QUASI-PERIODIC BREATHERS IN NEWTON'S CRADLE

CHUANFANG GE, JIANSHENG GENG, AND YINGFEI YI

Abstract. We consider the following parametrized Newton's cradle lattice with Hertzian interactions:

$$\ddot{x}_n + \beta_n^2 x_n = V'(x_{n+1} - x_n) - V'(x_n - x_{n-1}), \quad n \in \mathbb{Z},$$

where for a fixed integer $b \geq 0$, $\{\beta_n : |n| \leq b\}$ are positive parameters, $\{\beta_n : |n| > b\}$ are given positive constants, and $V(x) = \frac{1}{1+\alpha}|x|^{1+\alpha}$ is the Hertzian potential for a fixed real number $\alpha > \alpha_* := 12b + 25$. Corresponding to a large Lebesgue measure set of $(\beta_j)_{|j|\leq b} \in \mathbb{R}_+^{2b+1}$, we show the existence of a family of small amplitudes, linearly stable, quasi-periodic breathers for the Newton's cradle lattice which are quasi-periodic in time with $2b+1$ frequencies and localized in space with rate $\frac{1}{|n|^{1+\alpha}}$ as $|n| \gg 1$. To overcome obstacles in applying KAM method due to the finite smoothness of V especially when α is not an integer and to obtain a sharp estimate of the localization rate of the quasi-periodic breathers, the proof of our result uses the Jackson-Moser-Zehnder (JMZ) analytic approximation technique but with refined estimates on error bounds, depending on the smoothness and dimension, which provide crucial controls on the convergence of KAM iterations.

1. Introduction and main result

In this paper, we consider the existence of time quasi-periodic breathers in the following Newton's cradle lattices with Hertzian interaction:

$$\ddot{x}_n + \beta_n^2 x_n = V'(x_{n+1} - x_n) - V'(x_n - x_{n-1}), \quad x_n \in \mathbb{R},\ n \in \mathbb{Z}, \tag{1.1}$$

where $\beta_n's$ are positive constants and $V(x) = \frac{1}{1+\alpha}|x|^{1+\alpha}$, $\alpha > 0$, is the Hertzian interaction potential. We note that V is only of the class $C^{1+\alpha}$ for non-integer exponent α. As the uncoupled systems corresponding to (1.1) are harmonic oscillators, we treat a part of the natural frequencies β_n's as parameters in order to find quasi-periodic breathers. More precisely, let b be a fixed natural number and $\mathcal{O} \subseteq \mathbb{R}_+^{2b+1}$ be a bounded closed region. We regard

$$\beta := (\beta_j)_{|j|\leq b} \in \mathcal{O} \tag{1.2}$$

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

Key words and phrases. KAM theory; Birkhoff normal form; Newton's cradle, Jackson-Moser-Zehnder's analytic approximation.

The second author was partially supported by NSFC grant 11971012. The third author was partially supported by NSERC discovery grants 1257749 and RGPIN-2020-04451, a faculty development grant from University of Alberta, and a Scholarship from Jilin University.

as parameters, and let the remaining natural frequencies β_n, $|n| \geq b+1$, be fixed constants. For each set of the parameters, equations (1.1) form a Hamiltonian lattice with the Hamiltonian

$$H = \sum_{n\in\mathbb{Z}} \frac{1}{2}\dot{x}_n^2 + \frac{\beta_n^2}{2}x_n^2 + V(x_{n+1} - x_n). \tag{1.3}$$

The Newton's cradle lattice is a simplified model for granular chains consisting of a chain of identical spherical beads which are attached to linear pendula and interact nonlinearly via Hertz's contact forces. By transforming the Hamiltonian system with $\beta_n \equiv 1$, $n \in \mathbb{Z}$, into a discrete p-Schrödinger equation and using numerical simulations, James in [18] showed the existence of time periodic breathers when $\alpha \geq \frac{3}{2}$. Stability of these periodic breathers and their significance in physics were discussed in [19, 20]. These works raise both mathematically and physically interesting questions on the existence of time quasi-periodic breathers in the Newton's cradle lattices with Hertzian potentials.

Breathers and quasi-periodic breathers, as important coherent, localized structures or waves, have been largely found in Hamiltonian lattices such as discrete nonlinear Schrödinger equations, Fermi-Pasta-Ulam chains, Frenkel-Kontorova (FK) etc. We refer the readers to [1]-[7], [14, 16, 26] and references therein for developments in the subject. However, these existing works all deal with analytic Hamiltonians, and to the best of our knowledge, no result is yet known on the existence of quasi-periodic breathers for non-analytic lattice Hamiltonian systems. We remark that finite-dimensional KAM theory for full dimensional, quasi-periodic, invariant tori is well established in the non-analytic setting ([17, 21, 22, 23, 27, 28]). Based on Moser's theory, lower dimensional, quasi-periodic, invariant tori with a finite number of normal frequencies is also known for finite-smooth Hamiltonians ([12]). For non-analytic Hamiltonian PDEs, a Nash-Moser iteration method is developed in [8]-[10] to show the existence of lower dimensional, quasi-periodic, invariant tori in finite-smooth, defocusing, nonlinear Schrödinger equations with periodic boundary conditions. A lower bound of required smoothness in the Nash-Moser iterations is estimated in [13] for a class of nonlinear Schrödinger equations with external frequencies. While the Nash-Moser iteration method is an important technique in studying quasi-periodic solutions in finite-smooth, infinite-dimensional Hamiltonian systems including Hamiltonian lattices, it seems to have certain disadvantage for capturing sharp localization rates of quasi-periodic breathers (see Remark 1.1 (2) below).

In this paper, we use KAM method together with a refined Jackson-Moser-Zehnder (JMZ) analytic approximation technique to study quasi-periodic breathers for the parametrized Newton's cradle lattice (1.1). Our main result is stated as follows.

Main Theorem. *Consider* (1.1) *with* $\alpha > \alpha_* := 12b+25$ *and let parameters* β *lie in a bounded closed region* $\mathcal{O} \subseteq \mathbb{R}_+^{2b+1}$. *Then there exists a family*

$\mathcal{O}_\gamma \subset \mathcal{O}$, $0 < \gamma \ll 1$, with $\mathrm{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) = O(\gamma)$ *as* $\gamma \to 0$, *constants* $\varepsilon = \varepsilon(b, \gamma, \alpha) > 0$, $2 < \mathfrak{s}_* = \mathfrak{s}_*(b, \alpha) < \alpha + 1$, *and a Whitney smooth map* $\tilde{\omega} : \mathcal{O}_\gamma \to \mathbb{R}^{2b+1}$ *with* $\tilde{\omega} = \beta + O(\varepsilon)$, *such that for each* $\beta \in \mathcal{O}_\gamma$, (1.1) *admits a linearly stable, quasi-periodic breather* $x(t) = (x_n(t))_{n\in\mathbb{Z}}$ *with the following properties:*

1. *For each* $n \in \mathbb{Z}$, $x_n(t) = X_n(\tilde{\omega}(\beta)t)$, *where* $X_n : \mathbb{T}^{2b+1} \to \mathbb{R}$ *is a* $C^{\mathfrak{s}}$ *function for any* $2 \le \mathfrak{s} < \mathfrak{s}_*$;
2. $\sum_{n\in\mathbb{Z}} |x_n(0)|\langle n\rangle^{1+\alpha} \le \varepsilon$, *where* $\langle n\rangle = \max\{1, |n|\}$;
3. $x_n(t) \sim \frac{1}{|n|^{1+\alpha}}$ *as* $|n| \gg 1$ *uniformly in* t.

Remark 1.1. (1) The Main Theorem above actually holds for any finite number of oscillating sites rather than the $2b + 1$ adjacent sites considered, i.e., the oscillating lattice set $\mathcal{J}$ can be replaced by any set of N numbers in $\mathbb{Z}$ for a fixed natural number N and $\mathcal{O}$ can be replaced by any bounded closed region in $\mathbb{R}^N$. We choose $N = 2b + 1$ and $\mathcal{J} = \{n \in \mathbb{Z} : |n| \le b\}$ in this paper in order to give an explicit lower bound for the smoothness order α. The proof for the general situation will be similar to that for the Main Theorem above, except that, in the normal form reduction, the form of the finite term in the perturbation P in (3.12) needs to be adjusted accordingly.

(2) The Nash-Moser approach to KAM theory developed in [9, 10] can be also used to obtain quasi-periodic breathers for the parametrized Newton's cradle lattice (1.1). However, in order to apply the interpolation inequality to compensate the loss of regularities during iterations, this approach requires the consideration of solutions in the Banach space

$$\mathcal{H}^{s_1,s_2} = \{x = (x_n) = (\sum_{l\in\mathbb{Z}^{2b+1}} x_{n,l}e^{\mathrm{i}l\cdot\varphi}) : \sum_{n,l} |x_{n,l}|^2\langle n\rangle^{2s_1}\langle l\rangle^{2s_2} < \infty\}$$

where $\varphi \in \mathbb{T}^{2b+1}$ and $s_1 + s_2 = 1 + \alpha$ in which s_1, s_2 are the order of localization, smoothness of quasi-periodic breathers, respectively. A calculation shows that the essential smoothness order required to carry over this approach is $s_2 \ge 13b + 16$, implying that $s_1 \ge \alpha - 13b - 15$, i.e., the estimated localization rate of quasi-periodic breathers is at most $1/|n|^{\alpha-13b-15}$. As stated in the Main Theorem above, our approach yields the optimal localization rate $1/|n|^{1+\alpha}$ of quasi-periodic breathers.

(3) For each set of parameter β, (1.1) becomes an infinite-dimensional Hamiltonian system in the Banach space

$$\mathcal{H} = \{x = (x_n) : \sum_{n\in\mathbb{Z}} |x_n|\langle n\rangle^{\alpha+1} < \infty\}.$$

The Main Theorem above says that each quasi-periodic breather obtained lies in an invariant $(2b + 1)$-torus of $\mathcal{H}$ of the class $C^{\mathfrak{s}}$ for any $2 \le \mathfrak{s} < \mathfrak{s}_*$.

By reducing the Hamiltonian (1.3) into an action-angle-normal form, our approach of proving the Main Theorem is to construct a sequence of analytic approximations to the non-analytic perturbation of (1.3) such that only the analytic approximations participates in each KAM iteration and the loss of regularity of perturbation after each KAM iteration is recovered by shrinking the complex domain of the angle variables.

To carry out the KAM iterations for the non-analytic, infinite-dimensional Hamiltonian (1.3), we will prove an analytic approximation lemma which refines the JMZ Approximation Lemma in finite dimensions (e.g. [12, Lemma 2.1]) by giving an explicit error estimate of such an approximation in terms of the order of smoothness and the dimension of the domain of definition of the function under consideration. We note that such an explicit estimate is necessary in the infinite-dimensional situation, simply because at each KAM step one needs to consider a truncation of the perturbation, though defined in a finite-dimensional domain but with increased dimension from the one at the previous KAM step. The special structure of Hamiltonian (1.3) allows the consideration of a truncation of the perturbation at each KAM step by including two more normal variables from the one at the previous KAM step.

We also remark that the approach adopted in this paper should be extendable to study quasi-periodic breathers in non-parametrized Newton's cradle lattices with anharmonic local potentials which provide a rich set of natural frequencies, for instance in Hamiltonian lattices of the form

$$H = \sum_{n\in\mathbb{Z}} \frac{1}{2}\dot{x}_n^2 + \frac{1}{2}x_n^2 + \frac{1}{4}x_n^4 + V(x_{n+1} - x_n)$$

considered in [11] for periodic breathers. We will study this problem in a seperate work.

The rest of the paper is organized as follows. In Section 2, we give a refined JMZ analytic approximation lemma (Lemma 2.1) with concrete error estimates on the smoothness and dimension. In Section 3, we introduce action-angle-normal variables and derive a normal form for the Hamiltonian (1.3) along with its real analytic approximations. In Section 4, we sketch our KAM scheme for one KAM step along with the construction of the symplectic transformation for the approximated Hamiltonian at the same step. In Section 5, we give necessary estimates for the validity of KAM iterations and prove an iteration lemma. Measure estimates and the convergence of KAM iterations are shown in Section 6 to complete the proof of the Main Theorem.

Through the rest of the paper, for any complex number or vector w, $\bar{w}$ stands for its complex conjugate. For simplicity, we will use the same symbol $|\cdot|$ to denote the absolute value of a complex number and the norm of a vector space. However, if an integer vector is considered, then $|\cdot|$ is specified as the l^1-norm, and, if a matrix $A = (a_{ij})_{m\times n}$ is considered, then its norm is specified by

$$|A| = \max\{\sup_{1\le j\le n} \sum_{1\le i\le m} |a_{ij}|, \sup_{1\le i\le m} \sum_{1\le j\le n} |a_{ij}|\}.$$

For any complex vector or sequence $w = (w_n)$, we use $\|\cdot\|$ to denote the following weighted norm:

$$\|w\| =: \sum_n |w_n|\langle n\rangle^{1+\alpha}.$$

For any function $f \in C^{l,1}(X\times\mathcal{O})$, where X is a compact region in a Banach space, $l \in \mathbb{R}_+$, and $\mathcal{O}$ is the parameter region in (1.2), we will adopt the short notions

$$|f|_{C^l(X),\mathcal{O}} =: |f|_{C^{l,1}(X\times\mathcal{O})}, \qquad |f|_{X,\mathcal{O}} =: |f|_{C^{0,1}(X\times\mathcal{O})}.$$

If f is valued in a vector space, then the above short notions simply mean to apply the function norms to its components followed by the norm of the vector space. We also use the symbol "$\preccurlyeq$" to represent "$\le$" up to a constant that is independent of the iteration process. Finally, for any Hamiltonian function G, we denote by X_G the Hamiltonian vector field associated with G and by ϕ^t_G the corresponding Hamiltonian flow.

2. Analytic approximation

In carrying out KAM iterations for a non-analytic, infinite-dimensional Hamiltonian like (1.3), we provide in this section a refined, finite-dimensional JMZ analytic approximation lemma which gives an explicit estimate on the error of the approximation in terms of the order of smoothness and dimension. Such approximations will be applied to truncated Hamiltonian perturbations and participate in the KAM iterations instead of the truncated Hamiltonian perturbations.

Consider the Banach space

$$\ell = \left\{ x = \{x_n\}_{|n|>b} : \|x\| = \sum_{|n|>b} |x_n|\langle n\rangle^{\alpha+1} < \infty \right\}.$$

For given $0<\rho,\sigma<1$ sufficiently small and each $k\in\mathbb{Z}_+$, also consider the following subsets of $\mathbb{T}^{2b+1}\times\mathbb{R}^{2b+1}\times\ell\times\ell$ and its complexification:

$$\begin{aligned}
D_\rho &= \left\{ (x,y) = (\{x_n\},\{y_n\}) : \begin{array}{ll} x_n\in\mathbb{T},\ |n|\le b; & \sum_{|n|>b}|x_n|\langle n\rangle^{1+\alpha}\le\rho, \\ \sum_{|n|\le b}|y_n|\langle n\rangle^{1+\alpha}\le\rho; & \sum_{|n|>b}|y_n|\langle n\rangle^{1+\alpha}\le\rho \end{array} \right\}, \\
\Omega^k &= \left\{ (x,y) = (\{x_n\},\{y_n\}) : \begin{array}{l} x_n,y_n\in\mathbb{R},\ n\in\mathbb{Z};\quad x_n=y_n=0,\ \text{if } |n|>b+k, \\ \sum_{n\in\mathbb{Z}}|x_n|\langle n\rangle^{1+\alpha}+\sum_{n\in\mathbb{Z}}|y_n|\langle n\rangle^{1+\alpha}<\infty \end{array} \right\},
\end{aligned}$$

$$\Delta_\sigma^k = \left\{ (x,y)=(\{x_n\},\{y_n\}) : \begin{array}{ll} (\mathrm{Re}x, \mathrm{Re}y) \in D_\rho, & \\ \sum\limits_{|j|\le b} |\mathrm{Im}\, x_j| \langle j\rangle^{1+\alpha} \le \sigma; & \sum\limits_{b<|j|} |\mathrm{Im}\, x_j| \langle j\rangle^{1+\alpha} \le \sigma, \\ \sum\limits_{|j|\le b} |\mathrm{Im}\, y_j| \langle j\rangle^{1+\alpha} \le \sigma; & \sum\limits_{b<|j|} |\mathrm{Im}\, y_j| \langle j\rangle^{1+\alpha} \le \sigma, \\ |\mathrm{Im}\, x_j| = |\mathrm{Im}\, y_j| = 0, & \text{if } |j| \ge b+k+1 \end{array} \right\}.$$

Let ϕ_k be a radially symmetric, $C^\infty(\Omega^k)$ cut-off function satisfying

$$\phi_k(\varsigma) = 1, \quad \forall |\varsigma| \le 1, \quad \textit{and} \quad \phi_k(\varsigma) = 0, \quad \forall |\varsigma| \ge 2.$$

Our approximation lemma is stated as follows.

Lemma 2.1. *Let $f \in C^{l,1}(D_\rho \times \mathcal{O})$ for some $l > 2$. For each $k \in \mathbb{Z}_+$ and $\sigma > 0$, consider*

$$f_{\sigma,k}(\psi,\beta) := K_k^\sigma * f^k(\psi,\beta) = \sigma^{-(2b+1+2k)} \int_{\Omega^k} K_k(\frac{\psi - y}{\sigma}) f^k(y,\beta) dy,$$

where $f^k = f|_{\Omega^k \times \mathcal{O}}$ and $K_k = \hat{\phi_k}$ is the Fourier transform of ϕ_k. Then the following holds.

(1) *$f_{\sigma,k}(\psi,\beta)$ is real analytic in $\psi \in \Delta_\sigma^k$ and smooth in $\beta \in \mathcal{O}$. Moreover, there is a constant $c_1 > 0$, depending only on l and b, such that*

$$\begin{aligned} &\left| f_{\sigma,k} - \sum_{|\varrho| \le l-1} \frac{\partial^\varrho f^k(\mathrm{Re}\psi)}{\varrho!} (\mathrm{iIm}\psi)^\varrho \right|_{C^1(\Delta_\sigma^k),\mathcal{O}} \\ \le\ & c_1 |f|_{C^l(D_\rho),\mathcal{O}} \sigma^{l-1} (2b+1+2k)^{l+1}, \end{aligned}$$

where $\varrho = ((\varrho_j)_{|j|\le b+k}, \{0\})$ with $(\varrho_j)_{|j|\le b+k} \in \mathbb{N}^{2b+1+2k}$ and $\varrho! := \prod\limits_{|j|\le b+k} \varrho_j!$.

(2) *If*

$$\left| f^{k+1} - f^k \right|_{C^l(D_\rho),\mathcal{O}} \le c_0, \qquad k \in \mathbb{Z}_+ \tag{2.4}$$

for some constant $c_0 > 1$ depending only on l and b, then there is a constant $c_2 > 0$, depending only on l and b, such that

$$\begin{aligned} &\left| f_{\sigma_{k+1},k+1} - f_{\sigma_k,k} \right|_{C^1(\Delta_{\sigma_{k+1}}^{k+1}),\mathcal{O}} \\ \le\ & c_2 |f|_{C^l(D_\rho),\mathcal{O}} \sigma_k^{l-1} (2b+3+2k)^{l+1}, \quad k \in \mathbb{Z}_+, \end{aligned}$$

for any decreasing sequence $\{\sigma_k\}$ of positive numbers. Consequently, if $\{\sigma_k\}$ is a supper-exponentially decreasing sequence, then $\{f_{\sigma_k,k}\}$ converges uniformly to f, i.e., in $C^{0,1}(D_\rho \times \mathcal{O})$.

Proof. For simplicity, we only prove the lemma without treating the C^1-norm $|\cdot|_{\mathcal{O}}$ in parameters. The proof involving $|\cdot|_{\mathcal{O}}$-norm can be carried out similarly by taking into consideration of some straightforward estimates with respect to $\beta \in \mathcal{O}$. For simplicity, we use c to denote any intermediate positive constant which depends only on l and b.

Let $k \in \mathbb{Z}_+$ and $\sigma > 0$ be given. The real analyticity of $f_{\sigma,k}(\psi)$ on the complex neighborhood Δ_σ^k of $\mathbb{T}^{2b+1}$ is a standard fact (see e.g. [27]).

Write $l = p + \delta$, where p an integer and $\delta \in [0, 1)$. Then any $g \in C^l(D_\rho)$ has Taylor expansion of the form

$$g(\psi + \eta) = \sum_{0 \le |\varrho| \le p} \frac{1}{\varrho!} \eta^\varrho \partial^\varrho g(\psi) + R_g(\psi, \eta), \tag{2.5}$$

where

$$R_g(\psi, \eta) = \sum_{|\varrho| = p} \frac{p}{\varrho!} \eta^\varrho \int_0^1 (1 - \mu)^{p-1} [\partial^\varrho g(\psi + \mu\eta) - \partial^\varrho g(\psi)] d\mu.$$

Let $k \in \mathbb{Z}_+$ be given. For any polynomial P, since K_k is of Schwartz class, we have

$$\int_{\Omega^k} K_k(x) P(x) dx = P(0).$$

It follows from the Cauchy integral formula that

$$\int_{\Omega^k} K_k(x + \mathrm{i}\eta) P(x) dx = P(-\mathrm{i}\eta). \tag{2.6}$$

Using the change of variables $\xi = \frac{1}{\sigma}\mathrm{Re}(\psi - y) = \frac{1}{\sigma}\mathrm{Re}\,\psi - \frac{1}{\sigma}y$ and applying (2.5) and (2.6), we can re-write the function $f_{\sigma,k}$ as

$$\begin{aligned} f_{\sigma,k}(\psi) &= \int_{\Omega^k} K_k(\xi + \mathrm{i}\frac{1}{\sigma}\mathrm{Im}\,\psi) f(\mathrm{Re}\,\psi - \sigma\xi) d\xi \\ &= \sum_{0 \le |\varrho| \le p} \frac{1}{\varrho!} (\mathrm{iIm}\,\psi)^\varrho \partial^\varrho f^k(\mathrm{Re}\,\psi) + \hat{R}_{f^k}(\psi, \sigma), \end{aligned} \tag{2.7}$$

where

$$\hat{R}_{f^k}(\psi, \sigma) = \int_{\Omega^k} K_k(\xi + \mathrm{i}\frac{1}{\sigma}\mathrm{Im}\,\psi) R_{f^k}(\mathrm{Re}\,\psi, \sigma\xi) d\xi.$$

Obviously, the first term of in (2.7) is independent of σ. As for the estimate of the second term, since

$$\left| \partial^\varrho f^k(\mathrm{Re}\,\psi - \sigma\xi) - \partial^\varrho f^k(\mathrm{Re}\,\psi) \right| \le |f|_{C^l(D_\rho)} |\sigma\xi|^\delta,$$

we have

$$|\hat{R}_{f^k}(\psi, \sigma)| \le |f|_{C^l(D_\rho)} \sigma^l \sum_{|\varrho| = p} \frac{1}{\varrho!} \sup_{|\eta| < 1} \int_{\Omega^k} |K_k(\xi + \mathrm{i}\eta)\xi^\varrho| \, |\xi|^\delta d\xi. \tag{2.8}$$

It follows from (2.8) and the Cauchy integral formula that

$$\begin{aligned} &\left| f_{\sigma,k}(\psi) - \sum_{0 \le |\xi| \le p} \frac{\partial^\xi f^k(\mathrm{Re}\,\psi)}{\xi!} (\mathrm{iIm}\,\psi)^\xi \right|_{C^0(\Delta_\sigma^k)} \\ = &\sup_{\Delta_\sigma^k} \left| f_{\sigma,k}(x) - \sum_{0 \le |\varrho| \le p} \frac{\partial^\varrho f^k(\mathrm{Re}\,\psi)}{\varrho!} (\mathrm{iIm}\,\psi)^\varrho \right| \end{aligned}$$

$$
\begin{aligned}
&\leq \ |f|_{C^l(D_\rho)}\sigma^l \sum_{|\varrho|=p} \frac{1}{\varrho!} \sup_{|\eta|<1} \int_{\Omega^k} |K_k(\xi+\mathrm{i}\eta)\xi^\varrho|\,|\xi|^\delta d\xi \\
&\leq \ c|f|_{C^l(D_\rho)}\sigma^l(2b+1+2k)^l.
\end{aligned}
$$

Similarly, for any monotone decreasing sequence $\{\sigma_k\}$ of positive numbers, we have by (2.4) that

$$
\left|\sum_{|\varrho|\leq p} \frac{\partial^\varrho f^k(\mathrm{Re}\,\psi)}{\varrho!}(\mathrm{iIm}\,\psi)^\varrho - \sum_{|\varrho|\leq p} \frac{\partial^\varrho f^{k+1}(\mathrm{Re}\,\psi)}{\varrho!}(\mathrm{iIm}\,\psi)^\varrho\right| \leq c_0\sigma_{k+1}^l \tag{2.9}
$$

on $\Delta^{k+1}_{\sigma_{k+1}}$, and consequently,

$$
\begin{aligned}
&\left|f_{\sigma_{k+1},k+1} - f_{\sigma_k,k}\right|_{C^0(\Delta^k_{\sigma_{k+1}})} = \sup_{\Delta^{k+1}_{\sigma_{k+1}}} |f_{\sigma_k,k}(\psi) - f_{\sigma_{k+1},k+1}(\psi)| \\
\leq \ & c_0\sigma_{k+1}^l + \sup_{\Delta^k_{\sigma_{k+1}}} |\hat{R}_{f^k}(\psi,\sigma_{k+1})| + \sup_{\Delta^{k+1}_{\sigma_{k+1}}} |\hat{R}_{f^{k+1}}(\psi,\sigma_{k+1})| \\
\leq \ & c_0\sigma_{k+1}^l + |f|_{C^l(D_\rho)}\sigma_{k+1}^l \sum_{|\varrho|=p} \frac{1}{\varrho!} \sup_{|\eta|<1} \int_{\Omega^k} |K_k(\xi+\mathrm{i}\eta)\xi^\varrho|\,|\xi|^\delta d\xi \\
+ \ & |f|_{C^l(D_\rho)}\sigma_{k+1}^l \sum_{|\varrho|=p} \frac{1}{\varrho!} \sup_{|\eta|<1} \int_{\Omega^{k+1}} |K_{k+1}(\xi+\mathrm{i}\eta)\xi^\varrho|\,|\xi|^\delta d\xi \\
\leq \ & c|f|_{C^l(D_\rho)}\sigma_{k+1}^l(2b+3+2k)^l.
\end{aligned}
$$

To make estimates in the C^1 norm, we recall the following well-known Cauchy estimates: For any $0<\sigma'<\sigma$ and any k, if $g\in\mathbb{C}^1(\Delta^k_\sigma)$, then

$$
|g|_{C^1(\Delta^k_{\sigma'})} \leq \sum_{|\xi'|=1} (\sigma-\sigma')^{-1}|g|_{C^0(\Delta^k_\sigma)}. \tag{2.10}
$$

By considering previous estimates in $\Delta^k_{2\sigma}$ instead of Δ^k_σ, it follows from (2.9)- (2.10) that

$$
\begin{aligned}
&\left|f_{\sigma,k}(\psi) - \sum_{0\leq|\xi|\leq p-1} \frac{\partial^\xi f^k(\mathrm{Re}\,\psi)}{\xi!}(\mathrm{iIm}\,\psi)^\xi\right|_{C^1(\Delta^k_\sigma)} \\
\leq \ & (2b+1+2k)\sigma^{-1}\left|f_{\sigma,k}(x) - \sum_{0\leq|\xi|\leq p} \frac{\partial^\xi f^k(\mathrm{Re}\,\psi)}{\xi!}(\mathrm{iIm}\,\psi)^\xi\right|_{C^0(\Delta^k_{2\sigma})} \\
\leq \ & 2^l c\sigma^{-1}|f|_{C^l(D_\rho)}\sigma^l(2b+1+2k)^{l+1} \\
\leq \ & c|f|_{C^l(D_\rho)}\sigma^{l-1}(2b+1+2k)^{l+1}
\end{aligned}
$$

and

$$
\begin{aligned}
\left|f_{\sigma_{k+1},k+1} - f_{\sigma_k,k}\right|_{C^1(\Delta^{k+1}_{\sigma_{k+1}})} \ &\leq \ (2b+3+2k)\sigma_{k+1}^{-1}\,\left||f_{\sigma_{k+1},k+1} - f_{\sigma_k,k}\right|_{C^0(\Delta^{k+1}_{2\sigma_{k+1}})} \\
&\leq \ 2^l c\sigma_{k+1}^{-1}|f|_{C^l(D_\rho)}\sigma_{k+1}^l(2b+3+2k)^{l+1}
\end{aligned}
$$

$$\leq \quad c|f|_{C^l(D_\rho)}\sigma_{k+1}^{l-1}(2b+3+2k)^{l+1}.$$

Thus, if $\{\sigma_k\}$ is a super-exponentially decreasing sequent, then $\{f_{\sigma_k,k}\}$ is a Cauchy sequence in $C^{1,0}(\Delta_\rho \times \Omega)$ hence in $C^0(\Delta_\rho \times \Omega)$. $\square$

3. Normal form

To proceed with the KAM iterations, we first convert the Hamiltonian (1.3) into a normal form, which include the introduction of action-angle-normal variables and making analytic approximations using Lemma 2.1.

3.1. **Action-angle-normal variables.** Setting $\dot{x}_n = y_n$, $n \in \mathbb{Z}$, in (1.1), (1.3) takes the form

$$H = \sum_{n\in\mathbb{Z}} \frac{1}{2}y_n^2 + \frac{\beta_n^2}{2}x_n^2 + V(x_{n+1}-x_n).$$

Let $\varepsilon > 0$ be a small parameter. With the re-scalings $y_n, x_n \to \varepsilon y_n, \varepsilon x_n$, $n \in \mathbb{Z}$, $H \to \varepsilon^{-2}H$, the re-scaled Hamiltonian reads

$$H = \sum_{n\in\mathbb{Z}} \frac{1}{2}y_n^2 + \frac{\beta_n^2}{2}x_n^2 + \varepsilon^{\alpha-1}V(x_{n+1}-x_n).$$

Denote $\mathbb{Z}_1 = \{j \in \mathbb{Z} : |j| > b\}$. For a given value $a = (a_j)_{|j|\leq b} \in \mathbb{R}_+^{2b+1}$, we introduce the standard action-angle-normal variables

$$(I, \theta, \tilde{x}, \tilde{y}) = (\{I_j\}_{|j|\leq b}, \{\theta_j\}_{|j|\leq b}, \{x_n\}_{n\in\mathbb{Z}_1}, \{y_n\}_{n\in\mathbb{Z}_1}),$$

i.e.,

$$x_j = \sqrt{\frac{2}{\beta_j}}\sqrt{I_j + a_j}\sin\theta_j, \qquad y_j = \sqrt{2\beta_j}\sqrt{I_j + a_j}\cos\theta_j, \quad |j| \leq b.$$

Without loss of generality, we set $\beta_j = 1$, $j \in \mathbb{Z}_1$. Then, in terms of the action-angle-normal variables, the above Hamiltonian becomes

$$H = N + \varepsilon^{\alpha-1}P, \tag{3.11}$$

where

$$N = e + \langle \omega, I\rangle + \sum_{n\in\mathbb{Z}_1} \frac{1}{2}(x_n^2 + y_n^2)$$

with $e = \sum\limits_{|j|\leq b} \beta_j a_j$, $\omega = (\beta_j)_{|j|\leq b}$, and

$$\begin{aligned} &P(\theta, I, \tilde{x}, \tilde{y}; \beta) \\ = \quad &\sum_{j=-b}^{b-1} \frac{1}{1+\alpha}\left|\sqrt{\frac{2}{\beta_{j+1}}}\sqrt{I_{j+1}+a_{j+1}}\sin\theta_{j+1} - \sqrt{\frac{2}{\beta_j}}\sqrt{I_j+a_j}\sin\theta_j\right|^{1+\alpha} \\ &+\frac{1}{1+\alpha}\left|\sqrt{\frac{2}{\beta_{-b}}}\sqrt{I_{-b}+\beta_{-b}}\sin\theta_{-b} - x_{-b-1}\right|^{1+\alpha} \end{aligned}$$

$$+\frac{1}{1+\alpha}\left|x_{b+1}-\sqrt{\frac{2}{\beta_b}}\sqrt{I_b+\beta_b}\sin\theta_b\right|^{1+\alpha}+\sum_{n,n+1\in\mathbb{Z}_1}\frac{1}{1+\alpha}\left|x_{n+1}-x_n\right|^{1+\alpha}.$$

In the above, the parameter β lies in the prescribed bounded closed region $\mathcal{O}$ of $\mathbb{R}_+^{2b+1}$ as in the Main Theorem, the angular and action variables θ, I lie in $\mathbb{T}^{2b+1}, \mathbb{R}^{2b+1}$, respectively, and the normal variables $\tilde{x}, \tilde{y}$ lie in ℓ. For a fixed sufficiently small $\rho > 0$, if we consider the domain

$$\breve{D}_{2\rho}=\left\{(\theta,I,\tilde{x},\tilde{y})\in\mathbb{T}^{2b+1}\times\mathbb{R}^{2b+1}\times\ell\times\ell:\|I\|\le\rho,\|\tilde{x}\|\le 2\rho,\|\tilde{y}\|\le 2\rho\right\},$$

then it is clear that $P \in C^{1+\alpha,1}(\breve{D}_{2\rho}\times\mathcal{O})$. We recall that $\|\cdot\|$ above is the weighted norm defined at the end of Section 1.

3.2. **Analytic approximation.** For each $\nu \ge 1$, define

$$P^\nu := P|_{\Omega_\nu} = P(\theta, I, \{x_n\}_{b<|n|\le b+\nu}, \{0\}, \{y_n\}_{b<|n|\le b+\nu}, \{0\}).$$

Then

$$\begin{aligned}P^{\nu+1}-P^\nu=&\frac{1}{1+\alpha}\left|x_{b+\nu+1}-x_{b+\nu}\right|^{1+\alpha}+\frac{1}{1+\alpha}\left|x_{b+\nu+1}\right|^{1+\alpha}\\ &-\frac{1}{1+\alpha}\left|x_{b+\nu}\right|^{1+\alpha}+\frac{1}{1+\alpha}\left|x_{-b-\nu}-x_{-b-\nu-1}\right|^{1+\alpha}\\ &+\frac{1}{1+\alpha}\left|x_{-b-\nu-1}\right|^{1+\alpha}-\frac{1}{1+\alpha}\left|x_{-b-\nu}\right|^{1+\alpha}.\end{aligned}\tag{3.12}$$

It follows that the perturbation P satisfies the condition (2.4) in Lemma 2.1, i.e.,

$$|P^{\nu+1}-P^\nu|_{C^{1+\alpha}(\breve{D}(2\rho)),\mathcal{O}}\le C_0,$$

where C_0 is a constant depending only on α and b.

Let $\sigma_\nu \downarrow 0$, $\nu \ge 1$, be a supper-exponentially decreasing sequence to be specified later. For each $\nu \ge 1$, we consider the real analytic function

$$P_{2\sigma_\nu} := K_\nu^{2\sigma_\nu} * P^\nu.$$

Then by Lemma 2.1, $\{P_{2\sigma_\nu}\}$ is a sequence of real analytic approximations to P, i.e.,

$$P = P_{2\sigma_1} + \sum_{\nu\ge 2}(P_{2\sigma_\nu} - P_{2\sigma_{\nu-1}})$$

uniformly on $\breve{D}(2\rho)\times\mathcal{O}$.

For convenience, we complexitify the real variables $\tilde{x} = (x_n)_{n\in\mathbb{Z}_1}$ and $\tilde{y} = (y_n)_{n\in\mathbb{Z}_1}$ by linear transformations:

$$w_n=\frac{1}{\sqrt{2}}(x_n+\mathrm{i}y_n),\qquad \bar{w}_n=\frac{1}{\sqrt{2}}(x_n-\mathrm{i}y_n),\quad n\in\mathbb{Z}_1.\tag{3.13}$$

Under this transformation, the function N and P in Hamiltonian (3.11) becomes

$$\tilde{H} = \tilde{N} + \varepsilon^{\alpha-1}\tilde{P},$$

where

$$\tilde{N} = e + \langle \omega, I \rangle + \sum_{n \in \mathbb{Z}_1} w_n \bar{w}_n$$

and

$$\tilde{P}(\theta, I, w, \bar{w}; \beta) := P(\theta, I, (\frac{w_n + \bar{w}_n}{\sqrt{2}}), (\frac{w_n - \bar{w}_n}{\mathrm{i}\sqrt{2}}); \beta).$$

For each $\nu \geq 1$, since

$$\begin{array}{c} \sum_{b<|n|} |\mathrm{Im}\, w_n| \langle n \rangle^{1+\alpha} \leq \sigma_\nu \\ \sum_{b<|n|} |\mathrm{Im}\, \bar{w}_n| \langle n \rangle^{1+\alpha} \leq \sigma_\nu \end{array} \Longrightarrow \begin{array}{c} \sum_{b<|n|} |\mathrm{Im}\, x_n| \langle n \rangle^{1+\alpha} \leq \sqrt{2}\sigma_\nu \\ \sum_{b<|n|} |\mathrm{Im}\, y_n| \langle n \rangle^{1+\alpha} \leq \sqrt{2}\sigma_\nu, \end{array}$$

we see that the function

$$\tilde{P}_{\sigma_\nu}(\theta, I, w, \bar{w}; \beta) := P_{\sigma_\nu}(\theta, I, (\frac{w_n + \bar{w}_n}{\sqrt{2}}), (\frac{w_n - \bar{w}_n}{\mathrm{i}\sqrt{2}}); \beta)$$

is bounded, C^1 in $\beta \in \mathcal{O}$, and real-analytic in $\psi := (\theta, I, w, \bar{w}) \in \Delta^\nu_{\sigma_\nu}$, where

$$\Delta^\nu_{\sigma_\nu} = \left\{ \psi : \begin{array}{c} \mathrm{Re}(\theta, I, w, \bar{w}) \in \breve{D}_\rho, \\ \sum_{|j|\leq b} |\mathrm{Im}\, \theta_j| \langle j \rangle^{1+\alpha} \leq \sigma_\nu,\ \sum_{|j|\leq b} |\mathrm{Im}\, I_j| \langle j \rangle^{1+\alpha} \leq \sigma_\nu, \\ \sum_{b<|n|} |\mathrm{Im}\, w_n| \langle n \rangle^{1+\alpha} \leq \sigma_\nu, \\ \sum_{b<|n|} |\mathrm{Im}\, \bar{w}_n| \langle n \rangle^{1+\alpha} \leq \sigma_\nu, \\ |\mathrm{Im}\, w_n| = |\mathrm{Im}\, \bar{w}_n| = 0,\ |n| \geq b + \nu + 1 \end{array} \right\},$$

and moreover,

$$|\tilde{P}_{\sigma_\nu}|_{C^1(\Delta^\nu_{\sigma_\nu}), \mathcal{O}} \leq c(\alpha) |P_{\sigma_\nu}|_{C^1(\Delta^\nu_{2\sigma_\nu}), \mathcal{O}}.$$

Since by Lemma 2.1, there is a constant $C_1 > 0$ depending only on α and b such that

$$\begin{aligned} &\left| P_{\sigma_\nu} - P_{\sigma_{\nu+1}} \right|_{C^1(\Delta^{\nu+1}_{2\sigma_{\nu+1}}), \mathcal{O}} \\ \leq\ & C_1 |P|_{C^{1+\alpha}(\breve{D}(2\rho)), \mathcal{O}} (2\sigma_\nu)^{\alpha-1} (2b + 3 + 2\nu)^{1+\alpha}, \end{aligned} \tag{3.14}$$

we have

$$\begin{aligned} &\left| \tilde{P}_{\sigma_\nu} - \tilde{P}_{\sigma_{\nu+1}} \right|_{C^1(\Delta^{\nu+1}_{\sigma_{\nu+1}}), \mathcal{O}} \\ \leq\ & C_2 |P|_{C^{1+\alpha}(\breve{D}(2\rho)), \mathcal{O}} \sigma_\nu^{\alpha-1} (2b + 3 + 2\nu)^{1+\alpha}, \end{aligned} \tag{3.15}$$

where $C_2 > 0$ is a constant depending only on α and b. Hence, $\{\tilde{P}_{\sigma_\nu}\}$ is a sequence of uniform real analytic approximations to $\tilde{P}$.

4. KAM scheme

To highlight our main idea of the KAM iterations, we describe in this section one KAM step and leave all estimates to later sections. As mentioned in Section 1, only the analytic approximations $\{\tilde{P}_{\sigma_\nu}\}$ of $\tilde{P}$ $(= P)$ will participate in the KAM iterations. More precisely, we will construct a sequence of real analytic, symplectic transformations

$$\Phi_\nu := \Phi_{\nu-1} \circ \phi_\nu = \phi_1 \circ \cdots \circ \phi_\nu$$

so that

$$(N_\nu + \tilde{P}_{\sigma_\nu}) \circ \phi_\nu = N_{\nu+1} + P_{\nu+1} \tag{4.16}$$

for all $\nu = 1, 2, \cdots$, where $\{P_\nu\}$ is a sequence of real analytic perturbations to be constructed later, and for each $\nu \geq 1$, N_ν has the form

$$N_\nu = e_\nu + \langle \omega_\nu, I \rangle + \langle A_\nu Z_\nu, \bar{Z}_\nu \rangle + \sum_{|n| \geq b+\nu} w_n \bar{w}_n$$

with e_ν being a scalar-valued function, ω_ν being a $\mathbb{R}^{2b+1}$-valued function, and A_ν being a $2(\nu-1) \times 2(\nu-1)$ matrix-valued function, defined on a suitable parameter domain $\mathcal{O}_\nu$, and $Z_\nu = (w_n)_{b+1 \leq |n| \leq b+\nu-1}$. When $\nu = 1$, we simply define

$$N_1 = \tilde{N}, \quad P_1 = \tilde{P}_{\sigma_1}, \quad \mathcal{O}_1 := \mathcal{O}.$$

In fact, it can be shown easily by induction that (4.16) is equivalent to

$$(N_\nu + P'_\nu) \circ \phi_\nu = N_{\nu+1} + P_{\nu+1}, \qquad \nu = 1, 2, \cdots, \tag{4.17}$$

where $P'_1 = P_1 = \tilde{P}_{\sigma_1}$ and

$$P'_\nu = P_\nu + (\tilde{P}_{\sigma_\nu} - \tilde{P}_{\sigma_{\nu-1}}) \circ \Phi_{\nu-1}, \qquad \nu \geq 2.$$

It follows that

$$H_1 \circ \Phi_\nu + \sum_{j=2}^{\nu} (\tilde{P}_{\sigma_j} - \tilde{P}_{\sigma_{j-1}}) \circ \Phi_\nu = N_{\nu+1} + P'_{\nu+1},$$

i.e.,

$$(\tilde{N} + \tilde{P}_{\sigma_\nu}) \circ \Phi_\nu = N_{\nu+1} + P'_{\nu+1}, \quad \nu = 1, 2, \cdots. \tag{4.18}$$

On some phase and parameter domains and in suitable norms, we will show that Φ_ν and N_ν converge, say to Φ_∞ and N_∞ respectively, and P'_ν converges to 0. Therefore,

$$\tilde{H} \circ \Phi_\infty = N_\infty$$

which yields a family of quasi-periodic $(2b+1)$-tori of $\tilde{H}$.

We now outline the construction of the symplectic transformation and describe the transformed Hamiltonian for one KAM step, i.e., the νth step for a fixed $\nu \geq 1$.

4.1. **Truncation of the perturbation.** Consider the following Taylor-Fourier expansion of P'_ν:

$$\begin{aligned} P'_\nu &= \sum_{k,l,\varrho,\varrho'} P_{k,l,\varrho,\varrho'}(\beta)e^{\mathrm{i}\langle k,\theta\rangle} I^l Z_\nu^\varrho \bar{Z}_\nu^{\varrho'} \\ &+ \sum_{\substack{\varrho_n+\varrho'_m\geq 1\\ |m|=|n|=b+\nu,k,l,\varrho,\varrho'}} P_{k,l,\varrho,\varrho',\varrho_n,\varrho'_m}(\beta)e^{\mathrm{i}\langle k,\theta\rangle} I^l Z_\nu^\varrho \bar{Z}_\nu^{\varrho'} w_n^{\varrho_n}\bar{w}_m^{\varrho'_m}, \end{aligned}$$

where $Z_\nu = (w_n)_{b+1\leq|n|\leq b+\nu-1}$, $k\in\mathbb{Z}^{2b+1}$, $l\in\mathbb{N}^{2b+1}$, $\varrho_n, \varrho'_n\in\mathbb{N}$, and the multi-indices ϱ, ϱ' run over the set of integer vectors

$$\varrho \equiv (\cdots,\varrho_n,\cdots)_{b+1\leq|n|\leq b+\nu-1}, \quad \varrho' \equiv (\cdots,\varrho'_n,\cdots)_{b+1\leq|n|\leq b+\nu-1},$$

respectively. If $\nu=1$, then we simply let Z_1, ϱ, and ϱ' be null vectors.

With respect to a truncation order K_ν to be specified later, we consider the following truncation $R=R_\nu$ of P'_ν:

$$\begin{aligned} & R(\theta,I,Z_\nu,\bar{Z}_\nu,w,\bar{w}) \\ =& \sum_{\substack{|l|\leq 1\\ |k|\leq K_\nu}} P_{kl00}(\beta)e^{\mathrm{i}\langle k,\theta\rangle} I^l + \sum_{|k|\leq K_\nu} (\langle P_k^{10}(\beta), Z_\nu\rangle + \langle P_k^{10}(\beta),\bar{Z}_\nu\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\ +& \sum_{|k|\leq K_\nu} (\langle P_k^{20}(\beta)Z_\nu, Z_\nu\rangle + \langle P_k^{11}(\beta)Z_\nu,\bar{Z}_\nu\rangle + \langle P_k^{02}(\beta)\bar{Z}_\nu,\bar{Z}_\nu\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\ +& \sum_{\substack{|k|\leq K_\nu\\ |n|=b+\nu}} (\langle P_{kn}^{20}(\beta)w_n, Z_\nu\rangle + \langle P_{kn}^{02}(\beta)\bar{w}_n,\bar{Z}_\nu\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\ +& \sum_{\substack{|k|\leq K_\nu\\ |n|=b+\nu}} (\langle P_{kn}^{11}(\beta)w_n, \bar{Z}_\nu\rangle + \langle P_{kn}^{11}(\beta)\bar{w}_n, Z_\nu\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\ +& \sum_{\substack{|k|\leq K_\nu\\ |m|=|n|=b+\nu}} (P_{knm}^{20}(\beta)w_nw_m + P_{knm}^{11}(\beta)w_n\bar{w}_m + P_{knm}^{02}(\beta)\bar{w}_n\bar{w}_m)e^{\mathrm{i}\langle k,\theta\rangle} \\ +& \sum_{\substack{|k|\leq K_\nu\\ |n|=b+\nu}} (P_{kn}^{10}(\beta)w_n + P_{kn}^{01}(\beta)\bar{w}_n)e^{\mathrm{i}\langle k,\theta\rangle}, \end{aligned}$$

where P_{k*}^{ij}, $0\leq i,j\leq 2$, $|k|\leq K_\nu$, are numbers and vectors or matrices consisting of appropriate Taylor-Fourier coefficients of P'_ν.

To match the form of N_ν in order to define the new perturbation and derive the homological equation, we rewrite R as

$$R(\theta,I,Z_{\nu+1},\bar{Z}_{\nu+1},) = R_0+R_1+R_2,$$

where

$$R_0 = \sum_{\substack{|l|\leq 1\\ |k|\leq K_\nu}} P_{kl00}(\beta)e^{\mathrm{i}\langle k,\theta\rangle} I^l$$

$$
\begin{aligned}
R_1 &= \sum_{|k|\le K_\nu} (\langle R_k^{10}(\beta), Z_{\nu+1}\rangle + \langle R_k^{01}(\beta), \bar Z_{\nu+1}\rangle) e^{\mathrm{i}\langle k,\theta\rangle}\\
R_2 &= \sum_{|k|\le K_\nu} (\langle R_k^{20}(\beta) Z_{\nu+1}, Z_{\nu+1}\rangle + \langle R_k^{11}(\beta) Z_{\nu+1}, \bar Z_{\nu+1}\rangle + \langle R_k^{02}(\beta)\bar Z_{\nu+1}, \bar Z_{\nu+1}\rangle) e^{\mathrm{i}\langle k,\theta\rangle}
\end{aligned}
$$

with

$$
\begin{aligned}
R_k^{10} &= \begin{pmatrix} \bar P_k^{10} \\ \bar P_{kn}^{10} \end{pmatrix}_{|n|=b+\nu}, R_k^{01} = \begin{pmatrix} \bar P_k^{10} \\ \bar P_{kn}^{01} \end{pmatrix}_{|n|=b+\nu},\\
R_k^{20} &= \begin{pmatrix} P_k^{20} & \frac12 P_{kn}^{20} \\ \frac12 (P_{kn}^{20})^\top & (P_{knm}^{20})_{|m|=b+\nu} \end{pmatrix}_{|n|=b+\nu},\\
R_k^{02} &= \begin{pmatrix} P_k^{02} & \frac12 P_{kn}^{02} \\ \frac12 (P_{kn}^{02})^\top & (P_{knm}^{02})_{|m|=b+\nu} \end{pmatrix}_{|n|=b+\nu},\\
R_k^{11} &= \begin{pmatrix} P_k^{11} & P_{kn}^{11} \\ P_{k*n}^{11} & (P_{knm}^{11})_{|m|=b+\nu} \end{pmatrix}_{|n|=b+\nu}.
\end{aligned}
$$

4.2. **Construction of the symplectic transformation.** Denote

$$
H'_\nu = N_\nu + P'_\nu, \qquad H_\nu = N_\nu + P_\nu,\ \nu = 1,2,\cdots.
$$

For fixed ν, we will look for a real analytic Hamiltonian function $F = F_\nu$ such that the time-1 map $\phi_\nu =: \phi_F^1$ of the Hamiltonian vector field X_F transforms H'_ν into $H_{\nu+1}$.

We note by the second order Taylor formula that

$$
\begin{aligned}
H'_\nu\circ\phi_F^1 &= (N_\nu + R)\circ\phi_F^1 + (P'_\nu - R)\circ\phi_F^1\\
&= N_\nu + \{N_\nu, F\} + R + \int_0^1 (1-t)\{\{N_\nu, F\}, F\}\circ\phi_F^t dt\\
&\quad + \int_0^1 \{R, F\}\circ\phi_F^t dt + (P'_\nu - R)\circ\phi_F^1\\
&= N_{\nu+1} + P_{\nu+1} + \{N_\nu, F\} + R\\
&\quad - \sum_{|l|\le 1} P_{0l00} I^l - \langle R_0^{11} Z_{\nu+1}, \bar Z_{\nu+1}\rangle,
\end{aligned}
$$

where

$$
\begin{aligned}
N_{\nu+1} &= N_\nu + P_{0000} + \sum_{|l|=1} P_{0l00} I^l + \langle R_0^{11} Z_{\nu+1}, \bar Z_{\nu+1}\rangle, \qquad (4.19)\\
P_{\nu+1} &= \int_0^1 (1-t)\{\{N_{\nu+1}, F\}, F\}\circ\phi_F^t dt + \int_0^1 \{R, F\}\circ\phi_F^t dt\\
&+ (P'_\nu - R)\circ\phi_F^1.
\end{aligned}
$$

Thus, for the validity of

$$
H'_\nu\circ\phi_F^1 = H_{\nu+1},
$$

F needs to satisfy the following *homological equation*:

$$\{N_\nu, F\} + R - \sum_{|l|\le 1} P_{0l00} I^l - \langle R_0^{11} Z_{\nu+1}, \bar{Z}_{\nu+1}\rangle = 0. \tag{4.20}$$

4.3. **The solvability of the homological equation.** To solve a Hamiltonian F from the homological equation, we assume that F is of the form

$$F(\theta, I, z, \bar{z}) = F_0 + F_1 + F_2,$$

where

$$\begin{aligned}
F_0 &= \sum_{\substack{|l|\le 1\\ |k|\le K_\nu}} f_{kl00}(\beta) e^{\mathrm{i}\langle k,\theta\rangle} I^l,\\
F_1 &= \sum_{\substack{|k|\le K_\nu\\ b+1\le |n|\le b+\nu}} (f_n^{k10} w_n + f_n^{k01}\bar{w}_n) e^{\mathrm{i}\langle k,\theta\rangle}\\
&:= \sum_{|k|\le K_\nu} (\langle F_k^{10}, Z_{\nu+1}\rangle + \langle F_k^{01}, \bar{Z}_{\nu+1}\rangle) e^{\mathrm{i}\langle k,\theta\rangle}\ ,\\
F_2 &= \sum_{\substack{|k|\,\le\, K_\nu\\ b+1\,\le\, |n|, |m|\,\le\, b+\nu}} (f_{nm}^{k20} w_n w_m + f_{nm}^{k02}\bar{w}_n\bar{w}_m) e^{\mathrm{i}\langle k,\theta\rangle}\\
&\quad + \sum_{\substack{0\,<\,|k|\,\le\, K_\nu\\ b+1\,\le\, |n|, |m|\,\le\, b+\nu}} f_{nm}^{k11} w_n \bar{w}_m e^{\mathrm{i}\langle k,\theta\rangle}\\
&:= \sum_{|k|\le K_\nu} (\langle F_k^{20} Z_{\nu+1}, Z_{\nu+1}\rangle + \langle F_k^{02} \bar{Z}_{\nu+1}, \bar{Z}_{\nu+1}\rangle) e^{\mathrm{i}\langle k,\theta\rangle}\\
&\quad + \sum_{0<|k|\le K_\nu} \langle F_k^{11} Z_{\nu+1}, \bar{Z}_{\nu+1}\rangle e^{\mathrm{i}\langle k,\theta\rangle}\ .
\end{aligned}$$

We note that N_ν can be rewritten as

$$\begin{aligned}
N_\nu &= e_\nu + \langle \omega_\nu, I\rangle + \langle A_\nu Z_\nu, \bar{Z}_\nu\rangle + \sum_{|n|=b+\nu} w_n\bar{w}_n + \sum_{|n|>b+\nu} w_n\bar{w}_n\\
&:= e_\nu + \langle \omega_\nu, I\rangle + \langle \tilde{A}_\nu Z_{\nu+1}, \bar{Z}_{\nu+1}\rangle + \sum_{|n|>b+\nu} w_n\bar{w}_n,
\end{aligned}$$

where $Z_{\nu+1} = (w_n)_{b+1\le |n|\le b+\nu}$ and

$$\tilde{A}_\nu = \begin{pmatrix} A_\nu & 0\\ 0 & I_2 \end{pmatrix}$$

with I_2 being the 2×2 identity matrix.

Lemma 4.1. *The homological equation* (4.20) *is equivalent to*

$$
\begin{array}{l}
(\langle k,\omega_\nu\rangle) f_{kl00} = -\mathrm{i}P_{kl00},\ \ 0<|k|\le K_\nu,\ |l|\le 1,\\
(\langle k,\omega_\nu\rangle I-\tilde{A}_\nu))F_k^{10} = -\mathrm{i}R_k^{10},\ \ |k|\le K_\nu\\
(\langle k,\omega_\nu\rangle I+\tilde{A}_\nu)F_k^{01} = -\mathrm{i}R_k^{01},\ \ |k|\le K_\nu\\
(\langle k,\omega_\nu\rangle I-I\otimes\tilde{A}_\nu-\tilde{A}_\nu\otimes I)F_k^{20} = -\mathrm{i}R_k^{20},\ \ |k|\le K_\nu\\
(\langle k,\omega\rangle I+I\otimes\tilde{A}_\nu+\tilde{A}_\nu\otimes I)F_k^{02} = -\mathrm{i}R_k^{02},\ \ |k|\le K_\nu\\
(\langle k,\omega\rangle I+I\otimes\tilde{A}_\nu-\tilde{A}_\nu\otimes I)F_k^{11} = -\mathrm{i}R_k^{11},\ \ 0<|k|\le K_\nu.
\end{array}
\tag{4.21}
$$

Proof. Substituting F into (4.20) one sees that (4.20) is equivalent to the following equations

$$
\begin{array}{l}
\{N_\nu,F_0\}+R_0=\sum_{|l|\le 1} P_{0l00}I^l,\\
\{N_\nu,F_1\}+R_1=0,\\
\{N_\nu,F_2\}+R_2=\langle R_0^{11}Z_{\nu+1},\bar{Z}_{\nu+1}\rangle.
\end{array}
\tag{4.22}
$$

By comparing the coefficients, the first equation in (4.22) is obviously equivalent to the first equation in (4.21). Since

$$
\begin{aligned}
&\{N_\nu,F_1\}\\
=\ &\mathrm{i}\sum_{|k|\le K_\nu}(\langle((\langle k,\omega_\nu\rangle I-\tilde{A}_\nu)F_k^{10},Z_{\nu+1}\rangle+\langle((\langle k,\omega_\nu\rangle I+\tilde{A}_\nu)F_k^{01},\bar{Z}_{\nu+1}\rangle)e^{\mathrm{i}\langle k,\theta\rangle}
\end{aligned}
$$

and

$$
\begin{aligned}
&\{N_\nu,F_2\}\\
=\ &\mathrm{i}\sum_{|k|\le K_\nu}\langle((\langle k,\omega_\nu\rangle F_k^{20}-\tilde{A}_\nu F_k^{20}-F_k^{20}\tilde{A}_\nu)Z_{\nu+1},Z_{\nu+1}\rangle e^{\mathrm{i}\langle k,\theta\rangle}\\
&+\mathrm{i}\sum_{|k|\le K_\nu}\langle((\langle k,\omega_\nu\rangle F_k^{02}+\tilde{A}_\nu F_k^{02}+F_k^{02}\tilde{A}_\nu)\bar{Z}_{\nu+1},\bar{Z}_{\nu+1}\rangle e^{\mathrm{i}\langle k,\theta\rangle}\\
&+\mathrm{i}\sum_{0<|k|\le K_\nu}\langle((\langle k,\omega_\nu\rangle F_k^{11}+\tilde{A}_\nu F_k^{11}-F_k^{11}\tilde{A}_\nu)Z_{\nu+1},\bar{Z}_{\nu+1}\rangle e^{\mathrm{i}\langle k,\theta\rangle}\ ,
\end{aligned}
$$

the equivalence of between the remaining equations of (4.21) and (4.22) is established by comparing coefficients. □

Suppose that parameters involved in H'_ν, H_ν lie in some region $\mathcal{O}_\nu\subset\mathcal{O}$. To solve equations in (4.21), we consider the following sub-region of parameters:

$$
\mathcal{O}_{\nu+1}=\left\{\beta\in\mathcal{O}_\nu:\ \begin{array}{l}
|\langle k,\omega_\nu\rangle^{-1}|\le\frac{|K_\nu|^\tau}{\gamma},\quad k\neq 0,\\
|(\langle k,\omega_\nu\rangle I\pm\tilde{A}_\nu)^{-1}|\le\frac{(2\nu)^2|K_\nu|^\tau}{\gamma},\\
|(\langle k,\omega_\nu\rangle I+I\otimes\tilde{A}_\nu+\tilde{A}_\nu\otimes I)^{-1}|\le\frac{(2\nu)^4|K_\nu|^\tau}{\gamma},\\
|(\langle k,\omega_\nu\rangle I+I\otimes\tilde{A}_\nu-\tilde{A}_\nu\otimes I)^{-1}|\le\frac{(2\nu)^4|K_\nu|^\tau}{\gamma},\\
|k|\le K_\nu
\end{array}\right\},
$$

where $0<\gamma\ll 1$ is a fixed Diophantine constant and $\tau=2b+2$. We have the following result.

Proposition 4.1. (4.21) *is uniquely solvable on $\mathcal{O}_{\nu+1}$ to yield solutions satisfying*

$$\begin{aligned}
&|f_{kl00}|_{\mathcal{O}_{\nu+1}} \le \gamma^{-2}(|k|)^{2\tau+1}|P_{kl00}|_{\mathcal{O}_{\nu+1}},\ \ 0<|k|\le K_\nu,\ |l|\le 1,\\
&|F_k^{10}|_{\mathcal{O}_{\nu+1}} \le \gamma^{-2}(|k|)^{2\tau+1}(2\nu)^5|R_k^{10}|_{\mathcal{O}_{\nu+1}},\ \ |k|\le K_\nu,\\
&|F_k^{01}|_{\mathcal{O}_{\nu+1}} \le \gamma^{-2}(|k|)^{2\tau+1}(2\nu)^5|R_k^{01}|_{\mathcal{O}_{\nu+1}},\ \ |k|\le K_\nu,\\
&|F_k^{20}|_{\mathcal{O}_{\nu+1}} \le \gamma^{-2}(|k|)^{2\tau+1}(2\nu)^{10}|R_k^{20}|_{\mathcal{O}_{\nu+1}},\ \ |k|\le K_\nu,\\
&|F_k^{11}|_{\mathcal{O}_{\nu+1}} \le \gamma^{-2}(|k|)^{2\tau+1}(2\nu)^{10}|R_k^{11}|_{\mathcal{O}_{\nu+1}},\ \ 0<|k|\le K_\nu,\\
&|F_k^{02}|_{\mathcal{O}_{\nu+1}} \le \gamma^{-2}(|k|)^{2\tau+1}(2\nu)^{10}|R_k^{02}|_{\mathcal{O}_{\nu+1}},\ \ |k|\le K_\nu.
\end{aligned}$$

The proof of Proposition 4.1 follows easily from the following two elementary lammas on matrices.

Lemma 4.2. *Let A, B be $n\times n, m\times m$ real symmetric matrices, and C, X be $n\times m$ matrices. Then the matrix equation*

$$AX - XB = C$$

is solvable if and only if $I_m \otimes A - B\otimes I_n$ is nonsingular. Moreover, the solution of this matrix equation satisfies

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

Proof. See e.g. [25]. □

Lemma 4.3. *Let $A(\beta) = (a_{ij}(\beta))_{n\times n}$ be a family of invertible matrices depending differentially on the parameter $\beta\in O$, where O is a bounded region in an Euclidean space. If there are constants $M, L>0$ such that $|A^{-1}|\le L, |\partial_\beta A|\le M$, then*

$$|\partial_\beta A^{-1}| \le L^2 M.$$

Proof. Since $AA^{-1} = I$, differentiation yields that $(\partial_\beta A)A^{-1} + A(\partial_\beta A^{-1}) = 0$, i.e., $\partial_\beta A^{-1} = -A^{-1}(\partial_\beta A)A^{-1}$. Thus

$$|\partial_\beta A^{-1}| \le |A^{-1}|^2|\partial_\beta A| \le L^2 M.$$

□

5. KAM iterations

Following the KAM scheme outlined in the previous section, we now show details of the KAM iterations by providing estimates on the transformations and the transformed Hamiltonians. An iteration lemma will be given at the end of the section.

Given $r, s>0$, we denote the (r,s)-neighborhood of $\mathbb{T}^{2b+1}\times\{I=0\}\times\{w=0\}\times\{\bar w=0\}$ of $(\mathbb{C}^{2b+1}/\mathbb{Z}^{2b+1})\times\mathbb{C}^{2b+1}\times\mathbb{C}^{\mathbb{Z}_1}\times\mathbb{C}^{Z_1}$ by

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

We recall that $\|\cdot\|$ above is the weighted norm defined at the end of Section 1. For any vector-valued function

$$X = X(\theta, I, w, \bar w) = (X^\theta, X^I, \{X^{w_n}\}, \{X^{\bar w_n}\})$$

on $(\mathbb{C}^{2b+1}/\mathbb{Z}^{2b+1}) \times \mathbb{C}^{2b+1} \times \mathbb{C}^{\mathbb{Z}_1} \times \mathbb{C}^{Z_1}$ which depends on the parameter $\beta \in \mathcal{O}$ smoothly, we define its weighted norm on $D(r,s) \times \mathcal{O}$ by

$$\begin{aligned} \|X\|_{D(r,s),\mathcal{O}} &\equiv |X^I|_{D(r,s),\mathcal{O}} + \frac{1}{s^2}|X^\theta|_{D(r,s),\mathcal{O}} \\ &+ \frac{1}{s}\Big(\sum_{n\in\mathbb{Z}_1}|X^{w_n}|_{D(r,s),\mathcal{O}}\langle n\rangle^{1+\alpha} + \sum_{n\in\mathbb{Z}_1}|X^{\bar{w}_n}|_{D(r,s),\mathcal{O}}\langle n\rangle^{1+\alpha}\Big). \end{aligned} \tag{5.23}$$

Let $0 < q < 1 < \kappa < \frac{4}{3}$ be fixed constants satisfying

$$\frac{\kappa^2}{4-3\kappa} = \frac{\alpha-1}{\alpha_*-1}, \quad q = \frac{4-3\kappa}{2\tau+4},$$

where $\tau := 2b+2$ and $\alpha_* := 3(2\tau+4)+1$. Also let $\varepsilon_1 = \varepsilon$ and $s_1 > 0$ be given such that

$$\frac{\varepsilon_1^q}{2} > s_1 \geq \varepsilon_1^\kappa.$$

We will use the following iteration sequences for the KAM iterations:

$$r_\nu = \frac{1}{2}\sigma_\nu, \quad \sigma_\nu = \varepsilon_\nu^q, \quad s_{\nu+1} = \varepsilon_\nu^{\kappa-1}s_\nu,$$
$$K_\nu = K_0\frac{\kappa^\nu}{\sigma_\nu}, \quad \varepsilon_{\nu+1} = c\gamma^{-2}\varepsilon_\nu^\kappa,$$

$\nu = 1,2,\cdots$, where $\breve{D}_{2\rho}$ is the domain defined in Section 3, $K_0 \gg b+1$ and $c = c(b,\alpha,\alpha*) \gg 1$ are fixed constants, and $\gamma > 0$ is the fixed Diophantine constant given in Section 4. In fact, $\{\sigma_\nu\}$ is the sequence used for the analytic approximations in Sections 2 and 3 and K_ν's are the truncation orders used in Section 4. We note that ε_ν will be used to control the size of analytic perturbations in the νth KAM step, and, r_ν, s_ν are the size of phase domain $D(r_\nu, s_\nu)$ of the real analytic Hamiltonians H'_ν, H_ν at the νth KAM step.

When making ε sufficiently small, it is not hard to check that

$$s_\nu \geq \varepsilon_\nu^\kappa, \quad s_\nu < \sigma_\nu, \quad 2^{\nu+3}\varepsilon_\nu^{3(\kappa-1)} \ll 1. \tag{5.24}$$

5.1. **Estimates on the transformations.** For a given $\nu \geq 1$, we need to estimate the symplectic transformation $\phi_\nu = \phi_F^1$, where $F = F_\nu$ is the Hamiltonian satisfying the homological equation (4.20). To do so, we assume that

(A1) *the perturbations P'_ν and P_ν satisfy*

$$\|X_{P'_\nu}\|_{D(r_\nu,s_\nu),\mathcal{O}_\nu} \leq \varepsilon_\nu, \quad \|X_{P_\nu}\|_{D(r_\nu,s_\nu),\mathcal{O}_\nu} \leq \frac{\varepsilon_\nu}{2}.$$

We will use the induction to show later that the assumption (A1) holds for all $\nu = 1,2,\cdots$.

Lemma 5.1. *Denote $D_i = D(r_\nu, \frac{i}{4}s_\nu)$, $0 < i \leq 4$. If* (A1) *hold, then*

$$\|X_F\|_{D(\frac{r_\nu}{2},\frac{3}{4}s_\nu),\mathcal{O}_{\nu+1}} \preccurlyeq \gamma^{-2}r_\nu^{-(2\tau+2)}\varepsilon_\nu. \tag{5.25}$$

Proof. Let $R = R_\nu$ be the truncation of P'_ν defined in Section 4.1. We have by (A1) and the standard Cauchy estimate that

$$\|X_R\|_{D_3,\mathcal{O}_{\nu+1}} \le \varepsilon_\nu.$$

It follows from Proposition 4.1 that

$$\begin{aligned}
\frac{1}{s_\nu^2}\|F_\theta\|_{D(\frac{r_\nu}{2},\frac{3}{4}s_\nu),\mathcal{O}_{\nu+1}} \le\ & \frac{1}{s_\nu^2}\Big(\sum_{|k|\le K_\nu,|l|\le 1} |f_{kl00}|\cdot s_\nu^{2|l|}\cdot|k|\cdot e^{|k|r_\nu}\cdot e^{-|k|\frac{r_\nu}{2}} \\
&+ \sum_{|k|\le K_\nu} |F_k^{10}|\cdot|Z_{\nu+1}|\cdot|k|\cdot e^{|k|r_\nu}\cdot e^{-|k|\frac{r_\nu}{2}} \\
&+ \sum_{|k|\le K_\nu} |F_k^{01}|\cdot|\bar{Z}_{\nu+1}|\cdot|k|\cdot e^{|k|r_\nu}\cdot e^{-|k|\frac{r_\nu}{2}} \\
&+ \sum_{|k|\le K_\nu} |F_k^{20}|\cdot|Z_{\nu+1}|^2\cdot|k|\cdot e^{|k|r_\nu}\cdot e^{-|k|\frac{r_\nu}{2}} \\
&+ \sum_{0<|k|\le K_\nu} |F_k^{11}|\cdot|Z_{\nu+1}|\cdot|\bar{Z}_{\nu+1}|\cdot|k|\cdot e^{|k|r_\nu}\cdot e^{-|k|\frac{r_\nu}{2}} \\
&+ \sum_{|k|\le K_\nu} |F_k^{02}|\cdot|\bar{Z}_{\nu+1}|^2\cdot|k|\cdot e^{|k|r_\nu}\cdot e^{-|k|\frac{r_\nu}{2}}\Big) \\
\preccurlyeq\ & \gamma^{-2}r_\nu^{-(2\tau+2)}\|X_R\|_{D_3,\mathcal{O}_{\nu+1}} \preccurlyeq \gamma^{-2}r_\nu^{-(2\tau+2)}\varepsilon_\nu.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|F_I\|_{D(\frac{r_\nu}{2},\frac{3}{4}s_\nu),\mathcal{O}_{\nu+1}} &\le c\gamma^{-2}r_\nu^{-(2\tau+2)}\varepsilon_\nu, \\
\frac{1}{s_\nu}\|F_{Z_{\nu+1}}\|_{D(\frac{r_\nu}{2},\frac{3}{4}s_\nu),\mathcal{O}_{\nu+1}} &\le c\gamma^{-2}r_\nu^{-(2\tau+2)}\varepsilon_\nu \\
\frac{1}{s_\nu}\|F_{\bar{Z}_{\nu+1}}\|_{D(\frac{r_\nu}{2},\frac{3}{4}s_\nu),\mathcal{O}_{\nu+1}} &\preccurlyeq \gamma^{-2}r_\nu^{-(2\tau+2)}\varepsilon_\nu.
\end{aligned}$$

The lemma now follows from the definition of the weighted norm (5.23) and the above estimates. □

The following result particularly implies that $\phi_\nu : D(r_{\nu+1}, s_{\nu+1})\times\mathcal{O}_{\nu+1} \to D(r_\nu, s_\nu)\times\mathcal{O}_\nu$ is well defined.

Lemma 5.2. *Denote $\eta_\nu = \varepsilon_\nu^{\kappa-1}$ and $D_{i\eta_\nu} = D(\frac{i}{4}r_\nu, i\eta_\nu s_\nu)$, $1 \le i \le 4$. Then the following hold for all $|t| \le 1$:*

1. $\phi_F^t : D_{3\eta_\nu}\times\mathcal{O}_{\nu+1} \to D_{2\eta_\nu}\times\mathcal{O}_\nu$;
2. *There holds*

$$\|D\phi_F^t - Id\|_{D_{1\eta_\nu},\mathcal{O}_\nu} \preccurlyeq \gamma^{-2}r_\nu^{-(2\tau+2)}\varepsilon_\nu,$$

where the weighted norm on the left hand side is the one induced from that of (5.23) for the vector-valued function ϕ_F^t.

Proof. We note that

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

Hence, by (5.25), we have

$$\|\phi_F^t - id\|_{D(\frac{r_\nu}{2},\frac{3}{4}s_\nu),\mathcal{O}_{\nu+1}} \le \|X_F\|_{D(\frac{r_\nu}{2},\frac{3}{4}s_\nu),\mathcal{O}_{\nu+1}} \preccurlyeq \gamma^{-2} r_\nu^{-(2\tau+2)} \varepsilon_\nu. \tag{5.26}$$

Since by (5.24)

$$\gamma^{-2} r_\nu^{-(2\tau+2)} \varepsilon_\nu^{2(\kappa-1)} \ll 1,$$

(1) follows easily from (5.26).

Since F is a polynomial of degree 1 in I and degree 2 in z and $\bar{z}$, it follows from the definition of the weighted norm (5.23), (5.25) and the Cauchy estimate that

$$\|D^m F\|_{D_2,\mathcal{O}_{\nu+1}} \preccurlyeq \gamma^{-2} r_\nu^{-(2\tau+2)} \varepsilon_\nu, \quad m \ge 2. \tag{5.27}$$

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 $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, (2) follows immediately from (5.27). □

5.2. **Estimates on perturbations.** Recall and note that

$$\begin{aligned} P_1' &= P_1 = \tilde{P}_{\sigma_1}, \\ P_{\nu+1} &= \int_0^1 \{R_\nu(t), F_\nu\} \circ \phi_{F_\nu}^t dt + (P_\nu' - R_\nu) \circ \phi_{F_\nu}^1, \\ P_{\nu+1}' &= P_{\nu+1} + (\tilde{P}_{\sigma_{\nu+1}} - \tilde{P}_{\sigma_\nu}) \circ \Phi_\nu \end{aligned} \tag{5.28}$$

where $R_\nu(t) = (1-t)(N_{\nu+1} - N_\nu) + tR_\nu$, $\nu = 1, 2, \cdots$. To estimate the weighted norms of the Hamiltonian vector fields $X_{P_\nu'}, X_{P_\nu}$, we need the following lemma.

Lemma 5.3. *Let G_1, G_2 be two Hamiltonian functions on $D(r,s)$ for some $r, s > 0$ which depend on parameters in a region $\mathcal{O}$ of an Euclidean space smoothly. If there are constants $\varepsilon', \varepsilon'' > 0$ such that*

$$\|X_{G_1}\|_{D(r,s),\mathcal{O}} < \varepsilon', \ \|X_{G_2}\|_{D(r,s),\mathcal{O}} < \varepsilon'',$$

then there is a constant $c_* > 0$ independent on any parameter such that

$$\|X_{\{G_1,G_2\}}\|_{D(r-\sigma,\eta s),\mathcal{O}} < c_* \sigma^{-1} \eta^{-2} \varepsilon' \varepsilon''$$

for all $0 < \sigma < r, 0 < \eta \ll 1$.

Proof. See [15, Lemma 7.3]. □

We are ready to prove the following result.

Proposition 5.1. (A1) *holds for all $\nu = 1, 2, \cdots$.*

Proof. We prove the proposition via induction. For $\nu = 1$, we have by Lemma 2.1 (1) that

$$\begin{aligned}
|P_1'|_{C^1(\Delta^1_{\sigma_1}),\mathcal{O}} &= |P_1|_{C^1(\Delta^1_{\sigma_1}),\mathcal{O}} = |\tilde{P}_{\sigma_1}|_{C^1(\Delta^1_{\sigma_1}),\mathcal{O}} \\
&\leq \left| \tilde{P}_{\sigma_1} - \sum_{|\varrho|\leq\alpha-1} \frac{\partial^\varrho \tilde{P}_1(\operatorname{Re} X)}{\varrho!}(\mathrm{iIm}X)^\varrho \right|_{C^1(\Delta^1_{\sigma_1}),\mathcal{O}} \\
&+ \left| \sum_{|\varrho|\leq\alpha-1} \frac{\partial^\varrho \tilde{P}_1(\operatorname{Re} X)}{\varrho!}(\mathrm{iIm}X)^\varrho \right|_{C^1(\Delta^1_{\sigma_1}),\mathcal{O}} \\
&\preccurlyeq \varepsilon^{\alpha-1}\sigma_1^{\alpha-1} + \varepsilon^{\alpha-1} \preccurlyeq \varepsilon^{\alpha-1}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\|X_{P_1'}\|_{D(r_1,s_1),\mathcal{O}_1} &= \|X_{P_1}\|_{D(r_1,s_1),\mathcal{O}_1} = \|X_{\tilde{P}_{\sigma_1}}\|_{D(r_1,s_1),\mathcal{O}_1} \\
&\preccurlyeq s_1^{-2}\varepsilon^{\alpha-1} \preccurlyeq \frac{s_1^{-2}\varepsilon^{3\kappa}}{2} \preccurlyeq \frac{s_1}{2},
\end{aligned}$$

i.e., (A1) holds for $\nu = 1$.

Using induction, we assume that (A1) holds up to some $\nu \geq 1$. By (5.28), we have

$$X_{P_{\nu+1}} = \int_0^1 (\phi^t_{F_\nu})^* X_{\{R_\nu(t),F_\nu\}} dt + (\phi^1_{F_\nu})^* X_{(P_\nu'-R_\nu)}.$$

By the expression of $R_\nu(t)$, (4.19), and Cauchy estimates, we have

$$\|X_{R_\nu}(t)\|_{D_{1\eta_\nu},\mathcal{O}_\nu} \preccurlyeq \|X_{P_\nu'}\|_{D(r_\nu,s_\nu),\mathcal{O}_\nu} \preccurlyeq \varepsilon_\nu.$$

It follows from Lemmas 5.1, 5.3 that

$$\|X_{\{R_\nu(t),F_\nu\}}\|_{D_{2\eta_\nu},\mathcal{O}_\nu} \preccurlyeq \gamma^{-2} r_\nu^{-(2\tau+2)} \eta_\nu^{-2} \varepsilon_\nu^2. \tag{5.29}$$

Since R_ν is the truncation of the Taylor-Fourier series of P_ν' up to order K_ν in the angular variables and up to order 2 in the tangential and normal variables, we have

$$|P_\nu' - R_\nu|_{D_{2\eta_\nu},\mathcal{O}_\nu} \preccurlyeq \eta_\nu^2 s_\nu^2 e^{-\frac{K_\nu r_\nu}{2}}.$$

It follows that

$$\|X_{(P_\nu'-R_\nu)}\|_{D_{2\eta_\nu},\mathcal{O}_\nu} \preccurlyeq \eta_\nu^2 \preccurlyeq \eta_\nu \varepsilon_\nu. \tag{5.30}$$

By Lemma 5.1, we also have

$$\|D\phi^t_F\|_{D_{1\eta_\nu},\mathcal{O}_{\nu+1}} \leq 1 + \|D\phi^t_F - Id\|_{D_{1\eta_\nu},\mathcal{O}_{\nu+1}} \leq 2, \ \ |t| \leq 1. \tag{5.31}$$

It now follows from (5.29)-(5.31) that

$$\begin{aligned}
\|X_{P_{\nu+1}}\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}} &\preccurlyeq \gamma^{-2} r_\nu^{-(2\tau+2)} \eta_\nu^{-2} \varepsilon_\nu^2 + \eta_\nu \varepsilon_\nu \\
&\preccurlyeq \gamma^{-2} \varepsilon_\nu^{4-2\kappa-(2\tau+2)q}.
\end{aligned}$$

Since $q = \frac{4-3\kappa}{2\tau+4}$, it is obviously that $4 - 2\kappa - (2\tau + 2)q > \kappa$, and therefore,

$$\|X_{P_{\nu+1}}\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}} \le \frac{\varepsilon_{\nu+1}}{2}. \tag{5.32}$$

For each $j = 1, 2, \cdots, \nu$, it is straightforward to check that

$$2^{j+1}\gamma^{-2}r_j^{-2\tau-2}\varepsilon_j = 2^{j+1}\gamma^{-2}r_j^2\varepsilon_j^{3(\kappa-1)} \le 2^{j+3}\gamma^{-2}\varepsilon_j^{3(\kappa-1)} \le 1.$$

It then follows from Lemma 5.2 that

$$\begin{aligned}
&|D\Phi_\nu|_{\Delta^{\nu+1}_{\sigma_{\nu+1}},\mathcal{O}_{\nu+1}}\\
&\quad\le \|(D\phi_1 \circ \phi_2 \circ \cdots \circ \phi_\nu)(D\phi_2 \circ \phi_3 \circ \cdots \circ \phi_\nu)\cdots(D\phi_\nu)\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}}\\
&\preccurlyeq \prod_{j=1}^{\nu}[1 + \gamma^{-2}(r_j)^{-2\tau-2}\varepsilon_j] \preccurlyeq \prod_{j=1}^{\nu}(1 + \frac{1}{2^{j+1}}) \preccurlyeq 1.
\end{aligned} \tag{5.33}$$

Hence by Lemma 2.1,

$$\begin{aligned}
&|D[(\tilde{P}_{\sigma_{\nu+1}} - \tilde{P}_{\sigma_\nu}) \circ \Phi_\nu]|_{\Delta^{\nu+1}_{\sigma_{\nu+1}},\mathcal{O}_{\nu+1}}\\
\le\ & |D(\tilde{P}_{\sigma_{\nu+1}} - \tilde{P}_{\sigma_\nu})|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}} \cdot |D\Phi_\nu|_{\Delta^{\nu+1}_{\sigma_{\nu+1}},\mathcal{O}_{\nu+1}}\\
\preccurlyeq\ & |\tilde{P}_{\sigma_{\nu+1}} - \tilde{P}_{\sigma_\nu}|_{C^1(\Delta^{\nu+1}_{\sigma_{\nu+1}}),\mathcal{O}_{\nu+1}} \preccurlyeq \sigma_\nu^{\alpha-1}(2b + 3 + 2\nu)^{1+\alpha},
\end{aligned}$$

and consequently,

$$\begin{aligned}
&\|X_{(\tilde{P}_{\sigma_{\nu+1}}-\tilde{P}_{\sigma_\nu})\circ\Phi_\nu}\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}} \preccurlyeq s_{\nu+1}^{-2}|D[(\tilde{P}_{\sigma_{\nu+1}} - \tilde{P}_{\sigma_\nu}) \circ \Phi_\nu]|_{\Delta^{\nu+1}_{\sigma_{\nu+1}},\mathcal{O}_{\nu+1}}\\
&\preccurlyeq s_{\nu+1}^{-2}\sigma_\nu^{\alpha-1}(2b + 3 + 2\nu)^{1+\alpha} \preccurlyeq \varepsilon_\nu^{\alpha q-q}s_{\nu+1}^{-2}(2b + 3 + 2\nu)^{1+\alpha}\\
&\preccurlyeq \varepsilon_\nu^{3\kappa^2}s_{\nu+1}^{-2}(2b + 3 + 2\nu)^{1+\alpha} \le \frac{1}{2}\varepsilon_{\nu+1}^{3\kappa-2\kappa}(2b + 3 + 2\nu)^{1+\alpha}\\
&= \frac{\varepsilon_{\nu+1}}{2}\varepsilon_{\nu+1}^{\kappa-1}(2b + 3 + 2\nu)^{1+\alpha} \le \frac{\varepsilon_{\nu+1}}{2}.
\end{aligned}$$

Combining the above with (5.32), we have

$$\|X_{P'_{\nu+1}}\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}} \le \frac{\varepsilon_{\nu+1}}{2} + \frac{\varepsilon_{\nu+1}}{2} = \varepsilon_{\nu+1}.$$

The prove is now completed. □

5.3. **Iteration lemma.** The preceding analysis can be summarized as follows.

Proposition 5.2. (Iteration Lemma) *For ε sufficiently small, the KAM iterations give rise to the following sequences of parametrized Hamiltonians*

$$\begin{aligned}
&H'_\nu = N_\nu + P'_\nu, \qquad H_\nu = N_\nu + P_\nu,\\
&N_\nu = e_\nu + \langle\omega_\nu, I\rangle + \langle A_\nu Z_\nu, \bar{Z}_\nu\rangle + \sum_{|n|\ge b+\nu} w_n\bar{w}_n\\
&= e_\nu + \langle\omega_\nu(\beta), I\rangle + \langle\tilde{A}_\nu Z_{\nu+1}, \bar{Z}_{\nu+1}\rangle + \sum_{|n|>b+\nu} w_n\bar{w}_n
\end{aligned}$$

defined on $D(r_\nu, s_\nu)\times\mathcal{O}_\nu$ *and transformations* $\phi_\nu : D(r_{\nu+1}, s_{\nu+1})\times\mathcal{O}_{\nu+1}\to D(r_\nu, s_\nu)\times\mathcal{O}_\nu$*, for* $\nu = 1, 2, \cdots$*, which satisfy the following properties for each* ν*:*

(1) H'_ν, H_ν, ϕ_ν *are real analytic in phase variables and depend on parameters smoothly;*
(2) ϕ_ν *and is symplectic for each* $\beta\in\mathcal{O}_\nu$ *and* $H'_\nu\circ\phi_\nu = H_{\nu+1}$*;*
(3) *There hold*

$$|e_{\nu+1}-e_\nu|_{\mathcal{O}_{\nu+1}}\le\varepsilon_\nu,\ |\omega_{\nu+1}-\omega_\nu|_{\mathcal{O}_{\nu+1}}\le\varepsilon_\nu,\ |A_{\nu+1}-\tilde{A}_\nu|_{\mathcal{O}_{\nu+1}}\le\varepsilon_\nu; \tag{5.34}$$

$$\|X_{P_\nu}\|_{D(r_\nu,s_\nu),\mathcal{O}_\nu}\le\frac{\varepsilon_\nu}{2},\ \|X_{P'_\nu}\|_{D(r_\nu,s_\nu),\mathcal{O}_\nu}\le\varepsilon_\nu, \tag{5.35}$$

$$\|\Phi_{\nu+1}-\Phi_\nu\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}}\preccurlyeq\gamma^{-2}r_{\nu+1}^{-(2\tau+2)}\varepsilon_{\nu+1}, \tag{5.36}$$

where $\Phi_\nu = \phi_1\circ\cdots\circ\phi_\nu$.

Proof. (5.35) is concluded in Proposition 5.1. With the validity of (5.35), Lemma 5.2 holds and hence $\phi_\nu : D(r_{\nu+1}, s_{\nu+1})\times\mathcal{O}_{\nu+1}\to D(r_\nu, s_\nu)\times\mathcal{O}_\nu$ is well defined for all $\nu\ge 1$. The analytic properties and smooth dependence on parameters of the Hamiltonians and transformations for all $\nu\ge 1$ are clear from the above constructions. As a time-1 map of a Hamiltonian flow, each ϕ_ν is symplectic for any fixed parameter value.

It follows from (4.19) that

$$\begin{aligned} e_{\nu+1} &= e_\nu + P^\nu_{0000}, \\ \omega_{\nu+1} &= \omega_\nu + (P^\nu_{0l00})_{|l|=1}, \\ A_{\nu+1} &= \tilde{A}_\nu + \begin{pmatrix} P^{11\nu}_0 & P^{11\nu}_{0n} \\ P^{11\nu}_{0n} & (P^{11\nu}_{0nm})_{|m|=b+\nu} \end{pmatrix}_{|n|=b+\nu}. \end{aligned}$$

Hence (5.34) follows from (5.35).

Since

$$\begin{aligned} &\|\Phi_{\nu+1}-\Phi_\nu\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}} \\ &\preccurlyeq\|D\Phi_\nu\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}}\cdot\|\phi_{\nu+1}-id\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}}, \end{aligned}$$

(5.36) follows from (5.26) and (5.33). □

6. Proof of main theorem

6.1. **Measure estimate.** For fixed $0<\gamma\ll 1$, let $\mathcal{O}_\gamma = \bigcap_{\nu=1}^{\infty}\mathcal{O}_\nu$. We have the following result.

Proposition 6.1. meas$(\mathcal{O}\setminus\mathcal{O}_\gamma) = O(\gamma)$.

Proof. We follow arguments in [14]. Note that for each $\nu\ge 1$,

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu\setminus\Big(\bigcup_{|k|\le K_\nu}\mathcal{R}^{\nu,1}_k\bigcup\mathcal{R}^{\nu,2}_k\bigcup\mathcal{R}^{\nu,3}_k\bigcup\mathcal{R}^{\nu,4}_k\Big),$$

where for each $|k| \le K_\nu$,

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

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

$$\mathcal{R}_k^{\nu,3} = \left\{ \beta \in \mathcal{O}_\nu : |(\langle k, \omega_\nu(\beta)\rangle I + I \otimes \tilde{A}_\nu(\beta) + \tilde{A}_\nu(\beta) \otimes I)^{-1}| > \frac{(2\nu)^4 K_\nu^\tau}{\gamma} \right\},$$

$$\mathcal{R}_k^{\nu,4} = \left\{ \beta \in \mathcal{O}_\nu : |(\langle k, \omega_\nu(\beta)\rangle I \pm I \otimes \tilde{A}_\nu(\beta) - \tilde{A}_\nu(\beta) \otimes I)^{-1}| > \frac{(2\nu)^4 K_\nu^\tau}{\gamma} \right\}.$$

For fixed $\nu \ge 1$, $|k| \le K_\nu$, and $\beta \in \mathcal{O}_\nu$, denote

$$A_{k,\nu}(\beta) = \langle k, \omega_\nu(\beta)\rangle I + I \otimes \tilde{A}_\nu(\beta) + \tilde{A}_\nu(\beta) \otimes I, \quad \beta \in \mathcal{O}_\nu.$$

Then there is an orthonormal matrix P such that

$$A_{k,\nu} = P^\top \Lambda P$$

where $\Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_{2\nu^2})$ with λ_j, $j = 1, 2, \cdots, 2\nu^2$, being eigenvalues of $A_{k,\nu}(\beta)$. If $\min\{|\lambda_1|, |\lambda_2|, \cdots, |\lambda_{(2\nu)^2}\} \ge \frac{\gamma}{K_\nu^\tau}$, then

$$|A_{k,\nu}^{-1}(\beta)| \le |P|^2 |\Lambda^{-1}| \le (2\nu)^4 \frac{K_\nu^\tau}{\gamma}.$$

It follows that

$$\begin{aligned} \tilde{\mathcal{R}}_k^{\nu,3} &= \left\{ \beta \in \mathcal{O}_\nu : \exists \text{some eigenvalue } \lambda_j(\beta) \text{ of } A_{k,\nu}(\beta) \text{ s.t. } |\lambda_j(\beta)| < \frac{\gamma}{K_\nu^\tau} \right\} \\ &\supset \mathcal{R}_k^{\nu,3}. \end{aligned}$$

Let $\lambda_j(\beta)$ be an eigenvalue branch of $A_{k,\nu}(\beta)$. Then it is well-known that λ_j depends on β smoothly and there exists an unit eigenvector ψ_j associated with λ_j which also depends on β smoothly. Since

$$\lambda_j(\beta) = \langle (\langle k, \omega_\nu(\beta)\rangle I + I \otimes \tilde{A}_\nu(\beta) - \tilde{A}_\nu(\beta) \otimes I)\psi_j(\beta), \psi_j(\beta)\rangle,$$

we have by using the unity of ψ_j that

$$\partial_\beta \lambda_j(\beta) = \langle \partial_\beta (\langle k, \omega_\nu(\beta)\rangle I + I \otimes \tilde{A}_\nu(\beta) - \tilde{A}_\nu(\beta) \otimes I)\psi_j(\beta), \psi_j(\beta)\rangle,$$

and consequently,

$$\partial_\beta \lambda_j(\beta) = O(|k|).$$

A standard measure estimate then yields that

$$\text{meas}(\mathcal{R}_k^{\nu,3}) \le \text{meas}(\mathcal{R}_k^{\nu,3}) \preccurlyeq \frac{(2\nu)^2 \gamma}{K_\nu^{\tau+1}}$$

and thus

$$\text{meas}(\bigcup_{|k| \le K_\nu} R_k^{\nu,3}) \preccurlyeq \sum_{|k| \le K_\nu} \frac{(2\nu)^2 \gamma}{K_\nu^{\tau+1}} \preccurlyeq \frac{(2\nu)^2 \gamma}{K_\nu^{\tau-2b}} \preccurlyeq \frac{\gamma}{K_\nu^{\tau-2b-1}}.$$

Similarly,

$$\text{meas}(\bigcup_{|k|\le K_\nu} \mathcal{R}_k^{\nu,i}) \preccurlyeq \frac{\gamma}{K_\nu^{\tau-2b-1}}, \quad i = 1,3,4.$$

Now,

$$\begin{aligned}\text{meas}(\mathcal{O}\setminus\mathcal{O}_\gamma) &\le \text{meas}(\bigcup_{\nu\ge 1}\bigcup_{|k|\le K_\nu}(\mathcal{R}_k^{\nu,1}\bigcup\mathcal{R}_k^{\nu,2}\bigcup\mathcal{R}_k^{\nu,3}\bigcup\mathcal{R}_k^{\nu,4}))\\ &\preccurlyeq \sum_{\nu\ge 1}\frac{\gamma}{K_\nu^{\tau-2b-1}} \preccurlyeq \sum_{\nu\ge1}\frac{\gamma}{K_\nu} = O(\gamma).\end{aligned}$$

□

6.2. **Convergence.** By (4.18) and Proposition 5.2, we see that

$$(\tilde N + \tilde P_{\sigma_\nu})\circ\Phi_\nu = N_{\nu+1} + P'_{\nu+1}, \tag{6.37}$$

hold on $D(r_{\nu+1}, s_{\nu+1})\times\mathcal{O}_{\nu+1}$ for all $\nu = 1,2,\cdots$.

By (5.34), $e_\nu(\beta)$, $\omega_\nu(\beta)$, and $\tilde A_\nu(\beta)$ converge uniformly, as $\nu\to\infty$, say to $e_\infty(\beta)$, $\omega_\infty(\beta) = \beta + O(\varepsilon)$, and $\tilde A_\infty(\beta)$, respectively, on $\mathcal{O}_\gamma$. It also follows from (5.34) and the standard Whitney extension theory that $\omega_\infty(\beta)$ is Whitney smooth on $\mathcal{O}_\gamma$.

Denote $\mathbb{T}_* = \mathbb{T}^{2b+1}\times\{0\}\times\{0\}\times\{0\}\subset D(r_1,s_1)$ and

$$N_\infty = e_\infty(\beta) + \langle\omega_\infty(\beta), I\rangle + \langle A_\infty(\beta)Z_\infty, \bar Z_\infty\rangle,$$

where $Z_\infty = (w_n)_{|n|\ge b+1}$. Since, by (5.35), $X_{P'_\nu}\to 0$ as $\nu\to\infty$ uniformly on $\mathbb{T}_*\times\mathcal{O}_\gamma$, we have

$$X_{N_{\nu+1}+P'_{\nu+1}} \to X_{N_\infty}, \qquad \nu\to\infty,$$

uniformly on $\mathbb{T}_*\times\mathcal{O}_\gamma$, and consequently,

$$\phi^t_{N_{\nu+1}+P'_{\nu+1}} \to \phi^t_{N_\infty}, \quad \nu\to\infty \tag{6.38}$$

in $C^0(\mathbb{T}_*\times\mathcal{O}_\gamma, D(r_1,s_1))$ uniformly on compact subsets of $\mathbb{R}$.

By (3.15), $\tilde P_{\sigma_\nu}$ converges, as $\nu\to\infty$, to $\tilde P$ in $C^1(\breve D_\rho\times\mathcal{O}_\gamma)$ and hence

$$X_{\tilde N+\tilde P_{\sigma_\nu}} \to X_{\tilde H}, \quad \nu\to\infty$$

uniformly on $\breve D_\rho\times\mathcal{O}_\gamma$, and consequently,

$$\phi^t_{\tilde N+\tilde P_{\sigma_\nu}} \to \phi^t_{\tilde H}, \quad \nu\to\infty \tag{6.39}$$

in $C^0(\breve D_\rho\times\mathcal{O}_\gamma, D(r_1,s_1))$ uniformly on compact subsets of $\mathbb{R}$.

By (5.36), we also see that Φ_ν converges on $\mathbb{T}_*\times\mathcal{O}_\gamma$ as $\nu\to\infty$ uniformly to a near identity mapping Φ_∞ in $C^0(T_*\times\mathcal{O}_\gamma, D(r_1,s_1))$.

Now, it follows from (6.37)-(6.39) and the convergence of $\{\Phi_\nu\}$ that on $\mathbb{T}_*\times\mathcal{O}_\gamma$,

$$\phi^t_{\tilde H}\circ\Phi_\infty = \Phi_\infty\circ\phi^t_{N_\infty}, \quad t\in\mathbb{R}. \tag{6.40}$$

6.3. **Invariant tori.** For each $\beta \in \mathcal{O}_\gamma$, $\mathbb{T}_* \times \{\beta\}$ is clearly an invariant $(2b+1)$-torus of $\phi^t_{N_\infty}$ with flow $\{\theta + \omega_\infty(\beta)t\}$. It follows from (6.40) that $\Phi_\infty(\mathbb{T}_* \times \{\beta\})$ is an embedded, invariant, quasi-periodic, $(2b+1)$-torus of the original Hamiltonian flow $\phi^t_{\tilde{H}}$ with frequency $\omega_\infty(\beta)$. This torus is also linearly stable due to the constancy of the normal part of N_∞. Moreover, in virtual of the weighted norm used in the range $D(r_1, s_1)$ of $\Phi_\infty(\mathbb{T}_* \times \{\beta\})$, we see that the normal components of the quasi-periodic orbits $\Phi_\infty(\theta + \omega_\infty(\beta)t, \beta)$ is spatially localized as $\frac{1}{|n|^{1+\alpha}}$ as $|n| \gg 1$.

It remains to show the smoothness of the invariant torus $\Phi_\infty(\mathbb{T}_* \times \{\beta\})$ for each $\beta \in \mathcal{O}_\gamma$.

Denote

$$\mathfrak{s}_* = \frac{(\kappa-1)(\alpha-1)}{\kappa^2} + 2.$$

Then for any $\mathfrak{s} \in [2, \mathfrak{s}_*)$ and $i \in \mathbb{Z}_+^{2b+1}$ with $|i| = \mathfrak{s}$, we have by (5.36) and Cauchy estimates that

$$\begin{aligned} |\partial_\theta^i(\Phi_{\nu+1} - \Phi_\nu)|_{\mathbb{T}_*, \mathcal{O}_\gamma} &\le |\partial_\theta^i(\Phi_{\nu+1} - \Phi_\nu)|_{D(r_{\nu+1}, s_{\nu+1}), \mathcal{O}_\gamma} \\ \preccurlyeq\ \gamma^{-2}\varepsilon_{\nu+1}^{3(\kappa-1)} r_{\nu+1}^{-\mathfrak{s}} &\preccurlyeq \varepsilon_{\nu+1}^{3(\kappa-1)-q\mathfrak{s}} = \varepsilon_{\nu+1}^{3(\kappa-1)\frac{\mathfrak{s}_*-\mathfrak{s}}{\mathfrak{s}_*-2}}. \end{aligned}$$

It follows that $\{\Phi_\nu\}$ actually converges to Φ_∞ in $C^{\mathfrak{s},1}(\mathbb{T}_* \times \mathcal{O}_\gamma, D(r_1, s_1))$. Hence $\Phi_\infty(\mathbb{T}_* \times \{\beta\})$ is a $C^{\mathfrak{s}}$-torus for each $\beta \in \mathcal{O}_\gamma$.

The proof of the Main Theorem is now completed.

References

[1] Aubry, S., Anti-integrability in dynamical and variational problems. *Phys. D,* **86** (1995), no. 1-2, 284–296.

[2] Aubry, S., Discrete breathers: localization and transfer of energy in discrete Hamiltonian nonlinear systems. *Phys. D,* **216** (2006), no. 1, 1–30.

[3] Aubry, S., Johansson, M., Existence and stability of quasiperiodic breathers in the discrete nonlinear Schrödinger equation. *Nonlinearity* **10** (1997), no. 5, 1151–1178.

[4] Aubry, S., Kopidakis, G., Kadelburg, V. Variational proof for hard discrete breathers in some classes of Hamiltonian dynamical systems. *Discrete Contin. Dyn. Syst. Ser. B* **1** (2001), no. 3, 271–298.

[5] Bambusi, D., Asymptotic stability of breathers in some Hamiltonian networks of weakly coupled oscillators. *Commun. Math. Phys.* **324** (2013), no. 2, 515–547.

[6] Bambusi, D., Ponno, A., On metastability in FPU. *Commun. Math. Phys.* **264** (2006), no. 2, 539–561.

[7] Bambusi, D., Vella, D., Quasi periodic breathers in Hamiltonian lattices with symmetries. *Discrete Contin. Dyn. Syst. Ser. B* **2** (2002), no. 3, 389–399.

[8] Berti, M., Bolle, P., Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential, *J. Eur. Math. Soc.,* **15**(2013), 229–286.

[9] Berti, M., Bolle, P., A Nash-Moser approach to KAM theory, *Fields Institute Communications,* **special volume Hamiltonian PDEs and Applications,** 255–284, 2015.

[10] Berti, M., Kappeler, T., Montalto, R., Large KAM tori for perturbations of the defocusing NLS equation, *Astérisque,* **403** (2018,) 1–160

[11] Cuevas, J., James, G., Kevrekidis, P., Breathers in oscillator chains with Hertzian interactions. *Phys. D* **251** (2013), 39–59.

[12] Chierchia, L., Qian, D., Moser's theorem for lower dimensional tori, *J. Differ. Eqs.,* **206**(2004), 55–93.
[13] Gao, M., Zhang, K., Quasi-periodic solutions for the derivative nonlinear Schrödinger equation with finitely differentiable nonlinearities. *Chin. Ann. Math. Ser. B* **38** (2017), no. 3, 759C786.
[14] Geng, J., Viveros, J., Yi, Y., Quasi-periodic breathers in Hamiltonian networks of long-range coupling. *Phys. D,* **237** (2008), no. 22, 2866–2892.
[15] Geng, J., You, J., A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces. *Commun. Math. Phys.,* **262** (2006), 343–372.
[16] Geng, J., Yi, Y., A KAM theorem for Hamiltonian networks with long ranged couplings. *Nonlinearity* **20** (2007), no. 6, 1313–1342.
[17] Herman, M., Sur les courbes invariantes par les difféomorphismes de lanneau. Vol. **1**. (French) [On the curves invariant under diffeomorphisms of the annulus. Vol. **1**] With an appendix by Albert Fathi. With an English summary. *Astérisque,* 103-104. Société Mathématique de France, Paris, 1983, pp. i+221.
[18] James, G., Nonlinear waves in Newton's Cradle and the discrete $p-$Schrödinger equation, *Math. Models Methods Appl. Sci.,* **21** (2011), no. 11, 2335–2377.
[19] James, G., Periodic travelling waves and compactons in granular chains. *J. Nonlinear Sci.,* **22**(2012), no. 5, 813–848.
[20] James, G., Pelinovsky, D., Gaussian solitary waves and compactons in Fermi-Pasta-Ulam lattices with Hertzian potentials. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.,* **470** (2014), no. 2165, 20130462, 20 pp.
[21] Pöschel, J., Integrability of Hamiltonian systems on Cantor sets, *Comm. Pure Appl. Math.,* **35**(1982), 653–695.
[22] Pöschel, J., Quasi-periodic solutions for a nonlinear wave equation, *Comment. Math. Helvetici,* **71**(1996), 269–296.
[23] Pöschel, J., A KAM theorem for some nonlinear partial differential equations, *Ann. Sc. Norm. sup. Pisa CI. sci.,* **23**(1996), 119–148.
[24] Salamon, D., Zehnder, E., KAM theory in configuration space. *Comment. Math. Helv.* **64** (1989), no. 1, 84–132.
[25] You, J., Perturbation of lower dimensional tori for Hamiltonian systems. *J. Differ. Eqs.,* **152**(1999), 1–29.
[26] Yuan, X., Construction of quasi-periodic breathers via KAM technique. *Comm. Math. Phys.,* **226**(2002), no. 1, 61–100.
[27] Zehnder, E., Generalized implicit function theorems with applications to small divisor problems I, *Comm. Pure Appl. Math.* **28** (1975) 91–140.
[28] Zehnder, E., Generalized implicit function theorems with applications to small divisor problems II, *Comm. Pure Appl. Math.* **29** (1976) 49–113.

C. Ge: Department of Mathematical & Statistical Sci, Nanjing University of Information Science & Technology, Nanjing 210044, PRC
E-mail address: gechuanfang@126.com

J. Geng: Department of Mathematics, Nanjing University, Nanjing 210093, PRC
E-mail address: jgeng@nju.edu.cn

Y. Yi: Department of Mathematical & Statistical Sci, University of Alberta, Edmonton, Alberta, Canada T6G 2G1 and School of Mathematics, Jilin University, Changchun, 130012, PRC
E-mail address: yingfei@ualberta.ca