KAM FOR THE NONLINEAR BEAM EQUATION 1: SMALL-AMPLITUDE SOLUTIONS.

L. HAKAN ELIASSON, BENOÎT GRÉBERT, AND SERGEI B. KUKSIN

Abstract. In this paper we prove a KAM result for the non linear beam equation on the d-dimensional torus

$$u_{tt} + \Delta^2 u + mu + g(x,u) = 0\ , \quad t \in \mathbb{R},\ x \in \mathbb{T}^d, \qquad (*)$$

where $g(x,u) = 4u^3 + O(u^4)$. Namely, we show that, for generic m, many of the small amplitude invariant finite dimensional tori of the linear equation $(*)_{g=0}$, written as the system

$$u_t = -v, \quad v_t = \Delta^2 u + mu,$$

persist as invariant tori of the nonlinear equation $(*)$, re-written similarly. If $d \geq 2$, then not all the persisted tori are linearly stable, and we construct explicit examples of partially hyperbolic invariant tori. The unstable invariant tori, situated in the vicinity of the origin, create around them some local instabilities, in agreement with the popular belief in the nonlinear physics that small-amplitude solutions of space-multidimensonal hamiltonian PDEs behave in a chaotic way.

The proof uses an abstract KAM theorem from another our publication [15].

8/12/ 2015

Contents

1991 *Mathematics Subject Classification.*
Key words and phrases. Beam equation, KAM theory, hamiltonian systems.

1. Introduction

1.1. **The beam equation and the KAM for PDE theory.** The paper deals with small-amplitude solutions of the multi-dimensional nonlinear beam equation on the torus:

$$u_{tt} + \Delta^2 u + mu = -g(x,u)\,, \quad u = u(t,x),\ t \in \mathbb{R},\ x \in \mathbb{T}^d = \mathbb{R}^d/2\pi\mathbb{Z}^d, \tag{1.1}$$

where g is a real analytic function on $\mathbb{T}^d \times I$ for some neighbourhood I of the origin in $\mathbb{R}$, satisfying

$$g(x,u) = 4u^3 + g_0(x,u)\,, \qquad g_0 = O(u^4). \tag{1.2}$$

m is the mass parameter and we assume that $m \in [1,2]$. This equation is interesting by itself. Besides, it is a good model for the Klein–Gordon equation

$$u_{tt} - \Delta u + mu = -g(x,u), \qquad x \in \mathbb{T}^d, \tag{1.3}$$

which is among the most important equations of mathematical physics. We are certain that the ideas and methods of our work apply – with additional technical efforts – to eq. (1.3) (but the situation with the nonlinear wave equation $(1.3)_{m=0}$, as well as with the zero-mass beam equation, may be quite different).

Our goal is to develop a general KAM-theory for small-amplitude solutions of (1.1). To do this we compare them with time-quasiperiodic solution of the linearised at zero equation

$$u_{tt} + \Delta^2 u + mu = 0\,. \tag{1.4}$$

Decomposing real functions $u(x)$ on $\mathbb{T}^d$ to Fourier series

$$u(x) = \sum_{s\in\mathbb{Z}^d} u_s e^{is\cdot x} \ + \text{c.c.}$$

(here c.c. stands for "complex conjugated"), we write time-quasiperiodic solutions for (1.4), corresponding to a finite set of excited wave-vectors $\mathcal{A} \subset \mathbb{Z}^d$, $|\mathcal{A}| =: n$, as

$$u(t,x) = \sum_{s\in\mathcal{A}} (a_s e^{i\lambda_s t} + b_s e^{-i\lambda_s t}) e^{is\cdot x} + \text{c.c.}, \tag{1.5}$$

where $\lambda_s = \sqrt{|s|^4 + m}$. We examine these solutions and their perturbations in eq. (1.1) under the assumption that the action-vector $I = \{\frac{1}{2}(a_s^2 + b_s^2),\ s \in \mathcal{A}\}$ is small. In our work this goal is achieved provided that
- the finite set $\mathcal{A}$ is typical in some mild sense;
- the mass parameter m does not belong to a certain set of zero measure.

The linear stability of the obtained solutions for (1.1) is under control. If $d \geq 2$, and $|\mathcal{A}| \geq 2$, then some of them are linearly unstable.

The specific choice of a hamiltonian PDE with the mass parameter which we work with – the beam equation (1.1) – is sufficiently arbitrary. This is simply the easiest non-linear space-multidimensional equation from mathematical physics for which we can perform our programme of the KAM-study of small-amplitude solutions in space-multidimensional hamiltonian PDEs, and obtain for them the results, outlines above. We are certain that out picture of the KAM-behaviour of small solutions, as well as the method, developed to prove it, are sufficiently, general. In particular, we believe that out method applies to the Klein-Gordon equation (1.3).

Before to give exact statement of the result, we discuss the state of affairs in the KAM for PDE theory. The theory started in late 1980's and originally applied to 1d hamiltonian PDEs, see in [23, 24, 12]. The first works on this theory treated
a) perturbations of linear hamiltonian PDE, depending on a vector-parameter of the dimension, equal to the number of frequencies of the unperturbed quasiperiodic solution of the linear system (for solutions (1.5) this is $|\mathcal{A}|$).

Next the theory was applied to
b) perturbations of integrable hamiltonian PDE, e.g. of the KdV or Sine-Gordon equations, see [25].

In paper [6]
c) small-amplitude solutions of the 1d Klein-Gordon equation (1.3) with $g(x,u) = -u^3 + O(u^4)$ were treated as perturbed solutions of the Sine-Gordon equation,[1] and a singular version of the KAM-theory b) was developed to study them.

It was proved in [6] that for a.a. values of m and for any finite set $\mathcal{A}$ most of the small-amplitude solutions (1.5) for the linear Klein-Gordon equation (with $\lambda_s = \sqrt{|s|^2 + m}$) persist as linearly stable time-quasiperiodic solutions for (1.3). In [26] it was realised that it is easier to study small solutions of 1d equations like (1.3) not as perturbations of solutions for an integrable PDE, but rather as perturbations of solutions for a Birkhoff–integrable system, after the equation is normalised by a Birkhoff transformation. The paper [26] deals not with 1d Klein-Gordon equation (1.3), but with 1d NLS equation, which is similar to (1.3) for the problem under discussion; in [29] the method of [26] was applied to the 1d equation (1.3). The approach of [26] turned out to be very efficient and later was applied to many other 1d hamiltonian PDEs.

[1] Note that for suitable a and b we have $mu - u^3 + O(u^4) = a\sin bu + O(u^4)$. So the 1d equation (1.3) is the Sine-Gordon equation, perturbed by a small term $O(u^4)$.

Space-multidimensional KAM for PDE theory started 10 years later with the paper [8] and, next, publications [9] and [17]. The just mentioned works deal with parameter-depending linear equations (cf. a)). The approach of [17] is different from that of [8, 9] and allows to analyse the linear stability of the obtained KAM-solutions. Also see [4, 5]. Since integrable space-multidimensional PDE (practically) do not exist, then no multi-dimensional analogy of the 1d theory b) is available.

Efforts to create space-multidimensional analogies of the KAM-theory c) were made in [33] and [30, 31], using the KAM-techniques of [8, 9] and [17], respectively. Both works deal with the NLS equation. Their main disadvantage compare to the 1d theory c) is severe restrictions on the finite set $\mathcal{A}$ (i.e. on the class of unperturbed solutions which the methods allow to perturb). The result of [33] gives examples of some sets $\mathcal{A}$ for which the KAM-persistence of the corresponding small-amplitude solutions (1.5) holds, while the result of [30, 31] applies to solutions (1.5), where the set $\mathcal{A}$ is nondegenerate in certain very non-explicit way. The corresponding notion of non-degeneracy is so complicated that it is not easy to give examples of non-degenerate sets $\mathcal{A}$.

Some KAM-theorems for small-amplitude solutions of multidimensional beam equations (1.1) with typical m were obtained in [18, 19]. Both works treat equations with a constant-coefficient nonlinearity $g(x,u) = g(u)$, which is significantly easier than the general case (cf. the linear theory, where constant-coefficient equations may be integrated by the Fourier method). Similar to [33, 30, 31], the theorems of [18, 19] only allow to perturb solutions (1.5) with very special sets $\mathcal{A}$ (see also Appendix B). Solutions of (1.1), constructed in these works, all are linearly stable.

1.2. **Beam equation in the complex variables.** Introducing $v = u_t \equiv \dot{u}$ we rewrite (1.1) as

$$\begin{cases} \dot{u} &= -v, \\ \dot{v} &= \Lambda^2 u + g(x,u)\,, \end{cases} \tag{1.6}$$

where $\Lambda = (\Delta^2 + m)^{1/2}$. Defining $\psi(t,x) = \frac{1}{\sqrt{2}}(\Lambda^{1/2}u + i\Lambda^{-1/2}v)$ we get for the complex function $\psi(t,x)$ the equation

$$\frac{1}{i}\dot{\psi} = \Lambda\psi + \frac{1}{\sqrt{2}}\Lambda^{-1/2} g\left(x, \Lambda^{-1/2}\left(\frac{\psi + \bar{\psi}}{\sqrt{2}}\right)\right).$$

Thus, if we endow the space $L_2(\mathbb{T}^d, \mathbb{C})$ with the standard real symplectic structure, given by the two-form $-id\psi \wedge d\bar{\psi} = -d\tilde{u} \wedge d\tilde{v}$, where $\psi = \frac{1}{\sqrt{2}}(\tilde{u} + i\tilde{v})$, then equation (1.1) becomes a hamiltonian system

$$\dot{\psi} = i\,\partial H/\partial\bar{\psi}$$

with the hamiltonian function

$$H(\psi,\bar{\psi}) = \int_{\mathbb{T}^d} (\Lambda\psi)\bar{\psi}\mathrm{d}x + \int_{\mathbb{T}^d} G\left(x, \Lambda^{-1/2}\left(\frac{\psi+\bar{\psi}}{\sqrt{2}}\right)\right)\mathrm{d}x.$$

Here G is a primitive of g with respect to the variable u:

$$g = \partial_u G\,, \quad G(x,u) = u^4 + O(u^5)\,.$$

The linear operator Λ is diagonal in the complex Fourier basis

$$\{\varphi_s(x) = (2\pi)^{-d/2} e^{is\cdot x},\ s \in \mathbb{Z}^d\}.$$

Namely,

$$\Lambda\varphi_s = \lambda_s\varphi_s, \quad \lambda_s = \sqrt{|s|^4 + m}, \qquad \forall\, s \in \mathbb{Z}^d .$$

Let us decompose ψ and $\bar\psi$ in the basis $\{\varphi_s\}$:

$$\psi = \sum_{s\in\mathbb{Z}^d} \xi_s\varphi_s, \quad \bar\psi = \sum_{s\in\mathbb{Z}^d} \eta_s\varphi_{-s} .$$

We fix any $d^* > d/2$ and define the space

$$Y^C = \{(\xi,\eta) \in \ell^2(\mathbb{Z}^d,\mathbb{C})\times\ell^2(\mathbb{Z}^d,\mathbb{C}) \mid \sum_s \max(1,|s|^2)^{d^*}(|\xi_s|^2 + |\eta_s|^2) < \infty\}\,, \tag{1.7}$$

corresponding to the Fourier coefficients of complex functions $(\psi(x),\bar\psi(x))$ from the Sobolev space $H^{d^*}(\mathbb{T}^d,\mathbb{C}^2) =: H^{d^*}$. Let us endow Y^C with the complex symplectic structure $-i\sum_s \mathrm{d}\xi_s\wedge\mathrm{d}\eta_s$, and consider there the hamiltonian system

$$\begin{cases} \dot\xi_s &= i\frac{\partial H}{\partial\eta_s} \\ \dot\eta_s &= -i\frac{\partial H}{\partial\xi_s} \end{cases} \qquad s\in\mathbb{Z}^d\,, \tag{1.8}$$

where the hamiltonian function H is given by $H = H_2 + P$ with

$$H_2 = \sum_{s\in\mathbb{Z}^d} \lambda_s\xi_s\eta_s, \quad P = \int_{\mathbb{T}^d} G\left(x, \sum_{s\in\mathbb{Z}^d} \frac{\xi_s\varphi_s + \eta_{-s}\varphi_s}{\sqrt{2\lambda_s}}\right)\mathrm{d}x. \tag{1.9}$$

Then the beam equation (1.6), considered in the Sobolev space $\{(u,v) \mid (\psi,\bar\psi)\in H^{d^*}\}$, is equivalent to the hamiltonian system (1.8), restricted to the real subspace

$$Y^R := \{(\xi,\eta)\in Y^C \mid \eta_s = \bar\xi_s,\ s\in\mathbb{Z}^d\}. \tag{1.10}$$

The leading quartic part of P at the origin,

$$P_4 = \int_{\mathbb{T}^d} u^4\mathrm{d}x = \int_{\mathbb{T}^d}\left(\sum_{s\in\mathbb{Z}^d}\frac{\xi_s\varphi_s + \eta_{-s}\varphi_s}{\sqrt{2\lambda_s}}\right)^4\mathrm{d}x, \tag{1.11}$$

satisfies the *zero momentum condition*, i.e.

$$P_4 = \sum_{i,j,k,\ell\in\mathbb{Z}^d} C(i,j,k,\ell)(\xi_i+\eta_{-i})(\xi_j+\eta_{-j})(\xi_k+\eta_{-k})(\xi_\ell+\eta_{-\ell})\,,$$

where $C(i,j,k,\ell)\neq 0$ only if $i+j+k+\ell = 0$. If g does not depend on x, then P satisfies a similar property at any order. This condition turns out to be useful to restrict the set of small divisors that have to be controlled.

1.3. **Admissible and strongly admissible sets $\mathcal{A}$.** Let $\mathcal{A}$ be a finite subset of $\mathbb{Z}^d$, $|\mathcal{A}| =: n \geq 0$. We define

$$\mathcal{L} = \mathbb{Z}^d \setminus \mathcal{A}\,,$$

and decompose the spaces Y^C and Y^R as

$$Y^C = Y^C_{\mathcal{A}}\oplus Y^C_{\mathcal{L}}\,,\ Y^R = Y^R_{\mathcal{A}}\oplus Y^R_{\mathcal{L}}\,,\ \text{where } Y^C_{\mathcal{A}} = \{(\xi_a,\eta_a),\ a\in\mathcal{A} \mid (\xi,\eta)\in Y^C\}\,,\ \text{etc.}$$

Let us take a vector with positive components $I = (I_a)_{a\in\mathcal{A}} \in \mathbb{R}^n_+$. The n-dimensional real torus

$$T^n_I = \begin{cases} \xi_a = \bar\eta_a,\ |\xi_a|^2 = I_a, & a\in\mathcal{A} \\ \xi_s = \eta_s = 0, & s\in\mathcal{L}\,, \end{cases}$$

is invariant for the linear hamiltonian flow when $P = 0$ (i.e. $g = 0$ in (1.1)). Our goal is to prove the persistency of most of the tori T_I^n when the perturbation P turns on, assuming that the set of nodes $\mathcal{A}$ is *admissible* or *strongly admissible* in the sense, discussed below in this section.

Definition 1.1. *A finite set $\mathcal{A} \in \mathbb{Z}^d$, $|\mathcal{A}| =: n \geq 0$, is called admissible if*

$$j, k \in \mathcal{A},\ j \neq k \Rightarrow |j| \neq |k| .$$

Certainly if $n \leq 1$, then $\mathcal{A}$ is admissible.

For any n, large admissible sets $\mathcal{A}$ with n elements are typical in the following sense. For $R \geq 1$ denote by $B(R)$ the R-ball $\{x \in \mathbb{R}^d \mid |x| \leq R\}$, by $\mathbf{B}(R)$ – the integer ball $\mathbf{B}(R) = B(R) \cap \mathbb{Z}^d$, denote by $S(R)$ the sphere $S(R) = \partial B(R)$, and by $\mathbf{S}(R)$ – the integer sphere $\mathbf{S}(R) = S(R) \cap \mathbb{Z}^d$ (so $\mathbf{S}(R) = \emptyset$ if $R^2 \notin \mathbb{Z}$). Let $\xi^1, \ldots, \xi^n$, $\xi^j = \xi^{j\omega}$, be independent random variables, uniformly distributed in $\mathbf{B}(R)$, $R \geq 1$. Consider the event

$$\Omega_+ = \{\xi^i \neq \xi^j \quad \text{if} \quad i \neq j\} .$$

Then $\mathcal{A}^\omega = \{\xi^{1\omega}, \ldots, \xi^{n\omega}\}$, $\omega \in \Omega_+$, is an n-points random set. We will call it an *n-points random R-set.*

Obviously

$$\mathbb{P}\,\Omega_+ \geq 1 - C(n,d)R^{-d} .$$

Now consider the event

$$\Omega_1 = \{|\xi^i| \neq |\xi^j| \quad \text{for all} \quad i \neq j\} \subset \Omega_+ .$$

The conditional probability $\mathbb{P}\,(\Omega_1 \mid \Omega_+)$ is the probability that an n-points random R-set $\mathcal{A}^\omega$ is admissible. In Appendix E we show that

$$\mathbb{P}\,(\Omega_1 \mid \Omega_+) \geq 1 - C(n,d)R^{-1} . \tag{1.12}$$

So for any n and d

admissible n-points random R-sets with $R \gg 1$ are typical.

Now we define a subclass of admissible sets and start with a notation. For vectors $a, b \in \mathbb{Z}^d$ we write

$$a \angle b \quad \text{iff} \quad \#\{x \in \mathbf{S}(|a|) \mid |x - b| = |a - b|\} \leq 2 , \tag{1.13}$$

and

$$a \angle\!\angle b \quad \text{iff} \quad a \angle a + b .$$

Relation $a \angle b$ means that the integer sphere of radius $|b - a|$ with the centre at b intersects $\mathbf{S}(|a|)$ in at most two points. Obviously,

$$0 \angle b \quad \text{and} \quad 0 \angle\!\angle b \quad \forall\, b . \tag{1.14}$$

If $d = 2$, then for any a we have $a \angle b$ provided that $b \neq 0$, and $a \angle\!\angle b$ if $a + b \neq 0$.

Definition 1.2. *An admissible set $\mathcal{A}$ is called strongly admissible if either $|\mathcal{A}| \leq 1$, or $|\mathcal{A}| \geq 2$ and for any $a, b \in \mathcal{A}$, $a \neq b$, we have $a \angle\!\angle b$.*

Since $a+b\neq 0$ for any two different points of an admissible set, then

for $d\leq 2$ every admissible set is strongly admissible.

In high dimension this is not any more true, e.g. see the set (B.2) in Appendix B. Still strongly admissible n-points random R-sets with $R\gg 1$ are typical. Namely, consider again the random points $\xi^1,\dots,\xi^n$ in $\mathbf{B}(R)$, and consider the event

$$\Omega_2=\{\xi^i \angle\!\angle\, \xi^j \quad \text{for all} \quad i\neq j\}\,.$$

Then the random sets, corresponding to $\omega\in\Omega_1\cap\Omega_2$, are strongly admissible, and $\mathbb{P}\,(\Omega_1\cap\Omega_2\mid\Omega_+)$ is the probability that an n-point random R-set $\mathcal{A}^\omega$ is strongly admissible. Clearly

$$\mathbb{P}\,(\Omega\setminus\Omega_2)\leq n(n-1)(1-\mathbb{P}\,\{\xi^1 \angle\!\angle\, \xi^2\})\,. \tag{1.15}$$

In Appendix E we prove that

$$1-\mathbb{P}\,\{\xi^1 \angle\!\angle\, \xi^2\}\leq CR^{-\kappa}\,, \tag{1.16}$$

where $C=C(n,d)>0$ and $\kappa=\kappa(d)>0$ (e.g. $\kappa(3)=2/9$). By (1.12), (1.15) and (1.16), for any n and d

strongly admissible n-points random R-sets with $R\gg 1$ are typical.

1.4. **Statement of the main results.** We recall that $\mathcal{L}=\mathbb{Z}^d\setminus\mathcal{A}$ and define two subsets of $\mathcal{L}$, important for our construction:

$$\mathcal{L}_f=\{s\in\mathcal{L}\mid \exists\ a\in\mathcal{A} \text{ such that } |a|=|s|\}\,,\quad \mathcal{L}_\infty=\mathcal{L}\setminus\mathcal{L}_f. \tag{1.17}$$

Clearly $\mathcal{L}_f$ is a finite subset of $\mathcal{L}$. For example, if $d=1$ and $\mathcal{A}$ is admissible, then $\mathcal{A}\cap-\mathcal{A}\subset\{0\}$, so

$$\text{if } d=1, \text{ then } \mathcal{L}_f=-(\mathcal{A}\setminus\{0\}). \tag{1.18}$$

In a neighbourhood of an invariant torus T^n_I in the real space $\{(\xi_a=\bar\eta_a, a\in\mathcal{A})\}\subset\mathbb{C}^{2n}$ we introduce the real action-angle variables $(r_a,\theta_a)_\mathcal{A}$ by the relation

$$\xi_a=\sqrt{I_a+r_a}\,e^{i\theta_a}$$

(note that $-i\sum_{a\in\mathcal{A}}d\xi_a\wedge d\eta_a=-dI\wedge d\theta$). We will write

$$\xi^\mathcal{A}=\sqrt{I+r}\,e^{i\theta}\,,\ \eta^\mathcal{A}=\sqrt{I+r}\,e^{-i\theta}\,;\quad \xi^\mathcal{A}=\{\xi_a,a\in\mathcal{A}\}\,,\ \eta^\mathcal{A}=\{\eta_a,a\in\mathcal{A}\}\,. \tag{1.19}$$

We will often denote the internal frequencies by ω, i.e. $\lambda_s=\omega_s$ for $s\in\mathcal{A}$, and we will keep the notation λ_s for the external frequencies with $s\in\mathcal{L}=\mathbb{Z}^d\setminus\mathcal{A}$. Then the quadratic part of the Hamiltonian becomes, up to a constant,

$$H_2=\sum_{a\in\mathcal{A}}\omega_a r_a+\sum_{s\in\mathcal{L}}\lambda_s\xi_s\eta_s.$$

The perturbation P is an analytic function of all variables and reads

$$P(r,\theta,\xi,\eta)=\int_{\mathbb{T}^d}G(x,\hat u_{I,m}(r,\theta,\xi,\eta))\mathrm{d}x\,,$$

where $\hat u_{I,m}(r,\theta,\xi,\eta)$ is $u(x)=\Lambda^{-1/2}(\psi+\bar\psi)/\sqrt{2}$, expressed in the variables (r,θ,ξ_s,η_s):

$$\hat u_{I,m}=\sum_{s\in\mathcal{A}}\sqrt{I_a+r_a}\,\frac{e^{-i\theta_a}\varphi_a(x)+e^{i\theta_a}\varphi_{-a}(x)}{\sqrt{2}\,(|a|^4+m)^{1/4}}+\sum_{s\in\mathcal{L}}\frac{\xi_s\varphi_s(x)+\eta_{-s}\varphi_s(x)}{\sqrt{2}\,(|s|^4+m)^{1/4}}. \tag{1.20}$$

For any $I \in \mathbb{R}_+^n$, $m \in [1,2]$ and $\theta^0 \in \mathbb{T}^d$ the curve

$$r_a(t) = 0,\;\; \theta_a(t) = \theta_a^0 + t\omega_a \text{ for } a \in \mathcal{A}; \quad \xi_s(t) = \eta_s(t) = 0 \text{ for } s \notin \mathcal{A}\,, \tag{1.21}$$

is a solution of the linear beam equation (1.4), lying on the torus T_I^n. Our goal is to perturb the solutions (1.21) to solutions of the nonlinear equation (1.1). The first step is to put the nonlinear problem to a Birkhoff normal form in the vicinity of a small torus T_I^n. To do this we write $I = \nu\rho, \rho \in [c_*, 1]^{\mathcal{A}} =: \mathcal{D}$, where $0 < \nu \ll 1$ and $c_* \in (0, 1/2]$ is a fixed parameter, and in a small neighbourhood of T_I^n make a symplectic change of variables which simplifies the Hamiltonian $H = H_2 + P$. The corresponding result is obtained in Sections 3-4 and may be loosely stated as follows:

Theorem 1.3. *There exists a zero-measure Borel set $\mathcal{C} \subset [1,2]$ such that for any $m \notin \mathcal{C}$, any admissible set $\mathcal{A}$, $|\mathcal{A}| := n \geq 1$, any $c_* \in (0, 1/2]$ and any analytic nonlinearity* (1.2)*, there exist $\nu_0 > 0$ and $\beta_{*0} > 0$, and for any $0 < \nu \leq \nu_0$, $0 < \beta_* \leq \beta_{*0}$ there exists a closed domain $\tilde{Q} \subset \mathcal{D}$ which is a semi-analytic set,*[2] *such that* $\operatorname{meas}(\mathcal{D} \setminus \tilde{Q}) \leq C\nu^{\beta_*}$*, and for every $\rho \in \tilde{Q}$ there exists an analytic symplectic change of variables*

$$\tilde{\Phi}_\rho : (r', \theta', u, v) \mapsto (r, \theta, \xi, \eta),$$

C^∞–Whitney smooth in ρ, with the following property:

i) The transformed Hamiltonian $H_\rho = H \circ \tilde{\Phi}_\rho$ reads

$$\begin{aligned} H_\rho &= (\omega + \nu M\rho) \cdot r' + \frac{1}{2} \sum_{a \in \mathcal{L}_\infty} \Lambda_a(\rho)(u_a^2 + v_a^2) \\ &\quad + \frac{\nu}{2} \left(\sum_{b \in \mathcal{L}_f^e} \Lambda_b(\rho)\Big(u_b^2 + v_b^2\Big) + \Big\langle \widehat{K}(\rho) \begin{pmatrix} u^h \\ v^h \end{pmatrix}, \begin{pmatrix} u^h \\ v^h \end{pmatrix} \Big\rangle \right) + \tilde{f}(r', \theta', \tilde{\zeta}; \rho)\,, \end{aligned} \tag{1.22}$$

where $\mathcal{L} = \mathcal{L}_\infty \cup \mathcal{L}_f\,, \quad \mathcal{L}_f = \mathcal{L}_f^e \cup \mathcal{L}_f^h\,, \quad u^h = (u_a, a \in \mathcal{L}_f^h),\ v^h = (v_a, a \in \mathcal{L}_f^h)$, and the decomposition $\mathcal{L}_f = \mathcal{L}_f^e \cup \mathcal{L}_f^h$ depends on the component of the domain $\tilde{Q}$ (one of the sets $\mathcal{L}_f^e, \mathcal{L}_f^h$ may be empty). The matrix M is explicitly defined in (3.44)*, and each $\Lambda_a(\rho)$ is $C\nu(|a|+1)^{-2}$-close to λ_a. The function $\tilde{f}(\cdot;\rho)$ is analytic and is much smaller than the quadratic part.*

ii) The real symmetric matrix $\widehat{K}(\rho)$ smoothly depends on ρ and for all ρ satisfies $\|\widehat{K}(\rho)\| \leq C\nu^{-c_1\beta_*}$. If $\mathcal{L}_f^h \neq \emptyset$,[3] *then the hamiltonian operator $J\widehat{K}(\rho)$ is hyperbolic, and the moduli of real parts of its eigenvalues are bigger than $C^{-1}\nu^{c_2\beta_*}$. It may be complex-diagonalised by means of a smooth in ρ complex transformation $U(\rho)$ such that $\|U(\rho)\| + \|U(\rho)^{-1}\| \leq C\nu^{-c_3\beta_*}$.*

iii) The matrix $\widehat{K}(\rho)$ and the domain $\tilde{Q}$ do not depend on the component g_0 of the function g.

For exact statement of the normal form result see Theorem 4.6.

Applying to the normal form above an abstract KAM theorem for multidimensional PDEs, proved in [15], we obtain the main results of this work. To state them

[2]More precisely, there is a polynomial $\mathcal{R}$ of $\sqrt{\rho_j}$, $1 \leq j \leq n$, and a $\delta > 0$ such that $\tilde{Q} = \{\rho \in \mathcal{D} \mid \mathcal{R} \geq \delta\}$.

[3]otherwise the operator $J\widehat{K}$ is trivial.

we recall that a Borel subset $\mathfrak{J} \subset \mathbb{R}^n_+$ is said to have a positive density at the origin if

$$\liminf_{\nu \to 0} \frac{\operatorname{meas}(\mathfrak{J} \cap \{x \in \mathbb{R}^n_+ \mid \|x\| < \nu\})}{\operatorname{meas}\{x \in \mathbb{R}^n_+ \mid \|x\| < \nu\}} > 0 . \tag{1.23}$$

The set $\mathfrak{J}$ has the density one at the origin if the lim inf above equals one (so the ratio of the measures of the two sets converges to one as $\nu \to 0$).

Theorem 1.4. *There exists a zero-measure Borel set $\mathcal{C} \subset [1,2]$ such that for any strongly admissible set $\mathcal{A} \subset \mathbb{Z}^d$, $|\mathcal{A}| =: n \geq 1$, any analytic nonlinearity* (1.2)*, any constant $\alpha_* > 0$ and any $m \notin \mathcal{C}$ there exists a Borel set $\mathfrak{J} \subset \mathbb{R}^n_+$, having density one at the origin, with the following property:*
There exist constants $C, c > 0$, a continuous mapping $U : \mathbb{T}^n \times \mathfrak{J} \to Y^R = Y^R_{\mathcal{A}} \oplus Y^R_{\mathcal{L}}$ (see (1.10)*), analytic in the first argument, satisfying*

$$\big|U(\mathbb{T}^n \times \{I\}) - (\sqrt{I}\, e^{i\theta} , \sqrt{I}\, e^{-i\theta}, 0)\big|_{Y^R} \leq C|I|^{1-\alpha_*} \tag{1.24}$$

(see (1.19)*), and a continuous vector-function*

$$\omega' : \mathfrak{J} \to \mathbb{R}^n , \qquad |\omega'(I) - \omega - MI| \leq C|I|^{1+c\alpha_*} , \tag{1.25}$$

where the matrix M is the same as in (1.22)*, such that*

i) for any $I \in \mathfrak{J}$ and $\theta \in \mathbb{T}^n$ the parametrised curve

$$t \mapsto U(\theta + t\omega'(I), I) \tag{1.26}$$

is a solution of the beam equation (1.8)*. Accordingly, for each $I \in \mathfrak{J}$ the analytic n-torus $U(\mathbb{T}^n \times \{I\})$ is invariant for eq.* (1.8)*.*

ii) The set $\mathfrak{J}$ may be written as a countable disjoint union of positive-measure Borel sets $\mathfrak{J}_j$, such that the restrictions of the mapping U to the sets $\mathbb{T}^n \times \mathfrak{J}_j$ and of ω' to the sets $\mathfrak{J}_j$ are Whitney C^1-smooth.

iii) The solution (1.26) *is linearly stable if and only if in* (1.22) *the operator $\widehat{K}(\rho)$ is trivial (i.e. the set $\mathcal{L}^h_f$ is emty). The set $\mathfrak{J}_e$ of ρ's in $\mathfrak{J}$ for which $\widehat{K}(\rho)$ is trivial is of positive measure, and it equals $\mathfrak{J}$ if $d = 1$ or $|\mathcal{A}| = 1$. For $d \geq 2$ and for some choices of the set $\mathcal{A}$, $|\mathcal{A}| \geq 2$, the complement $\mathfrak{J} \setminus \mathfrak{J}_e$ has positive measure.*

We recall that for $d \leq 2$ every admissible set is strongly admissible, but for higher dimension this is not the case. Still if $\mathcal{A}$ is admissible and $d \geq 3$, then a weaker version of the theorem above is true:

Theorem 1.5. *There exists a zero-measure Borel set $\mathcal{C} \subset [1,2]$ such that for any admissible set $\mathcal{A} \subset \mathbb{Z}^d$ with $d \geq 3$ and $|\mathcal{A}| =: n \geq 1$, any analytic nonlinearity* (1.2)*, any constant $\alpha_* > 0$ and any $m \notin \mathcal{C}$ there exists a Borel set $\mathfrak{J} \subset \mathbb{R}^n_+$, having positive density at the origin, such that all assertions of Theorem 1.4 are true.*

Remark 1.6. 1) The torus T^n_I, invariant for the linear beam equation $(1.8)_{G=0}$, is of the size $\sim \sqrt{I}$. If $\alpha_* < 1/2$, then the constructed invariant torus $U(\mathbb{T}^n \times \{I\})$ of the nonlinear beam equation is a small perturbation of T^n_I since by (1.24) the Hausdorff distance between $U(\mathbb{T}^n \times \{I\})$ and T^n_I is smaller than $C|I|^{1-\alpha_*}$.

2) Our result applies to eq. (1.1) with any d. Notice that for d sufficiently large the global in time well-posedness of this equation is unknown.

3) The construction of solutions (1.26) crucially depends on certain equivalence relation in $\mathbb{Z}^d$, defined in terms of the set $\mathcal{A}$ (see (4.5)). This equivalence is trivial if $d = 1$ or $|\mathcal{A}| = 1$ and is non-trivial otherwise.

4) The operator $J\widehat{K}(\rho)$ is complex-conjugated to the hyperbolic part of a complex hamiltonian operator $iJK(\rho)$, corresponding to the complex Birkhoff normal form (3.14) for the beam equation. The operator K is symmetric and *real*. So it seems that "typically" (for some ρ) $iJK(\rho)$ has a nontrivial hyperbolic part, and accordingly $J\widehat{K}(\rho) \neq 0$. We cannot prove this, but discuss in Appendix B examples of sets $\mathcal{A}$ for which the operators iJK have nontrivial hyperbolic parts.

5) The solutions (1.26) of eq. (1.8), written in terms of the $u(x)$-variable as solutions $u(t,x)$ of eq. (1.1), are H^{d^*+1}-smooth as functions of x and analytic as functions of t. Here d^* is a parameter of the construction for which we can take any real number $> d/2$ (see (1.7)). The set $\mathfrak{J}$ depends on d^*, so the theorem's assertion does not imply immediately that the solutions $u(t,x)$ are C^∞–smooth in x. Still, since

$$-(\Delta^2 + m)u = u_{tt} + g(x,u),$$

where g is an analytic function, then the theorems imply by induction that the solutions $u(t,x)$ define analytic curves $\mathbb{R} \to H^p(\mathbb{T}^d)$, for any p. In particular, they are smooth functions.

Notation. *Abstract sets.* We denote a cardinality of a set X as $|X|$ or as $\#X$.

Matrices. For any matrix A, finite or infinite, we denote by ${}^tA$ the transposed matrix; in particular, ${}^t(a,b) = \begin{pmatrix} a \\ b \end{pmatrix}$. If A is a finite matrix, then $\|A\|$ stands for its operator-norm. By J we denote the symplectic matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ as well as various block-diagonal matrices $\operatorname{diag}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, while I stands for the identity matrix of any dimension.

Norms and pairings. For a linear space X of dimension $N \leq \infty$, interpreted as a space of real or complex sequences, we denote by $\langle \cdot, \cdot \rangle$ the natural bi-linear paring: if $X \ni v^j = (v_1^j, \dots, v_N^j)$, $j = 1,2$, then $\langle v^1, v^2 \rangle = \sum_j v_j^1 v_j^2$. Finite-dimensional spaces X as above and the lattices $\mathbb{Z}^N$ are given the Euclidean norm which we denote $|\cdot|$, and the corresponding distance. The tori are provided with the Euclidean distance. For $a \in \mathbb{Z}^N$ we denote $\langle a \rangle = \max(1, |a|)$.

Analytic mappings. We call analytic mappings between domains in complex Banach spaces *holomorphic* to reserve the name *analytic* for mappings between domains in real Banach spaces. A holomorphic mapping is called *real holomorphic* if it maps real-vectors of the space-domain to real vectors of the space-target. Note that when we work with spaces, formed by sequences of complex 2-vectors, we use two different reality conditions. The right one will be clear from the context. A mapping, defined on a closed subset of a Banach space is called analytic (or holomorphic) if it extends to an analytic (holomorphic) map, defined in some open neighbourhood of that set.

Parameters. Our functions depend on a parameter $\rho \in \mathcal{D}$, where $\mathcal{D} \subset \mathbb{R}^p$ is a compact set (or, more generally, a bounded Borel set) of positive Lebesgue measure, with a suitable $p \in \mathbb{N}$. Differentiability of functions on $\mathcal{D}$ is understood in the sense of Whitney. That is, $f \in C^k(\mathcal{D})$ if it extends to a C^k-smooth function $\tilde{f}$ on $\mathbb{R}^p$, and $|f|_{C^k(\mathcal{D})}$ is the infimum of $|\tilde{f}|_{C^k(\mathbb{R}^p)}$, taken over all C^k-extensions $\tilde{f}$ of f.

Acknowledgments. We are thankful for discussion to P. Milman, L. Parnovski and V. Šverák. Our research was supported by l'Agence Nationale de la Recherche through the grant ANR-10-BLAN 0102.

2. Small divisors

2.1. **Non resonance of basic frequencies.** In this subsection we assume that the set $\mathcal{A} \subset \mathbb{Z}^d$ is admissible, i.e. it only contains integer vectors with different norms (see Definition 1.1).
We consider the vector of basic frequencies

$$\omega \equiv \omega(m) = (\omega_a(m))_{a\in\mathcal{A}}\,, \quad m \in [1,2]\,, \tag{2.1}$$

where $\omega_a(m) = \lambda = \sqrt{|a|^4 + m}$. The goal of this section is to prove the following result:

Proposition 2.1. *Assume that $\mathcal{A}$ is an admissible subset of $\mathbb{Z}^d$ of cardinality n included in $\{a \in \mathbb{Z}^d \mid |a| \le N\}$. Then for any $k \in \mathbb{Z}^{\mathcal{A}} \setminus \{0\}$, any $\kappa > 0$ and any $c \in \mathbb{R}$ we have*

$$\text{meas}\left\{m \in [1,2] \ \mid\ \left|\sum_{a\in\mathcal{A}} k_a\omega_a(m) + c\right| \le \kappa\right\} \le C_n \frac{N^{4n^2}\kappa^{1/n}}{|k|^{1/n}}\,,$$

where $|k| := \sum_{a\in\mathcal{A}} |k_a|$ and $C_n > 0$ is a constant, depending only on n.

The proof follows closely that of Theorem 6.5 in [2] (also see [3]); a weaker form of the result was obtained earlier in [7]. Non of the constants C_j etc. in this section depend on the set $\mathcal{A}$.

Lemma 2.2. *Assume that $\mathcal{A} \subset \{a \in \mathbb{Z}^d \mid |a| \le N\}$. For any $p \le n = |\mathcal{A}|$, consider p points $a_1, \cdots, a_p$ in $\mathcal{A}$. Then the modulus of the following determinant*

$$D := \begin{vmatrix} \frac{d\omega_{a_1}}{dm} & \frac{d\omega_{a_2}}{dm} & . & . & . & \frac{d\omega_{a_p}}{dm} \\ \frac{d^2\omega_{a_1}}{dm^2} & \frac{d^2\omega_{a_2}}{dm^2} & . & . & . & \frac{d^2\omega_{a_p}}{dm^2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \frac{d^p\omega_{a_1}}{dm^p} & \frac{d^p\omega_{a_2}}{dm^p} & . & . & . & \frac{d^p\omega_{a_p}}{dm^p} \end{vmatrix}$$

is bounded from below:

$$|D| \ge CN^{-3p^2+p}\,,$$

where $C = C(p) > 0$ is a constant depending only on p.

Proof. First note that, by explicit computation,

$$\frac{d^j\omega_i}{dm^j} = (-1)^j \Upsilon_j \big(|i|^4 + m\big)^{\frac{1}{2}-j}\,, \qquad \Upsilon_j = \prod_{l=0}^{j-1} \frac{2l-1}{2}\,. \tag{2.2}$$

Inserting this expression in D, we deduce by factoring from each $l-th$ column the term $(|a_\ell|^4 + m)^{-1/2} = \omega_\ell^{-1}$, and from each $j-th$ row the term Υ_j that the determinant, up to a sign, equals

$$\left[\prod_{l=1}^{p}\omega_{a_\ell}^{-1}\right]\left[\prod_{j=1}^{p}\Upsilon_j\right]\times\begin{vmatrix}1 & 1 & 1 & . & . & . & 1\\ x_{a_1} & x_{a_2} & x_{a_3} & . & . & . & x_{a_p}\\ x_{a_1}^2 & x_{a_2}^2 & x_{a_3}^2 & . & . & . & x_{a_p}^2\\ . & . & . & . & . & . & .\\ . & . & . & . & . & . & .\\ . & . & . & . & . & . & .\\ x_{a_1}^p & x_{a_2}^p & x_{a_3}^p & . & . & . & x_{a_p}^p\end{vmatrix},$$

where we denoted $x_a := (|a|^4+m)^{-1} = \omega_a^{-2}$. Since $|\omega_{a_k}| \le 2|a_k|^2 \le 2N^2$ for every k, the first factor is bigger than $(2N^2)^{-p}$. The second is a constant, while the third is the Vandermond determinant, equal to

$$\prod_{1\le l<k\le p}(x_{a_\ell}-x_{a_k}) = \prod_{1\le l<k\le p}\frac{|a_k|^4-|a_\ell|^4}{\omega_{a_\ell}^2\omega_{a_k}^2} =: V\,.$$

Since $\mathcal{A}$ is admissible, then

$$|V| \ge \prod_{1\le l<k\le p}\frac{|a_k|^2+|a_\ell|^2}{\omega_{a_\ell}^2\omega_{a_k}^2} \ge \big(\frac{1}{4}\big)^{p(p-1)}N^{-3p(p-1)}\,,$$

where we used that each factor is bigger than $\frac{1}{16}N^{-6}$ using again that $|\omega_{a_k}| \le 2|a_k|^2 \le 2N^2$ for every k. This yields the assertion. $\square$

Lemma 2.3. *Let $u^{(1)},...,u^{(p)}$ be p independent vectors in $\mathbb{R}^p$ of norm at most one, and let $w\in\mathbb{R}^p$ be any non-zero vector. Then there exists $i\in[1,...,p]$ such that*

$$|u^{(i)}\cdot w| \ge C_p|w||\det(u^{(1)},\dots,u^{(p)})|\,.$$

Proof. Without lost of generality we may assume that $|w|=1$.

Let $|u^{(i)}\cdot w| \le a$ for all i. Consider the p-dimensional parallelogram Π, generated by the vector $u^{(1)},...,u^{(p)}$ in $\mathbb{R}^p$ (i.e., the set of all linear combinations $\sum x_j u^{(j)}$, where $0\le x_j\le 1$ for all j). It lies in the strip of width $2pa$, perpendicular to the vector w, and its projection to to the $p-1$-dimensional space, perpendicular to w, lies in the ball around zero of radius p. Therefore the volume of Π is bounded by $C_p p^{p-1}(2pa) = C_p' a$. Since this volume equals $|\det(u^{(1)},\dots,u^{(p)})|$, then $a \ge C_p|\det(u^{(1)},\dots,u^{(p)})|$. This implies the assertion. $\square$

Consider vectors $\frac{d^i\omega}{dm^i}(m)$, $1\le i\le n$, denote $K_i = |\frac{d^i\omega}{dm^i}(m)|$ and set

$$u^{(i)} = K_i^{-1}\frac{d^i\omega}{dm^i}(m), \qquad 1\le i\le n\,.$$

From (2.2) we see that[4] $K_i \le C_n$ for all $1\le i\le n$ (as before, the constant does not depend on the set $\mathcal{A}$). Combining Lemmas 2.2 and 2.3, we find that for any vector w and any $m\in[1,2]$ there exists $r=r(m)\le n$ such that

$$\begin{aligned}\Big|\frac{d^r\omega}{dm^r}(m)\cdot w\Big| = K_r\big|u^{(r)}\cdot w\big| &\ge K_rC_n|w|(K_1\dots K_n)^{-1}|D|\\ &\ge C_n|w|N^{-3n^2+n}\,.\end{aligned}\tag{2.3}$$

Now we need the following result (see Lemma B.1 in [14]):

[4] In this section C_n denotes any positive constant depending only on n.

Lemma 2.4. *Let $g(x)$ be a C^{n+1}-smooth function on the segment [1,2] such that $|g'|_{C^n} = \beta$ and $\max_{1\le k\le n}\min_x |\partial^k g(x)| = \sigma$. Then*

$$\mathrm{meas}\{x \mid |g(x)| \le \rho\} \le C_n\big(\frac{\beta}{\sigma}+1\big)\big(\frac{\rho}{\sigma}\big)^{1/n}.$$

Consider the function $g(m) = |k|^{-1}\sum_{a\in\mathcal{A}} k_a\omega_a(m) + |k|^{-1}c$. Then $|g'|_{C^n} \le C_n'$, and $\max_{1\le k\le n}\min_m |\partial^k g(m)| \ge C_n N^{-3n^2+n}$ in view of (2.3). Therefore, by Lemma 2.4,

$$\begin{aligned}\mathrm{meas}\{m \mid |g(m)| \le \frac{\kappa}{|k|}\} \le C_n N^{3n^2-n}\big(\frac{\kappa}{|k|}N^{3n^2-n}\big)^{1/n}\\ = C_n N^{3n^2+2n-1}\big(\frac{\kappa}{|k|}\big)^{1/n}.\end{aligned}$$

This implies the assertion of the proposition.

2.2. **Small divisors estimates.** We recall the notation (1.17), (2.1), and note the elementary estimates

$$\langle a\rangle^2 < \lambda_a(m) < \langle a\rangle^2 + \frac{m}{2\langle a\rangle^2} \qquad \forall\, a\in\mathbb{Z}^d,\ m\in[1,2], \tag{2.4}$$

where $\langle a\rangle = \max(1,|a|^2)$. In this section we study four type of linear combinations of the frequencies $\lambda_a(m)$:

$$\begin{aligned} D_0 =&\omega\cdot k, \quad k\in\mathbb{Z}^{\mathcal{A}}\setminus\{0\}\\ D_1 =&\omega\cdot k+\lambda_a, \quad k\in\mathbb{Z}^{\mathcal{A}},\ a\in\mathcal{L}\\ D_2^{\pm} =&\omega\cdot k+\lambda_a\pm\lambda_b, \quad k\in\mathbb{Z}^{\mathcal{A}},\ a,b\in\mathcal{L}.\end{aligned}$$

In subsequent sections they will become divisors for our constructions, so we call these linear combinations "divisors".

Definition 2.5. *Consider independent formal variables $x_0, x_1, x_2, \dots$. Now take any divisor of the form D_0, D_1 or $D_2^{\pm}$, write there each $\omega_a, a\in\mathcal{A}$, as λ_a, and then replace every $\lambda_a, a\in\mathbb{Z}^d$, by $x_{|a|^2}$. Then the divisor is called resonant if the obtained algebraical sum of the variables $x_j, j\ge 0$, is zero. Resonant divisors are also called trivial resonances.*

Note that a D_0-divisor cannot be resonant since $k\ne 0$ and the set $\mathcal{A}$ is admissible; a D_1-divisor $(k;a)$ is resonant only if $a\in\mathcal{L}_f$, $|k|=1$ and $\omega\cdot k = -\omega_b$, where $|a|=|b|$. Finally, a D_2^+-divisor or a D_2^- divisor with $k\ne 0$ may be resonant only when $(a,b)\in\mathcal{L}_f\times\mathcal{L}_f$, while the divisors D_2^- of the form $\lambda_a-\lambda_b$, $|a|=|b|$, all are resonant. So there are finitely many trivial resonances of the form D_0, D_1, D_2^+ and of the form D_2^- with $k\ne 0$, but infinitely many of them of the form D_2^- with $k=0$.

Our first aim is to remove from the segment $[1,2]=\{m\}$ a small subset to guarantee that for the remaining m's moduli of all non-resonant divisors admit positive lower bounds. Below in this section

(2.5) constants C, C_1 etc. depend on the admissible set $\mathcal{A}$, while the exponents c_1, c_2 etc depend only on $|\mathcal{A}|$. Borel sets $\mathcal{C}_\kappa$ etc. depend on the indicated arguments and $\mathcal{A}$.

We begin with the easier divisors D_0, D_1 and D_2^+.

Proposition 2.6. *Let $1 \ge \kappa > 0$. There exists a Borel set $\mathcal{C}_\kappa \subset [1,2]$ and positive constants C (cf. (2.5)), satisfying* meas $\mathcal{C}_\kappa \le C\kappa^{1/(n+2)}$, *such that for all $m \notin \mathcal{C}_\kappa$, all k and all $a, b \in \mathcal{L}$ we have*

$$|\omega \cdot k| \ge \kappa \langle k \rangle^{-n^2}, \qquad \textit{except if } k = 0, \tag{2.6}$$

$$|\omega \cdot k + \lambda_a| \ge \kappa \langle k \rangle^{-3(n+1)^3}, \qquad \textit{except if the divisor is a trivial resonance,} \tag{2.7}$$

$$|\omega \cdot k + \lambda_a + \lambda_b| \ge \kappa \langle k \rangle^{-3(n+2)^3}, \ \textit{except if the divisor is a trivial resonance.} \tag{2.8}$$

Here $\langle k \rangle = \max(|k|, 1)$.

Besides, for each $k \neq 0$ there exists a set $\mathfrak{A}^k_\kappa$ whose measure is $\le C\kappa^{1/n}$ such that for $m \notin \mathfrak{A}^k_\kappa$ we have

$$|\omega \cdot k + j| \ge \kappa \langle k \rangle^{-(n+1)n} \textit{for all } j \in \mathbb{Z}\ . \tag{2.9}$$

Proof. We begin with the divisors (2.6). By Proposition 2.1 for any non-zero k we have

$$\text{meas}\{m \in [1,2] \mid |\omega \cdot k| \le \kappa |k|^{-n^2}\} < C\kappa^{1/n} |k|^{-n-1/n}\ .$$

Therefore the relation (2.6) holds for all non-zero k if $m \notin \mathfrak{A}_0$, where meas $\mathfrak{A}_0 \le C\kappa^{1/n} \sum_{k \neq 0} |k|^{-n-1/n} = C\kappa^{1/n}$.

Let us consider the divisors (2.7). For $k = 0$ the required estimate holds trivially. If $k \neq 0$, then the relation, opposite to (2.7) implies that $|\lambda_a| \le C|k|$. So we may assume that $|a| \le C|k|^{1/2}$. If $|a| \notin \{|s| \mid s \in \mathcal{A}\}$, then Proposition 2.1 with $n := n+1$, $\mathcal{A} := \mathcal{A} \cup \{a\}$ and $N = C|k|^{1/2}$ implies that

$$\begin{aligned}&\text{meas}\{m \in [1,2] \mid |\omega \cdot k + \lambda_a| \le \kappa |k|^{-3(n+1)^3}\} \\ &\le C\kappa^{1/(n+1)} |k|^{2(n+1)^2 - 3(n+1)^2 - \frac{1}{n+1}} \le C\kappa^{1/(n+1)} |k|^{-(n+1)^2}\ .\end{aligned}$$

This relation with $n+1$ replaced by n also holds if $|a| = |s|$ for some $s \in \mathcal{A}$, but $\omega \cdot k + \lambda_a$ is not a trivial resonant. Since for fixed k the set$\{\lambda_a \mid |a|^2 \le C|k|\}$ has cardinality less than $2C|k|$, then the relation $|\omega \cdot k + \lambda_a| \le \kappa |k|^{-3(n+1)^3}$ holds for a fixed k and all a if we remove from [1,2] a set of measure $\le C\kappa^{1/(n+1)} |k|^{-(n+1)^2+1} \le C\kappa^{1/(n+1)} |k|^{-n-1}$. So we achieve that the relation (2.7) holds for all k if we remove from $[1,2]$ a set $\mathfrak{A}_1$ whose measure is bounded by $C\kappa^{1/(n+1)} \sum_{k \neq 0} |k|^{-n-1} = C\kappa^{1/(n+1)}$.

For a similar reason there exist a Borel set $\mathfrak{A}_2$ whose measure is bounded by $C\kappa^{1/(n+2)}$ and such that (2.8) holds for $m \notin \mathfrak{A}_2$. Taking $\mathcal{C}_\kappa = \mathfrak{A}_0 \cup \mathfrak{A}_1 \cup \mathfrak{A}_2$ we get (2.6)-(2.8). Proof of (2.9) is similar. □

Now we control divisors $D_2^- = \omega \cdot k + \lambda_a - \lambda_b$.

Proposition 2.7. *There exist positive constants C, c, c_- and for $0 < \kappa$ there is a Borel set $\mathcal{C}'_\kappa \subset [1,2]$ (cf. (2.5)), satisfying*

$$\text{meas } \mathcal{C}'_\kappa \le C\kappa^c, \tag{2.10}$$

such that for all $m \in [1,2] \setminus \mathcal{C}'_\kappa$, all $k \neq 0$ and all $a, b \in \mathcal{L}$ we have

$$R(k; a, b) := |\omega \cdot k + \lambda_a - \lambda_b| \ge \kappa |k|^{-c_-}, \tag{2.11}$$

except if the divisor is a trivial resonance

Proof. We may assume that $|b| \geq |a|$. We get from (2.4) that

$$|\lambda_a - \lambda_b - (|a|^2 - |b|^2)| \leq m|a|^{-2} \leq 2|a|^{-2}.$$

Take any $\kappa_0 \in (0,1]$ and construct the set $\mathfrak{A}^k_{\kappa_0}$ as in Proposition 2.6. Then $\operatorname{meas}\mathfrak{A}^k_{\kappa_0} \leq C\kappa_0^{1/n}$ and for any $m \notin \mathfrak{A}^k_{\kappa_0}$ we have

$$R := R(k;a,b) \geq \left|\omega\cdot k + |a|^2 - |b|^2\right| - 2|a|^{-2} \geq \kappa_0|k|^{-(n+1)n} - 2|a|^{-2}\,.$$

So $R \geq \frac{1}{2}\kappa_0|k|^{-(n+1)n}$ and (2.11) holds if

$$|b|^2 \geq |a|^2 \geq 4\kappa_0^{-1}|k|^{(n+1)n} =: Y_1.$$

If $|a|^2 \leq Y_1$, then

$$R \geq \lambda_b - \lambda_a - C|k| \geq |b|^2 - Y_1 - C|k| - 1.$$

Therefore (2.11) also holds if $|b|^2 \geq Y_1 + C|k| + 2$, and it remains to consider the case when $|a|^2 \leq Y_1$ and $|b|^2 \leq Y_1 + C|k| + 2$. That is (for any fixed non-zero k), consider the pairs (λ_a, λ_b), satisfying

$$|a|^2 \leq Y_1, \qquad |b|^2 \leq Y_1 + 2 + C|k| =: Y_2\,. \tag{2.12}$$

There are at most CY_1Y_2 pairs like that. Since the divisor $\omega\cdot k + \lambda_a - \lambda_b$ is not resonant, then in view of Proposition 2.1 with $N = Y_2^{1/2}$ and $|\mathcal{A}| \leq n+2$, for any $\tilde\kappa > 0$ there exists a set $\mathfrak{B}^k_{\tilde\kappa} \subset [1,2]$, whose measure is bounded by

$$C\tilde\kappa^{1/(n+2)}\kappa_0^{-c_1}|k|^{c_2}, \qquad c_j = c_j(n) > 0,$$

such that $R \geq \tilde\kappa$ if $m \notin \mathfrak{B}^k_{\tilde\kappa}$ for all pairs (a,b) as in (2.12) (and k fixed).

Let us choose $\tilde\kappa = \kappa_0^{2c_1(n+2)}$. Then $\operatorname{meas}\mathfrak{B}^k_{\tilde\kappa} \leq C\kappa_0^{c_1}|k|^{c_2}$ and $R \geq \kappa_0^{2c_1(n+2)}$ for a, b as in (2.12). Denote $\mathfrak{C}^k_{\kappa_0} = \mathfrak{A}^k_{\kappa_0} \cup \mathfrak{B}^k_{\tilde\kappa}$. Then $\operatorname{meas}\mathfrak{C}^k_{\kappa_0} \leq C\big(\kappa_0^{1/n} + \kappa_0^{c_1}|k|^{c_2}\big)$, and for m outside this set and all a, b (with k fixed) we have $R \geq \min\big(\frac{1}{2}\kappa_0|k|^{-(n+1)n}, \kappa_0^{2c_1(n+2)}\big)$. We see that if $\kappa_0 = \kappa_0(k) = 2\kappa^{c_3}|k|^{-c_4}$ with suitable $c_3, c_4 > 0$, then

$$\operatorname{meas}\big(\mathcal{C}'_\kappa = \cup_{k\neq 0}\mathfrak{C}^k_{\kappa_0}\big) \leq C\kappa^{c_3}\,,$$

and, if m is outside $\mathcal{C}'_\kappa$, $R(k;a,b) \geq \kappa|k|^{-c_-}$ with suitable $c_- > 0$. □

It remains to consider the divisors D_2^- with $k = 0$, i.e. $D_2^- = \lambda_a - \lambda_b$. Such a divisor is resonant if $|a| = |b|$.

Lemma 2.8. *Let $m \in [1,2]$ and the divisor $D_2^- = \lambda_a - \lambda_b$ is non-resonant, i.e. $|a| \neq |b|$. Then $|\lambda_a - \lambda_b| \geq \frac{1}{4}$.*

Proof. We have

$$|\lambda_a - \lambda_b| = \frac{\big||a|^4 - |b|^4\big|}{\sqrt{|a|^4 + m} + \sqrt{|b|^4 + m}} \geq \frac{|a|^2 + |b|^2}{\sqrt{|a|^4 + m} + \sqrt{|b|^4 + m}} \geq \frac{1}{4}.$$

□

By construction the sets $\mathcal{C}_\kappa$ and $\mathcal{C}'_\kappa$ decrease with κ. Let us denote

$$\mathcal{C} = \bigcap_{\kappa>0}(\mathcal{C}_\kappa \cup \mathcal{C}'_\kappa)\,. \tag{2.13}$$

From Propositions 2.6, 2.7 and Lemma 2.8 we get:

Proposition 2.9. *The set $\mathcal{C}$ is a Borel subset of $[1,2]$ of zero measure. For any $m \notin \mathcal{C}$ there exists $\kappa_* = \kappa_*(m) > 0$ such that the relations* (2.6)*,* (2.7)*,* (2.8) *and* (2.11) *hold with $\kappa = \kappa_*$.*

In particular, if $m \notin \mathcal{C}$, then any of the divisors

$$\omega \cdot s, \ \ \omega \cdot s \pm \lambda_a, \ \ \omega \cdot s \pm \lambda_a \pm \lambda_b, \quad s \in \mathbb{Z}^d, \ a, b \in \mathcal{L},$$

vanishes only if this is a trivial resonance. If it is not, then its modulus admits a qualified estimate from below.

The zero-measure Borel set $\mathcal{C}$ serves a fixed admissible set $\mathcal{A}$, $\mathcal{C} = \mathcal{C}_{\mathcal{A}}$. But since the set of all admissible sets is countable, then replacing $\mathcal{C}$ by $\cup_{\mathcal{A}} \mathcal{C}_{\mathcal{A}}$ we obtain a zero-measure Borel set which suits all admissible sets $\mathcal{C}$. For further purposes we modify $\mathcal{C}$ as follows:

$$\mathcal{C} =: \mathcal{C} \cup \{\tfrac{4}{3}, \tfrac{5}{3}\}\,. \tag{2.14}$$

3. The normal form

In Sections 3 and 4 we construct a symplectic change of variable that puts the Hamiltonian (1.9) to a normal form, suitable to apply the abstract KAM theorem that we have proved in [15]. Our notation mostly agrees with [15]. Constants in the estimates may depend on the dimension d, but this dependence is not indicated.

3.1. **Notation and statement of the result.**

We start with recalling some notation from [15]. Let $\mathcal{L}$ be any subset of $\mathbb{Z}^d$ (it is not excluded that $\mathcal{L} = \mathbb{Z}^d$). We fix any constant[5]

$$d^* > \tfrac{1}{2} d\,,$$

and for $\gamma \in [0,1]$ denote by $Y_\gamma^{\mathcal{L}}$ the following weighted complex ℓ_2-space

$$Y_\gamma^{\mathcal{L}} = \{\zeta^{\mathcal{L}} = \Big(\zeta_s = \begin{pmatrix} \xi_s \\ \eta_s \end{pmatrix} \in \mathbb{C}^2,\, s \in \mathcal{L}\Big) \mid \|\zeta^{\mathcal{L}}\|_\gamma < \infty\}, \tag{3.1}$$

where[6]

$$\|\zeta^{\mathcal{L}}\|_\gamma^2 = \sum_{s \in \mathcal{L}} |\zeta_s|^2 \langle s \rangle^{2d^*} e^{2\gamma|s|}, \qquad \langle s \rangle = \max(|s|, 1).$$

We will often drop the upper index $^{\mathcal{L}}$ and write Y_γ and ζ instead of $Y_\gamma^{\mathcal{L}}$ and $\zeta^{\mathcal{L}}$.

In a space $Y_\gamma = Y_\gamma^{\mathcal{L}}$ we define the complex conjugation as the involution

$$\zeta = {}^t(\xi, \eta) \mapsto {}^t(\bar\eta, \bar\xi)\,. \tag{3.2}$$

Accordingly, the real subspace of Y_γ is the space

$$Y_\gamma^R = Y_\gamma^{\mathcal{L}R} = \Big\{\zeta_s = \begin{pmatrix} \xi_s \\ \eta_s \end{pmatrix} \mid \eta_s = \bar\xi_s,\ s \in \mathcal{L}\Big\}\,. \tag{3.3}$$

Any mapping defined on (some part of) Y_γ with values in a complex Banach space with a given real part is called *real* if it gives real values to real arguments.

We denote by $\mathcal{M}_\gamma$ the set of infinite symmetric matrices $A : \mathcal{L} \times \mathcal{L} \to \mathcal{M}_{2\times 2}$ valued in the space of 2×2 matrices and satisfying

$$|A|_\gamma := \sup_{a,b \in \mathcal{L}} |A_a^b| \max([a-b], 1)^{d_*} e^{\gamma[a-b]} < \infty,$$

[5]The constants in the estimates below may depend on d^*, but this dependence never is indicated.

[6]We recall that $|\cdot|$ signifies the Euclidean norm.

where
$$[a-b]=\min(|a-b|,|a+b|)$$
(this is a pseudo-metric in $\mathbb{Z}^d$). Let us define the operator
$$D=\mathrm{diag}\{\langle s\rangle I, s\in\mathcal{L}\}$$
(here I stands for the identity 2×2-matrix). We denote by $\mathcal{M}_\gamma^D$ the set of infinite matrices $A\in\mathcal{M}_\gamma$ such that $DAD\in\mathcal{M}_\gamma$, and set
$$|A|_\gamma^D=|DAD|_\gamma=\sup_{a,b\in\mathcal{L}}\langle a\rangle\langle b\rangle|A_a^b|\max([a-b],1)^{d_*}e^{\gamma[a-b]}.$$
We note that in [15] instead of the norm $|\cdot|_\gamma^D$ we use the norm $|\cdot|_\gamma^\varkappa$ which for $\varkappa=2$ is "weakly equivalent" to $|\cdot|_\gamma^D$ in the sense that
$$|\cdot|_\gamma^D\le C_{\gamma'}|\cdot|_{\gamma'}^2\quad\forall\,\gamma'>\gamma\,,\qquad |\cdot|_\gamma^2\le C_{\gamma'}|\cdot|_{\gamma'}^D\quad\forall\,\gamma'>\gamma\,.$$

For a Banach space B (real or complex) we denote
$$\mathcal{O}_s(B)=\{x\in B\mid \|x\|_B<s\}\,,$$
and for $\sigma,\gamma,\mu\in(0,1]$ we set
$$\begin{aligned}\mathbb{T}_\sigma^n=&\{\theta\in\mathbb{C}^n/2\pi\mathbb{Z}^n\mid|\Im\theta|<\sigma\},\\ \mathcal{O}^\gamma(\sigma,\mu)=&\mathcal{O}_{\mu^2}(\mathbb{C}^n)\times\mathbb{T}_\sigma^n\times\mathcal{O}_\mu(Y_\gamma)=\{(r,\theta,\zeta)\},\\ \mathcal{O}^{\gamma\mathbb{R}}(\sigma,\mu)=&\mathcal{O}^\gamma(\sigma,\mu)\cap\{\mathbb{R}^n\times\mathbb{T}^n\times Y_\gamma^R\}.\end{aligned}$$
The introduced domains depend on the set $\mathcal{L}$. To indicate this dependence we will sometime write them as
$$\mathcal{O}^\gamma(\sigma,\mu)=\mathcal{O}^\gamma(\sigma,\mu)^{\mathcal{L}},\quad \mathcal{O}^{\gamma\mathbb{R}}(\sigma,\mu)=\mathcal{O}^{\gamma\mathbb{R}}(\sigma,\mu)^{\mathcal{L}}.$$
We will denote the points in $\mathcal{O}^\gamma(\sigma,\mu)$ as $x=(r,\theta,\zeta)$.

Example 3.1. If $\hat f=(\hat f_s,s\in\mathbb{Z}^d)\in Y_\sigma=Y_\sigma^{\mathbb{Z}^d}$, then the function $f(y)=\sum\hat f_se^{is\cdot y}$ is a holomorphic vector-function on $\mathbb{T}_\sigma^n$ and its norm is bounded by $C_d\|\hat f\|_\sigma$. Conversely, if $f:\mathbb{T}_\sigma^n\to\mathbb{C}^2$ is a bounded holomorphic function, then its Fourier coefficients satisfy $|\hat f_s|\le\mathrm{Const}\,e^{-|s|\sigma}$, so $\hat f\in Y_{\sigma'}^{\mathbb{Z}^d}$ for any $\sigma'<\sigma$.

Let $h:\mathcal{O}^0(\sigma,\mu)\times\mathcal{D}\to\mathbb{C}$ be a C^1-function, real holomorphic (see Notation) in the first variable $x=(r,\theta,\zeta)$, such that for all $0\le\gamma'\le\gamma$ and all $\rho\in\mathcal{D}$ the gradient-map
$$\mathcal{O}^{\gamma'}(\sigma,\mu)\ni x\mapsto\nabla_\zeta f(x,\rho)\in Y_\gamma$$
and the hessian-map
$$\mathcal{O}^{\gamma'}(\sigma,\mu)\ni x\mapsto\nabla_\zeta^2 f(x,\rho)\in\mathcal{M}_\gamma^D$$
also are real holomorphic. We denote this set of functions by $\mathcal{T}^{\gamma,D}(\sigma,\mu,\mathcal{D})=\mathcal{T}^{\gamma,D}(\sigma,\mu,\mathcal{D})^{\mathcal{L}}$.

For a function $h\in\mathcal{T}^{\gamma,D}(\sigma,\mu,\mathcal{D})$ we define the norm
$$[h]_{\sigma,\mu,\mathcal{D}}^{\gamma,D}$$
through
$$\sup_{\substack{0\le\gamma'\le\gamma\\ j=0,1}}\ \sup_{\substack{x\in\mathcal{O}^{\gamma'}(\sigma,\mu)\\ \rho\in\mathcal{D}}}\max(|\partial_\rho^jh(x,\rho)|,\mu\|\partial_\rho^j\nabla_\zeta h(x,\rho)\|_{\gamma'},\mu^2|\partial_\rho^j\nabla_\zeta^2h(x,\rho)|_{\gamma'}^D).\tag{3.4}$$

For any function $h \in \mathcal{T}^{\gamma,D}(\sigma,\mu,\mathcal{D})$ we denote by h^T its Taylor polynomial at $r=0,\zeta=0$, linear in r and quadratic in ζ:

$$h(x,\rho) = h^T(x,\rho) + O(|r|^2 + \|\zeta\|^3 + |r|\|\zeta\|).$$

For $\mathcal{L} = \mathbb{Z}^d$ we denote

$$\mathcal{T}^{\gamma,D}(\mu) = \{f \in \mathcal{T}^{\gamma,D}(\sigma,\mu,\mathcal{D})^{\mathbb{Z}^d} : f = f(\zeta)\} \tag{3.5}$$

(i.e., f is independent from θ, r and ρ). The norm (3.4), restricted to the space $\mathcal{T}^{\gamma,D}(\mu)$, will be denoted $[h]^{\gamma,D}_{\mu}$.

Let P be the hamiltonian function defined in (1.9).

Lemma 3.2. *$P \in \mathcal{T}^{\gamma_*,D}(\mu_*)$ for suitable $\gamma_*, \mu_* \in (0,1]$, depending on the nonlinearity $g(x,u)$.*

Lemma in proven in Appendix A.

The goal of this section is to get a normal form for the Hamiltonian $H_2 + P$ of the beam equation, written in the form (1.8), in toroidal domains in the spaces $Y_\gamma^{\mathbb{Z}^d}$ which are complex neighbourhoods of the finite-dimensional real tori

$$T_\rho = \{\zeta = ({}^t(\xi_s,\bar{\xi}_s), s\in\mathbb{Z}^d) \mid |\xi_a|^2 = \nu\rho_a \text{ if } a\in\mathcal{A},\ \xi_s = 0 \text{ if } s\in\mathcal{L}\}\,, \tag{3.6}$$

invariant for the linear equation. Here $\nu > 0$ is small and $\rho = (\rho_a, a\in\mathcal{A})$ is a vector-parameter of the problem, belonging to the domain $\mathcal{D} = [c_*,1]^{\mathcal{A}}$. We arbitrarily enumerate the points of $\mathcal{A}$, i.e. write $\mathcal{A}$ as

$$\mathcal{A} = \{a_1,\dots,a_n\}\,, \tag{3.7}$$

and accordingly write $\mathcal{D}$ as

$$\mathcal{D} = [c_*,1]^n\,. \tag{3.8}$$

In the vicinity of a torus (3.6) in the space $Y_\gamma^{\mathbb{Z}^d}$ we pass from the complex variables $(\zeta_a, a\in\mathcal{A})$, to the corresponding complex action-angles (I_a,θ_a), using the relations

$$\xi_a = \sqrt{I}_a e^{i\theta_a}, \qquad \eta_a = \sqrt{I}_a e^{-i\theta_a}\,, \quad a\in\mathcal{A}\,. \tag{3.9}$$

Note that in the variables (I,θ,ξ,η), where $I = (I_a, a\in\mathcal{A})$, $\xi = (\xi_b, b\in\mathcal{L})$ etc, the involution (3.2) reads

$$(I,\theta,\xi,\eta) \to (\bar{I},\bar{\theta},\bar{\eta},\bar{\xi})\,. \tag{3.10}$$

So a vector (I,θ,ξ,η) is real if $I=\bar{I}, \theta=\bar{\theta}, \xi=\bar{\eta}$.

The complex toroidal vicinities of the tori T_ρ (see (3.6)) in the space $Y_\gamma^{\mathbb{Z}^d}$ will be of the form

$$\mathbf{T}_\rho = \mathbf{T}_\rho(\nu,\sigma,\mu,\gamma) = \{\zeta \mid |I-\nu\rho| < \nu c_*^2\mu^2, |\Im\theta| < \sigma, \|\zeta^{\mathcal{L}}\|_\gamma < \nu^{1/2}c_*\mu\}\,, \tag{3.11}$$

where $I = (I_a, a\in\mathcal{A})$, $\theta = (\theta_a, a\in\mathcal{A})$ and $\zeta^{\mathcal{L}} = \{\zeta_s, s\in\mathcal{L}\}$. Since $c_* \le \rho_j \le 1$ for each j, then

$$\mathbf{T}_\rho(\nu,\sigma,\mu,\gamma)\cap Y_\gamma^R \subset \{\zeta\in Y_\gamma^R \mid \mathrm{dist}_\gamma(\zeta,T_\rho) < C\sqrt{\nu}\mu\} \tag{3.12}$$

if $\mu \le 1$, where $C>0$ is an absolute constant.

We recall (see (1.17)) that we have split the set $\mathcal{L}$ to the union $\mathcal{L} = \mathcal{L}_f \cup \mathcal{L}_\infty$. Accordingly, we will write vectors $\zeta^{\mathcal{L}} = \{\zeta_s, s\in\mathcal{L}\}$ as $\zeta^{\mathcal{L}} = (\zeta_f,\zeta_\infty)$, where

$$\zeta_f = \{\zeta_s, s\in\mathcal{L}_f\}, \qquad \zeta_\infty = \{\zeta_s, s\in\mathcal{L}_\infty\}\,.$$

Proposition 3.3. *There exists a zero-measure Borel set $\mathcal{C} \subset [1,2]$ such that for any admissible set $\mathcal{A}$, any $c_* \in (0,1/2]$ and $m \notin \mathcal{C}$ we can find real numbers $\gamma_*, \nu_0 \in (0,1]$, where γ_* depends only on $g(\cdot)$ and ν_0 depends on $\mathcal{A}, m, c_*$ and $g(\cdot)$, such that*

(i) For $0 < \nu \le \nu_0$ and $\rho \in \mathcal{D} = [c_*, 1]^n$ there exist real holomorphic transformations

$$\Phi_\rho : \mathcal{O}^\gamma\Big(\frac{1}{2}, \frac{c_*}{2\sqrt{2}}\Big)^{\mathcal{L}} \to \mathbf{T}_\rho(\nu, 1, 1, \gamma), \qquad 0 \le \gamma \le \gamma_*,$$

which do not depend on γ in the sense that they coincide on the set $\mathcal{O}^{\gamma_*}(\frac{1}{2}, \frac{c_*}{2\sqrt{2}})^{\mathcal{L}}$, and are diffeomorphisms on their images, analytically depending on ρ and transforming the symplectic structure $-id\xi \wedge d\eta$ on $\mathbf{T}_\rho(\nu,1,1,\gamma_*)$ to the 2-form

$$-\nu \sum_{\ell \in \mathcal{A}} dr_\ell \wedge d\theta_\ell \; - i \; \nu \sum_{a \in \mathcal{L}} d\xi_a \wedge d\eta_a .$$

The change of variable Φ_ρ is close to the scaling by the factor $\nu^{1/2}$ on the $\mathcal{L}_\infty$-modes but not on the $(\mathcal{A} \cup \mathcal{L}_f)$-modes, where it is close to a certain affine transformation, depending on θ. For each γ, Φ_ρ as a function of ρ holomorphically extends to the complex domain

$$\mathcal{D}_{c_1} = \{\rho \in \mathbb{C}^{\mathcal{A}} \mid |\Im \rho_j| < c_1,\ c_* - c_1 < \Re \rho_j < 1 + c_1\ \forall j \in \mathcal{A}\}, \quad 0 < c_1 < c_* . \tag{3.13}$$

(ii) Φ_ρ puts the Hamiltonian $H_2 + P$ (see (1.9)*) to a normal form in the following sense:* [7]

$$\frac{1}{\nu}(H_2 + P) \circ \Phi_\rho = \Omega(\rho) \cdot r + \sum_{a \in \mathcal{L}_\infty} \Lambda_a(\rho) \xi_a \eta_a + \frac{\nu}{2} \langle K(\rho) \zeta_f, \zeta_f \rangle + f(r, \theta, \zeta; \rho) . \tag{3.14}$$

Here the vector Ω and the scalars $\Lambda_a, a \in \mathcal{L}_\infty$, are affine functions of ρ. They are defined by relations (3.44), (3.45)*, and after the natural extension to the complex domain $\mathcal{D}_{c_1}$ satisfy there the estimates*

$$|\Omega(\rho) - \omega| \le C_1 \nu, \;\; |\Lambda_a(\rho) - \lambda_a(\rho)| \le C_1 \nu \langle a \rangle^{-2} . \tag{3.15}$$

(iii) K is a symmetric real matrix, acting on vectors ζ_f. It is a quadratic polynomial of $\sqrt{\rho} = (\sqrt{\rho}_1, \dots, \sqrt{\rho}_n)$, defined by relation (3.47)*, and satisfies*

$$\|K(\rho)\| \le C_2 \qquad \forall\, \rho \in \mathcal{D}_{c_1}. \tag{3.16}$$

The matrix does not depend on the component g_0 of the nonlinearity g.

(iv) The remaining term f belongs to $\mathcal{T}^{\gamma,D}(\frac{1}{2}, \frac{c_*}{2\sqrt{2}}, \mathcal{D})^{\mathcal{L}}$, analytically extends to $\rho \in \mathcal{D}_{c_1}$ and for each $0 \le \gamma \le \gamma_*$ this analytic extension satisfies

$$[f]^{\gamma, D}_{\frac{1}{2}, \frac{c_*}{2\sqrt{2}}, \mathcal{D}_{c_1}} \le C_2 \nu , \qquad [f^T]^{\gamma, D}_{\frac{1}{2}, \frac{c_*}{2\sqrt{2}}, \mathcal{D}_{c_1}} \le C_2 \nu^{3/2} . \tag{3.17}$$

The constants C_1 and c_1 depend only on $\mathcal{A}$ and c_*, while C_2 also depend on m and the function $g(x,u)$.

Note that (3.15) and the Cauchy estimate imply that

$$|\partial_\rho \Lambda_a(\rho)| \le C_3 \nu \langle a \rangle^{-2} \;\text{ for } a \in \mathcal{L}_\infty ,\ \rho \in \mathcal{D} . \tag{3.18}$$

The rest of this section is devoted to the proof of Proposition 3.3.

[7] The factor ν^{-1} in the l.h.s. of (3.14) corresponds to ν in the transformed symplectic structure in item (i). So the Hamiltonian of the transformed equations with respect to the symplectic structure $-dr \wedge d\theta - i\, d\xi \wedge d\eta$ is given by the r.h.s. of (3.14).

3.2. **Resonances and the Birkhoff procedure.** Let us write the quartic part $H^4 = H_2 + P_4$ of the Hamiltonian H (see (1.9), (1.11)) in the complex variables $\zeta = \zeta^{\mathbb{Z}^d} = \{{}^t(\xi_s, \eta_s), s \in \mathbb{Z}^d\}$:

$$\begin{aligned} H_2 =& \sum_{s\in\mathbb{Z}^d} \lambda_s \xi_s \eta_s, \\ P_4 =& (2\pi)^{-d} \sum_{(i,j,k,\ell)\in\mathcal{J}} \frac{(\xi_i+\eta_{-i})(\xi_j+\eta_{-j})(\xi_k+\eta_{-k})(\xi_\ell+\eta_{-\ell})}{4\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}, \end{aligned}$$

where $\mathcal{J}$ denotes the zero momentum set:

$$\mathcal{J} := \{(i,j,k,\ell) \subset \mathbb{Z}^d \mid i+j+k+\ell = 0\}.$$

We decompose $P_4 = P_{4,0} + P_{4,1} + P_{4,2}$ according to

$$\begin{aligned} P_{4,0} =& \frac{1}{4}(2\pi)^{-d} \sum_{(i,j,k,\ell)\in\mathcal{J}} \frac{\xi_i\xi_j\xi_k\xi_\ell + \eta_i\eta_j\eta_k\eta_\ell}{\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}, \\ P_{4,1} =& (2\pi)^{-d} \sum_{(i,j,k,-\ell)\in\mathcal{J}} \frac{\xi_i\xi_j\xi_k\eta_\ell + \eta_i\eta_j\eta_k\xi_\ell}{\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}, \\ P_{4,2} =& \frac{3}{2}(2\pi)^{-d} \sum_{(i,j,-k,-\ell)\in\mathcal{J}} \frac{\xi_i\xi_j\eta_k\eta_\ell}{\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}, \end{aligned}$$

and denote by R_5 the remainder term of the the nonlinearity P. I.e.

$$P = P_4 + R_5. \tag{3.19}$$

Finally we define

$$\mathcal{J}_2 = \{(i,j,k,\ell) \subset \mathbb{Z}^d \mid (i,j,-k,-\ell) \in \mathcal{J},\ \sharp\{i,j,k,\ell\} \cap \mathcal{A} \geq 2\}.$$

For later use we note that, by Proposition 2.9,

Lemma 3.4. *If $m \notin \mathcal{C}$, then there exists $\kappa(m) > 0$ such that for all $(i,j,k,\ell) \in \mathcal{J}_2$*

$$\begin{aligned} &|\lambda_i + \lambda_j + \lambda_k - \lambda_\ell| \geq \kappa(m); \\ &|\lambda_i + \lambda_j - \lambda_k - \lambda_\ell| \geq \kappa(m), \quad \textit{except if } \{|i|,|j|\} = \{|k|,|\ell|\}. \end{aligned}$$

For $\gamma \geq 0$ we consider the phase space $Y_\gamma = Y_\gamma^{\mathbb{Z}^d}$, defined as in Section 3.1, and endowed it with the symplectic structure $-i\sum d\xi_k \wedge d\eta_k$. Since $d^* > d/2$, then the spaces Y_γ are algebras with respect to the convolution, see Lemma 1.1 in [16]. This implies the following result, where $\langle\cdot,\cdot\rangle$ stands for the complex-bilinear paring of $\mathbb{C}^{2r}$ with itself:

Lemma 3.5. *Let $\gamma \geq 0$, $r \in \mathbb{N}$ and P^r be a real homogeneous polynomial on Y_γ of degree r,*

$$P^r(\zeta) = \sum_{(j_1,\dots j_r)\in(\mathcal{L})^r} \langle a_{j_1,\dots,j_r}, \zeta_{j_1}\otimes\cdots\otimes\zeta_{j_r}\rangle,$$

where $a_{j_1,\dots,j_r} \in \mathbb{C}^2\otimes\cdots\otimes\mathbb{C}^2$ (r times), $|a_{j_1,\dots,j_r}| \leq M$, and $a_{j_1,\dots,j_r} = 0$ unless $j_1 + \cdots + j_r = 0$. Then the gradient-map $\nabla P^r(\zeta)$ satisfies $\|\nabla P^r(\zeta)\|_\gamma \leq MC^{r-1}\|\zeta\|_\gamma^{r-1}$. So the flow-maps $\Phi^t_{P^r}$, $|t| \leq 1$, of the hamiltonian vector-field $X_{P^r} = iJ\nabla P^r$ are well defined real holomorphic mappings on a ball $B_\gamma(\delta) = \{\|\zeta\|_\gamma < \delta\}$, $\delta = \delta(M) > 0$, and satisfy there

$$\|\Phi^t_{P^r}(\zeta) - \zeta\|_\gamma \leq C_1\|\zeta\|_\gamma^{r-1}, \quad C_1 = C_1(M).$$

Corollary 3.6. *Consider the polynomial $Q^r(\zeta) = P^r(D^-(\zeta))$, where D^- is the diagonal matrix $diag\{|\lambda_s|^{-1/2}I\}$. Then the Hessian-map $\nabla^2_\zeta Q^r \in \mathcal{M}^D_\gamma$ and $|Q^r|^D_\gamma \leq MC^{r-2}\|\zeta\|^{r-2}_\gamma$ for any $\gamma \geq 0$. In particular $Q \in \mathcal{T}^{\gamma,D}(\mu)$ for any $0 < \mu \leq 1$ (see* (3.5)*).*

Note that the corollary applies to the monomials, forming P_4 (e.g. to P_4).

Proposition 3.7. *For $m \notin \mathcal{C}$ there exists a real holomorphic and symplectic change of variable τ in a neighbourhood of the origin in Y_γ that puts the Hamiltonian $H = H_2 + P$ into its partial Birkhoff normal form up to order five in the sense that it removes from P_4 all non-resonant terms, apart from those who are cubic or quartic in directions of $\mathcal{L}$. More precisely, for $0 \leq \gamma \leq \gamma_*$, where γ_* is as in Lemma 3.2, and for a suitable $\delta(m) \leq \mu_*$ (depending on m and $g(x,u)$), the mapping τ satisfies*

$$\|\tau^{\pm 1}(\zeta) - \zeta\|_\gamma \leq C(m)\|\zeta\|^3_\gamma \qquad \forall\, \zeta \in B_\gamma(\delta(m))\,. \tag{3.20}$$

It transforms the Hamiltonian $H_2 + P = H_2 + P_4 + R_5$ as follows:

$$(H_2 + P)\circ\tau = H_2 + Z_4 + Q^3_4 + R^0_6 + R_5\circ\tau\,, \tag{3.21}$$

where

$$Z_4 = \frac{3}{2}(2\pi)^{-d} \sum_{\substack{(i,j,k,\ell)\in\mathcal{J}_2\\ \{|i|,|j|\}=\{|k|,|\ell|\}}} \frac{\xi_i\xi_j\eta_k\eta_\ell}{\lambda_i\lambda_j},$$

and $Q^3_4 = Q_{4,1} + Q_{4,2}$ with[8]

$$\begin{aligned} Q_{4,1} &= (2\pi)^{-d} \sum_{(i,j,-k,\ell)\notin\mathcal{J}_2} \frac{\xi_i\xi_j\xi_k\eta_\ell + \eta_i\eta_j\eta_k\xi_\ell}{\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}},\\ Q_{4,2} &= \frac{3}{2}(2\pi)^{-d} \sum_{(i,j,k,\ell)\notin\mathcal{J}_2} \frac{\xi_i\xi_j\eta_k\eta_\ell}{\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}\,. \end{aligned}$$

The functions $Z_4, Q^3_4, R^0_6, R_5\circ\tau$ are real holomorphic on $B_\gamma(\delta(m))$. Besides R^0_6 and $R_5\circ\tau$ are, respectively, functions of order 6 and 5 at the origin. For any $0 < \mu \leq \delta(m)$ the functions Z_4, Q^3_4, R^0_6 and $R_5\circ\tau$ belong to $\mathcal{T}^{\gamma,D}(\mu)$ (see (3.5)*), and*

$$\big[Z_4\big]^{\gamma,D}_\mu + \big[Q^3_4\big]^{\gamma,D}_\mu \leq C\mu^4\,, \tag{3.22}$$

$$\big[R^0_6\big]^{\gamma,D}_\mu \leq C\mu^6\,, \tag{3.23}$$

$$\big[R_5\circ\tau\big]^{\gamma,D}_\mu \leq C\mu^5\,, \tag{3.24}$$

where C depends on $\mathcal{A}$, m and g.

[8]The upper index 3 signifies that Q^3_4 is at least cubic in the transversal directions $\{\zeta_a, a\in\mathcal{L}\}$.

Proof. We use the classical Birkhoff normal form procedure. We construct the transformation τ as the time one flow $\Phi^1_{\chi_4}$ of a Hamiltonian χ_4, given by

$$\begin{aligned}\chi_4 = &-\frac{i}{4}(2\pi)^{-d}\sum_{(i,j,k,\ell)\in\mathcal{J}}\frac{\xi_i\xi_j\xi_k\xi_\ell-\eta_i\eta_j\eta_k\eta_\ell}{(\lambda_i+\lambda_j+\lambda_k+\lambda_\ell)\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}\\ &-i(2\pi)^{-d}\sum_{(i,j,-k,\ell)\in\mathcal{J}_2}\frac{\xi_i\xi_j\xi_k\eta_\ell-\eta_i\eta_j\eta_k\xi_\ell}{(\lambda_i+\lambda_j+\lambda_k-\lambda_\ell)\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}\\ &-\frac{3i}{2}(2\pi)^{-d}\sum_{\substack{(i,j,k,\ell)\in\mathcal{J}_2\\ \{|i|,|j|\}\neq\{|k|,|\ell|\}}}\frac{\xi_i\xi_j\eta_k\eta_\ell}{(\lambda_i+\lambda_j-\lambda_k-\lambda_\ell)\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}\end{aligned} \tag{3.25}$$

By Lemma 3.4 and Lemma 3.5 for $m\notin\mathcal{C}$ the vector-field X_{χ_4} is real holomorphic in Y_γ and of order three at the origin. Hence $\tau=\Phi^1_{\chi_4}$ is a real holomorphic and symplectic change of coordinates, defined in $B_\gamma(\delta(m))$, a neighbourhood of the origin in Y_γ. By Lemma 3.5 it satisfies (3.20).

Since the Poisson bracket, corresponding to the symplectic form $-id\xi\wedge d\eta$ is $\{F,G\}=i\langle\nabla_\eta F,\nabla_\xi G\rangle-i\langle\nabla_\xi F,\nabla_\eta G\rangle$, and since $\nabla_{\eta_s}H_2=\lambda_s\xi_s$, $\nabla_{\xi_s}H_2=\lambda_s\eta_s$, then we calculate

$$\begin{aligned}\{H_2,\chi_4\}= &-\frac{1}{4}(2\pi)^{-d}\sum_{(i,j,k,\ell)\in\mathcal{J}}\frac{\xi_i\xi_j\xi_k\xi_\ell+\eta_i\eta_j\eta_k\eta_\ell}{\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}\\ &-(2\pi)^{-d}\sum_{(i,j,-k,\ell)\in\mathcal{J}_2}\frac{\xi_i\xi_j\xi_k\eta_\ell+\eta_i\eta_j\eta_k\xi_\ell}{\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}\\ &-\frac{3}{2}(2\pi)^{-d}\sum_{\substack{(i,j,k,\ell)\in\mathcal{J}_2\\ \{|i|,|j|\}\neq\{|k|,|\ell|\}}}\frac{\xi_i\xi_j\eta_k\eta_\ell}{\sqrt{\lambda_i\lambda_j\lambda_k\lambda_\ell}}.\end{aligned}$$

Therefore

$$\begin{aligned}(H_2+P_4)\circ\tau=&H_2+P_4-\{H_2,\chi_4\}-\{P_4,\chi_4\}\\ &+\int_0^1(1-t)\{\{H_2+P_4,\chi_4\},\chi_4\}\circ\Phi^t_{\chi_4}\mathrm{d}t\\ =&H_2+Z_4+Q_4^3+R_6^0\end{aligned}$$

with Z_4 and Q_4^3 as in the statement of the proposition and

$$R_6^0=\{P_4,\chi_4\}+\int_0^1(1-t)\{\{H_2+P_4,\chi_4\},\chi_4\}\circ\Phi^t_{\chi_4}\mathrm{d}t.$$

The reality of the functions Z_4 and Q_4^3 follow from the explicit formulas for them, while the inclusion of these functions to $\mathcal{T}^{\gamma,D}(\mu)$ for any $0<\mu\le 1$ and the estimate (3.22) hold by Corollary 3.6. Concerning R_6^0, by construction this is a holomorphic function of order ≥ 6 at the origin. Its reality follows from the equality (3.21), where all other functions are real. The inclusion $R_6^0\in\mathcal{T}^{\gamma,D}(\mu)$ for any $0<\mu\le\delta(m)$ and the estimate (3.23) follow from the following three facts:

(i) $\{H_2+P_4,\chi_4\}=Z_4+Q_4^3$ and χ_4 belong to $\mathcal{T}^{\gamma,D}(1)$ by Corollary 3.6.
(ii) $\{\mathcal{T}^{\gamma,D}(1),\mathcal{T}^{\gamma,D}(1)\}\in\mathcal{T}^{\gamma,D}(\frac{1}{2})$ (see Proposition 2.6 in [15]).
(iii) $\mathcal{T}^{\gamma,D}(\frac{1}{2})\circ\Phi^t_{\chi_4}\in\mathcal{T}^{\gamma,D}(\frac{1}{2}\delta(m))$.

In [15], Proposition 2.7, and [25], Lemma 10.7, the assertion (iii) is proven for a special class of Hamiltonians χ_4, but the proof easily generalises to Hamiltonian χ_4 as above.

Finally, since by Lemma 3.2 the function R_5 belongs to $\mathcal{T}^{\gamma,D}(\mu_*)$, then in view of (iii) $R_5 \circ \tau \in \mathcal{T}^{\gamma,D}(\frac{1}{2}\delta(m))$. Re-denoting $\frac{1}{2}\delta(m)$ to $\delta(m)$ we get (3.22)-(3.24). $\square$

Due to (3.20), if $\zeta \in \mathbf{T}_\rho(\nu, 1/2, 1/2, \gamma)$, $0 \le \gamma \le \gamma_*$, where $\nu \le C^{-1}\delta(m)^2$ and C is an absolute constant (see (3.11)), then $\|\tau^{\pm 1}(\zeta) - \zeta\|_\gamma \le C'(m)\nu^{\frac{3}{2}}$. Therefore

$$\tau^{\pm 1}(\mathbf{T}_\rho(\nu, 1/2, 1/2, \gamma)) \subset \mathbf{T}_\rho(\nu, 1, 1, \gamma)\,, \tag{3.26}$$

provided that $\nu \le C^{-1}\delta(m)^2$ and $\rho \in \mathcal{D}_{c_1}$, where $c_1 = c_1(\mathcal{A}, m, g(\cdot), c_*)$ is sufficiently small.

3.3. **Normal form, corresponding to admissible sets $\mathcal{A}$.** Everywhere in Section 3.3–4.5 the set $\mathcal{A}$ is assumed to be admissible in the sense of Definition 1.1.

The Hamiltonian Z_4 contains the integrable part formed by monomials of the form $\xi_i\xi_j\eta_i\eta_j = I_iI_j$ that only depend on the actions $I_n = \xi_n\eta_n$, $n \in \mathbb{Z}^d$. Denote it Z_4^+ and denote the rest Z_4^-. It is not hard to see that

$$Z_4^+ = \frac{3}{2}(2\pi)^{-d} \sum_{\ell \in \mathcal{A},\ k \in \mathbb{Z}^d} (4 - 3\delta_{\ell,k}) \frac{I_\ell I_k}{\lambda_\ell \lambda_k}. \tag{3.27}$$

To calculate Z_4^-, we decompose it according to the number of indices in $\mathcal{A}$: a monomial $\xi_i\xi_j\eta_k\eta_\ell$ is in Z_4^{-r} $(r = 0, 1, 2, 3, 4)$ if $(i, j, -k, -\ell) \in \mathcal{J}$ and $\sharp\{i, j, k, \ell\} \cap \mathcal{A} = r$. We note that, by construction, $Z_4^{-0} = Z_4^{-1} = \emptyset$.

Since $\mathcal{A}$ is admissible, then in view of Lemma 3.4 for $m \notin \mathcal{C}$ the set Z_4^{-4} is empty. The set Z_4^{-3} is empty as well:

Lemma 3.8. *If $m \notin \mathcal{C}$, then $Z_4^{-3} = \emptyset$.*

Proof. Consider any term $\xi_i\xi_j\eta_k\eta_\ell \in Z_4^{-3}$, i.e. $\{i, j, k, \ell\} \cap \mathcal{A} = 3$. Without lost of generality we can assume that $i, j, k \in \mathcal{A}$ and $\ell \in \mathcal{L}$. Furthermore we know that $i + j - k - \ell = 0$ and $\{|i|, |j|\} = \{|k|, |\ell|\}$. In particular we must have $|i| = |k|$ or $|j| = |k|$ and thus, since $\mathcal{A}$ is admissible, $i = k$ or $j = k$. Let for example, $i = k$. Then $|j| = |\ell|$. Since $i + j = k + \ell$ we conclude that $\ell = j$ which contradicts our hypotheses. $\square$

Recall that the finite set $\mathcal{L}_f \subset \mathcal{L}$ was defined in (1.17). The mapping

$$\ell : \mathcal{L}_f \to \mathcal{A}, \quad a \mapsto \ell(a) \in \mathcal{A} \text{ if } |a| = |\ell(a)|, \tag{3.28}$$

is well defined since the set $\mathcal{A}$ is admissible. Now we define two subsets of $\mathcal{L}_f \times \mathcal{L}_f$:

$$(\mathcal{L}_f \times \mathcal{L}_f)_+ = \{(a, b) \in \mathcal{L}_f \times \mathcal{L}_f \mid \ell(a) + \ell(b) = a + b\} \tag{3.29}$$

$$(\mathcal{L}_f \times \mathcal{L}_f)_- = \{(a, b) \in \mathcal{L}_f \times \mathcal{L}_f \mid a \ne b \text{ and } \ell(a) - \ell(b) = a - b\}. \tag{3.30}$$

Example 3.9. If $d = 1$, then in view of (1.18) $\ell(a) = -a$ and the sets $(\mathcal{L}_f \times \mathcal{L}_f)_\pm$ are empty. If d is any, but $\mathcal{A}$ is a one-point set $\mathcal{A} = \{b\}$, then $\mathcal{L}_f$ is the punched discrete sphere $\{a \in \mathbb{Z}^d \mid |a| = |b|, a \ne b\}$, $\ell(a) = b$ for each a, and the sets $(\mathcal{L}_f \times \mathcal{L}_f)_\pm$ again are empty. If $d \ge 2$ and $|\mathcal{A}| \ge 2$, then in general the sets $(\mathcal{L}_f \times \mathcal{L}_f)_\pm$ are non-trivial. See in Appendix B.

Obviously

$$(\mathcal{L}_f \times \mathcal{L}_f)_+ \cap (\mathcal{L}_f \times \mathcal{L}_f)_- = \emptyset\,. \tag{3.31}$$

For further reference we note that

Lemma 3.10. *If* $(a,b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+ \cup (\mathcal{L}_f \times \mathcal{L}_f)_-$ *then* $|a| \neq |b|$.

Proof. If $(a,b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+$ and $|a| = |b|$ then $\ell(a) = \ell(b)$ and we have

$$|a+b| = |2\ell(a)| = 2|a| = |a| + |b|$$

which is impossible since b is not proportional to a. If $(a,b) \in (\mathcal{L}_f \times \mathcal{L}_f)_-$ and $|a| = |b|$ then $\ell(a) = \ell(b)$ and we get $a-b = 0$ which is impossible in $(\mathcal{L}_f \times \mathcal{L}_f)_-$. $\square$

According to the decomposition $\mathcal{L} = \mathcal{L}_f \cup \mathcal{L}_\infty$, the space Y_γ, defined in (3.1), decomposes in the direct sum

$$Y_\gamma = Y_\gamma^f \oplus Y_\gamma^\infty, \quad Y_\gamma^f = \operatorname{span}\{\zeta_s, s \in \mathcal{L}_f\}\,, \;\; Y_\gamma^\infty = \overline{\operatorname{span}}\{\zeta_s, s \in \mathcal{L}_\infty\}\,. \tag{3.32}$$

Lemma 3.11. *For* $m \notin \mathcal{C}$ *the part* Z_4^{-2} *of the Hamiltonian* Z_4 *equals*

$$\begin{aligned} &3(2\pi)^{-d}\Big(\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_+} \frac{\xi_{\ell(a)}\xi_{\ell(b)}\eta_a\eta_b + \eta_{\ell(a)}\eta_{\ell(b)}\xi_a\xi_b}{\lambda_a\lambda_b} \\ &\qquad + 2\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_-} \frac{\xi_a\xi_{\ell(b)}\eta_{\ell(a)}\eta_b}{\lambda_a\lambda_b}\Big)\,. \end{aligned} \tag{3.33}$$

Proof. Let $\xi_i\xi_j\eta_k\eta_\ell$ be a monomial in Z_4^{-2}. We know that $(i,j,-k,-\ell) \in \mathcal{J}$ and $\{|i|,|j|\} = \{|k|,|\ell|\}$. If $i,j \in \mathcal{A}$ or $k,\ell \in \mathcal{A}$ then we obtain the finitely many monomials as in the first sum in (3.33). Now we assume that $i,\ell \in \mathcal{A}$ and $j,k \in \mathcal{L}$. Then we have that, either $|i| = |k|$ and $|j| = |\ell|$ which leads to finitely many monomials as in the second sum in (3.33). Or $i = \ell$ and $|j| = |k|$. In this last case, the zero momentum condition implies that $j = k$ which is not possible in Z_4^-. $\square$

3.4. **Eliminating the non integrable terms.** For $\ell \in \mathcal{A}$ we introduce the variables $(I_a, \theta_a, \zeta^{\mathcal{L}})$ as in (3.11). Now the symplectic structure $-id\xi \wedge d\eta$ reads

$$-\sum_{a\in\mathcal{A}} dI_a \wedge d\theta_a - id\xi^{\mathcal{L}} \wedge d\eta^{\mathcal{L}}\,. \tag{3.34}$$

In view of (3.27), (3.21) and Lemma 3.11, for $m \notin \mathcal{C}$ the transformed Hamiltonian may be written as (recall that $\omega = (\lambda_a, a \in \mathcal{A})$)

$$\begin{aligned} (H_2 + P)\circ\tau =& \omega\cdot I + \sum_{s\in\mathcal{L}} \lambda_s\xi_s\eta_s + \frac{3}{2}(2\pi)^{-d}\sum_{\ell\in\mathcal{A},\; k\in\mathbb{Z}^d} (4 - 3\delta_{\ell,k})\frac{I_\ell\xi_k\eta_k}{\lambda_\ell\lambda_k} \\ & + 3(2\pi)^{-d}\Big(\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_+} \frac{\xi_{\ell(a)}\xi_{\ell(b)}\eta_a\eta_b + \eta_{\ell(a)}\eta_{\ell(b)}\xi_a\xi_b}{\lambda_a\lambda_b} \\ & + 2\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_-} \frac{\xi_a\xi_{\ell(b)}\eta_{\ell(a)}\eta_b}{\lambda_a\lambda_b}\Big) \\ & + Q_4^3 + R_5^0\,, \qquad R_5^0 = R_5\circ\tau + R_6^0\,. \end{aligned}$$

The first line contains the integrable terms. The second and third lines contain the lower-order non integrable terms, depending on the angles θ; there are finitely many of them. The last line contains the remaining high order terms, where Q_4^3 is of total order (at least) 4 and of order 3 in the normal directions ζ, while R_5^0 is of total order at least 5. The latter is the sum of R_6^0 which comes from the Birkhoff normal form procedure (and is of order 6) and $R_5 \circ \tau$ which comes from the term of order 5 in the nonlinearity (1.2). Here I is regarded as a variable of order 2, while θ has zero order. The fourth line should be regarded as a perturbation.

To deal with the non integrable terms in the second and third lines, following the works on the finite-dimensional reducibility (see [13]), we introduce a change of variables

$$\Psi : (\tilde{I}, \tilde{\theta}, \tilde{\xi}, \tilde{\eta}) \mapsto (I, \theta, \xi, \eta)\,,$$

symplectic with respect to (3.34), but such that its differential at the origin is not close to the identity. It is defined by the following relations:

$$\begin{aligned} &I_\ell = \tilde{I}_\ell - \sum_{|a|=|\ell|,\ a\neq \ell} \tilde{\xi}_a \tilde{\eta}_a, \quad \theta_\ell = \tilde{\theta}_\ell \quad \ell \in \mathcal{A}\,; \\ &\xi_a = \tilde{\xi}_a e^{i\tilde{\theta}_{\ell(a)}}, \quad \eta_a = \tilde{\eta}_a e^{-i\tilde{\theta}_{\ell(a)}} \quad a \in \mathcal{L}_f\,; \qquad \xi_a = \tilde{\xi}_a, \quad \eta_a = \tilde{\eta}_a \quad a \in \mathcal{L}_\infty. \end{aligned}$$

For any $(\tilde{I}, \tilde{\theta}, \tilde{\zeta}) \in \mathbf{T}_\rho(\nu, \sigma, \mu, \gamma)$ denote by $y = \{y_l, l \in \mathcal{A}\}$ the vector, whose l-th component equals $y_l = \sum_{|a|=|l|\,, a\neq l} \tilde{\xi}_a \tilde{\eta}_a$. Then

$$|I - \tfrac{1}{2}\nu\rho^2| \leq |\tilde{I} - \tfrac{1}{2}\nu\rho^2| + |y| \leq c_*^2 \nu\mu^2 + \sum_{a\in\mathcal{L}_f} |\tilde{\xi}_a \tilde{\eta}_a| \leq 2c_*^2\nu\mu^2\,.$$

This implies that

$$\Psi^{\pm 1}\big(\mathbf{T}_\rho\big(\nu, \frac{1}{2}, \frac{1}{2\sqrt{2}}, \gamma\big)\big) \subset \mathbf{T}_\rho\big(\nu, \frac{1}{2}, \frac{1}{2}, \gamma\big)\,. \tag{3.35}$$

We abbreviate $\mathbf{T}_\rho\big(\nu, \frac{1}{2}, \frac{1}{2\sqrt{2}}, \gamma\big) =: \mathbf{T}_\rho$.
We note that $\Psi(T_I^n) = T_I^n$ and although Ψ is not close to the identity in general, it is close to the identity in variable (I, θ) in a neighbourhood of the T_I^n. Namely, denoting $\Psi(\tilde{I}, \tilde{\theta}, \tilde{\zeta}^{\mathcal{L}}) = (I, \theta, \zeta^{\mathcal{L}})$, we have

$$|\tilde{I}_a - I_a| \leq \|(\tilde{\zeta}^{\mathcal{L}})\|^2\,,\ a \in \mathcal{A},\ \theta = \tilde{\theta} \text{ and } \|\zeta^{\mathcal{L}}\|_\gamma = \|\tilde{\zeta}^{\mathcal{L}}\|_\gamma\,. \tag{3.36}$$

On the other hand, if $(\tilde{\xi}, \tilde{\eta}) \in \mathbf{T}_\rho$, then for $l \in \mathcal{A}$

$$\xi_l = \sqrt{I_l}\, e^{i\theta_l} = \sqrt{\tilde{I}_l}\, e^{i\tilde{\theta}_l} + O(\nu^{-1/2})\, O(|\zeta^{\mathcal{L}}|^2).$$

Therefore, dropping the tildes, we write the restriction to $\mathbf{T}_\rho$ of the transformed Hamiltonian as

$$\begin{aligned} H_1 :=& H\circ\tau\circ\Psi = \omega\cdot I + \sum_{a\in\mathcal{L}_\infty}\lambda_a\xi_a\eta_a \\ &+ 6(2\pi)^{-d}\sum_{\ell\in\mathcal{A},\ k\in\mathcal{L}}\frac{1}{\lambda_\ell\lambda_k}(I_\ell - \sum_{\substack{|a|=|\ell|\\ a\in\mathcal{L}_f}}\xi_a\eta_a)\xi_k\eta_k \\ &+ \frac{3}{2}(2\pi)^{-d}\sum_{\ell,k\in\mathcal{A}}\frac{4-3\delta_{\ell,k}}{\lambda_\ell\lambda_k}(I_\ell - \sum_{\substack{|a|=|\ell|\\ a\in\mathcal{L}_f}}\xi_a\eta_a)(I_k - \sum_{\substack{|a|=|k|\\ a\in\mathcal{L}_f}}\xi_a\eta_a) \\ &+ 3(2\pi)^{-d}\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_+}\frac{\sqrt{I_{\ell(a)}I_{\ell(b)}}}{\lambda_a\lambda_b}(\eta_a\eta_b+\xi_a\xi_b) \\ &+ 6(2\pi)^{-d}\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_-}\frac{\sqrt{I_{\ell(a)}I_{\ell(b)}}}{\lambda_a\lambda_b}\xi_a\eta_b + Q_4^{3'} + R_5^{0'} + \nu^{-1/2}R_5^{4'}\,. \end{aligned}$$

Here $Q_4^{3'}$ and $R_5^{0'}$ are the function Q_4^3 and R_5^0, transformed by Ψ (so the former satisfy the same estimates as the latter), while $R_5^{4'}$ is a function of forth order in the normal variables. Or, after a simplification:

$$\begin{aligned} H_1 =& \omega\cdot I + \sum_{a\in\mathcal{L}_\infty}\lambda_a\xi_a\eta_a + \frac{3}{2}(2\pi)^{-d}\sum_{\ell,k\in\mathcal{A}}\frac{4-3\delta_{\ell,k}}{\lambda_\ell\lambda_k}I_\ell I_k \\ &+ 3(2\pi)^{-d}\Big(2\sum_{\ell\in\mathcal{A},\ a\in\mathcal{L}_\infty}\frac{1}{\lambda_\ell\lambda_a}I_\ell\xi_a\eta_a - \sum_{\ell\in\mathcal{A},\ a\in\mathcal{L}_f}\frac{(2-3\delta_{\ell,|a|})}{\lambda_\ell\lambda_a}I_\ell\xi_a\eta_a\Big) \\ &+ 3(2\pi)^{-d}\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_+}\frac{\sqrt{I_{\ell(a)}I_{\ell(b)}}}{\lambda_a\lambda_b}(\eta_a\eta_b+\xi_a\xi_b) \\ &+ 6(2\pi)^{-d}\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_-}\frac{\sqrt{I_{\ell(a)}I_{\ell(b)}}}{\lambda_a\lambda_b}\xi_a\eta_b + Q_4^{3'} + R_5^{0'} + \nu^{-1/2}R_5^{4'}\,. \end{aligned} \tag{3.37}$$

We see that the transformation Ψ removed from $H\circ\tau$ the non-integrable lower-order terms on the price of introducing "half-integrable" terms which do not depend on the angles θ, but depend on the actions I and quadratically depend on finitely many variables ξ_a, η_a with $a\in\mathcal{L}_f$.

The Hamiltonian $H\circ\tau\circ\Psi$ should be regarded as a function of the variables $(I,\theta,\zeta^{\mathcal{L}})$. Abusing notation, below we drop the upper-index $\mathcal{L}$ and write $\zeta^{\mathcal{L}} = {}^t(\xi^{\mathcal{L}},\eta^{\mathcal{L}})$ as $\zeta = {}^t(\xi,\eta)$.

3.5. **Rescaling the variables and defining the transformation Φ.** Our aim is to study the Hamiltonian H_1 on the domains $\mathbf{T}_\rho = \mathbf{T}_\rho(\nu,\frac12,\frac{1}{2\sqrt2},\gamma)$, $0\le\gamma\le\gamma_*$ (see (3.35)). To do this we re-parametrise points of $\mathbf{T}_\rho$ by mean of the change of variables $(I,\theta,\xi,\eta) = \chi_\rho(\tilde r,\tilde\theta,\tilde\xi,\tilde\eta)$, where

$$I = \nu\rho + \nu\tilde r,\quad \theta = \tilde\theta,\quad \xi = \sqrt\nu\,\tilde\xi,\quad \eta = \sqrt\nu\,\tilde\eta\,.$$

Clearly,

$$\chi_\rho : \mathcal{O}^\gamma(\frac12,\frac{c_*}{2\sqrt2}) \to \mathbf{T}_\rho\,,$$

and in the new variables the symplectic structure reads

$$-\nu\sum_{\ell\in\mathcal{A}}\tilde{d}r_\ell\wedge d\tilde{\theta}_\ell\ -i\ \nu\sum_{a\in\mathcal{L}}d\tilde{\xi}_a\wedge d\tilde{\eta}_a.$$

Denoting

$$\Phi=\Phi_\rho=\tau\circ\Psi\circ\chi_\rho,$$

we see that this transformation is real holomorphic in $\rho\in\mathcal{D}_{c_1}$ for a suitable $c_1>0$. It satisfies all assertions of the item (i) of Proposition 3.3.
We also notice for later use that,using using (3.36) and (3.20), for $\mathfrak{z}=(r,\theta,z^{\mathcal{L}})\in\mathcal{O}^\gamma(\frac{1}{2},\frac{c_*}{2\sqrt{2}})$, $\zeta=\Phi_\rho(\mathfrak{z})=(\zeta^{\mathcal{A}},\zeta^{\mathcal{L}})$ satisfies or ν small enough

$$\|\zeta^{\mathcal{L}}\|_\gamma\le\nu^{1/2}\|z^{\mathcal{L}}\|_\gamma(1+C\left\|\nu^{1/2}\mathfrak{z}\right\|_\gamma^2)\le 2\nu^{1/2}\|z^{\mathcal{L}}\|_\gamma\,, \tag{3.38}$$

and

$$\left\|\zeta^{\mathcal{A}}-\nu^{1/2}\sqrt{\rho+r}e^{i\theta}\right\|\le(\sqrt{n}\nu^{1/2}\|z^{\mathcal{L}}\|_0)(1+C\left\|\nu^{1/2}\mathfrak{z}\right\|_0^2)\le 2\sqrt{n}\nu^{1/2}\|z^{\mathcal{L}}\|_0$$

thus

$$\left\|\zeta^{\mathcal{A}}-\nu^{1/2}\sqrt{\rho}e^{i\theta}\right\|\le(2\sqrt{n}\nu^{1/2}\|z^{\mathcal{L}}\|_0+\nu^{1/2}\frac{|r|}{2c_*})\le\frac{2}{c_*}\sqrt{n}\nu^{1/2}(\|z^{\mathcal{L}}\|_0+|r|)\,. \tag{3.39}$$

We have, dropping the tilde and forgetting the irrelevant constant $\nu(\omega\cdot\rho)$,

$$\begin{aligned}H\circ\Phi=\nu\Big[&\omega\cdot r+\sum_{a\in\mathcal{L}_\infty}\lambda_a\xi_a\eta_a+(2\pi)^{-d}\nu\Big(\frac{3}{2}\sum_{\ell,k\in\mathcal{A}}\frac{4-3\delta_{\ell,k}}{\lambda_\ell\lambda_k}\rho_\ell r_k\\&+6\sum_{\ell\in\mathcal{A},\ a\in\mathcal{L}_\infty}\frac{1}{\lambda_\ell\lambda_a}\rho_\ell\xi_a\eta_a-3\sum_{\ell\in\mathcal{A},\ a\in\mathcal{L}_f}\frac{(2-3\delta_{\ell,|a|})}{\lambda_\ell\lambda_a}\rho_\ell\xi_a\eta_a\\&+3\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_+}\frac{\sqrt{\rho_{\ell(a)}}\sqrt{\rho_{\ell(b)}}}{\lambda_a\lambda_b}(\eta_a\eta_b+\xi_a\xi_b)\\&+6\sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_-}\frac{\sqrt{\rho_{\ell(a)}}\sqrt{\rho_{\ell(b)}}}{\lambda_a\lambda_b}\xi_a\eta_b\Big)\Big]\\&+\Big((Q_4^{3'}+R_5^{0'}+\nu^{-1/2}R_5^{4'})(I,\theta,\sqrt{\nu}\zeta)\Big)\ |_{I=\nu\rho+\nu r}\ .\end{aligned} \tag{3.40}$$

So,

$$\nu^{-1}H\circ\Phi=h+f\,, \tag{3.41}$$

where $h\equiv h(I,\xi,\eta;\rho,\nu)$ is the quadratic part of the Hamiltonian, independent from the angle θ, and f is the perturbation, given by the last line in (3.40):

$$f=\nu^{-1}\Big((Q_4^{3'}+R_5^{0'}+\nu^{-1/2}R_5^{4'})(I,\theta,\nu^{1/2}\zeta)\Big)\ |_{I=\nu\rho+\nu r}\ . \tag{3.42}$$

We have

$$h=\Omega\cdot r+\sum_{a\in\mathcal{L}_\infty}\Lambda_a\xi_a\eta_a+\nu\langle K(\rho)\zeta_f,\zeta_f\rangle \tag{3.43}$$

where $\Omega = (\Omega_k)_{k\in\mathcal{A}}$ and

$$\Omega_k = \Omega_k(\rho,\nu) = \omega_k + \nu \sum_{\ell\in\mathcal{A}} M_k^\ell \rho_l, \quad M_k^\ell = \frac{3(4-3\delta_{\ell,k})}{(2\pi)^d \lambda_k \lambda_\ell}\,, \tag{3.44}$$

$$\Lambda_a = \Lambda_a(\rho,\nu) = \lambda_a + 6\nu(2\pi)^{-d} \sum_{\ell\in\mathcal{A}} \frac{\rho_\ell}{\lambda_\ell \lambda_a}\,. \tag{3.45}$$

Besides,

$$\zeta = (\zeta_a)_{a\in\mathcal{L}}, \quad \zeta_a = \begin{pmatrix} \xi_a \\ \eta_a \end{pmatrix}, \quad \zeta_f = (\zeta_a)_{a\in\mathcal{L}_f}\,,$$

and $K(\rho)$ is a symmetric complex matrix, acting in space

$$Y_\gamma^f = \{\zeta_f\} \simeq \mathbb{C}^{2|\mathcal{L}_f|}\,, \tag{3.46}$$

such that the corresponding quadratic form is

$$\begin{aligned} \langle K(\rho)\zeta_f, \zeta_f\rangle = & 3(2\pi)^{-d}\Big(\sum_{\ell\in\mathcal{A},\ a\in\mathcal{L}_f} \frac{(3\delta_{\ell,|a|}-2)}{\lambda_\ell\lambda_a} \rho_\ell \xi_a \eta_a \\ & + \sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_+} \frac{\sqrt{\rho_{\ell(a)}}\sqrt{\rho_{\ell(b)}}}{\lambda_a\lambda_b} (\eta_a\eta_b + \xi_a\xi_b) + \\ & 2 \sum_{(a,b)\in(\mathcal{L}_f\times\mathcal{L}_f)_-} \frac{\sqrt{\rho_{\ell(a)}}\sqrt{\rho_{\ell(b)}}}{\lambda_a\lambda_b} \xi_a\eta_b \Big). \end{aligned} \tag{3.47}$$

Note that the matrix M in (3.44) is invertible since

$$\det M = 3^n (2\pi)^{-dn} \big(\Pi_{k\in\mathcal{A}}\lambda_k\big)^{-2} \det\,(4-3\delta_{\ell,k})_{\ell,k\in\mathcal{A}} \neq 0\,.$$

Relation (3.15) follows from the explicit formulas (3.44)-(3.47), so the items (i) and (ii) of Proposition 3.3 are proven.

It is clear that the matrix $K(\rho)$ is analytic in $\rho\in\mathcal{D}_{c_1}$ (see definition in (3.13)) and satisfies (3.16). This proves (iii).

It remains to verify (iv). By Proposition 3.7 the function f belongs to the class $\mathcal{T}^{\gamma,D}(\frac12, \frac{c_*}{2\sqrt2}, \mathcal{D})$. Since the reminding term f has the form (3.42) then for $(r,\theta,\zeta)\in\mathcal{O}^\gamma(\frac12,\frac{c_*}{2\sqrt2})$ it satisfies the estimates

$$|f| \le C\nu\,, \quad \|\nabla_\zeta f\|_\gamma \le C\nu\,, \quad \|\nabla_\zeta^2 f\|_\gamma^D \le C\nu\,.$$

Now consider the f^T-component of f. Only the second term in (3.42) contributes to it and we have that

$$|f^T| + \|\nabla_\zeta f^T\|_\gamma + \|\nabla_\zeta^2 f^T\|_\gamma^D \le C\nu^{3/2}\,.$$

Recall that the function f depends on the parameter ρ through the substitution $I=\nu\rho+\nu r$. So f is analytic in ρ and holomorphically extends to a complex neighbourhood of $\mathcal{D}$ of order one, where it satisfies the estimates above with a modified constant C. Therefore by the Cauchy estimate the gradient of f in ρ satisfies in the smaller complex neighbourhood $\mathcal{D}_{c_1}$ the same estimates as above, again with a modified constant. This implies the assertion (iv) of the theorem.

We will provide the domain $\mathcal{O}^\gamma\Big(\frac12,\frac{c_*}{2\sqrt2}\Big)^{\mathcal{L}}$ with the coordinates (r,θ,ξ,η) with the symplectic structure $-\sum_{\ell\in\mathcal{A}} dr_\ell\wedge d\theta_\ell - i\sum_{a\in\mathcal{L}} d\xi_a\wedge d\eta_a$. Then the transformed hamiltonian system, constructed in Proposition 3.3 has the Hamiltonian, given by the r.h.s. of (3.14).

4. The final normalisation.

The normal form, provided by Proposition 3.3, has two disadvantages: it is written in the complex variables with the non-standard reality condition (3.3), while in the original equation (1.6) the reality condition is standard (namely, $u(t,x)$ and $v(t,x)$ are real functions), and – which is much more important – the hamiltonian operators $iJK(\rho)$, corresponding to different ρ, do not commute. In this section we pass in the normal form to the variables with the usual reality condition, construct a ρ-dependent transformation which diagonalises the hamiltonian operator, and examine the smoothness of this transformation as a function of ρ. So here we are concerned with analysis of the finite-dimensional linear hamiltonian system, corresponding to the Hamiltonian (3.47). In difference with the previous sections, now the parameter ρ will belong to subdomains $Q \subset \mathcal{D}$, which are closed semi-analytic sets, defined by single polynomial relation, $Q = \{\rho \mid P(\sqrt{\rho}) \geq \delta\}$. We recall that smoothness of functions on such domains is understood in the sense of Whitney. We keep assuming that the set $\mathcal{A}$ is admissible. We recall that we provide the phase-space with the symplectic structure $-\sum dr_\ell \wedge d\theta_\ell \ -i\sum d\xi_a \wedge d\eta_a$.

4.1. **Matrix $K(\rho)$.** Recalling (3.7) and (3.8), we write the symmetric matrix $K(\rho)$, defined by relation (3.47), as a block-matrix, polynomial in $\sqrt{\rho} = (\sqrt{\rho}_1, \dots, \sqrt{\rho}_n)$. We write it as $K(\rho) = K^d(\rho) + K^{n/d}(\rho)$. Here K^d is the block-diagonal matrix

$$
\begin{aligned}
&K^d(\rho) = \operatorname{diag}\left(\begin{pmatrix} 0 & \mu(a,\rho) \\ \mu(a,\rho) & 0 \end{pmatrix},\ a \in \mathcal{L}_f \right), \\
&\mu(a,\rho) = C_*\big(\frac{3}{2}\,\rho_{\ell(a)}\lambda_a^{-2} - \lambda_a^{-1}\sum_{l\in\mathcal{A}}\rho_l\lambda_l^{-1}\big)\,, \quad C_* = 3(2\pi)^{-d}.
\end{aligned}
\tag{4.1}
$$

Note that[9]

$$
\mu(a,\rho) \quad \text{is a function of } |a| \text{ and } \rho\,. \tag{4.2}
$$

The non-diagonal matrix $K^{n/d}$ has zero diagonal blocks, while for $a \neq b$ its block $K^{n/d}(\rho)^b_a$ equals

$$
C_*\frac{\sqrt{\rho_{l(a)}\rho_{l(b)}}}{\lambda_a\lambda_b}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \chi^+(a,b) + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \chi^-(a,b) \right),
$$

where

$$
\chi^+(a,b) = \begin{cases} 1, & (a,b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+, \\ 0, & \text{otherwise,} \end{cases}
$$

and χ^- is defined similar in terms of the set $(\mathcal{L}_f \times \mathcal{L}_f)_-$. In view of (3.31),

$$
\chi^+(a,b)\cdot\chi^-(a,b) \equiv 0.
$$

[9]Here and in similar situations below we do not mention the obvious dependence on the parameter $m \in [1,2]$.

Accordingly, the hamiltonian matrix $\mathcal{H}(\rho) = iJK(\rho)$ equals $\big(\mathcal{H}^d(\rho) + \mathcal{H}^{n/d}(\rho)\big)$, where

$$
\begin{aligned}
&\mathcal{H}^d(\rho) = i\,\mathrm{diag}\left(\begin{pmatrix} \mu(a,\rho) & 0 \\ 0 & -\mu(a,\rho) \end{pmatrix},\ a \in \mathcal{L}_f\right), \\
&\mathcal{H}^{n/d}(\rho)_a^b = iC_* \frac{\sqrt{\rho_{\ell(a)}}\sqrt{\rho_{\ell(b)}}}{\lambda_a\lambda_b}\Big[J\chi^+(a,b) + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\chi^-(a,b)\Big].
\end{aligned}
\tag{4.3}
$$

Note that all elements of the matrix $\mathcal{H}(\rho)$ are pure imaginary, and

$$
\text{if } (\mathcal{L}_f \times \mathcal{L}_f)_+ = \emptyset, \text{ then } -i\mathcal{H}(\rho) \text{ is real symmetric,} \tag{4.4}
$$

in which case all eigenvalues of $\mathcal{H}(\rho)$ are pure imaginary. In Appendix B we show that if $d \geq 2$, then, in general, the set $(\mathcal{L}_f \times \mathcal{L}_f)_+$ is not empty and the matrix $\mathcal{H}(\rho)$ may have hyperbolic eigenvalues.

Example 4.1. In view of Example 3.9, if $d = 1$ then the operator $\mathcal{H}^{n/d}$ vanishes. We see immediately that in this case $\mathcal{H}^d$ is a diagonal operator with simple spectrum.

Let us introduce in $\mathcal{L}_f$ the relation $\sim$, where

$$
a \sim b \ \text{ if and only if } \ a = b \ \text{ or } \ (a,b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+ \cup (\mathcal{L}_f \times \mathcal{L}_f)_- . \tag{4.5}
$$

It is easy to see that this is an equivalence relation. By Lemma 3.10

$$
a \sim b,\ a \neq b \ \Rightarrow\ |a| \neq |b| . \tag{4.6}
$$

The equivalence $\sim$, as well as the sets $(\mathcal{L}_f \times \mathcal{L}_f)_\pm$, depends only on the lattice $\mathbb{Z}^d$ and the set $\mathcal{A}$, not on the eigenvalues λ_a and the vector ρ. It is trivial if $d = 1$ or $|\mathcal{A}| = 1$ (see Example 3.9)) and, in general, is non-trivial otherwise. If $d \geq 2$ and $|\mathcal{A}| \geq 2$ it is rather complicated.

The equivalence relation divides $\mathcal{L}_f$ into equivalence classes, $\mathcal{L}_f = \mathcal{L}_f^1 \cup \cdots \cup \mathcal{L}_f^M$. The set $\mathcal{L}_f$ is a union of the punched spheres $\Sigma_a = \{b \in \mathbb{Z}^d \mid |b| = |a|, b \neq a\}$, $a \in \mathcal{A}$, and by (4.6) each equivalence class $\mathcal{L}_f^j$ intersects every punched sphere Σ_a at at most one point.

Let us order the sets $\mathcal{L}_f^j$ in such a way that for a suitable $0 \leq M_* \leq M$ we have
– $\mathcal{L}_f^j = \{b_j\}$ (for a suitable point $b_j \in \mathbb{Z}^d$) if $j \leq M_*$;
– $|\mathcal{L}_j^f| = n_j \geq 2$ if $j > M_*$.

Accordingly the complex space $Y^f = Y_0^f$ (see (3.32)) decomposes as

$$
Y^f = Y^{f1} \oplus \cdots \oplus Y^{fM}, \quad Y^{fj} = \operatorname{span}\{\zeta_s, s \in \mathcal{L}_f^j\}. \tag{4.7}
$$

Since each $\zeta_s, s \in \mathcal{L}_f$, is a 2-vector, then

$$
\dim Y^{fj} = 2|\mathcal{L}_j^j| := 2n_j\,, \qquad \dim Y^f = 2|\mathcal{L}_f| = 2\sum_{j=1}^M n_j := 2\mathbf{N}\,.
$$

So $\dim Y^{fj} = 2$ for $j \leq M_*$ and $\dim Y^{fj} \geq 4$ for $j > M_*$. In view of (4.6),

$$
|\mathcal{L}_f^j| = n_j \leq |\mathcal{A}| \qquad \forall j\,. \tag{4.8}
$$

We readily see from the formula for the matrix $\mathcal{H}(\rho) = iJK(\rho)$ that the spaces Y^{fj} are invariant for the operator $\mathcal{H}(\rho)$. So

$$
\mathcal{H}(\rho) = \mathcal{H}^1(\rho) \oplus \cdots \oplus \mathcal{H}^M(\rho)\,, \qquad \mathcal{H}^j = \mathcal{H}^{j\,d} + \mathcal{H}^{j\,n/d}, \tag{4.9}
$$

where $\mathcal{H}^j$ operates in the space Y^{fj} and $\mathcal{H}^{j\,d}$ and $\mathcal{H}^{j\,n/d}$ are given by the formulas (4.3) with $a,b \in \mathcal{L}_f^j$. The hamiltonian operator $\mathcal{H}^j(\rho)$ polynomially depends on $\sqrt{\rho}$. So its eigenvalues form an algebraic function of $\sqrt{\rho}$ (see (3.13)). Since the spectrum of $\mathcal{H}^j(\rho)$ is an even set, then we can write this algebraic function as $\{\pm i\Lambda_1^j(\rho), \dots, \pm i\Lambda_{n_j}^j(\rho)\}$ (the factor i is convenient for further purposes). The eigenvalues of $\mathcal{H}(\rho)$ are given by another algebraic function which we write as $\{\pm i\Lambda_m(\rho), 1 \le m \le \mathbf{N} = |\mathcal{L}_f|\}$. Accordingly,

$$\{\pm\Lambda_1(\rho), \dots, \pm\Lambda_{\mathbf{N}}(\rho)\} = \cup_{j\le M}\{\pm\Lambda_k^j(\rho), k \le n_j\}\,, \tag{4.10}$$

and $\Lambda_j = \Lambda_1^j$ for $j \le M_*$.

The functions Λ_k and Λ_k^j are defined up to multiplication by ± 1.[10] But if $j \le M_*$, then $\mathcal{L}_f^j = \{b_j\}$ and $\mathcal{H}^j = \mathcal{H}^{jd}$, so the spectrum of this operator is $\{\pm i\mu(b_j,\rho)\}$, where $\mu(b_j,\rho)$ is a well defined analytic function of ρ, given by the explicit formula (4.1). In this case we specify the choice of Λ_1^j:

$$\text{if } \ \mathcal{L}_f^j = \{b_j\}, \text{ we choose } \Lambda_1^j(\rho) = \mu(b_j,\rho). \tag{4.11}$$

So for $j \le M_*$, $\Lambda_j(\rho) = \mu(b_j,\rho)$ is a polynomial of $\sqrt{\rho}$, which depends only on $|b_j|$ and ρ.

Since the norm of the operator $K(\rho)$ satisfies (3.20), then

$$|\Lambda_r^j(\rho)| \le C_2 \quad \forall\,\rho,\ \forall\,r\,,\ \forall\,j\,. \tag{4.12}$$

Example 4.2. In view of (4.8), if $\mathcal{A}$ is a one-point set, $\mathcal{A} = \{a_*\}$, then all sets $|\mathcal{L}_f^j|$ are one-point. So $M_* = M = \mathbf{N}$ and

$$\{\pm\Lambda_1(\rho), \dots, \pm\Lambda_{\mathbf{N}}(\rho)\} = \{\pm\mu(a,\rho) \mid a \in \mathbb{Z}^d, |a| = |a_*|, a \ne a_*\}.$$

In this case the spectrum of the hamiltonian operator $\mathcal{H}(\rho)$ is pure imaginary, multiple and analytically depends on ρ.

Let $1 \le j_* \le n$ and $\mathcal{D}_0^{j_*}$ be the set

$$\mathcal{D}_0^{j_*} = \{\rho = (\rho_1, \dots, \rho_n) \mid c_* \le \rho_l \le c_{**} \ \text{ if } \ l \ne j_* \ \text{ and } \ 1 - c_{**} \le \rho_{j_*} \le 1\}\,, \tag{4.13}$$

where $0 < c_* \le \frac{1}{2}c_{**} < 1/4$. This is a subset of $\mathcal{D} = [c_*, 1]^n$ which lies in the $(\mathrm{Const}\, c_{**})$-vicinity of the point $\rho_* = (0, \dots, 1, \dots, 0)$ in $[0,1]^n$, where 1 stands on the j_*-th place. Since $K^{n/d}(\rho_*) = 0$, then $K(\rho_*) = K^d(\rho^*)$. Consider any equivalence class $\mathcal{L}_f^j$ and enumerate its elements as $b_1^j, \dots, b_{n_j}^j$ $(n_j \le n)$. For $\rho = \rho_*$ the matrix $\mathcal{H}^j(\rho_*)$ is diagonal with the eigenvalues $\pm i\mu(b_r^j, \rho_*), 1 \le r \le n_j$. It suggests that for c_{**} sufficiently small we may uniquely numerate the eigenvalues $\{\pm i\Lambda_r^j(\rho)\}$ $(\rho \in \mathcal{D}_0^{j_*})$ of the matrix $\mathcal{H}^j(\rho)$ in such a way that $\Lambda_r^j(\rho)$ is close to $\mu(b_r^j, \rho_*)$. Below we justify this possibility.

Take any $b \in \mathcal{L}_f$ and denote $\ell(b) = a_b \in \mathcal{A}$. If $a_b = a_{j_*}$, then

$$\mu(b,\rho_*) = C_*(\frac{3}{2}\lambda_{a_{j_*}}^{-2} - \lambda_{a_{j_*}}^{-2}) = \tfrac{1}{2}C_*\lambda_{a_{j_*}}^{-2}\,. \tag{4.14}$$

If $a_b \ne a_{j_*}$, then

$$\mu(b,\rho_*) = -C_*\lambda_{a(b)}^{-1}\lambda_{a_{j_*}}^{-1}. \tag{4.15}$$

[10] More precisely, if Λ_k is not real, then well defined is the quadruple $\{\pm\Lambda_k, \pm\bar{\Lambda}_k\}$; see below.

If $m \in [1,2]$ is different from $4/3$ and $5/3$, then it is easy to see that $2\lambda_a \neq \pm\lambda_{a'}$ for any $a, a' \in \mathcal{A}$. By (2.14) this implies that for $m \in [1,2] \setminus \mathcal{C}$ and for $b, b' \in \mathcal{L}_f$ such that $|b| \neq |b'|$ we have

$$|\mu(b,\rho_*)| \geq 2c^{\#}(m) > 0\,, \quad |\mu(b,\rho_*) \pm \mu(b',\rho_*)| \geq 2c^{\#}(m)\,,$$

and

$$|\mu(b,\rho)| \geq c^{\#}(m) > 0\,, \qquad |\mu(b,\rho) \pm \mu(b',\rho)| \geq c^{\#}(m) \quad \text{for} \;\; \rho \in \mathcal{D}_0^{j_*}\,, \tag{4.16}$$

if c_{**} is small. In particular, for each j the spectrum $\pm i\mu(b_r^j,\rho_*), 1 \leq r \leq n_j$ of the matrix $\mathcal{H}^j(\rho_*)$ is simple.

Lemma 4.3. *If $c_{**} \in (0,1/2)$ is sufficiently small,*[11] *then there exists $c^o = c^o(m) > 0$ such that for each r and j, $\Lambda_r^j(\rho)$ is a real analytic function of $\rho \in \mathcal{D}_0^{j_*}$, satisfying*

$$|\Lambda_r^j(\rho) - \mu(b_r^j,\rho)| \leq C\sqrt{c_{**}} \qquad \forall\, \rho \in D_0^{j_*}\,, \tag{4.17}$$

and

$$|\Lambda_r^j(\rho)| \geq c^o(m) > 0 \;\; \textit{and} \;\; |\Lambda_r^j(\rho) \pm \Lambda_l^j(\rho)| \geq c^o(m) \;\; \forall\, r \neq l, \forall\, j, \forall\, \rho \in D_0^{j_*}\,, \tag{4.18}$$

$$|\Lambda_{r_1}^{j_1}(\rho) + \Lambda_{r_2}^{j_2}(\rho)| \geq c^0(m) \quad \forall\, j_1, j_2, r_1, r_2 \quad \textit{and} \;\; \rho \in D_0^{j_*}\,. \tag{4.19}$$

In particular,

$$\Lambda_r^j \not\equiv 0 \quad \forall r; \quad \Lambda_r^j \not\equiv \pm\Lambda_l^j \quad \forall\, r \neq l\,. \tag{4.20}$$

The estimate (4.17) assumes that for $\rho \in \mathcal{D}_0^{j_*}$ we fix the sign of the function Λ_r^j by the following agreement:

$$\Lambda_j^r(\rho) \in \mathbb{R} \;\text{ and } \;\; \operatorname{sign}\Lambda_r^j(\rho) = \operatorname{sign}\mu(b_r^j,\rho) \;\; \forall \rho \in \mathcal{D}_0^{j_*}\,,\; \forall 1 \leq j_* \leq n\,,\; \forall\, r, j\,, \tag{4.21}$$

see (4.14), (4.15).

Below we fix any $c_{**} = c_{**}(\mathcal{A}, m, g(\cdot)) \in (0,1/2)$ such that the lemma's assertion holds, but the parameter $c_* \in (0, \frac{1}{2}c_{**}]$ will vary during the argument.

Proof. Since the spectrum of $\mathcal{H}^j(\rho_*)$ is simple and the matrix $\mathcal{H}^j(\rho)$ and the numbers $\mu(b_r^j,\rho)$ are polynomials of $\sqrt{\rho}$, then the basic perturbation theory implies that the functions $\Lambda_r^j(\rho)$ are real analytic in $\sqrt{\rho}$ in the vicinity of ρ_* and we have

$$|\mu(b_r^j,\rho_*) - \mu(b_r^j,\rho)| \leq C\sqrt{c_{**}}\,, \quad |\Lambda_r^j(\rho_*) - \Lambda_r^j(\rho)| \leq C\sqrt{c_{**}}\,,$$

so (4.17) holds. It is also clear that the functions $\Lambda_r^j(\rho)$ are analytic in $\rho \in \mathcal{D}_0^{j_*}$. Relations (4.17) and (4.16) (and the fact that $\mu(b,\rho)$ depends only on $|b|$ and ρ) imply (4.18) and (4.19) if $c_{**} > 0$ is sufficiently small. $\square$

Remark 4.4. The differences $|2\lambda_a - \lambda_b|$ can be estimated from below uniformly in a, b in terms of the distance from $m \in [1,2]$ to the points $4/3$ and $5/3$. So the constants $c^{\#}$ and c^o depend only on this distance, and they can be chosen independent from m if the latter belongs to the smaller segment $[1, 5/4]$.

[11] Its smallness only depends on $\mathcal{A}, m$ and $g(\cdot)$.

Contrary to (4.19), in general a difference of two eigenvalues $\Lambda_{r_1}^{j_1} - \Lambda_{r_2}^{j_2}$ may vanish identically. Indeed, if $j, k \le M_*$, then $\mathcal{L}_f^k$ and $\mathcal{L}_f^j$ are one-point sets, $\mathcal{L}_f^k = \{b_k\}$ and $\mathcal{L}_f^j = \{b_j\}$, and $\Lambda_1^j = \mu(b_j, \cdot)$, $\Lambda_1^k = \mu(b_k, \cdot)$. So if $|b_j| = |b_k|$, then $\Lambda_1^j \equiv \Lambda_1^k$ due to (4.2). In particular, in view of Example 4.2, if $n = 1$ then each $\mathcal{L}_f^j$ is a one-point set, corresponding to some point b_j of the same length. In this case all functions $\Lambda_k(\rho)$ coincide identically. But if $j \le M_* < k$, or if $\max j, k > M_*$ and the set $\mathcal{A}$ is strongly admissible (recall that everywhere in this section it is assumed to be admissible), then $\Lambda_{r_1}^{j_1} - \Lambda_{r_2}^{j_2} \not\equiv 0$. This is the assertion of the non-degeneracy lemma below, proved in Section 5.

Lemma 4.5. *Consider any two spaces $Y^{f\,r_1}$ and $Y^{f\,r_2}$ such that $r_1 \le r_2$ and $r_2 > M_*$. Then*

$$\Lambda_j^{r_1} \not\equiv \pm \Lambda_k^{r_2} \qquad \forall\, (r_1, j) \ne (r_2, k)\,, \tag{4.22}$$

provided that either $r_1 \le M_*$, or the set $\mathcal{A}$ is strongly admissible.

We recall that for $d \le 2$ all admissible sets are strongly admissible. For $d \ge 3$ non-strongly admissible sets exist. In Appendix B we give an example (B.2) of such a set for $d = 3$ and show that for it the relation (4.22) does not hold.

4.2. **Real variables.** Let us pass in (3.14) from the complex variables $\zeta^{\mathcal{L}} = (\xi^{\mathcal{L}}, \eta^{\mathcal{L}})$ to the real variables $\tilde\zeta^{\mathcal{L}} = (u^{\mathcal{L}}, v^{\mathcal{L}})$, where

$$\xi_l = \frac{1}{\sqrt{2}}\,(u_l + iv_l), \quad \eta_l = \frac{1}{\sqrt{2}}\,(u_l - iv_l), \qquad l \in \mathcal{L}\,, \tag{4.23}$$

and denote by Σ the mapping

$$\Sigma : (r, \theta, u^{\mathcal{L}}, v^{\mathcal{L}}) \mapsto (r, \theta, \zeta^{\mathcal{L}})\,. \tag{4.24}$$

Below we write $u, v, \tilde\zeta$ instead of $u^{\mathcal{L}}, v^{\mathcal{L}}, \tilde\zeta^{\mathcal{L}}$, and write $\tilde\zeta = \tilde\zeta^{\mathcal{L}}$ as $\tilde\zeta = (\tilde\zeta_f, \tilde\zeta_\infty)$, where $\tilde\zeta_f$ is formed by the components (u_l, v_l) of $\tilde\zeta$, belonging to the set $\mathcal{L}_f$, and similar with $\tilde\zeta_\infty$.

The new variables are real in the sense that now the reality condition, corresponding to the involution (3.2), becomes

$$\bar u_l = u_l, \quad \bar v_l = v_l \qquad \forall\, l \in \mathcal{L}\,,$$

and the composition $(r, \theta, u, v) \mapsto (r, \theta, \zeta^{\mathcal{L}}) \xrightarrow{\Phi_\rho} \zeta \mapsto (u(x), v(x))$ sends real vectors (r, θ, u, v) to real-valued functions $(u(x), v(x))$.

In the variables (r, θ, u, v) the symplectic form $-dr \wedge d\theta - i d\xi \wedge d\eta$ reads

$$\omega_2 = -dr \wedge d\theta - du \wedge dv,$$

The transformed Hamiltonian is $\widetilde K \widetilde K$

$$\begin{aligned} (H_2 + P) \circ \Phi_\rho = \Omega(\rho) \cdot r + \frac{1}{2} \sum_{a \in \mathcal{L}_\infty} \Lambda_a(\rho)(u_a^2 + v_a^2) \\ + \frac{\nu}{2} \langle \widetilde K(\rho) \tilde\zeta_f, \tilde\zeta_f \rangle + \tilde f(r, \theta, \tilde\zeta; \rho)\,. \end{aligned} \tag{4.25}$$

The assertion (ii)-(iv) of Proposition 3.3 in the new variables stay almost the same – we only note that the hamiltonian operator $\nu \mathcal{H}(\rho) = iJ\nu K$ now reads $J\nu\widetilde K$. Here $\langle \widetilde K \tilde\zeta_f, \tilde\zeta_f \rangle$ is the quadratic form $\langle K\zeta_f, \zeta_f \rangle$, written in the variables $\tilde\zeta_f$. So the spectrum of the operator $\mathcal{H}(\rho)$ equals that of the operator $J\widetilde K(\rho)$.

The transformation (4.23) obviously respects the decomposition (4.9), and

$$J\widetilde{K}(\rho)=L^1(\rho)\oplus\cdots\oplus L^M(\rho), \tag{4.26}$$

where each $L^j(\rho)$ is a real operator in the space Y^{fj}, corresponding to a set $\mathcal{L}_f^j$. We identify Y^{fj} with $\mathbb{C}^{2n_j}$, where $n_j=|\mathcal{L}_f^j|$, and identify the operator $L^j(\rho)$ with its matrix in $\mathbb{C}^{2n_j}$. This is an $(2n_j\times 2n_j)$-matrix which polynomially depends on $\sqrt{\rho}$ and is real for real ρ.

It is easy to see that any one-dimensional set $\mathcal{L}_f^j=\{b_j\}$ $j\le M_*$, contributes to the quadratic form $\frac{\nu}{2}\langle\widetilde{K}(\rho)\tilde{\zeta}_f,\tilde{\zeta}_f\rangle$ the term $\frac{\nu}{2}\mu(b_j,\rho)(u_{b_j}^2+v_{b_j}^2)$. So

$$\langle\widetilde{K}(\rho)\tilde{\zeta}_f,\tilde{\zeta}_f\rangle=\sum_{j=1}^{M_*}\mu(b_j,\rho)\Big(u_{b_j}^2+v_{b_j}^2\Big)+\langle\widetilde{K}_*(\rho)\tilde{\zeta}_f,\tilde{\zeta}_f\rangle\,, \tag{4.27}$$

where

$$J\widetilde{K}_*(\rho)=L^{M_*+1}(\rho)\oplus\cdots\oplus L^M(\rho),$$

and each operator L^r, $r\ge M_*+1$, is a block of size ≥ 2. The hamiltonian operators $L^j, j\le M_*$, corresponding to terms in the sum in (4.27), are given by the 2×2 matrices $L^j(\rho)=\mu(b_j,\rho)J$, and

$$J\widetilde{K}(\rho)=L^1(\rho)\oplus\cdots\oplus L^M(\rho)\,.$$

According to (4.10) and (4.27) we numerate the eigenvalues $\{\pm i\Lambda_j, 1\le j\le \mathbf{N}\}$ in such a way that

$$\Lambda_j(\rho)=\Lambda_1^j(\rho)=\mu(b_j,\rho)\quad\text{if}\quad 1\le j\le M_*\,. \tag{4.28}$$

4.3. **Removing singular values of the parameter ρ.** Due to Lemma 4.3 we know that for each j the eigenvalues $\{\pm i\Lambda_k^j(\rho),\ k\le n_j\}$, do not vanish identically in ρ and do not identically coincide. Now our goal is to quantify these statements by removing certain singular values of the parameter ρ. To do this let us first denote $P^j(\rho)=(\prod_l\Lambda_l^j(\rho))^2=\pm\det L^j(\rho)$ and consider the determinant

$$P(\rho)=\prod_j P^j(\rho)=\pm\det J\widetilde{K}(\rho)\,.$$

Recall that for an $R\times R$-matrix with eigenvalues $\kappa_1,\dots,\kappa_R$ (counted with their muiltiplicities) the discriminant of the determinant of this matrix equals the product $\prod_{i\ne j}(\kappa_1-\kappa_j)$. This is a polynomial of the matrix' elements.

Next we define a “poly-discriminant” $D(\rho)$, which is another polynomial of the matrix elements of $J\widetilde{K}(\rho)$. Its definition is motivated by Lemma 4.5, and it is different for the admissible and strongly admissible sets $\mathcal{A}$. Namely, if $\mathcal{A}$ is strongly admissible, then

– for $r=1,\dots,M_*$ define $D^r(\rho)$ as the discriminant of the determinant of the matrix $L^r(\rho)\oplus L^{M_*+1}(\rho)\oplus\cdots\oplus L^M(\rho)$;
– set $D(\rho)=D^1(\rho)\cdot\dots\cdot D^{M_*}(\rho)$.

This is a polynomial in the matrix coefficients of $J\widetilde{K}(\rho)$, so a polynomial of $\sqrt{\rho}$. It vanishes if and only if $\Lambda_m^r(\rho)$ equals $\pm\Lambda_k^l(\rho)$ for some r,l,m and k, where either $r,l\ge M_*+1$ and $m\ne k$ if $r=l$, or $r\le M_*$ and $m=1$.

If $\mathcal{A}$ is admissible, then we:

– for $l \le M_*, r \ge M_*+1$ define $D^{l,r}(\rho)$ as the discriminant of the determinant of the matrix $L^l(\rho) \oplus L^r(\rho)$;

– set $D(\rho) = \prod_{l \le M_*, r \ge M_*+1} D^{l,r}(\rho)$.

This is a polynomial in the matrix coefficients of $J\widetilde{K}(\rho)$, so a polynomial in $\sqrt{\rho}$. It vanishes if and only if $\Lambda_1^r(\rho)$ equals $\pm\Lambda_k^l(\rho)$ for some $r \le M_*$, some $l \ge M_*+1$ and some k, or if $\Lambda_k^l(\rho)$ equals $\pm\Lambda_m^l(\rho)$ for some $l \ge M_*+1$ and some $k \ne m$.

Finally, in the both cases we set

$$M(\rho) = \prod_{b \in \mathcal{L}_f} \mu(b,\rho) \prod_{\substack{b,b' \in \mathcal{L}_f \\ |b| \ne |b'|}} \big(\mu(b,\rho) - \mu(b',\rho)\big).$$

This also is a polynomial in $\sqrt{\rho}$ which does not vanish identically due to (4.16).

The set

$$X = \{\rho \mid P(\rho)\, D(\rho)\, M(\rho) = 0\}$$

is an algebraic variety, if written in the variable $\sqrt{\rho}$ (analytically diffeomorphic to the variable $\rho \in [c_*, 1]^{\mathcal{A}}$), and is non-trivial by Lemma 4.3. The open set $\mathcal{D} \setminus X$ is dense in $\mathcal{D}$ and is formed by finitely many connected components. Denote them $Q_1, \dots, Q_L$. For any component Q_l its boundary is a stratified analytic manifold with finitely many smooth analytic components of dimension $< n$, see [10, 22]. The eigenvalues $\Lambda_j(\rho)$ and the corresponding eigenvectors are locally analytic functions on the domains Q_l, but since some of these domains may be not simply connected, then the functions may have non-trivial monodromy, which would be inconvenient for us. But since each Q_l is a domain with a regular boundary, then by removing from it finitely many smooth closed hyper-surfaces we cut Q_l to a finite system of simply connected domains $Q_l^1, \dots, Q_l^{\hat{n}_l}$ such that their union has the same measure as Q_l and each domain Q_l^μ lies on one side of its boundary.[12] We may realise these cuts (i.e. the hyper-surfaces) as the zero-sets of certain polynomial functions of ρ. Denote by $R_1(\rho)$ the product of the polynomials, corresponding to the cuts made, and remove from $\tilde{Q}_l \setminus X$ the zero-set of R_1. This zero-set contains all the cuts we made (it may be bigger than the union of the cuts), and still has zero measure. Again, $(\tilde{Q}_l \setminus X) \setminus \{\text{zero-set of } R_1\}$ is a finite union of domains, where each one lies in some domain Q_l^r.

Intersections of these new domains with the sets $\mathcal{D}_0^{j_*}$ (see (4.13)) will be important for us by virtue of Lemma 4.3, and any fixed set $\mathcal{D}_0^{j_*}$, say $\mathcal{D}_0^1$, will be sufficient for out analysis. To agree the domains with $\mathcal{D}_0^1$ we note that the boundary of $\mathcal{D}_0^1$ in $\mathcal{D}$ is the zero-set of the polynomial

$$R_2(\rho) = (\rho_1 - (1 - c_{**}))(\rho_2 - c_{**}) \dots (\rho_n - c_{**})\,,$$

and modify the set X above to the set $\tilde{X}$,

$$\tilde{X} = \{\rho \in \mathcal{D} \mid \mathcal{R}(\rho) = 0\}\,, \qquad \mathcal{R}(\rho) = P(\rho) D(\rho) M(\rho) R_1(\rho) R_2(\rho)\,.$$

As before, $\mathcal{D} \setminus \tilde{X}$ is a finite union of open domains with regular boundary. We still denote them Q_l:

$$\mathcal{D} \setminus \tilde{X} = Q_1 \cup \dots \cup Q_{\mathbb{J}}\,, \qquad \mathbb{J} < \infty\,. \tag{4.29}$$

[12]For example, if $n = 2$ and $\tilde{Q}_l$ is the annulus $A = \{1 < \rho_1^2 + \rho_2^2 < 2\}$, then we remove from A not the interval $\{\rho_2 = 0, 1 < \rho_1 < 2\} =: J$ (this would lead to a simply connected domain which lies on both parts of the boundary J), but two intervals, J and $-J$.

A domain Q_j in (4.29) may be non simply connected, but since each Q_j belongs to some domain Q_l^r, then the eigenvalues $\Lambda_a(\rho)$ and the corresponding eigenvectors define in these domains single-valued analytic functions. Since every domain Q_l lies either in $\mathcal{D}_0^1$ or in its complement, we may enumerate the domains Q_l in such a way that

$$\mathcal{D}_0^1 \setminus \tilde{X} = Q_1 \cup \cdots \cup Q_{\mathbb{J}_1}, \quad 1 \le \mathbb{J}_1 \le \mathbb{J}. \tag{4.30}$$

The domains Q_l with $l \le \mathbb{J}_1$ will play a special role in our argument.

We naturally extend $\tilde{X}$ to a complex-analytic subset $\tilde{X}^c$ of $\mathcal{D}_{c_1}$ (see (3.13)), consider the set $\mathcal{D}_{c_1} \setminus \tilde{X}^c$, and for any $\delta > 0$ consider its closed sub-domain $\mathcal{D}_{c_1}(\delta)$,

$$\mathcal{D}_{c_1}(\delta) = \{\rho \in \mathcal{D}_{c_1} \mid |\mathcal{R}(\rho)| \ge \delta\} \subset \mathcal{D}_{c_1} \setminus \tilde{X}^c.$$

Since the factors, forming $\mathcal{R}$, are polynomials with bounded coefficients, then they are bounded in $\mathcal{D}_{c_1}$:

$$\|P\|_{C^1(\mathcal{D}_{c_1})} \le C_1, \dots, \|R_2\|_{C^1(\mathcal{D}_{c_1})} \le C_1. \tag{4.31}$$

So in the domain $\mathcal{D}_{c_1}(\delta)$ the norms of the factors $P, \dots, R_2$, making $\mathcal{R}$, are bounded from below by $C_2\delta$, and similar estimates hold for the factors, making P, D and M. Therefore, by the Kramer rule

$$\|(J\widetilde{K})^{-1}(\rho)\| \le C_1\delta^{-1} \qquad \forall \rho \in \mathcal{D}_{c_1}(\delta). \tag{4.32}$$

Similar for $\rho \in \mathcal{D}_{c_1}(\delta)$ we have

$$|\Lambda_k^j(\rho)| \ge C^{-1}\delta \qquad \forall j, k, \tag{4.33}$$

$$|\mu(b,\rho)| \ge C^{-1}\delta, \quad |\mu(b,\rho) - \mu(b',\rho)| \ge C^{-1}\delta \quad \text{if } \ b, b' \in \mathcal{L}_f \ \text{ and } |b| \ne |b'|, \tag{4.34}$$

and

$$|\Lambda_{k_1}^j(\rho) \pm \Lambda_{k_2}^r(\rho)| \ge C^{-1}\delta \quad \text{where } \ (j, k_1) \ne (r, k_2). \tag{4.35}$$

In (4.35) if the set $\mathcal{A}$ is strongly admissible, then the index j is any and $r \ge M_* + 1$, while if $\mathcal{A}$ is admissible, then either $j \le M_*$ (and so $k_1 = 1$) and $r \ge M_* + 1$, or $j = r \ge M_* + 1$. The functions $\Lambda_k^j(\rho)$ are algebraic functions on the complex domain $\mathcal{D}_{c_1}(\delta)$, but their restrictions to the real domains Q_l split to branches which are well defined analytic functions.

By Lemma D.1 (applied to the domain $\{\sqrt{\rho} \mid \rho \in [c_*, 1]^{\mathcal{A}}\}$)

$$\text{meas}(\mathcal{D} \setminus \mathcal{D}_{c_1}(\delta)) \le C\delta^{\beta_4}, \tag{4.36}$$

for some positive C and β_4. Denote $c_2 = c_1/2$, define set $\mathcal{D}_{c_2}$ as in (3.13) but replacing there c_1 with c_2, and denote $\mathcal{D}_{c_2}(\delta) = \mathcal{D}_{c_1}(\delta) \cap \mathcal{D}_{c_2}$. Obviously,

$$\text{the set } \mathcal{D}_{c_2}(2\delta) \text{ lies in } \mathcal{D}_{c_1}(\delta) \text{ with its } C^{-1}\delta\text{-vicinity}. \tag{4.37}$$

Consider the eigenvalues $\pm i\Lambda_k(\rho)$. They analytically depend on $\rho \in \mathcal{D}_{c_1}(\delta)$, where $|\Lambda_k| \le C_2$ for each $k \le \mathbf{N}$ by (4.12). In view of (4.37),

$$|\frac{\partial^l}{\partial \rho^l}\Lambda_k(\rho)| \le C_l\delta^{-l} \qquad \forall\, \rho \in \mathcal{D}_{c_2}(2\delta),\ l \ge 0,\ k \le \mathbf{N}, \tag{4.38}$$

by the Cauchy estimate.

4.4. **Diagonalising.** For real ρ the spectra of the operator $J\widetilde{K}(\rho)$ and of each operator $L^l(\rho)$ (see (4.26)) are invariant with respect to the involution $z \mapsto -z$ and the complex conjugation. When the spectrum of $J\tilde{K}$ does not contain zero, we have three possibilities for its eigenvalues $i\Lambda_j$:

a) $i\Lambda_j \in i\mathbb{R} \setminus \{0\}$;

b) $i\Lambda_j \in \mathbb{R} \setminus \{0\}$;

c) $i\Lambda_j \in \mathbb{C} \setminus (\mathbb{R} \cup i\mathbb{R})$.

The eigenvalues of the last type may be arranged in quadruples $\pm i\Lambda_j$, $\pm i\Lambda_{j+1}$, where $\Lambda_{j+1} = \bar{\Lambda}_j$. Due to (4.35) the type of an eigenvalue does not change while ρ stays in a connected component Q_l of $\mathcal{D} \setminus \tilde{X}$. We recall our agreement (4.28); in particular, for $j \le M_*$ the eigenvalues $\pm i\Lambda_j$ have the type a).

For $\rho \in Q_l$ denote by $\mathcal{L}_l^h$ the set of indices j such that Λ_j is of type b) or c), i.e. the corresponding eigenvalues $\pm i\Lambda_j(\rho)$ are hyperbolic (the cardinality of $\mathcal{L}_l^h$ may depend on l, but not on $\rho \in Q_l$). Now take any $\Lambda_j(\rho)$, where $\rho \in Q_l$, $j \in \mathcal{L}_l^h$ and write it as $\Lambda_r^m(\rho)$, $m \le M$ (see (4.10)). By the symmetries of the spectrum, $\overline{\Lambda_r^m}(\rho)$ equals $\Lambda_{r'}^m(\rho)$ for some r'. Therefore, by (4.35),

$$|\Im\Lambda_j(\rho)| \ge \frac{1}{2}C^{-1}\delta \quad \text{if } \Lambda_j(\rho) \text{ is of type b) or c) and } \rho \in \mathcal{D}_{c_1}(\delta)\,. \tag{4.39}$$

Finally, let us denote

$$\tilde{Q}_j = Q_j \cap \mathcal{D}_{c_1}(2\delta) \quad \text{for} \quad j \le \mathbb{J}\,.$$

These are closed domains, connected if $\delta \ll 1$, such that

$$\cup \tilde{Q}_l = \mathcal{D}_{c_1}(2\delta) \cap \mathcal{D}\,.$$

By construction, the restrictions to the eigenvalues $\Lambda_j(\rho)$ to the real domains $\tilde{Q}_l$ are single-valued analytic functions. Since $\mathcal{D}_{c_1}(2\delta)$ satisfies a natural version of the estimate (4.36), then $\text{meas}(\mathcal{D} \setminus \mathcal{D}_{c_1}(2\delta)) \le C'\delta^{\beta_4}$. So

$$\sum \text{meas}(Q_j \setminus \tilde{Q}_j) = \text{meas}(\mathcal{D} \setminus \mathcal{D}_{c_1}(2\delta)) \le C'\delta^{\beta_4}\,. \tag{4.40}$$

In view of (4.35) and (4.26), for $\rho \in \tilde{Q}_l$, $l \le \mathbb{J}$, the matrix $J\widetilde{K}(\rho)$ has complex eigenvectors $U_l(\rho), 1 \le l \le 2\mathbf{N}$, corresponding to the eigenvalues $\pm i\Lambda_j(\rho)$, which analytically depend on ρ. We normalise them to have unit length and numerate in such a way that the eigenvector U_{2l} corresponds to $i\Lambda_l$ and U_{2l-1} corresponds to $-i\Lambda_l$. We denote by $U(\rho)$ the complex matrix with the column-vectors $U_1(\rho), \dots, U_{2\mathbf{N}}(\rho)$. It is analytic in $\rho \in \tilde{Q}_l$ and diagonalises $J\widetilde{K}(\rho)$:

$$U(\rho)^{-1}J\widetilde{K}(\rho)(\rho)U(\rho) = L^0(\rho) = i\,\text{diag}\,\{\pm\Lambda_1(\rho), \dots, \pm\Lambda_n(\rho)\}. \tag{4.41}$$

Clearly $\|U(\rho)\| \le \sqrt{2\mathbf{N}}$. In view of (4.35) and Lemma C.1

$$\|U(\rho)^{-1}\| \le C_4\delta^{-\beta_5}, \quad \forall\, \rho \in \tilde{Q}_l; \quad \beta_5 = 2\mathbf{N} - 1. \tag{4.42}$$

Now we will modify $U(\rho)$ to a symplectic transformation which still diagonalises $J\widetilde{K}(\rho)$. Denote by ω_2 the symplectic form on the complex space Y^f, $\omega_2(v_1, v_2) = \langle Jv_1, v_2\rangle$, where $\langle\cdot,\cdot\rangle$ is the complex-bilinear form, and J is the symplectic matrix (see Notation).

It is well known (see [1, 28]) that $\omega_2(U_a, U_b) = 0$, unless $\Lambda_a = -\Lambda_b$. That is,

$$\omega_2(U_{2j-1}, U_k) = \delta_{2j,k}\pi_j(\rho), \tag{4.43}$$

where δ is the Kronecker symbol. It is clear that $|\pi_k(\rho)| \le 1$. For the same reason as in Appendix C (see there (C.6)), estimate (4.42) implies that

$$1 \ge |\pi_k(\rho)| \ge C_4^{-1}\delta^{\beta_5} \qquad \forall k,\ \forall \rho \in \tilde{Q}_l\,, \quad \beta_5 = 2\mathbf{N}-1\,. \tag{4.44}$$

Let us re-normalise the vectors $U_k(\rho), 1 \le k \le 2\mathbf{N}$, to vectors $\tilde{U}_k(\rho)$, where

$$\tilde{U}_{2j-1}(\rho) = \pi_j^{-1}(\rho)U_{2j-1}(\rho), \quad \tilde{U}_{2j}(\rho) = U_{2j}(\rho).$$

The modified conjugating operator is $\widetilde{U} = U \cdot \mathrm{diag}\,(\pi_1^{-1}, 1, \pi_2^{-1}, \dots, 1)$, and for its columns – the eigen-vectors $\tilde{U}_j$ – the relations (4.43) hold with $\pi_j \equiv 1$. So the transformation $\widetilde{U}(\rho)$ is symplectic, and we still have

$$\widetilde{U}(\rho)^{-1} J\widetilde{K}(\rho)\widetilde{U}(\rho) = L^0(\rho). \tag{4.45}$$

In view of (4.42), (4.44) it satisfies

$$\|\widetilde{U}(\rho)\| + \|\widetilde{U}(\rho)^{-1}\| \le C_5\delta^{-\beta_5}, \qquad \forall \rho \in \tilde{Q}_l\,,\ \ \forall l\,. \tag{4.46}$$

The operator $\widetilde{U}$ respects the decomposition (4.26) and equals a direct sum of M symplectic transformations, acting in the spaces Y^{fj}.

The operator $J\widetilde{K}(\rho)$ Whitney–smoothly depends on $\rho \in \tilde{Q}_l$ and its eigenvalues satisfy (4.35). Since the diagonalising transformation $\widetilde{U}$ respects the block-decomposition (4.26), satisfies (4.46), and since (4.35) holds for $r = j$ and any $k_1 \ne k_2$, then the basic perturbation theory for simple eigenvalues implies that $\widetilde{U}(\rho)$ smoothly depends on ρ and

$$\sup_{\rho\in\tilde{Q}_l} \left(\|\partial_\rho^j \widetilde{U}(\rho)\| + \|\partial_\rho^j \widetilde{U}(\rho)^{-1}\|\right) \le C_j\delta^{-\beta(j)} \qquad \forall\, j \ge 0\,. \tag{4.47}$$

The positive constants $\beta(j)$ depend only on $g(\cdot), m$ and $\mathcal{A}$. Their explicit form is not important for us.

Since for $j \le M_*$ the operator $L^j(\rho)$ is given by the 2×2-matrix $\mu(b_j,\rho)J$, then its normalised eigen-vectors are ${}^t(1,i)/\sqrt{2}$ and ${}^t(1,-i)/\sqrt{2}$. So the corresponding j-th block of the operator $\widetilde{U}$ is

$$\widetilde{U}(\rho)\mid_{Y^{fj}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} =: \Upsilon, \qquad 1 \le j \le M_*\,.$$

Accordingly the transformation $\widetilde{U}$ may be written as

$$\widetilde{U}(\rho) = \underbrace{(\Upsilon \oplus \dots \oplus \Upsilon)}_{M_*\ \text{terms}} \oplus \widetilde{U}_*(\rho)\,,$$

and similar $L^0(\rho) = \mathrm{diag}\{\pm i\mu(b_j,\rho)\} \oplus L^0_*(\rho)$.

In accordance with a), b) and c), we decompose further the complex diagonal operator L^0_* as

$$L^0_*(\rho) = L^a(\rho) \oplus L^b(\rho) \oplus L^c(\rho).$$

Re-ordering the eigenvalues $\pm i\Lambda_j$ we achieve that the elliptic eigenvalues, represented in the decomposition above as the eigenvalues of $L^a(\rho)$, are

$$\{\pm i\Lambda_r(\rho), M_* < r \le M_{**}\}, \quad M_{**} \le \mathbf{N}\,, \tag{4.48}$$

and the eigenvalues $\pm i\Lambda_r$, $r > M_{**}$, are hyperbolic (i.e., their real parts are non-zero).

The complex-diagonal operator $L^a(\rho)$ is elliptic, and a direct sum of operators Υ^{-1} transform it to a hamiltonian operator, corresponding to the real Hamiltonian

$$\frac{1}{2}\sum_{i\Lambda_j \text{ has type a)}} \pm\Lambda_j(\rho)(u_j^2+v_j^2)\,. \tag{4.49}$$

The sign $\pm$ depends on the Krein signature of the pair of eigenvalues $\pm i\Lambda_j(\rho)$, see [1]. Since the eigenvalues $i\Lambda_j(\rho)$ in (4.10) are defined up to multiplication by ±1, then changing the signs for some of them we achieve that all the signs in (4.49) are "+".[13] This argument does not apply to the domains $\mathcal{D}_0^{j_*}$ as in Lemma 4.3, where the sign of each $\Lambda_j(\rho)$ is fixed by the agreement (4.21). But for ρ in that domain still all the signs are plus. Indeed, for ρ in the vicinity of ρ_* this is true by Lemma 4.3 and (4.21). Let us take any point $\rho'\in\mathcal{D}_0^{j_*}$ and consider a deformation in $\mathcal{D}_0^{j_*}$ of any point, close to ρ_*, to ρ'. During the deformation the functions Λ_j stay real and do not vanish by Lemma 4.3, so the sign remains "+".

The operator $L^b(\rho)$ is real-diagonal hyperbolic, and we do not touch it. Consider the complex-diagonal hyperbolic operator $L^c(\rho)$. It splits to a direct sum of 4-dimensional diagonal operators, where each one has eigenvalues $(\pm a(\rho)\pm ib(\rho))$, $a,b\neq0$. As we show in Appendix C, a symplectic operator $(\widetilde{U}^{a,b})^{-1}$, satisfying the estimates (C.7), transforms this complex-diagonal hamiltonian operator to the operator with a real Hamiltonian which is the one-half of the quadratic form (C.4). This operator depends on ρ. In view of (4.35)

$$|a(\rho)|,|b(\rho)|\geq C^{-1}\delta \qquad \forall\,\rho\in\tilde{Q}_l\,,\ \forall\,l\,,$$

and for the same reason as above the derivatives $\partial_\rho^j\Big((\widetilde{U}^{a(\rho),b(\rho)})^{\pm1}\Big)$ satisfy estimates (4.47).

Now consider the operator

$$\widehat{U}(\rho)=\big(\Upsilon\oplus\cdots\oplus\Upsilon\big)\oplus\ \text{id}\ \oplus\Big(\bigoplus_j \widetilde{U}^{a_j(\rho),b_j(\rho)}\Big)\,,$$

where the first direct sum in the r.h.s. acts in the sub-space, corresponding to the eigenvalues $i\Lambda_j$ of type a), the identity transformation acts in the sub-space, corresponding to the eigenvalues of type b), and the last direct sum corresponds to the eigenvalues $\Lambda_j=a_j+ib_j$ of type c). The operator $(\widehat{U}(\rho))^{-1}$ transforms the complex-diagonal operator $L^0(\rho)$ to the hamiltonian operator, corresponding to a Hamiltonian

$$\frac{1}{2}\sum_{j=1}^{M_*}\mu(b_j,\rho)\Big(u_{b_j}^2+v_{b_j}^2\Big)+\frac{1}{2}\sum_{j=M_*+1}^{M_{**}}\Lambda_j(\rho)\Big(u_{b_j}^2+v_{b_j}^2\Big)+\frac{1}{2}\langle\widehat{K}(\rho)\tilde{\zeta}_f^h,\tilde{\zeta}_f^h\rangle\,. \tag{4.50}$$

Here the vector $\tilde{\zeta}_f^h$ is formed by the hyperbolic components of the vector $\tilde{\zeta}_f$, and the spectrum of the hamiltonian operator $J\widehat{K}(\rho)$ is formed by the hyperbolic eigenvalues of the operator $J\widetilde{K}(\rho)$. The operator $\widehat{U}(\rho)$ satisfies the estimates (4.47) with suitable exponents $\beta(j)$.

The operator $(\widehat{U}(\rho))^{-1}\circ\widetilde{U}(\rho)$ transforms the Hamiltonian (4.27) to the Hamiltonian above and also satisfies the estimates (4.47) (with modified exponents $\beta(j)$).

[13] So, in general, the functions $\Lambda_j(\rho)$ of type a) cannot be chosen positive.

4.5. **Final transformation.** We have

$$\text{meas}\cup_{j\in\mathbb{J}}\tilde{Q}_j = \text{meas}\,\mathcal{D}_{c_1}(2\delta)\cap(\mathcal{D}\setminus\tilde{X}) = \text{meas}\,\mathcal{D}_{c_1}(2\delta)\cap\mathcal{D}\,. \tag{4.51}$$

The number $\mathbb{J}$ does not depend on δ, but the closed domains $\tilde{Q}_j$ depend on it, $\tilde{Q}_j = \tilde{Q}_j(\delta)$, and for each j,

$$\tilde{Q}_j(\delta) \nearrow Q_j \quad \text{as} \quad \delta \to 0\,,$$

i.e., $\tilde{Q}_j(\delta_1) \subset \tilde{Q}_j(\delta_2)$ if $\delta_1 > \delta_2$, and $\bigcup_{\delta>0}\tilde{Q}_j(\delta) = Q_j$.

For any $j \le \mathbb{J}$ let us consider the operator $\widetilde{U}(\rho)$, $\rho \in \tilde{Q}_j$, as in (4.45) and denote it $\widetilde{U}_j(\rho)$. Now we define the operators $\widehat{\mathbf{U}}(\rho)$ and $\mathbf{U}(\rho)$, $\rho \in \cup\tilde{Q}_j$, by the following relations:

$$\widehat{\mathbf{U}}(\rho) = \widetilde{U}_j(\rho), \quad \mathbf{U}(\rho) = (\hat{U}_j(\rho))^{-1}\widetilde{U}_j(\rho) \quad \text{if } \rho \in \tilde{Q}_j\,.$$

With an eye on the relations (4.40) and (4.51), for any $\beta_* > 0$ and $\nu > 0$ we choose $\delta > 0$ such that

$$C\delta^{\beta_4} = \nu^{\beta_*}, \tag{4.52}$$

where C and β_4 are the constant and the exponent from (4.36). For convenience we denote (see (4.44))

$$\bar{c} = \frac{1}{\beta_4} \quad \text{and} \quad \hat{c} = \beta_5\bar{c} = (2\mathbf{N}-1)\bar{c}. \tag{4.53}$$

The Hamiltonian (4.25) is written in the variables $(r,\theta,\tilde{\zeta}^{\mathcal{L}})$, where $\tilde{\zeta}^{\mathcal{L}} = (\tilde{\zeta}_f, \tilde{\zeta}_\infty)$, according to the decomposition $\mathcal{L} = \mathcal{L}_f \cup \mathcal{L}_\infty$. Now we decompose the variable $\tilde{\zeta}_f$ further. Namely we write

$$\mathcal{L}_f = \mathcal{L}_f^e \cup \mathcal{L}_f^h\,, \tag{4.54}$$

and

$$\tilde{\zeta}_f = (\tilde{\zeta}_f^e, \tilde{\zeta}_f^h)\,, \quad \text{where } \tilde{\zeta}_f^e = (\tilde{\zeta}_b, b\in\mathcal{L}_f^e), \ \ \tilde{\zeta}_f^h = (\tilde{\zeta}_b, b\in\mathcal{L}_f^h)\,,$$

where the sets $\mathcal{L}_f^e$ and $\mathcal{L}_f^h$ correspond, respectively, to the elliptic and hyperbolic eigenvalues $i\Lambda_j(\rho)$. The sets $\mathcal{L}_f^e$, $\mathcal{L}_f^h$ and the decompositions above depend on the domain $\tilde{Q}_l$. In particular, in view of (4.21) and (4.30)

$$\mathcal{L}_f^h = \emptyset \quad \text{if} \quad \rho \in \mathcal{D}_0^1\,.$$

Recalling that $\mu(b_j,\rho) = \Lambda_j(\rho)$ if $1 \le j \le M_*$, we write the quadratic Hamiltonian (4.50), where $\rho \in \cup\tilde{Q}_l$, as

$$\frac{1}{2}\sum_{b\in\mathcal{L}_f^e}\Lambda_b(\rho)\Big(u_b^2+v_b^2\Big) + \frac{1}{2}\langle\widehat{K}(\rho)\tilde{\zeta}_f^h, \tilde{\zeta}_f^h\rangle\,.$$

The transformation $\mathbf{U}$, constructed above, acts on the variables $\tilde{\zeta}_f$. We extend it identically to the variables $(r,\theta,\tilde{\zeta}_\infty)$, and denote the extension (acting on the variables $(r,\theta,\tilde{\zeta}^{\mathcal{L}})$) as $\mathbf{U}\oplus\text{id}$. The normal form transformation from Proposition 3.3 and the constructions, made earlier in this section, jointly yield the transformation

$$\tilde{\Phi}_\rho = \Phi_\rho \circ \Sigma \circ (\mathbf{U}^{-1}\oplus\text{id})\,.$$

We recall that $\cup_{j\in\mathbb{J}}\tilde{Q}_j(\nu) = \mathcal{D}_{c_1}(2\delta)\cap\mathcal{D}$, denote $\mathcal{D}_{c_1}(2\delta)\cap\mathcal{D} = \tilde{Q}(\nu)$ and sum up properties of this transformation in the form of a normal form theorem:

Theorem 4.6. *There exists a zero-measure Borel set $\mathcal{C} \subset [1,2]$ such that for any admissible set $\mathcal{A}$ and any $m \notin \mathcal{C}$ there exist real numbers $\beta_{*0}, \nu_0, \gamma_* \in (0,1]$, $c_{**} \in (0,1/2]$, where γ_* depends only on $g(\cdot)$ and c_{**}, β_{*0} and ν_0 also depend on $\mathcal{A}$ and m, with the following property:*
For any $c_* \in (0, \frac{1}{2}c_{**}]$ the set $\mathcal{D} = [c_*, 1]^n$ has an algebraic subset $\tilde{X}$ which is a zero-set of a polynomial of $\sqrt{\rho}$, depending on $\mathcal{A}, m$, and for any $0 < \nu \le \nu_0$ and $0 < \beta_* \le \beta_{*0}$ there exists a closed semi-analytic set $\tilde{Q}(\nu) \subset \mathcal{D} \setminus \tilde{X}$, such that $\tilde{Q}(\nu) \nearrow \big(\mathcal{D} \setminus \tilde{X}\big)$ as $\nu \to 0$, and

$$\operatorname{meas}(\mathcal{D} \setminus \tilde{Q}(\nu)) \le C\nu^{\beta_*} . \tag{4.55}$$

For $\rho \in \tilde{Q} = \tilde{Q}(\nu)$ and $0 \le \gamma \le \gamma_*$ there exist real holomorphic transformations

$$\tilde{\Phi}_\rho : \mathcal{O}^\gamma(\tfrac{1}{2}, \mu) = \{(r,\theta,u,v)\} \to \mathbf{T}_\rho(\nu, 1, 1, \gamma) , \qquad \mu = c_*/2\sqrt{2} ,$$

which do not depend on γ in the sense that they coincide on the smallest set set $\mathcal{O}^{\gamma_*}(\frac{1}{2}, \mu)$. The transformations smoothly depend on ρ and satisfy $(\tilde{\Phi}_\rho)^*(-id\xi \wedge d\eta) = \nu(-dr\wedge d\theta - du \wedge dv)$. With respect to the symplectic structure $-dr\wedge d\theta - du\wedge dv$ the transformed system has the Hamiltonian $(H_2 + P) \circ \tilde{\Phi}_\rho =: H_\rho$ (see (1.9)), where

$$\begin{aligned} H_\rho &= \Omega(\rho)\cdot r + \frac{1}{2}\sum_{a\in\mathcal{L}_\infty} \Lambda_a(\rho)(u_a^2 + v_a^2) \\ &\quad + \frac{\nu}{2}\Big(\sum_{b\in\mathcal{L}_f^e} \Lambda_b(\rho)\Big(u_b^2 + v_b^2\Big) + \langle \widehat{K}(\rho)\tilde{\zeta}_f^h, \tilde{\zeta}_f^h\rangle\Big) + \tilde{f}(r,\theta,\tilde{\zeta};\rho). \end{aligned} \tag{4.56}$$

Here the real symmetric operator $\widehat{K}(\rho)$ acts in a space of dimension $2|\mathcal{L}_f^h|$, $\mathcal{L}_f^h = \mathcal{L}_f \setminus \mathcal{L}_f^e$. The decomposition $\mathcal{L}_f = \mathcal{L}_f^e \cup \mathcal{L}_f^h$ depends on the component of the domain $\mathcal{D} \setminus \tilde{X}$ which contains ρ, and for some of these components the set $\mathcal{L}_f^h$ is empty (so the operator $\widehat{K}(\rho)$ is trivial). Moreover,

i) the functions Ω and $\Lambda_a, a \in \mathcal{L}_\infty$, are the same as in Proposition 3.3, and the function $\tilde{f}$ satisfies

$$[\tilde{f}]^{\gamma,D}_{1/2,\mu,\tilde{Q}} \le C\nu^{-\hat{c}\beta_*}\nu , \quad [\tilde{f}^T]^{\gamma,D}_{1/2,\mu,\tilde{Q}} \le C\nu^{-\hat{c}\beta_*}\nu^{3/2} , \quad \mu = (c_*/2\sqrt{2}) . \tag{4.57}$$

The functions $\Lambda_b(\rho)$, $b \in \mathcal{L}_f^e$, are real analytic in $\tilde{Q}$ and

$$\|\Lambda_b\|_{C^r(\tilde{Q})} \le C_r \nu^{-r\bar{c}\beta_*} \ (r \ge 0), \quad \forall\, \rho \in \tilde{Q} . \tag{4.58}$$

They satisfy (4.33) and for some connected components of $\mathcal{D}\setminus\tilde{X}$ also satisfy (4.19).

ii) The operator $\widehat{K}(\rho)$ smoothly depends on $\rho \in \tilde{Q}$ and may be diagonalised by a complex symplectic operator $\widehat{\mathbf{U}}(\rho)$:

$$\widehat{\mathbf{U}}(\rho)^{-1} J \widehat{K}(\rho) \widehat{\mathbf{U}}(\rho) = i \operatorname{diag}\{\pm \tilde{\Lambda}_j(\rho), 1 \le j \le |\mathcal{L}_f^h|\} .$$

The eigenvalues $\tilde{\Lambda}_j(\rho)$ satisfy (4.58) and

$$|\Im \tilde{\Lambda}_j(\rho)| \ge C^{-1}\nu^{\bar{c}\beta_*} \qquad \forall\, \rho \in \tilde{Q} , \ \forall\, j . \tag{4.59}$$

The operator $\widehat{\mathbf{U}}(\rho)$ smoothly depends on ρ and satisfies

$$\sup_{\rho\in\tilde{Q}} \big(\|\partial_\rho^j \widehat{\mathbf{U}}(\rho)\| + \|\partial_\rho^j \widehat{\mathbf{U}}(\rho)^{-1}\|\big) \le C_j \nu^{-\beta_*\beta(j)} , \qquad \forall\, j \ge 0 . \tag{4.60}$$

Accordingly,

$$\sup_{\rho\in\tilde Q} \|\partial_\rho^j J\widehat K(\rho)\| \le C_j \nu^{-\beta_*\beta'(j)} \quad \textit{for} \quad j\ge 0 \tag{4.61}$$

(the exponents $\beta(j)$ and $\beta'(j)$ depend on $m, \mathcal{A}$ and j).

iii) The domains $\tilde Q(\nu)$ and the matrix $\widehat K(\rho)$ do not depend on the component g_0 of the nonlinearity g. The constants $C, C', \bar C$ etc and the factors $\bar c$, $\hat c$ in the exponents in the estimates above do not depend on $\nu\in(0,\nu_0]$.

5. Proof of the non-degeneracy Lemma 4.5

Consider the decomposition (4.9) of the hamiltonian operator $\mathcal{H}(\rho)$. To simplify notation, in this section we suspend the agreement that $|L_f^r|=1$ for $r\le M_*$, and changing the order of the direct summands achieve that the indices r_1 and r_2, involved in (4.22), are $r_1=1$ and $r_2=2$. For $r=1,2$ we will write elements of the set $\mathcal{L}_f^r$ as $a_j^r, 1\le j\le n_r$, and vectors of the space Y^{fr} as

$$\zeta=\left(\zeta_{a_j^r}=(\xi_{a_j^r},\eta_{a_j^r}), 1\le j\le n_r\right)=\left((\xi_{a_1^r},\eta_{a_1^r}),\dots,(\xi_{a_{n_r}^r},\eta_{a_{n_r}^r})\right). \tag{5.1}$$

Using (3.7) and abusing notation, we will regard the mapping $\ell:\mathcal{L}_f\to\mathcal{A}$ also as a mapping $\ell:\mathcal{L}_f\to\{1,\dots,n\}$. Consider the points $\ell(a_1^1),\dots,\ell(a_{n_1}^1)$ (they are different by (4.6)). Changing if needed the labelling (3.7) we achieve that

$$\{\ell(a_1^1),\dots,\ell(a_{n_1}^1)\ni 1\,. \tag{5.2}$$

We write the operator $\mathcal{H}^r$ as $\mathcal{H}^r=iM^r$, where

$$M^r(\rho)=JK^r(\rho)=JK^{r\,d}(\rho)+JK^{r\,n/d}(\rho)=:M^{r\,d}(\rho)+M^{r\,n/d}(\rho)\,,$$

and the real block-matrices $M^{r\,d}=i^{-1}\mathcal{H}^{r,\,d}$, $M^{r\,n/d}=i^{-1}\mathcal{H}^{r,\,n/d}$ are given by (4.3). Then $\{\pm\Lambda_j^r(\rho)\}$ are the eigenvalues of $M^r(\rho)$, and

$$M^{r\,d}(\rho)=\mathrm{diag}\,\left(\begin{pmatrix}\mu(a_j^r,\rho) & 0\\ 0 & -\mu(a_j^r,\rho)\end{pmatrix},\ 1\le j\le n_r\right),$$

where $\mu(a_j^r,\rho)$ is given by (4.1).

Renumerating the eigenvalues we achieve that in (4.22) (with $r_1=1, r_2=2$), $\Lambda_j^1=\Lambda_1^1$ and $\Lambda_k^2=\Lambda_1^2$. As in the proof of Lemma 4.3, consider the vector $\rho_*={}^t(1,0,\dots,0)$. Let us abbreviate

$$\mu(a,\rho_*)=\mu(a)\qquad\forall\,a\,,$$

where $\mu(a)$ depends only on $|a|$ by (4.2). In view of (4.3) $M^r(\rho_*)=M^{r\,d}(\rho_*)$ and thus $\Lambda_1^1(\rho_*)=\mu(a_1^1)$ and $\Lambda_1^2(\rho_*)=\mu(a_1^2)$, if we numerate the elements of $\mathcal{L}_f^1$ and $\mathcal{L}_f^2$ accordingly. As in the proof of Lemma 4.3, $\mu(|a_1^r|)$ equals $\frac12 C_*\lambda_{a_1^r}^{-2}$ or $-C_*\lambda_{\ell(a_1^r)}^{-1}\lambda_{a_1^r}^{-1}$. Therefore the relation $\mu(a_1^1)=\pm\mu(a_1^2)$ is possible only if the sign is "+" and $|a_1^1|=|a_1^2|$. So it remains to verify that under the lemma's assumption

$$\Lambda_1^1(\rho)\not\equiv\Lambda_1^2(\rho)\quad\text{if}\quad |a_1^1|=|a_1^2|\,. \tag{5.3}$$

Since $|a_1^1|=|a_1^2|$, then

$$\ell(a_1^1)=\ell(a_1^2)=:a_{j_\#}\in\mathcal{A}\quad\text{and}\ \Lambda_1^1(\rho_*)=\Lambda_1^2(\rho_*)=:\Lambda\,.$$

To prove that $\Lambda_1^1(\rho) \not\equiv \Lambda_1^2(\rho)$ we compare variations of the two functions around $\rho = \rho_*$. To do this it is convenient to pass from ρ to the new parameter $y = (y_j)_1^n$, defined by

$$y_j = \sqrt{\rho_j}, \quad j = 1, \cdots, n.$$

Abusing notation we will sometime write y_{a_j} instead of y_j. Take any vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, where $x_1 = 0$ and $x_j > 0$ if $j \geq 2$, and consider the following variation $y(\varepsilon)$ of $y_* = (1, 0, \cdots, 0)$:

$$y_j(\varepsilon) = \begin{cases} 1 & \text{if } j = 1, \\ \varepsilon x_j & \text{if } j \geq 2. \end{cases} \tag{5.4}$$

By (4.18), for small ε the real matrix $M^r(\varepsilon) := M^r(\rho(\varepsilon))$ $(r = 1, 2)$ has a simple eigenvalue $\Lambda_1^r(\varepsilon)$, close to Λ. We will show that for a suitable choice of vector x the functions $\Lambda_1^1(\varepsilon)$ and $\Lambda_1^2(\varepsilon)$ are different. More specifically, that their jets at zero of sufficiently high order are different.

Let r be 1 or 2. We denote $\Lambda(\varepsilon) = \Lambda_1^r(\rho(\varepsilon))$, $M(\varepsilon) = M^r(\rho(\varepsilon))$ and denote by $M^d(\varepsilon)$ and $M^{n/d}(\varepsilon)$ the diagonal and non-diagonal parts of $M(\varepsilon)$. The matrix $M^{n/d}(\varepsilon)$ is formed by 2×2-blocks

(5.5)
$$\left(M^{n/d}(\varepsilon)\right)_{a_k^r}^{a_j^r} = C_* \frac{y_{\ell(a_k^r)} y_{\ell(a_j^r)}}{\lambda_{a_k^r} \lambda_{a_j^r}} \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \chi^+(a_k^r, a_j^r) + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \chi^-(a_k^r, a_j^r) \right),$$

(note that if $j = k$, then the block vanishes).

For $\varepsilon = 0$, $M(0) = M^{rd}(0)$ is a matrix with the single eigenvalue $\Lambda(0) = \mu(a_1^r, \rho_*)$, corresponding to the eigen-vector $\zeta(0) = {}^t(1, 0, \ldots, 0)$. For small ε they analytically extend to a real eigenvector $\zeta(\varepsilon)$ of $M(\varepsilon)$ with the eigenvalue $\Lambda(\varepsilon)$, i.e.

$$M(\varepsilon)\zeta(\varepsilon) = \Lambda(\varepsilon)\zeta(\varepsilon), \qquad |\zeta(\varepsilon)| \equiv 1.$$

We abbreviate $\zeta = \zeta(0), M = M(0)$ and define similar $\dot\zeta, \ddot\zeta, \Lambda, \dot\Lambda \ldots$ etc, where the upper dot stands for $d/d\varepsilon$. We have

$$M = M^d = \operatorname{diag}\left(\mu(a_1^r), -\mu(a_1^r), \ldots, -\mu(a_{n_r}^r)\right), \tag{5.6}$$

$$\dot M^d = 0. \tag{5.7}$$

Since $(M(\varepsilon) - \Lambda(\varepsilon))\zeta(\varepsilon) \equiv 0$, then

$$(M(\varepsilon) - \Lambda(\varepsilon))\dot\zeta(\varepsilon) = -\dot M(\varepsilon)\zeta(\varepsilon) + \dot\Lambda(\varepsilon)\zeta(\varepsilon). \tag{5.8}$$

Jointly with (5.6) and (5.7) this relation with $\varepsilon = 0$ implies that

$$(M^d - \Lambda)\dot\zeta = -\dot M^{n/d}\zeta + \dot\Lambda\zeta. \tag{5.9}$$

In view of (5.6) we have $\langle (M^d - \Lambda)\dot\zeta, \zeta\rangle = 0$. We derive from here and from (5.9) that

$$\dot\Lambda = \langle \dot M^{n/d}\zeta, \zeta\rangle = 0. \tag{5.10}$$

Let us denote by π the linear projection $\pi : \mathbb{R}^{2n_r} \to \mathbb{R}^{2n_r}$ which makes zero the first component of a vector to which it applies. Then $M^d - \Lambda$ is an isomorphism of the space $\pi\mathbb{R}^{2n_r}$, and the vectors $\dot\zeta$ and $-\dot M\zeta + \dot\Lambda\zeta = \dot M^{n/d}\zeta$ belong to $\pi\mathbb{R}^{2n_r}$. So we get from (5.9) that

$$\dot\zeta = -(M^d - \Lambda)^{-1}\dot M^{n/d}\zeta, \tag{5.11}$$

where the equality holds in the space $\pi\mathbb{R}^{2n_r}$. Differentiating (5.8) we find that

$$(M(\varepsilon)-\Lambda(\varepsilon))\ddot\zeta(\varepsilon) = -\ddot M(\varepsilon)\zeta(\varepsilon) - 2\dot M(\varepsilon)\dot\zeta(\varepsilon) + \ddot\Lambda(\varepsilon)\zeta(\varepsilon) + 2\dot\Lambda(\varepsilon)\dot\zeta(\varepsilon)\,. \tag{5.12}$$

Similar to the derivation of (5.10) (and using that $\langle\zeta,\dot\zeta\rangle = 0$ since $|\zeta(\varepsilon)|\equiv 1$), we get from (5.12) and (5.10) that

$$\ddot\Lambda = \langle \ddot M\zeta,\zeta\rangle + 2\langle \dot M\dot\zeta,\zeta\rangle = \langle \ddot M\zeta,\zeta\rangle + 2\langle (M-\Lambda)^{-1}\dot M^{n/d}\zeta, {}^t(\dot M)\zeta\rangle\,. \tag{5.13}$$

Since for each ε and every j

$$\frac{d^2}{d\varepsilon^2}\rho_j(\varepsilon) = \frac{d^2}{d\varepsilon^2}y_j^2(\varepsilon) = 2x_j^2\,, \qquad \frac{d^2}{d\varepsilon^2}y_1(\varepsilon)y_j(\varepsilon) = 0\,,$$

and since $\langle \ddot M\zeta,\zeta\rangle = \langle \ddot M^d\zeta,\zeta\rangle$, then

$$\langle \ddot M\zeta,\zeta\rangle = \frac{d^2}{d\varepsilon^2}\mu(a_1^r,\rho(\varepsilon))\,|_{\varepsilon=0} = C_*\lambda_{a_{j_\#}}^{-1}\big(3\lambda_{a_{j_\#}}^{-1}x_{j_\#}^2 - 2\sum_{j=2}^{n} x_j^2\lambda_{a_j}^{-1}\big) =: k_1\,. \tag{5.14}$$

Note that k_1 does not depend on r.

Now consider the second term in the r.h.s. (5.13). For any $a,b\in\mathcal{L}_f^r$ we see that $\frac{d}{d\varepsilon}(y_{\ell(a)}(\varepsilon)y_{\ell(b)}(\varepsilon))\,|_{\varepsilon=0}$ is non-zero if exactly one of the numbers $\ell(a),\ell(b)$ is a_1, and this derivative equals $x_{\ell(c)}$, where $c\in\{a,b\}$, $\ell(c)\neq a_1$. Therefore, by (5.5),

$$\begin{aligned} &(\dot M^{n/d}\zeta)_{a_j^r} = \frac{C_*}{\lambda_{a_{j_\#}}}(\xi^o_{a_j^r}, -\eta^o_{a_j^r}), \qquad a_j^r\in\mathcal{L}_f^r\,,\\ &\xi^o_{a_j^r} = \frac{\varphi(a_1^r,a_j^r)}{\lambda_{a_j^r}}\chi^-(a_1^r,a_j^r), \quad \eta^o_{a_j^r} = \frac{\varphi(a_1^r,a_j^r)}{\lambda_{a_j^r}}\chi^+(a_1^r,a_j^r)\,, \end{aligned} \tag{5.15}$$

where $\varphi(a_1^r,a_1^r)=0$ and for $j\neq 1$

$$\varphi(a_1^r,a_j^r) = \begin{cases} x_{\ell(a_j^r)} & \text{if } \ j_\# = 1\,,\\ x_{j_\#} & \text{if } \ \ell(a_j^r) = a_1\,,\\ 0 & \text{if } \ j_\#\neq 1,\ \ell(a_j^r)\neq a_1\,. \end{cases}$$

Since $\chi^\pm(a_1^r,a_1^r)=0$, then $\xi^o_{a_1^r} = \eta^o_{a_1^r} = 0$.

In view of (3.31), at most one of the numbers $\xi^o_{a_j^r},\eta^o_{a_j^r}$ is non-zero. By (5.15),

$$((M-\Lambda)^{-1}\dot M^{n/d}\zeta)_{a_j^r} = \frac{C_*}{\lambda_{a_{j_\#}}}(\xi^{oo}_{a_j^r},\,\eta^{oo}_{a_j^r}), \tag{5.16}$$

where $\xi^{oo}_{a_j^r} = \eta^{oo}_{a_j^r} = 0$ if $j=1$, and otherwise

$$\xi^{oo}_{a_j^r} = \frac{\varphi(a_1^r,a_j^r)\chi^-(a_1^r,a_j^r)}{\lambda_{a_j^r}(\mu(a_j^r)-\mu(a_1^r))}, \qquad \eta^{oo}_{a_j^r} = \frac{\varphi(a_1^r,a_j^r)\chi^+(a_1^r,a_j^r)}{\lambda_{a_j^r}(\mu(a_j^r)+\mu(a_1^r))}\,.$$

Here $\mu(a_j^r) = \frac12 C_*\lambda_{a_1}^{-2}$ if $\ell(a_j^r) = a_1$ and $\mu(a_j^r) = -C_*\lambda_{a_l^r}^{-1}\lambda_{a_1}^{-1}$ if $\ell(a_j^r)\neq a_1$.

Similar,

$$({}^t\dot M\zeta)_{a_j^r} = \frac{C_*}{\lambda_{a_{j_\#}}}(\xi^o_{a_j^r},\eta^o_{a_j^r}),$$

so the second term in the r.h.s. of (5.13) equals

$$\frac{C_*^2}{\lambda^2_{a_{j_\#}}}\sum_{j=2}^{n_r}\frac{\varphi(a_1^r,a_j^r)^2}{\lambda^2_{a_j^r}}\Big(\frac{\chi^-(a_1^r,a_j^r)}{\mu(a_j^r)-\mu(a_1^r)} + \frac{\chi^+(a_1^r,a_j^r)}{\mu(a_j^r)+\mu(a_1^r)}\Big) =: k_2(r)\ . \tag{5.17}$$

Finally, we have seen that

$$\Lambda_1^r(\rho(\varepsilon)) = \Lambda_1^1(\rho_*) + \frac{1}{2}\varepsilon^2 k_1 + \frac{1}{2}\varepsilon^2 k_2(r) + O(\varepsilon^3), \quad r = 1, 2,$$

where k_1 does not depend on r. Since $a_1^r \sim a_j^r$ for each r and each j (see (4.5)), then for $j > 1$ at least one of the coefficients $\chi^{\pm}(a_1^r, a_j^r)$ is non-zero. As $\chi^+ \cdot \chi^- \equiv 0$, then

$$\frac{\chi^-(a_1^r, a_j^r)}{\mu(a_j^r) - \mu(a_1^r)} + \frac{\chi^+(a_1^r, a_j^r)}{\mu(a_j^r) + \mu(a_1^r)} \neq 0 \qquad \forall\, r,\ \ \forall\, j > 1\,. \tag{5.18}$$

We see that the sum, defining $k_2(r)$, is a non-trivial quadratic polynomial of the quantities $\varphi(a_1^r, a_j^r)$ if $n_r \geq 2$, and vanishes if $n_r = 1$.

The following lemma is crucial for the proof.

Lemma 5.1. *If the set $\mathcal{A}$ is strongly admissible and $|a| = |b|$, $a \neq b$, and $\chi^+(a, a') \neq 0$, $\chi^+(b, b') \neq 0$, or $\chi^-(a, a') \neq 0$, $\chi^-(b, b') \neq 0$, then $|a'| \neq |b'|$.*

Proof. Let first consider the case when $\chi^+ \neq 0$.
We know that $\ell(a) = \ell(b) =: a_{j_\#}$. Assume that $|a'| = |b'|$. Then $\ell(a') = \ell(b') =: a_{j_\flat} \in \mathcal{A}$. Denote $a_{j_\#} + a_{j_\flat} = c$. Then $c \neq 0$ since the set $\mathcal{A}$ is admissible. As $(a, a'), (b, b') \in (\mathcal{L}_f \times \mathcal{L}_f)_+$, then we have $|a_{j_\#} - c| = |a - c| = |b - c|$. As $|a_{j_\#}| = |a| = |b|$, then the three points $a_{j_\#}, a$ and b lie in the intersection of two circles, one centred in the origin and another centred in $c = a_{j_\#} + a_{j_\flat}$. Since $\mathcal{A}$ is strongly admissible, then $a_{j_\#} \angle c$ (see (1.13)). So among the three point two are equal, which is a contradiction. Hence, $|a'| \neq |b'|$ as stated.

The case $\chi^- \neq 0$ is similar. ∎

We claim that this lemma implies that

$$\Lambda_1^1(\rho(\varepsilon)) \not\equiv \Lambda_1^2(\rho(\varepsilon)) \quad \text{for a suitable choice of the vector } x \text{ in (5.4)}\,, \tag{5.19}$$

so (5.3) is valid and Lemma 4.5 holds. To prove (5.19) we consider two cases.

Case 1: $j_\# = 1$. Then $\varphi(a_1^r, a_j^r) = x_{\ell(a_j^r)}$. Denoting $\frac{C_*^2}{\lambda_{a_1}^2} \frac{x^2_{\ell(a_j^r)}}{\lambda^2_{a_j^r}} =: z_{\ell(a_j^r)}$ we see that $k_2(1)$ and $k_2(2)$ are linear functions of the variables $z_{a_1}, \ldots, z_{\ell_n}$.

i) Assume that $\chi^-(a_1^r, a_j^r) = 1$ for some $r \in \{1, 2\}$ and some $j > 1$. Denote $\ell(a_j^r) = a_{j_*}$. Then $j_* \neq j_\#$ and

$$k_2(r) = \frac{z_{a_{j_*}}}{\mu(a_j^r) - \mu(a_1^r)} + \ldots\,,$$

where $\ldots$ is independent from z_{j_*}. Now let $r' = \{1, 2\} \setminus \{r\}$, and find j' such that $\ell(a_{j'}^{r'}) = a_{j_*}$. If such j' does not exist, then $k_2(r')$ does not depend on z_{j_*}. Accordingly, for a suitable x we have $k_2(r) \neq k_2(r')$, and (5.19) holds. If $n_2 = 1$, then $r = 1$ and $r' = 2$. So j' does not exists and (5.19) is established.

If j' exists, then $n_1, n_2 \geq 2$, so the set $\mathcal{A}$ is strongly admissible. By Lemma 5.1 $\chi^-(a_1^{r'}, a_{j'}^{r'}) = 0$ since $\chi^-(a_1^r, a_j^r) = 1$ and

$$|a_1^r| = |a_1^{r'}|, \qquad |a_j^r| = |a_{j'}^{r'}|\,. \tag{5.20}$$

So

$$k_2(r') = z_{j_*} \frac{\chi^+(a_1^{r'}, a_{j'}^{r'})}{\mu(a_j^{r'}) + \mu(a_1^{r'})} + \ldots\,.$$

Since χ^+ equals 1 or 0, then using again (5.20) and the fact that $\mu(a)$ only depends on $|a|$, we see that $k_2(r) \neq k_2(r')$ for a suitable x, so (5.19) again holds.

ii) If $\chi^-(a_1^r, a_j^r) = 0$ for all j and r, then $\chi^+(a_1^r, a_j^r) = 1$ for some r and j. Define z_{j_*} as above. Then the coefficient in $k_2(r)$ in front of z_{j_*} is non-zero, while for $k_2(r')$ it vanishes. This is obvious if $n_{r'} = 1$. Otherwise $\mathcal{A}$ is strongly admissible and it holds by Lemma 5.1 (and since $\chi^- \equiv 0$). So (5.19) again holds.

Case 2: $j_\# \neq 1$. Then by (5.2) *there exists $a_j^1 \in \mathcal{L}_f^r$ such that $\ell(a_j^r) = a_1$. So $\chi^+(a_1^1, a_j^1) \neq 0$ or $\chi^-(a_1^1, a_j^1) \neq 0$.* Then $\varphi(a_1^1, a_j^1) = x_{a_{j_\#}}$, the sum in (5.17) is non-trivial and for the same reason as in Case 1 (5.19) holds.

This completes the proof of Lemma 4.5.

6. KAM

6.1. **An abstract KAM result.** We first recall the abstract KAM theorem from [15], adapting the result and the notation to the present context. Consider a Hamiltonian $H(r, \theta, u, v; \rho)$ of the form (4.56). Denote

$$\mathcal{L}_\infty \cup \mathcal{L}_f^e = \tilde{\mathcal{L}}_\infty, \quad \mathcal{L}_f^h = \tilde{\mathcal{L}}_f, \quad \nu \widehat{K}(\rho) = \mathbf{H}(\rho)\,, \tag{6.1}$$

and re-denote (see(3.45))

$$\nu \Lambda_b(\rho) =: \Lambda_b(\rho)\,, \quad \lambda_a := 0 \qquad \text{if} \quad b \in \mathcal{L}_f^{\mathrm{e}} \ \text{ and } \ a \in \mathcal{L}_f = \mathcal{L}_f^{\mathrm{e}} \cup \mathcal{L}_f^h\,. \tag{6.2}$$

Then the Hamiltonian reads

$$H = \Omega(\rho)\cdot r + \frac{1}{2} \sum_{a \in \tilde{\mathcal{L}}_\infty} \Lambda_a(\rho)(u_a^2 + v_a^2) + \frac{1}{2}\langle \mathbf{H}\tilde{\zeta}_f, \tilde{\zeta}_f\rangle + f(r, \theta, \tilde{\zeta}; \rho)\,, \quad \tilde{\zeta} = (u, v)\,. \tag{6.3}$$

Assume that the parameter ρ belongs to a closed ball in $\mathbb{R}^d$ of a radius at most one, which we denote $\mathcal{D}_0$. The Hamiltonian H is regarded as a perturbation of the quadratic Hamiltonian

$$h = \Omega(\rho)\cdot r + \frac{1}{2} \sum_{a \in \tilde{\mathcal{L}}_\infty} \Lambda_a(\rho)(u_a^2 + v_a^2) + \frac{1}{2}\langle \mathbf{H}(\rho)\tilde{\zeta}_f, \tilde{\zeta}_f\rangle\,.$$

We will assume that h satisfies the following assumptions A1 − A3, depending on constants

$$\delta_0, c, \beta > 0, \quad s_* \in \mathbb{N}\,, \tag{6.4}$$

where c is such that the set $\mathcal{A}$ is contained in the ball $\{|a| \leq c$.

For $a \in \tilde{\mathcal{L}}_f \cup \tilde{\mathcal{L}}_\infty \cup \{\emptyset\}$ we define

$$[a] = \begin{cases} \tilde{\mathcal{L}}_f & \text{if } \ a \in \tilde{\mathcal{L}}_f, \\ \mathcal{L}_f^e \cup \{b \in \mathcal{L}_\infty : |b| \leq c\} & \text{if } \ a \in \mathcal{L}_f^e \ \text{ or } a \in \mathcal{L}_\infty \ \text{ and } |a| \leq c\,, \\ \{b \in \mathcal{L}_\infty \mid |b| = |a|\} & \text{if } \ a \in \mathcal{L}_\infty \ \text{ and } |a| > c\,, \\ \{\emptyset\} & \text{if } \ a = \emptyset\,. \end{cases} \tag{6.5}$$

Hypothesis A1 (spectral asymptotic.) For all $\rho \in \mathcal{D}_0$ we have

(a) $|\Lambda_a| \geq \delta_0 \quad \forall\, a \in \tilde{\mathcal{L}}_\infty$;

(b) $|\Lambda_a - |a|^2| \leq c\langle a\rangle^{-\beta} \quad \forall\, a \in \tilde{\mathcal{L}}_\infty$;

(c) $\|(J\mathbf{H}(\rho))^{-1}\| \le \frac{1}{\delta_0}$, $\|(\Lambda_a(\rho)I - iJ\mathbf{H}(\rho))^{-1}\| \le \frac{1}{\delta_0}$ $\forall\, a \in \tilde{\mathcal{L}}_\infty$;

(d) $|\Lambda_a(\rho) + \Lambda_b(\rho)| \ge \delta_0$ for all $a, b \in \tilde{\mathcal{L}}_\infty$;

(e) $|\Lambda_a(\rho) - \Lambda_b(\rho)| \ge \delta_0$ if $a, b \in \tilde{\mathcal{L}}_\infty$ and $[a] \neq [b]$.

Hypothesis A2 (transversality). For each $k \in \mathbb{Z}^n \setminus \{0\}$ and every vector-function $\Omega'(\rho)$ such that $|\Omega' - \Omega|_{C^{s_*}(\mathcal{D})} \le \delta_0$ there exists a unit vector $\mathfrak{z} = \mathfrak{z}(k) \in \mathbb{R}^n$, satisfying

$$|\partial_{\mathfrak{z}} \langle k, \Omega'(\rho)\rangle| \ge \delta_0 \qquad \forall\, \rho \in \mathcal{D}_0. \tag{6.6}$$

Besides the following properties (i)-(iii) hold for each $k \in \mathbb{Z}^n \setminus \{0\}$:

(i) For any $a, b \in \tilde{\mathcal{L}}_\infty \cup \{\emptyset\}$ such that $(a, b) \neq (\emptyset, \emptyset)$, consider the following operator, acting on the space of $[a] \times [b]$-matrices [14]

$$L(\rho) : X \mapsto (\Omega'(\rho) \cdot k)X \pm Q(\rho)_{[a]}X + XQ(\rho)_{[b]}\,.$$

Here $Q(\rho)_{[a]}$ is the diagonal matrix $\mathrm{diag}\{\Lambda_{a'}(\rho) : a' \in [a]\}$, and $Q(\rho)_{[\emptyset]} = 0$. Then either

$$\|L(\rho)^{-1}\| \le \delta_0^{-1} \qquad \forall\, \rho \in \mathcal{D}_0\,, \tag{6.7}$$

or there exists a unit vector $\mathfrak{z}$ such that

$$|\langle v, \partial_{\mathfrak{z}} L(\rho) v\rangle| \ge \delta_0 \qquad \forall\, \rho \in \mathcal{D}_0\,,$$

for each vector v of unit length.

(ii) Denote $m = 2|\tilde{\mathcal{L}}_f|$ and consider the following operator in $\mathbb{C}^m$, interpreted as a space of row-vectors:

$$L(\rho, \lambda) : X \mapsto (\Omega'(\rho) \cdot k)X + \lambda X + iXJ\mathbf{H}(\rho)\,.$$

Then

$$\|L^{-1}(\rho, \Lambda_a)\| \le \delta_0^{-1} \qquad \forall\, \rho \in \mathcal{D}_0,\ \ a \in \tilde{\mathcal{L}}_\infty\,.$$

(iii) For any $a, b \in \tilde{\mathcal{L}}_f \cup \{\emptyset\}$ such that $(a, b) \neq (\emptyset, \emptyset)$, consider the operator, acting on the space of $[a] \times [b]$-matrices:

$$L(\rho) : X \mapsto (k \cdot \Omega'(\rho))X - iJ\mathbf{H}(\rho)_{[a]}X + iXJ\mathbf{H}(\rho)_{[b]}$$

(the operator $\mathbf{H}(\rho)_{[a]}$ equals $\mathbf{H}$ if $a \in \tilde{\mathcal{L}}_f$ and equals 0 if $a = \emptyset$, and similar with $\mathbf{H}(\rho)_{[b]}$). Then the following alternative holds: either $L(\rho)$ satisfies (6.7), or there exists an integer $1 \le j \le s_*$ such that

$$|\partial_{\mathfrak{z}}^j \det L(\rho)| \ge \delta_0 \|L(\rho)\|^{\dim - 1}_{C^j(\mathcal{D}_0)} \qquad \forall\, \rho \in \mathcal{D}_0\,. \tag{6.8}$$

Here $\dim = (\dim \tilde{\mathcal{L}}_f)^2$ if $a, b \in \tilde{\mathcal{L}}_f$ and $\dim = \dim \tilde{\mathcal{L}}_f$ if a or b is the empty set.

Hypothesis A3 (a Melnikov condition). There exist $\tau > 0$, $\rho_* \in \mathcal{D}_0$ and $C > 0$ such that

$$|\langle k, \Omega(\rho_*)\rangle - (\Lambda_a(\rho_*) - \Lambda_b(\rho_*))| \ge C|k|^{-\tau} \quad \forall\, k \in \mathbb{Z}^n, k \neq 0,\ \text{if } a, b \in \tilde{\mathcal{L}}_\infty \setminus [0] \tag{6.9}$$

(cf. (6.5)).

[14] so if $b = \emptyset$, this is the space $\mathbb{C}^{[a]}$.

Recall that the domains $\mathcal{O}^\gamma(\sigma,\mu)$ and the classes $\mathcal{T}^{\gamma,D}(\sigma,\mu,\mathcal{D}_0)$ were defined at the beginning of Section 3. Denote

$$\chi = |\partial_\rho\Omega(\rho)|_{C^{s_*-1}} + \sup_{a\in\tilde{\mathcal{L}}_\infty} |\partial_\rho\Lambda_a(\rho)|_{C^{s_*-1}} + \|\partial_\rho\mathbf{H}\|_{C^{s_*-1}}\,.$$

Consider the perturbation $f(r,\theta,\zeta;\rho)$ and assume that

$$\varepsilon = [f^T]^{\gamma,D}_{\sigma,\mu,\mathcal{D}_0} < \infty\,,\quad \xi = [f]^{\gamma,D}_{\sigma,\mu,\mathcal{D}_0} < \infty\,,$$

for some $\gamma,\sigma,\mu\in(0,1]$. We are now in position to state the abstract KAM theorem from [15]. More precisely, the result below follows from Corollary 3.7 of that work.

Theorem 6.1. *Assume that Hypotheses A1-A3 hold for $\rho\in\mathcal{D}_0$. Then there exist $\varepsilon_0,\kappa,\bar\beta,C>0$ such that if for a suitable $\aleph>0$ we have*

$$\chi,\xi\ = O(\delta_0^{1-\aleph}) \quad \text{and} \quad \varepsilon\le\varepsilon_0\delta_0^{1+\kappa\aleph} =: \varepsilon_*\,, \tag{6.10}$$

then there is a Borel set $\mathcal{D}'\subset\mathcal{D}_0$ with $meas(\mathcal{D}_0\setminus\mathcal{D}')\le C\varepsilon^{\bar\beta}$, and for all $\rho\in\mathcal{D}'$ the following holds:
There exists a C^{s_*}-smooth mapping

$$\mathfrak{F}:\mathcal{O}^0(\sigma/2,\mu/2)\times\mathcal{D}'\to\mathcal{O}^0(\sigma,\mu)\,,\quad (r,\theta,\tilde\zeta;\rho)\mapsto\mathfrak{F}_\rho(r,\theta,\tilde\zeta)\,,$$

defining for $\rho\in\mathcal{D}'$ real holomorphic symplectomorphisms $\mathfrak{F}_\rho:\mathcal{O}^0(\sigma/2,\mu/2)\to\mathcal{O}^0(\sigma,\mu)$, satisfying for any $x\in\mathcal{O}^0(\sigma/2,\mu/2)$, $\rho\in\mathcal{D}'$ and $|j|\le1$ the estimates

$$\|\partial^j_\rho(\mathfrak{F}_\rho(x)-x)\|_0\le C\frac{\varepsilon}{\varepsilon_*}\,,\qquad \|\partial^j_\rho(d\mathfrak{F}_\rho(x)-I)\|_{0,0}\le C\frac{\varepsilon}{\varepsilon_*}\,, \tag{6.11}$$

such that

$$H\circ\mathfrak{F}_\rho = \tilde\Omega(\rho)\cdot r + \frac12\langle\zeta,A(\rho)\zeta\rangle + g(r,\theta,\zeta;\rho)'. \tag{6.12}$$

Here

$$\partial_\zeta g = \partial_r g = \partial^2_{\zeta\zeta} g = 0 \ \text{ for } \ \zeta=r=0\,, \tag{6.13}$$

$\tilde\Omega=\tilde\Omega(\rho)$ is a new frequency vector satisfying

$$\|\tilde\Omega-\Omega\|_{\mathcal{C}^{s_*}}\le C\delta_0^{1+\aleph}\,, \tag{6.14}$$

and $A:\mathcal{L}\times\mathcal{L}\to\mathcal{M}_{2\times2}(\rho)$ is an infinite real symmetric matrix, belonging to $\mathcal{M}^D_0$. It is of the form $A=A_f\oplus A_\infty$, where

$$\|\partial^\alpha_\rho(A_f(\rho)-\mathbf{H}(\rho))\|\le C\delta_0^{1+\aleph}\,,\quad |\alpha|\le s_*\,. \tag{6.15}$$

The operator A_∞ is such that $(A_\infty)^b_a=0$ if $[a]\ne[b]$ (see (6.5)), and all eigenvalues of the hamiltonian operator JA_∞ are pure imaginary.

The exponent $\bar\beta$ and the constants ε_0 and κ depends only on $\mathcal{A}$, τ, s_* and $\dim\tilde{\mathcal{L}}_f$ (note that they do not depend on the radius of the ball $\mathcal{D}_0$). The constant C does not depend on $\delta_0,\aleph$, nor on the domain $\mathcal{D}_0$.

So for $\rho\in\mathcal{D}'$ the torus $\mathfrak{F}_\rho\big(\{0\}\times\mathbb{T}^n\times\{0\}\big)$ is invariant for the hamiltonian system with the Hamiltonian $H(\cdot;\rho)$ given by (6.3), and the hamiltonian flow on this torus is conjugated by the map $\mathfrak{F}_\rho$ with the linear flow, defined by the Hamiltonian (6.12) on the torus $(\{0\}\times\mathbb{T}^n\times\{0\})$.

Next we show that Theorem 6.1 applies to the Hamiltonian (4.56). To state the corresponding result we recall that in Theorem 4.6, assuming that $m\notin\mathcal{C}$, we put the beam equation to the normal form (4.56) when ρ belongs to the closed domain $\tilde Q(\nu)\subset\mathcal{D}=[c_*,1]^n$ and $0<\nu\le\nu_0$. The domain $\tilde Q(\nu)$ was constructed in Section 4 as the union $\tilde Q(\nu)=\cup_{j=1}^{\mathbb J}\tilde Q_j(\nu)$, where $\mathbb J$ does not depend on the small parameter ν and any domain $\tilde Q_j(\nu)$ lies in the corresponding connected component Q_j of the set $\mathcal D\setminus\tilde X$. The domains $\tilde Q_j(\nu)$ grow when ν decays and satisfy (4.55).

Let us define

$$\mathcal D_\nu=\begin{cases}\tilde Q(\nu) & \text{if } \mathcal A \text{ is strongly admissible,}\\ \tilde Q(\nu)\cap\mathcal D_0^1 & \text{if not}\end{cases}\tag{6.16}$$

(see (4.13)). We notice that $\mathcal D_\nu\subset\mathcal D_{\nu'}$ for $\nu\ge\nu'$.

Theorem 6.2. *There exists a zero-measure Borel set $\mathcal C\subset[1,2]$ such that for any admissible set $\mathcal A$ and any $m\notin\mathcal C$, may be found a real number $c_{**}\in(0,1/2]$, depending on $g(\cdot)$, $\mathcal A$ and m, such that for any $c_*\in(0,\frac12 c_{**}]$ there exist $\beta_{*0},\nu_0\in(0,1]$, depending on $g(\cdot)$, $\mathcal A$, m and c_*, with the following property:*
For any closed ball $\mathcal D_0\subset\mathcal D_{\nu_0}\subset\mathcal D=[c_*,1]^n$ and any $\nu\le\nu_0,\beta_*\le\beta_{*0}$ there exist a Borel set $\tilde{\mathcal D}_{0\nu}\subset\nu\mathcal D_0\subset\nu\mathcal D=[c_*\nu,\nu]^n$, depending on $\nu,g(\cdot),c_*,\mathcal A,m,\mathcal D_0$ and satisfying

$$\operatorname{meas}(\mathcal D_0\setminus\tilde{\mathcal D}_{0\nu})\le C\nu^{\bar\beta+n}\,,\qquad \bar\beta>0\,,\tag{6.17}$$

a C^1-mapping $U:\ \mathbb T^n\times\tilde{\mathcal D}_{0\nu}\to Y_0^{\mathbb Z^d\,R}=Y_0^{\mathcal A\,R}\times Y_0^{\mathcal L\,R}$, analytic in the first argument and satisfying

$$\left\|U(\theta,I)-(\sqrt I e^{i\theta},\sqrt I e^{-i\theta},0)\right\|_0\le C\nu^{1-c_1\beta_*}\,,\tag{6.18}$$

and a C^1-smooth vector function $\omega':\tilde{\mathcal D}_{0\nu}\to\mathbb R^n$, satisfying

$$\omega'(I)=\omega+M\,I+O(|I|^{1+c_2\beta_*})\,,\tag{6.19}$$

where the matrix M is given in (3.44), such that

i) for $I\in\tilde{\mathcal D}_{0\nu}$ and $\theta\in\mathbb T^n$ the curve

$$t\mapsto U(\theta+t\omega'(I),I)\tag{6.20}$$

is a solution of the beam equation (1.8). Accordingly, for each $I\in\tilde{\mathcal D}_{0\nu}$ the analytic torus $U(\mathbb T^n\times\{I\})$ is invariant for equation (1.8).

ii) The solution (6.20) is linearly stable if and only if the operator $\widehat K(\rho)$ in the normal form (4.56) is trivial (equivalently if the hamiltonian operator $iJK(\rho)$, corresponding to the Hamiltonian (3.47), is elliptic), and this stability does not depend on $\rho\in\tilde{\mathcal D}_{0\nu}$. In particular, this happens if in (6.16) $|\mathcal A|=1$ or $d=1$, or if $\rho\in\tilde Q\cap\mathcal D_0^1$.
The constants c_1, c_2 and C and the exponent $\bar\beta$ depend on $\mathcal A$, m,c_* and $g(\cdot)$, but not on the ball $\mathcal D_0\subset\mathcal D_{\nu_0}$.

6.2. **Proof of Theorem 6.2.** In this section we denote by C,C_1 etc and c,c_1 etc various constants, depending only on $g(\cdot),m,\mathcal A$ (not on $\mathcal D_0$ or ν).
First we verify that the Hypotheses A1-A3 of Theorem 6.1 are satisfied uniformly for $\rho\in\mathcal D_\nu$ (see (6.16)) and $\nu\le\nu_0$, where ν_0 is defined in Theorem 4.6:

Proposition 6.3. *There exists a zero-measure Borel set $\mathcal C\subset[1,2]$ such that for any admissible set $\mathcal A$ and any $m\notin\mathcal C$, may be found a real number $c_{**}\in(0,1/2]$, depending on $g(\cdot)$, $\mathcal A$ and m, such that for any $c_*\in(0,\frac12 c_{**}]$ there exists $\nu_0,\beta_{*0}\in(0,1]$, depending on $g(\cdot)$, $\mathcal A$, m and c_*, with the following property:*

For $0<\nu\le\nu_0$ and $0<\beta_*\le\beta_{*0}$ the system (4.56) *with the notation* (6.1), (6.2) *satisfies the Hypotheses A1-A3 of Theorem 6.1 for all $\rho\in\mathcal{D}_\nu\subset[c_*,1]^n$, where*

$$\delta_0=\nu^{1+\beta_*\alpha}\,,\quad c=2\max\{\langle a\rangle^3,a\in\mathcal{A}\},\quad \beta=2\,,\quad s_*=4\,|\tilde{\mathcal{L}}_f|^2 \tag{6.21}$$

with some

$$\alpha>(\bar c+\beta(0))\,. \tag{6.22}$$

The constant $\bar c$ and $\beta(0)$ are defined in Theorem 4.6, they depend only on m, $g(\cdot)$ and $\mathcal{A}$.

Proof. We will check the validity of the three hypotheses.
First we note that using (3.45), (2.4), (4.12), (4.33) and (4.52) we get

$$\tfrac12+\tfrac12|a|^2\le\Lambda_a\le 2|a|^2+1\,,\quad |\Lambda_a-\lambda_a|_{C^s(\mathcal{D}_0)}\le C_3\nu|a|^{-2}\qquad\forall\, s\,,\ \forall\, a\in\mathcal{L}_\infty\,, \tag{6.23}$$

$$C_1\nu^{1+\bar c\beta_*}\le|\Lambda_a|\le C_2\nu\qquad\forall\, a\in\mathcal{L}_f^e \tag{6.24}$$

(we recall (6.2)). The function $\Omega(\rho)\in\mathbb{R}^n$ is defined in (3.44), so

$$\Omega(\rho)=\omega+\nu M\rho,\qquad \det M\neq 0\,, \tag{6.25}$$

and $\mathbf{H}$ is a symmetric real linear operator in the space Y^f (cf (3.46)). Its norm satisfies

$$\|\mathbf{H}(\rho)\|\le C\nu^{1-\beta_*\beta'(0)}\,, \tag{6.26}$$

see Theorem 4.6.ii).
Hypothesis A1. Relations (a) and (b) immediately follow from (6.23), (6.24) and (6.22)

To prove (c) note that the operator $\widehat{\mathbf{U}}$ conjugates $J\mathbf{H}$ and $(\Lambda_a I-iJ\mathbf{H})$ with diagonal operators with the eigenvalues $\pm i\Lambda_j^h(\rho)$ and $\Lambda_a(\rho)\pm\Lambda_j^h(\rho)$, respectively. By (6.24) and (4.59) the norms of the eigenvalues are $\ge C^{-1}\nu^{1+\bar c\beta_*}$. Since the norms of $\widehat{\mathbf{U}}$ and its inverse are bounded by (4.60), then the required estimate follows by (6.22).

Condition (e) follows from (6.23), (6.24) and (6.5).

Now consider (d).[15] If $a\in\mathcal{L}$ and $b\in\mathcal{L}_\infty$, then again the relation follows from (6.23) and (6.24). Next, let $a,b\in\mathcal{L}_f$. Let us write Λ_a and Λ_b as Λ_r^j and Λ_m^k, $j\le k$. If $j=k$, then the condition follows from (4.35), (4.52) (from (4.33) if $m=r$). If $j\le M_*<k$, then it again follows from (4.35). If $j,k\le M_*$, then $\Lambda_r^j=\Lambda_1^j=\mu(b_j,\rho)$ and $\Lambda_m^k=\mu(b_m,\rho)$, so the relation follows from (4.34). Finally, let $j,k>M_*$. Then if the set $\mathcal{A}$ is strongly admissible, the required relation follows from (4.35), while if $\rho\in\mathcal{D}_0^1$, then it follows from (4.19).

Hypothesis A2. By (6.25), $\partial_{\mathfrak{z}}\Omega(\rho)=\nu M\mathfrak{z}$. Choosing

$$\mathfrak{z}=\frac{{}^tMk}{|{}^tMk|} \tag{6.27}$$

and using that $|\Omega'-\Omega|_{C^{s_*}}\le\delta_0$ we achieve that $\partial_{\mathfrak{z}}\langle k,\Omega'(\rho)\rangle\ge C\nu$, so (6.6) holds.

To verify (i) we restrict ourselves to the more complicated case when $a,b\neq\emptyset$. Then $L(\rho)$ is a diagonal operator with the eigenvalues

$$\lambda_{ab}^k:=\langle k,\Omega'(\rho)\rangle+\Lambda_a(\rho)\pm\Lambda_b(\rho)\quad a\in[a],\ b\in[b]\,.$$

[15] This is the only condition of Theorem 6.1 which we cannot verify for any domain $\tilde D_j$ without assuming that the set $\mathcal{A}$ is strongly admissible.

Clearly

$$|\lambda^k_{ab} - (\langle k,\omega\rangle + \lambda_a \pm \lambda_b)| \le C\nu|k|$$

(we recall (6.2)). Therefore by Propositions 2.6 and 2.7 the first alternative in (i) holds, unless

$$|k| \ge C\nu^{-c} \tag{6.28}$$

for some (fixed) $c > 0$. But if we choose $\mathfrak{z}$ as in (6.27), then $\partial_{\mathfrak{z}} L(\rho)$ becomes a diagonal matrix with the diagonal elements bigger than $|{}^tMk| - C\nu|k| - C_1\nu$. So if k satisfies (6.28), then the second alternative in (i) holds.

To verify (ii) we write $L(\rho,\Lambda_a)$ as

$$L = (\langle k,\Omega'\rangle + \Lambda_a(\rho))I + i\nu JH^h_0\,.$$

The transformation $\widehat{\mathbf{U}}$ conjugates L with the diagonal operator with the eigenvalues $\lambda^k_{aj} =: \langle k,\Omega'\rangle + \Lambda_a(\rho) \pm \nu i\Lambda^h_j$. In view of (4.59), $|\lambda^k_{aj}| \ge |\Im\lambda^k_{aj}| \ge C^{-1}\nu^{1+\bar c\beta_*}$. This implies (ii) by (6.22).

It remains to verify (iii). As before, we restrict ourselves to the more complicated case $a,b \in \tilde{\mathcal{L}}_f$. Let us denote

$$\lambda(\rho) := \langle k,\Omega'(\rho)\rangle = \langle k,\omega\rangle + \nu\langle k, M\rho\rangle + \langle k,(\Omega'-\Omega)(\rho)\rangle\,,$$

and write the operator $L(\rho)$ as

$$L(\rho) = \lambda(\rho)I + L^0(\rho)\,,\quad L^0(\rho)X = [X, iJ\mathbf{H}(\rho)]\,.$$

In view of (6.26),

$$\|L^0\|_{C^j} \le C_j\nu^{1-c_j\beta_*} \qquad \text{for } j \ge 0\,. \tag{6.29}$$

Now it is easy to see that if $|\langle k,\omega|\rangle \ge C(\nu^{1-c_0\beta_*} + \nu|k|)$ with a sufficiently big C, then the first alternative in (iii) holds.

So it remains to consider the case when

$$|\langle k,\omega\rangle| \le C(\nu^{1-c_0\beta_*} + \nu|k|)\,. \tag{6.30}$$

By Proposition 2.6 the l.h.s. is bigger than $\kappa|k|^{-n^2}$. Assuming that $\beta_{*0} \ll 1$, we derive from this and (6.30) that

$$|k| \ge C\nu^{-1/(1+n^2)}\,. \tag{6.31}$$

In view of (6.29)-(6.31), again if $\beta_{*0} \ll 1$, we have:

$$|\lambda(\rho)| \le C\nu(\nu^{-c_0\beta_*} + |k|) \le C_1\nu|k|\,, \tag{6.32}$$

$$|(\partial_\rho)^j\lambda(\rho)| \le C_j|k|\delta_0, \qquad 1 \le j \le s_*\,, \tag{6.33}$$

$$\|L\|_{C^j} \le C\nu(\nu^{-c_j\beta_*} + |k|) + C_j|k|\delta_0\,. \tag{6.34}$$

Denote $\det L(\rho) = D(\rho)$. Then

$$D(\rho) = \prod_{j,k\in\mathcal{L}^h_f}\prod_{\sigma_1,\sigma_2=\pm}(\lambda(\rho) + \sigma_1\nu\Lambda^h_j(\rho) - \sigma_2\nu\Lambda^h_k(\rho))\,.$$

Choosing $\mathfrak{z}$ as in (6.27) we get

$$\partial_{\mathfrak{z}}\lambda(\rho) \ge C^{-1}|k|\nu - |k|\delta_0 \ge \frac{1}{2}C^{-1}|k|\nu\,.$$

In view of (6.32), (6.33) and (6.29) this implies that

$$|\partial_{\mathfrak{z}}\lambda| \gtrsim |\lambda(\rho)|, \quad |\partial_{\mathfrak{z}}\lambda| \gg |(\partial_\rho)^j\lambda(\rho)|, \quad |\partial_{\mathfrak{z}}\lambda| \gg |(\partial_\rho)^j L^0|\,.$$

Let us denote $2\,|\tilde{\mathcal{L}}_f| = m$; then $s_* = m^2$. Chose in (6.8) $j = s_* = m^2$. Then $\partial_{\mathfrak{z}}^{s_*}D(\rho)$ is a small perturbation of $(\partial_{\mathfrak{z}}\lambda(\rho))^{m^2}$ since the latter is the leading term of the former: all other terms, forming $\partial_{\mathfrak{z}}^{s_*}D(\rho)$, are much smaller. So we get that

$$|\partial_{\mathfrak{z}}^{s_*}D(\rho)| \ge C_1^{-1}(|k|\nu)^{m^2}\,.$$

In the same time, in view of (6.34) the r.h.s. of (6.8) is bounded from above by

$$C_m\delta_0(\nu^{(m^2-1)(1-c_j\beta_*)} + \nu^{m^2-1}|k|^{m^2-1})\,.$$

This implies (6.8), if

$$(|k|\nu)^{m^2}\nu^{-1-\beta_*\alpha} \ge C_1C_m((|k|\nu)^{m^2-1} + \nu^{(m^2-1)(1-c_j\beta_*)})$$

which is achieved as soon as $\beta_{*0} \le \frac{m^2}{c_j(1+m^2)(m^2-1)}$ since α is positive.

Hypothesis A3. The required inequality follows from Proposition 2.7 since the divisor, corresponding to (6.9) where $a, b \notin \mathcal{L}_f$, cannot be resonant. □

Now we will use Proposition 6.3 to derive Theorem 6.2 from Theorem 6.1.

Let us take γ_*, μ, $\sigma = \frac{1}{2}$ and $\hat{c} > 0$ as in the Theorem 4.6 (see also (4.53)), and take $\alpha, \beta_{*0}, \nu_0 > 0$ and $\delta_0 = \nu^{1+\beta_*\alpha}$ as in Proposition 6.3. Since $\nu \le \nu_0$, then $\mathcal{D}_0 \subset \mathcal{D}_{\nu_0} \subset \mathcal{D}_\nu$ and the proposition applies. So the Hypotheses A1-A3 of Theorem 6.1 are fulfilled, and to show that the theorem is applicable to the Hamiltonian (4.56) with $\rho \in \mathcal{D}_0$ we have to verify that the quantities χ, ξ and ε meet the relation (6.10). To do this let us write the estimates (4.57) as

$$\xi := [\tilde{f}]^{\gamma,D}_{\sigma,\mu,\mathcal{D}_0} \le C\nu^{1-\hat{c}\beta_*}\,, \quad \varepsilon := [\tilde{f}^T]^{\gamma,D}_{\sigma,\mu,\mathcal{D}_0} \le C\nu^{3/2-\hat{c}\beta_*}\,,$$

and note that trivially $\chi \le C\nu^{1-\beta_*\beta'(s_*)}$. This implies (6.10), written as,

$$\chi \le C\delta_0^{1-\aleph}\,, \quad \xi \le C\delta_0^{1-\aleph}\,, \qquad \varepsilon \le \varepsilon_0\delta_0^{1+\kappa\aleph} =: \varepsilon_*\,,$$

if ν_0 is sufficiently small and

$$\frac{\alpha+\beta'(s_*)}{1+\beta_*\alpha}\beta_* \le \aleph\,, \quad \frac{\alpha+\hat{c}}{1+\beta_*\alpha}\beta_* \le \aleph\,, \quad \aleph < \frac{1/2-\beta_*\alpha-\hat{c}\beta_*}{\kappa(1+\beta_*\alpha)}\,. \tag{6.35}$$

Let us chose

$$\alpha = 2(\bar{c}+\beta(0))\,, \qquad \aleph = 4n\beta_*\big(\alpha+\hat{c}+\beta'(s_*)+\beta'(0)\big) = \beta_*\tilde{c}\,. \tag{6.36}$$

Then (6.22) and (6.35) hold if β_{*0} is sufficiently small, and the relations (6.10) are fulfilled.

Now we apply Theorem 6.1 to the Hamiltonian $\tilde{H}_\rho$ with $\rho \in \mathcal{D}_0$. We get that there exist positive constants $\bar{\beta}$, C_1 with the property that for $0 < \beta_* \le \beta_{*0}$ and $0 < \nu \le \nu_0$ there exist a Borel set $\mathcal{D}' \subset \mathcal{D}_0$, satisfying

$$\mathrm{meas}(\mathcal{D}_0 \setminus \mathcal{D}') \le C_1\varepsilon^{\bar{\beta}} \le C_1\nu^{(3/2-\hat{c}\beta_*)\bar{\beta}} \le C\nu^{\bar{\beta}},$$

and a C^{s_*}-smooth mapping $\mathfrak{F} : \mathcal{O}^0(1/4, \mu/2)\times\mathcal{D}' \to \mathcal{O}^0(1/2, \mu)$ such that for $\rho \in \mathcal{D}'$

$$H_\rho \circ \mathfrak{F}_\rho = \omega'(\rho)\cdot r + \frac{1}{2}\langle\zeta, A(\rho)\zeta\rangle + g(r,\theta,\zeta;\rho), \tag{6.37}$$

where $\partial_\zeta g = \partial_r g = \partial^2_{\zeta\zeta} g = 0$ for $\zeta = r = 0$. As a consequence, the torus $\mathfrak{F}_\rho(\{0\}\times\mathbb{T}^n\times\{0\})$ is invariant for the Hamiltonian H_ρ, and the the mapping $\mathfrak{F}_\rho$ linearises the hamiltonian flow on this torus, i.e. solutions of the hamiltonian equation on the torus read $\mathfrak{F}_\rho(0,\theta_0+t\tilde\Omega(\rho),0)$, $\theta_0\in\mathbb{T}^n$. So, setting $\tilde{\mathcal{D}}_{\nu 0}:=\nu\mathcal{D}'$, $\tilde{\mathcal{D}}_{\nu 0}=\{I\}$, we have $\mathrm{meas}(\nu\mathcal{D}_0\setminus\tilde{\mathcal{D}}_{\nu 0})\le C_1\nu^{\bar\beta+n}$, and defining the mapping U as

$$U:\mathbb{T}^n\times\tilde{\mathcal{D}}_{\nu 0}\to Y_0^{\mathbb{Z}^d R},\quad U(\theta,I)=\tilde\Phi_\rho\circ\mathfrak{F}_\rho(0,\theta,0)\quad\text{with }\rho=\nu^{-1}I\,,$$

where $\tilde\Phi_\rho$ is the mapping from Theorem 4.6, we obtain that the curve (6.20) is a solution of (1.8). As $\mathfrak{F}_\rho$ is close to the identity (see (6.11)), we deduce that

$$\mathrm{dist}(\mathfrak{F}_\rho(0,\cdot,0)(\mathbb{T}^n),T_I^n)\le C\frac{\varepsilon}{\varepsilon_*}\le C\nu^{\frac32-\hat c\beta_*-(1+\beta_*\alpha)(1+2\kappa\aleph)}\le C\nu^{\frac12-c_\flat\beta_*}\tag{6.38}$$

with $c_\flat=(2+3\kappa)\tilde c$. In particular, the torus $\mathfrak{F}_\rho(0,\cdot,0)(\mathbb{T}^n)$ lies in the domain $\mathbf{T}_\rho(\nu,1,1,0)$. In view of (6.13) it is invariant for the beam equation (1.8). Recall that

$$\tilde\Phi_\rho=\Phi_\rho\circ\Sigma\circ(\mathbf{U}^{-1}\oplus\mathrm{id})\,,$$

where $\mathbf{U}^{-1}\oplus\mathrm{id}$ only moves the $\mathcal{L}_f$-variables and satisfies (4.46), Σ is the change from the complex to the real variables and Φ_ρ satisfies (3.39). Combining (3.39), (4.46) and (6.38) we get that

$$\left\|U(I,\theta)^{\mathcal{A}}-(\sqrt{I}e^{i\theta},\sqrt{I}e^{-i\theta})\right\|\le C\,\nu^{\frac12}\ \nu^{-\hat c\beta_*}\ \nu^{\frac12-\hat c_\flat\beta_*}=C\,\nu^{1-(\hat c+c_\flat)\beta_*}$$

where $U(I,\theta)=(U(I,\theta)^{\mathcal{A}},U(I,\theta)^{\mathcal{L}})\in Y_0^{\mathbb{Z}^dR}=Y_0^{\mathcal{A}\,R}\times Y_0^{\mathcal{L}\,R}$, and $(\sqrt{I}e^{i\theta},\sqrt{I}e^{-i\theta})$ stands for the vector $\big((\xi_a,\eta_a),a\in\mathcal{A}\big),\xi_a\equiv\bar\eta_a$, as in (3.9).

Similarly we verify using (3.38), (4.46) and (6.11) that

$$\left\|U(I,\theta)^{\mathcal{L}}\right\|_0\le C\,\nu^{1-(\hat c+c_\flat)\beta_*}\,.$$

This proves (6.18) with $c_1=\hat c+c_\flat$. On the other hand (6.14) leads to

$$\|\Omega-\Omega'\|_{\mathcal{C}^1}\le C\nu^{(1+\beta_*\alpha)(1+\aleph)}\le C\nu^{1+c_2\beta_*}$$

with $c_2=\tilde c+2(\bar c+\beta(0))$. Thus defining $\omega'(I)=\Omega'(\nu^{-1}I)$ and using (3.44) we get

$$|\omega'(I)-\omega-M\,I|\le C\nu^{1+c_2\beta_*}.$$

Finally, since in Theorem 6.1 the infinite real symmetric matrix A is of the form $A=A_f\oplus A_\infty$, where

$$A_f(\rho)=\nu\widehat K(\rho)+O(\delta_0^{1+\aleph})=\nu\widehat K(\rho)+O(\nu^{(1+\beta_*\alpha)(1+\aleph)})$$

and A_∞ is a block–diagonal matrix such that all eigenvalues of the hamiltonian operator JA_∞ are pure imaginary (see (6.15)), then the linear stability of the constructed invariant torus $U(\mathbb{T}^n\times\{I\})$ is determined by the stability of the matrix $A_f(\rho)$, $\rho=\nu^{-1}I$. If the set $\mathcal{L}_f^h$ is not trivial, then by (4.59) the operator $J\widehat K$ admits an eigenvalue whose real part is larger than $\nu^{\bar c\beta_*}$. Let us denote $\nu^{\beta_*\beta'(0)}J\widehat K=L_1$, $\nu^{-1+\beta_*\beta_0'}JA_f=L_2$ ($\beta'(0)$ is defined in (4.61)). Then

i) $\|L_1\|\le C$ by (4.61),

ii) L_1 has an eigenvalues whose real part is $\ge\nu^{\beta_*(\bar c+\beta'(0))}$,

iii) $\|L_1-L_2\|\le C\nu^{-1+\beta_*\beta'(0)}\delta_0^{1+\aleph}\le C\nu^{\aleph}$.

By i), iii) and Lemma C.2 the distance between the spectra of the operators L_1 and L_2 is bounded by $C\nu^{\aleph/2n}\le C\nu^{2\beta^*(\hat c+\beta'(0))}$ where we used (6.36). Then in view of ii) the operator L_2 has an eigenvalue with a nontrivial real part. Accordingly, the hamiltonian operator JA_f is unstable.

6.3. **Proofs of Theorems 1.5 and 1.4.** Everywhere in this section "ball" means "closed ball". Let us fix $\beta_* > 0$ small enough so that Theorem 6.2 applies on any ball $\mathcal{D}_0 \subset \mathcal{D}_\nu$ (see (6.16)) with $\nu \le \nu_0$.

We start with a construction which allows to apply Theorem 6.2 to prove the two theorems. For any $\gamma > 0$ let us find $\nu' \equiv \nu'(\beta_*) > 0$ so small that

$$\sum_{j=1}^{\mathbb{J}} \operatorname{meas}(Q_j \setminus \tilde{Q}_j(\nu')) \le \frac{1}{4}\gamma \sum \operatorname{meas} Q_j = \frac{1}{4}\gamma \operatorname{meas} \mathcal{D} \tag{6.39}$$

(see (4.29) and (4.40)). Since each $\tilde{Q}_\nu$ is a component of a closed semi-analytic set, then its interior $\operatorname{Int} \tilde{Q}_\nu$ has the same measure as the set itself. By the Vitali theorem we can find a countable family of non-intersecting balls in $\operatorname{Int} \tilde{Q}_\nu$ such that their union feel up this domain up to a zero-measure set. Therefore in $\tilde{Q}_j(\nu')$ exist $N_j = N_j(\gamma)$ non-intersecting balls $B_j^1, \dots, B_j^{N_j}$, $B_j^r = B_j^r(\gamma)$, such that

$$\operatorname{meas}\big(\tilde{Q}_j(\nu') \setminus \cup_{r=1}^{N_j} B_j^r\big) \le \frac{1}{4\,\mathbb{J}}\gamma \operatorname{meas}(Q_j)\,, \qquad j = 1, \dots, \mathbb{J}\,. \tag{6.40}$$

Note that $\cup_{r=1}^{N_j} B_j^r \subset \tilde{Q}_j(\nu)$ if $\nu \le \nu'$.

Now to each ball B_j^r and every $\nu \le \nu'$ we apply Theorem 6.2 to construct a set $(B_j^r)'(\nu)$, corresponding to the tori, persisting in the perturbed equation, and for each j find $\nu_j \in (0, \nu']$ such that

$$\operatorname{meas}\big(\cup_{r=1}^{N_j} B_j^r(\gamma) \setminus \cup_{r=1}^{N_j} (B_j^r)'(\gamma,\nu)\big) \le \frac{1}{4\,\mathbb{J}}\gamma \operatorname{meas}(Q_j) \quad \text{if} \quad \nu \le \nu_j\,. \tag{6.41}$$

This ν_j depends on N_j.

Proof of Theorem 1.5. Let us consider the domains $Q_l \subset \mathcal{D}_0^1$. They correspond to $l \le \mathbb{J}_1$, see (4.30), and feel in $\mathcal{D}_0^1$ up to a set of zero measure. Since $c_* \le \frac{1}{2}c_{**} \le 1/4$, then

$$\mathcal{D}_0^1 \subset \{\rho \mid \tfrac{1}{2} \le \|\rho\| \le 2\}\,, \qquad \operatorname{meas} \mathcal{D}_0^1 \ge \tfrac{1}{2} c_{**}^n > 0\,. \tag{6.42}$$

Next in the construction above we choose $\gamma = 1/2$, find the corresponding $\nu_0 = \min(\nu_1, \cdots, \nu_{J_1}) > 0$ and for $\nu \le \nu_0$ construct the sets $(B_r^l)'(1/2, \nu)$, $l \le \mathbb{J}_1, r \le N_l$. Denote by $\mathcal{B}(\nu)$ their union, and denote $\mathcal{D}_\nu = \nu\mathcal{B}(\nu) \subset [0,\nu]^n$. Now we set

$$\mathfrak{J} = \cup_{j=0}^{\infty} \mathcal{D}_{\nu^{(j)}}\,, \qquad \nu^{(j)} = 5^{-j}\nu_0\,.$$

In view of (6.42), $\mathfrak{J}$ is a disjoint union of Borel sets. Since by (6.39)-(6.41) $\operatorname{meas}(\mathcal{D}_0^1 \setminus \mathcal{B}(\nu^{(j)}) \le \frac{3}{8} \operatorname{meas} \mathcal{D}_0^1$ and $\mathfrak{J} \cap [0, \nu^{(j)}]^n \supset \mathcal{D}_{\nu^{(j)}} = \nu^{(j)}\mathcal{B}(\nu^{(j)})$, we see that the $\liminf$ in (1.23) is $\ge \frac{3}{8} \operatorname{meas} \mathcal{D}_0^1$. So $\mathfrak{J}$ has a positive density at the origin.

Now we define the mapping $U : \mathbb{T}^n \times \mathfrak{J} \to Y^R$ by the relation

$$U\,|_{\nu^{(j)}(B_1^r)'(1/2,\nu^{(j)})} = U^{\nu^{(j)},(B_1^r)'(1/2,\nu^{(j)})} \qquad \forall\, r \le N_1,\ \ j \ge 0\,,$$

where $U^{\nu^{(j)},(B_1^r)'(1/2,\nu^{(j)})}$ is the mapping from Theorem 6.2, corresponding to $\nu = \nu^{(j)}$ and $\mathcal{D}_0 = B_1^r$, and define the mapping $\omega' : \mathfrak{J} \to \mathbb{R}^n$ similarly.

These maps obviously are continuous. Since for any vector $I \in \mathcal{D}_{\nu^{(j)}} \subset \mathfrak{J}$ the norm of I is equivalent to $\nu^{(j)}$, then by Theorem 6.2 the maps satisfy all assertions of Theorem 1.5, apart from those related to the triviality of the hyperbolic operator JK. If $d = 1$ or $|\mathcal{A}| = 1$, then $JK = 0$ by Examples 4.1 and 4.2. Examples of non-trivial operators JK (when $d \ge 2$ and $|\mathcal{A}| \ge 2$) are given in Appendix B. So Theorem 1.5 is proved.

Proof of Theorem 1.4. Since now the set $\mathcal{A}$ is strongly admissible, then Theorem 6.2 applies to every ball in every domain $\tilde{Q}_j(\nu)$, $j \le \mathbb{J}$. For any $\gamma > 0$ let us chose $c_* = c_*(\gamma)$ such that $\text{meas}([0,1]^n \setminus \mathcal{D}) \le \frac{1}{4}\gamma$. Next for each $j \le \mathbb{J}$ we find $\nu_j = \nu_j(\gamma)$ and the collection of balls $\{B_j^r(\gamma), r \le N_j\}$ and sets $\{(B_j^r)'(\gamma,\nu), r \le N_j, \nu \le \nu_j\}$ as in (6.40), (6.41). Denote

(6.43)

$$\nu(\gamma) = \min\{\nu_j(\gamma), j \le \mathbb{J}\}\,, \qquad \mathcal{B}(\gamma,\nu) = \cup_{j=1}^{\mathbb{J}} \cup_{r=1}^{N_j} (B_j^r)'(\gamma,\nu)\,, \quad 0 < \nu \le \nu(\gamma)\,.$$

Note that $\nu : (0,1] \to (0,1]$, $\gamma \mapsto \nu(\gamma)$, is a non-increasing function which goes to zero with γ. From (6.39), (6.40), (6.41) we have

$$\text{meas}([0,1]^n \setminus \mathcal{B}(\gamma,\nu)) \le \frac{1}{4}\gamma + \sum_{j=1}^{\mathbb{J}} \text{meas}(\tilde{Q}_j \setminus \cup_r (B_j^r)'(\gamma,\nu)) \le \frac{3}{4}\gamma\,, \tag{6.44}$$

for any $\nu \le \nu(\gamma)$.

Let $\nu_k = 2^{-k}$, $k \ge 1$, and let $\{\gamma_k\}$ be a non-increasing sequence of positive numbers, converging to zero so slowly that $\nu(\gamma_k) \ge \nu_k$. For $k \ge 1$ denote

$$K_k = [0,\nu_k]^n\,, \qquad \Gamma_k = \partial K_{k+1} \cap (0,\nu_k)^n$$

(Γ_k lies in the interior of K_k). Let O_k be an ε-vicinity of Γ_k in K_k ($\varepsilon > 0$) so small that $\text{meas}\, O_k \le \frac{1}{4}\gamma_k \nu_k^n$. For every k let $\mathcal{B}_k = \nu_k \mathcal{B}(\gamma_k,\nu_k) \subset K_k$. Then

$$\text{meas}(K_k \setminus \mathcal{B}_k) \le \frac{3}{4}\gamma_k \nu_k^n\,. \tag{6.45}$$

Finally, we set

$$\mathfrak{J}(m,\mathcal{A}) = \cup_{k=1}^{\infty} \big(\mathcal{B}_k \setminus (O_k \cup K_{k+1})\big)\,.$$

This is a disjoint union of Borel sets, and we derive from (6.45) that $\mathfrak{J}(m,\mathcal{A})$ has density one at the origin.

To construct the mappings $U : \mathbb{T}^n \times \mathfrak{J}(m,\mathcal{A}) \to Y^R$ and $\omega' : \mathfrak{J}(m,\mathcal{A}) \to \mathbb{R}^n$, for each $k \ge 1$ we define them on the sets $\nu_k (B_j^r)'(\gamma_k,\nu_k)$, forming the set $\mathcal{B}_k$, using Theorem 6.2. We do this exactly as above, when proving Theorem 1.5. Next we restrict the maps to the sets $\mathcal{B}_k \setminus (O_k \cup K_{k+1})$, forming $\mathfrak{J}(m,\mathcal{A})$. Our construction implies that the mappings are continuous. The estimates (1.24) and (1.25) with suitable constants C, c follow from Theorem 6.2, if we note that for $I \in K_k \setminus K_{k+1}$ the norm of I is equivalent to ν_k, provided that $\gamma_k \to 0$ sufficiently slow.

The analysis of the linear stability of the constructed solutions is the same as before. This proves Theorem 1.4.

Remark 6.4. Let $\mathcal{A} \subset \mathbb{Z}^d$ be an admissible set. The Hypothesis A1(d) is the only assumption of Theorem 6.1 which we cannot verify for the Hamiltonian (4.25) and all domains $\tilde{Q}_j$ if $d \ge 3$. Accordingly, if under the assumptions of Theorem 1.5 the Hypothesis A1(d) holds for all $\rho \in \tilde{Q}_j$ and every j, then the assertion of Theorem 1.4 is true for this $\mathcal{A}$. Similar, if the Hypothesis A1(d) holds for some domain $\tilde{Q}_j$, then the assertion of Theorem 6.2 is valid for any ball $\mathcal{D}_0 \subset \tilde{Q}_j$.

If for $\rho \in Q_j$ the operator $\widehat{K}$ is non-trivial and the assertion of Theorem 6.2 holds for balls $\mathcal{D}_0 \subset\in Q_j$ (e.g. the set $\mathcal{A}$ is strongly admissible), then the constructed KAM-solutions are linearly unstable. In this case the set $\mathfrak{J}$ in Theorem 1.5 contains a subset $\mathfrak{J}_h$ (corresponding to the scaling of the component $\tilde{Q}_j$), having positive density at the origin, filled in with linearly unstable KAM-solutions.

7. Conclusions.

The set $\mathfrak{A}(d,\mathcal{A})=U(\mathbb{T}^n\times\mathfrak{J})$, where $\mathcal{A}\subset\mathbb{Z}^d$ is an admissible set and U and $\mathfrak{J}$ are constructed in Theorems 1.5, 1.4, is invariant for the beam equation (written in the form (1.8)), and is filled in with its time-quasiperiodic solutions. The assertion ii) of Theorem 1.5 implies that the Hausdorff dimension of this set equals $2n$. Now let

$$\mathfrak{A}=\cup_{\mathcal{A}\subset\mathbb{Z}^k}\mathfrak{A}(d,\mathcal{A})\,.$$

This set is formed by time-quasiperiodic solutions of (1.1) and has infinite Hausdorff dimension. For $d=1$ it is linearly stable. But for $d\geq 2$ some solutions, forming the set (e.g. those, corresponding to $|\mathcal{A}|=1$) are linearly stable, while in view of the examples in Appendix B some others with $|\mathcal{A}|\geq 2$ are linearly unstable.

For $d\geq 2$ the unstable parts of the sets $\mathfrak{A}$ creates around them some local instabilities. It is unclear for us wether these instabilities have anything to do with the phenomenon of the energy cascade to high frequencies, predicted by the theory of wave turbulence for small-amplitude solutions of space-multidimensional hamiltonian PDEs. The linear instability of solutions and the energy cascade to high frequencies on various time-scales are now topics of major interest for the nonlinear PDE community, e.g. see in [11].

We note that the fact that KAM-solutions of high dimensional PDEs may be linearly unstable is not new: in [20] the instability of some KAM-solutions for the 2d cubic NLS equation was observed (see there Remark 1.1), while in [30, 31] algebraic reasons for the instability of KAM-solutions for multidimensional NLS equations were discussed.

Our study of the beam equation (1.1) leads to several natural questions. One is to find a sufficient condition for an admissible set $\mathcal{A}\subset\mathbb{Z}^d$, such that $d,|\mathcal{A}|\geq 2$, to guarantee that the hamiltonian operator $J\widehat{K}(\rho)$ in Theorem 1.3 is non-trivial for ρ in some component $\tilde{Q}_l$ of the set $\tilde{Q}$ (we recall that for some components of $\tilde{Q}$ it always is trivial). Cf. Remark 1.6.4).

If for some $\mathcal{A}$ this property is fulfilled and the assertion of Theorem 1.4 holds (e.g. the set $\mathcal{A}$ is strongly admissible), then by Remark 6.4 the set $\mathfrak{J}$ has a subset $\mathfrak{J}_h$, having positive density at the origin, such that for $\rho\in\mathfrak{J}_h$ the corresponding KAM-solutions of eq. (1.1) are linearly unstable.

Another question is to study the persistence of small-amplitude linear solutions (1.5) in the beam equation (1.1) for the case when the set $\mathcal{A}$ is not admissible.

A third question concerns the role of the Hypothesis A1(d) in Section 6.1. In the notation of that section, do the majority of the invariant tori $\mathbb{T}^n\times\{0\}\times\{0\}$ of the Hamiltonian h persist as invariant tori for the Hamiltonian H, if the condition A1(d) is violated and $\Lambda_a+\Lambda_b\equiv 0$ for some $a,b\in\tilde{\mathcal{L}}_\infty$?

We recall that the condition A1(d) is the only one which we can check for strongly admissible sets $\mathcal{A}$, but not for admissible.

Appendix A. Proof of Lemma 3.2

For any $\gamma\geq 0$ let us denote by Z_γ the space of complex sequences $v=(v_s,s\in\mathbb{Z}^d)$ with finite norm $\|v\|_\gamma$, defined by the same relation as the norm in the space Y_γ. For $v\in Z_\gamma$ we will denote by $\mathcal{F}(v)=u(x)$ the Fourier-transform of v, $u(x)=\sum v_s e^{is\cdot x}$. By Example 3.1 if $u(x)$ is a bounded real holomorphic function in $\mathbb{T}^n_{\sigma'}$, then $\mathcal{F}^{-1}u\in Z_\sigma$ for $\sigma<\sigma'$.

Let F be the Fourier-image of the nonlinearity g, i.e. $F(v)=\mathcal{F}^{-1}g(x,\mathcal{F}(v)(x))$.

Lemma A.1. *For sufficiently small $\mu_*>0,\gamma_*>0$ and for all $0\le\gamma\le\gamma_*$,*
i) F defines a real holomorphic mapping $\mathcal{O}_{\mu_*}(Z_\gamma)\to Z_\gamma$,
ii) ∇F defines a real holomorphic mapping $\mathcal{O}_{\mu_*}(Z_\gamma)\to M_\gamma$, where M_γ is the space of matrices $A:\mathbb{Z}^d\times\mathbb{Z}^d\to\mathbb{C}$, satisfying $|A|_\gamma:=\sup|A_a^b|\,e^{\gamma|a-b|}<\infty$.

Proof. i) For sufficiently small $\sigma',\mu>0$ the nonlinearity g defines a real holomorphic function $g:\mathbb{T}^d_{\sigma'}\times\mathcal{O}_\mu(\mathbb{C})\to\mathbb{C}$ and the norm of this function is bounded by some constant M. We may write it as $g(x,u)=\sum_{r=3}^\infty g_r(x)u^r$, where $g_r(x)=\frac{1}{r!}\frac{\partial^r}{\partial u^r}g(x,u)\,|_{u=0}$. So $g_r(x)$ is holomorphic in $x\in\mathbb{T}^d_{\sigma'}$ and by the Cauchy estimate $|g_r|\le M\mu^{-r}$. So

$$\|\mathcal{F}^{-1}g_r\|_\gamma\le C_\sigma M\mu^{-r}\quad\forall\,0\le\gamma\le\sigma\,,$$

for any $\sigma<\sigma'$. Cf. Example 3.1. We may write $F(v)$ as

$$F(v)=\sum_{r=3}^{\infty}(\mathcal{F}^{-1}g_r)\star\underbrace{v\star\cdots\star v}_{r}\,. \tag{A.1}$$

Since the space Z_γ is an algebra with respect to the convolution (see Lemma 1.1 in [16]), the r-th term of the sum is bounded as follows:

$$\|(\mathcal{F}^{-1}g_r)\star\underbrace{v\star\cdots\star v}_{r}\|_\gamma\le C_1C^{r+1}\mu^{-r}\|v\|_\gamma^r\,. \tag{A.2}$$

This implies the assertion with $\gamma_*=\sigma$ and a suitable $\mu_*>0$.

ii) For $r\ge3$ consider the r-th term in the sum for $g(x,u(x))$ and denote by G_r its Fourier-image, $G_r(v)=\mathcal{F}^{-1}(g_ru^r)$, $u=\mathcal{F}(v)$. Then

$$(\nabla G_r(v))_a^b=r(2\pi)^{-d}\int e^{-ia\cdot x}g_r(x)u^{r-1}e^{ib\cdot x}\,dx\,.$$

Applying (A.2) (with r convolutions instead of $r+1$) we see that

$$|(\nabla G_r(v))_a^b|\le C_2C^r\mu^{-r}\|v\|_\gamma^{r-1}\langle b-a\rangle^{-d^*}e^{-\gamma|b-a|}\,. \tag{A.3}$$

So $|\nabla G_r(v)|_\gamma\le C^r\mu^{-r}\|v\|_\gamma^{r-1}$, which implies the second assertion of the lemma. $\square$

Proof of Lemma 3.2. Let us consider the functional $P(\zeta)$ as in (1.9), and write it as $P(\zeta)=p\circ\Upsilon\circ D^{-1}\zeta$. Here D is the operator, defined in Section 3.1, Υ is the bounded operator

$$\Upsilon:Y_\gamma\to Z_\gamma,\qquad\zeta\to v,\;\; v_s=\frac{\xi_s+\eta_{-s}}{\sqrt{2}}\;\;\forall\,s,$$

and $p(v)=\int G(x,(\mathcal{F}^{-1}v)(x))\,dx$. Lemma A.1 with g replaced by G immediately implies that P is a real holomorphic function on $\mathcal{O}_{\mu_*}(Y_{\gamma_*})$ with suitable $\mu_*,\gamma_*>0$.

Next, since

$$\nabla P(\zeta)=D^{-1}\circ{}^t\Upsilon\circ\nabla p(\Upsilon\circ D^{-1}\zeta)\,,$$

where $\nabla P=F$ is the map in Lemma A.1, then ∇P defines a real holomorphic mapping $\mathcal{O}_{\mu_*}(Y_{\gamma_*})\to Y_{\gamma_*}$.
Further,

$$\nabla^2P(\zeta)=D^{-1}({}^t\Upsilon\;\nabla^2p(\Upsilon\circ D^{-1}\zeta)\;\Upsilon)D^{-1}\,.$$

Since for any $A \in M_\gamma$ the matrix ${}^t\Upsilon A \Upsilon$ is given by the relation

$$({}^t\Upsilon A\Upsilon)_a^b = \frac{1}{2}\sum_{a'=\pm a,\, b'=\pm b} A_{a'}^{b'}\,,$$

then $|D^{-1}({}^t\Upsilon A\Upsilon)D^{-1}|_\gamma^D \le 2|A|_\gamma$. So

$$|\nabla^2 P(\zeta)|_\gamma^D \le 2|\nabla^2 p(\zeta)|_\gamma = 2|\nabla F(\zeta)|_\gamma\,,$$

and in view of item ii) of Lemma A.1, the mapping

$$\nabla_\gamma^2 P : \mathcal{O}_{\mu_*}(Y_\gamma) \to \mathcal{M}_\gamma^D, \qquad 0 \le \gamma \le \gamma_*\,,$$

is real holomorphic and bounded in norm by a γ-independent constant. □

Appendix B. Examples

In this appendix we discuss some examples of hamiltonian operators $\mathcal{H}(\rho) = iJK(\rho)$ defined in (4.3), corresponding to various dimensions d and sets $\mathcal{A}$. In particular we are interested in examples which give rise to partially hyperbolic KAM solutions.

Examples with $(\mathcal{L}_f \times \mathcal{L}_f)_+ = \emptyset$.

As we noticed in (4.4), if $(\mathcal{L}_f \times \mathcal{L}_f)_+ = \emptyset$ then $\mathcal{H}$ is Hermitian, so the constructed KAM-solutions are linearly stable. This is always the case when $d = 1$.

When $d = 2$ and $\mathcal{A} = \{(k,0),(0,\ell)\}$ with the additional assumption that neither k^2 nor ℓ^2 can be written as the sum of squares of two natural numbers, we also have $(\mathcal{L}_f \times \mathcal{L}_f)_+ = \emptyset$.

Similar examples can be constructed in higher dimension, for instance for $d = 3$ we can take $\mathcal{A} = \{(1,0,0),(0,2,0)\}$ or $\mathcal{A} = \{(1,0,0),(0,2,0),(0,0,3)\}$.

We note that in [19] the authors perturb solutions (1.5), corresponding to set $\mathcal{A}$ for which $(\mathcal{L}_f \times \mathcal{L}_f)_+ = \emptyset$ and $(\mathcal{L}_f \times \mathcal{L}_f)_- = \emptyset$. This significantly simplifies the analysis since in that case there is no matrix K in the mormal form (3.14) and the unperturbed quadratic Hamiltonian is diagonal.

Examples with $(\mathcal{L}_f \times \mathcal{L}_f)_+ \neq \emptyset$. In this case hyperbolic directions may appear as we show below.

The choice $\mathcal{A} = \{(j,k),(0,-k)\}$ leads to $((j,-k),(0,k)) \in (\mathcal{L}_f \times \mathcal{L}_f)_+$.

Note that this example can be plunged in higher dimensions, e.g. the 3d-set $\mathcal{A} = \{(j,k,0),(0,-k,0)\}$ leads to a non trivial $(\mathcal{L}_f \times \mathcal{L}_f)_+$.

Examples with hyperbolic directions

Here we give examples of normal forms with hyperbolic eigenvalues, first in dimension two, then – in higher dimensions. That is, for the beam equation (1.1) we will find admissible sets $\mathcal{A}$ such that the corresponding matrices $iJK(\rho)$ in the normal form (3.14) have unstable directions. Then by Theorem 1.5 the time-quasiperiodic solutions of (1.1), constructed in the theorem, are linearly unstable.

We begin with dimension $d = 2$. Let

$$\mathcal{A} = \{(0,1),(1,-1)\}\,.$$

We easily compute using (3.29), (3.30) that

$$\mathcal{L}_f = \big\{(0,-1),(1,0),(-1,0),(1,1),(-1,1),(-1,-1)\big\}\,,$$

and

$$(\mathcal{L}_f \times \mathcal{L}_f)_+ = \{\big((0,-1),(1,1)\big);\big((1,1),(0,-1)\big)\}, \qquad (\mathcal{L}_f \times \mathcal{L}_f)_- = \emptyset.$$

So in this case the decomposition (4.9) of the hamiltonian operator $\mathcal{H}(\rho) = iJK(\rho)$ reads

$$\mathcal{H}(\rho) = \mathcal{H}_1(\rho) \oplus \mathcal{H}_1(\rho) \oplus \mathcal{H}_3(\rho) \oplus \mathcal{H}_4(\rho) \oplus \mathcal{H}_5(\rho)\,,$$

where $\mathcal{H}_1(\rho) \oplus \mathcal{H}_1(\rho) \oplus \mathcal{H}_3(\rho) \oplus \mathcal{H}_4(\rho)$ is a diagonal operator with purely imaginary eigenvalues and $\mathcal{H}_5(\rho)$ is an operator in $\mathbb{C}^4$ which may have hyperbolic eigenvalues. That is, now $M = 5$ and $M_* = 4$.

Let us denote $\zeta_1 = (\xi_1, \eta_1)$ (reps. $\zeta_2 = (\xi_2, \eta_2)$) the (ξ, η)-variables corresponding to the mode $(0, -1)$ (reps. $(1, 1)$). We also denote $\rho_1 = \rho_{(1,0)}$, $\rho_2 = \rho_{(1,-1)}$, $\lambda_1 = \sqrt{1+m}$ and $\lambda_2 = \sqrt{4+m}$. By construction $\mathcal{H}_5(\rho)$ is the restriction of the Hamiltonian $\langle K(m,\rho)\zeta_f, \zeta_f\rangle$ to the modes (ξ_1, η_1) and (ξ_2, η_2). We calculate using (3.47) that

$$\langle \mathcal{H}_5(\rho)(\zeta_1, \zeta_2), (\zeta_1, \zeta_2)\rangle = \beta(\rho)\xi_1\eta_1 + \gamma(\rho)\xi_2\eta_2 + \alpha(\rho)(\eta_1\eta_2 + \xi_1\xi_2)\,, \tag{B.1}$$

where

$$\alpha(\rho) = \frac{6}{4\pi^2}\frac{\sqrt{\rho_1\rho_2}}{\lambda_1\lambda_2}\,, \quad \beta(\rho) = \frac{3}{4\pi^2}\frac{1}{\lambda_1}\Big(\frac{\rho_1}{\lambda_1} - \frac{2\rho_2}{\lambda_2}\Big), \quad \gamma(\rho) = \frac{3}{4\pi^2}\frac{1}{\lambda_2}\Big(\frac{\rho_2}{\lambda_2} - \frac{2\rho_1}{\lambda_1}\Big)\,.$$

Thus the linear hamiltonian system, governing the two modes, reads[16]

$$\begin{cases} \dot\xi_1 & = -i(\beta\xi_1 + \alpha\eta_2) \\ \dot\eta_1 & = i(\beta\eta_1 + \alpha\xi_2) \\ \dot\xi_2 & = -i(\gamma\xi_2 + \alpha\eta_1) \\ \dot\eta_2 & = i(\gamma\eta_2 + \alpha\xi_1). \end{cases}$$

So the hamiltonian operator $\mathcal{H}_5$ has the matrix iM, where

$$M = \begin{pmatrix} -\beta & 0 & 0 & -\alpha \\ 0 & \beta & \alpha & 0 \\ 0 & -\alpha & -\gamma & 0 \\ \alpha & 0 & 0 & \gamma \end{pmatrix}.$$

We can calculate its characteristic polynomial of M explicitly to obtain after factorisation

$$\det(M - \lambda I) = \big(\lambda^2 + (\gamma - \beta)\lambda - \beta\gamma + \alpha^2\big)\big(\lambda^2 - (\gamma - \beta)\lambda - \beta\gamma + \alpha^2\big)\,.$$

Then we compute discriminant of the quadratic polynomial $\lambda^2 + (\gamma-\beta)\lambda - \beta\gamma + \alpha^2$,

$$\Delta = (\beta + \gamma)^2 - 4\alpha^2.$$

Choosing $\rho_1 = \rho_2 = \rho$ we get

$$\beta + \gamma = 3(2\pi)^{-2}\rho\Big(\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} - \frac{4}{\lambda_1\lambda_2}\Big), \quad \alpha = 6(2\pi)^{-2}\rho\frac{1}{\lambda_1\lambda_2}\,,$$

and

$$\Delta = \frac{9\rho}{(2\pi)^4}\Big(\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2}\Big)\Big(\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} - \frac{8}{\lambda_1\lambda_2}\Big) \le \frac{9\rho}{(2\pi)^4}\Big(\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2}\Big)\Big(\frac{1}{\lambda_1^2} - \frac{7}{\lambda_2^2}\Big)\,.$$

Thus, $\Delta < 0$ for all $m \in [1, 2]$, and M has eigenvalues with non vanishing imaginary parts. Accordingly, the hamiltonian operators $\mathcal{H}_5$ and $\mathcal{H}$ have hyperbolic directions. Actually, since the discriminant of the polynomial $\lambda^2 - (\gamma - \beta)\lambda - \beta\gamma + \alpha^2$ also equals Δ, the hamiltonian operator $\mathcal{H}_5$ has only hyperbolic directions.

[16]Recall that the symplectic two-form is: $-i\sum d\xi \wedge d\eta$.

This example can be generalised to any dimension $d \geq 3$. Let us do it for $d = 3$. Let

$$\mathcal{A} = \{(0,1,0),(1,-1,0)\}. \tag{B.2}$$

We verify that $\mathcal{L}_f$ contains 16 points, that $(\mathcal{L}_f \times \mathcal{L}_f)_- = \emptyset$ and

$$\begin{aligned}(\mathcal{L}_f \times \mathcal{L}_f)_+ = \{&((0,-1,0),(1,1,0)); ((1,1,0),(0,-1,0));\\ &((1,0,-1),(0,0,1)); ((0,0,1),(1,0,-1));\\ &((1,0,1),(0,0,-1)); ((0,0,-1),(1,0,1))\}\,.\end{aligned}$$

I.e. $(\mathcal{L}_f \times \mathcal{L}_f)_+$ contains three pairs of symmetric couples $(a,b),(b,a)$ which give rise to three non trivial 2×2-blocks in the matrix $\mathcal{H}$. Now $M = 13$, $M_* = 10$ and the decomposition (4.9) reads

$$\mathcal{H}(\rho) = \mathcal{H}_1(\rho) \oplus \cdots \oplus \mathcal{H}_{13}(\rho)\,.$$

Here $\mathcal{H}_1(\rho) \oplus \cdots \oplus \mathcal{H}_{10}(\rho)$ is the diagonal part of $\mathcal{H}$ with purely imaginary eigenvalues, while the operators $\mathcal{H}_{11}(\rho)$, $\mathcal{H}_{12}(\rho)$, $\mathcal{H}_{13}(\rho)$ correspond to non-diagonal 4×4–matrices.

Denoting $\rho_1 = \rho_{(0,1,0)}$ and $\rho_2 = \rho_{(1,-1,0)}$ we find that the restriction of the Hamiltonian $\langle K(m,\rho)\zeta_f,\zeta_f\rangle$ to the modes $(\xi_1,\eta_1) := (\xi_{(0,-1,0)},\eta_{(0,-1,0)})$ and $(\xi_2,\eta_2) := (\xi_{(1,1,0)},\eta_{(1,1,0)})$ is governed by the Hamiltonian $h_r(\rho_1,\rho_2)$ given in (B.1), as in the 2d case. Similarly the restrictions of the Hamiltonian $\langle K(m,\rho)\zeta_f,\zeta_f\rangle$ to the pair of modes $(\xi_{(1,0,-1)},\eta_{(1,0,-1)})$ and $(\xi_{(0,0,1)},\eta_{(0,0,1)})$ and to the pair of modes $(\xi_{(1,0,1)},\eta_{(1,0,1)})$ and $(\xi_{(0,0,-1)},\eta_{(0,0,-1)})$ are given by the same Hamiltonian (B.1). So $\mathcal{H}_{11}(\rho) \equiv \mathcal{H}_{12}(\rho) \equiv \mathcal{H}_{13}(\rho)$ and for $\rho_1 = \rho_2$ we have 3 hyperbolic directions, one in each block Y^{f11}, Y^{f12} and Y^{f13} (see (4.7)) with the same eigenvalues.

We notice that the eigenvalues are identically the same for all three blocks, thus the relation (4.22) is violated. This does not contradict Lemma 4.5 since the set (B.2) is not strongly admissible. Indeed, denoting $a = (0,1,0)$, $b = (1,-1,0)$ we see that $c := a + b = (1,0,0)$. So three points $(0,-1,0),(0,0,\pm 1) \in \mathbf{S}_{|a|}$ all lie at the distance $\sqrt{2}$ from c. Hence, it is not true that $a \angle\!\!\angle\, b$.

Appendix C. Some linear algebra

Lemma C.1. *Let L be an $N\times N$-complex matrix with eigenvalues $\lambda_1,\dots,\lambda_N$ such that $|\lambda_j - \lambda_k| \geq \delta > 0$ for all $j \neq k$, and the normalised eigenvectors $\xi_1,\dots,\xi_N$.*[17] *Consider the $N\times N$-matrix $U = (\xi_1\xi_2\dots\xi_N)$, so that*

$$U^{-1}LU = \mathit{diag}\{\lambda_1,\dots,\lambda_N\} =: \Lambda\,. \tag{C.1}$$

Then

$$\|U^{-1}\| \leq \sqrt{N}\,(2\delta^{-1}\|L\|)^{N-1}\,. \tag{C.2}$$

Proof. [18] Let $Ux = y$, where $\|y\| = 1$. We have to estimate the norm of x. To do this we will estimate the components x_j of that vector.

From (C.1) we have that

$$P(L)U = UP(\Lambda)\,, \tag{C.3}$$

[17] we recall that they should be regarded as column-vectors.

[18] We learned this short proof from V. Šverák.

for any polynomial P. Now for $j = 1, \dots, N$ consider the Lagrangian polynomials P_j,

$$P_j(z) = \frac{(z-\lambda_1)\dots\widehat{(z-\lambda_j)}\dots(z-\lambda_N)}{(\lambda_j-\lambda_1)\dots\widehat{(\lambda_j-\lambda_j)}\dots(\lambda_j-\lambda_N)}\,,$$

where the over-hat means that the corresponding factor is omitted. Then $P_j(\lambda_l) = \delta_{j,l}$. Therefore $P_j(\Lambda) = \operatorname{diag}(0,\dots,\underset{j}{1},\dots,0)$. Applying (C.3) with $P = P_j$ to the vector x we get:

$$P_j(L)y = UP_j(\Lambda)x = U\big({}^t(0,\dots,x_j,\dots,0)\big) = x_j\xi_j\,.$$

Therefore

$$|x_j| = \|P_j(L)y\| \le \|P_j(L)\| \le \frac{(2\|L\|)^{N-1}}{\delta^{N-1}}\,,$$

since $\|L-\lambda_j E\| \le 2\|L\|$. From this we find that $\|x\| \le \sqrt{N}(2\delta^{-1}\|L\|)^{N-1}\|y\|$, and the required estimate is established. □

As an example of applying estimate (C.2), consider in the symplectic space $\big(\mathbb{R}^4 = \{(p_1,p_2,q_1,q_2)\}, dp\wedge dq =: \omega_2\big)$ the symmetric matrix $A = A^{a,b}$, corresponding to the quadratic form

$$a(p_1q_1+p_2q_2)+b(p_1q_2-p_2q_1), \qquad a,b\neq 0\,, \tag{C.4}$$

and the hamiltonian operator JA. It has the eigenvalues $(\pm a\pm ib)$, see [1], Appendix 6. So the spectrum of JA is simple, and we can diagonalise it as in the lemma above: $U^{-1}JAU = \operatorname{diag}\{\pm a\pm ib\}$. Clearly $\|U\|\le 2$, and by (C.2)

$$\|U^{-1}\| \le 2(\|A\|/\min(|a|,|b|))^3 =: T\,.$$

Let us enumerate the eigenvalues $(\pm a\pm ib)$ as follows: $\lambda_1 = a+ib$, $\lambda_2 = -a+ib$, $\lambda_3 = -a-ib$, $\lambda_4 = a-ib$, and let $\xi_1,\dots,\xi_4$ be the corresponding eigenvectors. Then $\omega_2(\xi_a,\xi_b) = 0$, unless $\{a,b\} = \{1,3\}$ or $\{a,b\} = \{2,4\}$. Consider

$$\omega_2(\xi_1,\xi_3) =: t_{1,3}\,,\quad \omega_2(\xi_2,\xi_4) =: t_{2,4}\,.$$

Find a unit vector $\xi_1^d \in \mathbb{C}^4$ such that $\omega_2(\xi_1,\xi_1^d) = 1$, and decompose it as

$$\xi_1^d = x_1\xi_1+\dots+x_4\xi_4\,, \qquad x_j\in\mathbb{C}\,. \tag{C.5}$$

Then $\|x\| \le \|U^{-1}\| \le T$. We have

$$1 = \omega_2(\xi_1,\xi_1^d) = \sum x_j\omega_2(\xi_1,\xi_j) = x_3t_{1,3} \le \|x\|t_{1,3}\,. \tag{C.6}$$

So $1 \ge |t_{1,3}| \ge T^{-1}$. Similar $1\ge |t_{2,4}| \ge T^{-1}$.

Now let us modify the eigenvectors as follows:

$$\tilde\xi_1 = (t_{1,2})^{-1}\xi_1,\quad \tilde\xi_2 = (t_{2,4})^{-1}\xi_2,\quad \tilde\xi_3 = \xi_3,\quad \tilde\xi_4 = \xi_4\,.$$

Let $\{e_1 = e_{p_1}, e_2 = e_{p_2}, e_3 = e_{q_1}, e_4 = e_{q_2}\}$ be the standard euclidean base of $\mathbb{R}^4$. Then $\omega_2(e_a,e_b) = \omega_2(\tilde\xi_a,\tilde\xi_b)$ for all a,b, so the modified transformation $\widetilde U = U\operatorname{diag}\{(t_{1,2})^{-1},(t_{2,4})^{-1},1,1\}$ is symplectic. It still diagonalises JA, $\widetilde U^{-1}JA\widetilde U = \operatorname{diag}\{\pm a\pm ib\}$, and satisfies

$$\|\widetilde U^{-1}\| \le T, \qquad \|\widetilde U\| \le 2T\,. \tag{C.7}$$

Lemma C.2. *Consider two $N\times N$-matrices A_1, A_2, real or complex. Then the distance between their spectra is bounded by $C\|A_1-A_2\|^{1/N}$, where C depends only on N and the norms of the two matrices.*

Proof. Consider the characteristic polynomial of A_1. The classical Cartan theorem (see [27], Section 1.7) tells that the subset $S_\varepsilon(A_1)$ of the complex plain, where this polynomial is smaller than ε, may be covered by a finite collection of complex discs such that the sum of their radii equals $2e\,(\varepsilon)^{1/N}$. The set $S_\varepsilon(A_1)$ contains the eigenvalues of the matrix A_2 (i.e., the zeroes of its characteristic polynomial) if we chose $\varepsilon =$Const $\|A_1 - A_2\|$. This implies the assertion. ☐

Appendix D. An estimate for polynomial functions

Lemma D.1. *Let $F(x)$ be a non-trivial real polynomial of degree $\bar d$, restricted to a bounded domain $\mathcal{K} \subset \mathbb{R}^n$ with a piece-wise smooth boundary. Then there exists a positive constant C_F such that*

$$\text{meas}\{x \in K^n \mid |F(x)| < \varepsilon\} \le C_F \varepsilon^{1/\bar d}, \qquad \forall\, \varepsilon \in (0,1]\,. \tag{D.1}$$

Proof. By the compactness argument it suffices to prove this in the vicinity of any point $x^0 \in \mathcal{K} \subset \mathbb{R}^n$, where $F(x^0) = 0$. So we have reduced the problem to the case when

$$F : B^n_\varkappa := \{|x| < \varkappa\} \to \mathbb{R}, \qquad \varkappa > 0\,, \tag{D.2}$$

and F is a non-trivial polynomial of degree $\bar d$, $F(0) = 0$. For a unit vector $\xi \in R^n$ consider the polynomial of one variable $z \mapsto F(z\xi)$. For a generic ξ it has the form $C^F z^{\bar d} + \dots$, $C^F \ne 0$. Rotating the coordinate system we achieve that $\xi = (1, 0, \dots, 0)$. Denote

$$x = (x_1, \dots, x_n) = (x_1, \bar x), \qquad \bar x = (x_2, \dots, x_n)\,.$$

Then

$$F(x) = F(x_1, \bar x) = C_d(\bar x) x_1^d + \dots + C_0(\bar x)\,, \qquad C_d(0) = C^F\,,$$

where each C_j is a polynomial of $\bar x$ whose coefficients are bounded in terms of F. Decreasing $\varkappa$ if needed we achieve that

$$|C_d(\bar x)| \ge \tfrac{1}{2} C^F \qquad \forall\, \bar x \in B^{n-1}_\varkappa\,.$$

Lemma 2.4 with $n = \bar d$ applies to the function $x_1 \mapsto F(x_1, \bar x)$, $\bar x \in B^{n-1}_\varkappa$, and implies that

$$\text{meas}\{x_1 \in [-\varkappa, \varkappa] : |F(x_1; \bar x)| \le \varepsilon\} \le C'^F \varepsilon^{1/\bar d}\,.$$

Jointly with the Fubini theorem this inequality establishes for the function (D.2) estimate (D.1) with K^n replaced by $B^n_\varkappa$ and implies the assertion of the lemma. ☐

Appendix E. Admissible and strongly admissible random R-sets are typical

In this appendix we prove (1.12) and (1.16).

Proof of (1.12). Clearly

$$\mathbb{P}\,(\Omega \setminus \Omega_1) \le \binom{n}{2} \mathbb{P}\,\{|\xi^1| = |\xi^2|\}\,, \tag{E.1}$$

and

$$\mathbb{P}\,\{|\xi^1| = |\xi^2|\} = |\mathbf{B}(R)|^{-2} C^*\,, \qquad C^* = \sum_{\substack{(a,b) \in \mathbf{B}(R) \times \mathbf{B}(R) \\ |a| = |b|}} 1\,.$$

Denote by $B^+(R)$ the subset of $\mathbb{R}^d$ which is the union of standard 1-cubes with centres in points of $\mathbf{B}(R)$ and denote by $K(R)$ the cube $\{x \in \mathbb{R}^d \mid |x_j| \le R \ \forall j\}$. Then

$$C^* \le \int_{B^+(R)} \int_{B^+(R)} \chi_{|\,|x|-|y|\,|\le\sqrt{d}}\, dx\, dy \le \int_{K(R+1/2)} \int_{K(R+1/2)} \chi_{|\,|x|-|y|\,|\le\sqrt{d}}\, dx\, dy\,.$$

A straightforward (but a bit cumbersome) calculation shows that the r.h.s. is $\le C(d)R^{-1}$. Therefore $\mathbb{P}\,(\Omega\setminus\Omega_1) \le C(n,d)R^{-1}$. This and (E.1) implies (1.12).

Proof of (1.16). Let us denote $A_d = \pi^{d/2}/\Gamma(\frac{d+2}{2})$, where Γ is the gamma-function. Then, by the celebrated result of Vinogradov and Chen, for $d \ge 2$ we have

$$\big||\mathbf{B}(R)| - A_d R^d\big| \le C_{\theta_d} R^{\theta_d} \qquad \forall\, R > 0\,,$$

for any $\theta_d > d-2$ for $d \ge 4$ and $\theta_3 = 4/3$; e.g. see [32]. Since $|\mathbf{S}(R)| \le |\mathbf{B}(R+\varepsilon)| - |\mathbf{B}(R-\varepsilon)|$ for every $\varepsilon > 0$, then

$$\Gamma_{R,d} := |\mathbf{S}(R)| \le 2C_\theta R^\theta \qquad \forall\, R > 0\,, \tag{E.2}$$

with $\theta = \theta_d$ as above.[19]

Below we restrict ourselves to the case $d = 3$ since for higher dimension the argument is similar, but more cumbersome. We have that

$$1 - \mathbb{P}\,\{\xi^1 \angle\!\!\angle\, \xi^2\} = |\mathbf{B}(R)|^{-2} C^{**}\,, \quad C^{**} = \#\{(a,b) \in \mathbf{B}(R)\times\mathbf{B}(R) \mid \text{ not } a \angle\!\!\angle\, b\}\,, \tag{E.3}$$

and, denoting $a + b = c$, that

$$C^{**} \le \#\{(a,c) \in \mathbf{B}(2R)\times\mathbf{B}(2R) \mid \text{ not } a \angle c\}\,. \tag{E.4}$$

Now we will estimate the r.h.s. of (E.4), redenoting $2R$ back to R. That is, will estimate the cardinality of the set

$$X = \{(a,b) \in \mathbf{B}(R)\times\mathbf{B}(R) \mid \text{ not } a \angle b\}\,.$$

It is clear that $(a,b) \in X$, $a \ne 0$, iff there exist points $a', a'' \in \mathbf{S}(|a|)$ such that b lies in the line $\Pi_{a,a',a''}$, which is perpendicular to the triangle (a,a',a'') and passes through its centre, so it also passes through the origin. Let $v = v_{a,a',a''}$ be a primitive integer vector in the direction of $\Pi_{a,a',a''}$. For any $a \in \mathbb{Z}^d$, $a \ne 0$, denote

$$\Delta(a) = \big\{\, \{a', a''\} \subset \mathbf{S}(|a|)\setminus\{a\} \mid a' \ne a'' \big\}\,.$$

Then

$$|\Delta(a)| < \Gamma^2_{|a|,3} \le C_\theta^2 R^{2\theta}, \qquad \theta = \theta_3\,,$$

see (E.2). For a fixed $a \in \mathbf{B}(R)\setminus\{0\}$ consider the mapping

$$\Delta(a) \ni \{a', a''\} \mapsto v = v_{a,a',a''}\,.$$

It is clear that each direction $v = v_{a,a',a''}$ gives rise to at most $2R|v|^{-1}$ points b such that $(a,b) \in X$. So, denoting

$$X_a = \{b \in \mathbf{B}(R) \mid (a,b) \in X\}\,,$$

[19] It is known (see [21], Theorem 338) that $\Gamma_{R,2} \le C_\delta R^\delta$ for every $\delta > 0$. Writing $\Gamma_{R,3}$ as an integral in the counting measure $\sum_{s\in\mathbb{Z}^3}\delta(\cdot - s)$, $\Gamma_{R,3} = \int_{\mathbf{S}(R)} 1$, applying to this integral the Fubini theorem and the estimate for $\Gamma_{R',2}$, $0 \le R' \le R$, we find that $\Gamma_{R,3} \le C'_\delta R^{1+\delta}$ for each $\delta > 0$, which is better than the estimate, obtained from the Vinogradov–Chen result. But the latter is sufficient for us.

we have

$$|X_a| \le 2R \sum |v_{a,a',a''}|^{-1}\,, \quad \text{if } a \neq 0\,,$$

where the summation goes through all different vectors v, corresponding to various $\{a', a''\} \in \Delta(a)$. As $|v|^{-1}$ is the bigger the smaller $|v|$ is, we see that the r.h.s. is $\le 2R \sum_{v \in \mathbf{B}(R')\setminus\{0\}} |v|^{-1}$, where R' is any number such that $|\mathbf{B}(R')| \ge |\Delta(a)|$. Since $|\Delta(a)| \le \Gamma^2_{|a|,3}$, then choosing $R' = R'_a = C\Gamma^{2/3}_{|a|,3}$ we get for any $a \in \mathbf{B}(R)\setminus\{0\}$ that

$$|X_a| \le 2CR \sum_{\mathbf{B}(R'_a)\setminus\{0\}} |v|^{-1} \le C_1 R \int_{B(R'_a)} |x|^{-1}\,dx \le C_2 R (R'_a)^2 = C_3 R\, \Gamma^{4/3}_{|a|,3}\,.$$

By (1.14), $X_0 = \{0\}$. So

$$|X| = \sum_{a \in \mathbf{B}(R)} |X_a| \le 1 + CR \sum_{a \in \mathbf{B}(R)\setminus\{0\}} \Gamma^{4/3}_{|a|,3}\,.$$

Evoking the estimate (E.2) we finally get that

$$|X| \le C_1 R \sum_{a \in \mathbf{B}(R)\setminus\{0\}} |a|^{\frac{4}{3}\theta_3} \le C_2 R \int_{B(R)} |x|^{\frac{4}{3}\theta_3}\,dx \le C_3 R^{1+3+\frac{4}{3}\theta_3} = C_3 R^{5+7/9}\,.$$

Jointly with (E.3), (E.4) and the definition of the set X this implies the required relation (1.16) with $\kappa = 2/9$.

References

[1] V. I. Arnold. Mathematical methods in classical mechanics, 3rd edition. *Springer-Verlag*, Berlin, 1989.

[2] D. Bambusi, *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Physics **234** (2003), 253–283.

[3] D. Bambusi and B. Grébert, *Birkhoff normal form for PDE's with tame modulus.* Duke Math. J. 135 no. 3 (2006), 507-567.

[4] M. Berti, P. Bolle, *Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential*, Nonlinearity **25** (2012), 2579-2613.

[5] M. Berti, P. Bolle, *Quasi-periodic solutions with Sobolev regularity of NLS on T^d with a multiplicative potential*, J. Eur. Math. Soc. **15** (2013), 229-286.

[6] A. I. Bobenko and S. B. Kuksin, The nonlinear Klein-Gordon equation on an interval as a perturbed Sine-Gordon equation. *Comment. Math. Helv.*, **7**0, 1995, 63-112.

[7] J. Bourgain *Construction of approximative and almost-periodic solutions of perturbed linear Schrödinger and wave equations*, *GAFA* **6**, (1995), 201-235.

[8] J. Bourgain *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Shödinger equation*, *Ann. Math.* **1**48 (1998), 363-439.

[9] J. Bourgain *Green's function estimates for lattice Schrödinger operators and applications*, *Annals of Mathematical Studies*, Princeton, 2004.

[10] R. Benedetti and J-J Risler, Real algebraic and semi-algebraic sets. *Hermann*, Paris, 1990.

[11] J. Colliander, M. Keel, G. Staffilani, H. Takaoka and T. Tao, *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation*, *Inv. Math.* **1**81 (2010), 31-113.

[12] W. Craig *Problèmes de Petits Diviseurs dans les Équations aux Dérivées Partielles*, *Panoramas et Synthèses*, Société Mathématique de France, 2000.

[13] L.H Eliasson, *Almost reducibility of linear quasi-periodic systems*, in *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, 679-705, Proc. Sympos. Pure Math., 69, Amer. Math. Soc., Providence, RI, 2001.

[14] L.H Eliasson, *Perturbations of linear quasi-periodic systems*, in *Dynamical Systems and Small Divisors (Cetraro, Italy, 1998)*, 1-60, Lect. Notes Math. 1784, Springer, 2002.

[15] L.H. Eliasson, B. Grébert and S.B. Kuksin, KAM for the nonlinear beam equation 2: a normal form theorem. *Preprint* (2013).

[16] L.H. Eliasson and S.B. Kuksin. Infinite Töplitz-Lipschitz matrices and operators. *Z. Angew. Math. Phys.* **59** (2008), 24-50.
[17] L.H. Eliasson and S.B. Kuksin. KAM for the nonlinear Schrödinger equation. *Ann. Math* **172** (2010), 371-435.
[18] J. Geng and J. You. A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces. *Comm. Math. Phys.*, **262** (2006), 343–372.
[19] J. Geng and J. You. KAM tori for higher dimensional beam equations with constant potentials, *Nonlinearity*, **19** (2006), 2405–2423.
[20] J. Geng, X. Hu and J. You. An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation, *Adv. Math.*, **226** (2011), 5361–5402.
[21] G. H. Hardy and E. M. Wright. An Introduction to the Theory of Number. *Oxford University Press*, Oxford, 2008.
[22] S. Krantz and H. Parks. A premier of real analytic functions. *Birkhäuser*, Basel, 2002.
[23] S. B. Kuksin. Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum. *Funct. Anal. Appl.*, **21** (1987), 192–205.
[24] S. B. Kuksin. Nearly integrable infinite-dimensional Hamiltonian systems. *Lecture Notes in Mathematics, 1556.* Springer-Verlag, Berlin, 1993.
[25] S. B. Kuksin. Analysis of Hamiltonian PDEs. Oxford University Press, 2000.
[26] S. B. Kuksin and J. Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. Math.* **143** (1996), 149–179.
[27] B. Ya. Levin. Distribution of zeroes of entire functions. *Translations of Mathematical Monographs, 5.* AMS, Providence, 1964.
[28] J. Moser and C. L. Siegel. Lectures on celestial mechanics. *Springer-Verlag*, Berlin, 1971.
[29] J. Pöschel. Quasi-periodic solutions for a nonlinear wave equation. *Comment. Math. Helv.* **71** (1996) 269–296.
[30] C. Procesi and M. Procesi. A normal form of the nonlinear Schrdinger equation with analytic non–linearities, *Comm. Math. Phys* **312** (2012), 501-557.
[31] C. Procesi and M. Procesi. A KAM algorithm for the resonant nonlinear Schrödinger equation, *preprint 2013*.
[32] I. M. Vinogradov. An Introduction to the Theory of Numbers. *Pergamon Press*, London and New York, 1955.
[33] W.-M. Wang, Energy supercritical nonlinear Schrödinger equations: Quasiperiodic solutions. *Duke Math J., in press*.

Univ. Paris Diderot, Sorbonne Paris Cité, Institut de Mathématiques de Jussieu-Paris rive gauche, UMR 7586, CNRS, Sorbonne Universités, UPMC Univ. Paris 06, F-75013, Paris, France

E-mail address: `hakan.eliasson@imj-prg.fr`

Laboratoire de Mathématiques Jean Leray, Université de Nantes, UMR CNRS 6629, 2, rue de la Houssinière, 44322 Nantes Cedex 03, France

E-mail address: `benoit.grebert@univ-nantes.fr`

CNRS, Institut de Mathématiques de Jussieu-Paris rive gauche, UMR 7586, Univ. Paris Diderot, Sorbonne Paris Cité, Sorbonne Universités, UPMC Univ. Paris 06, F-75013, Paris, France

E-mail address: `sergei.kuksin@imj-prg.fr`