phi^{-1}\} = 0$. Furthermore, in view of Assumption (P3), $\{K \circ \Phi^{-1}, iH_2 \circ \Phi^{-1}\} = \{K, iH_2\} \circ \Phi^{-1} = 0$. Altogether we have shown that

$$\big\{P, iH_2 \circ \Phi^{-1}\big\} = \big\{Q + \epsilon \tilde{K}, iH_2 \circ \Phi^{-1}\big\} = 0$$

and Assumption (B2) follows.

It remains to check Assumption (B1). As already mentioned, the perturbation P consists of two parts

$$P = Q + \epsilon \tilde{K}.$$

In view of the definition (5.5), the Hamiltonian vector field of Q is given by

$$X_Q = (\partial_y Q, 0, \partial_v Q, -\partial_u Q).$$

To estimate the size of X_Q we apply Cauchy's estimate to each of its components. From the estimate (5.2) together with the bounds $|y| < r^2$ and $\|u\|_N + \|v\|_N < r$ one then gets that $\|X_Q\|^{\sup}_{r,N;D(s,r)\times \Pi_\iota} \leqslant$

cr^2. As Q analytically extends to some complex neighborhood of Π, again by Cauchy's estimate, one obtains a similar bound for the Lipschitz semi-norm of X_Q,

$$\|X_Q\|^{lip}_{r,N;D(s,r)\times\Pi_\iota} \leqslant cr^2.$$

Taking the weight factors in the norm $\|\cdot\|_{r,N}$ into account and using (5.3), one gets the following estimate for the second term in P, $\|X_{\tilde K}\|^{\sup}_{r,N;D(s,r)\times\Pi_\iota} \leqslant \frac{c}{r^2}$. Arguing as for Q, one obtains a bound of the same form for the Lipschitz semi-norm, $\|X_{\tilde K}\|^{lip}_{r,N;D(s,r)\times\Pi_\iota} \leqslant \frac{c}{r^2}$. Altogether, we thus have shown that for any $0<\alpha\leqslant M$ and $r>0$ small enough,

$$\|X_{Q+\epsilon\tilde K}\|^{\sup}_{r,N;D(s,r)\times\Pi_\iota} + \frac{\alpha}{M}\|X_{Q+\epsilon\tilde K}\|^{lip}_{r,N;D(s,r)\times\Pi_\iota} \leqslant C\left(r^2+\frac{\epsilon}{r^2}\right). \tag{5.6}$$

In particular, we have verified Assumption (B1) with V in (4.4) given by $D(s,r)$.

To meet the smallness condition (4.7) of Theorem 4.1 for $P=Q+\epsilon\tilde K$ choose r and α as follows

$$r^2=\sqrt{\epsilon}, \qquad \alpha=\frac{2C}{\gamma}\sqrt{\epsilon}, \tag{5.7}$$

with ϵ so small that $\alpha<1$. Here, C is taken from the preceding estimate, and γ is taken from Theorem 4.1. We then obtain

$$C\left(r^2+\frac{\epsilon}{r^2}\right)=2C\sqrt{\epsilon}=\gamma\alpha.$$

The estimate (5.6) then implies that (4.7) holds. The conclusions of Theorem 1.2 now follow from the ones of Theorem 4.1. Let us only comment on the measure theoretic statement of Theorem 1.2. By Theorem 4.1 and the choice (5.7) of α, for each Π_ι there exists $\Pi_{\iota,\epsilon}\subseteq\Pi_\iota$ so that

$$\operatorname{meas}(\Pi_\iota\setminus\Pi_{\iota,\epsilon})\to 0 \quad \text{as } \epsilon\to 0.$$

Finitely many sets Π_ι cover the parameter domain Π up to a set of measure η. By first choosing η and then ϵ small enough we can assure that

$$\operatorname{meas}\left(\Pi\setminus\bigcup_\iota \Pi_{\iota,\epsilon}\right)\to 0 \quad \text{as } \epsilon\to 0.$$

The proof of Theorem 1.2 is now complete. $\square$

6. Proof of Theorem 4.1

The aim of this section is to prove Theorem 4.1. It is based on the proof of a KAM theorem without symmetries presented in [18].

6.1. Linearized equation

In this subsection we study the linear system (4.8)

$$\{F,H\}+\hat H=R \quad \text{and} \quad \{F,S\}=0 \tag{6.1}$$

where H, S, R are given Hamiltonians and F, $\hat{H}$ are to be determined. It is convenient to introduce complex coordinates $w=(w_j)_{j\in B}$, $z=(z_j)_{j\in B}$ defined by

$$w=\frac{1}{\sqrt{2}}(u-iv) \quad \text{and} \quad z=\frac{1}{\sqrt{2}}(u+iv).$$

In these complex coordinates, the Hamiltonians $H\equiv H(y,w,z;\xi)$ and $S=S(y,w,z)$ are given by

$$H=\sum_{j\in A}\omega_j y_j+\sum_{j\in B}\Omega_j w_j z_j,$$
$$S=a+b\sum_{j\in A} j y_j+c\sum_{j\in B} j w_j z_j.$$

Here $\omega_j=\omega_j(\xi)$ and $\Omega_j=\Omega_j(\xi)$ depend on the parameter ξ and a,b,c are real constants with $b\neq 0, c\neq 0$. In the sequel we will assume that the constants a, b, c are given by $a=0$, $b=c=1$ – the case where $a\in\mathbb{R}$, $b,c\in\mathbb{R}\setminus\{0\}$ are arbitrary is proved in the same way. H is assumed to be *regular* on the domain $D(s,r)\times\Pi$ in the sense that for each $\xi\in\Pi$, $i_\xi H\equiv H(\cdot\,;\xi)$ is real analytic on $D(s,r)$ and $H(y,w,z;\cdot)$ is Lipschitz in ξ, uniformly on $D(s,r)$. The Hamiltonian $R=R(x,y,w,z;\xi)$ is also assumed to be regular on $D(s,r)\times\Pi$ and to be of the form

$$R=\sum_{2|l|+|m+n|\leqslant 2} R_{klmn}e^{ikx}y^l w^m z^n. \tag{6.2}$$

Here and in the sequel, a sum such as in (6.2) extends over all integer vectors $k\in\mathbb{Z}^A$, $l\in\mathbb{Z}^A_{\geqslant 0}$, and $m,n\in\mathbb{Z}^B_{\geqslant 0}$. Hence R is a polynomial in y,w,z of degree two – the y_j, $j\in A$, being variables of degree two – whose coefficients depend regularly on x and ξ in the sense above. Moreover, the Hamiltonian vector field $X_R\equiv X_R(x,y,w,z;\xi)$ associated with R is assumed to be a regular map

$$X_R: D(s,r)\times\Pi\to\mathcal{P}^N_{\mathbb{C}} \tag{6.3}$$

and R is assumed to satisfy the symmetry conditions

$$\{R,S\}=0. \tag{6.4}$$

The latter identity means that for any $k\in\mathbb{Z}^A$, $l\in\mathbb{Z}^A_{\geqslant 0}$, $m,n\in\mathbb{Z}^B_{\geqslant 0}$ and $\xi\in\Pi$

$$R_{klmn}\cdot\big(k\cdot\nu_A+(n-m)\cdot\nu_B\big)=0. \tag{6.5}$$

The mean value $[R]$ of R is defined by

$$[R]=\sum_{|l|+|m|=1} R_{0lmm}y^l w^m z^m.$$

Note that $[R]$ is of the same form as H. To shorten notation we drop the subscripts N and Π in $\|\cdot\|^{\sup}_{r,N;D(s,r)\times\Pi}$ and write $\|\cdot\|^{lip}_{r,D(s,r)}$ instead of $\|\cdot\|^{lip}_{r,N;D(s,r)\times\Pi}$ as well as $|\omega|^{lip}$, $|\Omega|^{lip}_{\ell^\infty,-\delta}$ instead of $|\omega|^{lip}_{\Pi}$, $|\Omega|^{lip}_{\Pi,\ell^\infty,-\delta}$. In the sequel, we will always assume that Ω satisfies condition (A2) of Section 4, i.e., $\Omega=\overline{\Omega}+\tilde{\Omega}$ where $\overline{\Omega}$ is independent of ξ with

$$\overline{\Omega}_j=|j|^d+\cdots \quad \text{for some } d>1$$

and $\tilde{\Omega} = \Omega - \overline{\Omega}$ is a Lipschitz map, $\tilde{\Omega} : \Pi \to \ell^{\infty,-\delta}(B, \mathbb{R})$ for some $0 \leqslant \delta < d - 1$. Finally, for any $e \in \mathbb{Z}^B$ with finite support we define

$$|e|_\delta := \sum_{j \in B} \langle j \rangle^\delta |e_j|.$$

Lemma 6.1. *Let $\alpha > 0$, $s > 0$, $r > 0$ and assume that H and R are regular on $D(s,r) \times \Pi$ and that R satisfies* (6.3) *and* (6.4). *Moreover assume that for any $\xi \in \Pi$ and any $(k, e) \in \mathcal{Z}$*

$$\left| k \cdot \omega(\xi) + e \cdot \Omega(\xi) \right| \geqslant \alpha A_k^{-1} \cdot 1 \vee |e|_\delta^{\frac{1}{2}} \tag{6.6}$$

where the sequence $(A_k)_{k \in \mathbb{Z}^A} \subseteq \mathbb{R}$ satisfies $A_k \geqslant 1$. Then the linear system (6.1) has a unique solution F, $\hat{H}$ when normalized by $[F] = 0$, $[\hat{H}] = \hat{H}$. The following estimates hold:

$$\|X_{\hat{H}}\|^{\sup}_{r, D(s,r)} \leqslant \|X_R\|^{\sup}_{r, D(s,r)}, \qquad \|X_{\hat{H}}\|^{lip}_{r, D(s,r)} \leqslant \|X_R\|^{lip}_{r, D(s,r)}$$

and for any $0 < \sigma \leqslant s$

$$\|X_F\|^{\sup}_{r, D(s-\sigma, r)} \leqslant \frac{16 B_\sigma}{\alpha} \|X_R\|^{\sup}_{r, D(s,r)},$$

$$\|X_F\|^{lip}_{r, D(s-\sigma, r)} \leqslant \frac{25 B_\sigma}{\alpha} \left(\|X_R\|^{lip}_{r, D(s,r)} + \frac{M}{\alpha} \|X_R\|^{\sup}_{r, D(s,r)} \right),$$

where $M \geqslant 1$ satisfies $M \geqslant |\omega|^{lip} + |\Omega|^{lip}_{\ell^{\infty,-\delta}}$ and $B_\sigma = (2^{|A|} \sum_{k \in \mathbb{Z}^A} \langle k \rangle^4 A_k^4 e^{-2|k|\sigma})^{\frac{1}{2}}$.

Proof. We are looking for solutions F and $\hat{H}$ of (6.1) which admit expansions of the form

$$F = \sum_{2|l| + |m+n| \leqslant 2} F_{klmn} e^{ik \cdot x} y^l w^m z^n$$

and

$$\hat{H} = \sum_{|k| + |m| = 1} \hat{H}_{0lmm} y^l w^m z^m.$$

Use that $\{x_j, y_j\} = 1$ for any $j \in A$ and $\{w_j, z_j\} = 1$ for any $j \in B$ and that all other brackets between coordinate functions vanish to conclude

$$\left\{ e^{ik \cdot x} y^l w^m z^n, y_j \right\} = i k_j e^{ik \cdot x} y^l w^m z^n \tag{6.7}$$

and

$$\left\{ e^{ik \cdot x} y^l w^m z^n, w_j z_j \right\} = i(n_j - m_j) e^{ik \cdot x} y^l w^m z^n. \tag{6.8}$$

It then follows that

$$\{F, H\} = \sum i F_{klmn} \big(k \cdot \omega + (n - m) \cdot \Omega \big) e^{ik \cdot x} y^l w^m z^n.$$

One then finds by comparison of coefficients that the system (6.1) admits the solution F and $\hat{H}$ given by

$$iF_{klmn} = \begin{cases} \frac{R_{klmn}}{k\cdot\omega+(n-m)\cdot\Omega}, & \text{if } R_{klmn} \neq 0 \text{ and } (k, n-m) \neq (0,0), \\ 0, & \text{otherwise,} \end{cases}$$

$$\hat{H}_{0lmm} = R_{0lmm}. \tag{6.9}$$

By (6.5), one has $R_{klmn}(k \cdot \nu_A + (n-m) \cdot \nu_B) = 0$. Thus the small divisor conditions (6.6) guarantee that F_{klmn} is well defined for any k, l, m, n. Furthermore $[F] = 0$ and $[\hat{H}] = \hat{H}$. When normalized in this way, F and $\hat{H}$ are uniquely determined. Clearly, one has $\{\hat{H}, S\} = 0$ and

$$\{F, S\} = \sum iF_{klmn}\big(k \cdot \nu_A + (n-m) \cdot \nu_B\big)e^{ik\cdot x}y^l w^m z^n$$

which by the definition of F equals

$$\sum \frac{R_{klmn}(k \cdot \nu_A + (n-m) \cdot \nu_B)}{k \cdot \omega + (n-m) \cdot \Omega} e^{ik\cdot x}y^l w^m z^n.$$

In view of (6.5) it then follows that $\{F, S\} = 0$. To derive the claimed estimates we decompose $R = R^0 + R^1 + R^2$ and write

$$R^0 = R^{00} = R^{000} + R^{001}; \qquad R^1 = R^{10} + R^{01}; \qquad R^2 = R^{20} + R^{11} + R^{02}$$

where R^a comprises all terms with $|m+n| = a$,

$$R^{000} = \sum R_{k000}e^{ik\cdot x}, \qquad R^{001} = \sum_{j\in A} \mathcal{R}_j^{001} y_j = \sum_{|l|=1} R_{kl00}e^{ik\cdot x}y^l$$

and R^{ab} are given by

$$R^{10} = \sum_{j\in B} \mathcal{R}_j^{10} w_j, \qquad R^{01} = \sum_{j\in B} \mathcal{R}_j^{01} z_j,$$

$$R^{20} = \sum_{i,j\in B} \mathcal{R}_{ij}^{20} w_i w_j, \qquad R^{11} = \sum_{i,j\in B} \mathcal{R}_{ij}^{11} z_i w_j, \qquad R^{02} = \sum_{i,j\in B} \mathcal{R}_{ij}^{02} z_i z_j.$$

The coefficients $\mathcal{R}_j^{ab}$ and $\mathcal{R}_{ij}^{ab}$ are given by the corresponding derivatives of R with respect to the components of w and z at $w = 0$, $z = 0$ and depend on x and ξ whereas the coefficients $\mathcal{R}_j^{001}$ are given by $\partial_{y_j} R|_{y=0}$ and also depend on x and ξ. So e.g. for any $j \in B$,

$$\mathcal{R}_j^{10} = \partial_{w_j} R|_{w=0,\, z=0} = \sum R_{k0m^j0}e^{ik\cdot x} \quad \text{and} \quad m^j = (\delta_{ji})_{\in B}.$$

The functions F and $\hat{H}$ are decomposed in a similar way. The linear system $\{F, H\} + \hat{H} = R$ then may be written as follows

$$\{F^{ab}, H\} = R^{ab} - [R^{ab}], \qquad \hat{H}^{ab} = [R^{ab}]$$

and it suffices to obtain the claimed estimates for each of the Hamiltonians F^{ab}, $\hat{H}^{ab}$ individually. By the definition of $\hat{H}^{ab}$, the claimed estimates for $\hat{H}$ held trivially. Concerning the terms F^{ab}, they all can be treated in a similar fashion. So we concentrate on F^{10} and F^{11} only. Let us begin with F^{10}. We want to estimate $X_{F^{10}} = (0, -\partial_x F^{10}, 0, i\partial_w F^{10})$ in terms of X_R. It is convenient to introduce the notation $\dot{\mathcal{R}} = (\dot{\mathcal{R}}_j)_{j\in B}$ with

$$\dot{\mathcal{R}}_j \equiv \mathcal{R}_j^{10} = \sum_k \dot{\mathcal{R}}_{kj} e^{ik\cdot x} = \partial_{w_j} R|_{w=0,\, z=0}.$$

By the definition of the norm $\|\cdot\|^{\sup}_{r,D(s,r)}$ one has

$$\|\dot{\mathcal{R}}\|^{\sup}_{D(s)} \leqslant r\|X_R\|^{\sup}_{r,D(s,r)} \tag{6.10}$$

where $D(s) := \{x \in \mathbb{C}^A/2\pi\mathbb{Z}^A\colon\ |\Im x| < s\}$. By assumption, $\dot{\mathcal{R}}\colon D(s) \to \ell^{2,N} \equiv \ell^{2,N}(B,\mathbb{C})$ is analytic and has a Fourier expansion with Fourier coefficients $(\dot{\mathcal{R}}_{kj})_{j\in B}$, $k \in \mathbb{Z}^A$, satisfying the L^2-bound

$$\sum_{k\in\mathbb{Z}^A} \big\|(\dot{\mathcal{R}}_{kj})_{j\in B}\big\|_N^2 e^{2|k|s} \leqslant 2^{|A|}\big(\|\dot{\mathcal{R}}\|^{\sup}_{D(s)}\big)^2.$$

Actually, due to the symmetry conditions (6.5), for each $k \in \mathbb{Z}^A$, $\dot{\mathcal{R}}_{kj} = 0$, and hence $\dot{\mathcal{F}}_{kj} = 0$, for any $j \in B$ except possibly for $j = j(k) = k\cdot \nu_A$.

For any $k \in \mathbb{Z}^A$, $\dot{\mathcal{R}}_{kj(k)}\colon \Pi \to \mathbb{C}$ is Lipschitz and the corresponding coefficient $\dot{\mathcal{F}}_{kj(k)}$ of $\dot{\mathcal{F}}$ is given by

$$i\dot{\mathcal{F}}_{kj(k)} = \frac{\dot{\mathcal{R}}_{kj(k)}}{k\cdot\omega - \Omega_{j(k)}}.$$

By the small divisors assumption (6.6), $|k\cdot\omega - \Omega_{j(k)}| \geqslant \alpha A_k^{-1}$ for any $k \in \mathbb{Z}^A$. Hence

$$\big\|(\dot{\mathcal{F}}_{kj})_{j\in B}\big\|_N \leqslant \frac{A_k}{\alpha}\big\|(\dot{\mathcal{R}}_{kj})_{j\in B}\big\|_N \tag{6.11}$$

and thus

$$\begin{aligned}
\|\dot{\mathcal{F}}\|^{\sup}_{D(s-\sigma)} &\leqslant \sum_k \big\|(\dot{\mathcal{F}}_{kj})_{j\in B}\big\|_N e^{|k|(s-\sigma)} \\
&\leqslant \frac{1}{\alpha}\Big(\sum_k A_k^2 e^{-2|k|\sigma}\Big)^{\frac12}\Big(\sum_k \big\|(\dot{\mathcal{R}}_{kj})_{j\in B}\big\|_N^2 e^{2|k|s}\Big)^{\frac12} \\
&\leqslant \frac{B_\sigma}{\alpha}\|\dot{\mathcal{R}}\|^{\sup}_{D(s)}
\end{aligned}$$

or, as $\dot{\mathcal{F}} = \partial_w F^{10}$,

$$\frac{1}{r}\big\|\partial_w F^{10}\big\|^{\sup}_{D(s-\sigma)} \leqslant \frac{B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}.$$

The other nonzero component of $X_{F^{10}}$ is given by $\partial_x F^{10} = \sum_k ik(\sum_{j\in B}\dot{\mathcal{F}}_{kj}w_j)e^{ik\cdot x}$. As by (6.11),

$$\Big|\sum_{j\in B}\dot{\mathcal{F}}_{kj}w_j\Big| \leqslant \big\|(\dot{\mathcal{F}}_{kj})_{j\in B}\big\|_N\|w\|_N \leqslant \frac{A_k}{\alpha}\big\|(\dot{\mathcal{R}}_{kj})_{j\in B}\big\|_N\|w\|_N,$$

one gets

$$\frac{1}{r}\left\|\partial_x F^{10}\right\|^{\sup}_{D(s-\sigma,r)} \leqslant \sum_k \frac{A_k}{\alpha}|k|\left\|(\dot{\mathcal{R}}_{kj})_{j\in B}\right\|_N e^{|k|(s-\sigma)} \leqslant \frac{B_\sigma}{\alpha}\|\dot{\mathcal{R}}\|^{\sup}_{D(s)}.$$

It then follows from (6.10) that

$$\frac{1}{r^2}\left\|\partial_x F^{10}\right\|^{\sup}_{D(s-\sigma,r)} \leqslant \frac{B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}.$$

Altogether, we have proved that

$$\|X_{F^{10}}\|^{\sup}_{r,D(s-\sigma,r)} \leqslant \frac{1}{r^2}\left\|\partial_x F^{10}\right\|^{\sup}_{D(s-\sigma,r)} + \frac{1}{r}\left\|\partial_w F^{10}\right\|^{\sup}_{r,D(s-\sigma,r)} \leqslant \frac{2B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(r,s)}.$$

Next we want to estimate $\|X_{F^{10}}\|^{lip}_{r,D(s-\sigma,r)}$. Let $\alpha_k := k\cdot\omega - \Omega_{j(k)}$ and $\Delta \equiv \Delta_{\xi\zeta}$ for $\xi,\zeta\in\Pi$. Then, for any $k\in\mathbb{Z}^A$ and with $j\equiv j(k)$,

$$i\Delta\dot{\mathcal{F}}_{kj} = \Delta\big(\alpha_k^{-1}\dot{\mathcal{R}}_{kj}\big) = \alpha_k^{-1}\Delta(\dot{\mathcal{R}}_{kj}) + \dot{\mathcal{R}}_{kj}\Delta\big(\alpha_k^{-1}\big)$$

and

$$-\Delta\big(\alpha_k^{-1}\big) = \frac{\Delta\alpha_k}{\alpha_k(\xi)\cdot\alpha_k(\zeta)} = \frac{k\cdot\Delta\omega - \Delta\Omega_j}{\alpha_k(\xi)\cdot\alpha_k(\zeta)}.$$

By the small divisors assumption (6.6), $|\alpha_k| \geqslant \alpha A_k^{-1}\langle j\rangle^{\frac{\delta}{2}}$. Recall that $\Omega_j = \overline{\Omega}_j + \tilde{\Omega}_j$, where $\Delta\Omega_j = \Delta\tilde{\Omega}_j = O(|j|^\delta)$. One then gets

$$\big|\Delta\alpha_k^{-1}\big| \leqslant \frac{A_k^2}{\alpha^2}\left(|k||\Delta\omega| + \frac{|\Delta\Omega_j|}{\langle j\rangle^\delta}\right)$$

and thus

$$\left\|(\Delta\dot{\mathcal{F}}_{kj})_{j\in B}\right\|_N \leqslant \frac{A_k}{\alpha}\left\|(\Delta\dot{\mathcal{R}}_{kj})_{j\in B}\right\|_N + \frac{A_k^2}{\alpha^2}\big(|k||\Delta\omega| + |\Delta\Omega|_{\ell^\infty,-\delta}\big)\left\|(\dot{\mathcal{R}}_{kj})_{j\in B}\right\|_N. \tag{6.12}$$

Summing up to the Fourier series as before we obtain

$$\|\Delta\dot{\mathcal{F}}\|_{D(s-\sigma)} \leqslant \frac{B_\sigma}{\alpha}\|\Delta\dot{\mathcal{R}}\|^{\sup}_{D(s)} + \frac{B_\sigma}{\alpha^2}\big(|\Delta\omega| + |\Delta\Omega|_{\ell^\infty,-\delta}\big)\|\dot{\mathcal{R}}\|^{\sup}_{D(s)}.$$

Dividing this inequality by $|\xi-\zeta|$ and taking the supremum over $\xi\neq\zeta$ in Π yields, with $\dot{\mathcal{F}} = \partial_w F^{10}$,

$$\frac{1}{r}\left\|\partial_w F^{10}\right\|^{lip}_{D(s-\sigma)} = \frac{1}{r}\|\dot{\mathcal{F}}\|^{lip}_{D(s-\sigma)} \leqslant \frac{B_\sigma}{\alpha}\left(\|X_R\|^{lip}_{r,D(s,r)} + \frac{M}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}\right)$$

where we used that $M \geqslant |\omega|^{lip} + |\Omega|^{lip}_{\ell^\infty,-\delta}$.

Now let us estimate the Lipschitz semi-norm of the other nonzero component $\partial_x F^{10}$ of $X_{F^{10}}$. Note that

$$\Delta\partial_x F^{10} = \sum_k ik\Big(\sum_{j\in B}\Delta\dot{\mathcal{F}}_{kj} w_j\Big)e^{ik\cdot x}.$$

Hence by (6.12)

$$\begin{aligned}
&\frac{1}{r}\big\|\Delta\partial_x F^{10}\big\|^{\sup}_{D(s-\sigma,r)}\\
&\quad\leqslant \sum_k |k|\big\|(\Delta\dot{\mathcal{F}}_{kj})_{j\in B}\big\|_N e^{|k|(s-\sigma)}\\
&\quad\leqslant \sum_k |k|e^{|k|(s-\sigma)}\bigg(\frac{A_k}{\alpha}\big\|(\Delta\dot{\mathcal{R}}_{kj})_{j\in B}\big\|_N + \frac{A_k^2}{\alpha^2}\big(|k||\Delta\omega| + |\Delta\Omega|_{\ell^\infty,-\delta}\big)\big\|(\dot{\mathcal{R}}_{kj})_{j\in B}\big\|_N\bigg)
\end{aligned}$$

and thus by the definition of B_σ

$$\frac{1}{r^2}\big\|\Delta\partial_x F^{10}\big\|^{\sup}_{D(s-\sigma,r)} \leqslant \frac{B_\sigma}{\alpha}\|\Delta\dot{\mathcal{R}}\|^{\sup}_{D(s)} + \frac{B_\sigma}{\alpha^2}\big(|\Delta\omega| + |\Delta\Omega|_{\ell^\infty,-\delta}\big)\|\dot{\mathcal{R}}\|^{\sup}_{D(s)}$$

leading as above to the estimate

$$\frac{1}{r^2}\big\|\partial_x F^{10}\big\|^{lip}_{D(s-\sigma,r)} \leqslant \frac{B_\sigma}{\alpha}\bigg(\|X_R\|^{lip}_{r,D(s,r)} + \frac{M}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}\bigg).$$

Altogether we have shown

$$\|X_{F^{10}}\|^{lip}_{r,D(s-\sigma,r)} \leqslant \frac{2B_\sigma}{\alpha}\bigg(\|X_R\|^{lip}_{r,D(s,r)} + \frac{M}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}\bigg).$$

Let us now turn our attention to the term $\ddot{F} = F^{11}$. We want to estimate

$$X_{F^{11}} = \big(0, -\partial_x F^{11}, -i\partial_z F^{11}, i\partial_w F^{11}\big).$$

Recall that $\ddot{R} \equiv R^{11} = \sum_{i,j}\mathcal{R}^{11}_{ij} z_i w_j$. For convenience, let $\ddot{\mathcal{R}}_{ij} := \mathcal{R}^{11}_{ij}$ and denote the operator corresponding to $(\ddot{\mathcal{R}}_{ij})_{i,j\in B}$ by $\ddot{\mathcal{R}}$. Note that $\ddot{\mathcal{R}}_{ij} = \partial_{w_j}\partial_{z_i} R|_{w=0,z=0}$. Due to the special form of R, $\ddot{\mathcal{R}}$ can be viewed as the Jacobian of $\partial_z R|_{z=0}$ with respect to w at $w=0$. In particular, it can be viewed as a linear operator on $\ell^{2,N} \equiv \ell^{2,N}(B,\mathbb{C})$. Hence by the Cauchy estimate for analytic maps between Banach spaces (cf. [18, Lemma A.3])

$$\|\ddot{\mathcal{R}}\|^{\sup}_{D(s)} \leqslant \frac{1}{r}\|\partial_z R\|^{\sup}_{D(s,r)} \leqslant \|X_R\|^{\sup}_{r,D(s,r)} \tag{6.13}$$

where $\|\ddot{\mathcal{R}}\|$ denotes the operator norm on $\ell^{2,N}(B,\mathbb{C})$. This is equivalent to the statement that $\tilde{\mathcal{R}} = (\langle i\rangle^{-N}\ddot{\mathcal{R}}_{ij}\langle j\rangle^N)_{i,j\in B}$ is a bounded operator on $\ell^2 \equiv \ell^2(B,\mathbb{C})$. Expanding $\tilde{\mathcal{R}}$ into its Fourier series with operator valued coefficients $\tilde{\mathcal{R}}_k = (\tilde{\mathcal{R}}_{k,ij})_{i,j\in B}$, $k\in\mathbb{Z}^A$, one gets as before

$$\sum_{k\in\mathbb{Z}^A}\|\tilde{\mathcal{R}}_k\|^2 e^{2|k|s} \leqslant 2^{|A|}\big(\|\tilde{\mathcal{R}}\|^{\sup}_{D(s)}\big)^2$$

where now $\|\tilde{\mathcal{R}}_k\|$ denotes the operator norm of $\tilde{\mathcal{R}}_k : \ell^2 \to \ell^2$. The corresponding coefficient $\tilde{\mathcal{F}}_k = (\tilde{\mathcal{F}}_{k,ij})_{i,j\in B}$ is given by (cf. (6.9))

$$i\tilde{\mathcal{F}}_{k,ij} = \begin{cases} \frac{\tilde{\mathcal{R}}_{k,ij}}{k\cdot\omega+\Omega_i-\Omega_j}, & \text{if } \tilde{\mathcal{R}}_{k,ij} \neq 0 \text{ and } |k|+|i-j| \neq 0, \\ 0, & \text{otherwise}. \end{cases} \tag{6.14}$$

Note that by the symmetry conditions (6.5), $\tilde{\mathcal{R}}_{k,ij} \neq 0$ implies that $j = i + k\cdot\nu_A$. Hence for any $k \in \mathbb{Z}^A$, in each row and in each column of the infinite matrix $(\tilde{\mathcal{R}}_{k,ij})_{i,j\in B}$ – and thus also of the infinite matrix $(\tilde{\mathcal{F}}_{k,ij})_{i,j\in B}$ – there is at most one nonzero entry. Therefore the operator norm of $\tilde{\mathcal{R}}_k = (\tilde{\mathcal{R}}_{k,ij})_{i,j\in B}$ can be computed to be

$$\|\tilde{\mathcal{R}}_k\| = \sup_{i,j\in B} |\tilde{\mathcal{R}}_{k,ij}|.$$

By (6.5)–(6.6), $\tilde{\mathcal{R}}_{k,ij} \neq 0$ with $|k|+|i-j| \neq 0$ implies that $|k\cdot\omega+\Omega_i-\Omega_j| \geqslant \alpha A_k^{-1}$. Hence $\|\tilde{\mathcal{F}}_k\| \leqslant \frac{A_k}{\alpha}\|\tilde{\mathcal{R}}_k\|$ uniformly on Π. Summing up over k leads to

$$\|\tilde{\mathcal{F}}\|^{\sup}_{D(s-\sigma)} \leqslant \frac{B_\sigma}{\alpha}\|\tilde{\mathcal{R}}\|^{\sup}_{D(s)}.$$

Going back to the operator norm of linear operators on $\ell^{2,N}$ one gets, in view of (6.13),

$$\begin{aligned} \frac{1}{r}\left\|\partial_z F^{11}\right\|^{\sup}_{D(s-\sigma,r)} &= \sup_{\|w\|_N<r}\frac{1}{r}\|\ddot{\mathcal{F}}w\|^{\sup}_{D(s-\sigma)} \leqslant \|\ddot{\mathcal{F}}\|^{\sup}_{D(s-\sigma)} \leqslant \frac{B_\sigma}{\alpha}\|\ddot{\mathcal{R}}\|^{\sup}_{D(s)} \\ &\leqslant \frac{B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}. \end{aligned} \tag{6.15}$$

Similarly one has

$$\frac{1}{r}\left\|\partial_w F^{11}\right\|^{\sup}_{D(s-\sigma,r)} \leqslant \frac{B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}.$$

To estimate $\frac{1}{r^2}\|\partial_x F^{11}\|^{\sup}_{D(s-\sigma,r)}$ note that $\partial_x F^{11} = i\sum_k k(\sum_{i,j}\ddot{\mathcal{F}}_{ij}z_i w_j)e^{ik\cdot x}$. As $|\sum_{i,j}\ddot{\mathcal{F}}_{ij}z_i w_j| \leqslant \|\ddot{\mathcal{F}}\|\|z\|_N\|w\|_N$, it follows from (6.15) and the definition of B_σ

$$\frac{1}{r^2}\left\|\partial_x F^{11}\right\|^{\sup}_{D(s-\sigma,r)} \leqslant \frac{B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}.$$

Altogether we thus have proved that

$$\|X_{F^{11}}\|^{\sup}_{r,D(s-\sigma,r)} \leqslant \frac{3B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}.$$

Next we want to estimate $\|X_{F^{11}}\|^{lip}_{r,D(s-\sigma,r)}$. The Lipschitz estimate of $\ddot{F}$ is obtained in a similar fashion as the one of $\dot{F}$. Indeed, let

$$j \equiv j(i,k) := i + k\cdot\nu_A \quad \text{and} \quad \alpha_{k,i} := k\cdot\omega + \Omega_i - \Omega_j.$$

Then

$$i\Delta\tilde{\mathcal{F}}_{k,ij} = \alpha_{k,i}^{-1}\Delta\tilde{\mathcal{R}}_{k,ij} + \tilde{\mathcal{R}}_{k,ij}\Delta\alpha_{k,i}^{-1}.$$

The small divisors assumption (6.6) then implies that

$$|\alpha_{k,i}| \geqslant \alpha A_k^{-1}\big(\langle i\rangle^\delta + \langle j\rangle^\delta\big)^{\frac{1}{2}}.$$

Using that

$$\frac{|\Delta(\Omega_i - \Omega_j)|}{\langle i\rangle^\delta + \langle j\rangle^\delta} \leqslant \frac{|\Delta\Omega_i|}{\langle i\rangle^\delta} + \frac{|\Delta\Omega_j|}{\langle j\rangle^\delta}$$

one gets as above

$$\big|\Delta\alpha_{k,i}^{-1}\big| \leqslant \frac{A_k^2}{\alpha^2}\left(|k||\Delta\omega| + \frac{|\Delta\Omega_i|}{\langle i\rangle^\delta} + \frac{|\Delta\Omega_j|}{\langle j\rangle^\delta}\right) \leqslant \frac{A_k^2}{\alpha^2}\big(|k||\Delta\omega| + 2|\Delta\Omega|_{\ell^\infty,-\delta}\big)$$

and therefore, uniformly on Π,

$$\|\Delta\tilde{\mathcal{F}}_k\| \leqslant \frac{A_k}{\alpha}\|\Delta\tilde{\mathcal{R}}_k\| + \frac{A_k^2}{\alpha^2}\big(|k||\Delta\omega| + 2|\Delta\Omega|_{\ell^\infty,-\delta}\big)\|\tilde{\mathcal{R}}_k\|.$$

Summing up over k this leads to

$$\|\Delta\tilde{\mathcal{F}}\|_{D(s-\sigma)}^{\sup} \leqslant \frac{B_\sigma}{\alpha}\|\Delta\tilde{\mathcal{R}}\|_{D(s)}^{\sup} + \frac{B_\sigma}{\alpha^2}\big(|\Delta\omega| + 2|\Delta\Omega|_{\ell^\infty,-\delta}\big)\|\tilde{\mathcal{R}}\|_{D(s)}^{\sup}.$$

Going back to the operator norm of linear operators on $\ell^{2,N}$ one gets

$$\|\Delta\ddot{\mathcal{F}}\|_{D(s-\sigma)}^{\sup} \leqslant \frac{B_\sigma}{\alpha}\|\Delta\ddot{\mathcal{R}}\|_{D(s)}^{\sup} + \frac{B_\sigma}{\alpha^2}\big(|\Delta\omega| + 2|\Delta\Omega|_{\ell^\infty,-\delta}\big)\|\ddot{\mathcal{R}}\|_{D(s)}^{\sup}. \tag{6.16}$$

Dividing this inequality by $|\xi - \zeta|$ and taking the supremum over $\xi \neq \zeta$ in Π yields

$$\|\ddot{\mathcal{F}}\|_{D(s-\sigma)}^{lip} \leqslant \frac{2B_\sigma}{\alpha}\left(\|\ddot{\mathcal{R}}\|_{D(s)}^{lip} + \frac{M}{\alpha}\|\ddot{\mathcal{R}}\|_{D(s)}^{\sup}\right).$$

Finally arguing as in (6.15) one concludes that

$$\frac{1}{r}\big\|\partial_z F^{11}\big\|_{D(s-\sigma,r)}^{lip} = \sup_{\|w\|_N<r}\frac{1}{r}\|\ddot{\mathcal{F}}w\|_{D(s-\sigma,r)}^{lip} \leqslant \frac{2B_\sigma}{\alpha}\left(\|X_R\|_{r,D(s,r)}^{lip} + \frac{M}{\alpha}\|X_R\|_{r,D(s,r)}^{\sup}\right). \tag{6.17}$$

Similarly one has

$$\frac{1}{r}\big\|\partial_w F^{11}\big\|_{D(s-\sigma,r)}^{lip} \leqslant \frac{2B_\sigma}{\alpha}\left(\|X_R\|_{r,D(s,r)}^{lip} + \frac{M}{\alpha}\|X_R\|_{r,D(s,r)}^{\sup}\right).$$

To estimate $\frac{1}{r^2}\|\partial_x F^{11}\|_{D(s-\sigma,r)}^{lip}$ note that

$$-i\partial_x\Delta F^{11} = \sum_k k\left(\sum_{i,j}\Delta\ddot{\mathcal{F}}_{ij}z_i w_j\right)e^{ik\cdot x}.$$

As $|\sum_{i,j}\Delta\ddot{\mathcal{F}}_{ij}z_i w_j| \leqslant \|\Delta\ddot{\mathcal{F}}\|\|z\|_N\|w\|_N$ it follows from (6.16) and the definition of B_σ

$$\frac{1}{r^2}\left\|\partial_x \Delta F^{11}\right\|^{\sup}_{D(s-\sigma,r)} \leqslant \sum_k |k|\|\Delta\ddot{\mathcal{F}}_k\|e^{|k|(s-\sigma)}$$
$$\leqslant \frac{B_\sigma}{\alpha}\|\Delta\ddot{\mathcal{R}}\|^{\sup}_{D(s)} + \frac{B_\sigma}{\alpha^2}\left(|\Delta\omega| + 2|\Delta\Omega|_{\ell^{\infty,-\delta}}\right)\|\ddot{\mathcal{R}}\|^{\sup}_{D(s)}.$$

With the same arguments which lead to (6.17) one then concludes that

$$\frac{1}{r^2}\left\|\partial_x F^{11}\right\|^{lip}_{D(s-\sigma,r)} \leqslant \frac{2B_\sigma}{\alpha}\left(\|X_R\|^{lip}_{r,D(s,r)} + \frac{M}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}\right).$$

Altogether we thus have proved that

$$\|X_{F^{11}}\|^{lip}_{r,D(s-\sigma,r)} \leqslant \frac{6B_\sigma}{\alpha}\left(\|X_R\|^{lip}_{r,D(s,r)} + \frac{M}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}\right).$$

All the other components F^{ab} admit the same type of estimates. More precisely, $\|X_{F^0}\|^{\sup}_{r,D(s-\sigma,r)}$, $\|X_{F^{20}}\|^{\sup}_{r,D(s-\sigma,r)}$, $\|X_{F^{11}}\|^{\sup}_{r,D(s-\sigma,r)}$ and $\|X_{F^{02}}\|^{\sup}_{r,D(s-\sigma,r)}$ are each bounded by $\frac{3B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}$ whereas $\|X_{F^{10}}\|^{\sup}_{r,D(s-\sigma,r)}$ and $\|X_{F^{01}}\|^{\sup}_{r,D(s-\sigma,r)}$ are bounded by $\frac{2B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}$. Altogether, one gets

$$\|X_F\|^{\sup}_{r,D(s-\sigma,r)} \leqslant \frac{16B_\sigma}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}.$$

Similarly, by the estimates above, one obtains

$$\|X_F\|^{lip}_{r,D(s-\sigma,r)} \leqslant \frac{25B_\sigma}{\alpha}\left(\|X_R\|^{lip}_{r,D(s,r)} + \frac{M}{\alpha}\|X_R\|^{\sup}_{r,D(s,r)}\right)$$

as claimed. □

Following [18], the estimates may be compactly written as follows. For $\lambda \geqslant 0$, define for a parameter dependent vector field $Y : D(s,r) \times \Pi \to \mathcal{P}^N_{\mathbb{C}}$ with components (Y_x, Y_y, Y_u, Y_v) and parameter $\xi \in \Pi$,

$$\|Y\|^{\lambda}_{r,D(s,r)} := \|Y\|^{\sup}_{r,D(s,r)} + \lambda\|Y\|^{lip}_{r,D(s,r)}.$$

Furthermore, let $\|Y\|^{*}_{r,D(s,r)}$ stand for either $\|Y\|^{\sup}_{r,D(s,r)}$ or $\|Y\|^{lip}_{r,D(s,r)}$.

Corollary 6.1. *Under the assumptions of Lemma* 6.1, *one has for* $0 < \sigma \leqslant s$ *and* $0 \leqslant \lambda \leqslant \frac{\alpha}{M}$

$$\|X_{\hat{H}}\|^{*}_{r,D(s,r)} \leqslant \|X_R\|^{*}_{r,D(s,r)},$$

and

$$\|X_F\|^{\lambda}_{r,D(s-\sigma,r)} \leqslant \frac{41B_\sigma}{\alpha}\|X_R\|^{\lambda}_{r,D(s,r)}.$$

Moreover, if $A_k = \langle k\rangle^\tau$, *then*

$$B_\sigma \leqslant b \cdot \sigma^{-(2\tau+|A|+2)} \tag{6.18}$$

with some constant $b \geqslant 1$ *depending only on* A *and* τ.

6.2. KAM step

At the νth step of the iteration scheme we are given a Hamiltonian $H_\nu + P_\nu$ where H_ν is in normal form and P_ν is a small perturbation satisfying $\{P_\nu, S\} = 0$. More precisely, H_ν and P_ν are assumed to be regular on $D(s_\nu, r_\nu) \times \Pi_\nu$, with $0 < s_\nu \leqslant s_{\nu-1}$ and $0 < r_\nu \leqslant r_{\nu-1}$ in the sense defined at the beginning of Section 6.1. Furthermore, $\Pi_\nu \subseteq \Pi$ is a compact subset and H_ν is of the form

$$H_\nu = \omega^\nu \cdot y + \Omega^\nu w \cdot z,$$

with $\omega^\nu = (\omega_j^\nu)_{j\in A}$ and $\Omega^\nu w = (\Omega_j^\nu w_j)_{j\in B}$ satisfying $|\omega^\nu|^{lip} + |\Omega^\nu|^{lip}_{\ell^\infty,-\delta} \leqslant M_\nu$ and the small divisors condition on Π_ν

$$\bigl|k \cdot \omega^\nu + e \cdot \Omega^\nu\bigr| \geqslant \alpha_\nu A_k^{-1} \cdot 1 \vee |e|_\delta^{\frac{1}{2}} \tag{6.19}$$

for any $(k, e) \in \mathcal{Z}$ where $A_k = \langle k\rangle^\tau$. The perturbation P_ν satisfies in addition the symmetry condition $\{P_\nu, S\} = 0$. In this subsection we now drop the index ν and write '+' for '$\nu + 1$' to simplify notation. Thus $P = P_\nu$ and $P_+ = P_{\nu+1}$ and so on. In the following, C stands for a constant which depends only on A and τ – actually the dependence on τ only enters through the constant b in (6.18). Furthermore we assume that the perturbation is so small that we can choose $0 < \eta < \frac{1}{16}$ and $0 < \sigma < \frac{s}{2}$ with $\sigma \leqslant 1$, such that

$$\|X_P\|^{\sup}_{r,D(s,r)} + \frac{\alpha}{M}\|X_P\|^{lip}_{r,D(s,r)} \leqslant \frac{\alpha\sigma^\kappa\eta^2}{c_0} \tag{6.20}$$

where $\kappa = 2\tau + |A| + 3$ and $c_0 \geqslant 1$ is a sufficiently large constant depending only on A and τ, which will be specified later and will enter the smallness condition of the perturbation P in Theorem 4.1, encoded in γ.

Approximation of P

We now approximate P by its Taylor polynomial R of degree two in y, w, and z of the form (6.2). This leads to corresponding approximations of the partial derivatives $\partial_x P$, $\partial_y P$, $\partial_w P$, and $\partial_z P$ which constitute the Hamiltonian vector field X_P. As in the proof of Lemma 6.1, we represent R in the form $\sum_{0\leqslant i+j\leqslant 2} R^{ij}$. The components of the Hamiltonian vector fields $X_{R^{ij}}$ can then be expressed in terms of the derivatives up to order 2 of components of X_P evaluated at $y = 0$, $w = 0$, $z = 0$. Since $P(\cdot;\xi)$ is analytic, Cauchy's estimate then leads to the estimate

$$\|X_R\|^*_{r,D(s,r)} \leqslant C\|X_P\|^*_{r,D(s,r)}, \tag{6.21}$$

where we recall that C stands for a constant which depends only on A and τ. Next we need to estimate how accurate X_R approximates X_P. We claim that

$$\|X_P - X_R\|^*_{\eta r,D(s,4\eta r)} \leqslant C\eta\|X_P\|^*_{r,D(s,r)}. \tag{6.22}$$

To prove this inequality note that

$$X_P - X_R = \bigl(\partial_y(P - R), -\partial_x(P - R), -i\partial_z(P - R), i\partial_w(P - R)\bigr).$$

Let us begin by estimating $\partial_y P - \partial_y R$. As $\partial_y R = \partial_y P|_{y=0,w=0,z=0}$ one has

$$\partial_y P-\partial_y R=\int_0^1 (y\cdot\partial_y)(\partial_y P)(x,ty,tw,tz)\,dt$$
$$+\int_0^1 (w\cdot\partial_w)(\partial_y P)(x,ty,tw,tz)\,dt+\int_0^1 (z\cdot\partial_z)(\partial_y P)(x,ty,tw,tz)\,dt.$$

Here $y\cdot\partial_y=\sum_{j\in A} y_j\partial_{y_j}$ and $w\cdot\partial_w$, $z\cdot\partial_z$ are defined similarly. By Cauchy's estimate one has

$$\big\|(y\cdot\partial_y)(\partial_y P)\big\|^{\sup}_{D(s,4\eta r)}\leqslant C\frac{(4\eta r)^2}{((1-4\eta)r)^2}\|\partial_y P\|^{\sup}_{D(s,r)}\leqslant C\eta^2\|\partial_y P\|^{\sup}_{D(s,r)}.$$

Similarly one gets

$$\big\|(w\cdot\partial_w)(\partial_y P)\big\|^{\sup}_{D(s,4\eta r)},\qquad \big\|(z\cdot\partial_z)(\partial_y P)\big\|^{\sup}_{D(s,4\eta r)}\leqslant C\eta\|\partial_y P\|^{\sup}_{D(s,r)}.$$

As $\|\partial_y P\|^{\sup}_{D(s,r)}\leqslant\|X_P\|^{\sup}_{r,D(s,r)}$, it then follows that

$$\|\partial_y P-\partial_y R\|^{\sup}_{D(s,4\eta r)}\leqslant C\eta\|X_P\|^{\sup}_{r,D(s,r)}.$$

In a similar way one shows that

$$\|\partial_y P-\partial_y R\|^{lip}_{D(s,4\eta r)}\leqslant C\eta\|X_P\|^{lip}_{r,D(s,r)}.$$

Next let us estimate the component $\partial_w P-\partial_w R$. Note that

$$\partial_w P(x,y,w,z)=\partial_w P(x,0,w,z)+\int_0^1 y\cdot\partial_y(\partial_w P)(x,ty,w,z)\,dt.$$

The error term $\int_0^1 y\cdot\partial_y(\partial_w P)(x,ty,w,z)\,dt$ is not part of X_R and Cauchy's estimate leads to

$$\frac{1}{\eta r}\left\|\int_0^1 y\cdot\partial_y(\partial_w P)(x,ty,w,z)\,dt\right\|^{\sup}_{D(s,4\eta r)}\leqslant\frac{1}{\eta r}\frac{(4\eta r)^2}{((1-4\eta)r)^2}\|\partial_w P\|^{\sup}_{D(s,r)}$$
$$\leqslant C\eta\frac{1}{r}\|\partial_w P\|^{\sup}_{D(s,r)}\leqslant C\eta\|X_P\|^{\sup}_{r,D(s,r)}.$$

Now expand $\partial_w P(x,0,w,z)$,

$$\partial_w P(x,0,w,z)-\partial_w P(x,0,0,0)=\int_0^1\frac{d}{dt}\partial_w P(x,0,tw,tz)\,dt=:I.$$

As $\frac{d}{dt}\partial_w P(x,0,tw,tz)=w\cdot\partial_w(\partial_w P)(x,0,tw,tz)+z\cdot\partial_z(\partial_w P)(x,0,tw,tz)$ we get

$$I=w\cdot\partial_w(\partial_w P)(x,0,0,0)+z\cdot\partial_z(\partial_w P)(x,0,0,0)+II+III+IV$$

where

$$II = \int_0^1 (1-s)(w\cdot\partial_w)\cdot(\tilde{w}\cdot\partial_w)(\partial_w P)(x,0,sw,sz)\,ds|_{\tilde{w}=w},$$

$$III = \int_0^1 (1-s)(z\cdot\partial_z)\cdot(\tilde{z}\cdot\partial_z)(\partial_w P)(x,0,sw,sz)\,ds|_{\tilde{z}=z},$$

$$IV = 2\int_0^1 (1-s)(z\cdot\partial_z)(w\cdot\partial_w)(\partial_w P)(x,0,sw,sz)\,ds.$$

The error terms II, III, IV are not part of X_R and by Cauchy's estimate for second derivatives one gets

$$\frac{1}{\eta r}\|II\|^{\sup}_{D(s,4\eta r)} \leqslant \frac{1}{\eta r}\frac{(\eta r)^2}{((1-4\eta)r)^2}\|\partial_w P\|^{\sup}_{D(s,r)} \leqslant C\eta\|X_P\|^{\sup}_{r,D(s,r)}.$$

For III and IV similar estimates are obtained. Altogether we then get

$$\frac{1}{\eta r}\|\partial_w P - \partial_w R\|^{\sup}_{D(s,4\eta r)} \leqslant C\eta\|X_P\|^{\sup}_{r,D(s,r)}.$$

In a similar way one shows that

$$\frac{1}{\eta r}\|\partial_w P - \partial_w R\|^{lip}_{D(s,4\eta r)} \leqslant C\eta\|X_P\|^{lip}_{r,D(s,r)}.$$

By the same arguments one also has

$$\frac{1}{\eta r}\|\partial_z P - \partial_z R\|^{*}_{D(s,4\eta r)} \leqslant C\eta\|X_P\|^{*}_{r,D(s,r)}.$$

Finally, we need to consider $\partial_x P - \partial_x R$. First expand $\partial_x P$ with respect to y,

$$\partial_x P(x,y,w,z) = \partial_x P(x,0,w,z) + (y\cdot\partial_y)(\partial_x P)(x,0,w,z) + V$$

where

$$V := \int_0^1 (1-t)(y\cdot\partial_y)(\tilde{y}\cdot\partial_y)(\partial_x P)(x,sy,w,z)\,ds|_{\tilde{y}=y}$$

is not part of $\partial_x R$. By Cauchy's estimate one gets

$$\frac{1}{(\eta r)^2}\|V\|^{\sup}_{D(s,4\eta r)} \leqslant \frac{C}{(\eta r)^2}\frac{(\eta r)^4}{((1-4\eta)r)^4}\|\partial_x P\|^{\sup}_{D(s,r)} \leqslant C\eta^2\frac{1}{r^2}\|\partial_x P\|^{\sup}_{D(s,r)} \leqslant C\eta^2\|X_P\|^{\sup}_{r,D(s,r)}.$$

As R is an affine function of y it follows that the term VI in the expansion

$$(y\cdot\partial_y)(\partial_x P)(x,0,w,z) = (y\cdot\partial_y)(\partial_x P)(x,0,0,0) + VI$$

is not part of $\partial_x R$ where

$$VI := \int_0^1 (w \cdot \partial_w)(y \cdot \partial_y)(\partial_x P)(x, 0, tw, tz)\, dt + \int_0^1 (z \cdot \partial_z)(y \cdot \partial_y)(\partial_x P)(x, 0, tw, tz)\, dt.$$

Arguing as above one has

$$\frac{1}{(\eta r)^2} \|VI\|^{\sup}_{D(s,4\eta r)} \leqslant \frac{C}{(\eta r)^2} \frac{(\eta r)^3}{((1-4\eta)r)^3} \|\partial_x P\|^{\sup}_{D(s,r)} \leqslant C\eta \cdot \|X_P\|^{\sup}_{r,D(s,r)}.$$

The remaining term $\partial_x P(x, 0, w, z)$ has to be expanded in w and z up to order 2. The remainder term VII can then be written in terms of integrals and Cauchy's estimate can be applied to show that

$$\frac{1}{(\eta r)^2} \|VII\|^{\sup}_{D(s,4\eta r)} \leqslant C\eta \|X_P\|^{\sup}_{r,D(s,r)}.$$

Altogether we thus have proved that

$$\|X_P - X_R\|^{\sup}_{\eta r,D(s,4\eta r)} \leqslant C\eta \|X_P\|^{\sup}_{r,D(s,r)}.$$

In a similar way one shows that

$$\|X_P - X_R\|^{lip}_{\eta r,D(s,4\eta r)} \leqslant C\eta \|X_P\|^{lip}_{r,D(s,r)}$$

and (6.22) is established.

Solution of linearized equation

Since the small divisors assumption (6.19) are supposed to hold, we can solve the linear system

$$\{F, H\} + \hat{H} = R, \qquad \{F, S\} = 0$$

with the help of Lemma 6.1. By Corollary 6.1 and the estimates (6.21) we obtain

$$\|X_{\hat{H}}\|^*_{r,D(s,r)} \leqslant C \|X_P\|^*_{r,D(s,r)} \tag{6.23}$$

and, for any $0 \leqslant \lambda \leqslant \frac{\alpha}{M}$,

$$\|X_F\|^{\lambda}_{r,D(s-\sigma,r)} \leqslant C\alpha^{-1}\sigma^{1-\kappa} \|X_P\|^{\lambda}_{r,D(s,r)} \tag{6.24}$$

where we recall that $\kappa = 2\tau + |A| + 3$. By the construction of F and the estimates of X_F of Lemma 6.1 it follows that X_F is a real analytic map $X_F : D(s-\sigma, r) \to \mathcal{P}^N_{\mathbb{C}}$ where $\mathcal{P}^N_{\mathbb{C}} = (\mathcal{P}^N_{\mathbb{C}}, \|\cdot\|_{r,N})$. At each point $\mathfrak{x} = (x, y, w, z) \in D(s-\sigma, r)$, the differential dX_F defines a bounded linear operator on $\mathcal{P}^N_{\mathbb{C}}$. Note that the $\|\cdot\|_{r,N}$-distance in $\mathcal{P}^N_{\mathbb{C}}$ between $D(s-2\sigma, \frac{r}{2})$ and the boundary of $D(s-\sigma, r)$ can be estimated from below by $\sigma \wedge \frac{1}{2} \geqslant \frac{\sigma}{2}$. Hence by Cauchy's estimate

$$\|dX_F\|^{\sup}_{r,D(s-2\sigma,\frac{r}{2})} \leqslant C\sigma^{-1} \|X_F\|^{\sup}_{r,D(s-\sigma,r)} \tag{6.25}$$

where for any $\mathfrak{x} \in D(s-2\sigma, \frac{r}{2})$, $\|d_{\mathfrak{x}} X_F\|$ denotes the operator norm on $\mathcal{P}^N_{\mathbb{C}}$,

$$\|d_{\mathfrak{x}} X_F\| = \sup_{\|Y\|_{r,N} \leqslant 1} \|d_{\mathfrak{x}} X_F \cdot Y\|_{r,N}.$$

Similarly, one sees that

$$\|dX_F\|^{lip}_{r,D(s-2\sigma,\frac{r}{2})} \leqslant C\sigma^{-1}\|X_F\|^{lip}_{r,D(s-\sigma,r)}. \tag{6.26}$$

Canonical transformation

The preceding estimates together with (6.20) and (6.24) imply that for any $0 \leqslant \lambda \leqslant \frac{\alpha}{M}$

$$\frac{1}{\sigma}\|X_F\|^{\lambda}_{r,D(s-\sigma,r)}, \|dX_F\|^{\lambda}_{r,D(s-2\sigma,\frac{r}{2})} \leqslant Cc_0^{-1}\eta^2. \tag{6.27}$$

Note that the $\|\cdot\|_{r,N}$-distance of $D(s-3\sigma, \frac{r}{4})$ to the boundary of $D(s-2\sigma, \frac{r}{2})$ is at least $\sigma \wedge \frac{1}{2^3} \geqslant \frac{\sigma}{8}$. Now choose c_0 in (6.20) sufficiently large to insure that for any $|t| \leqslant 1$ the flow X^t_F exists on $D(s-3\sigma, \frac{r}{4})$ and maps $D(s-3\sigma, \frac{r}{4})$ into $D(s-2\sigma, \frac{r}{2})$. Similarly, the flow X^t_F maps $D(s-4\sigma, \frac{r}{8})$ into $D(s-3\sigma, \frac{r}{4})$. By [14, Lemma A.4], together with the estimate (6.27) above we have

$$\left\|X^t_F - id\right\|^*_{r,D(s-3\sigma,\frac{r}{4})} \leqslant C\|X_F\|^*_{r,D(s-\sigma,r)}. \tag{6.28}$$

Since the $\|\cdot\|_{r,N}$ distance of $D(s-4\sigma, \frac{r}{8})$ to the boundary of $D(s-3\sigma, \frac{r}{4})$ is at least $\sigma \wedge \frac{1}{2^5} \geqslant \sigma/32$, it then follows from Cauchy's estimate that

$$\left\|dX^t_F - Id\right\|^*_{r,D(s-4\sigma,\frac{r}{8})} \leqslant C\sigma^{-1}\|X_F\|^*_{r,D(s-\sigma,r)}. \tag{6.29}$$

In particular, we notice that for any $-1 \leqslant t \leqslant 1$, $X^t_F : D(s-3\sigma, \frac{r}{4}) \times \Pi \to D(s-2\sigma, \frac{r}{2})$ is regular and for any $\xi \in \Pi$, $X^t_F(\cdot;\xi)$ defines a canonical coordinate transformation.

New Hamiltonian

Taking the pull back of $H + P$ by the canonical transformation $\Phi = X^t_F|_{t=1}$ one obtains the Hamiltonian $H_+ + P_+$, defined on $D(s-3\sigma, \frac{r}{4})$, where $H_+ = H + \hat{H}$ and, by (4.9)–(4.10)

$$P_+ = (P - R) \circ X^1_F + \int_0^1 \left\{(1-t)\hat{H} + tR, F\right\} \circ X^t_F \, dt.$$

We have already verified at the end of Section 4 that $S \circ X^1_F = S$ and $\{P_+, S\} = 0$. We now want to estimate the $\|\cdot\|_{r,N}$-norm of the vector field X_{P_+} in terms of the size of X_P. First note that X_{P_+} is given by

$$X_{P_+} = \left(X^1_F\right)^*(X_P - X_R) + \int_0^1 \left(X^t_F\right)^*[X_{(1-t)\hat{H}+tR}, X_F]\, dt.$$

It is shown in [14, pp. 130–132], that for any $0 \leqslant t \leqslant 1$, $0 < \eta < \frac{1}{16}$, $0 \leqslant \lambda \leqslant \frac{\alpha}{M}$ and any vector field $Y : D(s-2\sigma, 4\eta r) \to \mathcal{P}^N_{\mathbb{C}}$,

$$\left\| \left(X_F^t\right)^* Y \right\|_{\eta r, D(s-5\sigma, \eta r)}^{\lambda} \leqslant C \|Y\|_{\eta r, D(s-2\sigma, 4\eta r)}^{\lambda}. \tag{6.30}$$

By (6.22), one has $\|X_P - X_R\|_{\eta r, D(s, 4\eta r)}^{*} \leqslant C\eta \|X_P\|_{r, D(s,r)}^{*}$ and hence in view of (6.30),

$$\left\| \left(X_F^t\right)^* (X_P - X_R) \right\|_{\eta r, D(s-5\sigma, \eta r)}^{\lambda} \leqslant C\eta \|X_P\|_{r, D(s,r)}^{\lambda}. \tag{6.31}$$

It remains to consider the commutator $[X_G, X_F]$ where $G \equiv G(t) = (1-t)\hat{H} + tR$. Note that $[X_G, X_F] = dX_F \cdot X_G - dX_G \cdot X_F$. Hence at each point $\mathfrak{x}$ in $D(s-2\sigma, \frac{r}{2})$, $\|d_{\mathfrak{x}} X_F\|$ denotes the operator norm of $d_{\mathfrak{x}} X_F : \mathcal{P}_{\mathbb{C}}^N \to \mathcal{P}_{\mathbb{C}}^N$ with respect to the norm $\|\cdot\|_{r,N}$. By (6.24)–(6.26),

$$\frac{1}{\sigma} \|X_F\|_{r, D(s-\sigma, r)}^{\lambda}, \ \|dX_F\|_{r, D(s-2\sigma, \frac{r}{2})}^{\lambda} \leqslant C\alpha^{-1}\sigma^{-\kappa} \|X_P\|_{r, D(s,r)}^{\lambda}$$

whereas by (6.21), $\|X_G\|_{r, D(s,r)}^{*} \leqslant C \|X_P\|_{r, D(s,r)}^{*}$. By Cauchy's estimate one then also has

$$\|dX_G\|_{r, D(s-\sigma, \frac{r}{2})}^{*} \leqslant C\sigma^{-1} \|X_P\|_{r, D(s,r)}^{*}.$$

Combining the above estimates yields

$$\left\| [X_G, X_F] \right\|_{r, D(s-2\sigma, \frac{r}{2})}^{\sup} \leqslant C\alpha^{-1}\sigma^{-\kappa} \left(\|X_P\|_{r, D(s,r)}^{\sup} \right)^2. \tag{6.32}$$

Furthermore, as

$$\begin{aligned} \left\| [X_G, X_F] \right\|_{r, D(s-2\sigma, \frac{r}{2})}^{lip} \leqslant{}& \|dX_F\|_{r, D(s-2\sigma, \frac{r}{2})}^{lip} \|X_G\|_{r, D(s-2\sigma, \frac{r}{2})}^{\sup} \\ &+ \|dX_F\|_{r, D(s-2\sigma, \frac{r}{2})}^{\sup} \|X_G\|_{r, D(s-2\sigma, \frac{r}{2})}^{lip} + \|dX_G\|_{r, D(s-2\sigma, \frac{r}{2})}^{lip} \|X_F\|_{r, D(s-2\sigma, \frac{r}{2})}^{\sup} \\ &+ \|dX_G\|_{r, D(s-2\sigma, \frac{r}{2})}^{\sup} \|X_F\|_{r, D(s-2\sigma, \frac{r}{2})}^{lip} \end{aligned}$$

one also concludes that

$$\begin{aligned} \left\| [X_G, X_F] \right\|_{r, D(s-2\sigma, \frac{r}{2})}^{lip} \leqslant{}& C\alpha^{-1}\sigma^{-\kappa} \|X_P\|_{r, D(s,r)}^{lip} \cdot \|X_P\|_{r, D(s,r)}^{\sup} \\ &+ C\alpha^{-1}\sigma^{-\kappa} \cdot M\alpha^{-1} \cdot \left(\|X_P\|_{r, D(s,r)}^{\sup} \right)^2. \end{aligned} \tag{6.33}$$

Using that for any vector Y in $\mathcal{P}_{\mathbb{C}}^N$, $\|Y\|_{\eta r, N}^{*} \leqslant \eta^{-2} \|Y\|_{r,N}^{*}$ it then follows from (6.30), (6.32), and (6.33) and the fact that $4\eta r < \frac{r}{2}$

$$\begin{aligned} \left\| \left(X_F^t\right)^* [X_G, X_F] \right\|_{\eta r, D(s-5\sigma, \eta r)}^{\lambda} &\leqslant C\eta^{-2}\alpha^{-1}\sigma^{-\kappa} \|X_P\|_{r, D(s,r)}^{\sup} \|X_P\|_{r, D(s,r)}^{\lambda} \\ &\leqslant C\eta^{-2}\alpha^{-1}\sigma^{-\kappa} \left(\|X_P\|_{r, D(s,r)}^{\lambda} \right)^2 \end{aligned}$$

for any $0 \leqslant \lambda \leqslant \frac{\alpha}{M}$ and any $0 \leqslant t \leqslant 1$. Combined with (6.31) it leads to the following estimate of the new error term X_{P_+},

$$\|X_{P_+}\|_{\eta r, D(s-5\sigma, \eta r)}^{\lambda} \leqslant C\eta^{-2}\alpha^{-1}\sigma^{-\kappa} \left(\|X_P\|_{r, D(s,r)}^{\lambda} \right)^2 + C\eta \|X_P\|_{r, D(s,r)}^{\lambda}. \tag{6.34}$$

New normal form

We already have seen that $H_+ = H + \hat{H}$ where by (6.23), $\|X_{\hat{H}}\|^*_{r,D(s,r)} \leqslant C\|X_P\|^*_{r,D(s,r)}$. Note that $\hat{H}$ is of the form

$$\hat{H}(y, w, z; \xi) = \hat{\omega}(\xi) \cdot y + \hat{\Omega}(\xi) w \cdot z$$

and hence

$$|\hat{\omega}|^* \leqslant C\|X_P\|^*_{r,D(s,r)}. \tag{6.35}$$

Taking into account that $\frac{1}{r}\sup_{\|w\|_N < r} \|\hat{\Omega}(\xi) w\|_N = |\hat{\Omega}(\xi)|_{\ell^\infty}$, it also follows that

$$|\hat{\Omega}|^*_{\ell^\infty} \leqslant C\|X_P\|^*_{r,D(s,r)}. \tag{6.36}$$

In order to bound the small divisors for the new frequencies $\omega_+ = \omega + \hat{\omega}$ and $\Omega_+ = \Omega + \hat{\Omega}$ for $k \in \mathbb{Z}^A$ with $|k| \leqslant K$ with K to be chosen later in the proof. Observe that for any $(k, e) \in \mathcal{Z}$ with $|k| \leqslant K$, using that $|e| \leqslant 2$

$$|k \cdot \hat{\omega} + e \cdot \hat{\Omega}|^{\sup} \leqslant |k||\hat{\omega}|^{\sup} + |e||\hat{\Omega}|^{\sup}_{\ell^\infty} \leqslant (K+2)C\|X_P\|^{\sup}_{r,D(s,r)} \leqslant \hat{\alpha} A_k^{-1}$$

where $\hat{\alpha}$ satisfies $\hat{\alpha} > C\|X_P\|^{\sup}_{r,D(s,r)} \cdot (K+2)\max_{|k| \leqslant K} |A_k|$. It turns out that one can choose $\hat{\alpha}$ so that $\alpha_+ := \alpha - \hat{\alpha} > 0$ – see Lemma 6.3. With the small divisors assumption (6.19) it then follows that for any $(k, e) \in \mathcal{Z}$ with $|k| \leqslant K$, ω_+, Ω_+ satisfy on Π_ν

$$|k \cdot \omega_+ + e \cdot \Omega_+| \geqslant \alpha_+ A_k^{-1} \cdot 1 \vee |e|_\delta^{\frac{1}{2}}. \tag{6.37}$$

6.3. Iteration and proof of Theorem 4.1

To iterate the KAM step infinitely many times we now choose sequences for all the relevant parameters. Following [18], we choose a geometric sequence for σ, choose the η's to minimize the error estimate (6.34) and change α and M only slightly.

Let c_1 be twice the maximum of all those constants C obtained during the KAM step which depend only on $A \subset \mathbb{Z}$ and τ. For any $\nu \in \mathbb{Z}_{\geqslant 0}$ set

$$\alpha_\nu = \frac{\alpha_0}{2}\left(1 + 2^{-\nu}\right), \qquad M_\nu = M_0\left(2 - 2^{-\nu}\right), \qquad \lambda_\nu = \frac{\alpha_\nu}{M_\nu}$$

with $0 < \alpha_0 < 1$ and $M_0 \geqslant 1$ satisfying $M_0 \geqslant |\omega|^{lip}_{\Pi} + |\Omega|^{lip}_{\Pi,\ell^{\infty,-\delta}}$. Then $(\alpha_\nu)_{\nu \geqslant 0}$ is decreasing and $(M_\nu)_{\nu \geqslant 0}$ increasing. Hence $(\lambda_\nu)_{\nu \geqslant 0}$ is decreasing as well and

$$\lambda_0/4 \leqslant \lambda_\nu \leqslant \lambda_0. \tag{6.38}$$

Furthermore, with $\kappa = 2\tau + |A| + 3$,

$$\sigma_{\nu+1} = \frac{\sigma_\nu}{2}, \qquad \epsilon_{\nu+1} = \frac{c_1 \epsilon_\nu^{\frac{4}{3}}}{(\alpha_\nu \sigma_\nu^\kappa)^{\frac{1}{3}}}, \qquad \eta_\nu^3 = \frac{\epsilon_\nu}{\alpha_\nu \sigma_\nu^\kappa}$$

and

$$s_{\nu+1} = s_\nu - 5\sigma_\nu, \qquad r_{\nu+1} = \eta_\nu r_\nu, \qquad D_\nu = D(s_\nu, r_\nu).$$

The initial values s_0, σ_0 satisfy $0 < s_0 \leqslant 10$ and $\sigma_0 = s_0/40 \leqslant \frac{1}{4}$, implying that $s_0 > s_1 > \cdots > s_0/2$, and γ_0 the smallness condition

$$\gamma_0 \leqslant \left(c_0 + 2^{\kappa+3}c_1\right)^{-3} \tag{6.39}$$

where c_0 appears in (6.20). Finally let $K_\nu = K_0 2^\nu$ and $K_0^{\tau+1} = \frac{1}{c_1\gamma_0}$. The smallness condition of the perturbation is expressed by the inequality

$$\epsilon = \epsilon_0 \leqslant \gamma_0\alpha_0\sigma_0^\kappa. \tag{6.40}$$

Then one has the following bounds for the sequence $(\epsilon_\nu)_{\nu\geqslant 0}$.

Lemma 6.2. *For any $\nu \geqslant 0$,*

(i) $\epsilon_\nu \leqslant \gamma_0\alpha_\nu\sigma_\nu^\kappa 2^{-\nu}$;
(ii) $\epsilon_{\nu+1} \leqslant 2^{-\kappa-3}\epsilon_\nu$ *and* $\sum_0^\infty \epsilon_\nu \leqslant 2\epsilon_0$;
(iii) $\alpha_\nu^{-1}\sigma_\nu^{1-\kappa}\epsilon_\nu \leqslant \alpha_0^{-1}\sigma_0^{1-\kappa}\epsilon_0 2^{-\nu}$.

Proof. (i) The claimed estimate is proved by induction. For $\nu = 0$, the estimate holds by assumption (6.40). To prove the induction step, note that by definition, $\epsilon_{\nu+1} = c_1\epsilon_\nu(\alpha_\nu^{-1}\sigma_\nu^{-\kappa}\epsilon_\nu)^{1/3}$. By the induction hypotheses, $(\alpha_\nu^{-1}\sigma_\nu^{-\kappa}\epsilon_\nu)^{1/3} \leqslant (\gamma_0 2^{-\nu})^{1/3}$ and by the smallness condition of γ_0, one has

$$c_1\left(\alpha_\nu^{-1}\sigma_\nu^{-\kappa}\epsilon_\nu\right)^{1/3} \leqslant c_1\left(\gamma_0 2^{-\nu}\right)^{1/3} \leqslant c_1\frac{1}{c_1 2^{\kappa+3}} \leqslant 2^{-\kappa-3} \tag{6.41}$$

which together with the induction hypothesis implies

$$\epsilon_{\nu+1} \leqslant 2^{-\kappa-3}\cdot\gamma_0\alpha_\nu\sigma_\nu^\kappa 2^{-\nu} = \gamma_0\cdot 2^{-2}\alpha_\nu\cdot\left(2^{-1}\sigma_\nu\right)^\kappa\cdot 2^{-\nu-1} \leqslant \gamma_0\alpha_{\nu+1}\sigma_{\nu+1}^\kappa 2^{-\nu-1}.$$

(ii) By the definition of $\epsilon_{\nu+1}$, $\epsilon_{\nu+1}/\epsilon_\nu = c_1(\alpha_\nu^{-1}\sigma_\nu^{-\kappa}\epsilon_\nu)^{\frac{1}{3}}$. As by (6.41) $\epsilon_{\nu+1}/\epsilon_\nu \leqslant 2^{-\kappa-3}$ item (ii) follows.

(iii) The claimed estimate clearly holds in the case $\nu = 0$. To prove the induction step first note that by (ii), $\epsilon_{\nu+1}/\epsilon_\nu \leqslant 2^{-\kappa-3}$. Hence

$$\frac{\epsilon_{\nu+1}\alpha_{\nu+1}^{-1}\sigma_{\nu+1}^{1-\kappa}}{\epsilon_\nu\alpha_\nu^{-1}\sigma_\nu^{1-\kappa}} = \frac{\epsilon_{\nu+1}}{\epsilon_\nu}\cdot\frac{\alpha_\nu}{\alpha_{\nu+1}}\cdot\left(\frac{\sigma_{\nu+1}}{\sigma_\nu}\right)^{1-\kappa} \leqslant 2^{-3} \tag{6.42}$$

and the claimed estimate for $\epsilon_{\nu+1}$ then follows from the induction hypothesis. □

In [18], a version of the following Iterative Lemma is proved. It can be proved in the same way as in [18] and hence we omit its proof.

Lemma 6.3. *Suppose that $H_\nu + P_\nu$ is regular on $D_\nu \times \Pi_\nu$ in the sense defined at the beginning of Section* 6.1 *where $H_\nu = \omega^\nu(\xi)\cdot y + \Omega^\nu(\xi)w\cdot z$ is a regular Hamiltonian on $D_\nu \times \Pi_\nu$ in normal form satisfying $|\omega^\nu|_{\Pi_\nu}^{lip} + |\Omega^\nu|_{\Pi_\nu,\ell^{\infty,-\delta}}^{lip} \leqslant M_\nu$ and*

$$\left|k \cdot \omega^{\nu}(\xi) + e \cdot \Omega^{\nu}(\xi)\right| \geqslant \alpha_{\nu} A_k^{-1} \cdot 1 \vee |e|_{\delta}^{\frac{1}{2}}, \quad \forall \xi \in \Pi_{\nu},\ \forall (k, e) \in \mathcal{Z}, \tag{6.43}$$

and where P_{ν} satisfies $\{P_{\nu}, S\} = 0$ and

$$\|X_{P_{\nu}}\|_{r_{\nu}, D_{\nu}}^{\lambda_{\nu}} \leqslant \epsilon_{\nu}. \tag{6.44}$$

Then there exist a regular map $\Phi_{\nu+1} : D_{\nu+1} \times \Pi_{\nu} \to D_{\nu}$ with $\Phi_{\nu+1}(\cdot, \xi)$, being a real analytic symplectic coordinate transformation on $D_{\nu+1}$ for any $\xi \in \Pi_{\nu}$, and a closed subset $\Pi_{\nu+1}$ of Π_{ν}, $\Pi_{\nu+1} = \Pi_{\nu} \setminus \bigcup_{|k| > K_{\nu},\, (k,e) \in \mathcal{Z}} \mathcal{R}_{ke}^{\nu+1}(\alpha_{\nu+1})$, where

$$\mathcal{R}_{ke}^{\nu+1}(\alpha_{\nu+1}) = \left\{\xi \in \Pi_{\nu} \colon \left|k \cdot \omega^{\nu+1}(\xi) + e \cdot \Omega^{\nu+1}(\xi)\right| < \alpha_{\nu+1} A_k^{-1} \cdot 1 \vee |e|_{\delta}^{\frac{1}{2}}\right\}$$

such that $(H_{\nu} + P_{\nu}) \circ \Phi_{\nu+1} = H_{\nu+1} + P_{\nu+1}$ satisfies the same assumptions as $H_{\nu} + P_{\nu}$, but with $\nu + 1$ in place of ν.

Remark 6.1. We point out that the dependence of the set $\mathcal{R}_{ke}^{\nu+1}(\alpha_{\nu+1})$ on the perturbation P is not indicated in the notation. We will see in Section 6.4 that the measure of this set can be bounded in terms of $\alpha_{\nu+1}$ independently of the perturbation.

By (6.28)–(6.29) together with (6.24) and the assumption (6.44) we obtain the following estimates.

$$\frac{1}{\sigma_{\nu}} \|\Phi_{\nu+1} - id\|_{r_{\nu}, D_{\nu+1}}^{\lambda_{\nu}}, \|d\Phi_{\nu+1} - Id\|_{r_{\nu}, D_{\nu+1}}^{\lambda_{\nu}} \leqslant c_1 \alpha_{\nu}^{-1} \sigma_{\nu}^{-\kappa} \epsilon_{\nu}, \tag{6.45}$$

whereas by (6.35)–(6.36) together with assumption (6.44) one gets

$$\left|\omega^{\nu+1} - \omega^{\nu}\right|_{\Pi_{\nu}}^{\lambda_{\nu}}, \left|\Omega^{\nu+1} - \Omega^{\nu}\right|_{\Pi_{\nu}, \ell^{\infty}}^{\lambda_{\nu}} \leqslant c_1 \epsilon_{\nu}. \tag{6.46}$$

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. Given the assumptions of Theorem 4.1, we want to apply Lemma 6.3 (Iterative Lemma) with $\nu = 0$. Set

$$s_0 := s; \qquad r_0 := r; \qquad \alpha_0 := \alpha; \qquad M_0 := M; \qquad \epsilon_0 := \epsilon$$

and

$$D_0 = D(s, r); \qquad H_0 := H; \qquad P_0 := P$$

with s, r, α, M, ϵ, H, and P given as in Theorem 4.1. As in the Iterative Lemma, choose $\sigma_0 = s_0/40$, $\lambda_0 = \alpha_0/M_0$, and γ_0 and assume that $\epsilon_0 := \epsilon = \|X_{P_0}\|_{r_0, D_0}^{\lambda_0} \leqslant \gamma_0 \alpha_0 \sigma_0^{\kappa}$. Setting $\gamma := \gamma_0 \sigma_0^{\kappa}$ one then gets

$$\epsilon_0 = \|X_P\|_{r, D(s,r)}^{\lambda_0} \leqslant \gamma_0 \alpha_0 \sigma_0^{\kappa} = \alpha \gamma.$$

In particular, inequality (4.7) of Theorem 4.1 with γ chosen as above is satisfied. By Assumption (B2), $\{P_0, S\} = \{P, S\} = 0$. Furthermore for $\nu = 0$, the small divisors condition (6.43) holds on

$$\Pi_0 := \Pi \setminus \bigcup_{(k,e) \in \mathcal{Z}} \mathcal{R}_{ke}^0(\alpha_0)$$

where

$$\mathcal{R}_{ke}^0(\alpha_0) = \left\{\xi \in \Pi\colon \left|k\cdot\omega(\xi) + e\cdot\Omega(\xi)\right| < \alpha_0 A_k^{-1}\cdot 1\vee|e|_\delta^{\frac{1}{2}}\right\}.$$

Thus the Iterative Lemma applies and we obtain a decreasing sequence of domains $D_\nu \times \Pi_\nu$ and regular maps $\Phi_\nu : D_\nu \times \Pi_{\nu-1} \to D_{\nu-1}$, $\nu \geqslant 1$, with the properties listed in Lemma 6.3. Set $\Phi^\nu := \Phi_1 \circ \cdots \circ \Phi_\nu : D_\nu \times \Pi_{\nu-1} \to D_0$. In particular, $(H_0 + P_0)\circ\Phi^\nu = H_\nu + P_\nu$ and the estimates (6.45)–(6.46) hold. To prove the convergence of the sequence Φ^ν, note that the sequence $(r_\nu)_{\nu\geqslant 0}$ is decreasing. Thus for any $Y \in \mathcal{P}_{\mathbb{C}}^N$ one has $\|Y\|_{r_\nu,N} \leqslant \|Y\|_{r_{\nu+1},N}$. Hence for a linear operator $T : (\mathcal{P}_{\mathbb{C}}^N, \|\cdot\|_{r_{\nu+1},N}) \to (\mathcal{P}_{\mathbb{C}}^N, \|\cdot\|_{r_\nu,N})$

$$\|T\|_{r_{\nu+1},r_\nu} \equiv \sup_{\|\mathfrak{x}\|_{r_{\nu+1},N}\leqslant 1} \|T\mathfrak{x}\|_{r_\nu,N} \leqslant \sup_{\|\mathfrak{x}\|_{r_\nu,N}\leqslant 1} \|T\mathfrak{x}\|_{r_\nu,N} \equiv \|T\|_{r_\nu}.$$

By the mean value theorem one has

$$\left\|\Phi^{\nu+1} - \Phi^\nu\right\|_{r_0,D_{\nu+1}}^{\sup} \leqslant \left\|d\Phi^\nu\right\|_{r_\nu,r_0,D_\nu}^{\sup} \cdot \|\Phi_{\nu+1} - id\|_{r_\nu,D_{\nu+1}}^{\sup}$$

where for any $\mathfrak{x} \in D_{\nu+1}$, $d_{\mathfrak{x}}\Phi^\nu$ is viewed as a linear map $(\mathcal{P}_{\mathbb{C}}^N, \|\cdot\|_{r_\nu,N}) \to (\mathcal{P}_{\mathbb{C}}^N, \|\cdot\|_{r_0,N})$. In view of the chain rule $d\Phi^\nu = d\Phi_1 \circ \cdots \circ d\Phi_\nu$ and thus by the considerations above,

$$\left\|d\Phi^\nu\right\|_{r_\nu,r_0,D_\nu}^{\sup} \leqslant \prod_{\mu=1}^{\nu} \|d\Phi_\mu\|_{r_{\mu-1},D_\mu}^{\sup} \leqslant \prod_{\mu\geqslant 0}\left(1 + 2^{-\mu-2}\right) \leqslant 2$$

where we used that by (6.45), for any $\mu \geqslant 1$,

$$\|d\Phi_\mu\|_{r_\mu,D_\mu}^{\sup} \leqslant 1 + \|d\Phi_\mu - Id\|_{r_{\mu-1},D_\mu}^{\sup} \leqslant 1 + c_1\alpha_{\mu-1}^{-1}\sigma_{\mu-1}^{-\kappa}\epsilon_{\mu-1}$$

and by Lemma 6.2,

$$c_1\alpha_{\mu-1}^{-1}\sigma_{\mu-1}^{-\kappa}\epsilon_{\mu-1} \leqslant c_1\gamma_0 2^{-(\mu-1)} \leqslant c_1\frac{1}{(2^{\kappa+3}c_1)^3}2^{-\mu+1} \leqslant 2^{-\mu-2}.$$

Similarly, one argues for the Lipschitz semi-norm,

$$\left\|\Phi^{\nu+1} - \Phi^\nu\right\|_{r_0,D_{\nu+1}}^{lip} \leqslant \left\|d\Phi^\nu\right\|_{r_\nu,r_0,D_\nu}^{lip}\|\Phi_{\nu+1} - id\|_{r_\nu,D_{\nu+1}}^{\sup} + \left\|d\Phi^\nu\right\|_{r_\nu,r_0,D_\nu}^{\sup}\|\Phi_{\nu+1} - id\|_{r_\nu,D_{\nu+1}}^{lip}$$

and shows as for $\|d\Phi^\nu\|_{r_\nu,r_0,D_\nu}^{\sup}$ that $\|d\Phi^\nu\|_{r_\nu,r_0,D_\nu}^{lip}$ is uniformly bounded. As already pointed out at the beginning of this subsection, one has $\frac{\alpha_0}{4M_0} \leqslant \lambda_\nu \leqslant \lambda_0$ and hence

$$\left\|\Phi^{\nu+1} - \Phi^\nu\right\|_{r_0,D_{\nu+1}}^{\lambda_0} \leqslant C\|\Phi_{\nu+1} - id\|_{r_\nu,D_{\nu+1}}^{\lambda_\nu}.$$

Combined with (6.45) this leads to

$$\left\|\Phi^{\nu+1} - \Phi^\nu\right\|_{r_0,D_{\nu+1}}^{\lambda_0} \leqslant Cc_1\epsilon_\nu\alpha_\nu^{-1}\sigma_\nu^{1-\kappa}. \tag{6.47}$$

Therefore, $(\Phi^\nu)_{\nu\geqslant 1}$ converges uniformly on $\bigcap_{\nu\geqslant 0}(D_\nu \times \Pi_\nu) = D(s/2, 0) \times \Pi_*$ to a Lipschitz continuous family of real analytic torus embeddings $\Psi : \mathbb{T}^A \times \Pi_* \to \mathcal{M}^N$. Here $\Pi_* = \bigcap_{\nu\geqslant 0}\Pi_\nu$ and

$D_* \equiv D(s/2, 0) = D(s/2) \times \{0\} \times \{0\} \times \{0\} \subseteq \mathcal{M}^N$. Recall from the statement of Theorem 4.1 that $\Phi^0 \equiv \Psi_0$ denotes the trivial torus embedding $\mathbb{T}^A \times \Pi_* \to \mathbb{T}_0$. Then by (6.47)

$$\|\Psi - \Psi_0\|_{r_0, D_*}^{\lambda_0} \leqslant \sum_0^\infty \left\|\Phi^{\nu+1} - \Phi^\nu\right\|_{r_0, D_*}^{\lambda_0} \leqslant Cc_1 \sum_0^\infty \alpha_\nu^{-1} \sigma_\nu^{1-\kappa} \epsilon_\nu.$$

By Lemma 6.2, $\sum_0^\infty \alpha_\nu^{-1}\sigma_\nu^{1-\kappa}\epsilon_\nu \leqslant 2\alpha_0^{-1}(\frac{s_0}{40})^{1-\kappa}\epsilon_0$ and hence $\|\Psi - \Psi_0\|_{r_0, D_*}^{\lambda_0} \leqslant c\epsilon/\alpha$ as claimed in Theorem 4.1. Taking into account (6.40) and the estimate (6.46) one sees that the frequencies $\omega^\nu(\xi) \in \mathbb{R}^A$ and $\Omega^\nu(\xi) \in \ell^{\infty,-\delta}$ converge uniformly on Π_* to Lipschitz continuous functions $f : \Pi_* \to \mathbb{R}^A$ respectively $\Omega^* : \Pi_* \to \ell^{\infty,-\delta}$. Furthermore, letting ω^0 denote the frequency vector ω of the unperturbed Hamiltonian H, it follows that $f(\xi) - \omega(\xi) = \sum_{\nu=0}^\infty (\omega^{\nu+1}(\xi) - \omega^\nu(\xi))$ can be estimated as

$$|f - \omega|_{\Pi_*}^{\lambda_0} \leqslant C \sum_0^\infty \left|\omega^{\nu+1} - \omega^\nu\right|_{\Pi_\nu}^{\lambda_\nu} \leqslant C \sum_0^\infty \epsilon_\nu.$$

By Lemma 6.2(ii), $\sum_0^\infty \epsilon_\nu \leqslant 2\epsilon_0$. As $\epsilon_0 = \epsilon$ we thus have shown that $|f - \omega|_{\Pi_*}^{\lambda_0} \leqslant C\epsilon$ as claimed in Theorem 4.1. On the embedded tori, the flow of the perturbed Hamiltonian $H + P$ can be computed as follows. First note that

$$\begin{aligned} \left\|X_{H+P} \circ \Phi^\nu - d\Phi^\nu \cdot X_{H_\nu}\right\|_{r_0, D_\nu \times \Pi_*}^{\sup} &\leqslant \left\|d\Phi^\nu\right\|_{r_\nu, r_0, D_\nu \times \Pi_*} \left\|\left(\Phi^\nu\right)^* X_{H+P} - X_{H_\nu}\right\|_{r_\nu, D_\nu \times \Pi_*}^{\sup} \\ &\leqslant C\|X_{P_\nu}\|_{r_\nu, D_\nu \times \Pi_*}^{\sup}. \end{aligned}$$

In the limit, one thus obtains that $X_{H+P} \circ \Psi = d\Psi \cdot X_{H_*}$ on $D(s/2, 0)$ where

$$H_*(y, w, z; \xi) := f(\xi) \cdot y + \Omega^*(\xi) w \cdot z.$$

It thus follows that for any $x \in \mathbb{T}^A$ and $\xi \in \Pi_*$

$$X_{H+P}^t\left(\Psi(x; \xi)\right) = \Psi\left(x + tf(\xi); \xi\right)$$

as claimed in Theorem 4.1. It remains to show the claims of item (i) of Theorem 4.1, concerning the set $\Pi \setminus \Pi_*$. This will be done in the subsequent Section 6.4. □

6.4. Set of excluded parameters

The aim of this subsection is to prove item (i) of Theorem 4.1. While we again follow the line of arguments used in [13] and [18], there are notable differences due to the near resonances of the frequencies of the unperturbed Hamiltonian which we will point out in the course of the proof.

The KAM iteration leads to a decreasing sequence $(\Pi_\nu)_{\nu \geqslant 0}$ of closed subsets of the parameter space Π. Recall that $\Pi \setminus \Pi_* = \Pi \setminus (\bigcap_{\nu \geqslant 0} \Pi_\nu)$ where

$$\Pi_0 = \bigcup_{\substack{k \in \mathbb{Z}^A \\ (k,e) \in \mathcal{Z}}} \mathcal{R}_{ke}^0(\alpha_0) \quad \text{and} \quad \Pi_\nu = \bigcup_{\substack{|k| > K_\nu \\ (k,e) \in \mathcal{Z}}} \mathcal{R}_{ke}^\nu(\alpha_\nu) \quad \text{for } \nu \geqslant 1.$$

Recall that $\mathcal{Z} \subseteq \mathbb{Z}^A \times \mathbb{Z}^B$ is given by

$$\mathcal{Z} = \left\{(k, e) \in \mathbb{Z}^A \times \mathbb{Z}^B \setminus \left\{(0, 0)\right\}: |e| \leqslant 2;\ k \cdot \nu_A + e \cdot \nu_B = 0\right\},$$

K_ν is given by $K_\nu = (c_1\gamma_0)^{-\tau-1}2^\nu$, and for $(k,e) \in \mathbb{Z}^A \times \mathbb{Z}^B$,

$$\mathcal{R}^\nu_{ke}(\alpha_\nu) = \bigl\{\xi \in \Pi_{\nu-1}\colon \bigl|k\cdot\omega^\nu(\xi) + e\cdot\Omega^\nu(\xi)\bigr| < \alpha_\nu A_k^{-1}\cdot 1 \vee |e|_\delta^{\frac{1}{2}}\bigr\}$$

with $\Pi_{-1} = \Pi$. Here $\omega^\nu = (\omega^\nu_j)_{j\in A}$ and $\Omega^\nu = (\Omega^\nu_j)_{j\in B}$ are the frequencies obtained in the KAM iteration with $\omega^0 = \omega$, $\Omega^0 = \Omega$ denoting the ones of the unperturbed Hamiltonian H. We will prove that by choosing γ_0 sufficiently small – and hence K_0 sufficiently large – $\operatorname{meas}(\Pi \setminus \Pi_*)$ can be estimated as claimed. Note that the set Π_0 is defined in terms of the frequencies of H whereas for $\nu \geqslant 1$ the set Π_ν depends on the perturbation P. To estimate $\operatorname{meas}(\Pi \setminus \Pi_*)$ we need to make some preparations. It is convenient to extend the frequencies ω^ν, Ω^ν, defined on $\Pi_{\nu-1}$, to all of Π. Indeed, each component of $\omega^{\nu+1} - \omega^\nu : \Pi_\nu \to \mathbb{R}^A$ and of $\Omega^{\nu+1} - \Omega^\nu : \Pi_\nu \to \ell^\infty$ has a Lipschitz continuous extension from Π_ν to Π which preserves its minimum, maximum, and Lipschitz semi-norm – see e.g. [13, Lemma M.5]. Since we use the sup norm for $\omega^{\nu+1} - \omega^\nu$ and $\Omega^{\nu+1} - \Omega^\nu$ we obtain in this way extensions $(\omega^{\nu+1} - \omega^\nu)\check{}$ of $\omega^{\nu+1} - \omega^\nu$ and $(\Omega^{\nu+1} - \Omega^\nu)\check{}$ of $\Omega^{\nu+1} - \Omega^\nu$ to all of Π satisfying

$$\bigl|(\omega^{\nu+1} - \omega^\nu)\check{}\bigr|^{\lambda_\nu}_{\Pi} = \bigl|\omega^{\nu+1} - \omega^\nu\bigr|^{\lambda_\nu}_{\Pi_\nu}; \qquad \bigl|(\Omega^{\nu+1} - \Omega^\nu)\check{}\bigr|^{\lambda_\nu}_{\Pi,\ell^\infty} = \bigl|\Omega^{\nu+1} - \Omega^\nu\bigr|^{\lambda_\nu}_{\Pi_\nu,\ell^\infty}. \tag{6.48}$$

Now define $\check\omega^{\nu+1}$, $\check\Omega^{\nu+1}$ by telescoping sums

$$\check\omega^{\nu+1} = \omega + \sum_{\mu=0}^{\nu}\bigl(\omega^{\mu+1} - \omega^\mu\bigr)\check{} \quad\text{and}\quad \check\Omega^{\nu+1} = \Omega + \sum_{\mu=0}^{\nu}\bigl(\Omega^{\mu+1} - \Omega^\mu\bigr)\check{}.$$

Then $\check\omega^\nu : \Pi \to \mathbb{R}^A$ and $\check\Omega^\nu - \Omega : \Pi \to \ell^\infty$ are Lipschitz continuous extensions of ω^ν respectively $\Omega^\nu - \Omega$. Moreover, by Lemma 6.2(ii), $\sum_0^\infty \epsilon_\mu \leqslant 2\epsilon_0$ and hence it follows from (6.46) and (6.48) that for any $\nu \geqslant 0$,

$$\bigl|\check\omega^{\nu+1} - \omega\bigr|^{\lambda_\nu}_{\Pi}, \bigl|\check\Omega^{\nu+1} - \Omega\bigr|^{\lambda_\nu}_{\Pi,\ell^\infty} \leqslant c_1\sum_0^\nu \epsilon_\mu \leqslant 2c_1\epsilon_0.$$

As by (6.38), $\lambda_0/4 \leqslant \lambda_\nu \leqslant \lambda_0$, one then has

$$\bigl|\check\omega^\nu - \omega\bigr|^{\lambda_0/4}_{\Pi}, \bigl|\check\Omega^\nu - \Omega\bigr|^{\lambda_0/4}_{\Pi,\ell^\infty} \leqslant 2c_1\epsilon_0.$$

In particular,

$$\bigl|\check\omega^\nu - \omega\bigr|^{lip}_{\Pi}, \bigl|\check\Omega^\nu - \Omega\bigr|^{lip}_{\Pi,\ell^\infty} \leqslant \frac{8c_1\epsilon_0}{\lambda_0}.$$

Recall that $\lambda_0 = \alpha_0/M_0$, $\epsilon_0 \leqslant \gamma_0\alpha_0\sigma_0^\kappa$, $\sigma_0 \leqslant 1/4$, and $\gamma_0 \leqslant (2^{\kappa+3}c_1)^{-3}$. Thus

$$2c_1\epsilon_0 \leqslant \alpha_0/2 \quad\text{and}\quad \frac{8c_1\epsilon_0}{\lambda_0} \leqslant 8c_1\sigma_0^\kappa M_0\gamma_0.$$

By Assumption (A1), there exists a constant $1 \leqslant L < \infty$ satisfying

$$L \geqslant \bigl|\omega^{-1}\bigr|^{lip}_{\omega(\Pi)}. \tag{6.49}$$

Now require that γ_0 is chosen so small that in addition to (6.39), one has

$$8c_1\sigma_0^\kappa M_0\gamma_0 \leqslant 1/2L. \tag{6.50}$$

Then, for any $\nu \geqslant 0$,

$$\left|\check{\omega}^\nu - \omega\right|_{\Pi}^{\sup}, \left|\check{\Omega}^\nu - \Omega\right|_{\Pi,\ell^\infty}^{\sup} \leqslant \alpha/2 \quad \text{and} \quad \left|\check{\omega}^\nu - \omega\right|_{\Pi}^{lip}, \left|\check{\Omega}^\nu - \Omega\right|_{\Pi,\ell^\infty}^{lip} \leqslant 1/2L$$

and for any $(k,e) \in \mathbb{Z}^A \times \mathbb{Z}^B$, $\mathcal{R}_{ke}^\nu(\alpha_\nu)$ is contained in

$$\check{\mathcal{R}}_{ke}(\alpha_\nu) := \left\{\xi \in \Pi\colon \left|k \cdot \check{\omega}^\nu(\xi) + e \cdot \check{\Omega}^\nu(\xi)\right| < \alpha_\nu A_k^{-1} \cdot 1 \vee |e|_\delta^{\frac{1}{2}}\right\}.$$

In addition, we assume that $M = M_0 \geqslant 1$ bounds the frequencies,

$$|\omega|_{\Pi}^{\sup} + |\Omega - \overline{\Omega}|_{\Pi,\ell^{\infty,-\delta}}^{\sup} \leqslant M \quad \text{and} \quad |\omega|_{\Pi}^{lip} + |\Omega|_{\Pi,\ell^{\infty,-\delta}}^{lip} \leqslant M. \tag{6.51}$$

It turns out that we need not to distinguish between the different values of ν in $\check{\omega}^\nu$ and $\check{\Omega}^\nu$. In the sequel we only use the fact that $\check{\omega}^\nu$ and $\check{\Omega}^\nu$ are Lipschitz maps ω' and Ω', defined on Π, which satisfy the following inequalities

$$\left|\omega' - \omega\right|_{\Pi}^{\sup} + \left|\Omega' - \Omega\right|_{\Pi,\ell^\infty}^{\sup} \leqslant \alpha/2 \quad \text{and} \quad \left|\omega' - \omega\right|_{\Pi}^{lip} + \left|\Omega' - \Omega\right|_{\Pi,\ell^\infty}^{lip} \leqslant 1/2L. \tag{6.52}$$

Henceforth we consider functions ω', Ω' which satisfy these estimates – they may even depend on k and e – and for any $(k,e) \in \mathbb{Z}^A \times \mathbb{Z}^B$ define

$$\mathcal{R}'_{ke}(\alpha) = \left\{\xi \in \Pi\colon \left|k \cdot \omega'(\xi) + e \cdot \Omega'(\xi)\right| < \alpha A_k^{-1} \cdot 1 \vee |e|_\delta^{\frac{1}{2}}\right\}.$$

First we derive the following estimate for $\operatorname{meas}(\mathcal{R}'_{ke}(\alpha))$.

Lemma 6.4. *For any $(k,e) \in \mathbb{Z}^A \times \mathbb{Z}^B \setminus \{(0,0)\}$ with $|k| \geqslant 6LM|e|_\delta$*

$$\operatorname{meas}\bigl(\mathcal{R}'_{ke}(\alpha)\bigr) \leqslant 12L(LM\rho)^{|A|-1}\alpha|k|^{-\frac{1}{2}}A_k^{-1}$$

where $\rho = \operatorname{diam}(\Pi)$ denotes the diameter of Π.

Proof. Taking into account Assumption (A1) we introduce the unperturbed frequencies $\zeta = \omega(\xi)$ as new parameters with domain $\dot{\Pi} = \omega(\Pi)$ and consider the resonance zones $\dot{\mathcal{R}}_{ke} = \omega(\mathcal{R}'_{ke})$ in $\dot{\Pi}$. Writing $\dot{\omega}$ and $\dot{\Omega}$ for the pull back of ω' and Ω' by ω^{-1}, we then have by (6.49), (6.52)

$$|\dot{\omega} - id|_{\dot{\Pi}}^{lip} \leqslant \left|\omega' - \omega\right|_{\Pi}^{lip} \cdot \left|\omega^{-1}\right|_{\dot{\Pi}}^{lip} \leqslant 1/2.$$

In view of (6.49), (6.51), (6.52) and using that $L \geqslant 1$, $M \geqslant 1$, the Lipschitz semi-norm of $\dot{\Omega}$ can be bounded as follows

$$|\dot{\Omega}|_{\Pi,\ell^{\infty,-\delta}}^{lip} \leqslant \left(\left|\Omega' - \Omega\right|_{\Pi,\ell^{\infty,-\delta}}^{lip} + |\Omega|_{\Pi,\ell^{\infty,-\delta}}^{lip}\right)\left|\omega^{-1}\right|_{\dot{\Pi}}^{lip} \leqslant \left(\frac{1}{2L} + M\right)L \leqslant 2LM. \tag{6.53}$$

To estimate $\operatorname{meas}(\dot{\mathcal{R}}_{ke}(\alpha))$, let $g(\zeta) := k \cdot \dot{\omega}(\zeta) + e \cdot \dot{\Omega}(\zeta)$. Choose a vector $v \in \{-1,1\}^A$ such that $k \cdot v = |k|$ and write any vector in $\mathbb{R}^A$ as a linear combination of v and an element w in the orthogonal complement $v^\perp$ of v. Introduce the following affine function of the real variable r, $\zeta = \zeta(r) := rv + w$. For any $t > s$ with $\zeta(t)$, $\zeta(s)$ in $\dot{\Pi}$ one has

$$k \cdot \dot{\omega}(\zeta)\big|_s^t = k \cdot \zeta\big|_s^t + k \cdot \bigl(\dot{\omega}(\zeta) - \zeta\bigr)\big|_s^t \geqslant |k|(t-s) - \frac{1}{2}|k|(t-s) = \frac{1}{2}|k|(t-s). \tag{6.54}$$

Moreover by (6.53) and the assumption $6LM|e|_\delta \leqslant |k|$,

$$\begin{aligned} \left|e \cdot \dot{\Omega}(\zeta)|_s^t\right| &= \left|e \cdot \left(\dot{\Omega}\left(\zeta(t)\right) - \dot{\Omega}\left(\zeta(s)\right)\right)\right| \\ &\leqslant |e|_\delta |\dot{\Omega}|^{lip}_{\dot{\Pi},\ell^{\infty,-\delta}}(t-s) \leqslant 2LM|e|_\delta(t-s) \leqslant \frac{1}{3}|k|(t-s). \end{aligned} \tag{6.55}$$

Altogether we have shown that uniformly for $w \in v^\perp$ with $rv + w|_{r=t,s} \in \dot{\Pi}$,

$$g(rv+w)|_s^t \geqslant \frac{1}{6}|k|(t-s).$$

It follows that for each point $w \in v^\perp$ so that $rv + w \in \dot{\Pi}$ for some $r \in \mathbb{R}$, the set

$$\left\{r \in \mathbb{R}\colon\ rv + w \in \dot{\Pi};\ \left|g(rv+w)\right| < \eta\right\}$$

is contained in an interval I_w of length $\operatorname{meas}(I_w) \leqslant 12\eta|k|^{-1}$. With $\eta = \alpha A_k^{-1} \cdot 1 \vee |e|_\delta^{\frac{1}{2}}$ and Fubini's theorem one then concludes that

$$\operatorname{meas}\left(\dot{\mathcal{R}}_{ke}(\alpha)\right) \leqslant \frac{12\alpha A_k^{-1}}{|k|} \cdot 1 \vee |e|_\delta^{\frac{1}{2}} \cdot (\operatorname{diam} \dot{\Pi})^{|A|-1}.$$

As $6LM|e|_\delta \leqslant |k|$ and $LM \geqslant 1$ one then gets for any $e \in \mathbb{Z}^B$

$$\operatorname{meas}\left(\dot{\mathcal{R}}_{ke}(\alpha)\right) \leqslant 12\alpha|k|^{-\frac{1}{2}} A_k^{-1} \cdot (\operatorname{diam} \dot{\Pi})^{|A|-1}.$$

Going back to the original parameter domain Π by the inverse ω^{-1} of the frequency map and noting that $\operatorname{diam}(\dot{\Pi}) \leqslant |\omega|^{lip}_{\Pi}\rho$ with ρ denoting the diameter $\operatorname{diam}(\Pi)$ of Π it then follows that

$$\begin{aligned} \operatorname{meas}\left(\mathcal{R}'_{ke}(\alpha)\right) &\leqslant \left(\left|\omega^{-1}\right|^{lip}_{\dot{\Pi}}\right)^{|A|} \operatorname{meas}\left(\dot{\mathcal{R}}_{ke}(\alpha)\right) \\ &\leqslant 12L^{|A|}(M\rho)^{|A|-1}\alpha|k|^{-\frac{1}{2}} A_k^{-1} \end{aligned}$$

as claimed. □

It is convenient to introduce $\Lambda := \{e \in \mathbb{Z}^B\colon\ 1 \leqslant |e| \leqslant 2\}$ and

$$\Lambda_r := \left\{e \in \Lambda\colon\ e = (\delta_{aj})_{j\in\mathbb{Z}} - (\delta_{-aj})_{j\in\mathbb{Z}},\ a \in \mathbb{Z} \setminus \{0\}\right\}.$$

By a slight abuse of terminology we refer to Λ_r as the subset of resonant sites of Λ. It turns out that the estimates involving resonant sites have to be dealt with separately. Recall that the unperturbed frequencies satisfy $|\Omega - \overline{\Omega}|^{\sup}_{\Pi,\ell^{\infty,-\delta}} \leqslant M$ where for $j \in B$,

$$\overline{\Omega}_j = |j|^d + a_1|j|^{d_1} + \cdots + a_D|j|^{d_D}$$

with $d \equiv d_0 > d_1 > \cdots > d_D \geqslant 0$, $a_1, \ldots, a_D \in \mathbb{R}$, and $d > 1$, $0 \leqslant \delta < 1 \wedge (d-1)$.

Lemma 6.5. *There exists $E \geqslant 1$ depending only on M and the approximation $\overline{\Omega}$ of the unperturbed frequencies Ω so that for any $e \in \Lambda \setminus \Lambda_r$ with $|e|_* := \max_{j \in B}\{|j|\colon e_j \neq 0\} \geqslant E$*

$$\left|e \cdot \Omega'(\xi)\right| \geqslant \frac{1}{8}|e|_{d-1} \quad \forall \xi \in \Pi.$$

Proof. We only prove the claimed estimate for $e \in \Lambda \setminus \Lambda_r$ with $e \cdot \Omega'(\xi)$ of the form $\Omega_i'(\xi) - \Omega_j'(\xi)$ for some $i, j \in B$ with $j \neq -i$ which is the most subtle case. Write $\Omega_i' = \Omega_i + (\Omega_i' - \Omega_i)$ and $\Omega_i = \overline{\Omega}_i + (\Omega_i - \overline{\Omega}_i)$ and use that $|\Omega' - \Omega|^{\sup}_{\Pi, \ell^\infty} \leqslant \alpha/2$ and $|\Omega - \overline{\Omega}|^{\sup}_{\Pi, \ell^{\infty,-\delta}} \leqslant M$ to conclude that

$$\left|e \cdot \Omega'\right| \geqslant |\overline{\Omega}_i - \overline{\Omega}_j| - \alpha - M\left(\langle i\rangle^\delta + \langle j\rangle^\delta\right).$$

Without loss of generality assume that $i = |i| > j = |j|$. As $0 < \alpha < 1 \leqslant M$ and $i \geqslant 1$ it then follows that

$$\left|e \cdot \Omega'\right| \geqslant |\overline{\Omega}_i - \overline{\Omega}_j| - 3Mi^\delta \geqslant i^d - j^d - \sum_{l=1}^{D} |a_l|\left(i^{d_l} - j^{d_l}\right) - 3Mi^\delta.$$

For $j = 0$ we get, with $C = \sum_{l=1}^D |a_l|$,

$$\left|e \cdot \Omega'\right| \geqslant i^d - Ci^{d_1} - 3Mi^\delta.$$

Choosing $E \geqslant 1$ sufficiently large and using that in this case $|e|_{d-1} = |i|^{d-1} + 1$ and $d_1, \delta < d$ it then follows that for any e with $|e|_* \geqslant E$

$$\left|e \cdot \Omega'\right| \geqslant \frac{1}{8}|e|_{d-1} \quad \forall \xi \in \Pi.$$

If $j \geqslant 1$ note that $i^x - j^x$ is monotone increasing in $x \geqslant 0$ and hence

$$\left|e \cdot \Omega'\right| \geqslant i^d - j^d - C\left(i^{d_1} - j^{d_1}\right) - 3Mi^\delta.$$

Using that $i^d - j^d = (i-j)(i^{d-1} + j^{d-1}) + ji^{d-1} - ij^{d-1}$ and as $d \geqslant 1$,

$$i^d - j^d + ji^{d-1} - ij^{d-1} = i\left(i^{d-1} - j^{d-1}\right) + j\left(i^{d-1} - j^{d-1}\right) \geqslant 0$$

it then follows that

$$2\left(i^d - j^d\right) \geqslant (i-j)\left(i^{d-1} + j^{d-1}\right) \geqslant (i-j)i^{d-1}.$$

On the other hand

$$i^{d_1} - j^{d_1} = \int_j^i d_1 x^{d_1 - 1}\, dx \leqslant d_1 \cdot (i-j) \cdot 1 \vee i^{d_1 - 1}.$$

Altogether we then get

$$\begin{aligned}\left|e\cdot\Omega'\right| &\geqslant \frac{1}{4}(i-j)i^{d-1}\left(1-4Cd_1\frac{1\vee i^{d_1-1}}{i^{d-1}}\right)+\frac{1}{4}i^{d-1}\left(1-12Mi^{\delta-d+1}\right)\\ &\geqslant \frac{1}{8}\left(i^{d-1}+j^{d-1}\right)\left(2-4Cd_1\frac{1\vee i^{d_1-1}}{i^{d-1}}-12Mi^{\delta-d+1}\right).\end{aligned}$$

Choosing $E\geqslant 1$ larger, if necessary, it follows also in this case that for any e with $|e|_*\geqslant E$

$$\left|e\cdot\Omega'(\xi)\right|\geqslant\frac{1}{8}|e|_{d-1}\quad\forall\xi\in\Pi$$

as claimed. □

Lemma 6.6. *For any $e\in\Lambda\setminus\Lambda_r$ with $|e|_*\geqslant E$ and E given as in Lemma 6.5 and for any $k\in\mathbb{Z}^A$ and $0<\alpha<\frac{1}{16}$ with $\mathcal{R}'_{ke}(\alpha)\neq\emptyset$ one has*

$$|k|\geqslant\frac{1}{16}(1+M)^{-1}|e|_{d-1}.$$

Proof. Again we only prove the claimed estimate for $e\in\Lambda\setminus\Lambda_r$ with $e\cdot\Omega'(\xi)$ of the form $\Omega'_i(\xi)-\Omega'_j(\xi)$ for some $i,j\in B$ with $j\neq -i$ which is again the most subtle case. Since by assumption $\mathcal{R}'_{ke}(\alpha)\neq\emptyset$ there exists $\xi\in\Pi$ such that

$$\left|k\cdot\omega'(\xi)+\Omega'_i(\xi)-\Omega'_j(\xi)\right|<\alpha A_k^{-1}|e|_\delta^{\frac{1}{2}}.$$

By Lemma 6.5 one then concludes that for any $e\in\Lambda\setminus\Lambda_r$ with $|e|_*\geqslant E$

$$|k|\left|\omega'\right|_\Pi^{\sup}\geqslant\left|\Omega'_i-\Omega'_j\right|-\left|k\cdot\omega'+\Omega'_i-\Omega'_j\right|\geqslant\frac{1}{8}|e|_{d-1}-\alpha|e|_\delta^{\frac{1}{2}}.$$

As $\alpha<\frac{1}{16}$ and, by assumption, $d-1>\delta$, it then follows that $|k||\omega'|_\Pi^{\sup}\geqslant\frac{1}{16}|e|_{d-1}$. On the other hand, by (6.51) and (6.52),

$$\left|\omega'\right|_\Pi^{\sup}\leqslant\left|\omega'-\omega\right|_\Pi^{\sup}+|\omega|_\Pi^{\sup}\leqslant\frac{\alpha}{2}+M\leqslant 1+M$$

yielding $|k|\geqslant\frac{1}{16}(1+M)^{-1}|e|_{d-1}$ as claimed. □

Introduce

$$E_{nr}:=\left(2E^{d-1-\delta}\right)\vee\left(6\cdot 48\cdot LM(1+M)\right)\quad\text{and}\quad K_{nr}:=6LM\max_{|e|_{d-1-\delta}\leqslant E_{nr}}|e|_\delta$$

where the subscript index *nr* stands for ‘nonresonant’.

Lemma 6.7. *For any $0<\alpha<\frac{1}{16}$ and $(k,e)\in\mathbb{Z}^A\times\mathbb{Z}^B$ with $e\in\Lambda\setminus\Lambda_r$ and either $|k|\geqslant K_{nr}$ or $|e|_{d-1-\delta}\geqslant E_{nr}$,*

$$\operatorname{meas}\left(\mathcal{R}'_{ke}(\alpha)\right)\leqslant 12L(LM\rho)^{|A|-1}\alpha|k|^{-\frac{1}{2}}A_k^{-1}.$$

Proof. Without loss of generality assume that $\mathcal{R}'_{ke}(\alpha) \neq \emptyset$. If $|e|_{d-1-\delta} \geqslant E_{nr}$, then by the definition of E_{nr}, $|e|_{d-1-\delta} \geqslant 2E^{d-1-\delta}$ which in view of $0 \leqslant |e| \leqslant 2$ implies that $|e|_* \geqslant E$. By Lemma 6.6 one then gets $|k| \geqslant \frac{1}{16}(1+M)^{-1}|e|_{d-1}$. Note that $|e|_{d-1-\delta}|e|_\delta \leqslant 3|e|_{d-1}$. Together with the assumption $|e|_{d-1-\delta} \geqslant E_{nr} \geqslant 6 \cdot 48 \cdot LM(1+M)$ it then follows that

$$|k| \geqslant \frac{1}{16}(1+M)^{-1}\frac{1}{3} \cdot 6 \cdot 48 \cdot LM(1+M)|e|_\delta \geqslant 6LM|e|_\delta.$$

If $|e|_{d-1-\delta} < E_{nr}$, then $|k| \geqslant K_{nr} \geqslant 6LM|e|_\delta$. So in both cases, Lemma 6.4 applies, yielding the claimed estimate. $\square$

Next we treat the case of a resonant site, $e \in \Lambda_r$. Let $C_A = 1 \vee \max_{i \in A} |i|$ and introduce

$$E_r := 2(6LMC_A)^{(d-1-\delta)/(1-\delta)} \quad \text{and} \quad K_r := 6LM \max_{\substack{e \in \Lambda_r \\ |e|_{d-1-\delta} \leqslant E_r}} |e|_\delta$$

where the subscript index r stands for 'resonant'.

Lemma 6.8. *For any $(k, e) \in \mathcal{Z}$ with $e \in \Lambda_r$ and either $|k| \geqslant K_r$ or $|e|_{d-1-\delta} \geqslant E_r$*

$$\operatorname{meas}\big(\mathcal{R}'_{ke}(\alpha)\big) \leqslant 12L(LM\rho)^{|A|-1}\alpha|k|^{-\frac{1}{2}}A_k^{-1}.$$

We remark that in the proof of Lemma 6.8 the assumption $\delta < 1$ of Assumption (A2) is used in an essential way.

Proof. Note that for $e \in \Lambda_r$, $|e|_{d-1-\delta} = 2|i|^{d-1-\delta}$. Using that $0 \leqslant \delta < 1$ it then follows that $|e|_{d-1-\delta} \geqslant E_r$ implies $|i|^{1-\delta} \geqslant 6LMC_A$. On the other hand, as $(k, e) \in \mathcal{Z}$, it follows that $2|i| = |k \cdot \nu_A| \leqslant |k|C_A$ which then leads to

$$|k| \geqslant C_A^{-1}2|i| = C_A^{-1}2|i|^\delta|i|^{1-\delta} \geqslant 6LM|e|_\delta.$$

If $|e|_{d-1-\delta} < E_r$, then by assumption $|k| \geqslant K_r$ and hence $|k| \geqslant 6LM|e|_\delta$ as well. Thus in both cases we can again apply Lemma 6.4 to get the claimed estimate. $\square$

It is convenient to combine the statements of Lemma 6.4 for $e = 0$, Lemma 6.7, and Lemma 6.8. Introduce

$$E_* = E_r \vee E_{nr} \quad \text{and} \quad K_* = K_r \vee K_{nr}. \tag{6.56}$$

Corollary 6.2. *For any $(k, e) \in \mathcal{Z} \setminus \mathcal{Z}_*$ with $e \in \Lambda$ and for any $(k, 0) \in \mathcal{Z}$,*

$$\operatorname{meas}\big(\mathcal{R}'_{ke}(\alpha)\big) \leqslant 12L(LM\rho)^{|A|-1}\alpha|k|^{-\frac{1}{2}}A_k^{-1},$$

where

$$\mathcal{Z}_* := \big\{(k, e) \in \mathcal{Z}\colon\ 0 \leqslant |k| < K_*;\ 0 \leqslant |e|_{d-1-\delta} < E_*\big\}.$$

To continue, introduce for any $k \in \mathbb{Z}^A$ the resonance sets

$$\mathcal{R}'_k(\alpha) := \bigcup_{(k,e) \in \mathcal{Z} \setminus \mathcal{Z}_*} \mathcal{R}'_{ke}(\alpha).$$

Remark 6.1. Note that for any $0 < \alpha < 1/16$, $\mathcal{R}'_0(\alpha) = \emptyset$. Indeed, let $(0, e) \in \mathcal{Z} \setminus \mathcal{Z}_*$. If $e \in \Lambda_r$, then $0 = k \cdot \nu_A + e \cdot \nu_B = 2i$ for some $i \in \mathcal{Z}$ with $i, -i \in B$. But $i = 0$ contradicts that $e \in \Lambda_r$. It remains to treat the case $e \in \Lambda \setminus \Lambda_r$. The assumption $(0, e) \notin \mathcal{Z}_*$ implies that $|e|_{d-1-\delta} \geqslant E_*$. As $E_* \geqslant E_{nr} \geqslant 2E^{d-1-\delta}$ it follows that $|e|_* \geqslant E$. By Lemma 6.6 it then follows that $\mathcal{R}'_{0e}(\alpha) = \emptyset$.

The case $k \neq 0$ is treated in the following lemma. Recall that ρ denotes the diameter of Π.

Lemma 6.9. *Assume that* $0 < \alpha < 1/16$. *Then, for any* $k \in \mathbb{Z}^A \setminus \{0\}$,

$$\operatorname{meas}\bigl(\mathcal{R}'_k(\alpha)\bigr) \leqslant C\rho^{|A|-1}\alpha|k|^{-\frac{1}{2}+1\vee 2(d-1)^{-1}} A_k^{-1}$$

where C *is a constant depending on* L, M, A, d, δ *and the coefficients in the expansion of* $\overline{\Omega}$.

Proof. First note that

$$\mathcal{R}'_k(\alpha) = \mathcal{R}'^0_k(\alpha) \cup \mathcal{R}'^r_k(\alpha) \cup \mathcal{R}'^{nr}_k(\alpha)$$

where

$$\mathcal{R}'^0_k(\alpha) = \begin{cases} \mathcal{R}'_{k0}(\alpha) & \text{if } (k,0) \in \mathcal{Z} \setminus \mathcal{Z}_*, \\ \emptyset & \text{if } (k,0) \in \mathcal{Z}_*, \end{cases}$$

and

$$\mathcal{R}'^r_k(\alpha) = \bigcup_{\substack{e \in \Lambda_r \\ (k,e) \in \mathcal{Z} \setminus \mathcal{Z}_*}} \mathcal{R}'_{ke}(\alpha); \qquad \mathcal{R}'^{nr}_k(\alpha) = \bigcup_{\substack{e \in \Lambda \setminus \Lambda_r \\ (k,e) \in \mathcal{Z} \setminus \mathcal{Z}_*}} \mathcal{R}'_{ke}(\alpha).$$

By Corollary 6.2,

$$\operatorname{meas}\bigl(\mathcal{R}'^0_k(\alpha)\bigr) \leqslant 12L(LM\rho)^{|A|-1}\alpha|k|^{-\frac{1}{2}} A_k^{-1}. \tag{6.57}$$

Toward $\mathcal{R}'^r_k(\alpha)$ note that for $(k, e) \in \mathcal{Z} \setminus \mathcal{Z}_*$ with $e \in \Lambda_r$ it follows that

$$0 < 2|i| = |k \cdot \nu_A| \leqslant C_A|k|.$$

Hence

$$\sharp\bigl\{(k,e) \in \mathcal{Z} \setminus \mathcal{Z}_*\colon e \in \Lambda_r\bigr\} \leqslant C_A|k| \tag{6.58}$$

and thus again by Corollary 6.2,

$$\operatorname{meas}\bigl(\mathcal{R}'^r_k(\alpha)\bigr) \leqslant 12C_A L(LM\rho)^{|A|-1}\alpha|k| \cdot |k|^{-\frac{1}{2}} A_k^{-1}. \tag{6.59}$$

To estimate $\operatorname{meas}(\mathcal{R}'^{nr}_k(\alpha))$ we argue as follows. Consider $(k, e) \in \mathcal{Z} \setminus \mathcal{Z}_*$ with $\mathcal{R}'_{ke}(\alpha) \neq \emptyset$ and $e \in \Lambda \setminus \Lambda_r$. If $|e|_{d-1-\delta} \geqslant E_*$, then $|e|_* \geqslant E$ and hence by Lemma 6.6,

$$|k| \geqslant \frac{1}{16}(1+M)^{-1}|e|_{d-1}$$

and thus

$$\begin{aligned}&\sharp\big\{e \in \Lambda \setminus \Lambda_r\colon (k,e) \in \mathcal{Z};\ |e|_{d-1-\delta} \geqslant E_*;\ \mathcal{R}'_{ke}(\alpha) \neq \emptyset\big\}\\ &\quad\leqslant \sharp\big\{e \in \Lambda \setminus \Lambda_r\colon |e|_{d-1} \leqslant 16(1+M)|k|\big\}\\ &\quad\leqslant 3 \cdot 9 \cdot \big(16(1+M)\big)^{2(d-1)^{-1}} |k|^{2(d-1)^{-1}}. \end{aligned} \tag{6.60}$$

Finally

$$\sharp\big\{e \in \Lambda \setminus \Lambda_r\colon (k,e) \in \mathcal{Z} \setminus \mathcal{Z}_*;\ 1 \leqslant |e|_{d-1-\delta} \leqslant E_*\big\} \leqslant 3 \cdot \big(2E_*^{(d-1-\delta)^{-1}} + 1\big)^2.$$

Altogether, we then get again from Corollary 6.2 that $\operatorname{meas}(\mathcal{R}_k'^{nr}(\alpha))$ is bounded by

$$3\big(9\big(16(1+M)\big)^{2(d-1)^{-1}}|k|^{2(d-1)^{-1}} + \big(2E_*^{(d-1-\delta)^{-1}} + 1\big)^2\big) \cdot 12L(LM\rho)^{|A|-1}\alpha|k|^{-\frac{1}{2}}A_k^{-1}. \tag{6.61}$$

Combining (6.57), (6.59), (6.61) leads to the claimed estimate for $\operatorname{meas}(\mathcal{R}'_k(\alpha))$. $\square$

Proof of Theorem 4.1(i). First we need to choose the parameters K_0, τ, and α. Recall that K_0 is given by $K_0 = (c_1\gamma_0)^{-\tau-1}$ where γ_0 satisfies the smallness condition (6.39) and (6.50). In view of the definition, $A_k = \langle k\rangle^\tau$, and of Lemma 6.9, choose $\tau \geqslant |A| + \frac{1}{2} + 1 \vee 2(d-1)^{-1}$. Furthermore, if necessary, choose $0 < \gamma_0$ smaller so that $K_0 \geqslant K_*$ where K_* is given by (6.56). Finally let $0 < \alpha < 1/16$. With these choices we now estimate $\operatorname{meas}(\Pi \setminus \Pi_*)$. Write $\Pi \setminus \Pi_* = \bigcup_{i=1}^4 \Xi_\alpha^i$ where

$$\Xi_\alpha^1 = \bigcup_{\substack{|k|<K_*\\ (k,e)\in\mathcal{Z},\, e\neq 0}} \mathcal{R}^0_{ke}(\alpha_0), \qquad \Xi_\alpha^2 = \bigcup_{\substack{|k|<K_*\\ (k,0)\in\mathcal{Z}}} \mathcal{R}^0_{ke}(\alpha_0),$$

and

$$\Xi_\alpha^3 = \bigcup_{\substack{|k|\geqslant K_*\\ (k,e)\in\mathcal{Z}}} \mathcal{R}^0_{ke}(\alpha_0), \qquad \Xi_\alpha^4 = \bigcup_{\nu\geqslant 1} \bigcup_{\substack{|k|>K_\nu\\ (k,e)\in\mathcal{Z}}} \mathcal{R}^\nu_{ke}(\alpha_\nu).$$

We will estimate $\operatorname{meas}(\Xi_\alpha^i)$, $1 \leqslant i \leqslant 4$, separately. First note that in view of (6.58) and (6.60), for each $0 \leqslant |k| < K_*$, the set $\{e \in \Lambda\colon (k,e) \in \mathcal{Z};\ \mathcal{R}^0_{ke}(\alpha) \neq \emptyset\}$ is finite. Hence Ξ_α^1 is a finite union of resonance sets $\mathcal{R}^0_{ke}(\alpha)$. By its definition, $\mathcal{R}^0_{ke}(\alpha)$ is a closed subset of the compact set $\Pi \subseteq \mathbb{R}^A$ and monotone increasing with respect to α. Furthermore, by Assumption (A3), $\operatorname{meas}(\mathcal{R}^0_{ke}(0)) = 0$. Hence it follows that

$$\lim_{\alpha\to 0} \operatorname{meas}\big(\Xi_\alpha^1\big) = 0.$$

By Corollary 6.2

$$\operatorname{meas}\big(\Xi_\alpha^2\big) \leqslant \sum_{k\neq 0} 12L(LM\rho)^{|A|-1}\alpha|k|^{-\frac{1}{2}}A_k^{-1} \leqslant C\rho^{|A|-1}\alpha \cdot \sum_{k\neq 0}|k|^{-\frac{1}{2}}A_k^{-1} \tag{6.62}$$

and by Lemma 6.9,

$$\operatorname{meas}\big(\Xi_\alpha^3\big) \leqslant \operatorname{meas}\bigg(\sum_{|k|\geqslant K_*} \mathcal{R}^0_k(\alpha)\bigg) \leqslant \sum_{|k|\geqslant K_*} C\rho^{|A|-1}\alpha|k|^{-\frac{1}{2}+1\vee 2(d-1)^{-1}}A_k^{-1}. \tag{6.63}$$

By the choice of $\tau \geqslant |A| + \frac{1}{2} + 1 \vee 2(d-1)^{-1}$, one then gets

$$\begin{aligned}\operatorname{meas}\big(\Xi_\alpha^2\big) + \operatorname{meas}\big(\Xi_\alpha^3\big) &\leqslant C\rho^{|A|-1}\alpha\Big(\sum_{k\neq 0}|k|^{-\frac{1}{2}}A_k^{-1} + \sum_{|k|\geqslant K_*}|k|^{-\frac{1}{2}+1\vee 2(d-1)^{-1}}A_k^{-1}\Big)\\ &\leqslant 2C\rho^{|A|-1}\alpha\sum_{k\neq 0}\frac{1}{|k|^{1+|A|}} \leqslant CC'\rho^{|A|-1}\alpha \end{aligned}\tag{6.64}$$

where C' is a constant only depending on $|A|$. Toward $\operatorname{meas}(\Xi_\alpha^4)$, recall that $\mathcal{R}_{ke}^{\nu}(\alpha_\nu) \subseteq \mathcal{R}_{ke}^{\nu}(\alpha)$ as $\alpha_\nu \leqslant \alpha$, for any $\nu \geqslant 1$. As $K_* \leqslant K_0 < K_\nu$ for any $\nu \geqslant 1$ one concludes that for any nonempty resonance set $\mathcal{R}_{ke}^{\nu}(\alpha_\nu)$ in Ξ_α^4 one has $(k,e) \in \mathcal{Z} \setminus \mathcal{Z}_*$. Lemma 6.9 then implies that

$$\operatorname{meas}\big(\Xi_\alpha^4\big) \leqslant \sum_{\nu\geqslant 1}\sum_{|k|>K_\nu} C\rho^{|A|-1}\alpha|k|^{-\frac{1}{2}+1\vee 2(d-1)^{-1}}A_k^{-1} \leqslant CC'\rho^{|A|-1}\alpha\sum_{\nu\geqslant 1}\frac{1}{K_\nu}. \tag{6.65}$$

By the choice of K_ν, $K_\nu = K_0 2^\nu$, one then gets

$$\operatorname{meas}\big(\Xi_\alpha^4\big) \leqslant CC'\rho^{|A|-1}\alpha\sum_{\nu\geqslant 1}K_0^{-1}/2^\nu \leqslant CC'K_0^{-1}\rho^{|A|-1}\alpha.$$

In particular it follows that

$$\sum_{i=2}^{4}\operatorname{meas}\big(\Xi_\alpha^i\big) = O(\alpha), \quad \alpha \to 0.$$

This finishes the proof of item (i) of Theorem 4.1. □

Acknowledgment

The first author would like to thank Dario Bambusi for discussions and comments.

References

[1] J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, Int. Math. Res. Not. IMRN (1994) 475–497.
[2] J. Bourgain, Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Ann. of Math. 148 (1998) 363–439.
[3] J. Bourgain, Nonlinear Schrödinger Equations, Park City Lect., 1999.
[4] J. Bourgain, Green's Function Estimates for Lattice Schrödinger Operators and Applications, Ann. of Math. Stud., vol. 158, Princeton University Press, 2005.
[5] W. Craig, C.E. Wayne, Periodic solutions of nonlinear Schrödinger equations and Nash Moser method, ETH preprint, 1993, in: J. Semanis (Ed.), Hamiltonian Mechanics, Toruń, 1993, in: NATO Adv. Sci. Inst. Ser. B Phys., vol. 331, Plenum, New York, 1994, pp. 103–122.
[6] H.L. Eliasson, S.B. Kuksin, KAM for the nonlinear Schrödinger equation, Ann. of Math. 172 (2010) 371–435.
[7] J. Geng, J. You, A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions, J. Differential Equations 209 (2005) 1–56.
[8] J. Geng, J. You, A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces, Comm. Math. Phys. 262 (2006) 343–372.
[9] J. Geng, Y. Yi, Quasi-periodic solutions in a nonlinear Schrödinger equation, J. Differential Equations 233 (2007) 512–542.
[10] B. Grébert, T. Kappeler, Perturbations of the defocusing nonlinear Schrödinger equation, Milan J. Math. 71 (2003) 141–174.
[11] B. Grébert, T. Kappeler, Symmetries of the nonlinear Schrödinger equation, Bull. Soc. Math. France 130 (4) (2002) 603–618.
[12] B. Grébert, T. Kappeler, J. Pöschel, The Defocusing NLS Equation and Its Normal Form, EMS Ser. Lect. Math., EMS Publishing House, in press.
[13] T. Kappeler, J. Pöschel, KDV & KAM, Springer, 2003.

[14] S.B. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation, Ann. of Math. 143 (1996) 149–179.
[15] S.B. Kuksin, Analysis of Hamiltonian PDEs, Oxford University Press, Oxford, 2000.
[16] Z. Liang, Quasi-periodic solutions for 1D Schrödinger equations with the nonlinearity $|u|^{2p}u$, J. Differential Equations 244 (2008) 2185–2225.
[17] Z. Liang, J. You, Quasi-periodic solutions for 1D Schrödinger equations with higher order nonlinearity, SIAM J. Math. Anal. 36 (2005) 1965–1990.
[18] J. Pöschel, A KAM theorem for some nonlinear partial differential equations, Ann. Sc. Norm. Super. Pisa Cl. Sci. 23 (1996) 119–148.