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

Luigi Chierchia
Dipartimento di Matematica
Università Roma Tre
E-mail: luigi@matrm3.mat.uniroma3.it

Jiangong You*
Department of Mathematics
Nanjing University, Nanjing 210093, China
E-mail: jyou@nju.edu.cn

July 26, 2018

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
\[ u_{tt} - u_{xx} + V(x)u = f(u), \quad f(u) = O(u^2), \quad (1.1) \]
see Wayne [16], Kuksin [3] and Pöschel [14]. In such papers, KAM theory for lower dimensional tori [13], [12], [7] (i.e., invariant tori of dimension lower than the number

*The research was partly supported by NNSF of China and by the Italian CNR grant # 211.01.31.
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 Wayne [6] who established the existence of periodic solutions. The techniques used in [6] are based not on KAM theory, but rather on a generalization of the Lyapunov-Schmidt procedure and on techniques by Fröhlich and Spencer [8]. Recently, Craig and Wayne’s approach has been significantly improved by Bourgain [3], [4] 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 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.

The purpose of this paper is to show that, allowing for more general normal forms, one can indeed use KAM techniques to deal also with the multiple normal frequency case arising in PDE’s with periodic boundary conditions.

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 potential will be taken in a d-dimensional family of functions parameterized by a real d-vector $\xi$, $V(x) = V(x, \xi)$, satisfying general non-degenerate (“non-resonance–of–eigenvalue”) conditions. Then for “most” potentials in the family (i.e. for most $\xi$ 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 the case the operator A is positive definite.

We believe that the technique used in this paper can be generalized so as to cover 2D wave equations.

The paper is organized as follows: In section 2 we formulate a general infinite dimensional KAM Theorem designed to deal with multiple normal frequency cases; in section 3 we show how to apply the preceding KAM Theorem to the nonlinear wave equation (1.1) with periodic boundary conditions. The proof of the KAM Theorem is provided in sections 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. Let \( \mathcal{Z} \equiv \bigotimes_{n \in \mathbb{N}} \mathbb{C}^{d_n} \); the coordinates in \( \mathcal{Z} \) are given by \( z = (z_0, z_1, z_2, \cdots) \) with \( z_n \equiv (z_{1n}, \cdots, z_{dn}) \in \mathbb{C}^{d_n} \). Given two real numbers \( a, \rho \), 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

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

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

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

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

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

For positive numbers \( r, s \) we denote by

\[
D_{a, \rho}(r, s) = \{ (\theta, I, z) \in \mathcal{P}_{a, \rho} : |\text{Im} \theta| < r, |I| < s^2, |z|_{a, \rho} < s \}
\]

(2.1)

a complex neighborhood of \( \mathcal{T}^d \times \{ I = 0 \} \times \{ z = 0 \} \). 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

\[
F = \sum_{\alpha} F_\alpha z^\alpha,
\]

(2.2)

where the multi-index \( \alpha \) runs over the set \( \alpha \equiv (\alpha_0, \alpha_1, ...) \in \bigotimes_{n \in \mathbb{N}} \mathbb{N}^{d_n} \) with finitely many non-vanishing components; (iii) for each \( \alpha \), the function \( F_\alpha = F_\alpha(\theta, I, \xi) \) is real analytic in the variables \( (\theta, I) \in \{ |\text{Im} \theta| < r, |I| < s^2 \} \); (iv) for each \( \alpha \), the dependence of \( F_\alpha \) upon the parameter \( \xi \) is of class \( C^d_{\text{W}}(\mathcal{O}) \) for some \( \bar{d} > 0 \) (to be fixed later): here \( C^m_{\text{W}}(\mathcal{O}) \) denotes the class of functions which are \( m \) times differentiable on the closed set \( \mathcal{O} \) in the sense of Whitney [17].

1We use the notations \( \mathbb{N} = \{0, 1, 2, \cdots\}, \mathbb{Z}_+ = \{1, 2, \cdots\} \).

2Thus \( \exists n_0 > 0 \) such that \( z^\alpha \equiv \prod_{n=0}^{n_0} z_n^{\alpha_n} \equiv \prod_{n=0}^{n_0} \prod_{j=1}^{d_n} (z_{jn})^{\alpha_{jn}} \).
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:

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

where, if $F_{\alpha} = \sum_{k, l} F_{k\ell\alpha}(\xi) I^l e^{i(k, \theta)}$, $\|F_{\alpha}\|$ is defined as

$$\|F_{\alpha}\| \equiv \sum_{k, l} |F_{k\ell\alpha}| \|s^2l! e^{l|k|}\|, \quad |F_{k\ell\alpha}|_{\mathcal{O}} \equiv \max_{|p| \leq d^2} \left| \frac{\partial^p F_{k\ell\alpha}}{\partial \xi^p} \right|, \quad (2.4)$$

(the derivatives with respect to $\xi$ 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

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

where $J_{d_n}$ denotes the standard symplectic matrix $\begin{pmatrix} 0 & I_{d_n/2} \\ -I_{d_n/2} & 0 \end{pmatrix}$ and $i = \sqrt{-1}$; 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$^3$

$$\dot{\theta} = F_I, \quad \dot{I} = -F_\theta, \quad \dot{z}_n = i J_{d_n} F_{z_n}, \quad n \in \mathbb{N}. \quad (2.5)$$

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$^4$

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

**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 $\|X_F\|_{D_{a,\rho}(r, s), \mathcal{O}}$ 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

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

(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

\[ dI \wedge d\theta + i \sum_n \sum_{j=1}^{d_n/2} z_n^j \wedge dz_n^{j+d_n/2}. \]

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.,

\[ \frac{d\theta}{dt} = \omega, \quad \frac{dI}{dt} = 0, \quad \frac{dz_n}{dt} = iJ_{d_n} A_n z_n, \quad n \in \mathbb{N}, \]  

(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

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

(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 make 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

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

(2.10)

where \( \lambda_n \) are real and independent of \( \xi \) while \( B_n \) may depend on \( \xi \); furthermore, the behaviour of \( \lambda_n \)’s is assumed to be as follows

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

(2.11)

(A2) Gap condition: There exists \( \delta_1 > 0 \) such that

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

(\( \sigma(\cdot) \) 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 $d^2$ Whitney–smooth functions of $\xi$ with $C^d_{Wh}$-norm bounded by some positive constant $L$.

(A4) **Non-resonance condition:**

$$\text{meas}\{\xi \in \mathcal{O} : \langle k, \omega(\xi) \rangle \langle \langle k, \omega(\xi) \rangle + \lambda(\xi) \rangle \langle \langle k, \omega(\xi) \rangle + \lambda(\xi) + \mu(\xi) \rangle = 0 \} = 0,$$  \hspace{1cm} (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 $\|X_P\|_{D_{a,\rho}(r,s), \mathcal{O}} < \infty$ with $\bar{a} > a$. In fact, we assume that one of the following holds:

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

(such conditions correspond, respectively, to analytic or smooth solutions). In the case of $d = 1$ (i.e., the periodic solution case) one can allow $\bar{a} = a$.

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 $\varepsilon = \varepsilon(d, d, b, \delta, \delta_1, \bar{a} - a, L, \gamma)$ such that if $\|X_P\|_{D_{a,\rho}(r,s), \mathcal{O}} < \varepsilon$, then the following holds true. There exists a Cantor set $\mathcal{O}_\gamma \subset \mathcal{O}$ with $\text{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) \to 0$ as $\gamma \to 0$, and two maps (real analytic in $\theta$ and Whitney smooth in $\xi \in \mathcal{O}$)

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

such that for any $\xi \in \mathcal{O}_\gamma$ and $\theta \in T^d$ the curve $t \to \Psi(\theta + \bar{\omega}(\xi)t, \xi)$ is a quasi-periodic solution of the Hamiltonian equations governed by $H = N + P$. Furthermore, $\Psi(T^d, \xi)$ 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 $\lambda_i$ 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 $\gamma$ plays the role of the Diophantine constant for the frequency $\bar{\omega}$ in the sense that, there is $\tau > 0$ such that $\forall k \in \mathbb{Z}^d \setminus \{0\}$,

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

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

(iv) The regularity property $\bar{a} > a$ is needed for the measure estimates on $\mathcal{O} \setminus \mathcal{O}_\gamma$. As we already mentioned, such regularity requirement is not necessary for periodic solution case, i.e., $d = 1$. Thus the above theorem applies immediately to the construction of periodic solutions for nonlinear Schrödinger equations.

(v) The non-degeneracy condition (2.12) (which is stronger than Bourgain’s non-degenerate condition [11] but weaker than Melnikov’s one [12]) 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

\[ u(t) = u(t, x + 2\pi), \quad u_t(t) = u_t(t, x + 2\pi), \]  

\[ A = -u_{xx} + V(x, \xi)u, \quad x, t \in \mathbb{R}, \]

\[ u(t, x) = u(t, x + 2\pi), \quad u_t(t, x) = u_t(t, x + 2\pi), \]  

(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 \( \mu_n \) 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:

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

(3.2)

Following Kuksin \cite{kuk} and Bourgain \cite{bour}, we consider a family of real analytic (or smooth) potentials \( V(x, \xi) \), where the \( d \)-parameters \( \xi = (\xi_1, \cdots, \xi_d) \in O \subset \mathbb{R}^d \) are simply taken to be a given set of \( d \) frequencies \( \lambda_n \equiv \sqrt{\mu_n} \):

\[ \xi_i \equiv \sqrt{\mu_n} \equiv \lambda_n, \quad i = 1, \cdots, d \]  

(3.3)

where \( \mu_n \) are (positive) eigenvalues of \( A \).

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

\[ |\mu_k - \mu_h| > \delta_1, \]  

(3.4)

for all \( k > h \) except when \( k \) is even and \( h = k - 1 \) (in which case \( \mu_k \) and \( \mu_h \) might even coincide).

Notice that, in particular, having \( d \) eigenvalues as independent parameters excludes the constant potential case \( V \equiv \text{constant} \) (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 \cite{cra} and Bourgain \cite{bour, bour2}.

Equation (3.1) may be rewritten as

\[ \dot{u} = v, \quad \dot{v} + Au = f(u), \]  

(3.5)

\footnote{Plenty of such potentials may be constructed with, e.g., the inverse spectral theory.}
which, as well known, may be viewed as the (infinite dimensional) Hamiltonian equations
\[ \dot{u} = H \dot{v}, \quad \dot{v} = -H \dot{u}, \]
associated to the Hamiltonian
\[ H = \frac{1}{2}(v, v) + \frac{1}{2}(Au, u) + \int_I g(u) \, dx, \quad (3.6) \]
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 \[14\], 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 \(\lambda_n \equiv \sqrt{\mu_n}\). System \([3.5]\) is then formally equivalent to the lattice Hamiltonian equations
\[ \dot{q}_n = \lambda_n p_n, \quad \dot{p}_n = -\lambda_n q_n - \frac{\partial G}{\partial q_n}, \quad G \equiv \int_I g\left(\sum_{n \in \mathbb{N}} \frac{q_n}{\sqrt{\lambda_n}} \phi_n\right) dx \quad (3.7) \]
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 \[14\], 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
\[ t \in I \rightarrow (q(t), p(t)) \equiv \left(\{q_n(t)\}_{n \geq 0}, \{p_n(t)\}_{n \geq 0}\right) \]
be an analytic (respectively, smooth) solution of \([3.7]\) such that
\[ \sup_{t \in I} \sum_{n \in \mathbb{N}} \left(|q_n(t)| + |p_n(t)|\right)^a e^n < \infty \quad (3.8) \]
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: \(w_n = \frac{1}{\sqrt{2}}(q_n + ip_n), \bar{w}_n = \frac{1}{\sqrt{2}}(q_n - ip_n)\). Equations \([3.7]\) read then
\[ \dot{\bar{w}}_n = -i\lambda_n \bar{w}_n - \frac{1}{\sqrt{2}} \frac{\partial \tilde{G}}{\partial \bar{w}_n}, \quad \dot{w}_n = i\lambda_n \bar{w}_n + \frac{1}{\sqrt{2}} \frac{\partial \tilde{G}}{\partial w_n}, \quad (3.9) \]
where the perturbation \(\tilde{G}\) is given by
\[ \tilde{G}(w) = \int_I g\left(\sum_{n \in \mathbb{N}} \frac{w_n + \bar{w}_n}{\sqrt{2\lambda_n}} \phi_n\right) dx \quad (3.10) \]

---

6 Recall that, for simplicity, we assume that all eigenvalues \(\mu_n\) are positive.
7 Regularity refers to the components \(q_n\) and \(p_n\).
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,

\[
I_i = w_{n_i} \bar{w}_{n_i}, \quad i = 1, \cdots, d,
\]

so that the system (3.9) becomes

\[
\begin{align*}
\frac{d\theta_j}{dt} &= \omega_j + P_{\theta_j}, \quad \frac{dI_j}{dt} = -P_{\theta_j}, \quad j = 1, \cdots, d, \\
\frac{dw_n}{dt} &= -i\lambda_n w_n - iP_{w_n}, \quad \frac{d\bar{w}_n}{dt} = i\lambda_n \bar{w}_n + iP_{\bar{w}_n}, \quad n \neq n_1, n_2, \cdots, n_d,
\end{align*}
\]

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 \((\theta, I)\) variables and the frequencies \(\omega = (\omega_1, \cdots, \omega_d)\) coincide with the parameter \(\xi\) introduced in (3.3):

\[
\omega_i \equiv \xi_i = \lambda_{n_i}.
\]

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

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

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 \(0 < s < |I_0|\)) 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 \(n\), say \(n > \bar{n} > n_d\) and denote by \(z_0 = (\{ w_n \}_{0 \leq n < \bar{n}}, \{ \bar{w}_n \}_{0 \leq n < \bar{n}})\) the remaining conjugated variables. The Hamiltonian (3.13) takes the form

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

where

\[
A_n = \text{Diag}(\lambda_{2n-1}, \lambda_{2n}, \lambda_{2n-1}, \lambda_{2n}) \begin{pmatrix} 0 & I_d \\ I_d & 0 \end{pmatrix}
\]

\[
= \lambda_{2n} \begin{pmatrix} 0 & I_d \\ I_d & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \lambda_{2n-1} - \lambda_{2n} & 0 \\ 0 & 0 & 0 & 0 \\ \lambda_{2n-1} - \lambda_{2n} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
\]

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

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

\footnote{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

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

(3.15)

here the parameter $a$ is taken to be 0 (respectively, the parameter $\rho$ 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:

$$
d_n = 4, n \geq 1; d_0 = \bar{n} + 1 - d; \bar{d} = \max\{d_0, 4\}; b = 1; \delta = 2; \delta_1 \text{ choosen as in (3.4)}; \bar{a} - a = \frac{1}{2}.
$$

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 $\text{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) \to 0$ as $\gamma \to 0$, such that $\{\xi \in \mathcal{O}_\gamma\}$ has a family of small-amplitude (proportional to some power of $\gamma$), analytic (respectively, smooth) quasi-periodic solutions of the form

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

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

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 $\mu_n$’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\(^5\)). 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 $N_\nu$ is an “integrable normal form” and $P_\nu$ 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 $a$, $\rho$ and $\bar{a}$ are fixed throughout the proof and are therefore omitted in the notations.
so that the vector field $X_{H_v \circ \Phi_v}$ defined on $D(r_{v+1}, s_{v+1})$ satisfies

$$\|X_{H_v \circ \Phi_v} - X_{N_{v+1}}\|_{r_{v+1}, s_{v+1}, O_{v+1}} \leq \epsilon^v$$

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

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

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

(4.1)

defined in $D(r, s) \times O$; teh $A_n$'s are symmetric matrices. We assume that $\xi \in O$ satisfies

$$12$$

(for a suitable $\tau > 0$ to be specified later)

$$\|\langle k, \omega \rangle^{-1} - 1\| < \frac{|k| \tau}{\gamma}, \quad \|\langle k, \omega \rangle I_d + A_i J_d \rangle^{-1} - \frac{|k| \tau}{\gamma} d^2,$$

(4.2)

We also assume that

$$\max_{|p| \leq d^2} \left\| \frac{\partial^p A_n}{\partial x^p} \right\| \leq L,$$

(4.3)

on $O$, and

$$\|X_P\|_{r, s, O} \leq \epsilon.$$

(4.4)

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

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

(4.5)

where

$$\Gamma(t) = \sup_{u \geq 1} u^c e^{-\frac{1}{4} u t} \sim t^{-c}.$$

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 $O_+ \subset O$ and a change of variables $\Phi : D_+ \times O_+ = D(r_+, s_+) \times O_+ \rightarrow D(r, s) \times O$, 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 $\xi \in O_+$.

$^{12}$ 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

$$A \otimes B = (a_{ij}B) = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \cdots & \cdots & \cdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}.$$ 

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

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

4.1 Solving the linearized equation

Expand $P$ into the Fourier-Taylor series

$$P = \sum_{k,l,\alpha} P_{kl\alpha} e^{i(k,\theta)} I^l z^\alpha$$

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

$$R(\theta, I, z) \equiv P_0 + P_1 + P_2 \equiv \sum_{k,|l| \leq 1} P_{k0l} e^{i(k,\theta)} I^l + \sum_{k,|\alpha| = 1} P_{k0\alpha} e^{i(k,\theta)} z^\alpha,$$

with

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

It is convenient to rewrite $R$ as follows

$$R(\theta, I, z) = \sum_{k,|l| \leq 1} P_{k0l} e^{i(k,\theta)} I^l + \sum_{k,i} \langle R_k^i, z_i \rangle e^{i(k,\theta)} + \sum_{k,i,j} \langle R_{kji}^i, z_i, z_j \rangle e^{i(k,\theta)},$$

where $R_k^i, R_{kji}^i$ are respectively the $d_i$ vector and $(d_j \times d_i)$ matrix defined by

$$R_k^i = \int \frac{\partial P}{\partial z_i} e^{-i(k,\theta)} d\theta|_{z=0,I=0}, \quad R_{kji}^i = \frac{1 + \delta_j^i}{2} \int \frac{\partial^2 P}{\partial z_j \partial z_i} e^{-i(k,\theta)} d\theta|_{z=0,I=0}.$$

Note that $R_{kji}^i = (R_{ji}^i)^T$.

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

$$\|X_R\|_{r,s,o} \leq \|XP\|_{r,s,o} \leq \epsilon.$$

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

$$\|XP - R\|_{r,s,o} < c \epsilon_+.$$

Then we look for a special $F$, defined in domain $D_+ = D(r_+, s_+)$, such that the time one map $\phi^1_F$ 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

$$H \circ \phi^1_F = (N + R) \circ \phi^1_F + (P - R) \circ \phi^1_F = N + \{N, F\} + R + \frac{1}{2} \int_0^1 ds \int_0^s \{\{N + R, F\}, F\} \circ \phi^1_F dt + \{R, F\} + (P - R) \circ \phi^1_F = N_+ + P_+ + \{N, F\} + R - P_{000} - \langle \omega', I \rangle - \sum_{n \in \mathbb{N}} \langle R^0_{mn}, z_n \rangle,$$
where
\[ \omega' = \int \frac{\partial P}{\partial I} d\theta |_{I=0,z=0}, \quad R^0_{nn} = \int \frac{\partial^2 P}{\partial z^2_n} d\theta |_{I=0,z=0}, \]
\[ N_+ = N + P_{000} + \langle \omega', I \rangle + \sum_{n \in N} \langle R^0_{nn} z_n, z_n \rangle \]
\[ P_+ = \frac{1}{2} \int_0^1 ds \int_0^s \{ \{ N + R, F \}, F \} \circ X_t dt + \{ R, F \} + (P - R) \circ \phi^1. \]

We shall find a function \( F \) of the form
\[ F(\theta, I, z) = F_0 + F_1 + F_2 = \sum_{|l| \leq 1, |k| \neq 0} F_{kl0} e^{i(k, \theta)} I^l + \sum_{i \in N} \langle F_{ki}^k, z_i \rangle e^{i(k, \theta)} \]
\[ + \sum_{|k| + |i-j| \neq 0} \langle F^k_{ij}, z_i \rangle e^{i(k, \theta)}, \]  \hfill (4.12)
satisfying the equation
\[ \{ N, F \} + R - P_{000} - \langle \omega', I \rangle - \sum_{n \in N} \langle R^0_{nn} z_n, z_n \rangle = 0. \]  \hfill (4.13)

**Lemma 4.1** Equation (4.13) is equivalent to
\[ F_{kl0} = (i\langle k, \omega \rangle)^{-1} P_{kl0}, \quad k \neq 0, |l| \leq 1, \]
\[ (\langle k, \omega \rangle I_{di} + A_d J_{di}) F_{ki}^k = i R^k_i, \]
\[ (\langle k, \omega \rangle J_{di} + A_d J_{di}) F_{ij}^k - F_{ij}^k (J_{d_i} A_j) = i R^k_{ij}, \quad |k| + |i-j| \neq 0. \]  \hfill (4.14)

**Proof.** Inserting \( F \), defined in (4.12), into (4.13) one sees that (4.13) is equivalent to the following equations\(^{14}\)
\[ \{ N, F_0 \} + P_0 - \langle \omega', I \rangle = 0, \]
\[ \{ N, F_1 \} + P_1 = 0, \]
\[ \{ N, F_2 \} + P_2 - \sum_{n \in Z} \langle R^0_{nn} z_n, z_n \rangle = 0. \]  \hfill (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\(^{15}\)
\[ \{ N, F_1 \} = \langle \partial \gamma N, \partial \gamma F_1 \rangle + \langle \nabla_z N, J \nabla_z F_1 \rangle \]
\[ = \langle \partial \gamma N, \partial \gamma F_1 \rangle + \sum_i \langle \nabla_{z_i} N, i J_{d_i} \nabla z_i F_1 \rangle \]
\[ = i \sum_{k,i} (\langle k, \omega \rangle F_{ki}^k, z_i) + \langle A_i z_i, J_{d_i} F_{ki}^k \rangle e^{i(k, \theta)} \]
\[ = i \sum_{k,i} (\langle (k, \omega) I_{d_i} + A_i J_{d_i}) F_{ki}^k, z_i) e^{i(k, \theta)} . \]  \hfill (4.16)

\(^{14}\)Recall the definition of \( P_i \) in (4.6).
\(^{15}\)Recall the definition of \( N \) in (4.1).
It follows that $F^k_i$ are determined by the linear algebraic system

$$i(\langle k, \omega \rangle I_d + A_i J_d) F^k_i + R^k_i = 0, \quad i \in \mathbb{N}, k \in \mathbb{Z}^d$$

Similarly, from

$$\{N, F^2\} = \langle \partial_I N, \partial_\theta F^2 \rangle + \sum_i \langle \nabla_{z_i} N, i J_d \nabla_{z_i} F^2 \rangle = i \sum_{|k|+|i-j| \neq 0} \langle \langle \langle k, \omega \rangle I_d z_i + A_i J_d \rangle F^k_{ji} z_j \rangle e^{i(k,\theta)}$$

it follows that, $F^k_{ji}$ is determined by the following matrix equation

$$\langle \langle \langle k, \omega \rangle I_d + A_j J_d \rangle F^k_{ji} - F^k_{ji}(J_d A_i) \rangle z_i, z_j \rangle e^{i(k,\theta)} = ir^k_{ji}, \quad |k| + |i-j| \neq 0 (4.17)$$

where $F^k_{ji}, R^k_{ji}$ 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).

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

$$AX - XB = C, \quad (4.19)$$

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 [19] or [11], p. 256.

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

14
4.2 Estimates on the coordinate transformation

We proceed to estimate $X_F$ and $\Phi^1_F$. We start with the following

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

$$\|X_F\|_{D_2,O} < c \gamma^{-c}\Gamma(r-r_+).$$

(4.20)

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

$$\|F_{k0}\|_{O} \leq |(k, \omega)|^{-1}|P_{kl}| < c \gamma^{-c}|k|^c e^{-|k|(r-r_+)}e^{|s-2l|}, \quad k \neq 0,$$

$$\|F_i^k\|_{O} = \|((k, \omega)I_{d_1} + A_iJ_{d_1})^{-1}R^k_i\| \leq \|((k, \omega)I_{d_1} + A_iJ_{d_1})^{-1}\| \cdot \|R^k_i\|$$

$$\leq c \gamma^{-c}|k|^c |R^k_i|,$$

$$\|F_{ij}^k\|_{O} \leq \|((k, \omega)I_{d_i} + A_iJ_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_i}J_{d_j})^{-1}\| \cdot \|R^k_{ij}\|$$

$$\leq c \gamma^{-c}|k|^c |R^k_{ij}|, \quad |k| + |i - j| \neq 0.$$

(4.21)

Where $\| \cdot \|_O$ for matrix is similar to (2.4).

It follows that

$$\frac{1}{s^2}\|F_{\theta}\|_{D_2,O} \leq \frac{1}{s^2}\left(\sum_{|l| \leq 1} |F_{k0}| \cdot |F_i| \cdot |k| \cdot |e^{i(k, \theta)}| + \sum_{|l| \leq 1} |F^k_i| \cdot |z_i| \cdot |k| \cdot |e^{i(k, \theta)}| \right)$$

$$< c \gamma^{-c}\Gamma(r-r_+}\|X_R\|$$

$$< c \gamma^{-c}\Gamma(r-r_+),$$

(4.22)

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

$$\|F_i\|_{D_2,O} = \sum_{|l| \leq 1} |F_{k0}| \cdot |e^{i(k, \theta)}| < c \gamma^{-c}\Gamma(r-r_+).$$

Now we estimate $\|X_{F_1}\|_{D_2,O}$. Note that

$$\|F^1_{zi}\|_{D_2,O} = \|\sum_k F^k_{zi} e^{-i<k, \theta>}\|_{D_2,O}$$

$$< c \gamma^{-c}\Gamma \sum_{k,i} |R^k_i|^c |e^{i<k, \theta>} < c \gamma^{-c}\Gamma \frac{\partial P_i}{\partial z_i}.$$
Note that
\[ \left\| F_{z_i}^2 \right\|_{D_2,\mathcal{O}} = \left\| \sum_{k,j} (F_{ij}^k + (F_{ij}^k)^T) z_j e^{i(k,\theta)} \right\|_{D_2,\mathcal{O}} < c \gamma^{-c} \Gamma \left\| \frac{\partial P_2}{\partial z_i} \right\|. \]  
(4.24)

Similarly, we have
\[ \left\| X_{F_{z_i}}^2 \right\|_{D_2,\mathcal{O}} < c \gamma^{-c} \Gamma \epsilon. \]  
(4.25)

The conclusion of the lemma follows from the above estimates.

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

**Lemma 4.4** Let \( \eta = \epsilon^4 \), \( D_{\frac{1}{2}\eta} = D(r_+ + \frac{1}{2} \eta, \frac{1}{2} \eta) \), \( i = 1, 2 \). We then have
\[ \phi_t^F : D_{\frac{1}{2}\eta} \rightarrow D_{\eta}, \quad 0 \leq t \leq 1, \]  
(4.26)

if \( \epsilon \ll \left( \frac{1}{2} \gamma^{-c} \Gamma^{-1} \right)^{\frac{3}{2}} \). Moreover,
\[ \left\| D \phi_t^F - Id \right\|_{D_{\frac{1}{2}\eta}} < c \gamma^{-c} \Gamma \epsilon. \]  
(4.27)

**Proof.** Let
\[ \left\| D^m F \right\|_{D,\mathcal{O}} = \max \left\{ \left| \frac{\partial^{\alpha} F}{\partial \theta^i \partial I^l} \right|_{D,\mathcal{O}}, |\alpha| + |i| + |l| = m \geq 2 \right\}. \]

Note that \( F \) is polynomial in \( I \) of order 1, in \( z \) of order 2. From (4.25) and the Cauchy inequality, it follows that
\[ \left\| D^m F \right\|_{D,\mathcal{O}} < c \gamma^{-c} \Gamma \epsilon, \]  
(4.28)

for any \( m \geq 2 \).

To get the estimates for \( \phi_t^F \), we start from the integral equation,
\[ \phi_t^F = id + \int_0^t X_F \circ \phi_s^F \, ds \]
\[ \phi_t^F : D_{\frac{1}{2}\eta} \rightarrow D_{\eta}, \quad 0 \leq t \leq 1, \]  
(4.28)

since
\[ D \phi_t^F = Id + \int_0^1 (DX_F)D \phi_s^F \, ds = Id + \int_0^1 J(D^2 F)D \phi_s^F \, ds, \]

it follows that
\[ \left\| D \phi_t^F - Id \right\| \leq 2 \left\| D^2 F \right\| < c \gamma^{-c} \Gamma \epsilon. \]  
(4.29)

The estimates of second order derivative \( D^2 \phi_t^F \) follows from (4.28). \( \blacksquare \)

\textsuperscript{16}Recall the definition of the weighted norm in (2.6).
4.3 Estimates for the new normal form

The map \( \phi_1^F \) defined above transforms \( H \) into \( H_+ = N_+ + P_+ \) (see (4.11) and (4.13)) with

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

where

\[
e_+ = e + P_{000}, \quad \omega_+ = \omega + P_{000}(|l| = 1), \quad A_i^+ = A_i + 2R_{i+}^0.
\]

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

\[
|\omega_+ - \omega| < \varepsilon, \quad \|R_{i+}^0\| < \epsilon i^{-\delta}
\]

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

\[
\|(A_i^+)^{-1}\| \leq \frac{\|A_i^{-1}\|}{1 - 2\|A_i^{-1} R_{i+}^0\|} \leq 2\|A_i^{-1}\|,
\]

\[
\|(k, \omega + P_{000})I_{d_i} - J_{d_i} A_i^+\)^{-1}\| \leq \frac{\|(k, \omega)I_{d_i} + A_i J_{d_i}\)^{-1}\|}{1 - \|(k, \omega)I_{d_i} + A_i J_{d_i}\)^{-1}\|} \leq \left(\frac{|k|}{\gamma_+}\right)^{\bar{a}},
\]

provided \(|k|^2 \varepsilon < c (\gamma^2 - \gamma_+^2)\). Similarly, we have

\[
\|(k, \omega + P_{000})I_{d_j} (A_i^+ J_{d_j}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j} A_j^+)\)^{-1}\| \leq \left(\frac{|k|}{\gamma_+}\right)^{\bar{a}},
\]

provided \(|k|^2 \varepsilon < c (\gamma^2 - \gamma_+^2)\). This means that in the next KAM step, small denominator conditions are automatically satisfied for \(|k| < K\) where \( K^2 \varepsilon < c (\gamma^2 - \gamma_+^2)\).

The following bounds will be used later for the measure estimates.

\[
\left|\frac{\partial^l \langle \omega_+ - \omega \rangle}{\partial \xi^l}\right| \leq \varepsilon, \quad \left|\frac{\partial^l (A_i^+ - A_i)}{\partial \xi^l}\right| \leq c \epsilon i^{-\delta},
\]

for \(|l| \leq \bar{a}^2\) by the definition of the norm.

4.4 Estimates for the new perturbation

To complete the KAM step we have to estimate the new error term. By the definition of \( \phi_1^F \) and Lemma 4.4,

\[
H \circ \phi_1^F = N_+ + P_+
\]

is well defined in \( D_{\frac{2}{\gamma}}\). Moreover, we have the following estimates

\[
\|X_{P_+}\|_{D_{\frac{2}{\gamma}}} = \|X \int_0^1 dt \int_0^t \{ (N + R, F) \circ \phi_1^F + (R, F) + (P - R) \circ \phi_1^F \} D_{\frac{2}{\gamma}}\] 
\[\leq \|X \int_0^1 dt \int_0^t \{ (N + R, F) \circ \phi_1^F \} D_{\frac{2}{\gamma}} + \|X (P - R) \circ \phi_1^F \| D_{\frac{2}{\gamma}}\] 
\[\leq \|X \{ (N + R, F) \} \|_{D_{\gamma}} + \|X_{P-R}\|_{D_{\gamma}}\] 
\[< c \gamma^{-c} T^2 e^{\frac{4}{3}} \] < c \epsilon+ 
\] (4.36)
by (4.3) and Lemma 7.3.

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

$$
\|X_{p_+}\|_{r_+ s_+} = \|X_{p_+}\|_{D_{\frac{\gamma}{2}} \rho} \leq c \gamma^{-c} \Gamma^2 \eta \epsilon = c \epsilon_+.
$$

(4.37)

The KAM step is now completed.

5 Iteration Lemma and Convergence

For any given $s, \epsilon, r, \gamma$, we define, for all $\nu \geq 1$, the following sequences

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

(5.1)

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

Note that

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

is a well defined finite function of $r$.

5.1 Iteration Lemma

The analysis of the preceeding section can be summarized as follows.

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

$$
N_\nu = \epsilon_\nu + \langle \omega_\nu(\xi), I \rangle + \sum_{i \in \mathbb{N}} \langle A_\nu(\xi) z_i, z_i \rangle,
$$
be a normal form with parameters $\xi$ satisfying

$$\| \langle k, \omega_\nu \rangle^{-1} \| < \frac{|k|^\tau}{\gamma_\nu}, \quad \|(i(k, \omega_\nu)I_{d_i} + A_i^\nu J_{d_i})^{-1}\| < \left(\frac{|k|^\tau}{\gamma_\nu}\right)^d$$

$$\|(i(k, \omega_\nu)I_{d_id_j} + (A_i^\nu J_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j} A_j^\nu))^{-1}\| < \left(\frac{|k|^\tau}{\gamma_\nu}\right)^d$$  (5.2)

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

$$\left| \frac{\partial^2 (\omega_\nu - \omega_{\nu-1})}{\partial \xi^2} \right| \leq \epsilon_{\nu-1}, \quad \left| \frac{\partial^2 (A_i^\nu - A_i^{\nu-1})}{\partial \xi^2} \right| \leq \epsilon_{\nu-1}^{-1-\delta},$$

on $O_\nu$ in Whitney’s sense.

Finally, assume that

$$\|X_{P_\nu}\|^{\frac{\alpha, \rho}{D_\nu, O_\nu}} \leq \epsilon_\nu.$$

Then, there is a subset $O_{\nu+1} \subset O_\nu$,

$$O_{\nu+1} = O_\nu \cup \{|k| \geq K_{\nu+1} \mathcal{R}_{kij}^{\nu+1}(\gamma_\nu)\}$$

where

$$\mathcal{R}_{kij}^{\nu+1}(\gamma_\nu) = \left\{ \xi \in O_\nu : \|\langle k, \omega_{\nu+1} \rangle^{-1} \| \leq \frac{|k|^\tau}{\gamma_{\nu+1}}, \quad \|\langle k, \omega_{\nu+1} \rangle I_{2m} + (A_i^{\nu+1} J_{d_i})^{-1}\| \leq \left(\frac{|k|^\tau}{\gamma_\nu}\right)^d \right\}$$

with $\omega_{\nu+1} = \omega_\nu + P_{000}^\nu$, and a symplectic change of variables

$$\Phi_\nu : D_{\nu+1} \times O_{\nu+1} \to D_\nu,$$  (5.3)

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

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

satisfying

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

$$\|X_{P_{\nu+1}}\|^{\frac{\alpha, \rho}{D_{\nu+1}, O_{\nu+1}}} \leq \epsilon_{\nu+1}.\quad (5.6)$$

### 5.2 Convergence

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

$$\epsilon_0 = \epsilon, \gamma_0 = \gamma, s_0 = s, L_0 = L, N_0 = N, P_0 = P,$$
\[ \mathcal{O}_0 = \left\{ \xi \in \mathcal{O} : \left| \frac{(k, \omega)^{-1}}{\gamma} \right| \left\| (k, \omega) I_{d_i} + A_i J_d \right\|^{-1} \left\| \frac{(k, \omega)}{\gamma} \right\|^2 \right\}, \]

(with \( \epsilon \) and \( \gamma \) 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 [15], 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

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

Since

\[ \epsilon_{\nu+1} = c \gamma^{-\epsilon} \Gamma(r_{\nu} - r_{\nu+1}) \epsilon_\nu \leq (c \gamma^{-\epsilon} \Psi(r) \epsilon)^{(\frac{1}{3})^\nu}. \]

It follows that \( \epsilon_{\nu+1} \to 0 \) provided \( \epsilon \) is sufficiently small.

Let \( \phi^t_H \) be the flow of \( X_H. \) Since \( H \circ \Psi^\nu = H_{\nu+1}, \) we have that

\[ \phi^t_H \circ \Psi^\nu = \Psi^\nu \circ \phi^t_{H_{\nu+1}}. \]

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

\[ \phi^t_H \circ \Psi^\infty = \Psi^\infty \circ \phi^t_{H^\infty}, \]

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

\[ \phi^t_H(\Psi^\infty(\mathbb{T}^d \times \{\xi\})) = \Psi^\infty \phi^t_{N^\infty}(\mathbb{T}^d \times \{\xi\}) = \Psi^\infty(\mathbb{T}^d \times \{\xi\}), \]

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 \( \xi. \) The normal behaviour of the invariant torus is governed by the matrix \( A_\infty^\nu = \sum_{\nu \in \mathbb{N}} A_\nu^\nu. \]
Measure Estimates

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

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

the sets $\mathcal{R}^\nu_k, \mathcal{R}^\nu_{ki}, \mathcal{R}^\nu_{kij}$ being respectively

$$\{ \xi \in O_\nu : |\langle k, \omega_\nu \rangle - 1| > \frac{|k|^\tau}{\gamma_\nu} \},$$

$$\{ \xi \in O_\nu : \|M_1^{-1}\| > (\frac{|k|^\tau}{\gamma_\nu})^2 \},$$

and

$$\{ \omega \in O_\nu : \|M_2^{-1}\| > (\frac{|k|^\tau}{\gamma_\nu})^2 \},$$

(6.1)

where

$$M_1 = \langle k, \omega_\nu \rangle I_d_i + A^\nu_i J_d_i$$
$$M_2 = \langle k, \omega_\nu \rangle I_d_i d_j + (A^\nu_j J_d_j) \otimes I_d_i - I_d_j \otimes (J_d_i A^\nu_i).$$

(6.2)

We include in the set $\{ \xi \in O : \|M(\omega)^{-1}\| > C \}$ also the $\xi$'s for which $M$ is not invertible.

Remind that $\omega_\nu(\xi) = \xi + \sum_{\nu=1}^{\nu-1} P_{00}(\xi)$ with $|\sum P_{00}(\xi)|_{C_{\nu}} \leq \epsilon$, $A^\nu_i = A_i + 2\sum_{\nu} R^0_{ii}$ with $\|\sum_{\nu} R^0_{ii}\| = O(\epsilon_{i}^{-\delta})$.

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

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

Proof. As it is well known

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

The set $\mathcal{R}^\nu_{ki}$ is empty if $i > \text{const} |k|$, while, if $i \leq \text{const} |k|$, from Lemmata 6.6, 7.7 there follows that

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

We now give a detailed proof for the most complicated estimate, i.e., the estimate on the measure of the set $\mathcal{R}^\nu_{kij}$. Rewrite $\mathcal{M}_2$ as

$$\mathcal{M}_2 \equiv A_{ij} + B^\nu_{ij},$$

with

$$A_{ij} = \langle k, \omega_{\nu+1} \rangle I_d_i d_j + \lambda_j \text{Diag}(I_d_j/2, -I_d_j/2) \otimes I_d_i - \lambda_i I_d_j \otimes \text{Diag}(-I_d_i/2, I_d_i/2).$$

(6.3)

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

17 Recall (4.32), (5.3).
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

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

(6.4)

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

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

and

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

(6.5)

we have

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

(6.6)

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

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

(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

$$\{ \xi \in \mathcal{O}_\nu : \| (A + B^\nu)^{-1} \| > \frac{|k|^\nu}{\gamma_\nu} \gamma^2 \} \subset \{ \xi \in \mathcal{O}_\nu : \|D^{-1}\| > c \frac{|k|^\nu}{\gamma_\nu} \gamma^2 \}.$$  

(6.8)

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

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

(6.9)

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

$$\text{meas} \mathcal{R}^\nu_{kij} = \text{meas} \{ \xi \in \mathcal{O}_\nu : \| (A + B^\nu)^{-1} \| > \frac{|k|^\nu}{\gamma_\nu} \gamma^2 \} \leq \text{meas} \{ \xi \in \mathcal{O}_\nu : | \det D | < c \left( \frac{\gamma_\nu}{|k|^\nu} \right) \gamma^2 \} \leq c \left( \frac{\gamma_\nu}{|k|^\nu} \right) \gamma^2 < c \frac{\gamma}{|k|^\nu}.$$  

(6.10)

This proofs the lemma.
Lemma 6.2 If \( i > c |k| \), then \( R^\nu_{ki} = \emptyset \); If \( \max\{i,j\} > c |k|^{\frac{1}{b-1}} \), \( i \neq j \) for \( b > 1 \) or \( |i - j| > \text{const} \ |k| \) for \( b = 1 \), then \( R^\nu_{kij} = \emptyset \) 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 \( R^\nu_{kij} \). Notice that \( \max\{i,j\} > \text{const} \ |k|^{1-b} \) for \( b > 1 \) or \( |i - j| > \text{const} \ |k| \) for \( b = 1 \) implies

\[
|\lambda_i \pm \lambda_j| = (j^b - i^b)(1 + O(i^{-\delta} + j^{-\delta})) \\
\geq \frac{1}{2} |j^b - i^b|(1 + O(i^{-\delta} + j^{-\delta})) \geq \text{const} |k|. \tag{6.11}
\]

It follows that \( A_{ij} \) defined in (6.3) is invertible and

\[
\| (A_{ij})^{-1} \| < |k|^{-1}.
\]

By Neumann series, we have \( \| (A_{ij} + B^\nu_{ij})^{-1} \| < 2 |k|^{-1} \) for large \( k \) (say \( |k| > K_0 \)), i.e,

\( R^\nu_{kij} = \emptyset \).

Lemma 6.3 For \( b \geq 1 \), we have

\[
\text{meas} \left( \bigcup_{\nu \geq 0} R^\nu \right) = \text{meas} \left( \bigcup_{\nu, |k| > K_\nu, i,j} (R^\nu_k \cup R^\nu_{ki} \cup R^\nu_{kij}) \right) < c \gamma^{\frac{b}{1+b}}.
\]

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

\[
\text{meas} \left( \bigcup_{\nu} \bigcup_{|k| > K_\nu} R^\nu_{kij} \right),
\]

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

\[
\text{meas} \left( \bigcup_{i \neq j} R_{kij} \right) = \text{meas} \left( \bigcup_{i \neq j, i,j < C|k|^{\frac{1}{b-1}}} R^k_{ij} \right) < c \frac{|k|^{\frac{2-b}{b-1}} \gamma}{|k|^{\frac{1}{b-1}}} > c \frac{\gamma}{|k|^{\frac{1}{b-1}-\frac{2}{b-1}}}. \tag{6.12}
\]

For \( i = j \). As in Lemma 6.1, we can find \( Q_1, Q_2 \) so that (2.4) holds with the diagonal elements of \( A_{11} \) being \( k, \omega_{\nu} \geq 2 \lambda_i \) and \( A_{22} = k, \omega_{\nu} > I \). Repeating the arguments in Lemma 6.1, we get (6.3) and

\[
R^\nu_{kii} < \{ \xi : |\det D| < c \left( \frac{\gamma_{\nu}}{|k|^{\frac{1}{\tau-1}}} \right)^2 \}
\]

\[
= \{ \xi : \prod |\langle k, \omega_{\nu} \rangle + O(i^{-\delta})| < c \left( \frac{\gamma_{\nu}}{|k|^{\frac{1}{\tau-1}}} \right)^2 \}
\]

\[
< \{ \xi : |\langle k, \omega_{\nu} \rangle| < c \left( \frac{\gamma}{|k|^{\frac{1}{\tau-1}}} + \frac{1}{i^{\delta}} \right) \} \equiv Q_{ii}^k. \tag{6.13}
\]
Since $Q_{i_0}^k \subset Q_{i_0}^k$ for $i \geq i_0$, using (6.10), we find that
\[
\text{meas} \left( \bigcup_i R_{kii} \right) \leq \sum_{i<i_0} |R_{kii}| + |Q_{i_0}^k| < c \left( \frac{i_0 \gamma}{|k|^\tau - 1} + \frac{1}{i_0^\delta} \right) \]
for any $i_0$. Following Pöschel ([15]), we choose $i_0 = \left( \frac{|k|^\tau - 1}{\gamma} \right)^{1/\delta}$, so that
\[
\text{meas} \left( \bigcup_i R_{kii} \right) < c \left( \frac{\gamma}{|k|^\tau - 1} \right)^{1/\delta}. \tag{6.14}
\]
Let $\tau > \max\{d + 2 + \frac{2}{\varepsilon - 1}, (d + 1)^{\frac{1}{d+1}} + 1\}$. As in (6.12), (6.14), we find
\[
\text{meas} \left( \bigcup_{|k| > K_\nu} \bigcup_{i,j} R_{kij}^{(\gamma)} \right) = \text{meas} \left( \bigcup_{|k| > K_\nu} \bigcup_{i \neq j} R_{kij}^{(\gamma)} \right) + \text{meas} \left( \bigcup_{|k| > K_\nu} \bigcup_j R_{kij}^{(\gamma)} \right) < c K_\nu^{-1} \gamma^{1/\tau},
\]
The quantity $\text{meas}(\bigcup_{|k| > K_\nu} \bigcup_{i,j} R_{kij}^{(\gamma)})$ is then bounded by
\[
\sum_{\nu \geq 1} \text{meas}(\bigcup_{|k| > K_\nu} \bigcup_{i,j} R_{kij}^{(\gamma)}) < c \gamma^{1/\tau} \sum_{\nu \geq 0} K_\nu^{-1} < c \gamma^{1/\tau}, \tag{6.15}
\]
provided $\tau > \max\{d + 2 + \frac{2}{\varepsilon - 1}, (d + 1)^{\frac{1}{d+1}} + 1\}$. 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 $R_{kij}^1 = \emptyset$ for $m > C|k|$. Following the scheme of the above proof, we find
\[
\bigcup_{k,i,j} R_{kij} = \bigcup_{k,i,m} R_{k,i,i+m} = \bigcup_{k,m < C|k|} \bigcup_{i} R_{k,i,i+m} \subset \bigcup_{k,m < C|k|} (\bigcup_{i_0} R_{k,i_0,i_0+m} \cup Q_{k,i_0,i_0+m}). \tag{6.16}
\]
where
\[
Q_{k,i_0,i_0+m} = \{ \xi : |(k, \omega_\nu) + m| < c \left( \frac{\gamma}{|k|^\tau - 1} + \frac{1}{i_0^\delta} \right) \}.
\]
Again, taking $i_0^{1+\delta} = \frac{|k|^\tau - 1}{\gamma}$, we have, for fixed $k$,
\[
\left| \bigcup_{i,j} R_{kij}^k \right| < c \sum_{m < C|k|} \left( \frac{i_0 \gamma}{|k|^\tau - 1} + i_0^{-\delta} \right) < c \left| k \right| \left( \frac{\gamma}{|k|^\tau - 1} \right)^{1/\delta}. \tag{6.17}
\]
As in the case $b > 1$, we have that $\text{meas}(\bigcup_{|k| > K_\nu} \bigcup_{i,j} R_{kij}^{(\gamma)})$ is bounded by $O(\gamma^{1/\tau})$ if $\tau > (d + 1)^{\frac{1}{d+1}} + 1$.

**Remark** In (6.13), $|\det D| = \prod |(k, \omega) + O(i^{-\delta})|$ (guaranteed by the regularity property) is crucial for the proof. But it is not necessary for the periodic solution case, i.e., $d = 1$. Since $R_{k,i+m,i}^1 = \emptyset$ if $i > c > |k| \ll 1$ are sufficiently large.
7 Appendix

Proof of Proposition 7.1 From the hypotheses there follows that the eigenfunctions $\phi_n$ are analytic (respectively, smooth) and bounded with, in particular,

$$\sup \langle |\phi_n'| + |\phi_n''| \rangle \leq \text{const } \mu_n .$$

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

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

one has

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

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

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

$$u_{tt} + Au = \sum \frac{\dot{q}_n}{\sqrt{\lambda}} \phi_n + \frac{q_n}{\sqrt{\lambda_n}} A\phi_n = \sum \left( \int f(u) \phi_n \right) \phi_n = f(u) ,$$

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 $(FG)_{klp} = \sum_{l'} F_{k' l' p} G_{k' l' p'}$, we have that

$$\|FG\|_{D(r, s)} = \sup_D \sum_{klp} |(FG)_{klp}| \|y\|_l \|z\|_p^n e^{k|z|}$$

$$\leq \sup_D \sum_{klp} \sum_{l'} |F_{k-k' l-l' p-p'} G_{k' l' p'}| \|y\|_l \|z\|_p^n e^{k|z|}$$

$$= \|F\|_{D(r, s)} \|G\|_{D'(r, s)}$$

(7.2)

and the proof is finished. ■

Lemma 7.2 (Cauchy inequalities)

$$\|F_{h_i}\|_{D(r-\sigma, s)} \leq \text{const }^{-1} \|F\|_{D(r, s)} ,$$

and

$$\|F_1\|_{D(r, \frac{1}{2}s)} \leq 2 \frac{1}{s^a} \|F\|_{D(r, s)} , \quad \|F_{2_2n}\|_{D(r, \frac{1}{2}s)} \leq 2 \frac{n^a e^{n\rho}}{s} \|F\|_{D(r, s)}$$

25
Let \( \{ \cdot, \cdot \} \) is Poisson bracket of smooth functions
\[
\{ F, G \} = \sum_\frac{\partial F}{\partial \theta_i} \frac{\partial G}{\partial \theta_i} + \sum_{i \in \mathbb{N}} \frac{\partial F}{\partial z_i} i J_d \frac{\partial G}{\partial z_i}, \tag{7.3}
\]
where \( J_d \) are standard symplectic matrix in \( \mathbb{R}^{d_i} \).

**Lemma 7.3** If
\[
\| X_F \|_{r,s} < \epsilon', \| X_G \|_{r,s} < \epsilon'',
\]
then
\[
\| X_{\{F,G\}} \|_{r-\sigma,\eta s} < c \sigma^{-1} \eta^{-2} \epsilon' \epsilon'', \eta \ll 1.
\]

**Proof.** Note that
\[
\frac{d}{dz_n} \{ F, G \} = \langle F_{\theta z_n}, G_I \rangle + \langle F_{\theta}, G_{I z_n} \rangle - \langle F_{I z_n}, G_{\theta} \rangle - \langle F_{I}, G_{\theta z_n} \rangle + \sum_{i \in \mathbb{N}} (\langle F_{z_i z_n}, J_d G_{z_i} \rangle + \langle F_{z_i}, J_d G_{z_i z_n} \rangle) \tag{7.4}
\]
Since
\[
\| \langle F_{\theta z_n}, G_I \rangle \|_{D(r-\sigma,s)} < c \sigma^{-1} \| F_{z_n} \| \cdot \| G_y \|
\]
\[
\| \langle F_{\theta}, G_{I z_n} \rangle \|_{D(r-\sigma,\frac{s}{2})} < c s^{-2} \| F_{z_n} \| \cdot \| G_{z_n} \|
\]
\[
\| \langle F_{I z_n}, G_{\theta} \rangle \|_{D(r,\frac{s}{2})} < c \sigma^{-1} \| F_I \| \cdot \| G_{z_n} \|
\]
\[
\| \langle F_{I}, G_{\theta z_n} \rangle \|_{D(r-\sigma,s)} < c \sigma^{-1} \| F_I \| \cdot \| G_{z_n} \|
\]
\[
\| \langle F_{z_i z_n}, J_d G_{z_i} \rangle \|_{D(r,\frac{s}{2})} < c s^{-1} \| F_{z_n} \| \cdot \| G_{z_i} \| i^s e^{i \rho}
\]
\[
\| \langle F_{z_i}, J_d G_{z_i} \rangle \|_{D(r,\frac{s}{2})} < c s^{-1} \| F_{z_n} \| \cdot \| G_{z_i} \| i^s e^{i \rho} \tag{7.5}
\]
it follows from the definition of the weighted norm (see (2.6)), that
\[
\| X_{\{F,G\}} \|_{r-\sigma,\eta s} < c \| F_{z_n} \| \cdot \| G_{z_n} \|
\]
In particular, if \( \eta \sim \epsilon' \), \( \epsilon', \epsilon'' \sim \epsilon \), we have \( \| X_{\{F,G\}} \|_{r-\sigma,\eta s} \sim \epsilon^\frac{4}{3} \).

**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
\[
\| g \|_{\mathcal{O}} < c \gamma^{-c} |k|^c L
\]

**Proof.** The proof follows directly from the definition of the Whitney’s differentiability.
A Similar lemma for matrices holds:

**Lemma 7.5** Let \( O \) 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 \( \xi \in O \) bounded by \( L \). Then

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

is \( C^m \) Whitney-smooth with

\[
\|F\|_O < c \gamma^{-c}k|c|^c L,
\]

where \( M \) stands for either \( \langle k, \omega \rangle I_{d_i} + A_iJ_{d_i} \) if \( B \) is \((d_i \times d_i)\)-matrix, or \( \langle k, \omega \rangle I_{d_i,d_j} + (A_iJ_{d_i}) \otimes I_{d_j} - I_{d_i} \otimes (J_{d_j}A_j) \) 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. It is known \( \|M\| \) is equivalent to norm \( \max |a_{ij}| \). Since a constant depends only on the space dimension and two fixed norms is irrelevant, we will simply denote \( \|M\| = \max |a_{ij}| \).

**Lemma 7.6** Let \( M \) be a \( N \times N \) non-singular matrix with \( \|M\| < c |k| \), then

\[
\{ \omega : \|M^{-1}\| \geq h \} \subset \{ \omega : |\det M| < c \frac{|k|^{N-1}}{h} \}
\]

**Proof.** Firstly, we note that if \( M \) is a nonsingular \( N \times N \) matrix with elements bounded by \( |m_{ij}| \leq m \), by Cramer rule, the inverse of \( M \) is \( M^{-1} = \frac{1}{|M|}\text{adj}M \). Thus

\[
\|M^{-1}\| < c \frac{m^{N-1}}{|\det M|}
\]

where the constant depends on \( N \). In particular, if \( m = \text{const}|k|, |\text{Det}M| > \frac{|k|^{N-1}}{h} \) then

\[
\|M^{-1}\| < c \cdot h.
\]

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 [18] [20]. 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 \( I_h = \{ u : |g(u)| < h \} \). If for some constant \( d > 0 \), \( |g^{(m)}(u)| \geq d \) for all \( u \in I \), then \( \text{meas} (I_h) \leq ch^{\frac{1}{m}} \) where \( c = 2(2 + 3 + \cdots + m + d^{-1}) \).

For the proof of Lemma 3.1, we need the following

**Lemma 7.8**

\[
\sum_{j \in \mathbb{Z}} e^{-|n-j|\rho + |j|} \leq C e^{\rho|n|}, \quad \sum_{j,n \in \mathbb{Z}} |q_j| e^{-|n-j|\rho + |n|\rho} \leq C|q|_\rho
\]

if \( \rho < r, q \in \mathbb{Z}_\rho \) where \( C \) depends on \( r - \rho \).
Lemma 7.9

$$\sum_{j \in \mathbb{Z}} (1 + |n - j|)^{-K}|j|^a < c |n|^a, \sum_{j,n \in \mathbb{Z}} |g_j|(1 + |n - j|)^{-K}|n|^a \leq C|q|_a$$

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

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

Proof of Lemma 3.1}

Here we give a direct proof. It is clearly enough to consider the case of $f(u)$ being a monomial $u^{N+1}$ for some $N \geq 1$. From (3.10), one can see that the regularity of $G$ implies the regularity of $\hat{G}$. In what follows, we shall give the proof for $G$. Suppose that the potential $V(x)$ is analytic in $|\text{Im} x| < r$ (respectively, belongs to Sobolev space $H^K$) then the eigenfunctions are analytic in $|\text{Im} x| < r$ (respectively, belong to $H^{K+2}$). If we let $\phi_i(x) = \sum a_i^n e^{i(n,x)}$ then (see, e.g., [8])

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

Recall that

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

where

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

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

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

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

it follows that

$$\|G_{q_j}\|_{a+\frac{1}{2},\rho} = \|G_{q_0}\| + \sum_{j \geq 1} |G_{q_j}|j|^{a+\frac{1}{2}}e^{\rho}$$

$$< c \sum_{j,i_1, \cdots, i_N; n_0 + \cdots + n_N = 0} |a_j^{n_0}|j^{a}e^{j}\rho(\prod_{s=0}^N |a_s^{n_s}q_{i_s}|)$$

$$< c \sum_{j,i_1, \cdots, i_N; n_0 + \cdots + n_N = 0} (1 + |j - n_0|)^{-N}|j|^a e^{j\rho-|n_0-j|\rho} \left( \prod_{s=1}^N (1 + |n_s - i_s|)^{-K-2} e^{-|n_s-i_s|^r}|q_{i_s}| \right)$$

$$< c \sum_{i_1, \cdots, i_N; n_0 + \cdots + n_N = 0} |n_0|^a e^{n_0\rho} \left( \prod_{s=1}^N (1 + |n_s - i_s|)^{-K-2} e^{-|n_s-i_s|^r}|q_{i_s}| \right)$$

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

$$< c \sum_{i_1, \cdots, i_N; n_s = 1}^N (\sum_{s=1}^N n_s)^a e^{|\sum_{s=1}^N n_s|\rho} \left( \prod_{s=1}^N (1 + |n_s - i_s|)^{-K-2} e^{-|n_s-i_s|^r}|q_{i_s}| \right).$$
< c \sum_{i_1,\ldots,i_N; n_1,\ldots,n_N} \left( \prod_{s=1}^{N} (1 + |n_s - i_s|)^{-\frac{2}{K}} |n_s|^{a} e^{-|n_s - i_s| + |n_s|} \rho |q_{i_s}| \right) \\
< c \sum_{i_1,\ldots,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) \leq c \|q\|_{a,\rho}^{N}. \tag{7.6}
[12] Melnikov, V.K., On some cases of the conservation of conditionally periodic motions under a small change of the Hamiltonian function, *Soviet Mathematics Doklady*, 6, 1965, 1592-1596.

[13] Moser, J., Convergent series expansions for quasiperiodic motions, *Math. Ann.*, 169(1), 1967, 136-176.

[14] Pöschel, J., Quasi-periodic solutions for a nonlinear wave equation, *Comment. Math. Helvetici*, 71, 1996, 269-296.

[15] Pöschel, J., A KAM-Theorem for some Nonlinear Partial Differential Equations, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 23, 1996, 119-148.

[16] Wayne, C.E., Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, *Comm. Math. Phys.*, 127, 1990, 479-528.

[17] Whitney, H., Analytical extension of differentiable functions defined on closed set *Trans. A. M. S.*, 36, 1934, 63-89.

[18] Xu, J., You, J., Qiu, Q., Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z* 226, 1997, 375-387.

[19] You, J., Perturbation of lower dimensional tori for Hamiltonian systems, to appear in *J. Differential Equations*, 1998.

[20] You, J., A KAM theorem for hyperbolic-type degenerate lower dimensional tori in Hamiltonian systems, *Commun. Math. Phys.*, 192, 1998, 145-168.