Commun. Math. Phys. 211, 497 – 525 (2000)

Communications in
Mathematical Physics

KAM Tori for 1D Nonlinear Wave Equations with Periodic Boundary Conditions*

Luigi Chierchia[1], **Jiangong You**[2]

[1] Dipartimento di Matematica, Università Roma Tre, 00146 Rome, Italy.
E-mail: luigi@matrm3.mat.uniroma3.it
[2] Department of Mathematics, Nanjing University, Nanjing 210093, P.R. China. E-mail: jyou@ nju.edu.cn

Received: 2 August 1999 / Accepted: 7 January 2000

Abstract: In this paper, one-dimensional (1D) nonlinear wave equations

$$u_{tt} - u_{xx} + V(x)u = f(u),$$

with periodic boundary conditions are considered; V is a periodic smooth or analytic function and the nonlinearity f is an analytic function vanishing together with its derivative at $u = 0$. It is proved that for "most" potentials $V(x)$, the above equation admits small-amplitude periodic or quasi-periodic solutions corresponding to finite dimensional invariant tori for an associated infinite dimensional dynamical system. The proof is based on an infinite dimensional KAM theorem which allows for multiple normal frequencies.

1. Introduction and Results

In the 90's the celebrated KAM (Kolmogorov–Arnold–Moser) theory has been successfully extended to infinite dimensional settings so as to deal with certain classes of partial differential equations carrying a Hamiltonian structure, including, as a typical example, wave equations of the form

$$u_{tt} - u_{xx} + V(x)u = f(u), \quad f(u) = O(u^2); \tag{1.1}$$

see Wayne [17], Kuksin [10] and Pöschel [15]. In such papers, KAM theory for lower dimensional tori [14,13,8] (i.e., invariant tori of dimension lower than the number of degrees of freedom), has been generalized in order to prove the existence of small-amplitude quasi-periodic solutions for (1.1) subject to Dirichlet or Neumann boundary conditions (on a finite interval for odd and analytic nonlinearities f). The technically more difficult periodic boundary condition case has been later considered by Craig and

* This research was partly supported by NNSF of China, by the Italian CNR grant # 211.01.31 and by MURST (Project "Metodi Variazionali ed Equazioni Differenziali Non Lineari").

Wayne [7] who established the existence of periodic solutions. The techniques used in [7] are based not on KAM theory, but rather on a generalization of the Lyapunov-Schmidt procedure and on techniques by Fröhlich and Spencer [9]. Recently, Craig and Wayne's approach has been significantly improved by Bourgain [3–5] who obtained the existence of quasi-periodic solutions for certain kind of 1D and, most notably, 2D partial differential equations with periodic boundary conditions.

The technical reason why KAM theory has not been used to treat the periodic boundary condition case is related to the multiplicity of the spectrum of the associated Sturm-Liouville operator $A = -\frac{d^2}{dx^2} + V(x)$. Such multiplicity leads to some extra "small denominator" problems (related to the so called normal frequencies), which make the KAM analysis particularly delicate.

The purpose of this paper is to show that, improving the KAM machinery, one can indeed use KAM techniques to deal also with the multiple normal frequency case arising in PDE's with periodic boundary conditions (e.g., 1D wave equations).

The advantage of the KAM approach is, from one side, to possibly simplify proofs and, on the other side, to allow the construction of local normal forms close to the considered torus, which could be useful for a better understanding of the dynamics. For example, in general, one can easily check linear stability and the vanishing of Liapounov exponents.

A rough description of our results is as follows. Consider the periodic boundary problem for (1.1) with an analytic nonlinearity f and a real analytic (or smooth enough) potential V. Such a potential will be taken in a d-dimensional family of functions parameterized by a real d-vector ξ, $V(x) = V(x, \xi)$, satisfying general non-degenerate ("non–resonance–of–eigenvalue") conditions. Then for "most" potentials in the family (i.e. for most ξ in Lebesgue measure sense), there exist small-amplitude quasi-periodic solutions for (1.1) corresponding to d-dimensional KAM tori for the associated infinite dimensional Hamiltonian system. Moreover (as usual in the KAM approach) one obtains, for the constructed solutions, a local normal form which provides linear stability in case the operator A is positive definite.

Finally we hope that the technique used in this paper can be generalized so as to deal with more general situations such as, for example, 2D wave equations.

The paper is organized as follows: In Sect. 2 we formulate a general infinite dimensional KAM Theorem designed to deal with multiple normal frequency cases; in Sect. 3 we show how to apply the preceding KAM Theorem to the nonlinear wave Eq. (1.1) with periodic boundary conditions. The proof of the KAM Theorem is provided in Sects. 4–6. Some technical lemmata are proved in the Appendix.

2. An Infinite Dimensional KAM Theorem

In this section we will formulate a KAM Theorem in an infinite dimensional setting which can be applied to some 1D partial differential equations with periodic boundary conditions.

We start by introducing some notations.

2.1. Spaces. For $n \in \mathbb{N}$, let $d_n \in \mathbb{Z}_+$ be positive *even* integers[1]. Let $\mathcal{Z} \equiv \prod_{n\in\mathbb{N}} \mathbb{C}^{d_n}$: the coordinates in $\mathcal{Z}$ are given by $z = (z_0, z_1, z_2, \cdots)$ with $z_n \equiv (z_n^1, \cdots, z_n^{d_n}) \in \mathbb{C}^{d_n}$.

[1] We use the notations $\mathbb{N} = \{0, 1, 2, \cdots\}$, $\mathbb{Z}_+ = \{1, 2, \cdots\}$.

Given two real numbers a, ρ, we consider the (Banach) subspace of $\mathcal{Z}$ given by

$$\mathcal{Z}_{a,\rho} = \{z \in \mathcal{Z} : \ |z|_{a,\rho} < \infty\},$$

where the norm $|\cdot|_{a,\rho}$ is defined as

$$|z|_{a,\rho} = |z_0| + \sum_{n\in\mathbb{Z}_+} |z_n| n^a e^{n\rho},$$

(and the norm in $\mathbb{C}^{d_n}$ is taken to be the 1–norm $|z_n| = \sum_{j=1}^{d_n} |z_n^j|$).

In what follows, we shall consider either $a = 0$ and $\rho > 0$ or $a > 0$ and $\rho = 0$ (corresponding respectively to the analytic case or the finitely smooth case).

The role of complex neighborhoods in phase space of KAM theory will be played here by the set

$$\mathcal{P}_{a,\rho} \equiv \hat{\mathbb{T}}^d \times \mathbb{C}^d \times \mathcal{Z}_{a,\rho},$$

where $\hat{\mathbb{T}}^d$ is the complexification of the real torus $\mathbb{T}^d = \mathbb{R}^d/2\pi\mathbb{Z}^d$.

For positive numbers r, s we denote by

$$D_{a,\rho}(r,s) = \{(\theta, I, z) \in \mathcal{P}_{a,\rho} : |\mathrm{Im}\,\theta| < r, |I| < s^2, |z|_{a,\rho} < s\}, \tag{2.1}$$

a complex neighborhood of $\mathbb{T}^d \times \{I = 0\} \times \{z = 0\}$. Finally, we denote by $\mathcal{O}$ a given compact set in $\mathbb{R}^d$ with positive Lebesgue measure: $\xi \in \mathcal{O}$ will parameterize a selected family of potential $V = V(x, \xi)$ in (1.1).

2.2. Functions. We consider functions F on $D_{a,\rho}(r,s) \times \mathcal{O}$ having the following properties: (i) F is real for real arguments; (ii) F admits an expansion of the form

$$F = \sum_\alpha F_\alpha z^\alpha, \tag{2.2}$$

where the multi-index α runs over the set $\alpha \equiv (\alpha_0, \alpha_1, ...) \in \prod_{n\in\mathbb{N}} \mathbb{N}^{d_n}$ with finitely many non-vanishing components[2] α_n; (iii) for each α, the function $F_\alpha = F_\alpha(\theta, I, \xi)$ is real analytic in the variables $(\theta, I) \in \{|\mathrm{Im}\theta| < r, |I| < s^2\}$; (iv) for each α, the dependence of F_α upon the parameter ξ is of class $C_W^{\bar{d}^2}(\mathcal{O})$ for some $\bar{d} > 0$ (to be fixed later): here $C_W^m(\mathcal{O})$ denotes the class of functions which are m times differentiable on the closed set $\mathcal{O}$ in the sense of Whitney [18] (and the appearance of the square is due to later notational convenience).

The convergence of the expansion (2.2) in $D_{a,\rho}(r,s) \times \mathcal{O}$ will be guaranteed by assuming the finiteness of the following weighted norm:

$$\|F\|_{D_{a,\rho}(r,s),\mathcal{O}} \equiv \sup_{|z|_{a,\rho}\le s} \sum_\alpha \|F_\alpha\| \ |z^\alpha|, \tag{2.3}$$

[2] Thus $\exists\, n_0 > 0$ such that $z^\alpha \equiv \prod_{n=0}^{n_0} z_n^{\alpha_n} \equiv \prod_{n=0}^{n_0} \prod_{j=1}^{d_n} (z_n^j)^{\alpha_n^j}$.

where, if $F_\alpha = \sum_{k\in\mathbb{Z}^d, l\in\mathbb{N}^d} F_{kl\alpha}(\xi) I^l e^{\mathrm{i}\langle k,\theta\rangle}$, ($\langle\cdot,\cdot\rangle$ being the standard inner product in $\mathbb{C}^n$), $\|F_\alpha\|$ is short, here, for

$$\|F_\alpha\| \equiv \sum_{k,l} |F_{kl\alpha}|_{\mathcal{O}}\, s^{2|l|} e^{|k|r}, \quad |F_{kl\alpha}|_{\mathcal{O}} \equiv \max_{|p|\le \bar{d}^2} |\frac{\partial^p F_{kl\alpha}}{\partial \xi^p}|, \tag{2.4}$$

(the derivatives with respect to ξ are in the sense of Whitney).

The set of functions $F : D_{a,\rho}(r,s) \times \mathcal{O} \to \mathbb{C}$ verifying (i) - (iv) above with finite $\|\cdot\|_{D_{a,\rho}(r,s),\mathcal{O}}$ norm will be denoted by $\mathcal{F}_{D_{a,\rho}(r,s),\mathcal{O}}$.

2.3. Hamiltonian vector fields and Hamiltonian equations. To functions $F \in \mathcal{F}_{D_{a,\rho}(r,s),\mathcal{O}}$, we associate a Hamiltonian vector field defined as

$$X_F = (F_I, -F_\theta, \{\mathrm{i}J_{d_n} F_{z_n}\}_{n\in\mathbb{N}}),$$

where J_{d_n} denotes the standard symplectic matrix $\begin{pmatrix} 0 & I_{d_n/2} \\ -I_{d_n/2} & 0 \end{pmatrix}$ and $\mathrm{i} = \sqrt{-1}$; the derivatives of F are defined as the derivatives term–by–term of the series (2.2) defining F. The appearance of the imaginary unit is due to notational convenience and will be justified later by the use of complex canonical variables.

Correspondingly we consider the Hamiltonian equations[3]

$$\dot\theta = F_I, \quad \dot I = -F_\theta, \quad \dot z_n = \mathrm{i}J_{d_n} F_{z_n}, \quad n\in\mathbb{N}. \tag{2.5}$$

A solution of such equation is intended to be just a C^1 map from an interval to the domain of definition of F, $D_{a,\rho}(r,s)$, satisfying (2.5).

Given a real number $\bar a$, we define also a weighted norm for X_F by letting[4]

$$\|X_F\|^{\bar a,\rho}_{D_{a,\rho}(r,s),\mathcal{O}} \equiv \tag{2.6}$$
$$\|F_I\|_{D_{a,\rho}(r,s),\mathcal{O}} + \frac{1}{s^2}\|F_\theta\|_{D_{a,\rho}(r,s),\mathcal{O}} + \frac{1}{s}(\|F_{z_0}\|_{D_{a,\rho}(r,s),\mathcal{O}} + \sum_{n\in\mathbb{Z}_+} \|F_{z_n}\|_{D_{a,\rho}(r,s),\mathcal{O}} n^{\bar a} e^{n\rho}).$$

Notational Remark. In what follows, only the indices r, s and the set $\mathcal{O}$ will change while $a, \bar a, \rho$ will be kept fixed, therefore we shall usually denote $\|X_F\|^{\bar a,\rho}_{D_{a,\rho}(r,s),\mathcal{O}}$ by $\|X_F\|_{r,s,\mathcal{O}}$, $D_{a,\rho}(r,s)$ by $D(r,s)$ and $\mathcal{F}_{D_{a,\rho}(r,s),\mathcal{O}}$ by $\mathcal{F}_{r,s,\mathcal{O}}$.

---

[3] Dot stands for the time derivatives d/dt.

[4] The norm $\|\cdot\|_{D_{a,\rho}(r,s),\mathcal{O}}$ for scalar functions is defined in (2.3). For vector (or matrix–valued) functions $G : D_{a,\rho}(r,s)\times\mathcal{O} \to \mathbb{C}^m$, ($m < \infty$) is similarly defined as $\|G\|_{D_{a,\rho}(r,s),\mathcal{O}} = \sum_{i=1}^m \|G_i\|_{D_{a,\rho}(r,s),\mathcal{O}}$ (for the matrix–valued case the sum will run over all entries).

2.4. Perturbed Hamiltonians and the KAM result. The starting point will be a family of integrable Hamiltonians of the form

$$N = \langle \omega(\xi), I\rangle + \frac{1}{2}\sum_{n\in\mathbb{N}} \langle A_n(\xi) z_n, z_n\rangle, \tag{2.7}$$

where $\xi \in \mathcal{O}$ is a parameter, A_n is a $d_n \times d_n$ real symmetric matrix and $\langle\cdot,\cdot\rangle$ is the standard inner product; here the phase space $\mathcal{P}_{a,\rho}$ is endowed with the symplectic form $dI \wedge d\theta + \mathrm{i}\sum_n \sum_{j=1}^{d_n/2} z_n^j \wedge dz_n^{j+d_n/2}$.

For simplicity, we shall take, later, $\omega(\xi) \equiv \xi$.

For each $\xi \in \mathcal{O}$, the Hamiltonian equations of motion for N, i.e.,

$$\frac{d\theta}{dt} = \omega, \quad \frac{dI}{dt} = 0, \quad \frac{dz_n}{dt} = \mathrm{i}J_{d_n} A_n z_n, \quad n \in \mathbb{N}, \tag{2.8}$$

admit special solutions $(\theta, 0, 0) \to (\theta + \omega t, 0, 0)$ corresponding to an invariant torus in $\mathcal{P}_{a,\rho}$.

Consider now the perturbed Hamiltonians

$$H = N + P = \langle \omega(\xi), I\rangle + \frac{1}{2}\sum_{n\in\mathbb{N}} \langle A_n(\xi) z_n, z_n\rangle + P(\theta, I, z, \xi) \tag{2.9}$$

with $P \in \mathcal{F}_{r,s,\mathcal{O}}$.

Our goal is to prove that, for most values of parameter $\xi \in \mathcal{O}$ (in Lebesgue measure sense), the Hamiltonian $H = N + P$ still admits an invariant torus provided $\|X_P\|$ is sufficiently small.

In order to obtain this kind of result *we shall need the following assumptions on* A_n *and the perturbation* P:

(A1) *Asymptotics of eigenvalues.* There exist $\bar{d} \in \mathbb{N}$, $\delta > 0$ and $b \geq 1$ such that $d_n \leq \bar{d}$ for all n, and

$$A_n = \lambda_n \begin{pmatrix} 0 & I_{d_n/2}, \\ I_{d_n/2} & 0 \end{pmatrix} + B_n, \quad B_n = O(n^{-\delta}), \tag{2.10}$$

where λ_n are real and independent of ξ while B_n may depend on ξ; furthermore, the behaviour of λ_n's is assumed to be as follows

$$\lambda_n = n^b + o(n^b), \quad \frac{\lambda_m - \lambda_n}{m^b - n^b} = 1 + o(n^{-\delta}), \quad n < m. \tag{2.11}$$

(A2) *Gap condition.* There exists $\delta_1 > 0$ such that

$$\mathrm{dist}\left(\sigma(J_{d_i} A_i), \sigma(J_{d_j} A_j)\right) > \delta_1 > 0, \quad \forall i \neq j\ ;$$

($\sigma(\cdot)$ denotes "spectrum of $\cdot$").

Note that for large i, j, the *gap condition* follows from the asymptotic property.

(A3) *Smooth dependence on parameters.* All entries of B_n are $\bar{d}^2$ Whitney–smooth functions of ξ with $C_W^{\bar{d}^2}$-norm bounded by some positive constant L.

(A4) *Non-resonance condition.*

$$\text{meas}\{\xi \in \mathcal{O} : \quad \langle k, \omega(\xi)\rangle(\langle k, \omega(\xi)\rangle + \lambda)(\langle k, \omega(\xi)\rangle + \lambda + \mu) = 0\} = 0, \tag{2.12}$$

for each $0 \neq k \in \mathbb{Z}^d$ and for any $\lambda, \mu \in \bigcup_{n \in \mathbb{N}} \sigma(J_{d_n} A_n)$; meas $\equiv$ Lebesgue measure.

(A5) *Regularity of the perturbation*. The perturbation $P \in \mathcal{F}_{D_{a,\rho}(r,s),\mathcal{O}}$ is *regular* in the sense that $\|X_P\|^{\bar{a},\rho}_{D_{a,\rho}(r,s),\mathcal{O}} < \infty$ with $\bar{a} > a$. In fact, we assume that one of the following holds:

$$(a) \quad \rho > 0, \quad \bar{a} > a = 0; \quad (b) \quad \rho = 0, \quad \bar{a} > a > 0,$$

(such conditions correspond, respectively, to analytic or smooth solutions).

Now we can state our KAM result.

Theorem 1. *Assume that N in (2.7) satisfies (A1)–(A4) and P is regular in the sense of (A5) and let $\gamma > 0$. There exists a positive constant $\epsilon = \epsilon(d, \bar{d}, b, \delta, \delta_1, \bar{a} - a, L, \gamma)$ such that if $\|X_P\|^{\bar{a},\rho}_{D_{a,\rho}(r,s),\mathcal{O}} < \epsilon$, then the following holds true. There exists a Cantor set $\mathcal{O}_\gamma \subset \mathcal{O}$ with* meas$(\mathcal{O} \setminus \mathcal{O}_\gamma) \to 0$ *as $\gamma \to 0$, and two maps (real analytic in θ and Whitney smooth in $\xi \in \mathcal{O}$)*

$$\Psi : T^d \times \mathcal{O}_\gamma \to D_{a,\rho}(r, s) \subset \mathcal{P}_{a,\rho}, \quad \tilde{\omega} : \mathcal{O}_\gamma \to R^d,$$

such that for any $\xi \in \mathcal{O}_\gamma$ and $\theta \in T^d$ the curve $t \to \Psi(\theta + \tilde{\omega}(\xi)t, \xi)$ is a quasi-periodic solution of the Hamiltonian equations governed by $H = N + P$. Furthermore, $\Psi(\mathbb{T}^d, \xi)$ is a smoothly embedded d-dimensional H-invariant torus in $\mathcal{P}_{a,\rho}$.

Remarks. (i) For simplicity we shall in fact assume that *all eigenvalues λ_i of A_n are positive for all n's.* The case of some non-positive eigenvalues can be easily dealt with at the expense of a (even) heavier notation.

(ii) In the above case (i.e. positive eigenvalues), Theorem 1 yields *linearly stable* KAM tori.

(iii) The parameter γ plays the role of the Diophantine constant for the frequency $\tilde{\omega}$ in the sense that there is $\tau > 0$ such that $\forall k \in \mathbb{Z}^d \setminus \{0\}$,

$$\langle k, \tilde{\omega}\rangle > \frac{\gamma}{2|k|^\tau}.$$

Notice also that $\mathcal{O}_\gamma$ is claimed to be nonempty and big only for γ small enough.

(iv) The regularity property $\bar{a} > a$ is used only in estimating the measure of $\mathcal{O} \setminus \mathcal{O}_\gamma$. Such regularity requirement is not necessary for for constructing periodic solutions, i.e., $d = 1$. Thus *the above theorem applies to the construction of periodic solutions for 1-D nonlinear Schrödinger equations*.

(v) The non-degeneracy condition (2.12) (which is stronger than Bourgain's non-degenerate condition [4] but weaker than Melnikov's one [13]) covers the multiple normal frequency case: this is the technical reason that allows to treat PDE's with periodic boundary conditions.

3. Application to 1D Wave Equations

In this section we show how Theorem 1 implies the existence of quasi-periodic solutions for 1D wave equations with periodic boundary conditions.

Let us rewrite the wave equation (1.1) as follows:

$$u_{tt} + Au = f(u), \quad Au \equiv -u_{xx} + V(x,\xi)u, \; x,t \in \mathbb{R},$$
$$u(t,x) = u(t,x+2\pi), \quad u_t(t,x) = u_t(t,x+2\pi), \tag{3.1}$$

where $V(\cdot,\xi)$ is a *real–analytic* (*or smooth*) periodic potential parameterized by some $\xi \in \mathbb{R}^d$ (see below) and $f(u)$ is a *real–analytic* function near $u = 0$ with $f(0) = f'(0) = 0$.

As it is well known, the operator A with periodic boundary conditions admits an orthonormal basis of eigenfunctions $\phi_n \in L^2(\mathbb{T})$, $n \in \mathbb{N}$, with corresponding eigenvalues μ_n satisfying the following asymptotics for large n

$$\mu_{2n-1}, \mu_{2n} = n^2 + \frac{1}{2\pi}\int_{\mathbb{T}} V(x)dx + O(n^{-2}).$$

For simplicity, we shall consider the case of vanishing mean value of the potential V and assume that all eigenvalues are positive:

$$\int_{\mathbb{T}} V(x)dx = 0\,, \quad \mu_n \equiv \lambda_n^2 > 0\,, \quad \forall\, n. \tag{3.2}$$

Following Kuksin [10] and Bourgain [3], we consider a family of real analytic (or smooth) potentials $V(x,\xi)$, where *the d-parameters $\xi = (\xi_1, \cdots, \xi_d) \in \mathcal{O} \subset \mathbb{R}^d$ are simply taken to be a given set of d frequencies $\lambda_{n_i} \equiv \sqrt{\mu_{n_i}}$*:

$$\xi_i \equiv \sqrt{\mu_{n_i}} \equiv \lambda_{n_i}\,, \quad i = 1, \cdots, d \tag{3.3}$$

where μ_{n_i} are (positive) eigenvalues of[5] A.

We may also (and shall) require that there exists a positive $\delta_1 > 0$ such that

$$|\mu_k - \mu_h| > \delta_1\,, \tag{3.4}$$

for all $k > h$ except when k is even and $h = k-1$ (in which case μ_k and μ_h might even coincide).

Notice that, in particular, having d eigenvalues as *independent* parameters excludes the constant potential case $V \equiv$ constant (where, of course, all eigenvalues are double: $\mu_{2j-1} = \mu_{2j} = j^2 + V$). In fact, this case seems difficult to be handled by KAM approach even in the finite dimensional case. Such difficulty does not arise, instead, in the remarkable alternative approach developed by Craig, Wayne [7] and Bourgain [3,4].

Equation (3.1) may be rewritten as

$$\dot{u} = v, \quad \dot{v} + Au = f(u), \tag{3.5}$$

which, as is well known, may be viewed as the (infinite dimensional) Hamiltonian equations $\dot{u} = H_v$, $\dot{v} = -H_u$ associated to the Hamiltonian

$$H = \frac{1}{2}(v,v) + \frac{1}{2}(Au,u) + \int_{\mathbb{T}} g(u)\, dx, \tag{3.6}$$

[5] Plenty of such potentials may be constructed with, e.g., the inverse spectral theory.

where g is a primitive of $(-f)$ (with respect to the u variable) and $(\cdot,\cdot)$ denotes the scalar product in L^2.

As in [15], we introduce coordinates $q = (q_0, q_1, \cdots)$, $p = (p_0, p_1, \cdots)$ through the relations

$$u(x) = \sum_{n\in\mathbb{N}} \frac{q_n}{\sqrt{\lambda_n}}\phi_n(x), \quad v = \sum_{n\in\mathbb{N}} \sqrt{\lambda_n}\, p_n\phi_n(x),$$

where[6] $\lambda_n \equiv \sqrt{\mu_n}$. System (3.5) is then formally equivalent to the lattice Hamiltonian equations

$$\dot q_n = \lambda_n p_n, \quad \dot p_n = -\lambda_n q_n - \frac{\partial G}{\partial q_n}, \quad G \equiv \int_{\mathbb{T}} g(\sum_{n\in\mathbb{N}} \frac{q_n}{\sqrt{\lambda_n}}\phi_n)dx\ , \tag{3.7}$$

corresponding to the Hamiltonian function $H = \sum_{n\in\mathbb{N}} \lambda_n(q_n^2 + p_n^2) + G(q)$. Rather than discussing the above formal equivalence, we shall, following [15], use the following elementary observation (proved in the Appendix):

Proposition 3.1. *Let V be analytic (respectively, smooth), let I be an interval and let*

$$t \in I \to (q(t), p(t)) \equiv \Big(\{q_n(t)\}_{n\geq 0}, \{p_n(t)\}_{n\geq 0}\Big)$$

be an analytic (respectively, smooth[7]) solution of (3.7) such that

$$\sup_{t\in I} \sum_{n\in\mathbb{N}} \Big(|q_n(t)| + |p_n(t)|\Big)\, n^a\, e^{n\rho} < \infty \tag{3.8}$$

for some $\rho > 0$ and $a = 0$ (respectively, for $\rho = 0$ and a big enough). Then

$$u(t,x) \equiv \sum_{n\in\mathbb{N}} \frac{q_n(t)}{\sqrt{\lambda_n}}\phi_n(x),$$

is an analytic (respectively, smooth) solution of (3.1).

Before invoking Theorem 1 we still need some manipulations. We first switch to complex variables: $w_n = \frac{1}{\sqrt 2}(q_n + \mathrm{i}p_n)$, $\bar w_n = \frac{1}{\sqrt 2}(q_n - \mathrm{i}p_n)$. Equations (3.7) read then

$$\dot w_n = -\mathrm{i}\lambda_n w_n - \mathrm{i}\frac{\partial \tilde G}{\partial \bar w_n}, \quad \dot{\bar w}_n = \mathrm{i}\lambda_n \bar w_n + \mathrm{i}\frac{\partial \tilde G}{\partial w_n}, \tag{3.9}$$

where the perturbation $\tilde G$ is given by

$$\tilde G(w) = \int_{\mathbb{T}} g(\sum_{n\in\mathbb{N}} \frac{w_n + \bar w_n}{\sqrt{2\lambda_n}}\phi_n)dx. \tag{3.10}$$

Next we introduce standard action-angle variables $(\theta, I) = ((\theta_1, \cdots, \theta_d), (I_1, \cdots, I_d))$ in the $(w_{n_1}, \cdots, w_{n_d}, \bar w_{n_1}, \cdots, \bar w_{n_d})$-space by letting,

$$I_i = w_{n_i}\bar w_{n_i}, \quad i = 1, \cdots, d,$$

[6] Recall that, for simplicity, we assume that all eigenvalues μ_n are positive.

[7] Regularity refers to the components q_n and p_n.

so that the system (3.9) becomes

$$\frac{d\theta_j}{dt} = \omega_j + P_{I_j}, \quad \frac{dI_j}{dt} = -P_{\theta_j}, \quad j = 1, \cdots, d,$$
$$\frac{dw_n}{dt} = -\mathrm{i}\lambda_n w_n - \mathrm{i}P_{\bar{w}_n}, \quad \frac{d\bar{w}_n}{dt} = \mathrm{i}\lambda_n \bar{w}_n + \mathrm{i}P_{w_n}, \ n \neq n_1, n_2, \cdots, n_d, \tag{3.11}$$

where P is just $\tilde{G}$ with the $(w_{n_1}, \cdots, w_{n_d}, \bar{w}_{n_1}, \cdots, \bar{w}_{n_d})$-variables expressed in terms of the (θ, I) variables and the *frequencies* $\omega = (\omega_1, ..., \omega_d)$ coincide with the parameter ξ introduced in (3.3):

$$\omega_i \equiv \xi_i = \lambda_{n_i} \,. \tag{3.12}$$

The Hamiltonian associated to (3.11) (with respect to the symplectic form $dI \wedge d\theta + \mathrm{i}\sum_n dw_n \wedge d\bar{w}_n$) is given by

$$H = \langle \omega, I\rangle + \sum_{n \neq n_1, \cdots, n_d} \lambda_n w_n \bar{w}_n + P(\theta, I, w, \bar{w}, \xi). \tag{3.13}$$

Remark. Actually, in place of H in (3.13) one should consider the *linearization* of H around a given point I_0 and let I vary in a small ball B (of radius $0 < s \ll |I_0|$) in the "positive" quadrant $\{I_j > 0\}$. In such a way the dependence of H upon I is obviously analytic. For notational convenience we shall however do not report explicitly the dependence of H on I_0.

Finally, to put the Hamiltonian in the form (2.9) we couple the variables $(w_n, \bar{w}_n)$ corresponding to "closer" eigenvalues. More precisely, we let $z_n = (w_{2n-1}, w_{2n}, \bar{w}_{2n-1}, \bar{w}_{2n})$ for large[8] n, say $n > \bar{n} > n_d$ and denote by $z_0 = \left(\{w_n\}_{\substack{0 \le n \le \bar{n} \\ n \neq n_1, \dots, n_d}}, \{\bar{w}_n\}_{\substack{0 \le n \le \bar{n} \\ n \neq n_1, \dots, n_d}}\right)$ the remaining conjugated variables. The Hamiltonian (3.13) takes the form

$$H = \langle \omega, I\rangle + \frac{1}{2}\sum_{n \in \mathbb{N}} \langle A_n z_n, z_n\rangle + P(\theta, I, z, \xi), \tag{3.14}$$

where

$$A_n = \mathrm{Diag}(\lambda_{2n-1}, \lambda_{2n}, \lambda_{2n-1}, \lambda_{2n}) \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}$$
$$= \lambda_{2n} \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \lambda_{2n-1} - \lambda_{2n} & 0 \\ 0 & 0 & 0 & 0 \\ \lambda_{2n-1} - \lambda_{2n} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

for $n > n_d$, while $A_0 = \mathrm{Diag}(\{\lambda_n\}, \{\lambda_n\}; 1 \le n \le n_d, n \neq n_1, \cdots, n_d) \begin{pmatrix} 0 & I_{d_0} \\ I_{d_0} & 0 \end{pmatrix}$ with $d_0 = \bar{n} + 1 - d$.

The perturbation P in (3.14) has the following (nice) regularity property.

[8] Compare (A1).

Lemma 3.1. *Suppose that V is real analytic in x (respectively, belongs to the Sobolev space $H^k(\mathbb{T})$ for some $k \in \mathbb{N}$). Then for small enough $\rho > 0$ (respectively, $a > 0$), $r > 0$ and $s > 0$ one has*

$$\|X_P\|^{a+1/2,\rho}_{D_{a,\rho}(r,s),\mathcal{O}} = O(|z|^2_{a,\rho}) ; \tag{3.15}$$

here the parameter a is taken to be 0 (respectively, the parameter ρ is taken to be 0).

A proof of this lemma is given in the Appendix. In fact, X_P is even more "regular" (a fact, however, not needed in what follows): (3.15) holds with 1 in place of 1/2.

The Hamiltonian (3.14) is seen to satisfy all the assumptions of Theorem 1 with: $d_n = 4, n \geq 1$; $d_0 = \bar{n} + 1 - d$; $\bar{d} = \max\{d_0, 4\}$; $b = 1$; $\delta = 2$; δ_1 chosen as in (3.4); $\bar{a} - a = \frac{1}{2}$. Thus Theorem 1 yields the following

Theorem 2. *Consider a family of 1D nonlinear wave equation (3.1) parameterized by $\xi \equiv \omega \in \mathcal{O}$ as above with $V(\cdot, \xi)$ real-analytic (respectively, smooth). Then for any $0 < \gamma \ll 1$, there is a subset $\mathcal{O}_\gamma$ of $\mathcal{O}$ with $\text{meas}(\mathcal{O} \backslash \mathcal{O}_\gamma) \to 0$ as $\gamma \to 0$, such that $(3.1)_{\xi \in \mathcal{O}_\gamma}$ has a family of small-amplitude (proportional to some power of γ), analytic (respectively, smooth) quasi-periodic solutions of the form*

$$u(t,x) = \sum_n u_n(\omega'_1 t, \cdots, \omega'_d t)\phi_n(x),$$

where $u_n : \mathbb{T}^d \to \mathbb{R}$ and $\omega'_1, \cdots, \omega'_d$ are close to $\omega_1, \cdots, \omega_d$.

Remark. As mentioned above, our KAM theorem (which applies only to the case that not all the eigenvalues are multiple[9] and under the hypothesis that all μ_n's are positive) implies that the quasi-periodic solutions obtained are *linearly stable*. In the case that all the eigenvalues are double (as in the constant potential case), one should not expect linear stability (see the example given by Craig, Kuksin and Wayne [6]). We also notice that, essentially with only notational changes, the proof of the above theorem goes through in the case that some of the eigenvalues are negative.

4. KAM Step

Theorem 1 will be proved by a KAM iteration which involves an infinite sequence of change of variables.

At each step of the KAM scheme, we consider a Hamiltonian vector field with

$$H_\nu = N_\nu + P_\nu,$$

where N_ν is an "integrable normal form" and P_ν is defined in some set of the form[10] $D(s_\nu, r_\nu) \times \mathcal{O}_\nu$.

We then construct a map[11]

$$\Phi_\nu : D(s_{\nu+1}, r_{\nu+1}) \times \mathcal{O}_{\nu+1} \subset D(r_\nu, s_\nu) \times \mathcal{O}_\nu \to D(r_\nu, s_\nu) \times \mathcal{O}_\nu$$

9 Recall that we require that the torus frequencies are independent parameters.

10 Recall the notations from Section 2.

11 Recall that the parameters a, ρ and $\bar{a}$ are fixed throughout the proof and are therefore omitted in the notations.

so that the vector field $X_{H_\nu \circ \Phi_\nu}$ defined on $D(r_{\nu+1}, s_{\nu+1})$ satisfies

$$\|X_{H_\nu \circ \Phi_\nu} - X_{N_{\nu+1}}\|_{r_{\nu+1}, s_{\nu+1}, \mathcal{O}_{\nu+1}} \le \epsilon_\nu^\kappa$$

with some new normal form $N_{\nu+1}$ and for some fixed ν-independent constant $\kappa > 1$.

To simplify notations, in what follows, the quantities without subscripts refer to quantities at the ν^{th} step, while the quantities with subscripts $+$ denotes the corresponding quantities at the $(\nu+1)^{\text{th}}$ step. Let us then consider the Hamiltonian

$$H = N + P \equiv e + \langle \omega, I \rangle + \frac{1}{2} \sum_{n \in \mathbb{N}} \langle A_n z_n, z_n \rangle + P, \tag{4.1}$$

defined in $D(r, s) \times \mathcal{O}$; the A_n's are symmetric matrices. We assume that $\xi \in \mathcal{O}$ satisfies[12] (for a suitable $\tau > 0$ to be specified later)

$$|\langle k, \omega \rangle^{-1}| < \frac{|k|^\tau}{\gamma}, \quad \|(\langle k, \omega \rangle I_{d_i} + A_i J_{d_i})^{-1}\| < (\frac{|k|^\tau}{\gamma})^{\bar{d}},$$

$$\|(\langle k, \omega \rangle I_{d_i d_j} + (A_i J_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j} A_j))^{-1}\| < (\frac{|k|^\tau}{\gamma})^{\bar{d}^2}, \tag{4.2}$$

We also assume that

$$\max_{|p| \le \bar{d}^2} \|\frac{\partial^p A_n}{\partial \xi^p}\| \le L, \tag{4.3}$$

on $\mathcal{O}$, and

$$\|X_P\|_{r,s,\mathcal{O}} \le \epsilon. \tag{4.4}$$

We now let $0 < r_+ < r$, and define

$$s_+ = \frac{1}{2} s \epsilon^{\frac{1}{3}}, \quad \epsilon_+ = \gamma^{-\mathrm{c}} \Gamma(r - r_+) \epsilon^{\frac{4}{3}}, \tag{4.5}$$

where

$$\Gamma(t) \equiv \sup_{u \ge 1} u^{\mathrm{c}} e^{-\frac{1}{4} u t} \sim t^{-\mathrm{c}}$$

for $t > 0$. Here and later, the letter c denotes suitable (possibly different) constants that do not depend on the iteration step[13].

We now describe how to construct a set $\mathcal{O}_+ \subset \mathcal{O}$ and a change of variables Φ : $D_+ \times \mathcal{O}_+ = D(r_+, s_+) \times \mathcal{O}_+ \to D(r, s) \times \mathcal{O}$, such that the transformed Hamiltonian $H_+ = N_+ + P_+ \equiv H \circ \Phi$ satisfies all the above iterative assumptions with new parameters $s_+, \epsilon_+, r_+, \gamma_+, L_+$ and with $\xi \in \mathcal{O}_+$.

---

[12] The tensor product (or direct product) of two $m \times n$, $k \times l$ matrices $A = (a_{ij})$, B is a $(mk) \times (nl)$ matrix defined by

$$A \otimes B = (a_{ij} B) = \begin{pmatrix} a_{11} B & \cdots & a_{1n} B \\ \cdots & \cdots & \cdots \\ a_{n1} B & \cdots & a_{mn} B \end{pmatrix}.$$

$\|\cdot\|$ for matrix denotes the operator norm, i.e., $\|M\| = \sup_{|y|=1} |My|$. Recall that ω and the A_i's depend on ξ.

[13] Actually, here $c = \bar{d}^4 \tau + \bar{d}^2 \tau + \bar{d}^2 + 1$.

4.1. Solving the linearized equation. Expand P into the Fourier–Taylor series

$$P = \sum_{k,l,\alpha} P_{kl\alpha} e^{\mathrm{i}\langle k,\theta\rangle}\ I^l z^\alpha,$$

where $k \in \mathbb{Z}^d$, $l \in \mathbb{N}^d$ and $\alpha \in \otimes_{n\in\mathbb{N}}\mathbb{N}^{d_n}$ with finite many non-vanishing components.

Let R be the truncation of P given by

$$\begin{aligned} R(\theta, I, z) \equiv P_0 + P_1 + P_2 \equiv & \sum_{k,|l|\le 1} P_{kl0} e^{\mathrm{i}\langle k,\theta\rangle}\ I^l \\ & + \sum_{k,|\alpha|=1} P_{k0\alpha} e^{\mathrm{i}\langle k,\theta\rangle}\ z^\alpha + \sum_{k,|\alpha|=2} P_{k0\alpha} e^{\mathrm{i}\langle k,\theta\rangle}\ z^\alpha, \end{aligned} \tag{4.6}$$

with

$$2|l| + |\alpha| = 2\sum_{j=1,\cdots,d} l_j + \sum_{j\in\mathbb{N}} |\alpha_j| \le 2.$$

It is convenient to rewrite R as follows:

$$\begin{aligned} R(\theta, I, z) = & \sum_{k,|l|\le 1} P_{kl0} e^{\mathrm{i}\langle k,\theta\rangle}\ I^l \\ & + \sum_{k,i} \langle R_i^k, z_i\rangle e^{\mathrm{i}\langle k,\theta\rangle}\ + \sum_{k,i,j} \langle R_{ji}^k z_i, z_j\rangle e^{\mathrm{i}\langle k,\theta\rangle}\ , \end{aligned} \tag{4.7}$$

where R_i^k, R_{ji}^k are respectively the d_i vector and $(d_j \times d_i)$ matrix defined by

$$R_i^k = \int \frac{\partial P}{\partial z_i} e^{-\mathrm{i}\langle k,\theta\rangle} d\theta|_{z=0,I=0}, \quad R_{ji}^k = \frac{1+\delta_i^j}{2} \int \frac{\partial^2 P}{\partial z_j \partial z_i} e^{-\mathrm{i}\langle k,\theta\rangle} d\theta|_{z=0,I=0}. \tag{4.8}$$

Note that $R_{ij}^k = (R_{ji}^k)^T$.

Rewrite H as $H = N + R + (P - R)$. By the choice of s_+ in (4.5) and by the definition of the norms, it follows immediately that

$$\|X_R\|_{r,s,\mathcal{O}} \le \|X_P\|_{r,s,\mathcal{O}} \le \epsilon. \tag{4.9}$$

Moreover s_+, ϵ_+ are such that, in a smaller domain $D(r, s_+)$, we have

$$\|X_{P-R}\|_{r,s_+}\ <\ \mathrm{c}\,\epsilon_+. \tag{4.10}$$

Then we look for a special F, defined in domain $D_+ = D(r_+, s_+)$, such that the time one map ϕ_F^1 of the Hamiltonian vector field X_F defines a map from $D_+ \to D$ and transforms H into H_+.

More precisely, by second order Taylor formula, we have

$$\begin{aligned} H \circ \phi_F^1 &= (N + R) \circ \phi_F^1 + (P - R) \circ \phi_F^1 \\ &= N + \{N, F\} + R \\ &\quad + \frac{1}{2} \int_0^1 ds \int_0^s \{\{N + R, F\}, F\} \circ \phi_F^t dt + \{R, F\} + (P - R) \circ \phi_F^1. \\ &= N_+ + P_+ \\ &\quad + \{N, F\} + R - P_{000} - \langle \omega', I\rangle - \sum_{n\in\mathbb{N}} \langle R_{nn}^0 z_n, z_n\rangle, \end{aligned} \tag{4.11}$$

where

$$\omega' = \int \frac{\partial P}{\partial I} d\theta|_{I=0,z=0}, \quad R^0_{nn} = \int \frac{\partial^2 P}{\partial z_n^2} d\theta|_{I=0,z=0},$$
$$N_+ = N + P_{000} + \langle \omega', I\rangle + \sum_{n\in\mathbb{N}} \langle R^0_{nn} z_n, z_n\rangle,$$
$$P_+ = \frac{1}{2}\int_0^1 ds \int_0^s \{\{N+R, F\}, F\} \circ X_F^t dt + \{R, F\} + (P-R)\circ \phi_F^1.$$

We shall find a function F of the form

$$F(\theta, I, z) = F_0 + F_1 + F_2 = \sum_{|l|\le 1, |k|\neq 0} F_{kl0} e^{\mathrm{i}\langle k,\theta\rangle}\ I^l + \sum_{i\in\mathbb{N}} \langle F_i^k, z_i\rangle e^{\mathrm{i}\langle k,\theta\rangle}$$
$$+ \sum_{|k|+|i-j|\neq 0} \langle F_{ji}^k z_i, z_j\rangle e^{\mathrm{i}\langle k,\theta\rangle}\ , \tag{4.12}$$

satisfying the equation

$$\{N, F\} + R - P_{000} - \langle \omega', I\rangle - \sum_{n\in\mathbb{N}} \langle R^0_{nn} z_n, z_n\rangle = 0. \tag{4.13}$$

Lemma 4.1. *Equation (4.13) is equivalent to*

$$F_{kl0} = (\mathrm{i}\langle k, \omega\rangle)^{-1} P_{kl0}, \quad k\neq 0, |l|\le 1,$$
$$(\langle k,\omega\rangle I_{d_i} + A_{d_i} J_{d_i}) F_i^k = \mathrm{i} R_i^k, \tag{4.14}$$
$$(\langle k,\omega\rangle I_{d_i} + A_{d_i} J_{d_i}) F_{ij}^k - F_{ij}^k (J_{d_j} A_j) = \mathrm{i} R_{ij}^k, \ \ |k| + |i-j| \neq 0.$$

Proof. Inserting F, defined in (4.12), into (4.13) one sees that (4.13) is equivalent to the following equations[14]:

$$\{N, F_0\} + P_0 - \langle \omega', I\rangle = 0,$$
$$\{N, F_1\} + P_1 = 0,$$
$$\{N, F_2\} + P_2 - \sum_{n\in Z} \langle R^0_{nn} z_n, z_n\rangle = 0. \tag{4.15}$$

The first equation in (4.15) is obviously equivalent, by comparing the coefficients, to the first equation in (4.14). To solve $\{N, F_1\} + P_1 = 0$, we note that[15]

$$\begin{aligned}\{N, F_1\} &= \langle \partial_I N, \partial_\theta F_1\rangle + \langle \nabla_z N, J\nabla_z F_1\rangle \\ &= \langle \partial_I N, \partial_\theta F_1\rangle + \sum_i \langle \nabla_{z_i} N, \mathrm{i} J_{d_i} \nabla_{z_i} F_1\rangle \\ &= \mathrm{i}\sum_{k,i} (\langle\ \langle k,\omega\rangle F_i^k, z_i\rangle + \langle A_i z_i, J_{d_i} F_i^k\rangle) e^{\mathrm{i}\langle k,\theta\rangle} \\ &= \mathrm{i}\sum_{k,i} \langle (\langle k,\omega\rangle I_{d_i} + A_i J_{d_i}) F_i^k, z_i\rangle e^{\mathrm{i}\langle k,\theta\rangle}\ .\end{aligned} \tag{4.16}$$

[14] Recall the definition of P_i in (4.6).
[15] Recall the definition of N in (4.1).

It follows that F_i^k are determined by the linear algebraic system

$$\mathrm{i}(\langle k,\omega\rangle I_{d_i} + A_i J_{d_i})F_i^k + R_i^k = 0, \quad i \in \mathbb{N}, k \in \mathbb{Z}^d.$$

Similarly, from

$$\begin{aligned}
\{N, F_2\} &= \langle \partial_I N, \partial_\theta F_2\rangle + \sum_i \langle \nabla_{z_i} N, \mathrm{i} J_{d_i} \nabla_{z_i} F_2\rangle \\
&= \mathrm{i} \sum_{|k|+|i-j|\neq 0} (\langle\, \langle k,\omega\rangle F_{ji}^k z_i, z_j\rangle + \langle A_i z_i, J_{d_i}(F_{ji}^k)^T z_j\rangle + \langle A_j z_j, J_{d_j} F_{ji}^k z_i\rangle) e^{\mathrm{i}\langle k,\theta\rangle} \\
&= \mathrm{i} \sum_{|k|+|i-j|\neq 0} (\langle\, \langle k,\omega\rangle F_{ji}^k z_i, z_j\rangle + \langle (A_j J_{d_j} F_{ji}^k - F_{ji}^k J_{d_i} A_i) z_i, z_j\rangle) e^{\mathrm{i}\langle k,\theta\rangle} \\
&= \mathrm{i} \sum_{|k|+|i-j|\neq 0} \langle (\langle k,\omega\rangle F_{ji}^k + A_j J_{d_j} F_{ji}^k - F_{ji}^k J_{d_i} A_i) z_i, z_j\rangle e^{\mathrm{i}\langle k,\theta\rangle}
\end{aligned} \tag{4.17}$$

it follows that, F_{ji}^k is determined by the following matrix equation:

$$(\langle k,\omega\rangle I_{d_j} + A_j J_{d_j}) F_{ji}^k - F_{ji}^k (J_{d_i} A_i) = \mathrm{i} R_{ji}^k, \quad |k| + |i-j| \neq 0, \tag{4.18}$$

where F_{ji}^k, R_{ji}^k are $d_j \times d_i$ matrices defined in (4.12) and (4.7). Exchanging i, j we get the third equation in (4.14). □

The first two equations in (4.14) are immediately solved in view of (4.2). In order to solve the third equation in (4.14), we need the following elementary algebraic result from matrix theory.

Lemma 4.2. *Let A, B, C be respectively $n \times n$, $m \times m$, $n \times m$ matrices, and let X be an $n \times m$ unknown matrix. The matrix equation*

$$AX - XB = C, \tag{4.19}$$

is solvable if and only if $I_m \otimes A - B \otimes I_n$ is nonsingular. Moreover,

$$\|X\| \leq \|(I_m \otimes A - B \otimes I_n)^{-1}\| \cdot \|C\|.$$

In fact, the matrix equation (4.19) is equivalent to the (bigger) vector equation given by $(I \otimes A - B \otimes I)X' = C'$, where X', C' are vectors whose elements are just the list (row by row) of the entries of X and C. For a detailed proof we refer the reader to the Appendix in [20] or [12], p. 256.

Remark. Taking the transpose of the third equation in (4.14), one sees that $(F_{ij}^k)^T$ satisfies the same equation of F_{ji}^k. Then (by the uniqueness of the solution) it follows that $F_{ji}^k = (F_{ij}^k)^T$.

4.2. Estimates on the coordinate transformation. We proceed to estimate X_F and Φ_F^1. We start with the following

Lemma 4.3. *Let* $D_i = D(\frac{i}{4}s, r_+ + \frac{i}{4}(r - r_+))$, $0 < i \leq 4$. *Then*

$$\|X_F\|_{D_3,\mathcal{O}} < c\,\gamma^{-c}\Gamma(r - r_+)\epsilon. \tag{4.20}$$

Proof. By (4.2), Lemma 4.1 and Lemmata 7.4, 7.5 in the Appendix, we have

$$\begin{aligned}
|F_{kl0}|_{\mathcal{O}} &\leq |\langle k,\omega\rangle|^{-1}|P_{kl}| < c\,\gamma^{-c}|k|^c e^{-|k|(r-r_+)}\epsilon s^{2-2|l|}, \quad k \neq 0,\\
\|F_i^k\|_{\mathcal{O}} &= \|(\langle k,\omega\rangle I_{d_i} + A_i J_{d_i})^{-1} R_i^k\| \leq \|(\langle k,\omega\rangle I_{d_i} + A_i J_{d_i})^{-1}\| \cdot \|R_i^k\|\\
&< c\,\gamma^{-c}|k|^c|R_i^k|,\\
\|F_{ij}^k\|_{\mathcal{O}} &\leq \|(\langle k,\omega\rangle I_{d_i d_j} + (A_i J_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j} A_j))^{-1}\| \cdot \|R_{ij}^k\|\\
&< c\,\gamma^{-c}|k|^c\|R_{ij}^k\|, \ \ |k| + |i - j| \neq 0,
\end{aligned} \tag{4.21}$$

where $\|\cdot\|_{\mathcal{O}}$ for matrix is similar to (2.4).

It follows that

$$\begin{aligned}
\frac{1}{s^2}\|F_\theta\|_{D_2,\mathcal{O}} &\leq \frac{1}{s^2}\Big(\sum |f_{kl0}| \cdot |I^l| \cdot |k| \cdot |e^{\mathrm{i}\langle k,\theta\rangle}| + \sum |F_i^k| \cdot |z_i| \cdot |k| \cdot |e^{\mathrm{i}\langle k,\theta\rangle}|\\
&\quad + \sum \|F_{ij}^k\| \cdot |z_i| \cdot |z_j| \cdot |k| \cdot |e^{\mathrm{i}\langle k,\theta\rangle}|\Big)\\
&< c\,\gamma^{-c}\Gamma(r - r_+)\|X_R\|\\
&< c\,\gamma^{-c}\Gamma(r - r_+)\epsilon,
\end{aligned} \tag{4.22}$$

where $\Gamma(r - r_+) = \sup_k |k|^c e^{-|k|\frac{1}{4}(r-r_+)}$.

Similarly,

$$\|F_I\|_{D_2,\mathcal{O}} = \sum_{|l|\leq 1} |F_{kl0}| \cdot |e^{\mathrm{i}\langle k,\theta\rangle}| < c\,\gamma^{-c}\Gamma(r - r_+)\epsilon.$$

Now we estimate $\|X_{F^1}\|_{D_2,\mathcal{O}}$. Note that

$$\begin{aligned}
\|F_{z_i}^1\|_{D_2,\mathcal{O}} &= \|\sum_k F_i^k e^{-i<k,\theta>}\|_{D_2,\mathcal{O}}\\
&< c\,\gamma^{-c}\Gamma \sum_{k,i} |R_i^k| e^{|k|r} < c\,\gamma^{-c}\Gamma \|\frac{\partial P_1}{\partial z_i}\|.
\end{aligned} \tag{4.23}$$

It follows that

$$\begin{aligned}
\|X_{F^1}\|_{D_2,\mathcal{O}} &< c \sum_{i\in\mathbb{N}} \|F_{z_i}^1\|_{D_2,\mathcal{O}} i^a e^{i\rho}\\
&< c\,\gamma^{-c}\Gamma \sum_{i\in\mathbb{N}} \|\frac{\partial P_1}{\partial z_i}\| i^a e^{i\rho} < c\,\gamma^{-c}\Gamma\epsilon,
\end{aligned}$$

by the definition of the weighted norm.

Note that[16]

$$\|F^2_{z_i}\|_{D_2,\mathcal{O}} = \|\sum_{k,j}(F^k_{ij} + (F^k_{ij})^T)z_j e^{\mathrm{i}\langle k,\theta\rangle}\|_{D_2,\mathcal{O}}$$
$$< \mathrm{c}\,\gamma^{-\mathrm{c}}\Gamma\|\frac{\partial P_2}{\partial z_i}\|. \tag{4.24}$$

Similarly, we have

$$\|X_{F^2}\|_{D_2,\mathcal{O}} \; < \; \mathrm{c}\,\gamma^{-\mathrm{c}}\Gamma\epsilon. \tag{4.25}$$

The conclusion of the lemma follows from the above estimates. □

In the next lemma, we give some estimates for ϕ^t_F. The following formula (4.26) will be used to prove that our coordinate transformations is well defined. Inequality (4.27) will be used to check the convergence of the iteration.

Lemma 4.4. *Let* $\eta = \epsilon^{\frac{1}{3}}$, $D_{\frac{i}{2}\eta} = D(r_+ + \frac{i-1}{2}(r - r_+), \frac{i}{2}\eta s)$, $i = 1, 2$. *We then have*

$$\phi^t_F : D_{\frac{1}{2}\eta} \to D_\eta, \quad 0 \le t \le 1, \tag{4.26}$$

if $\epsilon \ll (\frac{1}{2}\gamma^{-\mathrm{c}}\Gamma^{-1})^{\frac{3}{2}}$. *Moreover,*

$$\|D\phi^1_F - Id\|_{D_{\frac{1}{2}\eta}} \; < \; \mathrm{c}\,\gamma^{-\mathrm{c}}\Gamma\epsilon. \tag{4.27}$$

Proof. Let

$$\|D^m F\|_{D,\mathcal{O}} = \max\{|\frac{\partial^{|i|+|l|+p}}{\partial\theta^i\partial I^l\partial z^\alpha}F|_{D,\mathcal{O}}, |i| + |l| + |\alpha| = m \ge 2\}.$$

Note that F is polynomial in I of order 1, in z of order 2. From[17] (4.25) and the Cauchy inequality, it follows that

$$\|D^m F\|_{D_1,\mathcal{O}} \; < \; \mathrm{c}\,\gamma^{-\mathrm{c}}\Gamma\epsilon, \tag{4.28}$$

for any $m \ge 2$.

To get the estimates for ϕ^t_F, we start from the integral equation,

$$\phi^t_F = id + \int_0^t X_F \circ \phi^s_F \, ds$$

so that $\phi^t_F : D_{\frac{1}{2}\eta} \to D_\eta$, $\quad 0 \le t \le 1$, as it follows directly from (4.28). Since

$$D\phi^1_F = Id + \int_0^1 (DX_F)D\phi^s_F \, ds = Id + \int_0^1 J(D^2F)D\phi^s_F \, ds,$$

it follows that

$$\|D\phi^1_F - Id\| \le 2\|D^2F\| \; < \; \mathrm{c}\,\gamma^{-\mathrm{c}}\Gamma\epsilon. \tag{4.29}$$

The estimates of second order derivative $D^2\phi^1_F$ follows from (4.28). □

---

[16] Recall (2.3), the definition of the norm.

[17] Recall the definition of the weighted norm in (2.6).

4.3. Estimates for the new normal form. The map ϕ_F^1 defined above transforms H into $H_+ = N_+ + P_+$(see (4.11) and (4.13)) with

$$N_+ = e_+ + \langle \omega_+, y\rangle + \frac{1}{2}\sum_{i\in\mathbb{Z}_+} \langle A_i^+ z_i, z_i\rangle, \tag{4.30}$$

where

$$e_+ = e + P_{000}, \quad \omega_+ = \omega + P_{0l0}(|l| = 1), \quad A_i^+ = A_i + 2R_{ii}^0. \tag{4.31}$$

Now we prove that N_+ shares the same properties with N. By the regularity of X_P and by Cauchy estimates, we have

$$|\omega_+ - \omega| < \epsilon, \quad \|R_{ii}^0\| < \epsilon i^{-\delta} \tag{4.32}$$

with $\delta = \bar{a} - a > 0$. It follows that

$$\|(A_i^+)^{-1}\| \le \frac{\|A_i^{-1}\|}{1 - 2\|A_i^{-1}R_{ii}^0\|} \le 2\|A_i^{-1}\|,$$

$$\|(\langle k, \omega + P_{0l00}\rangle I_{d_i} - J_{d_i}A_i^+)^{-1}\| \le \frac{\|(\langle k,\omega\rangle I_{d_i} + A_i J_{d_i})^{-1}\|}{1 - \|(\langle k,\omega\rangle I_{d_i} + A_i J_{d_i})^{-1}\|\epsilon} \le (\frac{|k|^\tau}{\gamma_+})^{\bar{d}}, \tag{4.33}$$

provided $|k|^{\bar{d}\tau}\epsilon \ < \ \mathrm{c}\ (\gamma^{\bar{d}} - \gamma_+^{\bar{d}})$.

Similarly, we have

$$\|(\langle k, \omega + P_{0l00}\rangle I_{d_i d_j} + (A_i^+ J_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j} A_j^+))^{-1}\| \le (\frac{|k|^\tau}{\gamma_+})^{\bar{d}^2}, \tag{4.34}$$

provided $|k|^{\bar{d}^2\tau}\epsilon \ < \ \mathrm{c}\ (\gamma^{\bar{d}^2} - \gamma_+^{\bar{d}^2})$. This means that in the next KAM step, small denominator conditions are automatically satisfied for $|k| < K$ where $K^{\bar{d}^2\tau}\epsilon \ < \ \mathrm{c}\ (\gamma^{\bar{d}^2} - \gamma_+^{4\bar{d}^2})$.

The following bounds wil be used later for the measure estimates:

$$|\frac{\partial^l(\omega_+ - \omega)}{\partial\xi^l}|_{\mathcal{O}} \le \epsilon, \quad |\frac{\partial^l(A_i^+ - A_i)}{\partial\xi^l}|_{\mathcal{O}} \ < \ \mathrm{c}\ \epsilon i^{-\delta}, \tag{4.35}$$

for $|l| \le \bar{d}^2$ (by definition of the norms).

4.4. Estimates for the new perturbation. To complete the KAM step we have to estimate the new error term.

By the definition of ϕ_F^1 and Lemma 4.4,

$$H \circ \phi_F^1 = N_+ + P_+$$

is well defined in $D_{\frac{1}{2}\eta}$. Moreover, we have the following estimates:

$$\begin{aligned}
\|X_{P_+}\|_{D_{\frac{1}{2}\eta}} &= \|X_{\int_0^1 dt \int_0^s \{\{N+R,F\},F\}\circ\phi_F^s + \{R,F\} + (P-R)\circ\phi_F^1}\|_{D_{\frac{1}{2}\eta}} \\
&\le \|X_{(\int_0^1 dt \int_0^t \{\{N+R,F\},F\}\circ\phi_F^s}\|_{D_{\frac{1}{2}\eta}} + \|X_{(P-R)\circ\phi_F^1}\|_{D_{\frac{1}{2}\eta}} \\
&\le \|X_{\{\{N+R,F\},F\}}\|_{D_\eta} + \|X_{P-R}\|_{D_\eta} \\
&< \mathrm{c}\,\gamma^{-\mathrm{c}}\Gamma^2\epsilon^{\frac{4}{3}} \ < \ \mathrm{c}\ \epsilon_+
\end{aligned} \tag{4.36}$$

by (4.9) and Lemma 7.3.

Thus, there exists a big constant c, independent of iteration steps, such that

$$\|X_{P_+}\|_{r_+,s_+} = \|X_{P_+}\|^{\bar a,\rho}_{D_{\frac{1}{2}\eta}} \le c\gamma^{-\mathrm{c}}\Gamma^2\eta\epsilon = c\epsilon_+. \tag{4.37}$$

The KAM step is now completed.

5. Iteration Lemma and Convergence

For any given s, ϵ, r, γ, we define, for all $\nu \ge 1$, the following sequences

$$\begin{aligned}
r_\nu &= r(1 - \sum_{i=2}^{\nu+1} 2^{-i}),\\
\epsilon_\nu &= c\gamma_\nu^{-\mathrm{c}}\Gamma(r_{\nu-1} - r_\nu)^2\epsilon_{\nu-1}^{\frac{4}{3}},\\
\gamma_\nu &= \gamma(1 - \sum_{i=2}^{\nu+1} 2^{-i}),\\
\eta_\nu &= \frac{1}{2}\epsilon_\nu^{\frac{1}{3}}, \quad L_\nu = L_{\nu-1} + \epsilon_{\nu-1},\\
s_\nu &= \frac{1}{2}\eta_{\nu-1}s_{\nu-1} = 2^{-\nu}(\prod_{i=0}^{\nu-1}\epsilon_i)^{\frac{1}{3}}s_0,\\
K_\nu &= \frac{c}{2}\left(\epsilon_{\nu-1}^{-1}(\gamma_{\nu-1}^{\bar d^2} - \gamma_\nu^{\bar d^2})\right)^{\frac{1}{\bar d^2\tau}},\\
D_\nu &= D_{a,\rho}(r_\nu, s_\nu),
\end{aligned} \tag{5.1}$$

where c is the constant in (4.37). The parameters $r_0, \epsilon_0, \gamma_0, L_0, s_0, K_0$ are defined respectively to be $r, \epsilon, \gamma, L, s, 1$.

Note that

$$\Psi(r) = \prod_{i=1}^{\infty}[\Gamma(r_{i-1} - r_i)]^{2(\frac{3}{4})^i},$$

is a well defined finite function of r.

5.1. Iteration Lemma. The preceding analysis may be summarized as follows.

Lemma 5.1. *Suppose that $\epsilon_0 = \epsilon(d, \bar d, \delta, \delta_1, \bar a - a, L, \tau, \gamma)$ is small enough. Then the following holds for all $\nu \ge 0$. Let*

$$N_\nu = e_\nu + \langle\omega_\nu(\xi), I\rangle + \sum_{i\in\mathbb{N}}\langle A_i^\nu(\xi)z_i, z_i\rangle,$$

be a normal form with parameters ξ satisfying

$$\begin{aligned}
&|\langle k, \omega_\nu\rangle^{-1}| < \frac{|k|^\tau}{\gamma_\nu}, \quad \|(\mathrm{i}\langle k, \omega_\nu\rangle I_{d_i} + A_i^\nu J_{d_i})^{-1}\| < (\frac{|k|^\tau}{\gamma_\nu})^{\bar d},\\
&\|(\mathrm{i}\langle k, \omega_\nu\rangle I_{d_id_j} + (A_i^\nu J_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j}A_j^\nu))^{-1}\| < (\frac{|k|^\tau}{\gamma_\nu})^{\bar d^2}
\end{aligned} \tag{5.2}$$

on a closed set $\mathcal{O}_\nu$ *of* R^n *for all* $k \neq 0, i, j \in \mathbb{Z}$. *Moreover, suppose that* $\omega_\nu(\xi)$, $A_i^\nu(\xi)$ *are* $C^{\bar{d}^2}$ *smooth and satisfy*

$$|\frac{\partial^{\bar{d}^2}(\omega_\nu - \omega_{\nu-1})}{\partial \xi^{\bar{d}^2}}| \leq \epsilon_{\nu-1}, \ |\frac{\partial^{\bar{d}^2}(A_i^\nu - A_i^{\nu-1})}{\partial \xi^{\bar{d}^2}}| \leq \epsilon_{\nu-1} i^{-\delta},$$

on $\mathcal{O}_\nu$ *(in Whitney's sense).*

Finally, assume that

$$\|X_{P_\nu}\|^{\bar{a},\rho}_{D_\nu,\mathcal{O}_\nu} \leq \epsilon_\nu.$$

Then, there is a subset $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu$,

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu \setminus \cup_{|k| \geq K_{\nu+1}} \mathcal{R}^{\nu+1}_{kij}(\gamma_\nu),$$

where

$$\mathcal{R}^{\nu+1}_{kij}(\gamma_{\nu+1}) = \left\{ \xi \in \mathcal{O}_\nu : | \begin{array}{c} |\langle k, \omega_{\nu+1}\rangle^{-1}| > \frac{|k|^\tau}{\gamma_\nu}, \quad \|(\langle k, \omega_\nu\rangle I_{2m} + (A_i^{\nu+1} J_{d_i})^{-1}\| \geq (\frac{|k|^\tau}{\gamma_\nu})^{\bar{d}}, \text{or} \\ \|(\langle k, \omega_{\nu+1} > I_{d_i d_j} + (A_j^{\nu+1} J_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j} A_j^{\nu+1}))^{-1}\| > (\frac{|k|^\tau}{\gamma_\nu})^{\bar{d}^2} \end{array} \right\},$$

with $\omega_{\nu+1} = \omega_\nu + P^\nu_{0l0}$, *and a symplectic change of variables*

$$\Phi_\nu : D_{\nu+1} \times \mathcal{O}_{\nu+1} \to D_\nu, \tag{5.3}$$

such that $H_{\nu+1} = H_\nu \circ \Phi_\nu$, *defined on* $D_{\nu+1} \times \mathcal{O}_{\nu+1}$, *has the form*

$$H_{\nu+1} = e_{\nu+1} + \langle \omega_{\nu+1}, I\rangle + \sum_{i \in \mathbb{N}} \langle A_i^{\nu+1} z_i, z_i\rangle + P_{\nu+1}, \tag{5.4}$$

satisfying

$$\max_{l \leq \bar{d}^2} |\frac{\partial^l(\omega_{\nu+1}(\xi) - \omega_\nu(\xi))}{\partial \xi^l}| \leq \epsilon_\nu, \quad \max_{|l| \leq \bar{d}^2} |\frac{\partial^l(A_i^{\nu+1}(\xi) - A_i^\nu)}{\partial \xi^l}| \leq \epsilon_\nu i^{-\delta}, \tag{5.5}$$

$$\|X_{P_{\nu+1}}\|^{\bar{a},\rho}_{D_{\nu+1},\mathcal{O}_{\nu+1}} \leq \epsilon_{\nu+1}. \tag{5.6}$$

5.2. Convergence. Suppose that the assumptions of Theorem 1 are satisfied. To apply the iteration lemma with $\nu = 0$, recall that

$$\epsilon_0 = \epsilon, \gamma_0 = \gamma, s_0 = s, L_0 = L, N_0 = N, P_0 = P,$$

$$\mathcal{O}_0 = \left\{ \xi \in \mathcal{O} : | \begin{array}{c} |\langle k, \omega\rangle^{-1}| < \frac{|k|^\tau}{\gamma}, \|(\langle k, \omega\rangle I_{d_i} + A_i J_{d_i})^{-1}\| < (\frac{|k|^\tau}{\gamma})^{\bar{d}}, \text{or} \\ \|(\langle k, \omega\rangle I_{d_i d_j} + (A_i J_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j} A_j))^{-1}\| < (\frac{|k|^\tau}{\gamma})^{\bar{d}^2} \end{array} \right\},$$

(with ϵ and γ small enough). Inductively, we obtain the following sequences:

$$\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu,$$
$$\Psi^\nu = \Phi_1 \circ \cdots \circ \Phi_\nu : D_{\nu+1} \times \mathcal{O}_\nu \to D_0, \nu \geq 0,$$
$$H \circ \Psi^\nu = H_{\nu+1} = N_{\nu+1} + P_{\nu+1}.$$

Let $\mathcal{O}_\gamma = \cap_{\nu=0}^{\infty} \mathcal{O}_\nu$. As in [16], thanks to Lemma 4.4, we may conclude that $N_\nu, \Psi^\nu, D\Psi^\nu, \omega_{\nu+1}$ converge uniformly on $D_\infty \times \mathcal{O}_\gamma = D(\frac{1}{2}r, 0, 0) \times \mathcal{O}_\gamma$ with

$$N_\infty = e_\infty + \langle \omega_\infty, I \rangle + \langle A_\infty z, z \rangle = e_\infty + \langle \omega_\infty, I \rangle + \sum_{i \in \mathbb{N}} \langle A_i^\infty z_i, z_i \rangle.$$

Since

$$\epsilon_{\nu+1} = c\gamma_\nu^{-\mathrm{c}} \Gamma(r_\nu - r_{\nu+1}) \epsilon_\nu \leq (c\gamma^{-\mathrm{c}} \Psi(r) \epsilon)^{(\frac{4}{3})^\nu}.$$

It follows that $\epsilon_{\nu+1} \to 0$ provided ϵ is sufficiently small.

Let ϕ_H^t be the flow of X_H. Since $H \circ \Psi^\nu = H_{\nu+1}$, we have that

$$\phi_H^t \circ \Psi^\nu = \Psi^\nu \circ \phi_{H_{\nu+1}}^t. \tag{5.7}$$

The convergence of $\Psi^\nu, D\Psi^\nu, \omega_{\nu+1}, X_{H_{\nu+1}}$ implies that one can take limit in (5.7) so as to get

$$\phi_H^t \circ \Psi^\infty = \Psi^\infty \circ \phi_{H_\infty}^t, \tag{5.8}$$

on $D(\frac{1}{2}r, 0, 0) \times \mathcal{O}_\gamma$, with

$$\Psi^\infty : D(\frac{1}{2}r, 0, 0) \times \mathcal{O}_\gamma \to \mathcal{P}_{a,\rho} \times \mathbb{R}^d.$$

From (5.8) it follows that

$$\phi_H^t(\Psi^\infty(\mathbb{T}^d \times \{\xi\})) = \Psi^\infty \phi_{N_\infty}^t(\mathbb{T}^d \times \{\xi\}) = \Psi^\infty(\mathbb{T}^d \times \{\xi\}),$$

for $\xi \in \mathcal{O}_\gamma$. This means that $\Psi^\infty(\mathbb{T}^d \times \{\omega\})$ is an embedded torus invariant for the original perturbed Hamiltonian system at $\xi \in \mathcal{O}_\gamma$. We remark here the frequencies $\omega_\infty(\xi)$ associated to $\Psi^\infty(\mathbb{T}^d \times \{\xi\})$ is slightly different from ξ. The normal behaviour of the invariant torus is governed by the matrix $A_i^\infty = \sum_{\nu \in \mathbb{N}} A_i^\nu$. $\square$

6. Measure Estimates

At each KAM step, we have to exclude the following resonant set of ξ's:

$$\mathcal{R}^\nu = \bigcup_{|k| > K_\nu, i, j} (\mathcal{R}_k^\nu \cup \mathcal{R}_{ki}^\nu \cup \mathcal{R}_{kij}^\nu),$$

the sets $\mathcal{R}_k^\nu, \mathcal{R}_{ki}^\nu, \mathcal{R}_{kij}^\nu$ being, respectively,

$$\{\xi \in \mathcal{O}_\nu : |\langle k, \omega_\nu \rangle^{-1}| > \frac{|k|^\tau}{\gamma_\nu}\}, \quad \{\xi \in \mathcal{O}_\nu : \|\mathcal{M}_1^{-1}\| > (\frac{|k|^\tau}{\gamma_\nu})^{\bar{d}}\},$$
$$\text{and } \{\omega \in \mathcal{O}_\nu : \|\mathcal{M}_2^{-1}\| > (\frac{|k|^\tau}{\gamma_\nu})^{\bar{d}^2}\}, \tag{6.1}$$

where

$$\mathcal{M}_1 = \langle k, \omega_\nu \rangle I_{d_i} + A_i^\nu J_{d_i},$$
$$\mathcal{M}_2 = \langle k, \omega_\nu \rangle I_{d_i d_j} + (A_j^\nu J_{d_j}) \otimes I_{d_i} - I_{d_j} \otimes (J_{d_i} A_i^\nu). \tag{6.2}$$

In the set $\{\xi \in \mathcal{O} : \|M(\omega)^{-1}\| > C\}$ are included also the ξ's for which M is not invertible. Recall that $\omega_\nu(\xi) = \xi + \sum_{j=0}^{\nu-1} P_{000}^j(\xi)$ with[18] $|\sum P_{000}^j(\xi)|_{C^{\bar{d}2}} \leq \epsilon$, $A_i^\nu = A_i + 2\sum_\nu R_{ii}^{0,\nu}$ with $\|\sum_\nu R_{ii}^{0,\nu}\| = O(\epsilon i^{-\delta})$.

Lemma 6.1. *There is a constant K_0 such that, for any i, j, and $|k| > K_0$,*

$$\text{meas}(\mathcal{R}_k^\nu \cup \mathcal{R}_{ki}^\nu \cup \mathcal{R}_{kij}^\nu) \ < \ \text{c}\ \frac{\gamma}{|k|^{\tau-1}}.$$

Proof. As it is well known

$$\text{meas}\ (\mathcal{R}_k^\nu) \leq \frac{\gamma_\nu}{|k|^\tau}.$$

The set $\mathcal{R}_{ki}^\nu$ is empty if $i >$ const $|k|$, while, if $i \leq$ const $|k|$, from Lemmata 7.6, 7.7 there follows that

$$\text{meas}(\mathcal{R}_{ki}^\nu) \ < \ \text{c}\ \frac{\gamma_\nu}{|k|^{\tau-1}}.$$

We now give a detailed proof for the most complicated estimate, i.e., the estimate on the measure of the set $\mathcal{R}_{kij}^\nu$. Note that the main part of $\mathcal{M}_2$ is diagonal[19]. In fact $\mathcal{M}_2$ can be rewritten as

$$\mathcal{M}_2 \equiv \mathcal{A}_{ij} + \mathcal{B}_{ij}^\nu,$$

with

$$\mathcal{A}_{ij} = \langle k, \omega_{\nu+1} \rangle I_{d_i d_j} + \lambda_j \,\text{Diag}(I_{d_j/2}, -I_{d_j/2}) \otimes I_{d_i} - \lambda_i I_{d_j} \otimes \text{Diag}\ (-I_{d_i/2}, I_{d_i/2}). \tag{6.3}$$

The matrix $\mathcal{A}_{ij}$ is diagonal with entries $\lambda_{kij} = \langle k, \omega_\nu \rangle \pm \lambda_i \pm \lambda_j$ in the diagonal, where λ_i, λ_j are given in (2.10) and $\pm$ sign depends on the position. $\mathcal{B}_{ij}^\nu$ is a matrix of size $O(i^{-\delta} + j^{-\delta})$ since $A_i^\nu = A_i + B_i + O(i^{-\delta}) = A_i + O(i^{-\delta})$ by (2.11) and (4.32).

In the rest of the proof we drop in the notation the indices i, j since they are fixed. Now either all $\lambda_{kij} \leq |k|$ or there are some diagonal elements $\lambda_{kij} > |k|$. We first consider the latter case. By permuting rows and columns, we can find two non-singular matrices Q_1, Q_2 with elements 1 or 0 such that

$$Q_1(\mathcal{A} + \mathcal{B}^\nu)Q_2 = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} + \begin{pmatrix} \tilde{B}_{11} & \tilde{B}_{12} \\ \tilde{B}_{21} & \tilde{B}_{22} \end{pmatrix}, \tag{6.4}$$

where A_{11}, A_{22} are diagonal matrices and A_{11} contains all diagonal elements λ_{kij} which are bigger than $|k|$. Moreover, defining Q_3, Q_4, D as

$$Q_3 = \begin{pmatrix} I & \tilde{0} \\ -\tilde{B}_{21}(A_{11} + \tilde{B}_{11})^{-1} & I \end{pmatrix}, \quad Q_4 = \begin{pmatrix} I & -(A_{11} + \tilde{B}_{11})^{-1}\tilde{B}_{12} \\ 0 & I \end{pmatrix},$$

[18] Recall (4.32), (5.5).

[19] Recall (2.10), (6.2).

and

$$D = A_{22} + \tilde{B}_{22} - \tilde{B}_{21}(A_{11} + \tilde{B}_{11})^{-1}\tilde{B}_{12} = A_{22} + O(i^{-\delta} + j^{-\delta}), \tag{6.5}$$

we have

$$Q_3 Q_1 \left(\mathcal{A} + \mathcal{B}^{\nu+1}\right) Q_2 Q_4 = \begin{pmatrix} A_{11} + B_{11} & 0 \\ 0 & D \end{pmatrix}. \tag{6.6}$$

For $\xi \in \mathcal{O}$ such that D is invertible, we have

$$(\mathcal{A} + \mathcal{B}^{\nu})^{-1} = Q_2 Q_4 \begin{pmatrix} (A_{11} + B_{11})^{-1} & 0 \\ 0 & D^{-1} \end{pmatrix} Q_3 Q_1. \tag{6.7}$$

Since the norm of $Q_1, Q_2, Q_3, Q_4, (A_{11} + B_{11})^{-1}$ are uniformly bounded, it follows from (6.7) that

$$\left\{\xi \in \mathcal{O}_\nu : \|(\mathcal{A} + \mathcal{B}^{\nu})^{-1}\| > \left(\frac{|k|^\tau}{\gamma_\nu}\right)^{\bar{d}^2}\right\} \subset \left\{\xi \in \mathcal{O}_\nu : \|D^{-1}\| \ > \mathrm{c} \left(\frac{|k|^\tau}{\gamma_\nu}\right)^{\bar{d}^2}\right\}. \tag{6.8}$$

If all $\lambda_{kij} \ < \ \mathrm{c}\ |k|$ we simply take $D = \mathcal{A} + \mathcal{B}^{\nu}$. Since all elements in D are of size $O(|k|)$, by Lemma 7.6 in the Appendix, we have

$$\left\{\xi \in \mathcal{O}_\nu : \|D^{-1}\| \ > \mathrm{c} \left(\frac{|k|^\tau}{\gamma_\nu}\right)^{\bar{d}^2}\right\} \subset \left\{\xi \in \mathcal{O}_\nu : |\det D| \ < \ \mathrm{c} \left(\frac{\gamma_\nu}{|k|^{\tau-1}}\right)^{\bar{d}^2}\right\}. \tag{6.9}$$

Let N denote the dimension of D (which is not bigger than $\bar{d}^2$). Since $D = A_{22} + O(i^{-\delta} + j^{-\delta})$, the N^{th} order derivative of $\det D$ with respective to some ξ_i is bounded away from zero by $\frac{1}{2d}|k|^N$ (provided that $|k|$ is bigger enough). From (6.8), (6.9) and Lemma 7.7, it follows that

$$\begin{aligned} \operatorname{meas} \mathcal{R}^{\nu}_{kij} &= \operatorname{meas}\left\{\xi \in \mathcal{O}_\nu : \left\|\left(\mathcal{A} + \mathcal{B}^{\nu}\right)^{-1}\right\| > \left(\frac{|k|^\tau}{\gamma_\nu}\right)^{\bar{d}^2}\right\} \\ &\le \operatorname{meas}\left\{\xi \in \mathcal{O}_\nu : |\det D| \ < \ \mathrm{c} \left(\frac{\gamma_\nu}{|k|^{\tau-1}}\right)^{\bar{d}^2}\right\} \\ &< \mathrm{c}\left(\frac{\gamma_\nu}{|k|^{\tau-1}}\right)^{\frac{\bar{d}^2}{N}} \ < \ \mathrm{c}\ \frac{\gamma}{|k|^{\tau-1}}. \end{aligned} \tag{6.10}$$

This proves the lemma. □

Lemma 6.2. *If $i \ > \mathrm{c}\ |k|$, then $\mathcal{R}^{\nu}_{ki} = \emptyset$; If $\max\{i, j\} \ > \mathrm{c}\ |k|^{\frac{1}{b-1}}, i \neq j$ for $b > 1$ or $|i - j| > \mathrm{const}\ |k|$ for $b = 1$, then $\mathcal{R}^{\nu}_{kij} = \emptyset$, where the constant* c *depends on the diameter of $\mathcal{O}$.*

Proof. As above, we only consider the most complicated case, i.e., the case of $\mathcal{R}^{\nu}_{kij}$. Notice that $\max\{i, j\} > \text{const } |k|^{\frac{1}{b-1}}$ for $b > 1$ or $|i - j| > \text{const } |k|$ for $b = 1$ implies

$$\begin{aligned} |\lambda_i \pm \lambda_j| &= (j^b - i^b)(1 + O(i^{-\delta} + j^{-\delta})) \\ &\geq \frac{1}{2}|j - i|(i^{b-1} + j^{b-1})(1 + O(i^{-\delta} + j^{-\delta})) \geq \text{const } |k|. \end{aligned} \quad (6.11)$$

It follows that $\mathcal{A}_{ij}$ defined in (6.3) is invertible and

$$\|(\mathcal{A}_{ij})^{-1}\| < |k|^{-1}.$$

By Neumann series, we have $\|(\mathcal{A}_{ij} + \mathcal{B}^{\nu}_{ij})^{-1}\| < 2|k|^{-1}$ for large k (say $|k| > K_0$), i.e, $\mathcal{R}^{\nu}_{kij} = \emptyset$. $\square$

Lemma 6.3. *For $b \geq 1$, we have*

$$\text{meas}\Big(\bigcup_{\nu \geq 0} \mathcal{R}^{\nu}\Big) = \text{meas} \bigcup_{\nu, |k| > K_\nu, i, j} (\mathcal{R}^{\nu}_k \cup \mathcal{R}^{\nu}_{ki} \cup \mathcal{R}^{\nu}_{kij}) \ < \ \text{c}\, \gamma^{\frac{\delta}{1+\delta}}.$$

Proof. The measure estimates for $\mathcal{R}^0$ comes from our assumption (2.12). We then consider the estimate

$$\text{meas}\Big(\bigcup_{\nu} \bigcup_{|k| > K_\nu} \bigcup_{i,j} \mathcal{R}^{\nu}_{kij}\Big),$$

which is the most complicate one.

Let us consider separately the case $b > 1$ and the case $b = 1$. We first consider $b > 1$. By Lemmata 6.1, 6.2, if $|k| > K_0$ and $i \neq j$, we have

$$\text{meas}\,\Big(\bigcup_{i \neq j} \mathcal{R}_{kij}\Big) = \text{meas}\,\Big(\bigcup_{i \neq j; i, j < C|k|^{\frac{1}{b-1}}} \mathcal{R}^{k}_{ij}\Big) \ < \ \text{c}\, \frac{|k|^{\frac{2}{b-1}}\gamma}{|k|^{\tau-1}} = \frac{\gamma}{|k|^{\tau-1-\frac{2}{b-1}}}. \quad (6.12)$$

For $i = j$. As in Lemma 6.1, we can find Q_1, Q_2 so that (6.4) holds with the diagonal elements of A_{11} being $< k, \omega_\nu > \pm 2\lambda_i$ and $A_{22} =< k, \omega_\nu > I$. Repeating the arguments in Lemma 6.1, we get (6.9) and

$$\begin{aligned} \mathcal{R}^{\nu}_{kii} &\subset \Big\{\xi : |\det D| \ < \ \text{c}\,\Big(\frac{\gamma_\nu}{|k|^{\tau-1}}\Big)^{\bar{d}^2}\Big\} \\ &= \Big\{\xi : \prod |\langle k, \omega_\nu \rangle + O\big(i^{-\delta}\big)| \ < \ \text{c}\,\Big(\frac{\gamma_\nu}{|k|^{\tau-1}}\Big)^{\bar{d}^2}\Big\} \\ &\subset \Big\{\xi : |\langle k, \omega_\nu \rangle| \ < \ \text{c}\,\Big(\frac{\gamma}{|k|^{\tau-1}} + \frac{1}{i^{\delta}}\Big)\Big\} \equiv \mathcal{Q}^{k}_{ii}. \end{aligned} \quad (6.13)$$

Since $\mathcal{Q}^k_{ii} \subset \mathcal{Q}^k_{i_0 i_0}$ for $i \geq i_0$, using (6.10), we find that

$$\text{meas}\,\Big(\bigcup_{i} \mathcal{R}_{kii}\Big) \leq \sum_{i < i_0} |\mathcal{R}_{kii}| + |\mathcal{Q}^k_{i_0 i_0}| \ < \ \text{c}\,\left(\frac{i_0 \gamma}{|k|^{\tau-1}} + \frac{1}{i_0^{-\delta}}\right)$$

for any i_0. Following Pöschel ([16]), we choose $i_0 = \big(\frac{|k|^{\tau-1}}{\gamma}\big)^{\frac{1}{1+\delta}}$, so that

$$\text{meas}\,\Big(\bigcup_i \mathcal{R}_{kii}\Big| \ < \ \text{c}\,\Big(\frac{\gamma}{|k|^{\tau-1}}\Big)^{\frac{\delta}{1+\delta}}. \tag{6.14}$$

Let $\tau > \max\{d+2+\frac{2}{b-1}, (d+1)^{\frac{1+\delta}{\delta}}+1\}$. As in (6.12), (6.14), we find

$$\begin{aligned}\text{meas}\,\Big(\bigcup_{|k|>K_\nu}\bigcup_{i,j}\mathcal{R}^\nu_{kij}(\gamma_\nu)\Big) &= \text{meas}\,\Big(\bigcup_{|k|>K_\nu}\bigcup_{i\neq j}\mathcal{R}^\nu_{kij}(\gamma_\nu)\Big)\\ +\text{meas}\,\Big(\cup_{|k|>K_\nu}\bigcup_i \mathcal{R}^\nu_{kii}(\gamma_{\nu+1})\Big) &< \ \text{c}\ K_\nu^{-1}\gamma^{\frac{\delta}{1+\delta}}.\end{aligned}$$

The quantity $\text{meas}\big(\bigcup_\nu\bigcup_{|k|>K_\nu}\bigcup_{i,j}\mathcal{R}^\nu_{kij}\big)$ is then bounded by

$$\sum_{\nu\geq 1}\text{meas}\Big(\bigcup_{|k|>K_\nu}\bigcup_{i,j}\mathcal{R}^\nu_{kij}(\gamma_\nu)\Big) \ < \ \text{c}\,\gamma^{\frac{\delta}{1+\delta}}\sum_{\nu\geq 0}K_\nu^{-1} \ < \ \text{c}\,\gamma^{\frac{\delta}{1+\delta}}, \tag{6.15}$$

provided $\tau > \max\{d+2+\frac{2}{b-1}, (d+1)^{\frac{1+\delta}{\delta}}+1\}$. This concludes the proof for $b>1$.

Consider now $b=1$. Without loss of generality, we assume $j \geq i$ and $j = i+m$. Note that Lemma 6.2 implies $\mathcal{R}^k_{ij} = \emptyset$ for $m > C|k|$. Following the scheme of the above proof, we find

$$\begin{aligned}\bigcup_{k,i,j}\mathcal{R}_{kij} &= \bigcup_{k,i,m}\mathcal{R}_{ki,i+m} = \bigcup_{k,m<C|k|}\bigcup_i \mathcal{R}_{ki,i+m}\\ &\subset \bigcup_{k,m<C|k|}\Big(\bigcup_{i<i_0}\mathcal{R}_{ki_0,i_0+m}\cup\mathcal{Q}_{ki_0,i_0+m}\Big),\end{aligned} \tag{6.16}$$

where

$$\mathcal{Q}_{ki_0,i_0+m} = \Big\{\xi : |\langle k,\omega_\nu\rangle + m| \ < \ \text{c}\,\Big(\frac{\gamma}{|k|^{\tau-1}}+\frac{1}{i^{-\delta}}\Big)\Big\}.$$

Again, taking $i_0^{1+\delta} = \frac{|k|^{\tau-1}}{\gamma}$, we have, for fixed k,

$$\begin{aligned}\Big|\bigcup_{i,j}\mathcal{R}^k_{ij}\Big| &< \ \text{c}\sum_{m<C|k|}\Big(\frac{i_0\gamma}{|k|^{\tau-1}}+i_0^{-\delta}\Big)\\ &< \ \text{c}\ |k|\Big(\frac{\gamma}{|k|^{\tau-1}}\Big)^{\frac{\delta}{1+\delta}}\ .\end{aligned} \tag{6.17}$$

As in the case $b>1$, we have that $\text{meas}(\bigcup_\nu\bigcup_{|k|>K_\nu}\bigcup_{i,j}\mathcal{R}^\nu_{kij})$ is bounded by $O(\gamma^{\frac{\delta}{1+\delta}})$ if $\tau > (d+1)^{\frac{1+\delta}{\delta}}+1$. □

Remark. In (6.13), $|\det D| = \prod|\langle k,\omega\rangle + O(i^{-\delta})|$ (guaranteed by the regularity property) is crucial for the proof.

7. Appendix

Proof of Proposition 3.1. From the hypotheses there follows that the eigenfuctions ϕ_n are analytic (respectively, smooth) and bounded with, in particular,

$$\sup_{\mathbb{R}}(|\phi_n'| + |\phi_n''|) \leq \text{ const } \mu_n .$$

Thus, the sum defining $u(t, x)$ is uniformly convergent in $I \times [0, 2\pi]$. Since

$$\frac{\partial G}{\partial q_n} = -\frac{1}{\sqrt{\lambda_n}} \int f(\sum_k \frac{q_k}{\sqrt{\lambda_k}} \phi_k)\phi_n ,$$

one has

$$|q_n| \leq \text{const } \frac{e^{-n\rho}}{n^a}, \quad |\dot{q}_n| \leq \text{const } \lambda_n \frac{e^{-n\rho}}{n^a} \leq \text{const } \frac{e^{-n\rho}}{n^{a-1}},$$
$$|\ddot{q}_n| \leq \text{const } \frac{e^{-n\rho}}{n^{a+1}}.$$

Thus (if a is big enough, in the smooth case) $u(t, x)$ is a C^2 function and

$$\begin{aligned} u_{tt} + Au &= \sum \frac{\ddot{q}_n}{\sqrt{\lambda}} \phi_n + \frac{q_n}{\sqrt{\lambda_n}} A\phi_n \\ &= \sum \Big(\int f(u)\phi_n \Big) \phi_n = f(u), \end{aligned} \tag{7.1}$$

where in the last equality we used the fact that $f(u)$ is a smooth periodic function. □

Lemma 7.1.

$$\|FG\|_{D(r,s)} \leq \|F\|_{D(r,s)} \|G\|_{D(r,s)}.$$

Proof. Since $(FG)_{klp} = \sum_l F_{k-k',l-l',p-p'} G_{k'l'p'}$, we have that

$$\begin{aligned} \|FG\|_{D(r,s)} &= \sup_D \sum_{klp} |(FG)_{klp}|\ |y|^l\ |z^\alpha| e^{|k|r} \\ &\leq \sup_D \sum_{klp} \sum_{l'} |F_{k-k',l-l',p-p'} G_{k'l'p'}|\ |y|^l |z^\alpha| e^{|k|r} \\ &= \|F\|_{D(r,s)} \|G\|_{D(r,s)} \end{aligned} \tag{7.2}$$

and the proof is finished. □

Lemma 7.2 (Cauchy inequalities**).**

$$\|F_{\theta_i}\|_{D(r-\sigma,s)} \leq c\sigma^{-1} \|F\|_{D(r,s)},$$

and

$$\|F_I\|_{D(r,\frac{1}{2}s)} \leq 2\frac{1}{s^2} \|F\|_{D(r,s)}, \quad \|F_{z_n}\|_{D(r,\frac{1}{2}s)} \leq 2\frac{n^a e^{n\rho}}{s} \|F\|_{D(r,s)}.$$

Let $\{\cdot,\cdot\}$ is Poisson bracket of smooth functions

$$\{F, G\} = \sum \left(\frac{\partial F}{\partial \theta_i}\frac{\partial G}{\partial I_i} - \frac{\partial F}{\partial I_i}\frac{\partial G}{\partial \theta_i}\right) + \sum_{i\in\mathbb{N}} \left\langle \frac{\partial F}{\partial z_i}, \mathrm{i} J_{d_i} \frac{\partial G}{\partial z_i}\right\rangle, \tag{7.3}$$

where J_{d_i} are standard symplectic matrix in $\mathbb{R}^{d_i}$.

Lemma 7.3. *If*

$$\|X_F\|_{r,s} < \epsilon', \|X_G\|_{r,s} < \epsilon'',$$

then

$$\|X_{\{F,G\}}\|_{r-\sigma,\eta s} \ < \ \mathrm{c}\, \sigma^{-1}\eta^{-2}\epsilon'\epsilon'', \quad \eta \ll 1.$$

Proof. Note that

$$\begin{aligned}\frac{d}{dz_n}\{F, G\} \ &= \langle F_{\theta z_n}, G_I\rangle + \langle F_\theta, G_{I z_n}\rangle - \langle F_{I z_n}, G_\theta\rangle - \langle F_I, G_{\theta z_n}\rangle \\ &\quad + \sum_{i\in\mathbb{N}} (\langle F_{z_i z_n}, J_{d_i} G_{z_i}\rangle + \langle F_{z_i}, J_{d_i} G_{z_i z_n}\rangle).\end{aligned} \tag{7.4}$$

Since

$$\begin{aligned}\|\langle F_{\theta z_n}, G_I\rangle\|_{D(r-\sigma,s)} \ &< \ \mathrm{c} \ \ \sigma^{-1}\|F_{z_n}\|\| \cdot \|G_y\|, \\ \|\langle F_\theta, G_{I z_n}\rangle\|_{D(r-\sigma,\frac12 s)} \ &< \ \mathrm{c} \ \ s^{-2}\|F_\theta\| \cdot \|G_{z_n}\|, \\ \|\langle F_{I z_n}, G_\theta\rangle\|_{D(r,\frac12 s)} \ &< \ \mathrm{c} \ \ s^{-2}\|F_{z_n}\| \cdot \|G_\theta\|, \\ \|\langle F_I, G_{\theta z_n}\rangle\|_{D(r-\sigma,s)} \ &< \ \mathrm{c} \ \ \sigma^{-1}\|F_I\| \cdot \|G_{z_n}\|, \\ \|\langle F_{z_i z_n}, J_{d_i} G_{z_i}\rangle\|_{D(r,\frac12 s)} \ &< \ \mathrm{c} \ \ s^{-1}\|F_{z_n}\| \cdot \|G_{z_i}\| i^a e^{i\rho}, \\ \|\langle F_{z_i z_n}, J_{d_i} G_{z_i}\rangle\|_{D(r,\frac12 s)} \ &< \ \mathrm{c} \ \ s^{-1}\|F_{z_n}\| \cdot \|G_{z_i}\| i^a e^{i\rho},\end{aligned} \tag{7.5}$$

it follows from the definition of the weighted norm (see (2.6)), that

$$\|X_{\{F,G\}}\|_{r-\sigma,\eta s} \ < \ \mathrm{c}\, \sigma^{-1}\eta^{-2}\epsilon'\epsilon''.$$

In particular, if $\eta \sim \epsilon^{\frac13}, \epsilon', \epsilon'' \sim \epsilon$, we have $\|X_{\{F,G\}}\|_{r-\sigma,\eta s} \sim \epsilon^{\frac43}$. □

Lemma 7.4. *Let $\mathcal{O}$ be a compact set in $\mathbb{R}^d$ for which (4.2) holds. Suppose that $f(\xi)$ and $\omega(\xi)$ are C^m Whitney-smooth function in $\xi \in \mathcal{O}$ with C^m_W norm bounded by L. Then*

$$g(\xi) \equiv \frac{f(\xi)}{\langle k, \omega(\xi)\rangle}$$

is C^m Whitney-smooth in $\mathcal{O}$ with[20]

$$\|g\|_{\mathcal{O}} \ < \ \mathrm{c}\, \gamma^{-\mathrm{c}}|k|^{\mathrm{c}} L.$$

Proof. The proof follows directly from the definition of Whitney's differentiability. □

[20] Recall the definition in (2.4).

A similar lemma for matrices holds:

Lemma 7.5. *Let $\mathcal{O}$ be a compact set in $\mathbb{R}^d$ for which (4.2) holds. Suppose that $B(\xi)$, $A_i(\xi)$ are C^m Whitney-smooth matrices and $\omega(\xi)$ is a Whitney-smooth function in $\xi \in \mathcal{O}$ bounded by L. Then*

$$C(\xi) = BM^{-1},$$

is C^m Whitney-smooth with

$$\|F\|_{\mathcal{O}} \;<\; \mathrm{c}\,\gamma^{-\mathrm{c}}|k|^{\mathrm{c}}L,$$

where M stands for either $\langle k,\omega\rangle I_{d_i} + A_i J_{d_i}$ if B is $(d_i \times d_i)$-matrix, or $\langle k,\omega\rangle I_{d_i d_j} + (A_i J_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j} A_j)$ if B is $(d_i d_j \times d_i d_j)$-matrix,

For a $N \times N$ matrix $M = (a_{ij})$, we denote by $|M|$ its determinant. Consider M as a linear operator on $(R^N, |\cdot|)$, where $|x| = \sum |x_i|$. Let $\|M\|$ be its operator norm and recall that $\|M\|$ is equivalent to the norm $\max |a_{ij}|$; thus disregarding a constant (depending only on dimensions) we will simply denote $\|M\| = \max |a_{ij}|$.

Lemma 7.6. *Let M be a $N \times N$ non-singular matrix with $\|M\| \;<\; \mathrm{c}\,|k|$, then*

$$\{\omega : \|M^{-1}\| > h\} \subset \left\{\omega : |\det M| \;<\; \mathrm{c}\,\frac{|k|^{N-1}}{h}\right\}.$$

Proof. First, note that if M is a nonsingular $N \times N$ matrix with elements bounded by $|m_{ij}| \le m$, its inverse is $M^{-1} = \frac{1}{|M|}\mathrm{adj}M$ so that

$$\|M^{-1}\| \;<\; \mathrm{c}\,\frac{m^{N-1}}{|\det M|}$$

with a constant depending on N. In particular, if $m = \mathrm{const}|k|$, $|\mathrm{Det}M| > \frac{|k|^{N-1}}{h}$, then

$$\|M^{-1}\| \;<\; \mathrm{c}\,h.$$

This proofs the lemma. □

In order to estimate the measure of $\mathcal{R}^{\nu+1}$, we need the following lemma, which has been proven in [19,21]. A similar estimate is also used by Bourgain [4].

Lemma 7.7. *Suppose that $g(u)$ is a C^m function on the closure $\bar{I}$, where $I \subset R^1$ is a finite interval. Let $I_h = \{u : |g(u)| < h\}$, $h > 0$. If for some constant $d > 0$, $|g^{(m)}(u)| \ge d$ for all $u \in I$, then* meas $(I_h) \le ch^{\frac{1}{m}}$, *where $c = 2(2+3+\cdots+m+d^{-1})$.*

For the proof of Lemma 3.1, we need the following

Lemma 7.8.

$$\sum_{j\in\mathbb{Z}} e^{-|n-j|r+\rho|j|} \le Ce^{\rho|n|}, \quad \sum_{j,n\in\mathbb{Z}} |q_j| e^{-|n-j|r+|n|\rho} \le C|q|_\rho$$

if $\rho < r$, $q \in \mathcal{Z}_\rho$ where C depends on $r-\rho$.

Lemma 7.9.

$$\sum_{j\in\mathbb{Z}}(1+|n-j|)^{-K}|j|^a \;<\; \mathrm{c}\,|n|^a,\ \sum_{j,n\in\mathbb{Z}}|q_j|(1+|n-j|)^{-k}|n|^a \le C|q|_a$$

if $K > a+1$, $q \in \mathcal{Z}_{a,\rho=0}$, *where* C *depends on* $K-a-1$.

The proofs of the above two lemmata are elementary and we omit them.

A direct proof of Lemma 3.1. It is clearly enough to consider the case of $f(u)$ equal to a monomial u^{N+1} for some $N \ge 1$. From (3.10), one can see that the regularity of G implies the regularity of $\tilde{G}$. In the following, we shall give the proof for G.

Suppose that the potential $V(x)$ is analytic in $|\mathrm{Im}x| < r$ (respectively, belongs to Sobolev space H^K) then the eigenfunctions are analytic in $|\mathrm{Im}x| < r$ (respectively, belong to H^{K+2}). If we let $\phi_i(x) = \sum a_i^n e^{\mathrm{i}\langle n,x\rangle}$, then (see, e.g.,[7])

$$|a_i^n| \;<\; \mathrm{c}\, e^{-|i-n|r} \quad \text{respectively} \quad |a_i^n| \;<\; \mathrm{c}\,(1+|n-i|^{-K-2}).$$

Recall that

$$G(q) = \sum_{i_0,\cdots,i_N} C_{i_0\cdots i_N}\frac{q_{i_0}\cdots q_{i_N}}{\sqrt{\lambda_{i_0}\cdots\lambda_{i_N}}},$$

where

$$C_{i_0\cdots i_N} = \int_{T^1}\phi_{i_0}\cdots\phi_{i_N}dx = \sum_{n_0+n_1+\cdots+n_N=0}(\prod_{s=0}^{N}a_{i_s}^{n_s}),$$

with $|a_{i_s}^{n_s}| \;<\; \mathrm{c}\, e^{-|i_s-n_s|r}$ (respectively, $|a_{i_s}^{n_s}| \;<\; \mathrm{c}\,(1+|n_s-i_s|^{-K-2})$.

In what follows, we assume either $a=0$, $\rho>0$ or $a>0$, $\rho=0$. Since

$$G_{q_j} = (N+1)\sum_{i_1,\cdots,i_N} C_{ji_1\cdots i_N}\frac{q_{i_1}\cdots q_{i_N}}{\sqrt{\lambda_j\lambda_{i_1}\cdots\lambda_{i_N}}},$$

it follows that

$$\begin{aligned}
\|G_q\|_{a+\frac12,\rho} &= \|G_{q_0}\| + \sum_{j\ge1}|G_{q_j}||j|^{a+\frac12}e^{j\rho}\\
&< \mathrm{c}\sum_{\substack{j,i_1,\cdots,i_N,\\ n_0+\cdots+n_N=0}}|a_j^{n_0}|j^a e^{|j|\rho}(\prod_{s=1}^{N}|a_{i_s}^{n_s}q_{i_s}|)\\
&< \mathrm{c}\sum_{\substack{j,i_1,\cdots i_N;\\ n_0+\cdots+n_N=0}}(1+|j-n_0|)^{-N}|j|^a e^{|j|\rho-|n_0-j|r}\left(\prod_{s=1}^{N}(1+|n_s-i_s|)^{-K-2}e^{-|n_s-i_s|r}|q_{i_s}|\right)\\
&< \mathrm{c}\sum_{\substack{i_1,\cdots,i_N;\\ n_0+\cdots+n_N=0}}|n_0|^a e^{|n_0|\rho}\left(\prod_{s=1}^{N}(1+|n_s-i_s|)^{-K-2}e^{-|n_s-i_s|r}|q_{i_s}|\right)
\end{aligned}$$

$$< c \sum_{\substack{i_1,\cdots,i_N;\\ n_1,\cdots,n_N}} (|\sum_{s=1}^{N} n_s|)^a e^{|\sum_{s=1}^N n_s|\rho} \left(\prod_{s=1}^{N}(1+|n_s-i_s|)^{-K-2} e^{-|n_s-i_s|r} |q_{i_s}|\right)$$

$$< c \sum_{\substack{i_1,\cdots,i_N;\\ n_1,\cdots,n_N}} \left(\prod_{s=1}^{N}(1+|n_s-i_s|)^{-K-2} |n_s|^a e^{-|n_s-i_s|r+|n_s|\rho} |q_{i_s}|\right)$$

$$< c \sum_{i_1,\cdots,i_N} \left(\prod_{s=1}^{N} |i_s|^a e^{|i_s|\rho} |q_{i_s}|\right)$$

$$< c \prod_{s=1}^{N} \left(\sum_{i_s} |i_s|^a e^{|i_s|\rho} |q_{i_s}|\right) < c\ |q|_{a,\rho}^N. \tag{7.6}$$

□

References

1. Arnold, V. I.: Proof of A. N. Kolmogorov's theorem on the preservation of quasi periodic motions under small perturbations of the Hamiltonian. Usp. Math. USSR **18**, 13–40 (1963)
2. Arnol'd, V. I.: Small denominators and problems of stability of motions in classical and celestial mechanics. Russ. Math. Surv. **18** 6, 85–191 (1963)
3. Bourgain, J.: *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*. International Mathematics Research Notices, 1994, pp. 475–497
4. Bourgain, J.: Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. Ann. of Math. To appear
5. Bourgain, J.: Nonlinear Schrödinger equations. To appear in *IAS-Park City Lectures(AMS)*
6. Craig, W.: KAM theory in infinite dimensions. Lectures in Applied Mathematics, **31**, 1996
7. Craig, W., and Wayne, C.E.: Newton's method and periodic solutions of nonlinear wave equations. Commun. Pure Appl. Math. **46**, 1409–1498 (1993)
8. Eliasson, L.H.: Perturbations of stable invariant tori for Hamiltonian systems. Ann. Sc. Norm. Sup. Pisa **15**, 115–147 (1988)
9. Fröhlich, J., and Spencer, T.: Absence of diffusion in the Anderson tight binding model for large disorder or lower energy. Commun. Math. Phy. **88**, 151–184 (1983)
10. Kuksin, S.B.: *Nearly integrable infinite dimensional Hamiltonian systems*. Lecture Notes in Mathematics, **1556**, Berlin: Springer, 1993
11. Kuksin, S.B., Pöschel, J.: Invariant cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation. Ann. of Math. **142**, 149–179 (1995)
12. Lancaster, P.: *Theory of Matrices*. New York and London: Academic Press LTD, 1969
13. Melnikov, V.K.: On some cases of the conservation of conditionally periodic motions under a small change of the Hamiltonian function. Sov. Math. Doklady **6**, 1592–1596 (1965)
14. Moser, J.: Convergent series expansions for quasiperiodic motions. Math. Ann. **169** (1), 136–176 (1967)
15. Pöschel, J.: Quasi-periodic solutions for a nonlinear wave equation. Comment. Math. Helvetici **71**, 269–296 (1996)
16. Pöschel, J.: A KAM-Theorem for some Nonlinear Partial Differential Equations. Ann. Scuola Norm. Sup. Pisa Cl. Sci. **23**, 119–148 (1996)
17. Wayne, C.E.: Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory. Commun. Math. Phys. **127**, 479–528 (1990)
18. Whitney, H.: Analytical extension of differentiable functions defined on closed set. Trans. A. M. S. **36**, 63–89 (1934)
19. Xu, J., You, J., Qiu, Q.: Invariant tori for nearly integrable Hamiltonian systems with degeneracy. Math. Z. **226**, 375–387 (1997)
20. You, J.: Perturbation of lower dimensional tori for Hamiltonian systems. J. Differ. Eqs. **152**, 1–29 (1999)
21. You, J.: A KAM theorem for hyperbolic-type degenerate lower dimensional tori in Hamiltonian systems. Commun. Math. Phys. **192**, 145–168 (1998)

Communicated by Ya. G. Sinai