A KAM THEOREM FOR HAMILTONIAN NETWORKS WITH LONG RANGED COUPLINGS

JIANSHENG GENG AND YINGFEI YI

Abstract. We consider Hamiltonian networks of long ranged and weakly coupled oscillators with variable frequencies. By deriving an abstract infinite dimensional KAM type of theorem, we show that for any given positive integer N and a fixed, positive measure set $\mathcal{O}$ of N variable frequencies, there is a subset $\mathcal{O}_* \subset \mathcal{O}$ of positive measure such that each $\omega \in \mathcal{O}_*$ corresponds to a small amplitude, quasi-periodic breather (i.e., a solution which is quasi-periodic in time and exponentially localized in space) of the Hamiltonian network with N-frequencies which are slightly deformed from ω.

1. Introduction and Main Result

Associated with the symplectic structure $\sum_n p_n \wedge q_n$, we consider Hamiltonian networks defined by real analytic Hamiltonians of the form

$$H = \sum_{n\in\mathbb{Z}} (\frac{p_n^2}{2} + V_n(q_n)) + W(\{q_n\}), \tag{1.1}$$

where V_n's are the on-site potentials satisfying $V_n(0) = V_n'(0) = 0$ and $V_n''(0) \equiv \frac{\beta_n^2}{2}$, $\beta_n > 0$, and W is a coupling potential. Hamiltonian networks have been used in solid state physics in describing the vibration of particles (atoms) in a lattice (see [10, 11, 20]) and also used to model DNA chains (see [10, 14, 34]). They also arise naturally as spatial discretization of Hamiltonian PDEs such as nonlinear wave equations.

Among the solutions of a Hamiltonian network, of particular physical interests are the so-called *breathers* or *quasi-periodic breathers*, which are self-localized, time periodic or quasi-periodic, solutions whose amplitudes decay at least exponentially as $|n| \to \infty$. Breathers or quasi-periodic breathers are often referred to as dynamical solitons or intrinsic localized modes in physics and they have been largely found via numerics in many physical models (see [10, 29] and references therein). The existence of breathers in Hamiltonian networks associated with Hamiltonians (1.1) was rigorously analyzed when $\beta_n \equiv \beta$ by Aubry [1, 2], Mackay–Aubry [25] for the inter-particle, nearest neighbor coupling potential

$$W(\{q_n\}) = \sum_n (q_{n+1} - q_n)^2$$

and by Bambusi [3] for the long–range coupling potential

$$W(\{q_n\}) = \sum_{n\neq m} \frac{1}{|n-m|^\alpha}(q_n - q_m)^2, \quad \alpha > 1.$$

Like in [25], breathers in the nearest neighbor coupling case can be studied near a fixed periodic orbit of the uncoupled Hamiltonian by certain continuation or perturbation arguments, provided that the couplings are "weak", and, no small divisors need to be considered in such perturbation problems. These perturbation techniques are also applicable in finding quasi-periodic breathers with

1991 *Mathematics Subject Classification.* Primary 37K60, 37K55.

Key words and phrases. Coupled oscillators, Hamiltonian networks, long ranged coupling, KAM theory, quasi-periodic breathers.

The second author is Partially supported by NSF grant DMS0204119.

two or three frequencies for certain models with symmetries (see Bambusi–Vella [4], Johansson–Aubry [19]). Using a modified KAM technique, the existence of quasi-periodic breathers with any finite number of frequencies was recently shown by Yuan [33] for the higher order, nearest-neighbor coupling potential

$$W(\{q_n\}) = \sum_n (q_{n+1} - q_n)^3. \tag{1.2}$$

Almost periodic breathers with infinitely many frequencies have also been investigated. Associated with the potential (1.2), Fröhlich-Spencer-Wayne [15] considered the case when the frequencies are non-negative random variables with smooth distribution of fast decay at infinity and showed that there is a set $\Omega \subset \mathbb{R}_+^\infty$ with positive probability measure such that each $\omega \in \Omega$ corresponds to an almost periodic breather with infinite many frequencies (see also Pöschel [28] for more general spatial structures).

In this paper, we will study the existence of quasi-periodic breathers for the Hamiltonian (1.1) with the following higher order, long-ranged coupling potential

$$W(\{q_n\}) = \frac{1}{3} \sum_{n \neq m} e^{-|n-m|^\alpha} (q_n - q_m)^3, \ \alpha \geq 1, \tag{1.3}$$

or equivalently the Hamiltonian network

$$\frac{d^2 q_n}{dt^2} + V_n'(q_n) = -\sum_{m \in \mathbb{Z}} e^{-|n-m|^\alpha} (q_n - q_m)^2. \tag{1.4}$$

For a given integer $N > 1$, we specify N integers $\{i_1, \cdots, i_N\}$ and let $\mathbb{Z}_1 = \mathbb{Z} \setminus \{i_1, \cdots, i_N\}$. We treat $\omega = (\beta_{i_1}, \cdots, \beta_{i_N})$ as parameters in a bounded closed region $\mathcal{O}$ in R_+^N and assume the following spectral gap condition:

SG) There exist $1 \leq d < \infty$ and $\gamma > 0$ such that

$$\{\beta_n\}_{n \in \mathbb{Z}_1} = \cup_{l=1}^\infty \Lambda_l$$

where Λ_l, $l = 1, 2, \cdots$, are sets satisfying

$$\#(\Lambda_l) \leq d, \qquad \text{for all } l,$$

and

$$|\beta_n - \beta_m| \geq \gamma, \qquad \text{for all } \beta_n \in \Lambda_l, \beta_m \in \Lambda_j, l \neq j.$$

We will show the following result.

Theorem A. *Assume* SG) *with γ sufficiently small. Then there exists a Cantor set $\mathcal{O}_\gamma \subset \mathcal{O}$, with* $\text{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) = O(\gamma)$*, such that for any $\omega \in \mathcal{O}_\gamma$, the Hamiltonian network* (1.4) *associated with ω admits a small amplitude, linearly stable, quasi-periodic breather $q(t) = (\{q_n(t)\})$ of N-frequency ω_* which is close to ω, and moreover, $|q_n| \sim e^{-|n|}$.*

The condition **SG**) clearly holds when $\beta_n = |n|$, $n \in \mathbb{Z}_1$. Comparing with the cases of nonlinear wave equations [12, 17, 26, 27] in which $\beta_n \sim |n|$, $n \in \mathbb{Z}$, the validity of Theorem A crucially depends on the coupling potential or perturbation (1.3) which admits a weaker regularity but a higher order perturbation.

For the case of nearest-neighboring coupled Hamiltonian networks, breathers were shown to be super-exponentially localized in space ([33]). This is due to the at most linear growth of the normal components in the normal form associated with the short-ranged coupling potential (1.2). Our result only asserts the exponential localization of quasi-periodic breathers due to the exponential growth of the normal components in the normal form associated with the exponentially weighted, long-ranged coupling potential (1.3). If the long-ranged coupling potential

$$W(\{q_n\}) = \frac{1}{3} \sum_{n \neq m} \frac{1}{|n-m|^\alpha} (q_n - q_m)^3, \ \alpha > 1$$

is considered instead, then the normal components in the associated normal form will have a super-exponential growth, and our method will equally applicable to yield quasi-periodic breathers which are localized like $\frac{1}{|n|^\alpha}$ in space.

Theorem A will be proved by using KAM (Kolmogorov-Arnold-Moser) method. In fact, we will present an abstract infinite dimensional KAM type of theorem from which Theorem A will follow. Such an infinite dimensional KAM theorem differs significantly from those for Hamiltonian PDEs like nonlinear Schödinger, wave, beam, and KdV equations studied by many authors using either KAM or CWB (Craig-Wayne-Bourgain) method (see [5, 6, 7, 8, 9, 12, 13, 16, 17, 18, 21, 22, 23, 26, 27, 30] and references therein). This is mainly due to the fact that, when the normal frequencies of a Hamiltonian network have linear growth, the perturbation (1.3) admits weaker regularity than those of Hamiltonian PDEs under KAM or Newton iterations.

Similar to the short-ranged coupling cases considered in [25, 33], it is also important to study the existence of quasi-periodic breathers for Hamiltonian networks with long-ranged coupling potentials and constant frequencies $\beta_n \equiv \beta$, $n \in \mathbb{Z}$, i.e., those formed by weakly coupled identical oscillators. As the KAM iteration mechanism and measure estimates for the constant frequency case significantly differ from the variable ones to be studied in this paper, we will consider the constant frequency case in a separate work.

The paper is organized as follows. In Section 2 we state an abstract infinite dimensional KAM theorem and prove Theorem A as a corollary. Sections 3 and 4 are devoted to the proof of the abstract KAM theorem. More precisely, in Section 3, we give detailed construction of the KAM iteration for one KAM step. We complete the proof of the abstract infinite dimensional KAM theorem in Section 4 by showing an iteration lemma, convergence, and measure estimate. Some technical lemmas are provided in the Appendix.

2. An Abstract KAM Theorem

In this section, we will formulate an abstract KAM theorem which can be applied to the Hamiltonian networks of long-ranged and weakly coupled oscillators with variable frequencies. Theorem A will be proved by using the abstract KAM theorem and normal form reductions.

2.1. **The abstract theorem.**

We begin with some notations. Let integers $N > 1$, $d \geq 1$, and real numbers $r, s > 0$ be given. We consider the complex neighborhood $D(r,s)$ of $\mathbb{T}^N \times \{0\} \times \{0\} \times \{0\} \subset \mathbb{T}^N \times R^N \times \ell^1 \times \ell^1$ defined by

$$D(r,s) = \{(\theta, I, w, \bar{w}) : |\mathrm{Im}\theta| < r,\ |I| < s^2,\ \|w\| < s,\ \|\bar{w}\| < s\},$$

where $|\cdot|$ denote the sup-norm of complex vectors and $\|\cdot\|$ denote the ℓ^1 norm. Also let $\mathcal{O}$ be a positive (Lebesgue) measure set in $\mathbb{R}^N$.

Let $F(\theta, I, w, \bar{w})$ be a real analytic function on $D(r,s)$ which depends on a parameter $\xi \in \mathcal{O}$, C^{d^2}-Whitney smoothly (i.e., C^{d^2} in the sense of Whitney). We expand F into the Taylor-Fourier series with respect to $\theta, I, w, \bar{w}$:

$$F(\theta, I, w, \bar{w}) = \sum_{\alpha,\beta} F_{\alpha\beta} w^\alpha \bar{w}^\beta,$$

where $\alpha \equiv (\cdots, \alpha_n, \cdots)$, $\beta \equiv (\cdots, \beta_n, \cdots)$, $\alpha_n, \beta_n \in \mathbb{N}$, are multi-indices with finitely many non-vanishing components, and

$$F_{\alpha\beta} = \sum_{k \in \mathbb{Z}^N, l \in \mathbb{N}^N} F_{kl\alpha\beta}(\xi) I^l e^{\mathrm{i}\langle k, \theta\rangle}.$$

We define the weighted norm of F by

$$\|F\|_{D(r,s),\mathcal{O}} \equiv \sup_{\substack{\|w\|<s \\ \|\bar{w}\|<s}} \sum_{\alpha,\beta} \|F_{\alpha\beta}\|\, |w^\alpha| |\bar{w}^\beta|,$$

where

$$\|F_{\alpha\beta}\| \equiv \sum_{k,l} |F_{kl\alpha\beta}|_{\mathcal{O}} s^{2|l|} e^{|k|r}, \quad |F_{kl\alpha\beta}|_{\mathcal{O}} \equiv \sup_{\xi\in\mathcal{O}} \{\max_{m\le d^2} |\partial_\xi^m F_{kl\alpha\beta}|\}.$$

In the above and also for the rest of the paper, derivatives in $\xi \in \mathcal{O}$ are taken in the sense of Whitney.

For a vector-valued function $G : D(r,s) \times \mathcal{O} \to \mathbb{C}^n$, $n < \infty$, we simply define its weighed norm by

$$\|G\|_{D(r,s),\mathcal{O}} \equiv \sum_{i=1}^{n} \|G_i\|_{D(r,s),\mathcal{O}}.$$

For the Hamiltonian vector field

$$X_F = (F_I, -F_\theta, \{\mathrm{i}F_{w_n}\}, \{-\mathrm{i}F_{\bar{w}_n}\})$$

associated with a Hamiltonian function F on $D(r,s) \times \mathcal{O}$, we define its weighted norm by

$$\|X_F\|_{D(r,s),\mathcal{O}} \equiv \|F_I\|_{D(r,s),\mathcal{O}} + \frac{1}{s^2}\|F_\theta\|_{D(r,s),\mathcal{O}} + \frac{1}{s}(\sum_n \|F_{w_n}\|_{D(r,s),\mathcal{O}} + \sum_n \|F_{\bar{w}_n}\|_{D(r,s),\mathcal{O}}).$$

Associated with the symplectic structure $dI \wedge d\theta + \mathrm{i}\sum_{n\in\mathbb{Z}} dw_n \wedge d\bar{w}_n$, we consider the following family of real analytic, parameterized Hamiltonians

$$\begin{aligned} H &= N + P, \\ N &= \langle \omega(\xi), I\rangle + \sum_{n\in\mathbb{Z}} \Omega_n w_n \bar{w}_n, \\ P &= P(\theta, I, w, \bar{w}, \xi), \end{aligned} \tag{2.1}$$

where $(I, \theta, w, \bar{w}) \in D(r,s)$, $\xi \in \mathcal{O}$, Ω_n's are positive and independent of ξ, and all ξ-dependence are of class C^{d^2} in the sense of Whitney.

It is clear that when $P = 0$, the unperturbed Hamiltonians N are completely integrable, admitting a family of quasi-periodic solutions $(\theta + \omega t, 0, 0, 0)$ corresponding to invariant N-tori in the phase space.

To study the persistence of some of these N-tori, we need the following assumptions on $\omega(\xi)$, Ω_n and the perturbation P:

$(A1)$ *Non-degeneracy of tangential frequencies*: There is a constant $\delta > 0$ such that

$$|\det(\frac{\partial\omega}{\partial\xi})| \ge \delta.$$

$(A2)$ *Gap conditions of normal frequencies*: There exist sufficiently small $\gamma > 0$ and sets Λ_l, $l = 1, 2, \cdots$, such that

$$\begin{aligned} &\{\Omega_n\}_{n\in\mathbb{Z}} = \cup_{l=1}^{\infty} \Lambda_l, \\ &\#(\Lambda_l) \le d, \qquad \text{for all } l, \\ &|\Omega_n - \Omega_m| \ge \gamma, \qquad \text{if } \Omega_n \in \Lambda_l, \Omega_m \in \Lambda_j, l \ne j. \end{aligned}$$

(*A*3) *Decay property of the perturbation*: $P = \breve{P} + \acute{P} + \grave{P}$, where $\breve{P} = \breve{P}(\theta, I, w, \bar{w}, \xi)$, $\acute{P} = \acute{P}(\theta, I, w, \bar{w}, \xi)$, $\grave{P} = \grave{P}(\theta, I, w, \bar{w}, \xi)$ are such that

$$\breve{P} = \breve{P}(\theta, I, 0, 0, \xi) + \sum_{\substack{n\in\mathbb{Z}\\ \alpha_n+\beta_n\geq 1}} \breve{P}_n(\theta, I, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n}, \quad \|\breve{P}_n(\theta, I, \xi)\| \leq e^{-|n|}; \tag{2.2}$$

$$\acute{P} = \sum_{\substack{n,m\in\mathbb{Z}, n\neq m\\ \alpha_n+\beta_n, \alpha_m+\beta_m\geq 1\\ \alpha_n+\beta_n+\alpha_m+\beta_m\geq 3}} \acute{P}_{nm}(\xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} w_m^{\alpha_m} \bar{w}_m^{\beta_m}, \quad \|\acute{P}_{nm}(\xi)\| \leq e^{-|n-m|}; \tag{2.3}$$

$$\grave{P} = \sum_{n\in\mathbb{Z}} O(|w_n|^3). \tag{2.4}$$

Our abstract KAM theorem states as the following.

Theorem B. *Consider the Hamiltonian* (2.1) *and assume* (A1)-(A3). *For a fixed* $\gamma > 0$ *sufficiently small, there exists a positive constant* $\varepsilon = \varepsilon(\mathcal{O}, d, \delta, N, \gamma, r, s)$ *such that if* $\|X_P\|_{D(r,s),\mathcal{O}} < \varepsilon$, *then the following holds. There exist Cantor sets* $\mathcal{O}_\gamma \subset \mathcal{O}$ *with* $\mathrm{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) = O(\gamma)$ *and maps*

$$\Psi : \mathbb{T}^N \times \mathcal{O}_\gamma \to D(r, s), \qquad \tilde{\omega} : \mathcal{O}_\gamma \to \mathbb{R}^N,$$

which are real analytic in θ *and* C^{d^2}*-Whitney smooth in* ξ *with* $\|\Psi - \Psi_0\|_{D(\frac{r}{2},0),\mathcal{O}_\gamma} \to 0$ *and* $|\tilde{\omega} - \omega|_{\mathcal{O}_\gamma} \to 0$ *as* $\gamma \to 0$, *where* Ψ_0 *is the trivial embedding:* $\mathbb{T}^N \times \mathcal{O} \to \mathbb{T}^N \times \{0, 0, 0\}$, *such that each* $\xi \in \mathcal{O}_\gamma$ *and* $\theta \in \mathbb{T}^N$ *corresponds to a linearly stable,* N*-frequency quasi-periodic solution* $\Psi(\theta + \tilde{\omega}(\xi)t, \xi) = (\theta + \tilde{\omega}(\xi)t, I(t), \{w_n(t)\})$ *of the Hamiltonian* (2.1). *Moreover,* $|w_n| \sim e^{-|n|}$.

Since our perturbation has a weaker regularity, the frequencies of these invariant tori are in general non-resonant instead of Diophantine.

Comparing with results on quasi-periodic solutions for Hamiltonian PDEs (see e.g. [5, 6, 7, 8, 9, 12, 13, 16, 17, 18, 21, 22, 23, 26, 27, 30]), the above theorem relaxes the linear or super-linear growth conditions on the normal frequencies Ω_n. Indeed, it is easy to see that the gap condition (A2) above is weaker than the linear or sup-linear growth conditions on the normal frequencies. The assumption (A3) is new but natural for networks of long ranged and weakly coupled oscillators.

It is not clear whether a Lyapunov center theorem is possible for a Hamiltonian network whose normal frequencies satisfy the gap condition (A2). At least, the above theorem assert a quasi-periodic type of Lyapunov center result in the sense of measure.

2.2. **Proof of Theorem A.** Recall that the Hamiltonian networks of long ranged, weakly coupled oscillators considered in Theorem A is described by the Hamiltonian

$$H = \sum_{n\in\mathbb{Z}} [\frac{p_n^2}{2} + V_n(q_n)] + \frac{1}{3} \sum_{n\neq m} e^{-|n-m|^\alpha} (q_n - q_m)^3, \quad \alpha \geq 1,$$

which, in terms of the Taylor expansion at $q = 0$, can be equivalently rewritten as

$$H = \sum_{n\in\mathbb{Z}} [\frac{p_n^2}{2} + \frac{\beta_n^2 q_n^2}{2}] + \frac{1}{3} \sum_{n\neq m} e^{-|n-m|^\alpha} (q_n - q_m)^3 + \sum_{n\in\mathbb{Z}} O(|q_n|^3).$$

Let $\varepsilon > 0$ be sufficiently small. With the re-scalings $p_n, q_n \to \varepsilon p_n, \varepsilon q_n$, the re-scaled Hamiltonian reads

$$\varepsilon^{-2} H(\varepsilon p, \varepsilon q) = \sum_{n\in\mathbb{Z}} [\frac{p_n^2}{2} + \frac{\beta_n^2 q_n^2}{2}] + \frac{\varepsilon}{3} \sum_{n\neq m} e^{-|n-m|^\alpha} (q_n - q_m)^3 + \varepsilon \sum_{n\in\mathbb{Z}} O(|q_n|^3).$$

Let N, $\{i_1, \cdots, i_N\}$, and $\mathbb{Z}_1 = \mathbb{Z} \setminus \{i_1, \cdots, i_N\}$ be as in Theorem A. For a given value $a = (a_1, \cdots, a_N) \in \mathbb{R}_+^N$, we introduce the standard action-angle-normal variables $(I, \theta, w, \bar{w}) =$

$(I,\theta,\{w_n\}_{n\in\mathbb{Z}_1},\{\bar{w}_n\}_{n\in\mathbb{Z}_1})\in\mathbb{R}^N\times\mathbb{T}^N\times\ell^1$, i.e.,

$$p_{i_j}=\sqrt{\beta_{i_j}}\sqrt{I_j+a_j}\cos\theta_j,\ \ q_{i_j}=\sqrt{\frac{1}{\beta_{i_j}}}\sqrt{I_j+a_j}\sin\theta_j,\ 1\le j\le N,$$

$$p_n=\frac{\sqrt{\beta_n}(w_n+\bar{w}_n)}{\sqrt{2}},\ \ q_n=\frac{w_n-\bar{w}_n}{\mathrm{i}\sqrt{2\beta_n}},\ n\in\mathbb{Z}_1.$$

Let $\xi=(\xi_1,\cdots,\xi_N)=(\beta_{i_1},\cdots,\beta_{i_N})$. Then in terms of the action-angle-normal variables the above Hamiltonian becomes

$$H=N+P=\langle\omega(\xi),I\rangle+\sum_{n\in\mathbb{Z}_1}\Omega_n w_n\bar{w}_n+P(\theta,I,w,\bar{w},\xi), \tag{2.5}$$

where

$$\omega(\xi)=(\omega_1(\xi),\cdots,\omega_N(\xi))=(\xi_1,\cdots,\xi_N),$$
$$\Omega_n=\beta_n,\ n\in\mathbb{Z}_1,$$

and

$$P=\breve{P}+\acute{P}+\grave{P}$$

satisfying

$$\begin{aligned}
\breve{P} &= \breve{P}(\theta,I,0,0,\xi)+\sum_{\substack{n\in\mathbb{Z}_1\\ \alpha_n+\beta_n\ge1}}\breve{P}_n(\theta,I,\xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}, \quad \|\breve{P}_n(\theta,I,\xi)\|\le e^{-|n|^{\alpha}}\le e^{-|n|};\\
\acute{P} &= \sum_{\substack{n,m\in\mathbb{Z}_1,n\neq m\\ \alpha_n+\beta_n,\alpha_m+\beta_m\ge1\\ \alpha_n+\beta_n+\alpha_m+\beta_m\ge3}}\acute{P}_{nm}(\xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}w_m^{\alpha_m}\bar{w}_m^{\beta_m}, \quad \|\acute{P}_{nm}(\xi)\|\le e^{-|n-m|^{\alpha}}\le e^{-|n-m|};\\
\grave{P} &= \sum_{n\in\mathbb{Z}_1}O(|w_n|^3).
\end{aligned}$$

It is also easy to see that we can choose appropriate $r,s>0$ such that $\|X_P\|_{D(r,s),\mathcal{O}}<\varepsilon$. Hence the Hamiltonian (2.5) satisfies all conditions of Theorem B, from which Theorem A follows.

3. KAM Step

In what follows, we will perform KAM iterations to (2.1) which involves infinite many successive steps, called KAM steps, of iterations, to eliminate lower order θ-dependent terms in P. Each KAM step will make the perturbation smaller than the previous one at a cost of excluding a small measure set of parameters. At the end, the KAM iterations will be shown to converge and the measure of the total excluding set will be shown to be small.

To begin with the KAM iteration, we set $\Omega_n^0=\Omega_n$, $n\in\mathbb{Z}_1$, $r_0=r$, $s_0=s$.

3.1. **Normal form.** We first convert the Hamiltonian (2.1) into a more convenient form in order to perform the KAM iteration. Let $\varepsilon_*\sim\varepsilon^{\frac{5}{4}}$ and choose a K_0 such that $K_0\sim|\ln\varepsilon_*|$. According to the forms of (2.3), (2.4) in the Assumption (A3), we can make s_0 smaller if necessary such that

$$\|X_{\acute{P}+\grave{P}}\|_{D(r_0,s_0)\times\mathcal{O}}\le\varepsilon_*.$$

We now treat the term $\breve{P}$. According to the form of (2.2) and the definition of the norm, we have

$$\begin{aligned}
\breve{P}&=\breve{P}(\theta,I,0,0,\xi)+\sum_{\alpha_n+\beta_n\ge1}\breve{P}_n(\theta,I,\xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}\\
&=\sum_{k,l}\breve{P}_{kl}I^l e^{\mathrm{i}\langle k,\theta\rangle}+\sum_{\substack{n,k,l\\ \alpha_n+\beta_n\ge1}}\breve{P}_n^{kl\alpha_n\beta_n}I^l e^{\mathrm{i}\langle k,\theta\rangle}\,w_n^{\alpha_n}\bar{w}_n^{\beta_n},
\end{aligned}$$

where

$$\|\breve{P}_{kl}\|\le e^{-|k|r_0},\quad \|\breve{P}_n^{kl\alpha_n\beta_n}\|\le e^{-|k|r_0}e^{-|n|}.$$

Thus we can make r_0, s_0 smaller such that

$$\|X_{\sum_{|k|>K_0,|l|\le 1}\breve{P}_{kl}I^l e^{\mathrm{i}\langle k,\theta\rangle}+\sum_{\substack{|n|>K_0 \text{ or } |k|>K_0\\ \alpha_n+\beta_n\ge 1}}\breve{P}_n^{k\alpha_n\beta_n}e^{\mathrm{i}\langle k,\theta\rangle}\, w_n^{\alpha_n}\bar{w}_n^{\beta_n}+O(|I|^2)+O(|w|^3)}\|\le \varepsilon_*.$$

Let

$$R=\sum_{|k|\le K_0,|l|\le 1}\breve{P}_{kl}I^l e^{\mathrm{i}\langle k,\theta\rangle}\;+\sum_{\substack{|n|\le K_0,|k|\le K_0\\ 1\le\alpha_n+\beta_n\le 2}}\breve{P}_n^{k\alpha_n\beta_n}e^{\mathrm{i}\langle k,\theta\rangle}\; w_n^{\alpha_n}\bar{w}_n^{\beta_n}.$$

We first construct a symplectic transformation $\Phi_*=\Phi^1_{F_*}$ defined as the time-1 map of the Hamiltonian flow associated to a Hamiltonian F_* of the form

$$\begin{aligned}F_* &= \sum_{0<|k|\le K_0,|l|\le 1}F_{kl00}I^l e^{\mathrm{i}\langle k,\theta\rangle}\;+\sum_{|k|\le K_0,|n|\le K_0}(F_n^{k10}w_n+F_n^{k01}\bar{w}_n)e^{\mathrm{i}\langle k,\theta\rangle}\\ &+\sum_{|k|\le K_0,|n|\le K_0}(F_{nn}^{k20}w_nw_n+F_{nn}^{k02}\bar{w}_n\bar{w}_n)e^{\mathrm{i}\langle k,\theta\rangle}\\ &+\sum_{0<|k|\le K_0,|n|\le K_0}F_{nn}^{k11}w_n\bar{w}_n e^{\mathrm{i}\langle k,\theta\rangle}\end{aligned}$$

such that all non-resonance terms $\breve{P}_{kl}I^l e^{\mathrm{i}\langle k,\theta\rangle}$, $0<|k|\le K_0,|l|\le 1$, $\breve{P}_n^{k\alpha_n\beta_n}e^{\mathrm{i}\langle k,\theta\rangle}\; w_n^{\alpha_n}\bar{w}_n^{\beta_n}$, $0<|k|\le K_0$, $\alpha_n+\beta_n\le 2$ are eliminated, and terms $\breve{P}_{0l}I^l$, $|l|\le 1$, $\breve{P}_{nn}^{011}w_n\bar{w}_n$, $|n|\le K_0$ are added to the normal form part of the new Hamiltonian. More precisely, let the coefficients of F_* satisfy the homological equation

$$\{N,F_*\}+R=\sum_{|l|\le 1}\breve{P}_{0l}I^l+\sum_{|n|\le K_0}\breve{P}_{nn}^{011}w_n\bar{w}_n. \tag{3.1}$$

It is easy to see that the homological equation (3.1) is solvable on the parameter set

$$\mathcal{O}_*=\left\{\xi\in\mathcal{O}:\begin{array}{ll}|\langle k,\omega\rangle|\ge\frac{\gamma}{K_0^\tau}, & 0<|k|\le K_0\\ |\langle k,\omega\rangle+\Omega_n|\ge\frac{\gamma}{K_0^\tau}, & |k|\le K_0,|n|\le K_0\\ |\langle k,\omega\rangle+2\Omega_n|\ge\frac{\gamma}{K_0^\tau}, & |k|\le K_0,|n|\le K_0\end{array}\right\}.$$

Hence we obtain the transformation Φ_* such that

$$\begin{aligned}H_* &= H\circ\Phi_*=N_*+\breve{P}_*+\acute{P}+\grave{P},\\ N_* &= e_*+\langle\omega_*(\xi),I\rangle+\sum_{|n|\le K_0}\Omega_n^*w_n\bar{w}_n+\sum_{|n|>K_0}\Omega_nw_n\bar{w}_n,\end{aligned}$$

where $\omega*=\omega+\breve{P}_{0l}(|l|=1)$, $\Omega_n^*=\Omega_n+\breve{P}_{nn}^{011}$, and

$$\breve{P}_*=\breve{P}^*(\theta,I,w_{n(|n|\le K_0)},\bar{w}_{n(|n|\le K_0)},\xi)+\sum_{\substack{|n|>K_0\\ \alpha_n+\beta_n\ge 1}}\breve{P}_n^*(\theta,I,w_{m(|m|\le K_0)},\bar{w}_{m(|m|\le K_0)},\xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}$$

satisfies

$$\|\breve{P}_n^*(\theta,I,w_{m(|m|\le K_0)},\bar{w}_{m(|m|\le K_0)},\xi)\|\le e^{-(|n|-K_0)}.$$

In the above, the first and the second term of $\breve{P}_*$ come from $P\circ\Phi_*$ and $\breve{P}\circ\Phi_*+\acute{P}\circ\Phi_*$ respectively. The decay property of $\breve{P}_n^*$ follows from the fact that Φ_* depends only on I, θ and w_m, $\bar{w}_m$ for $|m|\le K_0$.

Next, write

$$\begin{aligned}
&\sum_{\substack{|n|>K_0\\ \alpha_n+\beta_n\geq 1}} \breve{P}^*_n(\theta, I, w_{m(|m|\leq K_0)}, \bar{w}_{m(|m|\leq K_0)}, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n}\\
=& \sum_{\substack{|n|>5K_0\\ \alpha_n+\beta_n\geq 1}} \breve{P}^*_n(\theta, I, w_{m(|m|\leq K_0)}, \bar{w}_{m(|m|\leq K_0)}, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n}\\
+& \sum_{\substack{K_0<|n|\leq 5K_0\\ \alpha_n+\beta_n\geq 1}} \breve{P}^*_n(\theta, I, w_{m(|m|\leq K_0)}, \bar{w}_{m(|m|\leq K_0)}, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n}.
\end{aligned}$$

On the domain $D(r_0, s_0) \times \mathcal{O}_*$, it is easy to see that the norm of the vector field associated with the first term above is bounded by ε_*^2. To handle the second term above, we note that due to the gap condition (A2) of normal frequencies, when Ω_n, Ω_m belong to the same Λ_l, terms of the form $\breve{P}^*_{mn} w_m\bar{w}_n + \breve{P}^*_{nm} w_n\bar{w}_m$ are not able to be eliminated via solving a homological equation. Hence they need to be included in the normal form part of the Hamiltonian. More precisely, according to the assumption (A2), we let L_0 be the smallest positive integer such that $\{\Omega_n\}_{|n|\leq 5K_0}$ lie in $\Lambda_1, \cdots, \Lambda_{L_0}$. Let $\tilde{K}_0 = 5K_0$. Clearly $L_0 \leq \tilde{K}_0$.

Let

$$R_* = \sum_{|k|\leq \tilde{K}_0, |l|\leq 1} \breve{P}^*_{kl} I^l e^{\mathrm{i}\langle k,\theta\rangle} + \sum_{\substack{|m|\leq K_0, |n|\leq \tilde{K}_0, |k|\leq K_0\\ 1\leq \alpha_m+\beta_n\leq 2}} \breve{P}^{*k\alpha_m\beta_n}_{mn} e^{\mathrm{i}\langle k,\theta\rangle} \,(w_m^{\alpha_m}\bar{w}_n^{\beta_n} + \bar{w}_m^{\alpha_m} w_n^{\beta_n})$$

and let

$$\begin{aligned}
F_{**} =& \sum_{0<|k|\leq \tilde{K}_0, |l|\leq 1} f_{kl00} e^{\mathrm{i}\langle k,\theta\rangle}\, I^l + \sum_{|k|\leq \tilde{K}_0, |n|\leq \tilde{K}_0} (f_n^{k10} w_n + f_n^{k01}\bar{w}_n) e^{\mathrm{i}\langle k,\theta\rangle}\\
&+ \sum_{|k|\leq \tilde{K}_0, |m|\leq K_0, |n|\leq \tilde{K}_0} (f_{nm}^{k20} w_n w_m + f_{nm}^{k11} w_n\bar{w}_m + f_{nm}^{k02}\bar{w}_n\bar{w}_m) e^{\mathrm{i}\langle k,\theta\rangle}\\
&+ \sum_{|k|\leq \tilde{K}_0, K_0<|n|\leq \tilde{K}_0} (f_{nn}^{k20} w_n w_n + f_{nn}^{k02}\bar{w}_n\bar{w}_n) e^{\mathrm{i}\langle k,\theta\rangle}\\
&+ \sum_{0<|k|\leq \tilde{K}_0, K_0<|n|\leq \tilde{K}_0} f_{nn}^{k11} w_n\bar{w}_n e^{\mathrm{i}\langle k,\theta\rangle}
\end{aligned}$$

satisfy the homological equation

$$\{N_*, F_{**}\} + R_* = \sum_{|l|\leq 1} \breve{P}^*_{0l} I^l + \sum_{|m|\leq K_0, |n|\leq \tilde{K}_0} \breve{P}^{*011}_{nm} w_n\bar{w}_m + \sum_{K_0<|n|\leq \tilde{K}_0} \breve{P}^{*011}_{nn} w_n\bar{w}_n. \tag{3.2}$$

It is clear that the equation (3.2) is solvable on the domain

$$\mathcal{O}_0 = \left\{ \xi \in \mathcal{O}_* : \begin{array}{ll}
|\langle k,\omega_*\rangle| \geq \frac{\gamma}{\tilde{K}_0^\tau}, & 0<|k|\leq \tilde{K}_0\\
|\langle k,\omega_*\rangle + \Omega_n^*| \geq \frac{\gamma}{\tilde{K}_0^\tau}, & |k|\leq \tilde{K}_0, |n|\leq K_0\\
|\langle k,\omega_*\rangle + \Omega_n| \geq \frac{\gamma}{\tilde{K}_0^\tau}, & |k|\leq \tilde{K}_0, K_0<|n|\leq \tilde{K}_0\\
|\langle k,\omega_*\rangle + \Omega_m^* \pm \Omega_n^*| \geq \frac{\gamma}{\tilde{K}_0^\tau}, & |k|\leq \tilde{K}_0, |m|, |n|\leq K_0\\
|\langle k,\omega_*\rangle + \Omega_m^* \pm \Omega_n| \geq \frac{\gamma}{\tilde{K}_0^\tau}, & |k|\leq \tilde{K}_0, |m|\leq K_0, K_0<|n|\leq \tilde{K}_0\\
|\langle k,\omega_*\rangle + 2\Omega_n| \geq \frac{\gamma}{\tilde{K}_0^\tau}, & |k|\leq \tilde{K}_0, K_0<|n|\leq \tilde{K}_0
\end{array} \right\}.$$

Consider the symplectic transformation $\Phi_{**} = \Phi^1_{F_{**}}$. Then

$$\begin{aligned}
H_0 &= H_*\circ\Phi_{**} = N_0 + P^0,\\
N_0 &= e_0 + \langle \omega_0(\xi), I\rangle + \sum_{l=1}^{L_0} \langle A_l^0 z_l^0, \bar{z}_l^0\rangle + \sum_{|n|>\tilde{K}_0} \Omega_n^0 w_n\bar{w}_n,
\end{aligned}$$

where

$$\begin{aligned}
e_0 &= e_* + \breve{P}^*_{00},\\
\omega_0 &= \omega_* + \breve{P}^*_{0l}(|l|=1),\\
\sum_{l=1}^{L_0}\langle A^0_l z^0_l, \bar{z}^0_l\rangle &= [\sum_{|n|\le \tilde{K}_0}\Omega^{**}_n w_n\bar{w}_n + \sum_{|m|\le K_0,|n|\le\tilde{K}_0}(\breve{P}^{*011}_{mn}w_m\bar{w}_n + \breve{P}^{*011}_{nm}w_n\bar{w}_m)],\\
\Omega^{**}_n &= \Omega^*_n + \breve{P}^{*011}_{nn},\\
P^0 &= \breve{P}_0 + \acute{P}_0 + \grave{P}_0,\\
\breve{P}_0 &= \breve{P} = \breve{P}^0(\theta, I, w_{n(|n|\le\tilde{K}_0)}, \bar{w}_{n(|n|\le\tilde{K}_0)}, \xi) + \sum_{\substack{|n|>\tilde{K}_0\\ \alpha_n+\beta_n\ge 1}}\breve{P}^0_n(\theta, I, w_{m(|m|\le\tilde{K}_0)}, \bar{w}_{m(|m|\le\tilde{K}_0)}, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}\\
&\overset{\text{def}}{=} \breve{P}^0(\theta, I, z^0, \bar{z}^0, \xi) + \sum_{\substack{|n|>\tilde{K}_0\\ \alpha_n+\beta_n\ge 1}}\breve{P}^0_n(\theta, I, z^0, \bar{z}^0, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n},\\
&\|\breve{P}^0_n(\theta, I, z^0, \bar{z}^0, \xi)\| \le e^{-(|n|-\tilde{K}_0)}, \quad |n| > \tilde{K}_0,\\
\acute{P}_0 &= \acute{P} = \sum_{\substack{n\ne m\\ \alpha_n+\beta_n,\alpha_m+\beta_m\ge 1\\ \alpha_n+\beta_n+\alpha_m+\beta_m\ge 3}}\acute{P}^0_{nm}(\xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}w_m^{\alpha_m}\bar{w}_m^{\beta_m},\\
&\|\acute{P}^0_{nm}(\xi)\| \le e^{-|n-m|},\\
\grave{P}_0 &= \grave{P} = \sum_{n\in\mathbb{Z}} O(|w_n|^3),
\end{aligned}$$

with

$$z^0_l = (\cdots, w_n, \cdots)_{\substack{\Omega_n\in\Lambda_l\\ |n|\le\tilde{K}_0}}, \quad z^0 = (z^0_1, \cdots, z^0_{L_0}),$$

$\mathrm{diam}(A^0_l) \le d$, and $\|X_{P^0}\|_{D(r_0,s_0),\mathcal{O}_0} \le \varepsilon_*^{\frac{5}{4}} \overset{\text{def}}{=} \varepsilon_0$.

Suppose that after a νth KAM step, we arrive at a Hamiltonian

$$\begin{aligned}
&H \equiv H_\nu = N + P = N + \breve{P} + \acute{P}_0 + \grave{P}_0,\\
&N = N_\nu = \langle\omega(\xi), I\rangle + \sum_{l=1}^{L}\langle A_l z_l, \bar{z}_l\rangle + \sum_{|n|>\tilde{K}}\Omega^0_n w_n\bar{w}_n,\\
&\breve{P} = \breve{P}_\nu = \breve{P}(\theta, I, z^\nu, \bar{z}^\nu, \xi) + \sum_{\substack{|n|>\tilde{K}\\ \alpha_n+\beta_n\ge 1}}\breve{P}_n(\theta, I, z, \bar{z}, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}\\
&\quad = \breve{P}^\nu(\theta, I, z^\nu, \bar{z}^\nu, \xi) + \sum_{\substack{|n|>\tilde{K}_\nu\\ \alpha_n+\beta_n\ge 1}}\breve{P}^\nu_n(\theta, I, z^\nu, \bar{z}^\nu, \xi)w_n^{\alpha_n}\bar{w}_n^{\beta_n}
\end{aligned}$$

defined on a domain $D(r,s)\times\mathcal{O} = D(r_\nu, s_\nu)\times\mathcal{O}_\nu$, where $\tilde{K} = \tilde{K}_\nu$ is a positive constant, $L = L_\nu$ is the smallest positive integer such that $\{\Omega^0_n\}_{|n|\le\tilde{K}}$ lie in $\Lambda_1, \cdots, \Lambda_L$,

$$z_l = z^\nu_l = (\cdots, w_n, \cdots)_{\substack{\Omega_n\in\Lambda_l\\ |n|\le\tilde{K}}}, \quad \bar{z}_l = \bar{z}^\nu_l = (\cdots, \bar{w}_n, \cdots)_{\substack{\Omega_n\in\Lambda_l\\ |n|\le\tilde{K}}},$$
$$z = z^\nu = (z_1, \cdots, z_L), \quad \bar{z} = \bar{z}^\nu = (\bar{z}_1, \cdots, \bar{z}_L),$$

$P = P^\nu$, for some $\varepsilon = \varepsilon_\nu$, and

$$\|\breve{P}_n(\theta, I, z, \bar{z}, \xi)\|_{D(r,s),\mathcal{O}} \le e^{-(|n|-\tilde{K})}, \quad |n| > \tilde{K}.$$

It is clear that $L \le \tilde{K}$.

We will construct a symplectic transformation $\Phi = \Phi_\nu$, which, in smaller frequency and phase domains, carries the above Hamiltonian into the next KAM cycle. Below, all constants $c_1 - c_{12}$

are positive and independent of the iteration process. 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

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

We also use $\|\cdot\|$ to denote the operator matrix norm, i.e., for a matrix M, $\|M\|=\sup_{\|y\|=1}\|My\|$.

Let $\tilde K_+=\frac{7}{4}\tilde K+\frac14 K_0$. In this KAM step, Ω_n^0 with $\tilde K<|n|\le \tilde K_+$ will be added to the new normal matrix A_l^+ according to the assumption $(A2)$. In order to have a compact formulation when solving a homological equation, we rewrite N as

$$\begin{aligned} N &= e+\langle\omega(\xi),I\rangle+\sum_{l=1}^{L}\langle A_l z_l,\bar z_l\rangle\\ &+ \sum_{\tilde K<|n|\le\tilde K_+}\Omega_n^0 w_n\bar w_n+\sum_{|n|>\tilde K_+}\Omega_n^0 w_n\bar w_n\\ &\overset{\text{def}}{=} e+\langle\omega(\xi),I\rangle+\sum_{l=1}^{L_+}\langle\tilde A_l z_l^+,\bar z_l^+\rangle+\sum_{|n|>\tilde K_+}\Omega_n^0 w_n\bar w_n, \end{aligned}$$

where $\dim(\tilde A_l)\le d$, L_+ is the smallest positive integer such that $\{\Omega_n\}_{|n|\le\tilde K_+}$ lie in $\Lambda_1,\cdots,\Lambda_{L_+}$ (hence $L_+\le\tilde K_+$),

$$\tilde A_l=\begin{pmatrix} A_l & 0\\ 0 & \Omega_n^0\end{pmatrix},\quad \Omega_n^0\in\Lambda_l,\ \tilde K<|n|\le\tilde K_+,$$

$$z_l^+=(\cdots,w_n,\cdots)_{\substack{\Omega_n\in\Lambda_l\\ |n|\le\tilde K_+}},\quad \bar z_l^+=(\cdots,\bar w_n,\cdots)_{\substack{\Omega_n\in\Lambda_l\\ |n|\le\tilde K_+}},$$

$$z^+=(z_1^+,\cdots,z_{L_+}^+),\quad \bar z^+=(\bar z_1^+,\cdots,\bar z_{L_+}^+).$$

3.2. **Truncation.** We first expand $\breve P$ into the Taylor-Fourier series

$$\breve P=\sum_{k,l,\alpha,\beta}\breve P_{kl\alpha\beta}e^{\mathrm{i}\langle k,\theta\rangle}\ I^l z^\alpha\bar z^\beta+\sum_{\substack{k,l,n,\alpha,\beta\\ |n|>\tilde K,\alpha_n+\beta_n\ge1}}\breve P_{kln\alpha\beta}e^{\mathrm{i}\langle k,\theta\rangle}\ I^l z^\alpha\bar z^\beta w_n^{\alpha_n}\bar w_n^{\beta_n},$$

where $k\in\mathbb Z^N, l\in\mathbb N^N$ and the multi-index α (resp. β) runs over the set $\alpha\equiv(\alpha^1,\cdots,\alpha^l,\cdots,\alpha^L)$ for $\alpha^l=(\cdots,\alpha_m,\cdots)_{\substack{\Omega_m\in\Lambda_l\\ |m|\le\tilde K}}$, $\alpha_m\in\mathbb N$ (resp. $\beta\equiv(\beta^1,\cdots,\beta^l,\cdots,\beta^L)$ for $\beta^l=(\cdots,\beta_m,\cdots)_{\substack{\Omega_m\in\Lambda_l\\ |m|\le\tilde K}}$, $\beta_m\in\mathbb N$). Let R be the following truncation of $\breve P$:

$$\begin{aligned} R(\theta,I,z,\bar z,w,\bar w)=&\sum_{|k|\le K_+,|l|\le1}\breve P_{kl00}e^{\mathrm{i}\langle k,\theta\rangle}\ I^l\\ +&\sum_{|k|\le K_+,l\le L}(\langle\breve P_l^{k10},z_l\rangle+\langle\breve P_l^{k01},\bar z_l\rangle)e^{\mathrm{i}\langle k,\theta\rangle}+\sum_{|k|\le K_+,\tilde K<|n|\le\tilde K_+}(\breve P_n^{k10}w_n+\breve P_n^{k01}\bar w_n)e^{\mathrm{i}\langle k,\theta\rangle}\\ +&\sum_{|k|\le K_+,l,j\le L}(\langle\breve P_{jl}^{k20}z_l,z_j\rangle+\langle\breve P_{jl}^{k11}z_l,\bar z_j\rangle+\langle\breve P_{jl}^{k02}\bar z_l,\bar z_j)e^{\mathrm{i}\langle k,\theta\rangle}\\ +&\sum_{|k|\le K_+,l\le L,\tilde K<|n|\le\tilde K_+}(\langle\breve P_{nl}^{k20}z_l,w_n\rangle+\langle\breve P_{nl}^{k11}z_l,\bar w_n\rangle+\langle\breve P_{ln}^{k11}w_n,\bar z_l\rangle+\langle\breve P_{nl}^{k02}\bar z_l,\bar w_n\rangle)e^{\mathrm{i}\langle k,\theta\rangle}\\ +&\sum_{|k|\le K_+,\tilde K<|n|\le\tilde K_+}(\breve P_{nn}^{k20}w_nw_n+\breve P_{nn}^{k11}w_n\bar w_n+\breve P_{nn}^{k02}\bar w_n\bar w_n)e^{\mathrm{i}\langle k,\theta\rangle}, \end{aligned}$$

where $K_+=\frac34\tilde K+\frac14 K_0$.

Remark 3.1. *Due to decay property, terms in the Taylor-Fourier expansion of $\breve{P}$ corresponding to $|k| > K_+$ or $|n| > \tilde{K}_+$ are small enough to be postponed to the next KAM step. In addition, due to the decay property of $\breve{P}$ and the fact that $\acute{P}_0$ starts from third order terms, there are no coupling terms of the form $\sum_{\tilde{K}<|n|,|m|\le\tilde{K}_+}^{n\neq m} w_n\bar{w}_m$ in R. If $\acute{P}_0$ starts from second order terms, then couplings between different oscillators are so fast that we have to consider all normal frequencies - a case we are not able to handle with the method in this paper.*

According to the normal form N and the assumption $(A2)$, we may rewrite R as

$$\begin{aligned} R(\theta, I, z^+, \bar{z}^+) &= R_0 + R_1 + R_2 = \sum_{|k|\le K_+,|l|\le 1} P_{kl00} e^{\mathrm{i}\langle k,\theta\rangle} I^l \\ &+ \sum_{|k|\le K_+, l\le L_+} (\langle R_l^{k10}, z_l^+\rangle + \langle R_l^{k01}, \bar{z}_l^+\rangle) e^{\mathrm{i}\langle k,\theta\rangle} \\ &+ \sum_{|k|\le K_+, l,j\le L_+} (\langle R_{jl}^{k20} z_l^+, z_j^+\rangle + \langle R_{jl}^{k11} z_l^+, \bar{z}_j^+\rangle + \langle R_{jl}^{k02} \bar{z}_l^+, \bar{z}_j^+\rangle) e^{\mathrm{i}\langle k,\theta\rangle}, \end{aligned}$$

where,

$$R_l^{k10} = \begin{pmatrix} \breve{P}_l^{k10} \\ \breve{P}_n^{k10} \end{pmatrix}_{\tilde{K}<|n|\le\tilde{K}_+},$$

$$R_l^{k01} = \begin{pmatrix} \breve{P}_l^{k01} \\ \breve{P}_n^{k01} \end{pmatrix}_{\tilde{K}<|n|\le\tilde{K}_+},$$

$$R_{jl}^{k20} = \begin{pmatrix} \breve{P}_{jl}^{k20} & \breve{P}_{jn}^{k20} \\ \breve{P}_{nl}^{k20} & \breve{P}_{nn}^{k20} \end{pmatrix}_{\tilde{K}<|n|\le\tilde{K}_+},$$

$$R_{jl}^{k11} = \begin{pmatrix} \breve{P}_{jl}^{k11} & \breve{P}_{jn}^{k11} \\ \breve{P}_{nl}^{k11} & \breve{P}_{nn}^{k11} \end{pmatrix}_{\tilde{K}<|n|\le\tilde{K}_+},$$

$$R_{jl}^{k02} = \begin{pmatrix} \breve{P}_{jl}^{k02} & \breve{P}_{jn}^{k02} \\ \breve{P}_{nl}^{k02} & \breve{P}_{nn}^{k02} \end{pmatrix}_{\tilde{K}<|n|\le\tilde{K}_+}.$$

It is clear that $R_{lj}^{k20} = (R_{jl}^{k20})^\top$, $R_{lj}^{k11} = (R_{jl}^{k11})^\top$ and $R_{lj}^{k02} = (R_{jl}^{k02})^\top$.

Remark 3.2. *Note that R is dependent on the former $\tilde{K}_+$ normal components, and independent of the latter normal components. As a result, in the KAM step, we do not need to consider small divisors $\langle k,\omega\rangle + \Omega_n^0$, $\langle k,\omega\rangle \pm \Omega_n^0 \pm \Omega_m^0$, $|n| > \tilde{K}_+$ or $|m| > \tilde{K}_+$. This enables us to only consider finite small divisor conditions at each KAM step. Consequently, the excluded measure is sufficiently small at each step. As $\tilde{K}_+$ will increase along KAM iterations, we ultimately handle all small divisor conditions.*

Rewrite H as $H = N + R + (P - R)$. By the definition of norms, we immediately have

$$\|X_R\|_{D(r,s),\mathcal{O}} \le \|X_P\|_{D(r,s),\mathcal{O}} \le \varepsilon.$$

Note that

$$P - R = \sum_{|k|>K_+} \breve{P}_k(I, z, \bar{z}, \xi) e^{\mathrm{i}\langle k,\theta\rangle} + \sum_{|n|>\tilde{K}_+, \alpha_n+\beta_n\ge 1} \breve{P}_n(\theta, I, z, \bar{z}, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} + h.o.t.$$

where h.o.t. denotes the terms of the form $O(|I|^2 + |I||w| + |w|^3)$. Let $r_+ = \frac{r}{2} + \frac{r_0}{4}$, $\eta = \varepsilon^{\frac{1}{4}}$. Using the facts

$$\|\breve{P}_n(\theta, I, z, \bar{z}, \xi)\| \le e^{-(|n|-\tilde{K})}, \qquad \|\breve{P}_k(I, z, \bar{z}, \xi)\| \le e^{-|k|r},$$

we have that if

C1) $e^{-K_+\frac{r-r_+}{2}} \le \varepsilon^{\frac{5}{4}}$, then

$$\|X_{P-R}\|_{D(r_++\frac{r-r_+}{2},\eta s),\mathcal{O}} \le \sum_{|k|>K_+} e^{-|k|\frac{r-r_+}{2}} + \sum_{|n|>\tilde{K}_+} e^{-(|n|-\tilde{K})} + h.o.t. \le c_1\varepsilon^{\frac{5}{4}}. \tag{3.3}$$

3.3. **The homological equation.** We now look for a Hamiltonian F, defined in a domain $D_+ = D(r_+, s_+)$, such that the time-1 map $\Phi = \Phi_F^1$ of the Hamiltonian vector field X_F defines a map from D_+ to D and transforms H into H_+ in the next KAM cycle. Let F have the form

$$\begin{aligned} F(\theta, I, z^+, \bar{z}^+) &= F_0 + F_1 + F_2 \\ &= \sum_{0<|k|\le K_+,|l|\le 1} F_{kl00}e^{\mathrm{i}\langle k,\theta\rangle}\, I^l + \sum_{|k|\le K_+,|n|\le \tilde{K}_+} (f_n^{k10}w_n + f_n^{k01}\bar{w}_n)e^{\mathrm{i}\langle k,\theta\rangle} \\ &+ \sum_{|k|\le K_+,|n|,|m|\le \tilde{K}_+} (f_{nm}^{k20}w_nw_m + f_{nm}^{k02}\bar{w}_n\bar{w}_m)e^{\mathrm{i}\langle k,\theta\rangle} \\ &+ \sum_{|k|\le K_+,|n|,|m|\le \tilde{K}_+} f_{nm}^{k11}w_n\bar{w}_m e^{\mathrm{i}\langle k,\theta\rangle} \\ &= \sum_{0<|k|\le K_+,|l|\le 1} F_{kl00}e^{\mathrm{i}\langle k,\theta\rangle}\, I^l + \sum_{|k|\le K_+,l\le L_+} (\langle F_l^{k10}, z_l^+\rangle + \langle F_l^{k01}, \bar{z}_l^+\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\ &+ \sum_{|k|\le K_+,l,j\le L_+} (\langle F_{jl}^{k20}z_l^+, z_j^+\rangle + \langle F_{jl}^{k02}\bar{z}_l^+, \bar{z}_j^+\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\ &+ \sum_{\substack{|k|\le K_+,l,j\le L_+\\ |k|+|l-j|\ne 0}} \langle F_{jl}^{k11}z_l^+, \bar{z}_j^+\rangle e^{\mathrm{i}\langle k,\theta\rangle} \end{aligned} \tag{3.4}$$

and satisfy the homological equation

$$\{N, F\} + R = \breve{P}_{0000} + \langle \omega', I\rangle + \sum_{l=1}^{L_+}\langle R_{ll}^{011}z_l^+, \bar{z}_l^+\rangle, \tag{3.5}$$

where

$$\omega' = \int \frac{\partial \breve{P}}{\partial I}d\theta|_{z^+=\bar{z}^+=0,I=0}.$$

In the rest of the sub-section, $|k|$ is always bounded by K_+, $|n|,|m|$ are always bounded by $\tilde{K}_+$, and l, j are always bounded by L_+.

Lemma 3.1. *Equation* (3.5) *is equivalent to*

$$\begin{aligned} &\langle k,\omega\rangle F_{kl00} = \mathrm{i}\breve{P}_{kl00}, \qquad k\ne 0, |l|\le 1, \\ &(\langle k,\omega\rangle I - \tilde{A}_l)F_l^{k10} = \mathrm{i}R_l^{k10}, \\ &(\langle k,\omega\rangle I + \tilde{A}_l)F_l^{k01} = \mathrm{i}R_l^{k01}, \\ &(\langle k,\omega\rangle I - \tilde{A}_j)F_{jl}^{k20} - F_{jl}^{k20}\tilde{A}_l = \mathrm{i}R_{jl}^{k20}, \\ &(\langle k,\omega\rangle I + \tilde{A}_j)F_{jl}^{k11} - F_{jl}^{k11}\tilde{A}_l = \mathrm{i}R_{jl}^{k11}, \quad |k| + |l-j| \ne 0, \\ &(\langle k,\omega\rangle I + \tilde{A}_j)F_{jl}^{k02} + F_{jl}^{k02}\tilde{A}_l = \mathrm{i}R_{jl}^{k02}. \end{aligned} \tag{3.6}$$

Proof. It is clear that (3.5) is equivalent to the following equations

$$\begin{aligned} &\{N, F_0\} + R_0 = \breve{P}_{0000} + \langle\omega', I\rangle, \\ &\{N, F_1\} + R_1 = 0, \\ &\{N, F_2\} + R_2 = \sum_{l=1}^{L_+}\langle R_{ll}^{011}z_l^+, \bar{z}_l^+\rangle. \end{aligned} \tag{3.7}$$

By comparing the coefficients, the first equation in (3.7) is obviously equivalent to the first equation in (3.6). Since

$$\begin{aligned}
\{N,F_1\} &= \mathrm{i}\sum_{k,l}(\langle\, \langle k,\omega\rangle F_l^{k10}, z_l^+\rangle - \langle \tilde{A}_l z_l^+, F_l^{k10}\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \mathrm{i}\sum_{k,l}(\langle\, \langle k,\omega\rangle F_l^{k01}, \bar{z}_l^+\rangle + \langle \tilde{A}_l \bar{z}_l^+, F_l^{k01}\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\
&= \mathrm{i}\sum_{k,l}\langle(\langle k,\omega\rangle I - \tilde{A}_l)F_l^{k10}, z_l^+\rangle e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \mathrm{i}\sum_{k,l}\langle(\langle k,\omega\rangle I + \tilde{A}_l)F_l^{k01}, \bar{z}_l^+\rangle e^{\mathrm{i}\langle k,\theta\rangle}\ ,
\end{aligned}$$

and,

$$\begin{aligned}
\{N,F_2\} &= \mathrm{i}\sum_{k,l,j}(\langle\, \langle k,\omega\rangle F_{jl}^{k20} z_l^+, z_j^+\rangle - \langle F_{jl}^{k20} z_l^+, \tilde{A}_j z_j^+\rangle - \langle \tilde{A}_l z_l^+, (F_{jl}^{k20})^T z_j^+\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \mathrm{i}\sum_{|k|+|l-j|\neq 0}(\langle\, \langle k,\omega\rangle F_{jl}^{k11} z_l^+, \bar{z}_j^+\rangle + \langle F_{jl}^{k11} z_l^+, \tilde{A}_j \bar{z}_j^+\rangle - \langle \tilde{A}_l z_l^+, (F_{jl}^{k11})^T \bar{z}_j^+\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \mathrm{i}\sum_{k,l,j}(\langle\, \langle k,\omega\rangle F_{jl}^{k02} \bar{z}_l^+, \bar{z}_j^+\rangle + \langle F_{jl}^{k02} \bar{z}_l^+, \tilde{A}_j \bar{z}_j^+\rangle + \langle \tilde{A}_l \bar{z}_l^+, (F_{jl}^{k02})^T \bar{z}_j^+\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\
&= \mathrm{i}\sum_{k,l,j}(\langle\, \langle k,\omega\rangle F_{jl}^{k20} z_l^+, z_j^+\rangle - \langle(\tilde{A}_j F_{jl}^{k20} + F_{jl}^{k20}\tilde{A}_l) z_l^+, z_j^+\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \mathrm{i}\sum_{|k|+|l-j|\neq 0}(\langle\, \langle k,\omega\rangle F_{jl}^{k11} z_l^+, \bar{z}_j^+\rangle + \langle(\tilde{A}_j F_{jl}^{k11} - F_{jl}^{k11}\tilde{A}_l) z_l^+, \bar{z}_j^+\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \mathrm{i}\sum_{k,l,j}(\langle\, \langle k,\omega\rangle F_{jl}^{k02} \bar{z}_l^+, \bar{z}_j^+\rangle + \langle(\tilde{A}_j F_{jl}^{k02} + F_{jl}^{k02}\tilde{A}_l) \bar{z}_l^+, \bar{z}_j^+\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\
&= \mathrm{i}\sum_{k,l,j}\langle(\langle k,\omega\rangle F_{jl}^{k20} - \tilde{A}_j F_{jl}^{k20} - F_{jl}^{k20}\tilde{A}_l) z_l^+, z_j^+\rangle e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \mathrm{i}\sum_{|k|+|l-j|\neq 0}\langle(\langle k,\omega\rangle F_{jl}^{k11} + \tilde{A}_j F_{jl}^{k11} - F_{jl}^{k11}\tilde{A}_l) z_l^+, \bar{z}_j^+\rangle e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \mathrm{i}\sum_{k,l,j}\langle(\langle k,\omega\rangle F_{jl}^{k02} + \tilde{A}_j F_{jl}^{k02} + F_{jl}^{k02}\tilde{A}_l) \bar{z}_l^+, \bar{z}_j^+\rangle e^{\mathrm{i}\langle k,\theta\rangle}\ ,
\end{aligned}$$

we see from the second and the third equation in (3.7) that F_l^{k10}, F_l^{k01}, F_{jl}^{k20}, F_{jl}^{k11}, F_{jl}^{k02} satisfy the respective equations in (3.6). □

Let

$$\mathcal{O}_+ = \left\{ \xi\in\mathcal{O} : \begin{array}{ll}
|\langle k,\omega\rangle^{-1}| \le \left(\frac{K_+^\tau}{\gamma}\right)^{d^2}, & 0<|k|\le K_+ \\
\|(\langle k,\omega\rangle I + \tilde{A}_l)^{-1}\| \le \left(\frac{K_+^\tau}{\gamma}\right)^{d^2}, & |k|\le K_+, l\le L_+ \\
\|(\langle k,\omega\rangle I + \tilde{A}_l\otimes I + I\otimes\tilde{A}_j)^{-1}\| \le \left(\frac{K_+^\tau}{\gamma}\right)^{d^2}, & |k|\le K_+, l,j\le L_+ \\
\|(\langle k,\omega\rangle I + \tilde{A}_l\otimes I - I\otimes\tilde{A}_j)^{-1}\| \le \left(\frac{K_+^\tau}{\gamma}\right)^{d^2}, & |k|+|l-j|\neq 0, \\
 & |k|\le K_+, l,j\le L_+
\end{array}\right\}.$$

The first three equations in (3.6) can be immediately solved on $\mathcal{O}_+$. The solvability of the remaining equations in (3.6) follows from the following elementary algebraic result from matrix theory.

Lemma 3.2. *Let A, B, C be $n \times n, m \times m, n \times m$ matrices respectively, and let X be an $n \times m$ unknown matrix. The matrix equation*

$$AX - XB = C, \tag{3.8}$$

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

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

Proof. See [24, 32]. □

By taking the transpose of the fourth equation in (3.6), one sees that $(F_{jl}^{k20})^\top$ satisfies the same equation as F_{lj}^{k20}. We have by the uniqueness of solution that $F_{lj}^{k20} = (F_{jl}^{k20})^\top$. Similarly, $F_{lj}^{k11} = (F_{jl}^{k11})^\top$ and $F_{lj}^{k02} = (F_{jl}^{k02})^\top$. Hence, the Hamiltonian F is uniquely determined on $\mathcal{O}_+$.

We proceed to estimate X_F and Φ_F^1. .

Lemma 3.3. *Let $D_i = D(r_+ + \frac{i}{4}(r - r_+), \frac{i}{4}s)$, $0 < i \leq 4$. If*

C2) $K_+^{\tau d^4} \leq \varepsilon^{-\frac{1}{4}}$,

then there is a constant $c_2 > 0$ such that

$$\|X_F\|_{D_3, \mathcal{O}_+} \leq c_2 \gamma^{-d^4} (r - r_+)^{-N} \varepsilon^{\frac{3}{4}}.$$

Proof. By the definition of $\mathcal{O}_+$, Lemma 3.1, Lemma 3.2 and Lemma 5.5, Lemma 5.6 in the Appendix, we have

$$\begin{aligned}
|F_{kl00}|_{\mathcal{O}_+} &\leq |\langle k, \omega \rangle|^{-d^2} |\breve{P}_{kl00}|_{\mathcal{O}_+} \leq \gamma^{-d^4} K_+^{\tau d^4} |\breve{P}_{kl00}|_{\mathcal{O}_+}, \quad k \neq 0; \\
\|F_l^{k10}\|_{\mathcal{O}_+} &\leq \gamma^{-d^4} K_+^{\tau d^4} \|R_l^{k10}\|_{\mathcal{O}_+}; \\
\|F_l^{k01}\|_{\mathcal{O}_+} &\leq \gamma^{-d^4} K_+^{\tau d^4} \|R_l^{k01}\|_{\mathcal{O}_+}; \\
\|F_{jl}^{k20}\|_{\mathcal{O}_+} &\leq \gamma^{-d^4} K_+^{\tau d^4} \|R_{jl}^{k20}\|_{\mathcal{O}_+}; \\
\|F_{jl}^{k11}\|_{\mathcal{O}_+} &\leq \gamma^{-d^4} K_+^{\tau d^4} \|R_{jl}^{k11}\|_{\mathcal{O}_+}, \ |k| + |l - j| \neq 0; \\
\|F_{jl}^{k02}\|_{\mathcal{O}_+} &\leq \gamma^{-d^4} K_+^{\tau d^4} \|R_{jl}^{k02}\|_{\mathcal{O}_+}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\frac{1}{s^2} \|F_\theta\|_{D_3, \mathcal{O}_+} &\leq \frac{1}{s^2} \Big(\sum_{k, |l| \leq 1} |F_{kl00}| \cdot s^{2|l|} \cdot |k| \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} \\
&+ \sum_{k,l} (\|F_l^{k10}\| \cdot \|z_l^+\|) \cdot |k| \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} \\
&+ \sum_{k,l} (\|F_l^{k01}\| \cdot \|\bar{z}_l^+\|) \cdot |k| \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} \\
&+ \sum_{k,l,j} \|F_{jl}^{k20}\| \cdot \|z_l^+\| \|z_j^+\| \cdot |k| \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} \\
&+ \sum_{|k| + |l - j| \neq 0} \|F_{jl}^{k11}\| \cdot \|z_l^+\| \|\bar{z}_j^+\| \cdot |k| \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} \\
&+ \sum_{k,l,j} \|F_{jl}^{k02}\| \cdot \|\bar{z}_l^+\| \|\bar{z}_j^+\| \cdot |k| \cdot e^{|k|(r - \frac{1}{4}(r - r_+))} \Big)
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{\gamma^{-d^4}K_+^{\tau d^4}}{s^2}\Big(\sum_{k,|l|\leq 1}|\breve{P}_{kl00}|\cdot s^{2|l|}\cdot|k|\cdot e^{|k|(r-\frac{1}{4}(r-r_+))}\\
&+ \sum_{k,l}(\|R_l^{k10}\|\cdot\|z_l^+\|)\cdot|k|\cdot e^{|k|(r-\frac{1}{4}(r-r_+))}\\
&+ \sum_{k,l}(\|R_l^{k01}\|\cdot\|\bar{z}_l^+\|)\cdot|k|\cdot e^{|k|(r-\frac{1}{4}(r-r_+))}\\
&+ \sum_{k,l,j}\|R_{jl}^{k20}\|\cdot\|z_l^+\|\|z_j^+\|\cdot|k|\cdot e^{|k|(r-\frac{1}{4}(r-r_+))}\\
&+ \sum_{|k|+|l-j|\neq 0}\|R_{jl}^{k11}\|\cdot\|z_l^+\|\|\bar{z}_j^+\|\cdot|k|\cdot e^{|k|(r-\frac{1}{4}(r-r_+))}\\
&+ \sum_{k,l,j}\|R_{jl}^{k02}\|\cdot\|\bar{z}_l^+\|\|\bar{z}_j^+\|\cdot|k|\cdot e^{|k|(r-\frac{1}{4}(r-r_+))}\Big)\\
&\leq c_3\gamma^{-d^4}(r-r_+)^{-N}K_+^{\tau d^4}\|X_R\|\\
&\leq c_3\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{3}{4}}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|F_I\|_{D_3,\mathcal{O}_+} &= \sum_{|l|=1}|F_{kl00}|e^{|k|(r-\frac{1}{4}(r-r_+))}\leq c_4\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{3}{4}}.\\
\|X_{F_1}\|_{D_3,\mathcal{O}_+} &\leq \frac{1}{s}\Big(\sum_n\|F_{1_{w_n}}\|+\sum_n\|F_{1_{\bar{w}_n}}\|\Big)\leq\frac{1}{s}\Big(\sum_l\|F_{1_{z_l^+}}\|+\sum_l\|F_{1_{\bar{z}_l^+}}\|\Big)\\
&\leq c_5\gamma^{-d^4}(r-r_+)^{-N}K_+^{\tau d^4}\|X_{R_1}\|\leq c_5\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{3}{4}}.\\
\|X_{F_2}\|_{D_3,\mathcal{O}_+} &\leq \frac{1}{s}\Big(\sum_n\|F_{2_{w_n}}\|+\sum_n\|F_{2_{\bar{w}_n}}\|\Big)\leq\frac{1}{s}\Big(\sum_l\|F_{2_{z_l^+}}\|+\sum_l\|F_{2_{\bar{z}_l^+}}\|\Big)\\
&\leq c_6\gamma^{-d^4}(r-r_+)^{-N}K_+^{\tau d^4}\|X_{R_2}\|\leq c_6\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{3}{4}}.
\end{aligned}$$

The proof is now completed by adding the estimates above together. □

Let $D_{i\eta}=D(r_++\frac{i}{4}(r-r_+),\frac{i}{4}\eta s), 0<i\leq 4$.

Lemma 3.4. *If*

C3) $c_2\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{1}{2}}<1$,

then

(3.9) $$\Phi_F^t: D_{2\eta}\to D_{3\eta},\quad -1\leq t\leq 1,$$

and moreover,

(3.10) $$\|D\Phi_F^t-I\|_{D_{1\eta}}<c_7\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{3}{4}}.$$

Proof. Let

$$\|D^mF\|_{D,\mathcal{O}_+}=\max\{\|\frac{\partial^{|i|+|l|+|\alpha|+|\beta|}}{\partial\theta^i\partial I^l\partial(z^+)^\alpha\partial(\bar{z}^+)^\beta}F\|_{D,\mathcal{O}_+},|i|+|l|+|\alpha|+|\beta|=m\geq 2\}.$$

We note that F is a polynomial of order 1 in I and of order 2 in z^+, $\bar{z}^+$. It follows from Lemma 3.3 and the Cauchy inequality that

$$\|D^mF\|_{D_2,\mathcal{O}_+}<c_8\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{3}{4}},$$

for any $m\geq 2$.

Using the integral equation

$$\Phi_F^t=id+\int_0^t X_F\circ\Phi_F^s\,ds$$

and Lemma 3.3, we easily see that $\Phi_F^t : D_{2\eta} \to D_{3\eta}, \quad -1 \le t \le 1$. Since

$$D\Phi_F^t = Id + \int_0^t (DX_F) D\Phi_F^s \, ds = Id + \int_0^t J(D^2 F) D\Phi_F^s \, ds,$$

where J denotes the standard symplectic matrix. Let $c_7 = 2c_8$. It follows that

$$\|D\Phi_F^t - I\| \le 2\|D^2 F\| \le c_7 \gamma^{-d^4} (r - r_+)^{-N} \varepsilon^{\frac{3}{4}}.$$

□

3.4. **The new Hamiltonian.** Let $\Phi = \Phi_F^1$, $s_+ = \frac{1}{8}\eta s$, $D_+ = D(r_+, s_+)$, and

$$\begin{aligned} N_+ &= e_+ + \langle \omega_+, I \rangle + \sum_{l=1}^{L_+} \langle A_l^+ z_l^+, \bar{z}_l^+ \rangle + \sum_{|n| > \tilde{K}_+} \Omega_n^0 w_n \bar{w}_n, \\ P^+ &= \breve{P}_+ + \acute{P}_0 + \grave{P}_0, \end{aligned}$$

where

$$\begin{aligned} e_+ &= e + \breve{P}_{0000}, \ \omega_+ = \omega + \breve{P}_{0l00}(|l| = 1), \\ A_l^+ &= \tilde{A}_l + R_{ll}^{011}, \quad l \le L_+, \\ z_l^+ &= (\cdots, w_n, \cdots)_{|n| \le \tilde{K}_+}, \quad \bar{z}_l^+ = (\cdots, \bar{w}_n, \cdots)_{|n| \le \tilde{K}_+}, \\ z^+ &= (z_1^+, \cdots, z_{L_+}^+), \quad \bar{z}^+ = (\bar{z}_1^+, \cdots, \bar{z}_{L_+}^+), \\ \breve{P}_+ &= \int_0^1 (1-t)\{\{N, F\}, F\} \circ \Phi_F^t dt + \int_0^1 \{R, F\} \circ \Phi_F^t dt + (\breve{P} - R) \circ \Phi_F^1 + \int_0^1 \{\acute{P}_0 + \grave{P}_0, F\} \circ \Phi_F^t dt. \end{aligned}$$

Then $\Phi : D_+ \times \mathcal{O}_+ \to D$, and, by the second order Taylor formula,

$$\begin{aligned} H_+ &\equiv H \circ \Phi = (N + R) \circ \Phi + (P - R) \circ \Phi \\ &= N + \{N, F\} + R + \int_0^1 (1-t)\{\{N, F\}, F\} \circ \Phi_F^t dt \\ &+ \int_0^1 \{R, F\} \circ \Phi_F^t dt + (\breve{P} - R) \circ \Phi_F^1 + (\acute{P}_0 + \grave{P}_0) \circ \Phi_F^1 \\ &= N + \{N, F\} + R + \breve{P}_+ + \acute{P}_0 + \grave{P}_0 \\ &= N_+ + P^+ + \{N, F\} + R - \breve{P}_{0000} - \langle \omega', I \rangle - \sum_{l=1}^{L_+} \langle R_{ll}^{011} z_l^+, \bar{z}_l^+ \rangle \\ &= N_+ + P^+. \end{aligned}$$

Below, we show that the new Hamiltonian H_+ enjoys similar properties as H.

By the Assumptions of $\breve{P}$, we have that there is a constant $c_9 > 0$ such that

$$|\omega_+ - \omega|_{\mathcal{O}_+} \le c_9 \varepsilon, \quad \|A_l^+ - \tilde{A}_l\|_{\mathcal{O}_+} \le c_9 \varepsilon.$$

Denote $R(t) = (1-t)(N_+ - N) + tR$. We can rewrite P^+ as

$$\begin{aligned} P^+ &= \int_0^1 (1-t)\{\{N, F\}, F\} \circ \Phi_F^t dt + \int_0^1 \{R, F\} \circ \Phi_F^t dt + (P - R) \circ \Phi_F^1 \\ &= \int_0^1 \{R(t), F\} \circ \Phi_F^t dt + (P - R) \circ \Phi_F^1. \end{aligned}$$

Hence

$$X_{P^+} = \int_0^1 (\Phi_F^t)^* X_{\{R(t), F\}} dt + (\Phi_F^1)^* X_{(P-R)}.$$

By Lemma 3.4, if

C4) $c_7 \gamma^{-d^4} (r - r_+)^{-N} \varepsilon^{\frac{3}{4}} \le 1,$

then
$$\|D\Phi_F^t\|_{D_{1\eta}} \le 1 + \|D\Phi_F^t - I\|_{D_{1\eta}} \le 2, \quad -1 \le t \le 1.$$
By Lemma 5.4 and (3.3), we also have
$$\|X_{\{R(t),F\}}\|_{D_{2\eta}} \le c_{10}\gamma^{-d^4}(r-r_+)^{-N}\eta^{-2}\varepsilon^{\frac{7}{4}},$$
$$\|X_{(P-R)}\|_{D_{2\eta}} \le c_1\varepsilon^{\frac{5}{4}}.$$
Let $c_0 = \max\{c_1, \cdots, c_{10}, c_{11}, c_{12}\}$, where c_{11}, c_{12} will be defined later, and let
$$\varepsilon_+ = 4c_0\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{5}{4}}.$$
Then
$$\|X_{P^+}\|_{D_+,\mathcal{O}_+} \le 2c_1\varepsilon^{\frac{5}{4}} + 2c_{10}\gamma^{-d^4}(r-r_+)^{-N}\varepsilon^{\frac{5}{4}} \le \varepsilon_+.$$
We now exam the decay property of $\breve{P}_+$. More precisely, write
$$\breve{P}_+ = \breve{P}^+(\theta, I, z^+, \bar{z}^+, \xi) + \sum_{|n|>\tilde{K}_+, \alpha_n+\beta_n\ge 1} \breve{P}_n^+(\theta, I, z^+, \bar{z}^+, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n}.$$
We show that
$$\|\breve{P}_n^+(\theta, I, z^+, \bar{z}^+, \xi)\|_{D_+,\mathcal{O}_+} \le e^{-(|n|-\tilde{K}_+)}, \quad |n| > \tilde{K}_+.$$
Since F only involves the normal components $w_n, \bar{w}_n$ for $|n| \le \tilde{K}_+$, so does $\{N, F\}$. Hence $\int_0^1 (1-t)\{\{N,F\},F\}\circ\Phi_F^t dt$ only involves the normal components $w_n, \bar{w}_n$ for $|n| \le \tilde{K}_+$. Recall that
$$\grave{P}_0 = \sum_n O(|w_n|^3).$$
Hence $\{\grave{P}_0, F\}$ only involves the normal components $w_n, \bar{w}_n$ for $|n| \le \tilde{K}_+$, so does $\int_0^1 \{\grave{P}_0, F\}\circ\Phi_F^t dt$. Since R is a truncation of $\breve{P}$, we only need to consider the terms $(\breve{P} - R)$ and $\int_0^1 \{\breve{P} + \acute{P}_0, F\}\circ\Phi_F^t dt$. Recall that
$$\breve{P} = \breve{P}(\theta, I, z, \bar{z}, \xi) + \sum_{|n|>\tilde{K}, \alpha_n+\beta_n\ge 1} \breve{P}_n(\theta, I, z, \bar{z}, \xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n},$$
$$\|\breve{P}_n(\theta, I, z, \bar{z}, \xi)\|_{D(r,s),\mathcal{O}_+} \le e^{-(|n|-\tilde{K})}.$$
$$\acute{P}_0 = \sum_{\substack{n\ne m \\ \alpha_n+\beta_n, \alpha_m+\beta_m \ge 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m\ge 3}} \acute{P}_{nm}^0(\xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n} w_m^{\alpha_m}\bar{w}_m^{\beta_m},$$
$$\|\acute{P}_{nm}^0(\xi)\|_{D(r,s),\mathcal{O}_+} \le e^{-|n-m|}.$$
Since R only involves the normal components $w_n, \bar{w}_n$ for $|n| \le \tilde{K}_+$, the terms corresponding to the normal components $w_n, \bar{w}_n$ for $|n| > \tilde{K}_+$ in $\breve{P} - R$ are just those corresponding for $|n| > \tilde{K}_+$ in $\breve{P}$, for which we have the decay property
$$\|\breve{P}_n(\theta, I, z, \bar{z}, \xi)\|_{D(r_+,s_+),\mathcal{O}_+} \le e^{-(|n|-\tilde{K})} \le e^{-(|n|-\tilde{K}_+)}.$$
To prove the decay estimates of $\int_0^1 \{\breve{P} + \acute{P}_0, F\}\circ\Phi_F^t dt$, we only need to consider the terms corresponding to the normal components $w_n, \bar{w}_n$ for $|n| > \tilde{K}_+$. Since F is independent of normal components $w_n, \bar{w}_n$ for $|n| > \tilde{K}_+$, so is $\int_0^1 \{\breve{P}(\theta, I, z, \bar{z}, \xi), F\}\circ\Phi_F^t dt$. Similarly,
$$\int_0^1 \{ \sum_{\substack{n\ne m, |n|, |m|\le \tilde{K}_+ \\ \alpha_n+\beta_n, \alpha_m+\beta_m\ge 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m\ge 3}} \acute{P}_{nm}^0(\xi) w_n^{\alpha_n}\bar{w}_n^{\beta_n} w_m^{\alpha_m}\bar{w}_m^{\beta_m}, F\}\circ\Phi_F^t dt$$

is independent of the normal components $w_n, \bar{w}_n$ for $|n| > \tilde{K}_+$. Thus, it remains to consider terms

$$\begin{aligned}
&\int_0^1 \{ \sum_{|n|>\tilde{K},\alpha_n+\beta_n\geq 1} \breve{P}_n(\theta, I, z, \bar{z}, \xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n}, F\} \circ \Phi_F^t dt \\
= &\int_0^1 \sum_{|n|>\tilde{K},\alpha_n+\beta_n\geq 1} \{\breve{P}_n(\theta, I, z, \bar{z}, \xi), F\} \circ \Phi_F^t w_n^{\alpha_n} \bar{w}_n^{\beta_n} dt \\
= &\sum_{|n|>\tilde{K},\alpha_n+\beta_n\geq 1} (\int_0^1 \{\breve{P}_n(\theta, I, z, \bar{z}, \xi), F\} \circ \Phi_F^t dt) w_n^{\alpha_n} \bar{w}_n^{\beta_n},
\end{aligned} \tag{3.11}$$

and

$$\begin{aligned}
&\int_0^1 \{ \sum_{\substack{n\neq m, |n|>\tilde{K}_+, |m|\leq \tilde{K}_+ \\ \alpha_n+\beta_n, \alpha_m+\beta_m\geq 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m\geq 3}} \acute{P}^0_{nm}(\xi) w_n^{\alpha_n} \bar{w}_n^{\beta_n} w_m^{\alpha_m} \bar{w}_m^{\beta_m}, F\} \circ \Phi_F^t dt \\
= &\int_0^1 \sum_{|n|>\tilde{K}_+} \{ \sum_{\substack{|m|\leq \tilde{K}_+ \\ \alpha_n+\beta_n, \alpha_m+\beta_m\geq 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m\geq 3}} \acute{P}^0_{nm}(\xi) w_m^{\alpha_m} \bar{w}_m^{\beta_m}, F\} \circ \Phi_F^t w_n^{\alpha_n} \bar{w}_n^{\beta_n} dt \\
= &\sum_{|n|>\tilde{K}_+} (\int_0^1 \{ \sum_{\substack{|m|\leq \tilde{K}_+ \\ \alpha_n+\beta_n, \alpha_m+\beta_m\geq 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m\geq 3}} \acute{P}^0_{nm}(\xi) w_m^{\alpha_m} \bar{w}_m^{\beta_m}, F\} \circ \Phi_F^t dt) w_n^{\alpha_n} \bar{w}_n^{\beta_n}.
\end{aligned} \tag{3.12}$$

Let

$$\tilde{P}_n = \breve{P}_n(\theta, I, z^+, \bar{z}^+, \xi) + \sum_{\substack{|m|\leq \tilde{K}_+ \\ \alpha_n+\beta_n, \alpha_m+\beta_m\geq 1 \\ \alpha_n+\beta_n+\alpha_m+\beta_m\geq 3}} \acute{P}^0_{nm}(\xi).$$

We combine (3.11) and (3.12) to consider decay property of

$$\sum_{|n|>\tilde{K},\alpha_n+\beta_n\geq 1} (\int_0^1 \{\tilde{P}_n, F\} \circ \Phi_F^t dt) w_n^{\alpha_n} \bar{w}_n^{\beta_n}.$$

By relaxing decay properties of $e^{-(|n|-\tilde{K})}$, $e^{-|n-m|}$ to $e^{-(|n|-\tilde{K}_+)}$, we have by Lemma 5.3 that

$$\|\{\tilde{P}_n, F\}\|_{D(r-\sigma,\frac{1}{2}s)} \leq c_{11}\gamma^{-d^4}(r-r_+)^{-N}\sigma^{-1}s^{-2}\varepsilon^{\frac{3}{4}}e^{-(|n|-\tilde{K}_+)}.$$

It follows from Cauchy estimate that

$$\|X_{\{\tilde{P}_n,F\}}\|_{D(r-2\sigma,\frac{1}{4}s)} \leq c_{12}\gamma^{-d^4}(r-r_+)^{-N}\sigma^{-2}s^{-4}\varepsilon^{\frac{3}{4}}e^{-(|n|-\tilde{K}_+)}.$$

Hence by Lemma 3.4, if

C5) $c_{11}\gamma^{-d^4}(r-r_+)^{-N}\eta^{-2}\varepsilon^{\frac{3}{4}} \leq \frac{1}{2}$,

C6) $c_{12}c_2(\gamma^{-d^4}(r-r_+)^{-N}\eta^{-2}\varepsilon^{\frac{3}{4}})^2 \leq \frac{1}{2}$,

then

$$\begin{aligned}
&\|\int_0^1 \{\tilde{P}_n, F\} \circ \Phi_F^t dt\|_{D(r_+,s_+)} \\
\leq\ &\|\{\tilde{P}_n, F\} \circ \Phi_F^t\|_{D(r_+,s_+)} \\
\leq\ &\|\{\tilde{P}_n, F\}\|_{D(r_+,s_+)} + \|\{\tilde{P}_n, F\} \circ \Phi_F^t - \{\tilde{P}_n, F\}\|_{D(r_+,s_+)} \\
\leq\ &\|\{\tilde{P}_n, F\}\|_{D(r_+,s_+)} + \|X_{\{\tilde{P}_n,F\}}\|_{D_{2\eta}}\|\Phi_F^t - id\|_{D_{1\eta}} \\
\leq\ &c_{11}\gamma^{-d^4}(r-r_+)^{-N}\eta^{-2}\varepsilon^{\frac{3}{4}}e^{-(|n|-\tilde{K}_+)} + c_{12}c_2(\gamma^{-d^4}(r-r_+)^{-N}\eta^{-2}\varepsilon^{\frac{3}{4}})^2 e^{-(|n|-\tilde{K}_+)} \\
\leq\ &e^{-(|n|-\tilde{K}_+)}.
\end{aligned}$$

This completes one step of KAM iterations. □

4. Proof of Theorem B

Let $r_0, s_0, \varepsilon_0, \gamma, K_0, \tilde{K}_0, \mathcal{O}_0, H_0$ be given at the beginning of Section 3. For each $\nu = 0, 1, \cdots$, we label all index-free quantities by ν and label all $+$-indexed quantities by $\nu + 1$. This defines, for all $\nu = 1, 2, \cdots$, the following sequences:

$$\begin{aligned}
r_\nu &= r_0(1 - \sum_{i=2}^{\nu+1} 2^{-i}),\\
\varepsilon_\nu &= 4c_0\gamma^{-d^4}(r_{\nu-1} - r_\nu)^{-N}\varepsilon_{\nu-1}^{\frac{5}{4}},\\
s_\nu &= \frac{1}{8}\eta_{\nu-1}s_{\nu-1} = 2^{-3\nu}(\prod_{i=0}^{\nu-1}\varepsilon_i)^{\frac{1}{4}}s_0,\ \ \eta_\nu = \varepsilon_\nu^{\frac{1}{4}},\\
K_\nu &= 4K_{\nu-1},\\
\tilde{K}_\nu &= \tilde{K}_{\nu-1} + K_\nu,\\
D_\nu &= D(r_\nu, s_\nu),\\
\tilde{D}_\nu &= D(r_{\nu+1} + \frac{1}{4}(r_\nu - r_{\nu+1}), \frac{1}{4}\eta_\nu s_\nu),\\
H_\nu &= N_\nu + P_\nu,\\
N_\nu &= e_\nu + \langle \omega_\nu(\xi), I\rangle + \sum_{l=1}^{L_\nu}\langle A_l^\nu z_l^\nu, \bar{z}_l^\nu\rangle + \sum_{|n|>\tilde{K}_\nu}\Omega_n^0 w_n\bar{w}_n,\\
\mathcal{O}_\nu &= \left\{\xi\in\mathcal{O}_{\nu-1}:\begin{array}{ll}
|\langle k,\omega_{\nu-1}\rangle^{-1}| \le \left(\frac{K_\nu^\tau}{\gamma}\right)^{d^2}, & 0<|k|\le K_\nu\\
\|(\langle k,\omega_\nu\rangle I + \tilde{A}_l^{\nu-1})^{-1}\| \le \left(\frac{K_\nu^\tau}{\gamma}\right)^{d^2}, & |k|\le K_\nu, l\le L_\nu\le\tilde{K}_\nu\\
\|(\langle k,\omega_{\nu-1}\rangle I + \tilde{A}_l^{\nu-1}\otimes I + I\otimes\tilde{A}_j^{\nu-1})^{-1}\| \le \left(\frac{K_\nu^\tau}{\gamma}\right)^{d^2}, & |k|\le K_\nu, l,j\le L_\nu\\
\|(\langle k,\omega_{\nu-1}\rangle I + \tilde{A}_l^{\nu-1}\otimes I - I\otimes\tilde{A}_j^{\nu-1})^{-1}\| \le \left(\frac{K_\nu^\tau}{\gamma}\right)^{d^2}, & \begin{array}{l}|k|+|l-j|\ne 0,\\ |k|\le K_\nu, l,j\le L_\nu\end{array}
\end{array}\right\},
\end{aligned}$$

where L_ν is the smallest positive integer such that $\{\Omega_n^0\}_{|n|\le\tilde{K}_\nu}$ lie in $\Lambda_1, \cdots, \Lambda_{L_\nu}$, and,

$$\tilde{A}_l^{\nu-1} = \begin{pmatrix} A_l^{\nu-1} & 0\\ 0 & \Omega_n^0\end{pmatrix},\quad \Omega_n^0\in\Lambda_l,\ \tilde{K}_{\nu-1} < |n| \le \tilde{K}_\nu.$$

4.1. Iteration Lemma.

The preceding analysis may be summarized as follows.

Lemma 4.1. *If ε is sufficiently small, then the following holds for all $\nu = 0, 1, \cdots$.*

a) *H_ν is real analytic on $D_\nu \times \mathcal{O}_\nu$,*

$$\begin{aligned}
N_\nu &= e_\nu + \langle\omega_\nu(\xi), I\rangle + \sum_{l=1}^{L_{\nu+1}}\langle\tilde{A}_l^\nu z_l^{\nu+1}, \bar{z}_l^{\nu+1}\rangle + \sum_{|n|>\tilde{K}_{\nu+1}}\Omega_n^0 w_n\bar{w}_n,\\
P^\nu &= \check{P}_\nu + \acute{P}_0 + \grave{P}_0,
\end{aligned}$$

and moreover,

$$
\begin{aligned}
&|\omega_{\nu+1}-\omega_\nu|_{\mathcal{O}_\nu} \le c_0\varepsilon_\nu,\\
&\|\tilde A_l^{\nu+1}-\tilde A_l^{\nu}\|_{\mathcal{O}_\nu} \le c_0\varepsilon_\nu,\\
&\|X_{P^\nu}\|_{D_\nu,\mathcal{O}_\nu} \le \varepsilon_\nu,\\
&\breve P_\nu = \breve P^\nu(\theta,I,z^\nu,\bar z^\nu,\xi) + \sum_{|n|>\tilde K_\nu,\alpha_n+\beta_n\ge 1} \breve P_n^\nu(\theta,I,z^\nu,\bar z^\nu,\xi) w_n^{\alpha_n}\bar w_n^{\beta_n},\\
&\acute P_0 = \sum_{\substack{n\ne m\\ \alpha_n+\beta_n,\alpha_m+\beta_m\ge 1\\ \alpha_n+\beta_n+\alpha_m+\beta_m\ge 3}} \acute P_{nm}^0(\xi) w_n^{\alpha_n}\bar w_n^{\beta_n} w_m^{\alpha_m}\bar w_m^{\beta_m},\\
&\grave P_0 = \sum_n O(|w_n|^3),
\end{aligned}
$$

with

$$
\begin{aligned}
&\|\breve P_n^\nu(\theta,I,z^\nu,\bar z^\nu,\xi)\|_{D_\nu,\mathcal{O}_\nu} \le e^{-(|n|-\tilde K_\nu)},\\
&\|\acute P_{nm}^0(\xi)\|_{D_\nu,\mathcal{O}_\nu} \le e^{-|n-m|}.
\end{aligned}
$$

b) *There is a symplectic transformation*

$$\Phi_\nu : \tilde D_\nu \times \mathcal{O}_{\nu+1} \to D_\nu$$

such that

$$H_{\nu+1} = H_\nu \circ \Phi_\nu.$$

Proof. It is sufficient to verify the conditions **C1**)–**C6**) for all $\nu = 0,1,\cdots$, which are easily seen to follow from the following conditions

D1) $\frac{2}{r_\nu - r_{\nu+1}} \ln \frac{1}{\varepsilon_\nu^{\frac54}} \le K_{\nu+1} \le \frac{1}{\varepsilon_\nu^{\frac{1}{4\tau d^4}}}$,

D2) $c_0\gamma^{-d^4}(r_\nu - r_{\nu+1})^{-N}\varepsilon_\nu^{\frac14} \le \frac12$

for all $\nu = 0,1,\cdots$.

We first let ε (hence ε_0) be sufficiently small such that

$$\varepsilon_0 < \min\{\frac{\gamma^{5d^4} r_0^{5N}}{2^{9N+6}c_0^5}\frac{1}{\Psi(r_0)}, \frac{\delta}{2}\},$$

where

$$\Psi(r_0) = \prod_{i=1}^\infty [(r_{i-1}-r_i)^{-5N}]^{(\frac45)^i}$$

which is easily seen to be well-defined. Then

$$c_0\gamma^{-d^4}(r_0 - r_1)^{-N}\varepsilon_0^{\frac14} \le \frac12,$$

i.e., **D2**) holds for $\nu = 0$. Recall that

$$\frac{8}{r_0}\ln\frac{1}{\varepsilon_0^{\frac54}} \le K_0 \le \frac{1}{\varepsilon_0^{\frac{1}{4\tau d^4}}}.$$

We see that **D1**) also holds for $\nu = 0$. Now, for any $\nu \ge 1$, we have by induction that

$$
\begin{aligned}
& c_0\gamma^{-d^4}(r_\nu - r_{\nu+1})^{-N}\varepsilon_\nu^{\frac14} = c_0\gamma^{-d^4}(r_\nu - r_{\nu+1})^{-N}(4c_0\gamma^{-d^4}(r_{\nu-1}-r_\nu)^{-N}\varepsilon_{\nu-1}^{\frac54})^{\frac14}\\
\le\ & (2^{4N+2}c_0^5\gamma^{-5d^4}(r_{\nu-1}-r_\nu)^{-5N}\varepsilon_{\nu-1}^{\frac54})^{\frac14} \le (2^{4N+2}c_0^5\gamma^{-5d^4}\Psi(r_0)\varepsilon_0)^{\frac14(\frac54)^\nu}\\
\le\ & (\frac{r_0^{5N}}{2^{5N+4}})^{\frac14(\frac54)^\nu} \le \frac12,
\end{aligned}
$$

and

$$\frac{2}{r_\nu - r_{\nu+1}} \ln \frac{1}{\varepsilon_\nu^{\frac{5}{4}}} \le \frac{2^{\nu+3}}{r_0} \ln \frac{1}{\varepsilon_0^{(\frac{5}{4})^\nu}} \le \frac{3^\nu}{r_0} \ln \frac{1}{\varepsilon_0} \le 4^\nu K_0$$
$$= K_{\nu+1} \le \frac{1}{\varepsilon_0^{(\frac{6}{5})^\nu \frac{1}{4\tau d^4}}} \le \frac{1}{\varepsilon_\nu^{\frac{1}{4\tau d^4}}},$$

i.e., **D1**) and **D2**) hold true. □

4.2. **Convergence.** Let $\Psi^\nu = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{\nu-1}$, $\nu = 1, 2, \cdots$. Inductively, we have that $\Psi^\nu : \tilde{D}_\nu \times \mathcal{O}_{\nu+1} \to D_0$ and

$$H_0 \circ \Psi^\nu = H_\nu = N_\nu + P^\nu$$

for all $\nu = 1, 2, \cdots$.

Let $\tilde{\mathcal{O}} = \cap_{\nu=0}^\infty \mathcal{O}_\nu$. Applying Lemma 4.1 and standard arguments (e.g. [22, 27]) we conclude that $H_\nu, e_\nu, N_\nu, P^\nu, \Psi^\nu, D\Psi^\nu, \omega_\nu$ converge uniformly on $D(\frac{1}{2}r_0, 0) \times \tilde{\mathcal{O}}$, say to, $H_\infty, e_\infty, N_\infty, P^\infty, \Psi^\infty$, $D\Psi^\infty, \omega_\infty$ respectively. It is clear that

$$N_\infty = e_\infty + \langle \omega_\infty, I \rangle + \sum_{l=1}^\infty \langle A_l^\infty z_l^\infty, \bar{z}_l^\infty \rangle.$$

Since

$$\varepsilon_\nu = 4c_0 \gamma^{-d^4} (r_{\nu-1} - r_\nu)^{-N} \varepsilon_{\nu-1}^{\frac{5}{4}} \le (4c_0 \gamma^{-d^4} \Psi(r_0) \varepsilon_0)^{(\frac{5}{4})^\nu},$$

we have by Lemma 4.1 that

$$X_{P^\infty}|_{D(\frac{1}{2}r_0,0)\times\tilde{\mathcal{O}}} \equiv 0.$$

Let Φ_H^t denote the flow of any Hamiltonian vector field X_H. Since $H_0 \circ \Psi^\nu = H_\nu$, we have that

$$\Phi_{H_0}^t \circ \Psi^\nu = \Psi^\nu \circ \Phi_{H_\nu}^t. \tag{4.1}$$

The uniform convergence of $\Psi^\nu, D\Psi^\nu, X_{H_\nu}$ imply that one can pass the limit in the above to conclude that

$$\Phi_{H_0}^t \circ \Psi^\infty = \Psi^\infty \circ \Phi_{H_\infty}^t,$$

on $D(\frac{1}{2}r_0, 0) \times \tilde{\mathcal{O}}$. It follows that

$$\Phi_{H_0}^t(\Psi^\infty(\mathbb{T}^N \times \{\xi\})) = \Psi^\infty \Phi_{N_\infty}^t(\mathbb{T}^N \times \{\xi\}) = \Psi^\infty(\mathbb{T}^N \times \{\xi\}),$$

for all $\xi \in \tilde{\mathcal{O}}$. Hence $\Psi^\infty(\mathbb{T}^N \times \{\xi\})$ is an embedded invariant torus of the original perturbed Hamiltonian system at $\xi \in \tilde{\mathcal{O}}$. The frequencies $\omega_\infty(\xi)$ associated with $\Psi^\infty(\mathbb{T}^N \times \{\xi\})$ are slightly deformed from the unperturbed ones $\omega(\xi)$, and, the normal behaviors of the invariant tori $\Psi^\infty(\mathbb{T}^N \times \{\xi\})$ are governed by their respective normal frequency matrices A_l^∞. □

4.3. **Measure estimates.** For each $\nu = 1, 2, \cdots$, let

$$\mathcal{R}_k^\nu(\gamma) = \left\{ \xi \in \mathcal{O}_{\nu-1} : |\langle k, \omega_{\nu-1} \rangle^{-1}| > (\frac{K_\nu^\tau}{\gamma})^{d^2} \right\},$$

$$\mathcal{R}_{kl}^\nu(\gamma) = \left\{ \xi \in \mathcal{O}_{\nu-1} : \|(\langle k, \omega_{\nu-1} \rangle I + \tilde{A}_l^{\nu-1})^{-1}\| > (\frac{K_\nu^\tau}{\gamma})^{d^2} \right\},$$

$$\mathcal{R}_{klj}^{\nu+}(\gamma) = \{ \xi \in \mathcal{O}_{\nu-1} : \|(\langle k, \omega_{\nu-1} \rangle I + \tilde{A}_l^{\nu-1} \otimes I + I \otimes \tilde{A}_j^{\nu-1})^{-1}\| > (\frac{K_\nu^\tau}{\gamma})^{d^2} \},$$

and,

$$\mathcal{R}_{klj}^{\nu-}(\gamma) = \{ \xi \in \mathcal{O}_{\nu-1} : \|(\langle k, \omega_{\nu-1} \rangle I + \tilde{A}_l^{\nu-1} \otimes I - I \otimes \tilde{A}_j^{\nu-1})^{-1}\| > (\frac{K_\nu^\tau}{\gamma})^{d^2} \}, \quad |k| + |l - j| \ne 0.$$

Then

$$\mathcal{O}_\nu \subseteq \mathcal{O}_{\nu-1} \backslash ((\bigcup_{|k| \le K_\nu} \mathcal{R}_k^\nu(\gamma)) \bigcup (\bigcup_{|k| \le K_\nu, l \le \tilde{K}_\nu} \mathcal{R}_{kl}^\nu(\gamma)) \bigcup (\bigcup_{|k| \le K_\nu, l, j \le \tilde{K}_\nu} \mathcal{R}_{klj}^{\nu\pm}(\gamma))),$$

for all $\nu = 1, 2, \cdots$. Consider the resonant sets

$$\mathcal{R}^\nu = (\bigcup_{|k|\le K_\nu} \mathcal{R}^\nu_k(\gamma))\bigcup(\bigcup_{|k|\le K_\nu, l\le \tilde K_\nu} \mathcal{R}^\nu_{kl}(\gamma))\bigcup(\bigcup_{|k|\le K_\nu, l,j\le \tilde K_\nu} \mathcal{R}^{\nu\pm}_{klj}(\gamma)).$$

It is clear that

$$\mathcal{O}\setminus\tilde{\mathcal{O}} \subseteq \bigcup_{\nu\ge 1}\mathcal{R}^\nu.$$

Lemma 4.2. *There is a constant $C_1 > 0$ such that*

$$\mathrm{meas}(\mathcal{R}^\nu_k(\gamma)\bigcup\mathcal{R}^\nu_{kl}(\gamma)\bigcup\mathcal{R}^{\nu\pm}_{klj}(\gamma)) \le C_1\frac{\gamma}{K_\nu^{\tau-1}}$$

for all $|k| \le K_\nu$, $l, j \le \tilde K_\nu$, and $\nu = 1, 2, \cdots$.

Proof. The proof follows from arguments in [31]. For simplicity, we only estimate the measures of $\mathcal{R}^{\nu-}_{klj}(\gamma)$. Measure estimates for $\mathcal{R}^\nu_k$, $\mathcal{R}^\nu_{kl}$, and $\mathcal{R}^{\nu+}_{klj}(\gamma)$ can be obtained similarly.

We note that

$$\langle k, \omega_\nu(\xi)\rangle I + \tilde A_l^{\nu-1}\otimes I - I\otimes\tilde A_j^{\nu-1} = \langle k, \omega_{\nu-1}(\xi)\rangle I + \mathrm{diag}(\Omega_{n_1}-\Omega_{m_1},\cdots,\Omega_{n_r}-\Omega_{m_r}) + W(\xi), \tag{4.2}$$

where Ω_{n_1}, Ω_{m_1}, $\cdots$, Ω_{n_r}, Ω_{m_r} are unperturbed frequencies independent of parameters, and $\dim(\tilde A_l^{\nu-1}\otimes I) = r \le d^2$, and $\|W(\xi)\|_{\mathcal{O}_{\nu-1}} \le \varepsilon_0$.

In the case $k = 0$ and $l \ne j$, since by assumption (A2) $|\Omega_{n_j} - \Omega_{m_j}| > \gamma$, $j = 1, \cdots, r$, we have by the standard Neumann series expansion that $\|\tilde A_l^{\nu-1}\otimes I - I\otimes\tilde A_j^{\nu-1})^{-1}\| < \frac{2}{\gamma} < (\frac{K^\tau_{\nu+1}}{\gamma})^{d^2}$ as $\varepsilon \ll 1$, i.e., $\mathcal{R}^{\nu-}_{0lj}(\gamma) = \emptyset$.

In the case $k \ne 0$, we have by Lemma 5.7 that

$$\mathcal{R}^{\nu 2}_{klj} \subset \{\xi\in\mathcal{O}_{\nu-1} : |\det(\langle k, \omega_{\nu-1}(\xi)\rangle I + \tilde A_l^{\nu-1}\otimes I - I\otimes\tilde A_j^{\nu-1})| < (\frac{\gamma}{K_\nu^{\tau-1}})^{d^2}\}.$$

Denote

$$g(\xi) = \det(\langle k, \omega_{\nu-1}(\xi)\rangle I + \tilde A_l^{\nu-1}\otimes I - I\otimes\tilde A_j^{\nu-1}).$$

Then it follows from (4.2) that

$$g(\xi) = \prod_{i=1}^r(\langle k, \omega_{\nu-1}(\xi)\rangle + \Omega_{n_i} - \Omega_{m_i}) + \sum_\alpha a_\alpha(\langle k, \omega_{\nu-1}(\xi)\rangle + \Omega_n - \Omega_m)^\alpha,$$

where $\|a_\alpha\|_{\mathcal{O}_{\nu-1}} \le \varepsilon_0$, the multi-index α runs over the set $\alpha = (\alpha_1, \cdots, \alpha_r)$, $\alpha_i = 0$ or 1, and $\sum_{i=1}^r \alpha_i \le r-1$. Due to the choice of ε_0, we have that

$$|\partial^r_\xi g(\xi)| \ge \prod_{i=1}^r |\langle k, \partial_\xi\omega_\nu(\xi)\rangle| - \varepsilon_0|k|^{r-1} \ge (\delta-\varepsilon_0)|k|^r - \varepsilon_0|k|^{r-1} \ge \frac{\delta}{2}|k|^r.$$

Hence by Lemma 5.8,

$$\mathrm{meas}(\mathcal{R}^{\nu-}_{klj}) \le C_1[(\frac{\gamma}{K_\nu^{\tau-1}})^{d^2}]^{\frac{1}{r}} \le C_1\frac{\gamma}{K_\nu^{\tau-1}}.$$

□

Lemma 4.3.

$$\mathrm{meas}(\mathcal{O}\setminus\tilde{\mathcal{O}}) \le \mathrm{meas}(\bigcup_{\nu\ge1}\mathcal{R}^\nu) = O(\gamma).$$

Proof. By Lemma 4.2, there exists a constant $C_2 > 0$ such that

$$\begin{aligned}\mathrm{meas}(\bigcup_{|k|\le K_\nu, l,j\le\tilde K_\nu}\mathcal{R}^{\nu\pm}_{klj}) &\le \sum_{|k|\le K_\nu, l,j\le\tilde K_\nu} C_1\frac{\gamma}{K_\nu^{\tau-1}}\\ &\le 4C_1\sum_{|k|\le K_\nu}\frac{\gamma}{K_\nu^{\tau-3}} \le C_2\frac{\gamma}{K_\nu^{\tau-b-3}}.\end{aligned}$$

Similarly, there are constants $C_3 > 0$, $C_4 > 0$ such that

$$\text{meas}(\bigcup_{|k|\leq K_\nu} \mathcal{R}_k^\nu(\gamma) \leq C_3 \frac{\gamma}{K_\nu^{\tau-b-3}},$$

and,

$$\text{meas}(\bigcup_{|k|\leq K_\nu, l\leq \tilde{K}_\nu} \mathcal{R}_{kl}^\nu(\gamma)) \leq C_4 \frac{\gamma}{K_\nu^{\tau-b-3}}.$$

Let $\tau \geq b+4$. We have that

$$\begin{aligned} & \text{meas}(\mathcal{O}\setminus\tilde{\mathcal{O}}) \leq \text{meas}(\bigcup_{\nu\geq 1}\mathcal{R}^\nu) \\ = \ & \text{meas}[\bigcup_{\nu\geq 1}((\bigcup_{|k|\leq K_\nu}\mathcal{R}_k^\nu)\bigcup(\bigcup_{|k|\leq K_\nu, l\leq\tilde{K}_\nu}\mathcal{R}_{kl}^\nu)\bigcup(\bigcup_{|k|\leq K_\nu, l,j\leq \tilde{K}_\nu}\mathcal{R}_{klj}^{\nu\pm}))] \\ = \ & O(\sum_{\nu\geq 0}\frac{\gamma}{K_{\nu+1}}) = O(\gamma). \end{aligned}$$

□

This completes the measure estimate.

5. Appendix

Lemma 5.1.

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

Proof. Since $(FG)_{kl\alpha\beta} = \sum_{k',l',\alpha',\beta'} F_{k-k',l-l',\alpha-\alpha',\beta-\beta'}G_{k'l'\alpha'\beta'}$, we have

$$\begin{aligned} \|FG\|_{D(r,s),\mathcal{O}} &= \sup_{\substack{\|w\|<s\\ \|\bar{w}\|<s}} \sum_{k,l,\alpha,\beta} |(FG)_{kl\alpha\beta}|s^{2l}|w^\alpha||\bar{w}^\beta|e^{|k|r} \\ &\leq \sup_{\substack{\|w\|<s\\ \|\bar{w}\|<s}} \sum_{k,l,\alpha,\beta}\sum_{k',l',\alpha',\beta'} |F_{k-k',l-l',\alpha-\alpha',\beta-\beta'}G_{k'l'\alpha'\beta'}|s^{2l}|w^\alpha||\bar{w}^\beta|e^{|k|r} \\ &\leq \|F\|_{D(r,s),\mathcal{O}}\|G\|_{D(r,s),\mathcal{O}}. \end{aligned}$$

□

Lemma 5.2. (*Generalized Cauchy inequalities*)

$$\begin{aligned} \|F_\theta\|_{D(r-\sigma,s)} &\leq \frac{1}{\sigma}\|F\|_{D(r,s)}, \\ \|F_I\|_{D(r,\frac{1}{2}s)} &\leq \frac{4}{s^2}\|F\|_{D(r,s)}, \\ \|F_w\|_{D(r,\frac{1}{2}s)} &\leq \frac{2}{s}\|F\|_{D(r,s)}, \\ \|F_{\bar{w}}\|_{D(r,\frac{1}{2}s)} &\leq \frac{2}{s}\|F\|_{D(r,s)}. \end{aligned}$$

Proof. See [27]. □

Let $\{\cdot,\cdot\}$ denote the Poisson bracket of smooth functions:

$$\{F,G\} = \langle\frac{\partial F}{\partial I},\frac{\partial G}{\partial\theta}\rangle - \langle\frac{\partial F}{\partial\theta},\frac{\partial G}{\partial I}\rangle + \mathrm{i}\sum_n(\frac{\partial F}{\partial w_n}\frac{\partial G}{\partial\bar{w}_n} - \frac{\partial F}{\partial\bar{w}_n}\frac{\partial G}{\partial w_n}).$$

Lemma 5.3. *There exists a constant $c>0$ such that if*

$$\|F_n\|_{D(r,s)} < e^{-|n|},\ \|G\|_{D(r,s)} < \varepsilon,$$

then

$$\|\{F_n,G\}\|_{D(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-2}\|F_n\|_{D(r,s)}\|G\|_{D(r,s)} \leq c\sigma^{-1}s^{-2}\varepsilon e^{-|n|}.$$

Proof. By Lemmas 5.1, 5.2,

$$
\begin{aligned}
\|\langle F_{n_I}, G_\theta\rangle\|_{D(r-\sigma,\frac{1}{2}s)} &< 4\sigma^{-1}s^{-2}\|F_n\|\cdot\|G\|,\\
\|\langle F_{n_\theta}, G_I\rangle\|_{D(r-\sigma,\frac{1}{2}s)} &< c\sigma^{-1}s^{-2}\|F_n\|\cdot\|G\|,\\
\|\sum_m F_{n_{w_m}} G_{\bar w_m}\|_{D(r,\frac{1}{2}s)} &\le \sum_m \|F_{n_{w_m}}\|_{D(r,\frac{1}{2}s)}\|G_{\bar w_m}\|_{D(r,\frac{1}{2}s)}\\
&\le \|F_{n_w}\|_{D(r,\frac{1}{2}s)}\|G_{\bar w}\|_{D(r,\frac{1}{2}s)}\\
&\le 4s^{-2}\|F_n\|\cdot\|G\|,\\
\|\sum_m F_{n_{\bar w_m}} G_{w_m}\|_{D(r,\frac{1}{2}s)} &\le \sum_m \|F_{n_{\bar w_m}}\|_{D(r,\frac{1}{2}s)}\|G_{w_m}\|_{D(r,\frac{1}{2}s)}\\
&\le \|F_{n_{\bar w}}\|_{D(r,\frac{1}{2}s)}\|G_w\|_{D(r,\frac{1}{2}s)}\\
&\le 4s^{-2}\|F_n\|\cdot\|G\|.
\end{aligned}
$$

It follows that

$$\|\{F_n, G\}\|_{D(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-2}\|F_n\|_{D(r,s)}\|G\|_{D(r,s)} \le c\sigma^{-1}s^{-2}\varepsilon e^{-|n|}.$$

□

Lemma 5.4. *There exists a constant $c>0$ such that if*

$$\|X_F\|_{D(r,s)} < \varepsilon',\ \|X_G\|_{D(r,s)} < \varepsilon''$$

for some $\varepsilon', \varepsilon'' > 0$, then

$$\|X_{\{F,G\}}\|_{D(r-\sigma,\eta s)} < c\sigma^{-1}\eta^{-2}\varepsilon'\varepsilon'',$$

for any $0<\sigma<r$ and $0<\eta\ll 1$. In particular, if $\eta\sim\varepsilon^{\frac{1}{4}}$, $\varepsilon'\sim\varepsilon, \varepsilon''\sim\varepsilon^{\frac{3}{4}}$, then

$$\|X_{\{F,G\}}\|_{D(r-\sigma,\eta s)} \sim \varepsilon^{\frac{5}{4}}.$$

Proof. See [16]. □

The following lemmas can be found in the Appendix of [12, 32]. □

Lemma 5.5. *Let $\mathcal{O}$ be a compact set in $\mathbb{R}^N$ for which small divisor conditions hold. Suppose that $\omega(\xi)$ are C^{d^2} Whitney-smooth functions in $\xi\in\mathcal{O}$ with derivative bounded by L and $f(\xi)$ are C^{d^2} Whitney-smooth functions in $\xi\in\mathcal{O}$ with $C_W^{d^2}$ norm bounded by L. Then*

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

is C^{d^2} Whitney-smooth in $\mathcal{O}$ with

$$\|g\|_{\mathcal{O}} < c\gamma^{-d^4}K_+^{\tau d^4}L.$$

Lemma 5.6. *Let $\mathcal{O}$ be a compact set in $\mathbb{R}^N$ for which small divisor conditions hold. Suppose that $A_l(\xi)$, $R_l(\xi)$ are respectively C^{d^2} Whitney-smooth matrices and vectors, and $\omega(\xi)$ is a C^{d^2} Whitney-smooth function with derivatives bounded by L. Then*

$$F_l(\xi) = M^{-1}R_l(\xi)$$

is C^{d^2} Whitney-smooth with

$$\|F_l\|_{\mathcal{O}} \le c\gamma^{-d^4}K_+^{\tau d^4}L,$$

where M stands for either $\langle k,\omega\rangle I + \tilde A_l$ or $\langle k,\omega\rangle I \pm \tilde A_l\otimes I \pm I\otimes \tilde A_j$.

Lemma 5.7. *Let M be a $r \times r$ non-singular matrix with $\|M\| < |k|$. Then*

$$\{\xi : \|M^{-1}\| > h\} \subseteq \{\xi : |\det M| < \frac{|k|^{r-1}}{h}\}.$$

Lemma 5.8. *Suppose that $g(u)$ is a C^m function on the closure $\bar{I}$, where $I \subset \mathbb{R}$ is a finite interval. Let $I_h = \{u : |g(u)| < h\}$, $h > 0$. If for some constant $d > 0$, $|g^{(m)}(u)| \geq d$ for all $u \in I$, then* $\mathrm{meas}(I_h) \leq ch^{\frac{1}{m}}$, *where $c = 2(2 + 3 + \cdots + m + d^{-1})$.*

References

[1] S. Aubry, Anti–integrability in dynamical and variational problems, *Physica D* **86** (1995), 284–296

[2] S. Aubry, Breathers in nonlinear lattices: existence, linear stability and quantization, *Physica D* **103** (1997), 201–250

[3] D. Bambusi, Exponential stability of breathers in Hamiltonian networks of weakly coupled oscillators, *Nonlinearity* **9** (1996), 433–457

[4] D. Bambusi and D. Vella, Quasi periodic breathers in Hamiltonian lattices with symmetries, *Discrete Cont. Dyn. Sys. Ser. B* **2** (2002), 389–399

[5] J. Bourgain, Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, *Annals of Math.* **148** (1998), 363–439

[6] J. Bourgain, Construction of periodic solutions of nonlinear wave equations in higher dimension, *Geom. Funct. Anal.* **5** (1995), 629–639

[7] J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE., *Int. Math. Res. Notices* (1994), 475–497

[8] J. Bourgain, Green's function estimates for lattice Schrödinger operators and applications. *Annals of Mathematics Studies,* **158,** Princeton University Press, Princeton, NJ, 2005

[9] J. Bourgain, Nonlinear Schrödinger equations, *Park City Series* **5,** American Mathematical Society, Providence, Rhode Island, 1999

[10] O. M. Braun and Y. S. Kivshar, Nonlinear dynamics of the Frenkel-Kontorova model, Phys. Reports **306** (1998), 1-108

[11] A. Carpio, L. L. Bonilla, Oscillatory wave fronts in chains of coupled nonlinear oscillators, *Phys. Rev. E* **67** (2003), 656-621

[12] L. Chierchia and J. You, KAM tori for 1D nonlinear wave equations with periodic boundary conditions, *Commun. Math. Phys.* **211** (2000), 498-525

[13] W. Craig and C. E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, *Comm. Pure. Appl. Math.* **46** (1993), 1409-1498

[14] T. Dauxois, M. Peyrard and C. R. Willis, Localised breather-like solution in a discrete Klein-Gordon model and application to DNA, *Physica D* **57** (1992), 267-282

[15] J. Fröhlich, T. Spencer and C. E. Wayne, Localization in disordered nonlinear dynamical systems, *J. Stat. Phys.* **42** (1986), 247–274

[16] J. Geng and J. You, A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions, *J. Differentail Equations* **209**(2005), 1-56

[17] J. Geng and J. You, A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces, *Commun. Math. Phys.* **262**(2006), 343-372

[18] J. Geng and Y. Yi, Quasi-periodic solutions in a nonlinear Schrödinger equation, *submitted*

[19] M. Johansson and S. Aubry, Existence and stability of quasiperiodic breathers in the discrete nonlinear Schrödinger equation, *Nonlinearity* **10** (1997), 1151–1178

[20] O. Kresse and L. Truskinovsky, Mobility of lattice defects: discrete and continuum approaches, *J. Mech. Phys. Solids,* **51** (2003), 1305-1332

[21] S. B. Kuksin, Hamiltonian perturbations of infinite–dimensional linear systems with an imaginary spectrum, *Funct. Anal. Appl.* **21** (1987), 192-205

[22] S. B. Kuksin, Nearly integrable infinite dimensional Hamiltonian systems, Lecture Notes in Mathematics, **1556**, Springer, Berlin, 1993

[23] S. B. Kuksin and J. Pöschel, Invariant Cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation, *Annals of Math.* **143** (1996), 149-179

[24] Y. Li and Y. Yi, Persistence of lower dimensional tori of general types in Hamiltonian systems, *Trans. Amer. Math. Soc.* **357** No. 4 (2004), 1565-1600

[25] R. S. Mackay and S. Aubry, Proof of existence of breathers of time–reversible or Hamiltonian networks of weakly coupled oscillators, *Nonlinearity* **7** (1994), 1623–1643

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

[27] J. Pöschel, A KAM Theorem for some nonlinear partial differential equations, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **23** (1996), 119-148
[28] J. Pöschel, Small divisors with spatial structure in infinite dimensional Hamiltonian systems, *Comm. Math. Phys.* **127** (1990), 351–393
[29] E. W. Prohofsky, K. C. Lu, L. L. Van Zandt, and B. F. Putnam, Breathing models and induced resonant melting of the double helix, *Phys. Lett. A* **70** (1979), 492-494
[30] C. E. Wayne, Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, *Commun. Math. Phys.* **127** (1990), 479-528
[31] J. Xu and J. You, Persistence of lower-dimensional tori under the first Melnikov's non-resonance condition, *J. Math. Pures Appl.* **80** (2001), 1045-1067
[32] J. You, Perturbation of lower dimensional tori for Hamiltonian systems, *J. Differential Equations* **152** (1999), 1-29
[33] X. Yuan, Construction of quasi–periodic breathers via KAM technique, *Commun. Math. Phys.* **226** (2002), 61–100
[34] C. T. Zhang, Soliton excitations in deoxyribonucleic acid (DNA) double holices, *Phys. Rev. A* **35** (1987), 886-891

J. Geng: Department of Mathematics, Nanjing University, Nanjing 210093, P.R.China
E-mail address: `jianshenggeng@yahoo.com (J. Geng)`

Y. Yi: School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332
E-mail address: `yi@math.gatech.edu (Y. Yi)`