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

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

ABSTRACT. In this paper we prove a KAM result for the nonlinear beam equation on the d-dimensional torus

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

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

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

persist as invariant tori of the nonlinear equation $(\ast)$, 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 nonlinear physics that small-amplitude solutions of space-multidimensional hamiltonian PDEs behave in a chaotic way.

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

24/1/ 2015

CONTENTS

1. Introduction \hspace{1cm} 2
  1.1. The beam equation and the KAM for PDE theory \hspace{1cm} 2
  1.2. Statement of the main result \hspace{1cm} 4
2. Small divisors \hspace{1cm} 8
  2.1. Non resonance of basic frequencies \hspace{1cm} 8
  2.2. Small divisors estimates \hspace{1cm} 10
3. The normal form \hspace{1cm} 13
  3.1. Notation and statement of the theorem \hspace{1cm} 13
  3.2. Resonances and the Birkhoff procedure \hspace{1cm} 16
  3.3. Normal form for admissible sets $\mathcal{A}$ \hspace{1cm} 20
  3.4. Eliminating the non integrable terms \hspace{1cm} 21
  3.5. Rescaling the variables and defining the transformation $\Phi$ \hspace{1cm} 23
  3.6. Real variables and final normalisation \hspace{1cm} 25
4. KAM \hspace{1cm} 27
  4.1. An abstract KAM result \hspace{1cm} 27
  4.2. KAM for the beam equation \hspace{1cm} 29
Appendix A. Proof of Lemma \hspace{1cm} 31
Appendix B. Examples \hspace{1cm} 32

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:

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

where \( m > 0 \) is the mass and \( g \) is a real analytic function on \( \mathbb{T}^d \times I \) for some neighbourhood \( I \) of the origin in \( \mathbb{R} \), satisfying

\[ g(x,u) = 4u^3 + O(u^4). \]

This equation is interesting by itself. Besides, it is a good model for the Klein–Gordon equation

\[ u_{tt} - \Delta u + mu = -g(x,u), \quad x \in \mathbb{T}^d, \]

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 these solutions with time-quasiperiodic solution of the linearised at zero equation

\[ u_{tt} + \Delta^2 u + mu = 0. \]

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

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

(here c.c. stands for “complex conjugated”), we write time-quasiperiodic solutions for (1.4), corresponding to a finite set of wave-vectors \( A \subset \mathbb{Z}^d \), as

\[ u(t,x) = \sum_{s \in A} (a_s e^{i \lambda_s t} + b_s e^{-i \lambda_s t}) e^{i s \cdot x} + \text{c.c.}, \]

where \( \lambda_s = \sqrt{|s|^4 + m} \). We wish to establish that for most small values of the action-vector \( \{ \frac{1}{2}(a_s^2 + b_s^2), \quad s \in A \} \), the solution (1.5) persists as a time-quasiperiodic solution of (1.1). In our work this goal is achieved provided that

- the finite set \( A \) 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 \( d \geq 2 \), then some of them are linearly unstable.

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 [20, 21, 11]. 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 |A|).

Next the theory was applied to

b) perturbations of integrable Hamiltonian PDE, e.g. of the KdV or Sine-Gordon equations, see [22].

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 $A$ most of 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 [22] 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 [23] 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 [26] the method of [23] was applied to the 1d equation (1.3). The approach of [23] turned out to be very efficient and later was applied to many other 1d Hamiltonian PDEs.

Space-multidimensional KAM for PDE theory started 10 years later with the paper [8] and, next, publications [9] and [16]. The just mentioned works deal with parameter-depending linear equations (cf. a)). The approach of [16] 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 [29] and [27, 28], using the KAM-techniques of [8, 9] and [16], respectively. Both works deal with the NLS equation. Their main disadvantage compare to the 1d theory c) is severe restrictions on the finite set $A$. The result of [29] gives examples of some sets $A$ for which the KAM-persistence of the corresponding small-amplitude solutions (1.5) holds, while the result of [27, 28] applies to solutions (1.5), where the set $A$ 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 $A$.

Some KAM-theorems for small-amplitude solutions of multidimensional beam equations (1.1) with typical $m$ were obtained in [17, 18]. 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 [29, 27, 28], the theorems of [17, 18] only allow to perturb solutions (1.5) with very special sets $A$. Solutions of (1.1), constructed in these works, all are linearly stable.

---

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 $O(u^4)$. 
1.2. Statement of the main result. Introducing $v = u_t ≡ \dot{u}$ we rewrite (1.1) as

(1.6) \[
\begin{align*}
\dot{u} &= -v, \\
\dot{v} &= \Lambda^2 u + g(x,u),
\end{align*}
\]

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

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

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

\[
\dot{\psi} = i\frac{\partial H}{\partial \overline{\psi}}
\]

with the Hamiltonian function

\[
H(\psi, \overline{\psi}) = \int_{\mathbb{T}^d} (\Lambda \psi)\overline{\psi} dx + \int_{\mathbb{T}^d} G\left(x, \Lambda^{-1/2}\left(\frac{\psi + \overline{\psi}}{\sqrt{2}}\right)\right) dx.
\]

where $G$ 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 $\Lambda$ is diagonal in the complex Fourier basis

\[
\{\varphi_s(x) = (2\pi)^{-d/2}e^{isx}, s \in \mathbb{Z}^d\}.
\]

Namely,

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

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

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

On the space $P_C := \ell^2(\mathbb{Z}^d, \mathbb{C}) \times \ell^2(\mathbb{Z}^d, \mathbb{C})$, endowed with the complex symplectic structure $-i\sum d\xi_s \wedge d\eta_s$, we consider the Hamiltonian system

(1.7) \[
\begin{align*}
\dot{\xi}_s &= i\frac{\partial H}{\partial \eta_s}, \\
\dot{\eta}_s &= -i\frac{\partial H}{\partial \xi_s},
\end{align*}
\]

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

(1.8) \[
H_2 = \sum_{s \in \mathbb{Z}^d} \lambda_s \xi_s \eta_s, \quad P = \int_{\mathbb{T}^d} G\left(x, \sum_{s \in \mathbb{Z}^d} \frac{\xi_s \varphi_s + \eta_{-s} \overline{\varphi}_s}{\sqrt{2\lambda_s}}\right) dx.
\]

The beam equation (1.1) is then equivalent to the Hamiltonian system (1.7), restricted to the real subspace

\[
P_R := \{(\xi, \eta) \in \ell^2(\mathbb{Z}^d, \mathbb{C}) \times \ell^2(\mathbb{Z}^d, \mathbb{C}) \mid \eta_s = \overline{\xi}_s, s \in \mathbb{Z}^d\}.
\]

The leading part of $P$ at the origin,

(1.9) \[
P_4 = \int_{\mathbb{T}^d} u^4 dx = \int_{\mathbb{T}^d} \left(\sum_{s \in \mathbb{Z}^d} \frac{\xi_s \varphi_s + \eta_{-s} \overline{\varphi}_s}{\sqrt{2\lambda_s}}\right)^4 dx,
\]
satisfies the zero momentum condition, i.e.

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

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.

Let \( \mathcal{A} \) be a finite subset of \( \mathbb{Z}^d \), \( |\mathcal{A}| = n \), and let us take a vector with positive components \( I = (I_a)_{a \in \mathcal{A}} \in \mathbb{R}_+^n \). The \( n \)-dimensional real torus

\[ T^n_I = \left\{ \begin{array}{l} \xi_a = \bar{\eta}_a, \ |\xi_0|^2 = I_a, \ a \in \mathcal{A} \\ \xi_s = \eta_s = 0, \ s \in \mathbb{Z}^d \setminus \mathcal{A}, \end{array} \right. \]

is invariant for the linear Hamiltonian flow when \( P = 0 \) (i.e. \( g = 0 \) in (1.11)). Our goal is to prove the persistency of most of the tori \( T^n_I \) when the perturbation \( P \) turns on, assuming that the set of nodes \( \mathcal{A} \) is admissible in the following sense:

**Definition 1.1.** A finite set \( \mathcal{A} \in \mathbb{Z}^d \) is called admissible if

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

The admissible sets \( \mathcal{A} \) are typical in the sense that if we take at random \( n \) integer points in the integer cube \( K_N^d = \{ a \in \mathbb{Z}^d : -N \leq a_j \leq N \forall j \} \), then the probability \( \pi(n, d, N) \) that the obtained \( n \)-points set is not admissible decays with \( N \) as \( N^{-1} \). Indeed, to get a random \( n \)-tuple in \( K_N^d \) we cut the solid cube \( K_N^d = \{ x \in \mathbb{R}^d : -N \leq x_j \leq N + 1 \} \) to \( (2N + 1)^d \) integer cubes of unit size and parametrise each cube by its lower left edge, which is a point in \( K_N^d \). Next we take independent random variables \( \xi^1, \ldots, \xi^n \), uniformly distributed in \( K_N^d \). They belong to some \( n \) unit cubes which define \( n \) random points in \( K_N^d \). The probability that the corresponding \( n \)-points set is not admissible is less that the probability that the difference between the lengths of some two vectors \( \xi^j, \xi^k \) is \( \leq 2\sqrt{d} \). Therefore

\[ \pi(n, d, N) \leq \mathbb{P}\{ ||\xi^j|| - ||\xi^k|| \leq 2\sqrt{d} \text{ for some } j \neq k \} \]

\[ \leq \frac{n(n-1)}{2} \mathbb{P}\{ ||\xi^1|| - ||\xi^2|| \leq 2\sqrt{d} \} \]

\[ \leq \frac{n(n-1)}{2(2N+1)^d} \int_{(K_N^d)^n} \chi_{||x_1|| - ||x_2|| \leq 2\sqrt{d}} \, dx^1 \ldots dx^n. \]

A straightforward (but a bit cumbersome) calculation shows that the r.h.s. is \( \leq \text{Const} \, N^{-1} \).

We denote

\[ \mathcal{L} = \mathbb{Z}^d \setminus \mathcal{A} \]

and set

\[ \mathcal{L}_f = \{ s \in \mathcal{L} | \exists a \in \mathcal{A} \text{ such that } |a| = |s| \}. \]

\[ \mathcal{L}_\infty = \mathcal{L} \setminus \mathcal{L}_f. \]

Clearly \( \mathcal{L}_f \) is a finite subset of \( \mathcal{L} \).

**Example 1.2.** If \( d = 1 \) and \( \mathcal{A} \) is admissible, then \( \mathcal{A} \cap -\mathcal{A} \subset \{0\} \) and \( \mathcal{L}_f = -(\mathcal{A} \setminus \{0\}) \).
In a neighbourhood of an invariant torus $T^n$ in $\mathbb{C}^{2n} = \{(\xi_a = \bar{\eta}_a, a \in \mathcal{A})\}$, we introduce the action-angle variables $(r_a, \theta_a)_{\mathcal{A}}$ by the relation

$$\xi_a = \sqrt{(I_a + r_a)(\cos \theta_a + i \sin \theta_a)}$$

(note that $-i \sum_{a \in \mathcal{A}} d\xi_a \wedge d\eta_a = -dI \wedge d\theta$). We will often denote the internal frequencies by $\omega$, i.e. $\lambda_s = \omega_s$ for $s \in \mathcal{A}$, and we will keep the notation $\lambda_s$ for the external frequencies with $s \in \mathcal{L} = \mathbb{Z}^d \setminus \mathcal{A}$.

The quadratic part of the Hamiltonian then 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 is a function of all variables and reads

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

where $\hat{u}_{I,m}(r, \theta, \xi, \eta)$ is $u(x) = (\psi + \eta)/\sqrt{2}$, expressed in the variables $(r, \theta, \xi_s, \eta_s)$:

$$(1.12) \quad \hat{u}_{I,m} = \sum_{a \in \mathcal{A}} \sqrt{I_a + r_a} e^{-i\theta_a \varphi_a(x)} + \sum_{s \in \mathcal{L}} \frac{\xi_s \varphi_s(x) + \eta_s \varphi_s(x)}{\sqrt{2}(|s|^4 + m)^{1/4}}.$$

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

$$r_a(t) = 0, \quad \theta_a(t) = \theta^0_a + t \omega_a$$

for $a \in \mathcal{A}$, is a solution of the linear beam equation (1.14), lying on the torus $T^n$. Our main theorem analyses persistence of these solutions in eq. (1.1) for typical vectors $I$.

**Theorem 1.3.** Assume that the nonlinearity $g(x, u) = 4u^3 + O(u^4)$ is analytic, and that the set $\mathcal{A}$, $|\mathcal{A}| = n$, is admissible. Then there exists a zero-measure Borel set $\mathcal{C} \subset [1, 2]$ and a Borel function $v_0 : [1, 2] \to \mathbb{R}$, strictly positive outside $\mathcal{C}$, such that for $m \notin \mathcal{C}$ and $0 < \nu \leq v_0(m)$

1) we can find a Borel set $\mathcal{D}_m \subset [\nu, 2\nu]^n$ asymptotically of full measure as $\nu \to 0$, i.e. satisfying \(\text{meas}([\nu, 2\nu]^n \setminus \mathcal{D}) \leq C(m)\nu^{n+\alpha}\) with some $\alpha = \alpha(\mathcal{A}) > 0$, and a mapping

$$U : \mathbb{T}^n \times \mathcal{D}_m \to \mathcal{P}_\mathbb{R} \subset \ell^2(\mathbb{Z}^d, \mathcal{C}) \times \ell^2(\mathbb{Z}^d, \mathcal{C}),$$

analytic in the first argument, such that

$$(1.13) \quad \text{dist} \left( U(\mathbb{T}^n \times \{I\}), T^n \right) \leq C(m, s, \mathcal{A})\nu^{\beta}, \quad \beta = \beta(\mathcal{A}) > 0,$$

and a vector-function $\omega' = \omega'_m : [0, \nu]^n \to \mathbb{R}^n$, $||\omega' - \omega||_{C^1} \leq C(m)\nu^{\beta}$, such that, for $I \in \mathcal{D}_m$ and $\theta \in \mathbb{T}^n$ the curve

$$(1.14) \quad t \mapsto U(\theta + t\omega'(I), I)$$

is a solution of the beam equation (1.14). Accordingly, for each $I \in \mathcal{D}_m$ the analytic $n$-torus $U(\mathbb{T}^n \times \{I\})$ is invariant for eq. (1.14).

2) A solution (1.14) is linearly unstable if certain matrix $iJK$, explicitly constructed in terms of the set $\mathcal{A}$ (see (3.4))), is unstable. This never happens if $d = 1$, while for $d \geq 2$ for some choices of the set $\mathcal{A}$ the solution is linearly unstable.

**Amplifications.** Relation (1.13) remains true if dist is distance with respect to the stronger norm $|| \cdot ||_0$, defined in Section 3.1.

As we explained above, the assumption that the set $\mathcal{A}$ is admissible is mild, at last when $|\mathcal{A}| \gg 1$. Still it is restrictive, and the union of the time-quasiperiodic
solutions of the linear beam equation \( (1.7)_{G=0} \), corresponding to admissible sets \( \mathcal{A} \), is not dense in the phase-space \( \ell^2(\mathbb{Z}^d, \mathbb{C}) \times \ell^2(\mathbb{Z}^d, \mathbb{C}) =: \ell^2 \times \ell^2 \). Accordingly, in difference with the 1d case (see \[23\]), the union of all KAM-solutions \( (1.14) \), constructed in Theorem 1.3, is not asymptotically dense at the origin of \( \ell^2 \times \ell^2 \).

We will deduce Theorem 1.3 from a normal form Theorem 4.1 (more involved than the 1d normal forms in \[23, 26\]), and an abstract KAM theorem for multidimensional PDEs, proved in \[14\]. Note that our result applies to eq. (1.1) with any \( d \), and that for \( d \) sufficiently large the global in time well-posedness of this equation is unknown.

For \( d \geq 2 \) many of the small-amplitude time-quasiperiodic solutions of the beam equation \( (1.1) \), constructed in Theorem 1.3, are linearly unstable. Their closures are unstable finite-dimensional invariant tori of the equation, situated in the vicinity of the origin, which creates around them some local instabilities. It is unclear whether 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 \[10\].

We note that the fact that KAM-solutions of high dimensional PDEs may be linearly unstable is not new: in \[19\] the instability of some KAM-solutions for the 2d cubic NLS equation was observed (see there Remark 1.1), while in \[27, 28\] algebraic reasons for the instability of KAM-solutions for various multidimensional NLS equations were discussed.

**Notation.**

- **Matrices.** For any matrix \( A \), finite or infinite, we denote by \( ^tA \) the transposed matrix; in particular, \( ^t(a, b) = \begin{pmatrix} a \\ b \end{pmatrix} \). By \( J \) we denote the matrix \( \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \) as well as various block-diagonal matrices \( \text{diag} \left( \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right) \).

- **Norms and pairings.** By \( \langle \cdot, \cdot \rangle \) we denote complex-linear pairing of complex spaces of finite or infinite dimension. All finite-dimensional spaces we consider are given the Euclidean norm which we denote \( | \cdot | \), and the corresponding distance. The tori are provided with the Euclidean distance.

- **Analytic mappings.** We call analytic mappings between domains in complex Banach spaces holomorphic to reserve the name analytic for real-analytic mappings. A holomorphic mapping is called real holomorphic if it maps real-vectors of the space-domain to real vectors of the space-target.

- **Parameters.** Our functions depend on parameters \( \rho \in \mathcal{D} \), 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 may be extended 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.

\[2\] I.e., this union does not intersect some non-trivial open cones in \( \ell^2 \times \ell^2 \) with vertexes at the origin.
2. Small divisors

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

We consider the vector of basic frequencies

\[
\omega \equiv \omega(m) = (\omega_\alpha(m))_{\alpha \in A}, \quad m \in [1, 2],
\]

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

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

\[
\text{meas} \left\{ m \in [1, 2] \mid \left| \sum_{\alpha \in A} k_\alpha \omega_\alpha(m) + c \right| \leq \kappa \right\} \leq C_n^{|k|^1/n},
\]

where \( |k| := \sum_{\alpha \in A} |k_\alpha| \) 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]. All the constants \( C_j \) etc. in this section do not depend on the set \( A \).

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

\[
D := \begin{vmatrix}
\frac{d\omega_{a_1}}{dm} & \frac{d\omega_{a_2}}{dm} & \cdots & \frac{d\omega_{a_p}}{dm} \\
\frac{d^2\omega_{a_1}}{dm^2} & \frac{d^2\omega_{a_2}}{dm^2} & \cdots & \frac{d^2\omega_{a_p}}{dm^2} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{d^p\omega_{a_1}}{dm^p} & \frac{d^p\omega_{a_2}}{dm^p} & \cdots & \frac{d^p\omega_{a_p}}{dm^p}
\end{vmatrix}
\]

is bounded from below:

\[
|D| \geq CN^{-3p^2+p},
\]

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

**Proof.** First note that, by explicit computation,

\[
\frac{d^l \omega_i}{dm^l} = (-1)^j \gamma_j (|a|^4 + m)^{\frac{l}{2} - 1}, \quad \gamma_j = \prod_{l=0}^{j-1} \frac{2l - 1}{2}.
\]

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

\[
\prod_{i=1}^p \omega_i^{-1} \prod_{j=1}^p \gamma_j \times \begin{vmatrix}
1 & 1 & 1 & \cdots & 1 \\
x_{a_1} & x_{a_2} & x_{a_3} & \cdots & x_{a_p} \\
x_{a_1}^2 & x_{a_2}^2 & x_{a_3}^2 & \cdots & x_{a_p}^2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
x_{a_1}^p & x_{a_2}^p & x_{a_3}^p & \cdots & x_{a_p}^p
\end{vmatrix},
\]
where we denoted $x_a := (|a|^4 + m)^{-1} = \omega_a^{-2}$. Since $|\omega_a| \leq 2|a_k|^2 \leq 2N^2$ 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

$$
\prod_{1 \leq t < k \leq p} (x_{at} - x_{ak}) = \prod_{1 \leq t < k \leq p} \frac{|a_t|^4 - |a_r|^4}{\omega_{a_t}^2 \omega_{a_k}^2} =: V.
$$

Since $A$ is admissible, then

$$|V| \geq \prod_{1 \leq t < k \leq p} \frac{|a_t|^2 + |a_r|^2}{\omega_{a_t}^2 \omega_{a_k}^2} \geq \left(\frac{1}{4}\right)^{p(p-1)} N^{-3p(p-1)},$$

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

**Lemma 2.3.** Let $u^{(1)}, \ldots, u^{(p)}$ 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 $i \in \{1, \ldots, p\}$ such that

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

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

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

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

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

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

$$|\frac{d^r \omega}{dm^r}(m) \cdot w| = K_{r} |u^{(r)} \cdot w| \geq K_{r} C_n |w|(K_1 \cdots K_n)^{-1} |D| \geq C_n |w| N^{-3n^2 + n}.$$

Now we need the following result (see Lemma B.1 in [13]):

**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 $\max_{1 \leq k \leq n} \min_{x} |\partial^k g(x)| = \sigma$. Then

$$\text{meas}\{x \mid |g(x)| \leq \rho\} \leq C_n \left(\frac{\beta}{\sigma} + 1\right) \left(\frac{\rho}{\sigma}\right)^{1/n}.$$

\^{3}In this section $C_n$ denotes any positive constant depending on $n$. 
Consider the function $g(m) = |k|^{-1} \sum_{a \in A} k_a \omega_a(m) + |k|^{-1} c$. Then $|g'|_{C^0} \leq C_n$ and $\max_{1 \leq k \leq n} \min_m |\partial^k g(m)| \geq C_n N^{-3n^2 + n}$ in view of (2.3). Therefore, by Lemma 2.4
\[
\text{meas} \{m \mid |g(m)| \leq \frac{K}{|k|} \} \leq C_n N^{3n^2 - n} \left( \frac{K}{|k|} N^{3n^2 - n} \right)^{1/n} = C_n N^{3n^2 + 2n - 1} \left( \frac{K}{|k|} \right)^{1/n}.
\]
This implies the assertion of the proposition.

2.2. Small divisors estimates. We recall the notation (1.10), (1.11), (2.1), and note the elementary estimate
\[
\max(1, |a|^2) < \lambda_0(m) < |a|^2 + \frac{m}{2|a|^2} \quad \forall a \in \mathbb{Z}^d, \ m \in [1, 2].
\]
In this section we study four type of linear combinations of the frequencies $\lambda_a(m)$:
\[
D_0 = \omega \cdot k, \quad k \in \mathbb{Z}^A \setminus \{0\},
\]
\[
D_1 = \omega \cdot k + \lambda_a, \quad k \in \mathbb{Z}^A, \ a \in \mathcal{L},
\]
\[
D_2^\pm = \omega \cdot k + \lambda_a \pm \lambda_b, \quad k \in \mathbb{Z}^A, \ a, b \in \mathcal{L}.
\]
In subsequent sections they will become divisors for our constructions, so we call these linear combinations "divisors".

Definition 2.5. Let $k \in \mathbb{Z}^A$ and $a, b \in \mathcal{L}$. Then $k$ is called $D_0$ resonant if $k = 0$;
(k; a) is $D_1$ resonant if $|a| = |s|$ for some $s \in A$ and $-\omega \cdot k = \omega_s$, so that $\omega \cdot k + \lambda_a \equiv 0$;
(k; a, b) is $D_2^\pm$ resonant if $|a| = |s|$, $|b| = |s'|$ with $s, s' \in A$ and $-\omega \cdot k = \omega_s \pm \omega_{s'}$, or $\omega = 0$, $|a| = |b|$ and the sign "±" in "-", so that $\omega \cdot k + \lambda_a \pm \lambda_b \equiv 0$.
The union of these three groups of linear combinations of frequencies is called the set of trivial resonances.

Note that $(k; a)$ can be $D_1$ resonant only when $a \in \mathcal{L}_f$, and $(k; a, b)$ can be $D_2^\pm$ resonant only when $(a, b) \in \mathcal{L}_f \times \mathcal{L}_f$. So there are only finitely many trivial resonances.

Our first aim is to remove from the segment $[1, 2] = \{m\}$ a small subset to guarantee that for the remaining $m$’s the moduli of the divisors $D_0, D_1, D_2^\pm$ admit a positive lower bound, except for the trivial resonances. Below in this section constants $C, C_1$ etc. depend on the admissible set $\mathcal{A}$,
\[
\text{while the exponents } c_1, c_2 \text{ etc depend only on } |\mathcal{A}|. \text{ Borel sets } C_k \text{ etc. depend on the indicated arguments and } \mathcal{A}.
\]
We begin with the easier divisors $D_0, D_1$ and $D_2^\pm$.

Proposition 2.6. Let $1 \geq \kappa > 0$. There exists a Borel set $C_\kappa \subset [1, 2]$ and positive constants $C$ (cf. (2.3)), satisfying $\text{meas } C_\kappa \leq C \kappa^{1/(n+2)}$, such that for all $m \notin C_\kappa$, all $k$ and all $a, b \in \mathcal{L}$ we have
\[
|\omega \cdot k| \geq \kappa |k|^{-n^2}, \quad \text{except if } k \text{ is } D_0 \text{ resonant, i.e. } k = 0,
\]
\[
|\omega \cdot k + \lambda_a| \geq \kappa |k|^{-3(n+1)^3}, \quad \text{except if } (k; a) \text{ is } D_1 \text{ resonant},
\]
Here \((k) = \max(|k|, 1)\).

Besides, for each \(k \neq 0\) there exists a set \(A_k^k\) whose measure is \(< C^1/n\) such that for \(m \notin A_k^k\) we have

\[
|\omega \cdot k + j| \geq \kappa(k)^{-(n+1)n} \text{ for all } j \in \mathbb{Z}.
\]

**Proof.** We begin with the divisors \((2.6)\). By Proposition 2.1 for any non-zero \(m / \mathbb{A}^k\) we have

\[
|\omega \cdot k + j| \geq \kappa(k)^{-(n+1)n} \text{ for all } j \in \mathbb{Z}.
\]

This relation with \((2.9)\) holds for all non-zero \(k \neq 0\), and therefore the required estimate holds trivially. We may assume that \(|\omega| \geq |k|^{1/2}\). If \(|\omega| \notin \{|s| : s \in A\}\), then Proposition 2.1 with \(n := n + 1, A := A \cup \{a\}\) and \(N = Ck^{1/2}\) implies that

\[
\text{meas}\{m \in [1, 2] \mid |\omega \cdot k + \lambda| \leq \kappa|k|^{-3(n+1)^3} \} \leq C^{1/(n+1)}|k|^{2(n(n+1)^2-3(n+1)^2-1/3)} \leq C^{1/(n+1)}|k|^{-n-1} = C^{k^{1/(n+1)}}.
\]

This relation with \(n + 1\) replaced by \(n\) also holds if \(|\alpha| = |s|\) for some \(s \in A\), but \(\omega \cdot k + \lambda\) is not a trivial resonant. Since for fixed \(k\) the set \(|\lambda_a \mid |a|^2 \leq C|k|\) has cardinality less than \(2C |k|\), then the relation \(|\omega \cdot k + \lambda| \leq \kappa|k|^{-3(n+1)^3}\) holds for a fixed \(k\) and all \(a\) if we remove from \([1, 2]\) a set of measure \(\leq C^{1/(n+1)}|k|^{-n-1}\). So we achieve that the relation \((2.7)\) holds for all \(k\) if we remove from \([1, 2]\) a set \(A_1\) whose measure is bounded by \(C^{1/(n+1)}|k|^{-n-1} = C^{k^{1/(n+1)}}\).

For similar reason there exist a Borel set \(A_2\) whose measure is bounded by \(C^{k^{1/(n+2)}}\) and such that \((2.8)\) holds for \(m \notin A_2\). Taking \(C_\kappa = A_0 \cup A_1 \cup A_2\) we get \((2.10)-(2.8)\). Proof of \((2.10)\) is similar.

Now we control divisors \(D^- = \omega \cdot k + \lambda_a - \lambda_b\).

**Proposition 2.7.** There exist positive constants \(C, c, c_\kappa\) and for \(0 < \kappa \leq C^{-1}\) there is a Borel set \(C_\kappa \subset [1, 2]\) (cf. \((2.6)\)), satisfying

\[
\text{meas} C_\kappa \leq C \kappa^{c_\kappa},
\]

such that for all \(m \in [1, 2] \setminus C_\kappa\), all \(k \neq 0\) and all \(a, b \in \mathcal{L}\) we have

\[
R(k; a, b) := |\omega \cdot k + \lambda_a - \lambda_b| \geq \kappa|k|^{-c_\kappa},
\]

except if \((k; a, b)\) is \(D^-\)-resonant.

**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 \(A^k_{\kappa_0}\) as in Proposition 2.6. Then \(\text{meas} A^k_{\kappa_0} \leq C^{1/n}\) and for any \(m \notin A^k_{\kappa_0}\) we have

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

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 $(\lambda_a, \lambda_b)$, satisfying
\begin{equation}
(2.12) \quad |a|^2 \leq Y_1, \quad |b|^2 \leq Y_1 + 2 + C|k| =: Y_2.
\end{equation}

There are at most $CY_1Y_2$ pairs like that. Since $(k; a, b)$ is not $D^-_\infty$ resonant, then in view of Proposition 2.1 with $N = Y_2^{1/2}$ and $|A| \leq n + 2$, for any $k > 0$ there exists a set $\mathfrak{B}_k \subset [1, 2]$, whose measure is bounded by

\[ Ck^{1/(n+2)}\kappa_0^{-c_1}|k|^{-c_2}, \quad c_j = c_j(n) > 0, \]

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

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

\[ \text{meas}(\mathfrak{C}_\kappa') \leq Ck^{c_3}, \]

and, if $m$ is outside $\mathfrak{C}_\kappa$, $R(k; a, b) \geq \kappa|k|^{-c_4}$ with suitable $c_4 > 0$. \qed

It remains to consider the divisors $D^-_\infty$ with $k = 0$:

**Lemma 2.8.** Let $m \in [1, 2]$ and $a, b \in \mathcal{L}$, $|a| \neq |b|$, then $|\lambda_a - \lambda_b| \geq \frac{1}{4}$.

**Proof.** We have

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

\[ \square \]

By construction the sets $\mathcal{C}_\kappa$ and $\mathcal{C}_\kappa'$ decrease with $\kappa$. Let us denote
\begin{equation}
(2.13) \quad \mathcal{C} = \bigcap_{\kappa > 0} (\mathcal{C}_\kappa \cup \mathcal{C}_\kappa').
\end{equation}

Then $\text{meas}\mathcal{C} = 0$, and 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 divisors
\[ \omega \cdot s, \quad \omega \cdot s \pm \lambda_a, \quad \omega \cdot s \pm \lambda_a \pm \lambda_b, \quad s \in \mathbb{Z}^d, \quad a, b \in \mathcal{L}, \]

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

In this section we construct a symplectic change of variable that puts the Hamiltonian \( H \) to a normal form, suitable to apply the abstract KAM theorem that we have proved in [14]. Our notation mostly agrees with [14]. Constants in the estimates may depend on the dimension \( d \), but this dependence is not indicated.

3.1. Notation and statement of the theorem. We start with recalling some notation from [14]. Let us fix any constant \( d^* > \frac{d}{2} \), and for \( \gamma \in [0, 1] \) denote by \( Y_\gamma \) the following weighted complex \( \ell_2 \)-space

\[
Y_\gamma = \{ \zeta = \left( \begin{array}{c} \xi_s \\ \eta_s \end{array} \right) \in \mathbb{C}^2, s \in \mathcal{L} \} \big| \| \zeta \|_\gamma < \infty \},
\]

where

\[
\| \zeta \|_\gamma^2 = \sum_{s \in \mathcal{L}} |\xi_s|^2 \langle s \rangle ^{2d^*} e^{2\gamma|s|}, \quad \langle s \rangle = \max(|s|, 1).
\]

In a space \( Y_\gamma \) we define the complex conjugation as the involution

\[
\zeta = t(\xi, \eta) \rightarrow t(\overline{\eta}, \overline{\xi}).
\]

Accordingly, the real subspace of \( Y_\gamma \) is the space

\[
Y_\gamma^R = \left\{ \zeta_s = \left( \begin{array}{c} \xi_s \\ \eta_s \end{array} \right) : \eta_s = \overline{\xi_s}, s \in \mathcal{L} \right\}.
\]

Any mapping defined on (some part of) \( Y_\gamma \) 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 \( A : \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{M}_{2 \times 2} \) valued in the space of \( 2 \times 2 \) matrices and satisfying

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

where

\[
[a - b] = \min(|a - b|, |a + b|).
\]

Let us define the operator

\[
D = \text{diag}\{ \langle s \rangle I, s \in \mathcal{L} \}
\]

(Here \( I \) stands for the identity \( 2 \times 2 \)-matrix). We denote by \( \mathcal{M}_D^\gamma \) the set of infinite matrices \( A \in \mathcal{M}_\gamma \) such that \( DAD \in \mathcal{M}_\gamma \), and set

\[
|A|_D^\gamma = |DAD|_\gamma = \sup_{a,b \in \mathcal{L}} \langle a \rangle \langle b \rangle |A_{ab}| \max([a - b], 1)^{d^*} e^{\gamma[a-b]}.
\]

We note that in [14] the norm \( |\cdot|_D^\gamma \) is denoted \( |\cdot|_\gamma^x \) with \( x = 2 \). Similar with other objects below whose notation involves the index \( D \).

For a Banach space \( B \) (real or complex) we denote

\[
\mathcal{O}_s(B) = \{ x \in B \mid \| x \|_B < s \},
\]

\[4\text{We recall that } |\cdot| \text{ signifies the Euclidean norm.} \]
We denote
\[ T^a_\sigma = \{ \theta \in \mathbb{C}^n / 2\pi \mathbb{Z}^n \mid |\Im \theta| < \sigma \}, \]
the contrary, if \( f \) is independent from \( s \), \( x \), \( r \), \( \theta \), \( \zeta \) = 0, linear in \( r \) and quadratic in \( \zeta \):
\[ h(x, \rho) = h^T(x, \rho) + O(|r|^2 + \|\zeta\|^3 + |r||\zeta|). \]
We denote
\[ T^{r \cdot} (\mu) = \{ f(\zeta) : f \in T^{r \cdot} (\sigma, \mu, D) \} \]
and for \( \sigma, \gamma, \mu \in (0, 1] \) we set
\[ T^a_\sigma = \{ \theta \in \mathbb{C}^n / 2\pi \mathbb{Z}^n \mid |\Im \theta| < \sigma \}, \]
\[ O^\gamma (\sigma, \mu) = O_{\mu^2} (\mathbb{C}^n \times T^a_\sigma \times O_\nu (Y_\gamma) = \{(r, \theta, \zeta)\}), \]
\[ O^{\gamma R} (\sigma, \mu) = O^\gamma (\sigma, \mu) \cap \{ \mathbb{R}^n \times T^n \times Y^R_\gamma \}. \]
We will denote points in \( O^\gamma (\sigma, \mu) \) as \( x = (\theta, r, \zeta) \).

The spaces \( Y^\sigma \) are important since functions with Fourier coefficients from \( Y^\sigma \) are holomorphic in \( T^a_\sigma \):

**Example 3.1.** If \( \hat{f} = (\hat{f}_s, s \in \mathbb{Z}^d) \in Y^\sigma \), then the function \( f(x) = \sum \hat{f}_s e^{i s \cdot x} \) is a holomorphic vector-function on \( T^a_\sigma \) and its norm is bounded by \( C_0 ||f||_\sigma \). On the contrary, if \( f : T^a_\sigma \to \mathbb{C}^n \) is a bounded holomorphic function, then its Fourier coefficients satisfy \( |\hat{f}_s| \leq \text{Const} \cdot e^{-s|\sigma|} \), so \( \hat{f} \in Y_\sigma \), for any \( \sigma' < \sigma \).

Let \( h : O^0 (\sigma, \mu) \times D \to \mathbb{C} \) be a \( C^1 \)-function, real holomorphic (see Notation) in the first variable \( x = (r, \theta, \zeta) \), such that for all \( 0 \leq \gamma' \leq \gamma \) and \( \rho \in D \) the gradient-map
\[ O^{\gamma'} (\sigma, \mu) \ni x \mapsto \nabla_\zeta f(x, \rho) \in Y_\gamma \]
and the hessian-map
\[ O^{\gamma'} (\sigma, \mu) \ni x \mapsto \nabla^2_\zeta f(x, \rho) \in M^{D}_{\gamma} \]
also are real holomorphic. We denote this set of functions by \( T^{\gamma \cdot} (\sigma, \mu, D) \). For a function \( h \in T^{\gamma \cdot} (\sigma, \mu, D) \) we define the norm
\[ [h]^{\gamma \cdot} = \sup_{0 \leq \gamma' \leq \gamma} \sup_{\rho \in D} \sup_{x \in O^{\gamma'} (\sigma, \mu)} \max(|\partial^j_\rho h(x, \rho)|, \mu^j|\partial^j_\rho \nabla_\zeta h(x, \rho)|^2, \mu^2 |\partial^j_\rho \nabla^2_\zeta h(x, \rho)|^2). \]
For any function \( h \in T^{\gamma \cdot} (\sigma, \mu, D) \) we denote by \( h^T \) its Taylor polynomial at \( r = 0, \zeta = 0 \), linear in \( r \) and quadratic in \( \zeta \):
\[ h(x, \rho) = h^T(x, \rho) + O(|r|^2 + \|\zeta\|^3 + |r||\zeta|). \]
We denote
\[ T^{\gamma \cdot} (\mu) = \{ f(\zeta) : f \in T^{\gamma \cdot} (\sigma, \mu, D) \} \]
\( (f \) is independent from \( \theta, r \) and \( \rho \); norm in \( T^{\gamma \cdot} (\mu) \) will be denoted \([h]^{\gamma \cdot} \).

Let \( P \) be the Hamiltonian function defined in \( [13] \).

**Lemma 3.2.** \( P \in T^{\gamma \cdot} (\mu^*) \) 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 \([13] \), in toroidal domains in the spaces \( Y_\gamma = \{ \zeta_s, s \in \mathbb{Z}^d \} \) which are neighbourhoods of the finite-dimensional real tori
\[ T_\rho = \{ \zeta = (\zeta_1, \ldots, \zeta_d), s \in \mathbb{Z}^d : |\zeta_s|^2 = \nu A_\rho^2 \text{ if } a \in A, \ z_s = 0 \text{ if } s \in \mathcal{L} \}, \]
KAM for the nonlinear beam equation

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

$$\mathcal{D} = [1, 2]^4.$$ 

In the vicinity of a torus we pass from the variables $(\zeta_a, a \in \mathcal{A})$, to the corresponding (complex) action-angles $(I_a, \theta_a)$, using the relations

$$\xi_a = \sqrt{T_a} e^{i\theta_a}, \quad \eta_a = \sqrt{T_a} e^{-i\theta_a}.$$

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

$$ (I, \theta, \xi, \eta) \rightarrow (\bar{I}, \bar{\theta}, \bar{\eta}, \bar{\xi}) $$

So a vector $(I, \theta, \xi, \eta)$ is real if $I = \bar{I}, \theta = \bar{\theta}, \xi = \bar{\eta}$. The toroidal vicinities of the tori $T_\rho$ (see (3.5)) will be of the form

$$ T_\rho = T_\rho(\nu, \sigma, \mu, \gamma) = \{ \zeta \in \mathbb{C} : |I - \nu \rho| < \nu \mu^2, |\Re \theta| < \sigma, |\Im \zeta| < \nu^{1/2} \mu \}, $$

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 $2 \geq \rho_j \geq 1$ for each $j$, then

$$ T_\rho(\nu, \sigma, \mu, \gamma) \cap Y_\gamma \subset \{ \zeta \in Y_\gamma : \text{dist}_{\gamma}(\zeta, T_\rho) < C \sqrt{\nu} \mu \} $$

if $\mu \leq \frac{1}{2}$, where $C > 0$ is an absolute constant.

**Theorem 3.3.** Let $\mathcal{A}$ be an admissible set. Then there exists a zero-measure set $\mathcal{C} \subset [1, 2]$, depending only on $\mathcal{A}$, and for each $m \in \mathcal{C}$ there exist real numbers $\gamma_\ast, \nu_0 \in (0, 1]$, where $\gamma_\ast$ depends only on $g(\cdot)$ and $\nu_0$ depends on $\mathcal{A}, m$ and $g(\cdot)$, such that

(i) For $0 < \nu \leq \nu_0$, $0 \leq \gamma \leq \gamma_\ast$ and $\rho \in \mathcal{D}$ there exists real holomorphic transformations

$$ \Phi_\rho : \mathcal{O}^\gamma(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}) \rightarrow T_\rho(\nu, 1, 1, \gamma), \quad 0 \leq \gamma \leq \gamma_\ast, $$

which coincide on the set $\mathcal{O}^\gamma(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}})$ and are diffeomorphisms on their images, analytically depending on $\rho$ and transforming the symplectic structure $-id\xi \wedge d\eta$ on $T_\rho(\nu, 1, 1, \gamma)$ to

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

The change of variable $\Phi_\rho$ is close to the scaling by the factor $\nu^{1/2}$ on the $\mathcal{L}_\infty$-modes but not on the $(\mathcal{A} \cup \mathcal{L})$-modes, where it is close to a certain affine transformation, depending on $\theta$. As a function of $\rho$, $\Phi_\rho$ holomorphically extends to the domain

$$ \mathcal{D}_{c_1} = \{ \rho \in \mathbb{C}^\mathcal{A} : |\Im \rho_j| < c_1, 1 - c_1 < \Re \rho_j < 2 + c_1 \forall j \in \mathcal{A} \}, \; c_1 > 0. $$

\footnote{so the index $\gamma$ does not enter the notation of the transformations.}
(ii) $\Phi_\rho$ puts the Hamiltonian function $H_2 + P$ to a normal form in the following sense:

$$
\frac{1}{\nu} (H_2 + P) \circ \Phi_\rho = \Omega(\rho) \cdot r + \sum_{a \in \mathcal{L}_\infty} \Lambda_a(\rho) \xi_\alpha \eta_\alpha
$$

(3.10)

$$
+ \frac{1}{\nu} \nu^2 (K(\rho) \zeta_f, \zeta_f) + f(r, \theta, \zeta; \rho).
$$

Here the vector $\Omega$ and the scalars $\Lambda_a$, $a \in \mathcal{L}_\infty$, are affine functions of $\rho$, while the symmetric complex matrix $K$ is a quadratic polynomial of $\sqrt{\rho} = (\sqrt{\rho}_1, \ldots, \sqrt{\rho}_n)$. They are defined by relations (3.37), and satisfies the estimates (3.38), and after the natural extension to $\mathcal{D}_{c_1}$ satisfy the estimates

$$
|\Omega(\rho) - \omega| \leq C_1 \nu, \quad |\Lambda_a(\rho) - \lambda_a(\rho)| \leq C_1 \nu |a|^{-2}, \quad \|K(\rho)\| \leq C_1.
$$

(3.11)

(iii) The remaining term $f$ belongs to $\mathcal{T}^{\gamma,D}(\frac{1}{2^2}, \frac{1}{2^2}, D)$ for each $0 \leq \gamma \leq \gamma_*$, and satisfies

$$
[f]^{\gamma,D}_{\frac{1}{2^2}, \frac{1}{2^2}} \leq C_2 \nu, \quad [f_1^{T}]^{\gamma,D}_{\frac{1}{2^2}, \frac{1}{2^2}} \leq C_2 \nu^{3/2}.
$$

(3.12)

$f$ is real holomorphic in $\rho \in \mathcal{D}_{c_1}$, and the estimates (3.12) hold for $f$ uniformly in $\rho \in \mathcal{D}_{c_1}$.

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

**Remark 3.4.** Properties of the Hamiltonian operator $L(\rho) := iJK(\rho)$ are crucial to study the behaviour of the beam equation in the toroidal domains $\mathcal{T}_\rho$. In Section 3.6 we show that for typical $\rho$ (i.e. for $\rho$ outside a small subset of $\mathcal{D}$) this operator is invertible. We can also prove that it decomposes to a direct sum $L(\rho) = L_1(\rho) \oplus \cdots \oplus L_\mathcal{J}(\rho)$, such that the linear spaces, where the linear mappings $L_j(\rho)$ operate, do not depend on $\rho$. Moreover, we know that for typical $\rho$ the operators $L_j(\rho)$ have simple spectrum (so $L(\rho)$ does not have Jordan cells). We also know that for $d = 2$ and for typical $\rho$ the whole operator $L(\rho)$ has simple spectrum. Unfortunately, we cannot establish this property if $d \geq 3$, and believe that, indeed when $d \geq 3$, for some admissible sets $\mathcal{A}$ the spectrum of $L(\rho)$ is multiple identically in $\rho$. It makes the proof of the KAM-theorem for the beam equation, given in Section 3.3 significantly more complicated in the sense that it has to evoke a rather sophisticated KAM-theorem, proven for this end in [14].

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

### 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)) in the complex variables $\zeta_\alpha = \iota(\xi_\alpha, \eta_\alpha)$:

$$
H_2 = \sum_{s \in \mathbb{Z}^d} \lambda_s \xi_s \eta_s,
$$

$$
P_4 = (2\pi)^{-d} \sum_{(i,j,k,\ell) \in \mathcal{J}} \frac{(\xi_i + \eta_i)(\xi_j + \eta_j)(\xi_k + \eta_k)(\xi_\ell + \eta_\ell)}{4\sqrt{\lambda_i \lambda_j \lambda_k \lambda_\ell}}.
$$

The factor $\nu^{-1}$ in the l.h.s. of (3.10) corresponds to $\nu$ in the transformed symplectic structure in item (i). So the Hamiltonian of the transformed equations with respect to the symplectic structure $-d\nu \wedge d\theta - i d\xi \wedge d\eta$ is given by the r.h.s. of (3.10).
Lemma 3.6. This implies the following result, where $\langle \cdot, \cdot \rangle$ stands for the complex-bilinear paring of $\mathbb{C}^2$ with itself:

For $(i, j, k, \ell) \in \mathbb{Z}^4$ we consider the linear combinations of the eigenvalues

\[ \Omega_0(i, j, k, \ell) = \lambda_i + \lambda_j + \lambda_k + \lambda_\ell, \]
\[ \Omega_1(i, j, k, \ell) = \lambda_i + \lambda_j + \lambda_k - \lambda_\ell, \]
\[ \Omega_2(i, j, k, \ell) = \lambda_i + \lambda_j - \lambda_k - \lambda_\ell. \]

They depend on $m$ since each $\lambda_j$ does.

**Definition 3.5.** A monomial $\xi_i \xi_j \xi_k \eta_\ell$ or $\eta_i \eta_j \eta_k \xi_\ell$ is called resonant if $\Omega_1(i, j, k, \ell) = 0$, in which case we denote $(i, j, k, \ell) \in \mathcal{R}_1$. A monomial $\xi_i \xi_j \eta_\ell \eta_k$ is called resonant if $\Omega_2(i, j, k, \ell) = 0$, in which case we denote $(i, j, k, \ell) \in \mathcal{R}_2$. We set $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$.

Finally we define

\[ \mathcal{J}_2 = \{(i, j, -k, -\ell) \in \mathcal{J} \mid \sharp \{i, j, k, \ell\} \cap A \geq 2\} \]

and denote by $\mathcal{J}_2^c$ the complementary set.

For later use we note that, by Proposition 3.9, if $m \notin \mathcal{C}$, then

\[ (i, j, k, \ell) \in \mathcal{R} \cap \mathcal{J}_2 \iff \begin{cases} (i, j, k, \ell) \in \mathcal{R}_2 \cap \mathcal{J}_2 \\ \{i, j\} = \{k, \ell\} \end{cases} \]

For $\gamma \geq 0$ we consider the phase space $Y_\gamma$, defined as in Section 3.1 with $\mathcal{L} = \mathbb{Z}^d$, and endowed it with the symplectic structure $-i \sum d\xi_k \wedge d\eta_k$. Since $d^* > d/2$, then the spaces $Y_\gamma$ are algebras with respect to the convolution, see Lemma 1.1 in [13]. This implies the following result, where $\langle \cdot, \cdot \rangle$ stands for the complex-bilinear paring of $\mathbb{C}^2r$ with itself:

**Lemma 3.6.** Let $\gamma \geq 0$, $r \in \mathbb{N}$ and $P^r$ be a real homogeneous polynomial on $Y_\gamma$ of degree $r$,

\[ P^r(\zeta) = \sum_{(j_1, \ldots, j_r) \in \mathcal{C}^r} a_{j_1, \ldots, j_r} \zeta_{j_1} \otimes \cdots \otimes \zeta_{j_r}, \]

where $a_{j_1, \ldots, j_r} \in \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ ($r$ times), $|a_{j_1, \ldots, j_r}| \leq M$, and $a_{j_1, \ldots, j_r} = 0$ unless $j_1 + \cdots + j_r = 0$. Then the gradient-map $\nabla P^r(\zeta)$ satisfies $\|\nabla P^r(\zeta)\|_{\gamma} \leq M^{r-1} \|\zeta\|^r_{\gamma}$. So the flow-maps $\Phi_{P^r_r}, |t| \leq 1$, of the Hamiltonian vector-field $X_{P^r} = i\nabla P^r$ are
well defined real holomorphic mappings on a ball $B_\gamma(\delta) = \{ ||\zeta||_\gamma < \delta \}$, $\delta = \delta(M) > 0$, and satisfy there
\[
\| \Phi_{\gamma,\psi}(\zeta) - \zeta \|_\gamma \leq C_1 \| \zeta \|_\gamma^{-1}, \quad C_1 = C_1(M).
\]

**Corollary 3.7.** Consider the polynomial $Q'(\zeta) = P'(\zeta)$, where $D^- = \text{diag}\{ |\lambda_\alpha|^{-1/2}I \}$. Then the Hessian-map $\nabla^2_-|_{Q'} \in M_\gamma^3$ and $|Q'|^2_\gamma \leq MC^{-2}|\zeta|^{\gamma-2}$ for any $\gamma \geq 0$. In particular $Q \in T^{\gamma,D}(\mu)$ for any $0 < \mu \leq 1$ (see (3.11)).

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

**Proposition 3.8.** For $m \notin \mathcal{C}$ there exists a real holomorphic and symplectic change of variable $\tau$ in a neighbourhood of the origin in $Y_\gamma$ that puts the Hamiltonian $H_4$ 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 $\gamma_*$ is as in Lemma 3.2, and for a suitable $\delta(m) \leq \delta_*$ (depending on $m$ and $g(x,u)$), the mapping $\tau$ satisfies
\[
\| \tau^\pm_1(\zeta) - \zeta \|_\gamma \leq C(m) \| \zeta \|_\gamma^3 \quad \forall \zeta \in B_\gamma(\delta(m)).
\]

It transforms the Hamiltonian $H_2 + P = H_2 + P_4 + R_5$ as follows:
\[
(H_2 + P) \circ \tau = H_2 + Z_4 + Q_4^3 + R_6^0 + R_5 \circ \tau,
\]
where
\[
Z_4 = \frac{3}{2} (2\pi)^{-d} \sum_{(i,j,k,l) \in J_2 \cap R_2} \frac{\xi_i \xi_j \eta_k \eta_l}{\lambda_i \lambda_j},
\]
and $Q_4^3 = Q_{4,1} + Q_{4,2}$ with $\text{7}$
\[
Q_{4,1} = (2\pi)^{-d} \sum_{(i,j,-k,l) \in J_2^5} \frac{\xi_i \xi_j \lambda_k \eta_l + \eta_k \eta_l \lambda_i \lambda_j}{\sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}},
\]
\[
Q_{4,2} = \frac{3}{2} (2\pi)^{-d} \sum_{(i,j,k,l) \in J_2^5} \frac{\xi_i \xi_j \eta_k \eta_l}{\sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}}.
\]
The functions $Z_4, Q_4^3, R_6^0, R_5 \circ \tau$ are real holomorphic on $B_\gamma(\delta(m))$, besides $R_6^0$ 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 $Z_4, Q_4^3, R_6^0$ and $R_5 \circ \tau$ belong to $T^{\gamma,D}(\mu)$ (see (3.11)), and
\[
[Z_4]_{\mu}^{\gamma,D} + [Q_4^3]_{\mu}^{\gamma,D} \leq C \mu^4,
\]
\[
[R_6^0]_{\mu}^{\gamma,D} \leq C \mu^6,
\]
\[
[R_5 \circ \tau]_{\mu}^{\gamma,D} \leq C \mu^5,
\]
where $C$ depends on $A$, $m$ and $g$.

---

$^7$The upper index 3 signifies that $Q_4^3$ is at least cubic in the transversal directions $\{ \zeta_\alpha, \alpha \in \mathcal{L} \}$. 
Proof. We use the classical Birkhoff normal form procedure. We construct the transformation \( \tau \) as the time one flow \( \Phi_{\chi_4}^t \) of a Hamiltonian \( \chi_4 \), given by

\[
\chi_4 = -\frac{i}{4}(2\pi)^{-d} \sum_{(i,j,k,l) \in J} \xi_i \xi_j \xi_k \xi_l - \eta_i \eta_j \eta_k \eta_l \Omega_0(i,j,k,l) \sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}
\]

\[-i(2\pi)^{-d} \sum_{(i,j,-k,l) \in J_2} \xi_i \xi_j \eta_k \eta_l - \eta_i \eta_j \xi_k \xi_l \Omega_1(i,j,k,l) \sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}
\]

\[-\frac{3}{2}(2\pi)^{-d} \sum_{(i,j,k,l) \in J_2 \setminus R_2} \xi_i \xi_j \eta_k \eta_l \Omega_2(i,j,k,l) \sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}.
\]

By Propositions 2.9 relation 3.14 and Lemma 3.6 for \( m \not\in \mathbb{C} \) the vector-field \( X_{\chi_4} \) is real holomorphic in \( Y_1 \) and of order three at the origin. Hence \( \tau = \Phi_{\chi_4}^t \) is a real holomorphic and symplectic change of coordinates, defined in \( B_{\gamma}(\delta(m)) \), a neighbourhood vicinity of the origin in \( Y_1 \). By Lemma 3.6 it satisfies (3.14).

Since the Poisson bracket, corresponding to the symplectic form \(-id\xi \wedge d\eta\) is \( \{F,G\} = i(\nabla_\xi F, \nabla_\xi G) - i(\nabla_\eta F, \nabla_\eta G) \), and since \( \nabla_\eta H_2 = \lambda_4 \xi_4, \nabla_\xi H_2 = \lambda_4 \eta_4 \), then we calculate

\[
\{H_2, \chi_4\} = -\frac{1}{4}(2\pi)^{-d} \sum_{(i,j,k,l) \in J} \xi_i \xi_j \xi_k \xi_l + \eta_i \eta_j \eta_k \eta_l \sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}
\]

\[-(2\pi)^{-d} \sum_{(i,j,-k,l) \in J_2} \xi_i \xi_j \eta_k \eta_l + \eta_i \eta_j \xi_k \xi_l \sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}
\]

\[-\frac{3}{2}(2\pi)^{-d} \sum_{(i,j,k,l) \in J_2 \setminus R_2} \xi_i \xi_j \eta_k \eta_l \sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}.
\]

Therefore

\[
(H_2 + P_4) \circ \tau = H_2 + P_4 - \{H_2, \chi_4\} - \{P_4, \chi_4\}
\]

\[+ \int_0^1 (1-t)\{\{H_2 + P_4, \chi_4\}, \chi_4\} \circ \Phi_{\chi_4}^t dt
\]

\[= H_2 + Z_4 + Q_4^3 + R_6^0
\]

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

\[
R_6^0 = \{P_4, \chi_4\} + \int_0^1 (1-t)\{\{H_2 + P_4, \chi_4\}, \chi_4\} \circ \Phi_{\chi_4}^t dt.
\]

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 \( T^{\gamma,D}(\mu) \) for any \( 0 < \mu \leq 1 \) and the estimate (3.14) hold by Corollary 3.7. Concerning \( R_6^0 \), by construction this is a holomorphic function of order \( \geq 6 \) at the origin. Its reality follows from the equality (3.16), where all other functions are real. The inclusion \( R_6^0 \in T^{\gamma,D}(\mu) \) for any \( 0 < \mu \leq \delta(m) \) and the estimate (3.15) follow from the following three facts:

(i) \( \{H_2 + P_4, \chi_4\} = Z_4 + Q_4^3 \) and \( \chi_4 \) belong to \( T^{\gamma,D}(1) \) by Corollary 3.7

(ii) \( \{T^{\gamma,D}(1), T^{\gamma,D}(1)\} \in T^{\gamma,D}(\frac{1}{2}) \) (see Proposition 2.6 in [13]).

(iii) \( T^{\gamma,D}(\frac{1}{2})\circ \Phi_{\chi_4}^{t} \in T^{\gamma,D}(\delta(m)) \). In [13], Proposition 2.7, and [22], Lemma 10.7, this result is proven for a special class of Hamiltonians \( \chi_4 \), but the proof easily generalises to Hamiltonian \( \chi_4 \) as above.
Finally, since by Lemma 3.2 the function \( R_5 \) belongs to \( T^{\gamma, \ldots} (\mu_*) \), then in view of (iii) \( R_5 \circ \tau \in T^{\gamma, \ldots} (\frac{1}{\delta} \delta(m)) \). Re-denoting \( \frac{1}{\delta} \delta(m) \) to \( \delta(m) \) we get (3.17)-(3.19). □

Due to (3.15), if \( \zeta \in T_\rho(\nu, 1/2, 1/2, \gamma) \), \( 0 \leq \gamma \leq \gamma_* \), where \( \nu \leq C^{-1} \delta(m)^2 \) and \( C \) is an absolute constant (see (3.7)), then (3.21)

\[
\tau^{+1}(T_\rho(\nu, 1/2, 1/2, \gamma)) \subseteq T_\rho(\nu, 1, 1, \gamma),
\]

provided that \( \nu \leq C^{-1} \delta(m)^2 \) and \( \rho \in D_{c_1} \), where \( c_1 = c_1(A, m, g(\cdot)) \) is sufficiently small.

3.3. Normal form for admissible sets \( A \). Everywhere in this section the set \( A \) is admissible in the sense of Definition 1.1.

The Hamiltonian \( Z_4 \) contains the integrable part formed by monomials of the form \( \xi_i \xi_j \eta_k \eta_l \) 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

\[
Z_4^+ = \frac{3}{2} (2\pi)^{-d} \sum_{\ell \in A, k \in \mathbb{Z}^d} (4 - 3\delta_{\ell, k}) \frac{I_l I_k}{\lambda_\ell \lambda_k}.
\]

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

Since \( A \) is admissible, then in view of (3.14) for \( m \notin \mathcal{C} \) the set \( Z_4^{-3} \) is empty. The set \( Z_4^{-3} \) is empty as well:

**Lemma 3.9.** If \( m \notin \mathcal{C} \), then \( Z_4^{-3} = \emptyset \).

**Proof.** Consider any term \( \xi_i \xi_j \eta_k \eta_l \in Z_4^{-3} \), i.e. \( \{i, j, k, \ell\} \cap A = 3 \). Without lost of generality we can assume that \( i, j, k \in A \) and \( \ell \in \mathcal{L} \). Furthermore we know that \( i + j - k - \ell = 0 \) and

\[
\lambda_i + \lambda_k = \lambda_j + \lambda_\ell.
\]

By (3.14) we must have \( |i| = |k| \) or \( |j| = |k| \) and thus, since \( A \) 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. □

Recall that the finite set \( \mathcal{L}_f \subset \mathcal{L} \) was defined in (1.10). The mapping

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

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

\[
(\mathcal{L}_f \times \mathcal{L}_f)_+ := \{(a, b) \in \mathcal{L}_f \times \mathcal{L}_f \mid \ell(a) + \ell(b) = a + b\}
\]

\[
(\mathcal{L}_f \times \mathcal{L}_f)_- := \{(a, b) \in \mathcal{L}_f \times \mathcal{L}_f \mid a \neq b \text{ and } \ell(a) - \ell(b) = a - b\}.
\]

**Example 3.10.** If \( d = 1 \), then in view of Example 1.2 \( \ell(a) = -a \) and the sets \((\mathcal{L}_f \times \mathcal{L}_f)_\pm\) are empty.

For \( d \geq 2 \), in general, the sets \((\mathcal{L}_f \times \mathcal{L}_f)_\pm\) both are non-trivial, see Appendix B. Obviously \((\mathcal{L}_f \times \mathcal{L}_f)_+ \cap (\mathcal{L}_f \times \mathcal{L}_f)_- = \emptyset \). For further reference we note that

**Lemma 3.11.** 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 $(\mathcal{L}_f \times \mathcal{L}_f)_-$. \hfill \Box

According to the decomposition $\mathcal{L} = \mathcal{L}_f \cup \mathcal{L}_\infty$, the space $Y_\gamma$, defined in (3.1), decomposes in the direct sum
$$(3.27) \quad Y_\gamma = Y^f_\gamma \oplus Y^\infty_\gamma, \quad Y^f_\gamma = \text{span}\{\zeta_s, s \in \mathcal{L}_f\}, \quad Y^\infty_\gamma = \text{span}\{\zeta_s, s \in \mathcal{L}_\infty\}.$$

**Lemma 3.12.** Assume that $\mathcal{A}$ is admissible. Then for $m \notin \mathcal{C}$ the part $Z^{-2}_4$ of the Hamiltonian $Z^{-1}_4$ equals
$$3(2\pi)^{-d}\sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+} \frac{\xi_{\ell(a)}\xi_{\ell(b)}\varpi_{\gamma_a}\varpi_{\gamma_b}}{\lambda_a\lambda_b} + 2\sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_-} \frac{\xi_{a}\xi_{\ell(b)}\varpi_{\gamma(a)}\varpi_{\gamma(b)}}{\lambda_a\lambda_b}.$$  

(3.28)

Proof. Let $\xi_i\xi_j\varpi_k\varpi_l$ be a monomial in $Z^{-2}_4$. We know that $(i, j, -k, -\ell) \in \mathcal{J}$ and satisfies (3.23). In view of (3.14) we must have
$$(3.29) \quad \{|i|, |j|\} = \{|k|, |\ell|\}.$$

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 (3.28). Now we assume that
$$i, \ell \in \mathcal{A} \quad \text{and} \quad j, k \in \mathcal{L}.$$

Then from (3.29) we have that, either $|i| = |k|$ and $|j| = |\ell|$ which leads to finitely many monomials as in the second sum in (3.28). Or $i = \ell$ and $|j| = |k|$. In this last case, the zero momentum condition implies that $j = \ell$ which is not possible in $Z^{-1}_4$. \hfill \Box

### 3.4. Eliminating the non integrable terms.

For $\ell \in \mathcal{A}$ we introduce the variables $(I_a, \theta_a, \zeta^\ell_a)$ as in (3.7). Now the symplectic structure $-i d\xi \wedge d\eta$ reads
$$(3.30) \quad - \sum_{a \in \mathcal{A}} dI_a \wedge d\theta_a - i d\xi^\ell \wedge d\eta^\ell.$$

In view of (3.22), (3.10) and Lemma 3.12 for $m \notin \mathcal{C}$ the transformed Hamiltonian may be written as (recall that $\omega = (\lambda_a, a \in \mathcal{A})$)
$$(H_2 + P) \circ \tau = \omega \cdot I + \sum_{s \in \mathcal{L}} \lambda_s \xi_s \eta_s + \frac{3}{2} (2\pi)^{-d} \sum_{\ell \in \mathcal{A}, k \in \mathbb{Z}^d} (4 - 3\delta_{\ell, k}) \frac{I_{\ell}\xi_{\ell}\eta_{k}}{\lambda_{\ell}\lambda_k} \xi_{\ell} \eta_{k} + 3(2\pi)^{-d} \sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+} \frac{\xi_{\ell(a)}\xi_{\ell(b)}\varpi_{\gamma(a)}\varpi_{\gamma(b)}}{\lambda_a\lambda_b} + 2 \sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_-} \frac{\xi_{a}\xi_{\ell(b)}\varpi_{\gamma(a)}\varpi_{\gamma(b)}}{\lambda_a\lambda_b} + Q^1_2 + R^0_5, \quad R^0_5 = R_5 \circ \tau + R^0_6.$$  

The first line contains the integrable terms. The second and third lines contain the lower-order non integrable terms, depending on the angles $\theta$; there are finitely
many of them. The last line contains the remaining high order terms, where \( Q^3_d \) is of total order (at least) 4 and of order 3 in the normal directions \( \zeta \), while \( R_0^d \) is of total order at least 5. The latter is the sum of \( R_0^d \) 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 [12]. Here \( I \) is regarded as a variable of order 2, while \( \theta \) has zero order. The fourth line should be regarded as a perturbation.

To deal with the non integrable terms in the second line, following the works on the finite-dimensional reducibility (see [12]), we introduce a change of variables

\[
\Psi : (\tilde{\zeta}, \tilde{\eta}) \mapsto (\xi, \eta),
\]
symplectic with respect to (3.30), but such that its differential at the origin is not close to the identity. It is defined by the following relations:

\[
\begin{align*}
\xi_a &= \tilde{\xi}_a e^{i \tilde{\theta}(a)}, & \eta_a &= \tilde{\eta}_a e^{-i \tilde{\theta}(a)} & a \in L_f, \\
I_{\ell} &= \tilde{I}_\ell - \sum_{|a|=|\ell|, \ a \neq \ell} \tilde{\xi}_a \tilde{\eta}_a, & \theta_{\ell} &= \tilde{\theta}_{\ell} & \ell \in A, \\
\xi_a &= \tilde{\xi}_a, & \eta_a &= \tilde{\eta}_a & a \in L_\infty.
\end{align*}
\]

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

\[
|I - \nu \rho| \leq |\tilde{I} - \nu \rho| + |y| \leq \nu \mu^2 + \sum_{a \in L_f} |\tilde{\xi}_a \tilde{\eta}_a| \leq 2 \nu \mu^2.
\]

This implies that

\[
(3.31) \quad \Psi^{\pm 1}(T_\rho(\nu, 1, 2, \frac{1}{2\sqrt{2}}, \gamma)) \subset T_\rho(\nu, 1, 2, \frac{1}{2\sqrt{2}}, \gamma).
\]

Denote \( T_\rho(\nu, 1, 2, \frac{1}{2\sqrt{2}}, \gamma) =: T_\rho \).

If \( (\tilde{\zeta}, \tilde{\eta}) \in T_\rho \), then for \( l \in A \)

\[
\xi_l = \sqrt{l} I_l e^{i \theta_l} = \sqrt{l} I_l e^{i \theta_l} + O(\nu^{-1/2}) O(|\zeta |^2).
\]

Therefore, dropping the tildes, we write the restriction to \( T_\rho \) of the transformed Hamiltonian as

\[
H_1 := H \circ \tau \circ \Psi = \omega \cdot I + \sum_{a \in L_\infty} \lambda_a \xi_a \eta_a
+ 6(2\pi)^{-d} \sum_{\ell, k \in L} \frac{1}{\lambda_\ell \lambda_k} (I_{\ell} - \sum_{|a|=|\ell| \atop a \in L_f} \xi_a \eta_a) \xi_k \eta_k
+ \frac{3}{2} (2\pi)^{-d} \sum_{\ell, k \in A} \frac{4 - 3 \delta_{\ell, k}}{\lambda_\ell \lambda_k} (I_{\ell} - \sum_{|a|=|\ell| \atop a \in L_f} \xi_a \eta_a) (I_k - \sum_{|a|=|k| \atop a \in L_f} \xi_a \eta_a)
+ 3(2\pi)^{-d} \sum_{(a, b) \in L_f \times L_f^+} \sqrt{l(a) l(b)} \frac{\lambda_a \lambda_b}{\lambda_a \lambda_b} (\eta_a \eta_b + \xi_a \xi_b)
+ 6(2\pi)^{-d} \sum_{(a, b) \in L_f \times L_f^-} \sqrt{l(a) l(b)} \frac{\lambda_a \lambda_b}{\lambda_a \lambda_b} \xi_a \eta_b + Q^3_d + R_0^d + \nu^{-1/2} B_0^{d'}. \]
Here $Q_3'$ and $R_0'$ are the function $Q_1^4$ and $R_0^5$, transformed by $\Psi$ (so the former satisfy the same estimates as the latter), while $R_4'$ is a function of forth order in the normal variables. Or, after a simplification:

\[
H_1 = \omega \cdot I + \sum_{a \in \mathcal{L}_\infty} \lambda_a \xi_a \eta_a + \frac{3}{2} (2\pi)^{-d} \sum_{\ell, k \in \mathcal{A}} \frac{4 - 3\delta_{\ell, k}}{\lambda_\ell \lambda_k} I_\ell I_k \\
+ 3(2\pi)^{-d} \left( \sum_{\ell \in \mathcal{A}, a \in \mathcal{L}_\infty} \frac{1}{\lambda_\ell \lambda_a} I_\ell \xi_a \eta_a - \sum_{\ell \in \mathcal{A}, a \in \mathcal{L}_f} (2 - 3\delta_{\ell, |a|}) \frac{1}{\lambda_\ell \lambda_a} I_\ell \xi_a \eta_a \right) \\
+ 3(2\pi)^{-d} \sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+} \frac{\sqrt{I(a) I(b)}}{\lambda_a \lambda_b} (\eta_a \eta_b + \xi_a \xi_b) \\
+ 6(2\pi)^{-d} \sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_-} \frac{\sqrt{I(a) I(b)}}{\lambda_a \lambda_b} \xi_a \eta_b + Q_3' + R_0' + \nu^{-1/2} R_4'
\]  
(3.32)

We see that the transformation $\Psi$ 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 $\theta$, but depend on the actions $I$ and quadratically depend on finitely many variables $\xi_a, \eta_a$ 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 $\zeta^\mathcal{L} = t^I(\xi, \eta)$ as $\zeta = t^I(\xi, \eta)$.

3.5. Rescaling the variables and defining the transformation $\Phi$. Our aim is to study the Hamiltonian $H_1$ on the domains $T_{\rho} = T_{\rho}(\nu, \frac{1}{2}, \frac{1}{2}, \gamma)$, $0 \leq \gamma \leq \gamma^*$ (see (3.31)). To do this we re-parametrise points of $T_{\rho}$ by mean of the change of variables $(I, \theta, \xi, \eta) = \chi_\rho(\tilde{\nu}, \tilde{\theta}, \tilde{\xi}, \tilde{\eta})$, where

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

Clearly,

$$\chi_\rho : \mathcal{O}^\gamma(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}) \to T_{\rho},$$ 

and in the new variables the symplectic structure reads

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

Denoting

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

we see that this transformation is real holomorphic in $\rho \in D_{c_1}$ for a suitable $c_1 > 0$. It satisfies all assertions of the item (i) of Theorem 3.3.
We have:

\[
H \circ \Phi = \nu \left[ \omega \cdot r + \sum_{a \in \mathcal{L}_\infty} \lambda_a \xi_a \eta_a + (2\pi)^{-d} \nu \left( \frac{3}{2} \sum_{\ell, k \in A} \frac{4 - 3\delta_{\ell, k}}{\lambda_\ell \lambda_k} \rho_\ell \xi_a \eta_a \right) + 6 \sum_{\ell \in A, a \in \mathcal{L}_\infty} \frac{1}{\lambda_\ell \lambda_a} \rho_\ell \xi_a \eta_a - 3 \sum_{\ell, k \in A, a \in \mathcal{L}_f} \frac{(2 - 3\delta_{\ell, |a|})}{\lambda_\ell \lambda_a} \rho_\ell \xi_a \eta_a \right]
\]

\[
+ 3 \sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+} \frac{\sqrt{P_{\ell(a)}(a) P_{\ell(b)}(b)}}{\lambda_\ell \lambda_b} (\eta_a \xi_b + \xi_a \eta_b)
\]

\[
+ 6 \sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_-} \frac{\sqrt{P_{\ell(a)}(a) P_{\ell(b)}(b)}}{\lambda_\ell \lambda_b} \xi_a \eta_b)
\]

\[
+ \left( (Q^\alpha + R^\nu_\delta + \nu^{-1/2} R^\nu_\delta^*) (I, \theta, \sqrt{\nu} \xi) \right) |_{I=\nu \rho + \nu \tau}.
\]

So,

\[
\nu^{-1} H \circ \Phi = h + f,
\]

where \( h \equiv h(I, \xi, \eta; \rho, \nu) \) is the quadratic part of the Hamiltonian, independent from the angle \( \theta \), and \( f \) is the perturbation, given by the last line in (3.33):

\[
f = \nu^{-1} \left( (Q^\alpha + R^\nu_\delta + \nu^{-1/2} R^\nu_\delta^*) (I, \theta, \nu^{1/2} \xi) \right) |_{I=\nu \rho + \nu \tau}.
\]

We have

\[
h = \Omega \cdot r + \sum_{a \in \mathcal{L}_\infty} \Lambda_a \xi_a \eta_a + \nu(K(\rho) \zeta_f, \zeta_f)
\]

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

\[
\Omega_k = \Omega_k(\rho, \nu) = \omega_k + \nu \sum_{\ell \in A} M^\ell_k \rho_\ell, \quad M^\ell_k = \frac{3(4 - 3\delta_{\ell, k})}{(2\pi)^d \lambda_\ell \lambda_k},
\]

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

Besides,

\[
\zeta = (\zeta_a)_{a \in \mathcal{L}}, \quad \zeta_a = \left( \begin{array}{c} \xi_a \\ \eta_a \end{array} \right), \quad \zeta_f = (\zeta_a)_{a \in \mathcal{L}_f},
\]

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

\[
Y_f = \{ \zeta_f \} \simeq \mathbb{C}^{2|\mathcal{L}_f|},
\]

such that the corresponding quadratic form is

\[
\langle K(\rho) \zeta_f, \zeta_f \rangle = 3(2\pi)^{-d} \left( \sum_{\ell \in A, a \in \mathcal{L}_f} \frac{(3\delta_{\ell, |a|} - 2)}{\lambda_\ell \lambda_a} \rho_\ell \xi_a \eta_a \right)
\]

\[
+ \sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+} \frac{\sqrt{P_{\ell(a)}(a) P_{\ell(b)}(b)}}{\lambda_\ell \lambda_b} (\eta_a \eta_b + \xi_a \xi_b) +
\]

\[
2 \sum_{(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_-} \frac{\sqrt{P_{\ell(a)}(a) P_{\ell(b)}(b)}}{\lambda_\ell \lambda_b} \xi_a \eta_b).
\]
Note that the matrix $M$ in (3.37) is invertible since
\[ \det M = 3^n (2\pi)^{-dn} \left( \prod_{k \in A} \lambda_k \right)^{-2} \det \left( 4 - 3\delta_{\ell,k} \right)_{\ell,k \in A} \neq 0. \]
Relation (3.11) immediately follow from the explicit formulas (3.37)-(3.40), so the items (i) and (ii) of Theorem 3.3 are proven.

It remains to verify (iii). By Proposition 3.8 the function $f$ belongs to $T^{\gamma,D} \left( \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, D \right)$. Since the reminding term $f$ has the form (3.35) then for $(r, \theta, \zeta) \in O^{\gamma}(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}})$ it satisfies the estimates
\[ |f| \leq C\nu^{3/2}, \quad \|\nabla_{\zeta} f\|_{\gamma} \leq C\nu, \quad \|\nabla_{\zeta}^2 f\|_{D} \leq C\nu. \]
Now consider the $f^T$-component of $f$. Only the second term in (3.35) contributes to it and we have that
\[ |f^T| + \|\nabla_{\zeta} f^T\| + \|\nabla_{\zeta}^2 f^T\|_{D} \leq C\nu^{3/2}. \]
Recall that the function $f$ depends on the parameter $\rho$ through the substitution $I = \nu \rho + \nu r$. So $f$ is analytic in $\rho$ and holomorphically extends to a complex neighbourhood of $D$ of order one, where it satisfies the estimates above with a modified constant $C$. Therefore by the Cauchy estimate the gradient of $f$ in $\rho$ satisfies in the smaller complex neighbourhood $D_{c1}$ the same estimates as above, again with a modified constant. This implies the assertion (iii) of the theorem.

3.6. Real variables and final normalisation. The normal form, provided by Theorem 3.3 has two disadvantages: it is complex (while the original equation is real), and the Hamiltonian operator $iJK(\rho)$ may degenerate for some $\rho$. In this section we remove these flaws.

Matrix $K(\rho)$. The symmetric matrix $K(\rho)$, defined by relation (3.40), is a block-matrix, which is a quadratic polynomial in $\sqrt{\rho} = (\sqrt{\rho_1}, \ldots, \sqrt{\rho_n})$. We write it as $K(\rho) = K^d(\rho) + K^{n/d}(\rho)$. Here $K^d$ is the diagonal part of $K$, which is a block-matrix
\[
K^d(\rho) = \text{diag} \left( \begin{pmatrix} 0 & \mu(a) \\ \mu(a) & 0 \end{pmatrix}, a \in \mathcal{L}_f \right), \tag{3.41}
\]
\[ \mu(a) = C_* \left( \frac{3}{2} \rho_{\ell(a)} \lambda_{a}^{-2} - \lambda_{a}^{-1} \sum_{l \in A} \rho_{l} \lambda_{l}^{-1} \right), \quad C_* = 3 (2\pi)^{-d}. \]
The non-diagonal part $K^{n/d}$ has zero diagonal blocks, while for $a \neq b$ its block $K^{n/d}(\rho)^b_{b}$ equals
\[ C_* \frac{\sqrt{P_{\ell(a)}P_{\ell(b)}}}{\lambda_{a}\lambda_{b}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \chi^+(a,b) + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \chi^-(a,b), \]
where
\[ \chi^+(a,b) = \begin{cases} 1, & (a,b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+, \\ 0, & \text{otherwise}, \end{cases} \]
and $\chi^-$ is defined similar in terms of the set $(\mathcal{L}_f \times \mathcal{L}_f)_-$.

Lemma 3.13. The function $\det(iJK(\rho))$ is a polynomial of $\sqrt{\rho}$ which does not vanish identically.
Proof. We only need to check that \( \det(iJK(\rho)) \neq 0 \). Let us enumerate the elements of \( A \) as \( a_1, a_2, \ldots, a_n \), where \( |a_1| < |a_2| < \cdots < |a_n| \), and enumerate elements of \( \mathcal{L}_f \) as \( b_1, \ldots, b_N \), where \( |b_1| \leq |b_2| \leq \cdots \leq |b_N| \). Then \( l(b_j) = a_{l(j)} \), for some sequence \( 1 \leq \sigma(1) \leq \sigma(2) \leq \cdots \leq \sigma(N) \leq n \). To simplify notation assume that \( \sigma(1) = 1 \) (i.e., \( a_1 \neq 0 \)). Denote \( \sqrt{\rho} = y \in \mathbb{R}^n \). Then \( K^{n/d} \) and \( K^d \) are polynomial functions of \( y \). Consider \( y_* = (1, 0, \ldots, 0) \). Then \( K^{n/d}(y_*) = 0 \) and the numbers \( \mu(b)(y_*, b \in \mathcal{L}_f) \), take two values: \( \frac{1}{2}C_\sigma \lambda_{a_1}^{-2} \) and \( -C_\sigma \lambda_{a_1}^{-2} \). So the matrix \( K(y_*) = K^d(y_*) \) is non-degenerate, and \( \det(iJK(\rho)) \neq 0 \). \( \square \)

Real variables. Let us pass in \((3.10)\) from the complex variables \( \text{Real variables} \). \( \right \)

\[ (3.42) \]

\[ \xi_l = \frac{1}{\sqrt{2}}(u_l + iv_l), \quad \eta_l = \frac{1}{\sqrt{2}}(u_l - iv_l), \quad l \in \mathcal{L}, \]

keeping \( r \) and \( \theta \) unchanged. We denote this change of variable as

\[ \zeta = \Sigma(\tilde{\zeta}). \]

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

\[ \bar{u}_l = u_l, \quad \bar{v}_l = v_l \quad \forall l \in \mathcal{L}. \]

In the variables \((r, \theta, u, v)\) the symplectic form \(-dr \land d\theta - id\xi \land d\eta\) reads

\[ \omega_2 = -dr \land d\theta - du \land dv, \]

and the transformed Hamiltonian is

\[ H(r, \theta, \tilde{\zeta}; \rho) := (H_2 + P) \circ \Phi_\rho \circ \Sigma = \Omega(\rho) \cdot r + \frac{1}{2} \sum_{a \in \mathcal{L}_f} \Lambda_a(\rho)(u_a^2 + v_a^2) + \frac{\nu}{2}(H_0(\rho)\tilde{\zeta}_f, \tilde{\zeta}_f) + \tilde{f}(r, \theta, \tilde{\zeta}; \rho), \]

where \( \tilde{\zeta}_f = (4(u_a, v_a), a \in \mathcal{L}_f) \) and \( (H_0\tilde{\zeta}_f, \tilde{\zeta}_f) \) is the quadratic form \( \langle K\tilde{\zeta}_f, \tilde{\zeta}_f \rangle \), written in the variables \( \tilde{\zeta}_f \). So the spectrum of the operator \( JH_0(\rho) \) equals that of the operator \( iJK(\rho) \). By Lemma \((3.13)\) \( \det JH_0(\rho) = \det iJK(\rho) \) is a non-trivial polynomial of the vector \( \sqrt{\rho} \). For any \( \delta > 0 \) denote

\[ (3.44) \]

\[ D_\delta = \{ \rho \in D : |\det JH_0(\rho)| \geq \delta \}. \]

Since the transformation

\[ D = [1, 2]^n \to \mathbb{R}^n, \quad \rho \mapsto \sqrt{\rho} \]

is a diffeomorphism which changes the measure of a subset of \( D \) by a factor, bounded from below and from above by some absolute positive constants, then in view of Lemma \((3.1)\)

\[ (3.45) \]

\[ \text{meas} (D \setminus D_\delta) \leq C\delta^\tilde{c}, \]

where \( \tilde{c} > 0 \) depends only on \( A \) and \( d \). This estimate and Theorem \((3.3)\) imply

Proposition 3.14. Under the assumptions of Theorem \((3.3)\) there exists a real holomorphic transformation

\[ \tilde{\Phi}_\rho = \Phi_\rho \circ \Sigma, \quad \tilde{\Phi}_\rho : \mathcal{O}^\gamma \left( \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}} \right) \to T_\rho(\nu, 1, 1, \gamma), \quad 0 \leq \gamma \leq \gamma_*, \]
Here the functions $\Omega$ possesses the form (3.43), and the transformed Hamiltonian $H$ of the form (3.44) satisfies the estimates for $\|r\|_{\Omega}$, provided that $\max(\Omega)$, $\|\xi,\eta\|_{\gamma}$.

Using (3.15) we finally get that the transformation $\tilde{\Phi}_{\rho} = \tau \circ \Psi \circ \chi_{\rho} \circ \Sigma$ sends the vicinity of the torus $\{0\} \times T^n \times \{0\}$ to the vicinity of $T_{\rho}$:

$$\text{dist}_{\gamma} \left( \tilde{\Phi}_{\rho}(r, \theta, \xi, \eta), T_{\rho} \right) \leq C(m) \sqrt{\nu} \left( \nu + \max(\sqrt{|r|}, \|\xi,\eta\|_{\gamma}) \right),$$

provided that $\max(\sqrt{|r|}, \|\xi,\eta\|_{s}) \leq 1/2$.

4. KAM

4.1. An abstract KAM result. We first recall the abstract KAM theorem from [14], adapting the result and the notation to the present context. Consider the Hamiltonian $H$ of the form (3.45), which depends on a parameter $\rho \in D_0 \subseteq \mathbb{R}^n$, regarding it as a perturbation of the quadratic Hamiltonian

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

Here the functions $\Lambda_a(\rho), a \in \mathcal{L}_{\infty}$, and $\Omega(\rho) \in \mathbb{R}^n$ are defined in (3.37), (3.38), so

$$\Omega(\rho) = \omega + \nu M \rho, \quad \det M \neq 0,$$

and $H$ is a symmetric linear operator in the space $Y^f$ (see (3.39)). Denote $M = \dim Y^f$.

We will assume that $h$ satisfies the following assumptions A1 and A2, depending on constants

$$C', \delta_0, \epsilon' \in (0, 1], \quad \beta_1 \geq 2, \beta_2 > 0, \quad s_* \in \mathbb{N}.$$

**Hypothesis A1** (spectral asymptotic.) For all $\rho \in D_0$ we have

(i) $\Lambda_a \geq \epsilon'$, $|\Lambda_a - |a|^2| \leq C'(a)^{-\beta_2} \forall a \in \mathcal{L}_{\infty}$;

(ii) $|\Lambda_a(\rho) \pm \Lambda_b(\rho)| \geq C' \max(|a|^{-\beta_2}, |b|^{-\beta_2}), \quad a, b \in \mathcal{L}_{\infty}, \quad |a| \neq |b|$;

(iii) $\|(JH(\rho))^{-1}\| \leq \frac{1}{e}, \quad \|\Lambda_a(\rho)I - iJH(\rho))^{-1}\| \leq \frac{1}{e} \forall a \in \mathcal{L}_{\infty}.$

**Hypothesis A2** (transversality). For each $k \in \mathbb{Z}^n \setminus \{0\}$ and every vector-function $\Omega' \rho$ such that $|\Omega' - \Omega|C_{\ast}(\rho) \leq \delta_0$ the following properties hold:

(i) for any $a \in \mathcal{L}_{\infty}$ consider the function $L(\rho) = \Omega'(\rho) \cdot k + \Lambda_a(\rho)$. Then it possesses the transversally property: either

$$|L(\rho)| \geq \delta_0 \quad \forall \rho \in D_0.$$
or there exists a unit vector \( \zeta = \zeta(k) \in \mathbb{R}^n \) such that
\[
|\partial_{\zeta} L(\rho)| \geq \delta_0 \quad \forall \rho \in \mathcal{D}_0.
\]

Here \( \partial_{\zeta} \) denotes the directional derivative in the direction \( \zeta \).

(i') For any \( a, b \in L_\infty \), the two functions \( L_{\pm}(\rho) = \Omega'(\rho) \cdot k + \Lambda_a(\rho) \pm \Lambda_b(\rho) \) possess the same transversality property as \( L(\rho) \) in (i).

(ii) For any \( \lambda \in \mathbb{R} \) consider the linear operator \( L(\rho, \lambda) \) in the space \( Y^f \):
\[
L(\rho, \lambda) : X \mapsto (\Omega'(\rho) \cdot k)X + \lambda X + iXJH(\rho) ,
\]
and denote \( P(\rho, \lambda) = \det L(\rho, \lambda) \). Then either
\[
\| L^{-1}(\rho, \lambda) \| \leq \delta_0^{-1} \quad \forall \rho \in \mathcal{D}_0, \ a \in L_\infty ,
\]
or there exists a unit vector \( \zeta = \zeta(k) \in \mathbb{R}^n \) such that
\[
|\partial_{\zeta} P(\rho, \lambda_a(\rho))| \geq |\partial_{\zeta} P(\rho, \lambda_a(\rho))\partial_{\zeta} \lambda_a(\rho)| + \delta_0 |L(\cdot, \lambda_a(\cdot))|^M_{C_1(\mathcal{D}_0)} ,
\]
for all \( \rho \in \mathcal{D}_0 \) and \( a \in L_\infty \).

(iii) Consider the linear operator in the space \( Y^f \),
\[
L(\rho) : X \mapsto (\Omega'(\rho) \cdot k)X - iJH(\rho)X .
\]
Then it possesses the following transversality property: either \( \| L(\rho) \|^{-1} \| \leq \delta_0^{-1} \) for all \( \rho \), or there exists a unit vector \( \zeta = \zeta(k) \) and an integer \( 1 \leq j \leq s_* \) such that
\[
|\partial_{\zeta}^j \det L(\rho)| \geq \delta_0 |L(\rho)|^{M-1}_{C_j} , \quad \forall \rho \in \mathcal{D}_0 .
\]

(iii') Consider the \( \rho \)-depending linear operator in the space of all linear transformations \( M \) of \( Y^f \):
\[
M \mapsto (\Omega'(\rho) \cdot k)M - iJH(\rho)M + iMJH(\rho) .
\]
Then it possesses the same transversality property as the operator \( L(\rho) \) in (iii).

Recall that the domains \( \mathcal{O}^\gamma(\sigma, \mu) \) and the classes \( \mathcal{T}^{\gamma,D}(\sigma, \mu, \mathcal{D}) \) were defined at the beginning of Section 3. Denote
\[
\chi = |\partial_\rho \Omega(\rho)|_{C^{\gamma_*-1}} + \sup_{s \in L_\infty} |\partial_\rho \Lambda_a(\rho)|_{C^{\gamma_*-1}} + ||\partial_\rho H||_{C^{\gamma_*-1}} .
\]

Consider a perturbation \( f(r, \theta, \zeta; \rho) \) and assume that
\[
\varepsilon = |f^{T}|_{\gamma, \sigma, \mu, \mathcal{D}} < \infty , \quad \xi = |f^{D}|_{\gamma, \sigma, \mu, \mathcal{D}} < \infty ,
\]
for some \( \gamma, \sigma, \mu \in (0, 1] \). We are now in position to state the abstract KAM theorem from [13].

\textbf{Theorem 4.1.} Assume that Hypotheses A1, A2 hold for \( \rho \in \mathcal{D}_0 . \) Then there exists \( \tilde{c} = \tilde{c}(s_*) \) and \( \tilde{c} = \tilde{c}(\beta, \beta_2) \) such that if for a suitable \( \tilde{c} \) there is given in terms of a functional class which we did not defined in his work, but its validity easily follows from Lemma 7.8.

\footnote{The theorem below is a bit weakened version of the result in [14]. In particular, in [14] the restriction on the perturbation \( f \) is given in terms of a functional class which we did not defined in his work, but its validity easily follows from Lemma 7.8.}
then there is a Borel set $D' \subset D_0$ with $\text{meas}(D_0 \setminus D') \leq C \varepsilon^{\beta_3}$, $\beta_3 > 0$, and for all $\rho \in D'$ the following holds:

There exists a real holomorphic symplectomorphism $\mathfrak{F}_\rho : \mathcal{O}^0(\sigma/2, \mu/2) \to \mathcal{O}^0(\sigma, \mu)$, satisfying

$$\|\mathfrak{F} - \text{id}\|_{0, D'} \leq C \delta^c, \quad c > 0.$$  

such that

$$H \circ \mathfrak{F}_\rho = \hat{\Omega}(\rho) \cdot r + \frac{1}{2} \zeta, A(\rho) \zeta + g(r, \theta, \zeta; \rho),$$

where $\partial_\zeta g = \partial_r g = \partial^2_{\zeta \zeta} g = 0$ for $\zeta = r = 0$. Here $\hat{\Omega} = \hat{\Omega}(\rho)$ is a new frequency vector and $A : \mathcal{L} \times \mathcal{L} \to \mathcal{M}_{2 \times 2}(\rho)$ is an infinite real symmetric matrix, belonging to $\mathcal{M}_{0}^{D'}$. It is of the form $A = A_\infty + A_0$, where

$$\|A_\infty(\rho) - H_0(\rho)\| \leq C' c'.$$

The operator $A_\infty$ is such that $A_{\infty; ab} = 0$ if $|a| \neq |b|$, and all eigenvalues of the Hamiltonian operator $J A_\infty$ are pure imaginary.

The constants $C, C'$ and exponents $c$ and $c'$ depend on the set $\mathcal{A}$, constants in (4.2) and $\gamma, \sigma, \mu$.

### 4.2. KAM for the Beam Equation

In this section we prove Theorem 1.3. By $C, C_1$ etc we denote various constants, depending only on $m$ and $A$.

In Proposition 3.11, assuming that $m \notin \mathcal{C}$, we put the beam equation in the normal form (3.33), where $\rho \in D_\delta$. To the Hamiltonian (3.33), where $\rho \in D_0$ and $D_0$ is a suitable subset of $D_\delta$, we are going to apply Theorem 4.1 with $H = \nu H_0$. Let us choose $\gamma, \sigma, \mu$ as in the proposition. Then

$$\varepsilon = [f^T]_{\gamma, D} \leq C_2 \nu^{3/2}.$$  

We chose

$$\delta_0 = \nu^{1+c}, \quad c' = \nu^{1+2c}, \quad \bar{c} > 0$$

Now we will show the Hamiltonian $H_\rho$ as in (3.43) with $\delta = \nu^{c'}$ meets Hypotheses A1, A2 of Theorem 4.1 with parameters, specified in (4.7), provided that $\bar{c}$ is sufficiently small.

Using (3.38) and (4.1) we get

$$|A_\infty - \lambda_\infty|_{C^1(\mathcal{D}_\delta)} \leq C_3 \nu |a|^{-2}, \quad |\Omega - \omega|_{C^1(\mathcal{D}_\delta)} \leq C_3 \nu.$$  

This and (2.4) imply (i) and (ii) in A1. Since $\|H_0\| \leq C$, then by the Kramer rule and the definition of the set $\mathcal{D}_\delta$ with $\delta = \nu^{c'}$ (see (3.44)), we have $\|(JH)^{-1}\| \leq C_4 \nu^{-1 - c}$ for $\rho \in \mathcal{D}_\delta$. So the first relation in (iii) also holds. Since $\lambda_\infty(\rho) \geq 1$ and $|H| \leq C \nu$, then the second relation holds as well.

Now we verify A2. Consider the function $L(\rho)$ as in (i). By (4.1) and (4.8),

$$|\partial_\lambda L| \geq \nu |M^3 \cdot k| - \delta_0 |k| - C |a|^{-2} \nu.$$  

Choosing

$$\delta = \frac{\nu |M^1 \cdot k|}{|M^1|},$$

we achieve that

$$|\partial_\lambda L| \geq \nu |M^3 \cdot k| - \delta_0 |k| - C |a|^{-2} \nu.$$  

This is bigger than $\nu$ if $|k| \geq C_5$. But if $|k| \leq C_5$, then in view of Proposition 2.9 $|L| \geq C_6 - C \gamma \nu$. So (i) holds if $\nu \ll 1$. 
To prove (i′) consider $L_-(\rho) = \Omega'(\rho) \cdot k + \Lambda_\alpha(\rho) - \Lambda_\beta(\rho)$ (the case of the sign + is easier). We may assume that $|a_2| \geq |a_2|$. First let $L_-$ be such that $\omega \cdot k + \lambda_\alpha - \lambda_\beta$ is a trivial resonance. That is, $\omega \cdot k = -\omega_{n_1} + \omega_{n_2}$, where $\omega_{n_1} = \lambda_\alpha$, $\omega_{n_2} = \lambda_\beta$. There are only finitely many divisors $L_-$ like that. Using (4.3) and (4.38) we see that by removing from $D_\delta$ a set $\tilde{D}$ of measure $\leq C\delta_0/\nu = C\nu^2$ we achieve that $|R| \geq \delta_0$ for all divisors of this type.

Now let $L_-$ does not correspond to a trivial resonance. Choosing $\tilde{z}$ as in (1.9), we have

$$|\partial_\tilde{z}L_-| \geq \nu|\tilde{M}k| - \delta_0|k| - C|a_2|^2\nu.$$ 

This is bigger than $C_1^{-1}\nu$, unless

$$|k| \leq C_2 \quad \text{and} \quad |a_2| \leq C_3.$$

But in this case, by Proposition 2.9

$$|L_-| \geq C_5|k|^{-\epsilon} - C_4\nu \geq C_3C_2 - C_4\nu.$$ 

So (i′) is fulfilled if $\nu \ll 1$, for $\rho \in D_\delta \setminus \tilde{D}$.

To verify (ii), we note that

$$\|L(\rho, \lambda_\alpha) - (\omega \cdot k + \lambda_\alpha)I\| \leq C\nu.$$

So in view of Proposition 2.9 $\|L(\rho, \lambda_\alpha)^{-1}\| \leq C_1^{-1}$ if $\nu \ll 1$.

Proofs of (iii) and (iii′) are the same in both cases the operator $L$ differs from $(\omega \cdot k)I$ at most by $C\nu$.

Now the Hypotheses A1, A2 are verified. To apply Theorem 14 it remains to verify (1.9), but these relations with a suitable $R > 0$ immediately follow from (1.6) and (1.7). Accordingly, Theorem 11 applies with $D_0 = D_\delta \setminus \tilde{D}$. This application provides the final (third) normal form for the beam equation, written in the form (1.10) with $\rho \in D_0$, where the first two normal forms are given by Theorem 3.8 and Proposition 5.11. Since $\text{mes}(D \setminus \tilde{D}) \leq \text{mes}(D_\delta \setminus \tilde{D}) \leq C'\delta$, we get the first assertion of Theorem 1.4 (and the amplification) with $U(\theta, I) = \Phi_I \circ \tilde{\Phi}_I$ and $\omega_m(I) = \tilde{\Omega}(I)$, where the estimate (1.13) follows from (3.40) and the bound on $||\tilde{\Phi}_I - \text{id}||_\gamma$.

The fact that the linearised equation has no less unstable directions than the matrix $iJK$ (or, equivalently, the matrix $JH_0$) follows from (1.5) since $c' \ll \nu$. The last assertion follows from the calculation in Appendix B below.

**Appendix A. Proof of Lemma 3.2**

For any $\gamma \geq 0$ let us denote by $Z_\gamma$ 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 $Y_\gamma$. For $v \in Z_\gamma$ we will denote by $F(v) = u(x)$ the Fourier-transform of $v$, $u(x) = \sum v_s e^{i\sigma x}$. By Example 3.1 if $u(x)$ is a bounded analytic function in $T^n_{\sigma'}$, then $F^{-1}u \in Z_{\sigma}$ for $\sigma < \sigma'$. Let $F$ be the Fourier-image of the nonlinearity $g$, i.e.

$$F(v) = F^{-1}g(x, F(v)(x)).$$
Lemma A.1. For sufficiently small $\mu_s > 0$, $\gamma_s > 0$ and for all $0 \leq \gamma \leq \gamma_s$,

i) $F$ defines an analytic mapping $O_{\mu_s}(Z_\gamma) \to Z_\gamma$,

ii) $\nabla F$ defines an analytic mapping $O_{\mu_s}(Z_\gamma) \to M_\gamma$, where $M_\gamma$ is the space of matrices $A : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{C}$, satisfying $|A|_\gamma := \sup |A_b^a| e^{\gamma|a-b|} < \infty$.

Proof. i) For sufficiently small $\sigma', \mu > 0$ the nonlinearity $g$ defines a real holomorphic function $g : \mathbb{T}_\sigma' \times O_{\mu}(\mathbb{C}) \to \mathbb{C}$ 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 analytic in $x \in \mathbb{T}_\sigma$ and by the Cauchy estimate $|g_r| \leq M\mu^{-r}$. So

$$\|F^{-1}g_r\|_{\gamma} \leq C_\sigma M\mu^{-r} \quad \forall 0 \leq \gamma \leq \sigma,$$

for any $\sigma < \sigma'$. Cf. Example 3.1. We may write $F(v)$ as

$$(A.1) \quad F(v) = \sum_{r=3}^{\infty} (F^{-1}g_r) \ast v \ast \cdots \ast v.$$

Since the space $Z_\gamma$ is an algebra with respect to the convolution (see Lemma 1.1 in [13]), the $r$-th term of the sum is bounded as follows:

$$(A.2) \quad \|((F^{-1}g_r) \ast v \ast \cdots \ast v)\|_{\gamma} \leq C_1 C_r+1 \mu^{-r} \|v\|_{\gamma}^r.$$

This implies the assertion with $\gamma_s = \sigma$ and a suitable $\mu_s > 0$.

ii) For $r \geq 3$ consider the $r$-th term in the sum for $g(x, u(x))$ and denote by $G_r$ its Fourier-image, $G_r(v) = F^{-1}(g_ru^r)$, $u = F(v)$. Then

$$\langle \nabla G_r(v) \rangle_a^b = r(2\pi)^{-d} \int e^{-ia \cdot x} g_r(x)u^{r-1}e^{ib \cdot x} \, dx.$$

Applying (A.2) (with $r$ convolutions instead of $r + 1$) we see that

$$(A.3) \quad \|\nabla G_r(v)\|_{\gamma} \leq C_2 C_r \mu^{-r} \|v\|_{\gamma}^{r-1} \langle b-a \rangle^{-d} e^{-\gamma|b-a|}.$$

So $\|\nabla G_r(v)\|_{\gamma} \leq C' \mu^{-r} \|v\|_{\gamma}^{r-1}$, which implies the second assertion of the lemma. \[\square\]

Proof of Lemma 3.4. Let us consider the functional $P(\zeta)$ as in [13], and write it as

$$P(\zeta) = p \circ \Upsilon \circ D^{-1}\zeta.$$

Here $D$ is the operator, defined in Section 3.1 $\Upsilon$ is the bounded operator

$$\Upsilon : Y_\gamma \to Z_\gamma, \quad \zeta \mapsto v, \quad v_s = \frac{\xi_s + \eta_s}{\sqrt{2}} \quad \forall s,$$

and $p(v) = \int G(x, (F^{-1}v)(x)) \, dx$. Lemma A.1 with $g$ replaced by $G$ immediately implies that $P$ is an analytic function on $O_{\mu_s}(Y_\gamma)$ with suitable $\mu_s, \gamma_s > 0$.

Next, since

$$\nabla P(\zeta) = D^{-1} \circ \Upsilon \circ \nabla p(\Upsilon \circ D^{-1}\zeta),$$

where $\nabla P = F$ is the map in Lemma A.1 then $\nabla P$ defines a real holomorphic mapping $O_{\mu_s}(Y_\gamma) \to Y_\gamma$.

Further,

$$\nabla^2 P(\zeta) = D^{-1}(\Upsilon \nabla^2 p(\Upsilon \circ D^{-1}\zeta) \Upsilon) D^{-1}.$$
Since for any $A \in M_n$, the matrix $\gamma^T A \gamma$ is given by the relation

$$\left(\gamma^T A \gamma\right)_a^b = \frac{1}{2} \sum_{a' = \pm a, b' = \pm b} A_{a'b'}^\gamma,$$

then $|D^{-1}(\gamma^T A \gamma)D^{-1}|^D \leq 2|A|_\gamma$. So

$$|\nabla^2 \rho(\zeta)|^D \leq 2|\nabla^2 p(\zeta)|_\gamma = 2|\nabla F(\zeta)|_\gamma,$$

and in view of item ii) of Lemma B.1, the mapping

$$\nabla^2 \rho : O_\mu, (Y_\gamma) \to \mathcal{M}^D_\gamma, \quad 0 \leq \gamma \leq \gamma_*,$$

is real holomorphic and bounded in norm by a $\gamma$-independent constant. \hfill \Box

**APPENDIX B. EXAMPLES**

In this appendix we explore some different configurations for the Hamiltonian operator $L(\rho) = iJK(\rho)$, according to the dimension $d$ and the set $\mathcal{A}$.

**Examples with** $(\mathcal{L}_f \times \mathcal{L}_f)_+ = \emptyset$.
As we noticed in section 3, if $(\mathcal{L}_f \times \mathcal{L}_f)_+ = \emptyset$ then $L$ is Hermitian so there is no hyperbolic feature, i.e. the KAM tori are linearly stable.

For instance the choice $d = 2$ and $\mathcal{A} = \{(k, 0), (0, \ell)\}$ with the additional assumption that no $k^2$ no $\ell^2$ are the sum of two squares, yields $(\mathcal{L}_f \times \mathcal{L}_f)_+ = \emptyset$.

These examples can be plunged in higher dimension, for instance $\mathcal{A} = \{(1, 0, 0), (0, 2, 0)\}$ or $\mathcal{A} = \{(1, 0, 0), (0, 2, 0), (0, 0, 3)\}$

**Examples with** $(\mathcal{L}_f \times \mathcal{L}_f)_+ \neq \emptyset$. In this case hyperbolic directions may appear as we can see below.

The choice $\mathcal{A} = \{(j, k), (0, -k)\}$ leads to $\{(j, -k), (0, k)\} \in (\mathcal{L}_f \times \mathcal{L}_f)_+$. Again this example can be plunged in higher dimension.

**The particular case** $\text{card} \mathcal{A} = 2$

When $\text{card} \mathcal{A} = 2$ we have a complete description of the different possibilities:

**Lemma B.1.** When $\text{card} \mathcal{A} = 2$ a node $a \in \mathcal{L}_f$ cannot belong both to a pair $(a, b)$ in $(\mathcal{L}_f \times \mathcal{L}_f)_+$ and to a pair $(a, c)$ in $(\mathcal{L}_f \times \mathcal{L}_f)_-$ with $b \neq c$.

As a consequence, the Hamiltonian matrix $L$ decomposes in a direct sum of matrices:

$L(\rho) = L^1(\rho) \oplus \cdots \oplus L^M(\rho)$ where each $L^j$ is

(i) either a block of dimension two which is diagonal and gives rise to linearly stable tori (when the block contains only one node).

(ii) either a block of dimension four which is the sum of a diagonal part and a symmetric part and which gives rise to linearly stable tori (when the block contains two nodes $a, b$ with $(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_-$).

(iii) either a block of dimension four which is the sum of a diagonal part and an antisymmetric part and which may give rise to two elliptic directions and two hyperbolic directions (when the block contains two nodes $a, b$ with $(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+$, see an explicit example below).

**Proof.** Assume that $(a, b)$ and $(a, c)$ are in $(\mathcal{L}_f \times \mathcal{L}_f)_+$. Then, since $\text{card} \mathcal{A} = 2$, necessarily $\ell(b) = \ell(c)$ which leads to $a + c = \ell(a) + \ell(b) = a + b$ and thus $b = c$. The case when $(a, b)$ and $(a, c)$ are in $(\mathcal{L}_f \times \mathcal{L}_f)_-$ is similar.

Now assume that $(a, b) \in (\mathcal{L}_f \times \mathcal{L}_f)_+$ and $(a, c) \in (\mathcal{L}_f \times \mathcal{L}_f)_-$. Then, since $\text{card} \mathcal{A} = 2$, necessarily $\ell(b) = \ell(c)$. On the other hand we get $(b, c) \in (\mathcal{L}_f \times \mathcal{L}_f)_+$ but this is impossible by virtue of Lemma 3.11. \hfill \Box
An example with hyperbolic directions

In this appendix we present an explicit example in dimension \( d = 2 \), corresponding to the case (iii) in Lemma 13.1. That is, for the 2d beam equation (1.1) we will find an admissible set \( \mathcal{A} \) such that the corresponding matrix \( iJK(\rho) \) in the normal form (3.30) has an unstable direction. Then by item 2) of Theorem 13.3 the time-quasiperiodic solutions of (1.1), constructed in the theorem, are linearly unstable.

Let
\[
\mathcal{A} = \{(0,1),(1,-1)\}
\]
we easily compute using (3.25), (3.26)
\[
(\mathcal{L}_f \times \mathcal{L}_f)_+ = \{(0,-1),(1,1)\},
\]
\[
(\mathcal{L}_f \times \mathcal{L}_f)_- = \emptyset.
\]

We consider the transformed Hamiltonian \( h + f \) of the beam equation, given by (3.31), (3.35) and (3.36), and wish to prove that for some choice of \( \rho \) and \( m \) the Hamiltonian operator \( iJK(m,\rho) \) has hyperbolic directions.

Let us denote \((\xi_1,\eta_1)\) (reps. \((\xi_2,\eta_2)\)) the \((\xi,\eta)\)-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 \(\lambda_2 = \sqrt{4 + m}\). Let \(h_r\) be the restriction of the Hamiltonian \((K(m,\rho)\xi_f,\xi_f)\) to the modes \((\xi_1,\eta_1)\) and \((\xi_2,\eta_2)\). We notice that these two modes do not interact with other modes in the quadratic part and we calculate using (3.40) that
\[
h_r = \beta \xi_1 \eta_1 + \gamma \xi_2 \eta_2 + \alpha (\eta_1 \eta_2 + \xi_1 \xi_2)
\]
with
\[
\alpha = 6(2\pi)^{-2} \frac{\sqrt{\rho_1 \rho_2}}{\lambda_1 \lambda_2}, \quad \beta = 3(2\pi)^{-2} \left( \frac{\rho_1}{\lambda_1} - \frac{2 \rho_2}{\lambda_2} \right), \quad \gamma = 3(2\pi)^{-2} \left( \frac{\rho_2}{\lambda_2} - \frac{2 \rho_1}{\lambda_1} \right).
\]
Thus the linear Hamiltonian system governing the two modes read\(^9\)
\[
\begin{cases}
\dot{\xi}_1 = -i(\beta \xi_1 + \alpha \eta_2) \\
\dot{\eta}_1 = i(\beta \eta_1 + \alpha \xi_2) \\
\dot{\xi}_2 = -i(\gamma \xi_2 + \alpha \eta_1) \\
\dot{\eta}_2 = i(\gamma \eta_2 + \alpha \xi_1).
\end{cases}
\]

Let us denote this vector-field as \(M(\rho)(\xi_1,\eta_1,\xi_2,\eta_2)\). Then the Hamiltonian operator \(L = iJK(\rho)\) admits the decomposition
\[
L(\rho) = M(\rho) \oplus N(\rho),
\]
where \(N\) corresponds to the diagonal operator \(iJK^d\) (see (3.31)) when the two nodes \((0,-1)\) and \((1,1)\) are removed from the set \(\mathcal{L}_f\). Now let us calculate the spectrum of the matrix
\[
-iM = \begin{pmatrix}
-\beta & 0 & 0 & -\alpha \\
0 & \beta & \alpha & 0 \\
0 & -\alpha & -\gamma & 0 \\
\alpha & 0 & 0 & \gamma
\end{pmatrix}.
\]
Its characteristic polynomial is
\[
\det(-iM - \lambda I) = \left(\lambda^2 + (\gamma - \beta)\lambda - \beta\gamma + \alpha^2\right)
\left(\lambda^2 - (\gamma - \beta)\lambda - \beta\gamma + \alpha^2\right).
\]
\(^9\)Recall that the symplectic two-form is:
\[-i \sum d\xi \wedge d\eta.\]
And the discriminant of the polynomial $\lambda^2 + (\gamma - \beta)\lambda - \beta \gamma + \alpha^2$ equals
\[
\Delta = (\beta + \gamma)^2 - 4\alpha^2.
\]
Now we choose $\rho_1 = \rho_2 = \rho$ and we get
\[
\beta + \gamma = 3(2\pi)^{-2}\rho \left( \frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} - \frac{8}{\lambda_1 \lambda_2} \right), \quad \alpha = 6(2\pi)^{-2}\rho \frac{1}{\lambda_1 \lambda_2}.
\]
Then we compute
\[
\Delta = \frac{9\rho}{(2\pi)^4} \left( \frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} \right) \left( \frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} - \frac{8}{\lambda_1 \lambda_2} \right) \leq \frac{9\rho}{(2\pi)^4} \left( \frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} \right) \left( \frac{1}{\lambda_1^2} - \frac{7}{\lambda_2^2} \right)
\]
and we verify that $\Delta < 0$ for all $m \in [1, 2]$.

Therefore $M$ has eigenvalues with non vanishing real part. This implies that the Hamiltonian operator $\mathcal{H}$ has hyperbolic directions.

**Appendix C. An estimate for polynomial functions.**

We will need the following classical result (see [24], Section 1.7):

**Cartan’s theorem.** Let $P_n(z)$ be a complex polynomial of degree $n$ with the leading coefficient $K$. Then for any $\varepsilon > 0$ the set $\{ z \in \mathbb{C} : |P_n(z)| < \varepsilon \}$ may be covered by a finite collection of complex discs such that the sum of their radii equals $2\varepsilon (\varepsilon / K)^{1/n}$.

**Lemma C.1.** Let $F(x)$ be a non-trivial real polynomial of degree $\bar{d}$, restricted to the cube $K^n = [0, 1]^n$. Then there exists a positive constant $C_F$ such that
\[
\text{meas}\{ x \in K^n : |F(x)| < \varepsilon \} \leq C_F \varepsilon^{1/\bar{d}}, \quad \forall \varepsilon \in (0, 1].
\]

**Proof.** By the compactness argument it suffices to prove this in the vicinity of any point $x_0 \in K^n$, where $F(x_0) = 0$. So we have reduced the problem to the case when
\[
F : B_\rho := \{ |x| < \rho \} \to \mathbb{R}, \quad \rho > 0,
\]
and $F$ is a non-trivial polynomial of degree $\bar{d}$. Rotating the coordinate system we achieve that the function $x_1 \mapsto F(x_1, 0, \ldots, 0)$ does not vanish identically. Denote
\[
x = (x_1, \ldots, x_n) = (x_1, \bar{x}), \quad \bar{x} = (x_2, \ldots, x_n),
\]
and write
\[
F(x) = \sum_{j=1}^m f_j(\bar{x})x_1^j, \quad 1 \leq m \leq \bar{d}.
\]
Let
\[
f_0(0) = \cdots = f_{k-1}(0) = 0, \quad f_k(0) \neq 0,
\]
where $1 \leq k \leq \bar{d}$. By the Cartan theorem for any $\bar{x}$
\[
\text{meas}\{ x_1 \in \mathbb{R} : |F(x_1, \bar{x})| < \varepsilon \} \leq C_F \varepsilon^{1/\bar{d}}.
\]
Jointly with the Fubini theorem this inequality establishes for the function (C.2) estimate (C.1) and implies the assertion of the lemma. \[ \Box \]
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., 70, 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 Schrödinger equation, Ann. Math. 148 (1998), 363-439.
[9] J. Bourgain Green’s function estimates for lattice Schrödinger operators and applications, Annals of Mathematical Studies, Princeton, 2004.
[10] 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. 181 (2010), 31-113.
[11] W. Craig Problèmes de Petits Diviseurs dans les Équations aux Dérivées Partielles, Panoramas et Synthèses, Société Mathématique de France, 2000.
[12] 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.
[13] 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.
[14] L.H Eliasson, B. Grébert and S.B. Kuksin, KAM for the nonlinear beam equation 2: a normal form theorem. Preprint (2013).
[15] L.H. Eliasson and S.B. Kuksin. Infinite Töplitz-Lipschitz matrices and operators. Z. Angew. Math. Phys. 59 (2008), 24-50.
[16] L.H. Eliasson and S.B. Kuksin. KAM for the nonlinear Schrödinger equation. Ann. Math 172 (2010), 371-435.
[17] J. Geng and J. You. A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces. Comm. Math. Phys., 262 (2006), 343–372.
[18] J. Geng and J. You. KAM tori for higher dimensional beam equations with constant potentials, Nonlinearity, 19 (2006), 2405–2423.
[19] 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.
[20] S. B. Kuksin. Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum. Funct. Anal. Appl., 21 (1987), 192–205.
[21] S. B. Kuksin. Nearly integrable infinite-dimensional Hamiltonian systems. Lecture Notes in Mathematics, 1556. Springer-Verlag, Berlin, 1993.
[22] S. B. Kuksin. Analysis of Hamiltonian PDEs. Oxford University Press, 2000.
[23] S. B. Kuksin and J. Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. Ann. Math. 143 (1996), 149–179.
[24] B. Ya. Levin. Distribution of zeroes of entire functions. Translations of Mathematical Monographs, 5. AMS, Providence, 1964.
[25] J. Moser and C. L. Siegel. Lectures on celestial mechanics. Springer-Verlag, Berlin, 1971.
[26] J. Pöschel. Quasi-periodic solutions for a nonlinear wave equation. Comment. Math. Helv. 71 (1996) 269–296.
[27] C. Procesi and M. Procesi. A normal form of the nonlinear Schrödinger equation with analytic non–linearities, Comm. Math. Phys 312 (2012), 501-557.
[28] C. Procesi and M. Procesi. A KAM algorithm for the resonant nonlinear Schrödinger equation, preprint 2013.
[29] W.-M. Wang, Nonlinear Schrödinger equations on the torus. *Monograph, 156 pp*, 2014 (submitted).

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