AIP | Journal of Mathematical Physics

Real analytic quasi-periodic solutions for the derivative nonlinear Schrödinger equations

Jiansheng Geng and Jian Wu

Citation: J. Math. Phys. 53, 102702 (2012); doi: 10.1063/1.4754822
View online: http://dx.doi.org/10.1063/1.4754822
View Table of Contents: http://jmp.aip.org/resource/1/JMAPAQ/v53/i10
Published by the American Institute of Physics.

JOURNAL OF MATHEMATICAL PHYSICS **53**, 102702 (2012)

Real analytic quasi-periodic solutions for the derivative nonlinear Schrödinger equations

Jiansheng Geng[a)] and Jian Wu[b)]
Department of Mathematics, Nanjing University, Nanjing 210093, People's Republic of China

(Received 19 March 2012; accepted 10 September 2012; published online 4 October 2012)

In this paper, we show that one dimension derivative nonlinear Schrödinger equation admits a whitney smooth family of small amplitude, real analytic quasi-periodic solutions with two Diophantine frequencies. The proof is based on a partial Birkhoff normal form reduction and an abstract infinite dimensional Kolmogorov-Arnold-Moser (KAM) theorem. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4754822]

I. INTRODUCTION AND MAIN RESULT

Consider the derivative nonlinear Schrödinger equation

$$\mathrm{i}u_t - u_{xx} - \mathrm{i}(|u|^4 u)_x = 0, \tag{1.1}$$

with the periodic boundary condition

$$u(t, x) = u(t, x + 2\pi). \tag{1.2}$$

The equation defines an infinite dimensional Hamiltonian system

$$u_t = \{u(x), H\} = \frac{d}{dx}\frac{\partial H}{\partial \bar{u}},$$

associated with the Hamiltonian

$$H = \int_0^{2\pi} \frac{1}{3}|u|^6 - \mathrm{i}u_x\bar{u}\, dx.$$

Let $\mu_j, \phi_j, j \in \mathbb{Z}$, denote the eigenvalues, eigenfunctions of $-\partial_{xx}$, and γ_j are fixed weights (see Ref. 1). The problem of the existence of time quasi-periodic solution for (1.1) is to find, for a given integer $N > 1$, a solution of the form

$$u = \sum_{j\neq 0} \gamma_j q_j \phi_j,$$

such that all q_j, $j \in \mathbb{Z}$, are quasi-periodic with the same N-frequencies. Then the Hamiltonian is

$$H = H(q, \bar{q}) = H(\sum_{j\neq 0} \gamma_j q_j \phi_j, \sum_{j\neq 0} \bar{\gamma}_j \bar{q}_j \bar{\phi}_j). \tag{1.3}$$

Motivated by the classical Kolmogorov-Arnold-Moser (KAM) theory in finite dimensional Hamiltonian systems, to obtain quasi-periodic solutions, for an integer $N > 1$, one finds a parameter space $O \subset \mathbb{R}^N$ and (symplectic) action-angle-normal coordinates $I = (I_1, \ldots, I_N)$,

a)E-mail: jgeng@nju.edu.cn.
b)E-mail: wujian0987@gmail.com.

$\theta = (\theta_1, \ldots, \theta_N)$, $z = (z_j)_{j\in\mathbb{Z}}$ such that the Hamiltonian (1.3) can be transformed into a parametrized Hamiltonian normal form

$$H(I, \theta, z, \bar{z}, \xi) = \langle \omega(\xi), I\rangle + \sum_{j\in\mathbb{Z}} \bar{\Omega}_j(\xi) z_j \bar{z}_j + P(I, \theta, z, \bar{z}, \xi), \quad \xi \in \mathcal{O}.$$

Denote d as the order of $\bar{\Omega}_j$ and δ as the order of the Hamiltonian vector field X_P. We use d to measure the growth rate of the eigenvalue $\bar{\Omega}_j$ and δ to measure the unboundedness of X_P.

When $\delta \le 0$, the vector field X_P is called bounded perturbation, the existence of quasi-periodic solutions of such PDEs has been deeply and widely studied. There are plenty of papers dedicated to such these issues. See Refs. 2–6, 10, and 12 for example.

When $\delta > 0$, the vector field X_P is called unbounded perturbation. In order to obtain the existence of quasi-periodic solutions for such PDEs, it is reasonable to assume

$$\delta \le d - 1,$$

according to a well known example, which was referred in Klainerman[8] and Lax.[11]

When $0 < \delta < d - 1$, the corresponding KAM theorem is due to Kuksin.[9] Kuksin's theorem is used to prove the existence of time quasi-periodic solutions of KdV equation subject to periodic boundary condition. See also Kappeler-Pöschel.[7] For such unbounded perturbations, Bambusi-Graffi[1] gave another KAM theorem to prove that the time dependent linear Schrödinger equation has only pure point spectrum.

When $0 < \delta = d - 1$, the nonlinearity of the PDE is the strongest. Liu-Yuan[17] gave a theorem, which generalized Kuksin's theorem from $\delta < d - 1$ to $\delta \le d - 1$. For the homological equations of variable coefficients

$$-\,\mathrm{i}\partial_\omega u + \lambda u + \mu(\theta) u = p(\theta), \quad |\mathrm{Im}\theta| < s, \tag{1.4}$$

they developed some new estimate for the solution u covering not only the case $\delta < d - 1$, but also the limit case $\delta = d - 1$. Using the generalized Kuksin's theorem, Zhang-Gao-Yuan[19] established a KAM theorem for an infinite dimensional reversible system with unbounded perturbation and obtained many C^∞ (**not real analytic**) quasi-periodic solutions for the following derivative nonlinear Schrödinger equations:

$$\begin{cases} iu_t + u_{xx} + |u_x|^2 u = 0, & (t, x) \in \mathbb{R} \times [0, \pi], \\ u(t, 0) = 0 = u(t, \pi), \end{cases} \tag{1.5}$$

which are reversible instead of Hamiltonian.

According to the generalized Kuksin's theorem, Liu-Yuan[18] established an improved KAM theorem such that the range of application is extended to a class of derivative nonlinear Schrödinger equations which are Hamiltonian

$$\mathrm{i}u_t + u_{xx} - M_\sigma u + \mathrm{i} f(u, \bar{u}) u_x = 0, \quad f(-u, -\bar{u}) = -f(u, \bar{u}), \tag{1.6}$$

with Dirichlet boundary conditions. Compared with the Eq. (1.1), the above Eq. (1.6) has the external parameters and the nonlinearity is a little artificial, which constitutes our main motivation why we consider the quasi-periodic solutions of the Eq. (1.1).

In Ref. 18, the authors mentioned that for a class of derivative nonlinear Schrödinger equations

$$\mathrm{i}u_t - u_{xx} - \mathrm{i}(|u|^2 u)_x = 0,$$

subject to periodic boundary conditions, the multiplicity of the normal frequency Ω_j is essentially equal to 1, i.e., $\Omega_j^\sharp = 1$. According to the KAM theorem, which was established by Liu-Yuan,[17] one can obtain that the Eq. (1.1) with periodic boundary conditions admit many C^∞ (**not real analytic**) quasi-periodic solutions with N Diophantine frequencies. It should be worth emphasizing that our obtained quasi-periodic solutions in this paper are real analytic. In this sense, our results cannot be covered by the KAM theorem established by Liu-Yuan.[17]

In our paper, we consider that $N = 2$. Together with compact form and the gauge invariant property, the homological equation has the following forms:

$$-\mathrm{i}\partial_\omega u + \lambda u = p(\theta).$$

Compared with the linearized Eq. (1.4), we see that $\mu(\theta) = 0$. Thus, our normal form is independent of the angle variables θ, which is also the main difference from Kuksin's theorem[9] (see also Kappeler-Pöschel[7]) and Liu-Yuan's theorem.[17] Afterwards, using an abstract KAM theorem with angle independent normal form, we obtain the real analytic quasi-periodic solution for the derivative nonlinear Schrödinger Equation (1.1).

Assume that

$$[u] = 0, \quad \text{where } [u] \text{ is the mean value of the function } u, \tag{1.7}$$

then the main result is described as follows:

Theorem 1. *Consider the derivative nonlinear Schrödinger Equation (1.1) with periodic boundary condition (1.2) and (1.7). Fix n_1, n_2 satisfying $n_1 - n_2$ is odd. Then there exists a Cantor subset $\widetilde{\mathcal{O}} = \widetilde{\mathcal{O}}(n_1, n_2) \subset \mathbb{R}^2_+$ of positive Lebesgue measure, such that each $\xi \in \widetilde{\mathcal{O}}$ corresponds to a real analytic quasi-periodic solution*

$$u(t, x) = \sum_{j=1}^{2} \sqrt{2\pi |n_j| \xi_j} e^{\mathrm{i}(\omega_{*j} t + n_j x)} + O(|\xi|^{\frac{5}{2}})$$

of (1.1), (1.2), (1.7) with the Diophantine frequencies

$$\omega_{*j} = n_j^2 + O(|\xi|^2), \quad 1 \le j \le 2.$$

Moreover, the quasi-periodic solutions u are linearly stable and depend on ξ Whitney smoothly.

Remark 1.1. Note that the solution which we obtain in Theorem 1 is real analytic. However, the quasi-periodic solution for the Eq. (1.6) in Ref. 18 is just only smooth, i.e., $u \in C^\infty$. The reason is the construction of the radius of Imθ. *In Ref. 18, the number s must be defined as the following manner: in the mth KAM iteration, $s_m = 2^{-m}$, so $s_m \to 0$ as $m \to \infty$. But in this paper, we only need $s_m \to \frac{s}{2}$ as $m \to \infty$ (see Sec. VIII for details).*

Remark 1.2. The assumption $[u] = 0$ is reasonable. Otherwise, writing $u = c + v$ with $[v] = 0$ and $c = [u]$, also let $v = \sum_{j \neq 0} \gamma_j q_j \phi_j$, the Hamiltonian then takes the form $H = H_c + \frac{2\pi}{3} c^6$ with

$$H_c = \sum_{j \neq 0} (2\pi \sigma_j j^2 + 4\pi c^4 \sigma_j j) |q_j|^2 + \sum_{j \neq 0} 2\pi c^4 \sigma_j j (q_j q_{-j} + \bar{q}_j \bar{q}_{-j})$$
$$+ \sum_{k,l,m} G^1_{klm} \sqrt{|klm|} q_k \bar{q}_l \bar{q}_m + \cdots + \sum_{i,j,k,l,m,n} G_{ijklmn} \sqrt{|ijklmn|} q_i q_j q_k \bar{q}_l \bar{q}_m \bar{q}_n.$$

The terms of the form $q_j q_{-j}$, $\bar{q}_j \bar{q}_{-j}$ will make the perturbation large by action of the Poisson bracket and we cannot deal with this case.

The rest of the paper is devoted to the proof of the main result. Here, we only treat the case $N = 2$. Section II is a preliminary section in which we define the phase space, admissible b-index, compact form and gauge invariant property. In Sec. III, we will introduce the Hamiltonian setting corresponding to the Eq. (1.1). In Sec. IV, we will derive a partial Birkhoff normal form of order six for the lattice Hamiltonian. In Sec. V, we will verify some conditions about frequencies and the perturbation. In Secs. VI–IX, we will introduce an infinite dimensional KAM theorem for constructing quasi-periodic solutions and cover the proof of this KAM theorem. Some lemmas necessary are given in the Appendix.

II. PRELIMINARY

A. Phase space

For any integer $p \ge 0$, we introduce the phase space

$$\mathcal{H}^{a,p} = \{u \in L^2(S^1, \mathbb{C}) : \|u\|_{a,p} < \infty\}$$

of complex valued functions on $S^1 = \mathbb{R}/2\pi\mathbb{Z}$, where

$$\|u\|_{a,p}^2 = \sum_{j\in\mathbb{Z}, j\neq 0} |\widehat{u}_j|^2 |j|^{2p} e^{2a|j|}$$

is defined in terms of the Fourier transform $\widehat{u}$ of u, $u(x) = \sum_{j\in\mathbb{Z}, j\neq 0} \widehat{u}_j e^{\mathrm{i}jx}$.

Let

$$\ell^{a,p} = \{q = (q_j) : q_j \in \mathbb{C},\ j \neq 0,\ \|q\|_{a,p} < \infty\}$$

be the space of all complex sequences with

$$\|q\|_{a,p}^2 = \sum_{j\in\mathbb{Z}, j\neq 0} |q_j|^2 |j|^{2p} e^{2a|j|}.$$

The convolution $w * z$ of two such sequences is defined by $(w * z) = \sum_m w_{n-m} z_m$.

Lemma 2.1. For $a > 0$, $p > \frac{1}{2}$, The space $\ell^{a,p}$ is a Banach algebra with respect to convolution of sequences, and

$$\|w * z\|_{a,p} \leq c\|w\|_{a,p}\|z\|_{a,p},$$

with a constant c depending only on p.

For the proof, see Ref. 13. ■

B. Admissible *b*-index

An ordered b-index $\{n_1, \ldots, n_b\}$ is said to be *admissible* if whenever i, j, k, l, m, n are integers such that $i + j + k = l + m + n$, $\{i, j, k\} \neq \{l, m, n\}$, and at least four of them lie in $\{n_1, \ldots, n_b\}$, then $\mu_i + \mu_j + \mu_k - \mu_l - \mu_m - \mu_n \neq 0$, where $\mu_k = k^2$. We let $\mathcal{J}$ denote the set of all admissible b-index. It is known that for any given $b > 1$, $\mathcal{J}$ is an infinite set (see the Appendix of Ref. 13). In particular, when $b = 2$,

$$\mathcal{J} = \{\{n_1, n_2\} : n_1 - n_2 \text{ is odd }\}.$$

Remark 2.1. Without the loss of generality, we assume that $n_1 > n_2 > 0$ for simplicity.

C. Compact form

Given $\{n_1, n_2\} \in J$. A real analytic function

$$F = F(\theta, I, z, \bar{z}) = \sum_{k,\alpha,\beta} F_{k\alpha\beta}(I) e^{i\langle k,\theta\rangle} z^\alpha \bar{z}^\beta$$

is said to admit a *compact form with respect to n_1, n_2* if

$$F_{k\alpha\beta} = 0, \quad \text{whenever } k_1 n_1 + k_2 n_2 + \sum_n (\alpha_n - \beta_n) n \neq 0,$$

where $k = (k_1, k_2) \in \mathbb{Z}^2$ and $\alpha = (\cdots, \alpha_n, \cdots)$, $\beta = (\cdots, \beta_n, \cdots)$, $\alpha_n, \beta_n \in \mathbb{N}$, with finitely many nonzero components of positive integers. In the case that $k_1 = k_2 = 0$, we simply say that F has a *compact form.*

Consider the Poisson bracket

$$\{F, G\} = 2\pi[\langle\frac{\partial F}{\partial\theta}, \frac{\partial G}{\partial I}\rangle - \langle\frac{\partial F}{\partial I}, \frac{\partial G}{\partial\theta}\rangle + i\sum_{n\in\mathbb{Z}} \sigma_n(\frac{\partial F}{\partial z_n}\frac{\partial G}{\partial \bar{z}_n} - \frac{\partial F}{\partial \bar{z}_n}\frac{\partial G}{\partial z_n})],$$

where $\sigma_n = \mathrm{Sgn}(n)$ is the sign of n.

Lemma 2.2. Given $\{n_1, n_2\} \in J$ and consider two real analytic functions $F(\theta, I, z, \bar{z})$, $G(\theta, I, z, \bar{z})$. If both F and G have compact forms with respect to n_1, n_2, then so does $\{F, G\}$.

Proof: Let

$$F = \sum_{A_1} F_{k_1\alpha_1\beta_1}(I)e^{\mathrm{i}\langle k_1,\theta\rangle} z^{\alpha_1}\bar{z}^{\beta_1}, \quad G = \sum_{A_2} G_{k_2\alpha_2\beta_2}(I)e^{\mathrm{i}\langle k_2,\theta\rangle} z^{\alpha_2}\bar{z}^{\beta_2},$$

where

$$A_i = \{(k_i, \alpha_i, \beta_i) : k_{i_1}n_1 + k_{i_2}n_2 + \sum_n(\alpha_{in} - \beta_{in})n = 0\}.$$

then

$$\begin{aligned}
\{F, G\} = {} & 2\pi \sum_{A_1}\sum_{A_2}\langle \mathrm{i}k_1, \frac{\partial G_{k_2\alpha_2\beta_2}(I)}{\partial I}\rangle F_{k_1\alpha_1\beta_1}(I)e^{\mathrm{i}\langle k_1,\theta\rangle} z^{\alpha_1}\bar{z}^{\beta_1} e^{\mathrm{i}\langle k_2,\theta\rangle} z^{\alpha_2}\bar{z}^{\beta_2} \\
& -2\pi \sum_{A_1}\sum_{A_2}\langle \mathrm{i}k_2, \frac{\partial F_{k_1\alpha_1\beta_1}(I)}{\partial I}\rangle G_{k_2\alpha_2\beta_2}(I)e^{\mathrm{i}\langle k_1,\theta\rangle} z^{\alpha_1}\bar{z}^{\beta_1} e^{\mathrm{i}\langle k_2,\theta\rangle} z^{\alpha_2}\bar{z}^{\beta_2} \\
& +2\pi\mathrm{i} \sum_n \sigma_n\alpha_{1n}\beta_{2n}(\sum_{A_1} F_{k_1\alpha_1\beta_1}(I)e^{\mathrm{i}\langle k_1,\theta\rangle} z^{\alpha_1-e_n}\bar{z}^{\beta_1})(\sum_{A_2} G_{k_2\alpha_2\beta_2}(I)e^{\mathrm{i}\langle k_2,\theta\rangle} z^{\alpha_2}\bar{z}^{\beta_2-e_n}) \\
& -2\pi\mathrm{i} \sum_n \sigma_n\alpha_{2n}\beta_{1n}(\sum_{A_1} F_{k_1\alpha_1\beta_1}(I)e^{\mathrm{i}\langle k_1,\theta\rangle} z^{\alpha_1}\bar{z}^{\beta_1-e_n})(\sum_{A_2} G_{k_2\alpha_2\beta_2}(I)e^{\mathrm{i}\langle k_2,\theta\rangle} z^{\alpha_2-e_n}\bar{z}^{\beta_2}),
\end{aligned}$$

where for $n \in \mathbb{Z}$, e_n is the multi-index whose nth component is 1 and other components are all 0. Since all terms above have compact forms with respect with n_1, n_2, so does $\{F, G\}$. ■

D. Gauge invariant property

A real analytic function

$$F(\theta, I, z, \bar{z}) = \sum_{k,\alpha,\beta} F_{k\alpha\beta}(I)e^{i\langle k,\theta\rangle} z^\alpha\bar{z}^\beta$$

is said to have *gauge invariant property* if

$$F_{k\alpha\beta} = 0, \quad \text{whenever } k_1 + k_2 + \sum_n(\alpha_n - \beta_n) \neq 0.$$

Lemma 2.3. Consider two real analytic functions $F(\theta, I, z, \bar{z})$, $G(\theta, I, z, \bar{z})$. If both F and G have gauge invariant property, then so does $\{F, G\}$.

The proof of this lemma is similar to Lemma 2.2, so we omit it. ■

Lemma 2.4. Given $\{n_1, n_2\} \in \mathcal{J}$ and consider a real analytic function $F(\theta, I, z, \bar{z})$. If F has compact form with respect to n_1, n_2 and gauge invariant property, then F contains no terms of the form $e^{i\langle k,\theta\rangle}z_n\bar{z}_n$ with $k \neq 0$.

Proof: Consider $F(\theta, I, z, \bar{z}) = \sum_{k,\alpha,\beta} F_{k\alpha\beta}(I)e^{i\langle k,\theta\rangle}z^\alpha\bar{z}^\beta$ with $\alpha = \beta = e_n$, where e_n denotes the nth component being 1 and the other components being 0. Since F has the compact form, we have

$$k_1n_1 + k_2n_2 + n - n = 0,$$

similarly, because of the gauge invariant property, we have

$$k_1 + k_2 + 1 - 1 = 0.$$

Together with $n_1 \neq n_2$, we obtain that $k_1 = k_2 = 0$. So, Lemma 2.4 is proved. ■

Remark 2.2. A crucial idea in the proof of our main theorem is to show that compact forms and the gauge invariant property will be preserved along KAM iterations. These properties enable the consideration of essentially finite small divisors and simplify the homologic equation in each KAM step.

We denote

$$\mathcal{A} = \{P : P = \sum_{k\in\mathbb{Z}^2, l\in\mathbb{N}^2,\alpha,\beta} P_{kl\alpha\beta}(\xi)e^{\mathrm{i}\langle \mathrm{k},\theta\rangle} I^l z^\alpha \bar{z}^\beta\},$$

where k, α, β have the following relations:

$$k_1n_1 + k_2n_2 + \sum_n(\alpha_n - \beta_n)n = 0, \quad k_1 + k_2 + \sum_n(\alpha_n - \beta_n) = 0.$$

III. HAMILTONIAN OF THE DERIVATIVE NONLINEAR SCHRÖDINGER EQUATION

In this section, we will give the Hamiltonian setting of the derivative nonlinear Schrödinger Equation (1.1). To set the stage, we endow $\mathcal{H}^{a,p}$ with the Poisson structure

$$\{F, G\} = \int_{S^1} \frac{\partial F}{\partial u(x)}(\frac{d}{dx}\frac{\partial G}{\partial \bar{u}(x)}) + \frac{\partial F}{\partial \bar{u}(x)}(\frac{d}{dx}\frac{\partial G}{\partial u(x)})\, dx,$$

where F, G are differentiable functions on $\mathcal{H}^{a,p}$ with L^2-gradients in $\mathcal{H}^{a,p-1}$. The Hamiltonian corresponding to the Eq. (1.1) with the periodic boundary condition (1.2) and (1.7) is then given by

$$u_t = \{u(x), H\} = \frac{d}{dx}\frac{\partial H}{\partial \bar{u}},$$

where

$$H = \int_0^{2\pi} \frac{1}{3}|u|^6 - \mathrm{i}u_x\bar{u}\, dx. \tag{3.1}$$

To write the Hamiltonian system more explicitly as an infinite dimensional system, we introduce infinitely many coordinates $q = (q_j)_{j\neq 0}$ by writing

$$u = \mathcal{F}(q) = \sum_{j\neq 0} \gamma_j q_j \phi_j, \tag{3.2}$$

where $\{\phi_j = \sqrt{\frac{1}{2\pi}}e^{\mathrm{i}jx}\}$ is an orthonormal basis, $\gamma_j = \sqrt{2\pi|j|}$ are fixed positive weights, the sequence $q = (q_j)_{j\neq 0}$ is an element of the Banach space $\ell^{a,p}$. Due to the choice of the weights, we have an isomorphism $\mathcal{F} : \ell^{a,p+\frac{1}{2}} \to \mathcal{H}^{a,p}$ for each $p \geq 1$.

The phase space $\ell^{a,p} \times \ell^{a,p}$ is endowed with the Poisson structure

$$\{F, G\} = 2\pi\mathrm{i}\sum_{j\neq 0}\sigma_j(\frac{\partial F}{\partial q_j}\frac{\partial G}{\partial \bar{q}_j} - \frac{\partial F}{\partial \bar{q}_j}\frac{\partial G}{\partial q_j}), \quad \sigma_j = \mathrm{Sgn}(j),$$

and the equations of motion in the new coordinates are given by

$$\dot{q}_j = \mathrm{i}\sigma_j\frac{\partial H}{\partial \bar{q}_j}, \quad j \neq 0.$$

This is most easily seen by observing that

$$\frac{\partial F}{\partial u} = \sum_{j\neq 0}\frac{\partial F}{\partial q_j}\frac{\partial q_j}{\partial u} = \sum_{j\neq 0}\frac{\partial F}{\partial q_j}\gamma_j^{-1}\phi_j^{-1},$$

and calculating $\{F, G\}$ on $\mathcal{H}^{a,p}$. Since the transformed Poisson structure is nondegenerate, it also defines a symplectic structure $\mathrm{i}\sum_{j\neq 0}\sigma_j \mathrm{d}q_j \wedge \mathrm{d}\bar{q}_j$.

The Hamiltonian expressed in the new coordinates q is determined by inserting the expansion (3.2) of q into the Hamiltonian (3.1). Using for simplicity the same symbol for the Hamiltonian as a function of q, we obtain

$$H = \Lambda + G \tag{3.3}$$

with

$$\Lambda = \sum_{j\neq 0} 2\pi\sigma_j j^2 |q_j|^2,$$

$$G = \sum_{i,j,k,l,m,n} G_{ijklmn}\sqrt{|ijklmn|}q_i q_j q_k \bar{q}_l \bar{q}_m \bar{q}_n = \sum_{\alpha,\beta} G_{\alpha\beta} q^\alpha \bar{q}^\beta, \tag{3.4}$$

where

$$G_{ijklmn} = \frac{8}{3}\pi^3 \int_0^{2\pi} \phi_i\phi_j\phi_k\bar{\phi}_l\bar{\phi}_m\bar{\phi}_n \, \mathrm{d}x. \tag{3.5}$$

Remark 3.1. By taking advantage of the special forms of G in (3.4) and G_{ijklmn} in (3.5), we know that $G \in \mathcal{A}$, i.e., $G_{\alpha\beta} = 0$ whenever $\sum_n(\alpha_n - \beta_n)n \neq 0$ or $\sum_n(\alpha_n - \beta_n) \neq 0$. Especially, $G_{ijkijk} = \frac{2}{3}\pi$.

Now we consider the regularity of the gradient of G.

Lemma 3.1. For $a > 0$ and $p > \frac{3}{2}$, the gradient $G_{\bar{q}}$ is a real analytic map from a neighborhood of the origin of $\ell^{a,p}$ into $\ell^{a,p-1}$, with

$$\|G_{\bar{q}}\|_{a,p-1} \leq c\|q\|^5_{a,p}.$$

Proof: Let $G_{\bar{q}} = (\{\frac{\partial G}{\partial \bar{q}_n}\})$, where

$$\frac{\partial G}{\partial \bar{q}_n} = \sqrt{|n|}\sum_{i+j+k-l-m=n}\frac{2\pi}{3}\sqrt{|ijklm|}q_i q_j q_k \bar{q}_l \bar{q}_m.$$

Defining $g_n = \frac{2\pi}{3}\sum_{i+j+k-l-m=n}\sqrt{|i|}q_i\sqrt{|j|}q_j\sqrt{|k|}q_k\sqrt{|l|}\bar{q}_l\sqrt{|m|}\bar{q}_m$, $g = (g_n)$ and $w = (\sqrt{|k|}q_k)$, we see that

$$g = cw * w * w * \bar{w} * \bar{w}.$$

Also, we have

$$\begin{aligned}\|w\|^2_{a,p-\frac{1}{2}} &= \sum_k (\sqrt{|k|}q_k)^2 |k|^{2(p-\frac{1}{2})}e^{2a|k|},\\ &= \sum_k |q_k|^2|k|^{2p}e^{2a|k|},\\ &= \|q\|^2_{a,p},\end{aligned}$$

from Lemma 2.1,

$$\|g\|_{a,p-\frac{1}{2}} \leq c\|q\|^5_{a,p},$$

then

$$\begin{aligned}\|\frac{\partial G}{\partial \bar{q}}\|^2_{a,p-1} &= \sum_n (\sqrt{|n|}g_n)^2|n|^{2(p-1)}e^{2a|n|},\\ &= \sum_n |g_n|^2|n|^{2(p-\frac{1}{2})}e^{2a|n|},\\ &= \|g\|^2_{a,p-\frac{1}{2}},\\ &\leq c\|q\|^{10}_{a,p}.\end{aligned}$$

Hence,

$$\|G_{\bar{q}}\|_{a,p-1} \leq c\|q\|^5_{a,p}. \qquad \blacksquare$$

IV. PARTIAL BIRKHOFF NORMAL FORM

In this section, we transform the Hamiltonian (3.3) into a partial Birkhoff normal form. For fixed $\{n_1, n_2\} \in \mathcal{J}$, we define the index sets Δ_*, $* = 0, 1, 2$, and Δ_3 in the following way: Δ_* is the set of index (i, j, k, l, m, n) such that there exist right $*$ components not in $\{n_1, n_2\}$. Δ_3 is the set of the index (i, j, k, l, m, n) such that there exist at least three components not in $\{n_1, n_2\}$. Define the resonance set $\mathcal{N} = \{(i, j, k, i, j, k)\} \cap \Delta_0$, $\mathcal{M} = \{(i, j, k, i, j, k)\} \cap \Delta_2$.

Lemma 4.1. Let $(i, j, k, l, m, n) \in (\Delta_0 \setminus \mathcal{N}) \cup \Delta_1 \cup (\Delta_2 \setminus \mathcal{M})$. *If*

$$i + j + k = l + m + n,$$

then

$$\mu_i + \mu_j + \mu_k - \mu_l - \mu_m - \mu_n = i^2 + j^2 + k^2 - l^2 - m^2 - n^2 \neq 0.$$

For the proof, see Ref. 6. $\blacksquare$

Furthermore, the analogue of Lemma 4.1 is more delicate:

Lemma 4.2. Let $(i, j, k, l, m, n) \in \Delta_2 \setminus \mathcal{M}$. *If*

$$i + j + k = l + m + n, \quad |n| \geq \max\{|3n_1 - n_2|, |3n_2 - n_1|\},$$

then

$$|\mu_i + \mu_j + \mu_k - \mu_l - \mu_m - \mu_n| = |i^2 + j^2 + k^2 - l^2 - m^2 - n^2| \geq \frac{|n|}{2}.$$

Proof: We assume without the loss of generality that $\{i, j, k\} \cap \{l, m, n\} = \emptyset$. Since $(i, j, k, l, m, n) \in \Delta_2 \setminus \mathcal{M}$, we either have (a) $i = j = k = n_1$, $l = n_2$, and $m, n \neq n_1, n_2$; or (b) $i = j = n_1$, $l = m = n_2$, and $k, n \neq n_1, n_2$.

In case (a), since

$$i + j + k - l - m - n = 3n_1 - n_2 - m - n = 0,$$

we have $m = 3n_1 - n_2 - n$. Consequently,

$$\begin{aligned}
|\mu_i + \mu_j + \mu_k - \mu_l - \mu_m - \mu_n| &= |3n_1^2 - n_2^2 - n^2 - (3n_1 - n_2 - n)^2|,\\
&= |2(n - \frac{3n_1 - n_2}{2})^2 + \frac{3}{2}(n_1 - n_2)^2|,\\
&\geq 2(|n| - |\frac{3n_1 - n_2}{2}|)^2.
\end{aligned}$$

If $|n| \geq |3n_1 - n_2|$, we obtain

$$2(|n| - |\frac{3n_1 - n_2}{2}|)^2 \geq \frac{|n|^2}{2} \geq \frac{|n|}{2}.$$

In case (b), since

$$i + j + k - l - m - n = 2n_1 - 2n_2 + k - n = 0,$$

we have $k = n - 2n_1 + 2n_2$. Consequently,

$$\begin{aligned}|\mu_i + \mu_j + \mu_k - \mu_l - \mu_m - \mu_n| &= |2n_1^2 - 2n_2^2 - n^2 + (n - 2n_1 + 2n_2)^2|,\\ &= |6n_1^2 + 2n_2^2 - 4nn_1 + 4nn_2 - 8n_1n_2|,\\ &= |2(n_2 - n_1)(2n - (3n_1 - n_2))|,\\ &\geq 4\big||n| - |\frac{3n_1 - n_2}{2}|\big|.\end{aligned}$$

If $|n| \geq |3n_1 - n_2|$, we obtain

$$4\big||n| - |\frac{3n_1 - n_2}{2}|\big| \geq 2|n| \geq \frac{|n|}{2}.$$

So, we complete the proof. ■

For our convenience, rewrite $G = G^0 + G^1 + G^2 + \widehat{G}$, where

$$G^* = \frac{2}{3}\pi \sum_{\substack{i+j+k-l-m-n=0\\(i,j,k,l,m,n)\in\Delta_*}} \sqrt{|ijklmn|}\, q_i q_j q_k \bar{q}_l \bar{q}_m \bar{q}_n, \quad * = 0, 1, 2,$$

and

$$\widehat{G} = \frac{2}{3}\pi \sum_{\substack{i+j+k-l-m-n=0\\(i,j,k,l,m,n)\in\Delta_3}} \sqrt{|ijklmn|}\, q_i q_j q_k \bar{q}_l \bar{q}_m \bar{q}_n.$$

Lemma 4.3. Given $\{n_1, n_2\} \in \mathcal{J}$, there exists a real analytic, symplectic change of coordinates Γ in a neighborhood of the origin of $\ell^{a,p}$, which transforms the Hamiltonian (3.3) into the partial Birkhoff normal form

$$H \circ \Gamma = \Lambda + \bar{G} + \widehat{G} + K,$$

such that the corresponding Hamiltonian vector fields $X_{\bar{G}}$, $X_{\widehat{G}}$, and X_K are real analytic from a neighborhood of the origin in $\ell^{a,p}$ to $\ell^{a,p-1}$, where

$$\begin{aligned}\Lambda &= \sum_{j\neq 0} 2\pi\sigma_j j^2 |q_j|^2,\\ \bar{G} &= \frac{2}{3}\pi(n_1^3|q_{n_1}|^6 + n_2^3|q_{n_2}|^6) + 6\pi(n_1^2 n_2 |q_{n_1}|^4|q_{n_2}|^2 + n_1 n_2^2 |q_{n_1}|^2 |q_{n_2}|^4)\\ &\quad + 6\pi \sum_{n\neq 0, n_1, n_2} (n_1^2|q_{n_1}|^4 + n_2^2|q_{n_2}|^4 + 4n_1n_2|q_{n_1}|^2|q_{n_2}|^2)|n||q_n|^2,\\ \widehat{G} &= \frac{2}{3}\pi \sum_{\substack{i+j+k-l-m-n=0\\(i,j,k,l,m,n)\in\Delta_3}} \sqrt{|ijklmn|}\, q_i q_j q_k \bar{q}_l \bar{q}_m \bar{q}_n,\\ \|K_{\bar{q}}\|_{a,p-1} &\leq c\|q\|_{a,p}^9.\end{aligned}$$

Moreover, $K(q, \bar{q}) \in \mathcal{A}$.

Proof: Let $\Gamma = X_F^t\,|_{t=1}$ be the time 1-map of the flow of the Hamiltonian vector field X_F given by the Hamiltonian

$$\begin{aligned}F &= F^0 + F^1 + F^2\\ &= \sum_{i,j,k,l,m,n} F^0_{ijklmn} q_i q_j q_k \bar{q}_l \bar{q}_m \bar{q}_n\\ &\quad + \sum_{i,j,k,l,m,n} F^1_{ijklmn} q_i q_j q_k \bar{q}_l \bar{q}_m \bar{q}_n\\ &\quad + \sum_{i,j,k,l,m,n} F^2_{ijklmn} q_i q_j q_k \bar{q}_l \bar{q}_m \bar{q}_n\end{aligned}$$

with coefficients

$$\mathrm{i}F^0_{ijklmn}=\begin{cases}\frac{\sqrt{|ijklmn|}}{6\pi(i^2+j^2+k^2-l^2-m^2-n^2)} & i+j+k-l-m-n=0,(i,j,k,l,m,n)\in\Delta_0\setminus\mathcal{N},\\ 0 & otherwise,\end{cases}$$

$$\mathrm{i}F^1_{ijklmn}=\begin{cases}\frac{\sqrt{|ijklmn|}}{6\pi(i^2+j^2+k^2-l^2-m^2-n^2)} & i+j+k-l-m-n=0,(i,j,k,l,m,n)\in\Delta_1,\\ 0 & otherwise,\end{cases}$$

$$\mathrm{i}F^2_{ijklmn}=\begin{cases}\frac{\sqrt{|ijklmn|}}{6\pi(i^2+j^2+k^2-l^2-m^2-n^2)} & i+j+k-l-m-n=0,(i,j,k,l,m,n)\in\Delta_2\setminus\mathcal{M},\\ 0 & otherwise.\end{cases}$$

It follows from Lemma 4.1, Lemma 4.2, and the definition of the sets Δ_*, $*=0,1,2$ that F is well defined.

To show the analyticity of the transformation, we note that there exists a constant $c>0$ such that for each $n\in\mathbb{Z}$,

$$|\frac{\partial F}{\partial\bar{q}_n}|\le c\sum_{i+j+k-l-m=n}|q_iq_jq_k\bar{q}_l\bar{q}_m|=c(q*q*q*\bar{q}*\bar{q})_n.$$

Hence, by Lemma 2.1

$$\|F_{\bar{q}}\|_{a,p}\le c\|q*q*q*\bar{q}*\bar{q}\|_{a,p}\le c\|q\|^5_{a,p}.$$

The analyticity of $F_{\bar{q}}$ then follows from that of each of its component and its local boundedness (see the Appendix A of Ref. 16).

Since $\Gamma=X^1_F$, then

$$H\circ\Gamma=\Lambda+\bar{G}+\widehat{G}+K,$$

where

$$\begin{aligned}\bar{G}&=G^0+G^1+G^2+\{\Lambda,F\},\\ &=\frac{2}{3}\pi(n_1^3|q_{n_1}|^6+n_2^3|q_{n_2}|^6)+6\pi(n_1^2n_2|q_{n_1}|^4|q_{n_2}|^2+n_1n_2^2|q_{n_1}|^2|q_{n_2}|^4)\\ &\quad+6\pi\sum_{n\neq0,n_1,n_2}(n_1^2|q_{n_1}|^4+n_2^2|q_{n_2}|^4+4n_1n_2|q_{n_1}|^2|q_{n_2}|^2)|n||q_n|^2,\\ \widehat{G}&=\frac{2}{3}\pi\sum_{\substack{i+j+k-l-m-n=0\\(i,j,k,l,m,n)\in\Delta_3}}\sqrt{|ijklmn|}q_iq_jq_k\bar{q}_l\bar{q}_m\bar{q}_n,\\ K&=\{G,F\}+\frac{1}{2!}\{\{\Lambda,F\},F\}+\frac{1}{2!}\{\{G,F\},F\}\\ &\quad+\cdots+\frac{1}{n!}\{\cdots\{\Lambda,\underbrace{F\}\cdots,F}_{n}\}+\frac{1}{n!}\{\cdots\{G,\underbrace{F\}\cdots,F}_{n}\}+\cdots.\end{aligned}$$

It is clear that $\|K_{\bar{q}}\|_{a,p-1}\le c\|q\|^9_{a,p}$. To show that K has a compact form, we note that since G has a compact form, so does F. Hence, by Lemma 2.2, $\{G,F\}$ has a compact form. Note that Λ is already in a compact form. Repeated applications of Lemma 2.2 show that all terms in K have compact forms, so does K. Similarly, using Lemma 2.3, we can get that K has the gauge invariant property. Hence, $K\in\mathcal{A}$. ∎

Now our Hamiltonian is $\widetilde{H}=\Lambda+\bar{G}+\widehat{G}+K$. Introduce the symplectic polar and complex coordinates by setting

$$\begin{cases}q_{n_j}=\sqrt{I_j+\xi_j}e^{\mathrm{i}\theta_j}, & j=1,2,\\ q_j=z_j, & j\in\mathbb{Z}_1=\mathbb{Z}\setminus\{0,n_1,n_2\},\end{cases}$$

depending on parameters $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2_+$, where $(I, \theta) \in \mathbb{R}^2 \times \mathbb{T}^2$ be the standard action-angle variables in the $(q_{n_1}, q_{n_2}, \bar{q}_{n_1}, \bar{q}_{n_2})$-space around ξ. Then one gets

$$\mathrm{i}\sum_{j\neq 0}\sigma_j \mathrm{d}q_j \wedge \mathrm{d}\bar{q}_j = \sum_{j=1,2}\mathrm{d}\theta_j \wedge \mathrm{d}I_j + \mathrm{i}\sum_{j\in\mathbb{Z}_1}\sigma_j \mathrm{d}z_j \wedge \mathrm{d}\bar{z}_j,$$

and the Poisson bracket

$$\{F, G\} = 2\pi\sum_{j=1,2}\left(\frac{\partial F}{\partial\theta_j}\frac{\partial G}{\partial I_j} - \frac{\partial F}{\partial I_j}\frac{\partial G}{\partial\theta_j}\right) + 2\pi\mathrm{i}\sum_{j\in\mathbb{Z}_1}\sigma_j\left(\frac{\partial F}{\partial z_j}\frac{\partial G}{\partial\bar{z}_j} - \frac{\partial F}{\partial\bar{z}_j}\frac{\partial G}{\partial z_j}\right). \tag{4.1}$$

The new Hamiltonian

$$\widetilde{H} = \Lambda + \bar{G} + \widehat{G} + K = \langle\tilde{\omega}(\xi), I\rangle + \sum_{j\in\mathbb{Z}_1}\widetilde{\Omega}_j(\xi)z_j\bar{z}_j + \widetilde{P},$$

where $\tilde{\omega}(\xi) = (\tilde{\omega}_1(\xi), \tilde{\omega}_2(\xi))$ with

$$\tilde{\omega}_1(\xi) = 2\pi n_1^2 + 6\pi n_1(n_1\xi_1 + n_2\xi_2)^2 - 4\pi n_1^3\xi_1^2,$$

$$\tilde{\omega}_2(\xi) = 2\pi n_2^2 + 6\pi n_2(n_1\xi_1 + n_2\xi_2)^2 - 4\pi n_2^3\xi_2^2,$$

$$\widetilde{\Omega}_j(\xi) = 2\pi\sigma_j[j^2 + (6(n_1\xi_1 + n_2\xi_2)^2 - 3(n_1^2\xi_1^2 + n_2^2\xi_2^2))j], \quad j \in \mathbb{Z}_1,$$

$$\begin{aligned}\widetilde{P} = {} & K + O(|I|^3) + O(|\xi||I|^2) + O(|I|^2\sum_{j\in\mathbb{Z}_1}|j||z_j|^2) + O(|\xi||I|\sum_{j\in\mathbb{Z}_1}|j||z_j|^2)\\ & + O(|\xi|^{\frac{3}{2}}\sum_{i=1}^{2}\sum_{k+l-m=n_i}\sqrt{|k|}|z_k|\sqrt{|l|}|z_l|\sqrt{|m|}|\bar{z}_m|)\\ & + O(|\xi|\sum_{k+l-m-n=0}\sqrt{|k|}|z_k|\sqrt{|l|}|z_l|\sqrt{|m|}|\bar{z}_m|\sqrt{|n|}|\bar{z}_n|).\end{aligned} \tag{4.2}$$

Now, let $\varepsilon > 0$ be sufficiently small. Rescaling ξ_j by $\varepsilon^2\sqrt{\xi_j}$, $j = 1, 2$, $z, \bar{z}$ by $\varepsilon^2 z, \varepsilon^2\bar{z}$, and I by $\varepsilon^4 I$, one obtains the rescaled Hamiltonian

$$H(I, \theta, z, \bar{z}, \xi) = \varepsilon^{-8}\widetilde{H}(\varepsilon^4 I, \theta, \varepsilon^2 z, \varepsilon^2\bar{z}, \varepsilon^2\sqrt{\xi_1}, \varepsilon^2\sqrt{\xi_2}), \tag{4.3}$$

$$= \langle\omega^*(\xi), I\rangle + \sum_{j\in\mathbb{Z}_1}\Omega_j^*(\xi)z_j\bar{z}_j + \varepsilon P^*(I, \theta, z, \bar{z}, \xi), \tag{4.4}$$

where $\omega^*(\xi) = (\omega_1^*(\xi), \omega_2^*(\xi))$ with

$$\omega_1^*(\xi) = \varepsilon^{-4}2\pi n_1^2 + 6\pi n_1(n_1\sqrt{\xi_1} + n_2\sqrt{\xi_2})^2 - 4\pi n_1^3\xi_1,$$

$$\omega_2^*(\xi) = \varepsilon^{-4}2\pi n_2^2 + 6\pi n_2(n_1\sqrt{\xi_1} + n_2\sqrt{\xi_2})^2 - 4\pi n_2^3\xi_2,$$

$$\Omega_j^*(\xi) = \varepsilon^{-4}2\pi\sigma_j j^2 + 6\pi\sigma_j j[2(n_1\sqrt{\xi_1} + n_2\sqrt{\xi_2})^2 - (n_1^2\xi_1 + n_2^2\xi_2)], \quad j \in \mathbb{Z}_1,$$

$$vspace* -10pt\, P^*(\theta, I, z, \bar{z}, \xi) = \varepsilon^{-9}\widetilde{P}(\varepsilon^4 I, \theta, \varepsilon^2 z, \varepsilon^2\bar{z}, \varepsilon^2\sqrt{\xi_1}, \varepsilon^2\sqrt{\xi_2}).$$

Note that the nonlinear Schrödinger Equation (1.1) has another conserved quantity $\int_0^{2\pi}|u|^2\mathrm{d}x = \sum_{j\neq 0}\gamma_j^2|q_j|^2 = c$, i.e.,

$$2\pi n_1|q_{n_1}|^2 + 2\pi n_2|q_{n_2}|^2 + \sum_{j\in\mathbb{Z}_1}2\pi|j||q_j|^2 = c.$$

The above rescaling yields that

$$\varepsilon^2(n_1 I_1 + n_2 I_2) + n_1\sqrt{\xi_1} + n_2\sqrt{\xi_2} + \varepsilon^2 \sum_{j\in\mathbb{Z}_1} |j||z_j|^2 = c,$$

i.e.,

$$n_1\sqrt{\xi_1} + n_2\sqrt{\xi_2} = c - \varepsilon^2(n_1 I_1 + n_2 I_2 + \sum_{j\in\mathbb{Z}_1} |j||z_j|^2) = c + O(\varepsilon^2).$$

Let $\omega(\xi) = (\omega_1(\xi), \omega_2(\xi))$, $\bar{\Omega} = (\bar{\Omega}_j)_{j\neq 0, n_1, n_2}$, where

$$\omega_i(\xi) = \varepsilon^{-4} 2\pi n_i^2 + 6\pi c^2 n_i - 4\pi n_i^3 \xi_i, \quad i = 1, 2, \tag{4.5}$$

$$\bar{\Omega}_j(\xi) = \varepsilon^{-4} 2\pi\sigma_j j^2 + 6\pi\sigma_j j(2c^2 - n_1^2\xi_1 - n_2^2\xi_2), \quad j \in \mathbb{Z}_1.$$

We can rewrite (4.3) as

$$H(I, \theta, z, \bar{z}, \xi) = \langle\omega(\xi), I\rangle + \sum_{j\in\mathbb{Z}_1} \bar{\Omega}_j(\xi) z_j \bar{z}_j + P(I, \theta, z, \bar{z}, \xi, \varepsilon), \tag{4.6}$$

where

$$\begin{aligned} P = {} & \varepsilon P^* - 12\pi c\varepsilon^2(n_1 I_1 + n_2 I_2 + \sum_{j\in\mathbb{Z}_1} |j||z_j|^2)(n_1 I_1 + n_2 I_2 + 2\sum_{j\in\mathbb{Z}_1} |j||z_j|^2) \\ & + 6\pi\varepsilon^4(n_1 I_1 + n_2 I_2 + \sum_{j\in\mathbb{Z}_1} |j||z_j|^2)^2(n_1 I_1 + n_2 I_2 + 2\sum_{j\in\mathbb{Z}_1} |j||z_j|^2). \end{aligned} \tag{4.7}$$

Remark 4.1. From the above discussion, it is clear that the $z-$ independent terms in $P(I, \theta, z, \bar{z})$ are bounded and $\|P_z\|_{a,p-1} \leq \varepsilon$.

Remark 4.2. Because of the form of the Poisson bracket (4.1), define $\Omega = (\Omega_j)_{j\in\mathbb{Z}_1}$, where $\Omega_j = \sigma_j\bar{\Omega}_j = \varepsilon^{-4} 2\pi j^2 + 6\pi j(2c^2 - n_1^2\xi_1 - n_2^2\xi_2)$.

V. VERIFYING SOME CONDITIONS

In the following paragraphs, we will make use of the KAM method to prove the existence of the quasi-periodic solutions. To make it quantitative, we introduce the following notations and spaces.

The phase space is $\mathcal{P}^{a,p} = \mathbb{T}^2 \times \mathbb{R}^2 \times \ell^{a,p} \times \ell^{a,p}$ with the coordinates $(\theta, I, z, \bar{z})$, where $\mathbb{T}^2 = \mathbb{R}^2/2\pi\mathbb{Z}^2$. With $\mathcal{P}^{a,p}_{\mathbb{C}}$ we denote the complexification of the phase space $\mathcal{P}^{a,p}$. Denote a neighborhood of $\mathbb{T}^2_0 = \mathbb{T}^2 \times \{I = 0\} \times \{z = 0\} \times \{\bar{z} = 0\}$ by

$$\begin{aligned} D(s, r) &= \{(\theta, I, z, \bar{z}) : |\mathrm{Im}\,\theta| < s, |I| < r^2, \|z\|_{a,p} < r, \|\bar{z}\|_{a,p} < r\} \\ &\subset \mathbb{C}^2 \times \mathbb{C}^2 \times \ell^{a,p}_{\mathbb{C}} \times \ell^{a,p}_{\mathbb{C}} \\ &= \mathcal{P}^{a,p}_{\mathbb{C}}, \end{aligned}$$

where $|\cdot|$ denotes the sup-norm of complex vectors, and $\ell^{a,p}_{\mathbb{C}}$ is the complexification of $\ell^{a,p}$.

Let $\mathcal{O}$ be a neighborhood of the origin in $\mathbb{R}^2_+$. Denote by $\Delta_{\xi\zeta}$ the difference operator in the variable $\xi \in \mathcal{O}$

$$\Delta_{\xi\zeta} = f(\cdot, \xi) - f(\cdot, \zeta). \tag{5.1}$$

For $l = (l_1, \ldots, l_k) \in \mathbb{Z}^k$, we denote by $|l| = \sum_{j=1}^k |l_j|$ its length, and $\langle l\rangle_\delta = \max(1, |\sum_{j\neq 0}|j|l_j| \cdot \sum_{j\neq 0}|j|^\delta|l_j|)$. We set $\Pi = \{(k, l) \neq 0, |l| \leq 2\} \subset \mathbb{Z}^2 \times \mathbb{Z}^\infty$.

In preparation for the proof of Theorem 1, we first give some propositions about the frequencies and the perturbation P. The first two propositions concern the frequencies of the Hamiltonian in the normal form (4.6).

Proposition 5.1 (Nondegeneracy). The map $\xi \mapsto \omega(\xi)$ is a homeomorphism from $\mathcal{O}$ to its image, which is Lipschitz continuous and its inverse also. Moreover, for $\forall(k, l) \in \Pi$, the resonance set

$$\mathcal{R}_{kl} = \{\xi \in \mathcal{O} : \langle k, \omega(\xi)\rangle + \langle l, \Omega(\xi)\rangle = 0\} \tag{5.2}$$

has Lebesgue measure zero, and there exists a constant $m > 0$ such that for all $\xi \in \mathcal{O}$,

$$|\langle l, \Omega(\xi)\rangle| \geq m\langle l\rangle_{d-1}, \ \forall 1 \leq |l| \leq 2. \tag{5.3}$$

Proof: Rewrite $\omega(\xi)$, $\Omega(\xi)$ as $\omega(\xi) = \alpha + A\xi$, $\Omega(\xi) = \beta + B\xi$, where

$$\alpha = (\varepsilon^{-4}2\pi n_1^2 + 6\pi c^2 n_1, \varepsilon^{-4}2\pi n_2^2 + 6\pi c^2 n_2), \quad A = \begin{pmatrix} -4\pi n_1^3 & 0 \\ 0 & -4\pi n_2^3 \end{pmatrix},$$

and

$$\beta = (\varepsilon^{-4}2\pi j^2 + 12c^2\pi j)_{j\in\mathbb{Z}_1}, \quad B = \begin{pmatrix} -6\pi n_1^2 j & -6\pi n_2^2 j \\ \vdots & \vdots \end{pmatrix}_{j\in\mathbb{Z}_1}. \tag{5.4}$$

Since $\det A = 16\pi^2 n_1^3 n_2^3 \neq 0$, we have that the map $\xi \mapsto \omega(\xi)$ is a homeomorphism and $|\langle k, \omega(\xi)\rangle| \not\equiv 0$, for $k \neq 0$.

For the remaining nonresonant conditions, one has to check that $\langle \alpha, k\rangle + \langle \beta, l\rangle \neq 0$ or $Ak + B^T l \neq 0$ for $1 \leq |l| \leq 2$. Suppose $Ak + B^T l = 0$, for some $k \in \mathbb{Z}$ and $1 \leq |l| \leq 2$. We let $l_i, l_j, i \neq j$, $i \neq 0, j \neq 0$ be the components of l. Then

$$2k_1 n_1 = 2k_2 n_2 = -3(l_i i + l_j j).$$

Because the perturbation has the compact form, then

$$k_1 n_1 + k_2 n_2 + l_i i + l_j j = 0,$$

as a consequence,

$$2k_1 n_1 - \frac{2}{3}k_1 n_1 = \frac{4}{3}k_1 n_1 = 0,$$

hence

$$k_1 = 0 = k_2, l_i i + l_j j = 0.$$

The integer solutions to the equation $l_i i + l_j j = 0$ with $1 \leq |l| \leq 2$ are $l_i = l_j = \pm 1$ and $i = -j$. Because the perturbation has the gauge invariant property, then the term with $k_1 = k_2 = 0$ and $l_i = l_j = \pm 1$ will not appear in the perturbation. Consequently, nonresonant conditions are satisfied. ■

Proposition 5.2 (Spectral asymptotics). There exists $d > 1$ and $\delta \leq d - 1$ such that the following holds. First,

$$|\Omega_i - \Omega_j| \geq m(|i|^{d-1} + |j|^{d-1}), \tag{5.5}$$

for all $i \neq j$ uniformly on $\mathcal{O}$ with constant $m > 0$. Second, the functions

$$\xi \to \frac{\Omega_j(\xi)}{|j|^\delta}$$

are uniformly Lipschitz on $\mathcal{O}$ for $j \neq 0$.

Proof: First, thanks to (5.4), we have that for $i, j \in \mathbb{Z}_1, i \neq j, \exists\, m > 0, \ s.t.$

$$\begin{aligned} |\Omega_i - \Omega_j| &\geq |\varepsilon^{-4}2\pi i^2 + 6\pi i(2c^2 - n_1^2\xi_1 - n_2^2\xi_2) - \varepsilon^{-4}2\pi j^2 - 6\pi j(2c^2 - n_1^2\xi_1 - n_2^2\xi_2)| \\ &\geq m(|i| + |j|). \end{aligned}$$

Second,

$$\begin{aligned}|\frac{\Omega_j(\xi)}{|j|^\delta}-\frac{\Omega_j(\zeta)}{|j|^\delta}| &= |6\pi\frac{j(n_1^2(\xi_1-\zeta_1)+n_2^2(\xi_2-\zeta_2))}{|j|^\delta}| \\ &\le 6\pi\max\{n_1^2,\ n_2^2\}|\xi-\zeta|,\end{aligned}$$

here, $d=2$, $\delta=1$. Then Proposition 5.2 is fulfilled. ∎

J. Pöschel [see Corollary C and its proof in Ref. 15] proves that there exists a finite set $\mathcal{O}_0\subset\mathcal{O}$ with meas$(\mathcal{O}\backslash\mathcal{O}_0)\to 0$ when $\alpha\to 0$, such that for all $\xi\in\mathcal{O}_0$,

$$|\langle k,\omega(\xi)\rangle+\langle l,\Omega(\xi)\rangle|\ge\frac{\alpha\langle l\rangle_\delta}{|k|^\tau},\ \forall(k,l)\in\Pi,\quad k\neq 0. \tag{5.6}$$

In the sequel, we will need the distance

$$|\Omega-\Omega'|_{-\delta,\mathcal{O}}=\sup_{\xi\in\mathcal{O}}\sup_{j\in\mathbb{Z}_1} j^{-\delta}|\Omega_j(\xi)-\Omega'_j(\xi)|,$$

and the semi-norm

$$|\Omega|^{\mathcal{L}}_{-\delta,\mathcal{O}}=\sup_{\substack{\xi,\zeta\in\mathcal{O}\\ \xi\neq\zeta}}\sup_{j\in\mathbb{Z}_1} j^{-\delta}\frac{|\Delta_{\xi\zeta}\Omega_j|}{|\xi-\zeta|}.$$

Finally, we get

$$|\omega|^{\mathcal{L}}_{\mathcal{O}}+|\Omega|^{\mathcal{L}}_{-\delta;\mathcal{O}}=M(=4\pi n_1^3+6\pi n_1^2).$$

The third proposition is concerned with the perturbed Hamiltonian vector field, $X_P=(\partial_\theta P,-\partial_I P,\{i\sigma_j\partial_{z_j}P\},\{-i\sigma_j\partial_{\bar z_j}P\})$. In order to describe the proposition, we then define the weighted norms

$$\|X_P\|_{r,q;D(s,r)\times\mathcal{O}}:=\sup_{D(s,r)}(|P_I|+\frac{1}{r^2}|P_\theta|+\frac{1}{r}\|P_z\|_{a,p}+\frac{1}{r}\|P_{\bar z}\|_{a,p}),$$

and

$$\|X_P\|^{\mathcal{L}}_{r,q;D(s,r)\times\mathcal{O}}:=\sup_{\substack{\xi,\zeta\in\mathcal{O}\\ \xi\neq\zeta}}\frac{\|\Delta_{\xi\zeta}X_P\|_{r,q}}{|\xi-\zeta|},$$

where $\Delta_{\xi\zeta}X_P=X_P(\cdot,\xi)-X_P(\cdot,\zeta)$.

Proposition 5.3 (Regularity of perturbation). There exists a neighborhood $D(s,r)$ of $\mathbb{T}^2_0$ in $\mathcal{P}^{a,p}_{\mathbb{C}}$ such that P is defined on $D(s,r)\times\mathcal{O}$, and its Hamiltonian vector field defines a map

$$X_P:D(s,r)\times\mathcal{O}\to\mathcal{P}^{a,q}_{\mathbb{C}},$$

where $q\ge 0$ satisfies $p-q\le\delta\le d-1$. Moreover, $X_P(\cdot,\xi)$ is real analytic on $D(s,r)$ for each $\xi\in\mathcal{O}$, and $X_P(w,\cdot)$ is uniformly Lipschitz on $\mathcal{O}$ for each $w\in D(s,r)$.

Proof: We first show that $P_z\in\ell^{a,p-1}$. From (4.7), it is clear that $\|P_z\|_{a,p-1}\le c\varepsilon\|z\|_{a,p}$. The other components of X_P can be handled in the same way, and we get $X_P\in\mathcal{P}^{a,q}_{\mathbb{C}}$, here, $q=p-1$, $d=2$ and $\delta=1$. We now turn to the Lipschitz norms. Because of the form of $\widetilde{P}$ in (4.2) and P in (4.7), we know that $\|X_P(w,\cdot)\|^{\mathcal{L}}_{r,p-1}\le c\varepsilon\|z\|_{a,p}$. So, X_P is Lipschitz continuous on $\mathcal{O}$, for all $w\in D(s,r)$. ∎

We will show the special form of P, the last proposition is then the following:

Proposition 5.4 (The special form of the perturbation). The perturbation P in Hamiltonian (4.6) belongs to $\mathcal{A}$.

Proof: Consider the Taylor-Fourier expansion of P : $P = \sum_{k,\alpha,\beta} P_{k\alpha\beta}(I)e^{\mathrm{i}\langle k,\theta\rangle} z^\alpha \bar{z}^\beta$. It follows from $\widehat{G}$, $K \in \mathcal{A}$ that $P \in \mathcal{A}$, i.e.,

$$P_{k\alpha\beta}(I) = 0,$$

whenever

$$k_1 n_1 + k_2 n_2 + \sum_{n\in\mathbb{Z}_1} (\alpha_n - \beta_n) n \neq 0,$$

or

$$k_1 + k_2 + \sum_{n\in\mathbb{Z}_1} (\alpha_n - \beta_n) \neq 0.$$

Together with Lemma 2.4, this particularly implies that P contains no terms of the form $e^{i\langle k,\theta\rangle} z_j \bar{z}_{-j}$ with $|j| > \frac{1}{2}\max\{|n_1|, |n_2|\}|k|$ and $e^{i\langle k,\theta\rangle} z_j \bar{z}_j$ with $k \neq 0$.

Thus, we give the following lemma through the analysis above:

Lemma 5.1. The perturbation P in Hamiltonian (4.6) satisfies propositions (5.1)-(5.4) and $\exists$ $\varepsilon_0(\varepsilon) > 0$ *small enough ,such that*

$$\|X_P\|_{r,q;D(s,r)\times\mathcal{O}} + \frac{\alpha}{M}\|X_P\|^{\mathcal{L}}_{r,q;D(s,r)\times\mathcal{O}} \leq \varepsilon_0.$$

VI. AN INFINITE DIMENSIONAL KAM THEOREM

Theorem 1 is a direct result of the following Theorem 2. Now consider the perturbed Hamiltonian

$$H = N + P = \langle\omega(\xi), I\rangle + \sum_{j\in\mathbb{Z}_1} \bar{\Omega}_j(\xi) z_j \bar{z}_j + P(\theta, I, z, \bar{z}, \xi). \tag{6.1}$$

Our purpose is to prove that the Hamilton system determined by Hamiltonian H admits quasi-periodic solutions provided $\|X_P\|_{r,q;D(s,r)\times\mathcal{O}} + \frac{\alpha}{M}\|X_P\|^{\mathcal{L}}_{r,q;D(s,r)\times\mathcal{O}}$ is sufficiently small.

Now we are ready to state our KAM theorem.

Theorem 2. *Assume that the Hamiltonian (6.1) satisfies the following hypotheses:*

(1) Frequencies satisfy Proposition 5.1 and Proposition 5.2;

(2) The perturbation P satisfies Proposition 5.3 and Proposition 5.4.

Then for sufficiently small $\alpha > 0$, *there exists a positive* $\varepsilon_0(\varepsilon) > 0$ *depending on* $\Omega, \alpha, d, \delta, m$ *such that*

$$\|X_P\|_{r,q;D(s,r)\times\mathcal{O}} + \frac{\alpha}{M}\|X_P\|^{\mathcal{L}}_{r,q;D(s,r)\times\mathcal{O}} \leq \varepsilon_0,$$

then the following holds true:

There exists a nonempty $\widetilde{\mathcal{O}}$ *of* $\mathcal{O}$ *such that for* $\forall\xi \in \widetilde{\mathcal{O}}$, *there exists a real analytic symplectic map* $\Psi(\cdot,\xi) : D(\frac{s}{2}, \frac{r}{2}) \to D(s, r)$ *which transforms H into the following form* $H \circ \Psi = N_* + P_*$, *where*

$$N_* = \langle\omega_*(\xi), I\rangle + \sum_{j\in\mathbb{Z}_1} \bar{\Omega}_{*j}(\xi) z_j \bar{z}_j,$$

and

$$P_* = \sum_{2|l|+|\alpha|+|\beta|\geq 3} P_{*kl\alpha\beta}(\theta) I^l z^\alpha \bar{z}^\beta.$$

Hence, for $\xi \in \widetilde{\mathcal{O}}, \Psi(\mathbb{T}^2, \xi)$ *is a real analytic invariant torus with frequency* ω_* *satisfying*

$$|\omega_*(\xi) - \omega(\xi)|_{\mathcal{O}} + \frac{\alpha}{M}|\omega_*(\xi) - \omega(\xi)|^{\mathcal{L}}_{\mathcal{O}} \leq c\varepsilon_0.$$

Moreover, each embedding is real analytic and

$$\|\Psi-\Psi_0\|^{\sup}_{r,p;D(\frac{s}{2})\times\widetilde{O}}+\frac{\alpha}{M}\|\Psi-\Psi_0\|^{L}_{r,p;D(\frac{s}{2})\times\widetilde{O}}\le\frac{c\varepsilon_0}{\alpha},$$

where $\Psi_0:\mathbb{T}^2\times O\to\mathbb{T}^2_0,\ (\theta,\xi)\mapsto(\theta,0,0,0)$ *is the trivial embedding for each* ξ, *and* c *is a positive constant. Also, we have* $\mathrm{meas}(\mathcal{O}\setminus\widetilde{\mathcal{O}})\to 0$ *as* $\alpha\to 0$.

In Secs. VII–IX, we will give the proof of Theorem 2.

VII. THE KAM STEP

In order to give a clear proof of Theorem 2, we start with stating the first KAM step in detail. For the duration of the rest of the sections, we denote "+" for the next step to simplify notation. Also, throughout the whole paper, we use letters c, C to denote suitable (possibly different) constants that do not depend on the iteration steps.

A. First KAM step

Let us consider the Hamiltonian (6.1) defined in $D(s,r)\times\mathcal{O}$. We assume that $\xi\in\mathcal{O}$ satisfies

$$\mathrm{DC}=\begin{cases}|\langle k,\omega(\xi)\rangle|\ge\frac{\alpha}{|k|^\tau}, & k\neq 0,\\ |\langle k,\omega(\xi)\rangle+\Omega_j(\xi)|\ge\frac{\alpha|j|^{1+\delta}}{|k|^\tau}, & \\ |\langle k,\omega(\xi)\rangle+\Omega_i(\xi)+\Omega_j(\xi)|\ge\frac{\alpha(|i|+|j|)(|i|^\delta+|j|^\delta)}{|k|^\tau}, & \\ |\langle k,\omega(\xi)\rangle+\Omega_i(\xi)-\Omega_j(\xi)|\ge\frac{\alpha(|i|^\delta+|j|^\delta)}{|k|^\tau}, & |k|+\big||i|-|j|\big|\neq 0.\end{cases}\tag{7.1}$$

We now let $0<s_+<s$ and define

$$r_+=\eta r,\quad \varepsilon_+=c(\alpha^{-1}(s-s_+)^{-(2\tau+3)})^{\frac{1}{3}}\varepsilon_0^{\frac{4}{3}}.\tag{7.2}$$

Now we describe how to construct a symplectic transformation $\Phi:D_+\times\mathcal{O}=D(s_+,r_+)\times\mathcal{O}\to D(s,r)\times\mathcal{O}$ such that the new Hamiltonian $H_+=N_++P_+$ with new parameters ε_+, s_+,α_+,r_+.

1. Solving the linearized equations

Expand P into the Fourier-Taylor series

$$P=\sum_{k\in\mathbb{Z}^2,l\in\mathbb{N}^2,\alpha,\beta}P_{kl\alpha\beta}(\xi)e^{\mathrm{i}\langle k,\theta\rangle}I^lz^\alpha\bar z^\beta.\tag{7.3}$$

Let R be the truncation of P given by

$$R(\theta,I,z,\bar z)=\sum_{2|l|+|\alpha+\beta|\le 2}\sum_{k\in\mathbb{Z}^2}P_{kl\alpha\beta}(\xi)e^{\mathrm{i}\langle k,\theta\rangle}I^lz^\alpha\bar z^\beta.\tag{7.4}$$

The mean value of such a Hamiltonian is defined as

$$[R]=\sum_{|l|+|\alpha|=1}R_{0l\alpha\alpha}I^lz^\alpha\bar z^\alpha,\tag{7.5}$$

and is of the same form as N.

At each step of KAM iteration, when we assume the small divisor conditions have been satisfied, we can look for a function F defined in $D(s_+,r_+)=D_+$ such that the time one-map X_F^1 of the Hamiltonian vector field X_F defines a map $D_+\to D$ and transforms H into H_+. The idea of the KAM step is to find, iteratively, an adequate function F so that the new error term has a small quadratic part. Namely, thanks to the Taylor formula, we can write

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

In view of the previous equation, we define the new normal form by $N_+ = N + \widehat{N}$, where $\widehat{N}$ satisfies the so-called homological equation (the unknown are F and $\widehat{N}$)

$$\{F, N\} + \widehat{N} = R. \tag{7.6}$$

Once the homological equation is solved, we define the new perturbation term P_+ by

$$P_+ = (P - R) \circ X_F^1 + \int_0^1 \{R(t), F\} \circ X_F^t dt, \tag{7.7}$$

where $R(t) = (1-t)\widehat{N} + tR$. The following result shows that it is possible to solve Eq. (7.6) under the condition (5.3) and the Diophantine condition (7.1).

Lemma 7.1. The homological Equation (7.6) has a solution F, $\widehat{N}$ which is unique with $[F] = 0$, $[\widehat{N}] = \widehat{N}$,$F$ is regular on $D(s, r) \times \mathcal{O}$ in the above sense, and satisfies for $0 < \sigma < s$ the estimates

$$\begin{aligned}
\|X_F\|^{\sup}_{r,p;D(s-\sigma,r)\times\mathcal{O}} &\le \frac{C}{\alpha\sigma^{2\tau+3}} \|X_R\|^{\sup}_{r,q;D(s,r)\times\mathcal{O}}, \\
\|X_F\|^{L}_{r,p;D(s-\sigma,r)\times\mathcal{O}} &\le \frac{C}{\alpha\sigma^{2\tau+3}} (\|X_R\|^{L}_{r,q;D(s,r)\times\mathcal{O}} + \frac{M}{\alpha} \|X_R\|^{\sup}_{r,q;D(s,r)\times\mathcal{O}}),
\end{aligned}$$

and

$$\begin{aligned}
\|X_{\widehat{N}}\|^{\sup}_{r,q;D(s-\sigma,r)\times\mathcal{O}} &\le C\|X_R\|^{\sup}_{r,q;D(s,r)\times\mathcal{O}}, \\
\|X_{\widehat{N}}\|^{L}_{r,q;D(s-\sigma,r)\times\mathcal{O}} &\le C(\|X_R\|^{L}_{r,q;D(s,r)\times\mathcal{O}} + \frac{M}{\alpha} \|X_R\|^{\sup}_{r,q;D(s,r)\times\mathcal{O}}).
\end{aligned}$$

Proof: Decompose $R = R^0 + R^1 + R^2$, where R^j comprises all terms in the expansion of R with $|\alpha + \beta| = j$. Decompose similarly F, N, and $\widehat{N}$, where necessarily $N^1 = 0$ and $\widehat{N}^1 = 0$ by normalization. Comparing coefficients the linearized equation decomposes into

$$\begin{aligned}
&\{F^0, N^0\} + \widehat{N}^0 = R^0, \\
&\{F^1, N\} = R^1, \\
&\{F^2, N\} + \widehat{N}^2 = R^2.
\end{aligned} \tag{7.8}$$

We will see that with the chosen normalization and the Diophantine conditions these equations determine $\widehat{N}^0$, F^0, F^1, and then $\widehat{N}^2$, F^2 uniquely.

The first equation is independent of $z, \bar{z}$ and amounts to the classical, finite-dimension partial differential equation

$$-2\pi \mathrm{i}\partial_\omega F^0 + \widehat{N}^0 = R^0, \quad \partial_\omega = \sum_{1\le i\le 2} \omega_i \partial_{\theta_i}.$$

This leads to $\widehat{N}^0 = [R^0]$ and $-2\pi\mathrm{i}\partial_\omega F^0 = R^0 - [R^0]$ with $[F^0] = 0$. Their estimates are standard and of the same form—indeed much better—than the ones for F^1, F^2, and $\widehat{N}^2$ obtained below. For later reference, we record that

$$\begin{aligned}
\|X_{F^0}\|^{\sup}_{r,p;D(s-2\sigma,r)\times\mathcal{O}} &\le \frac{C}{\alpha\sigma^{2\tau+3}} \|X_R\|^{\sup}_{r,q;D(s,r)\times\mathcal{O}}, \\
\|X_{F^0}\|^{L}_{r,p;D(s-4\sigma,r)\times\mathcal{O}} &\le \frac{C}{\alpha\sigma^{2\tau+3}} (\|X_R\|^{L}_{r,q;D(s,r)\times\mathcal{O}} + \frac{M}{\alpha} \|X_R\|^{\sup}_{r,q;D(s,r)\times\mathcal{O}}).
\end{aligned}$$

Note that X_{F^0} does not have any $z, \bar{z}-$component, so $\|X_{F^0}\|_{r,p}$ does not depend on p.

Consider the second equation in (7.8). Writing

$$R^1 = R^{10} + R^{01} = \langle \mathcal{R}^{10}, z\rangle + \langle \mathcal{R}^{01}, z\rangle$$

and similarly F^1 it decomposes into

$$\{F^{ij}, N\} = R^{ij}, \quad i + j = 1,$$

and it suffices to study each equation individually. We have $\mathcal{R}^{10} = R_z \mid_{z,\bar{z}=0}$ and thus

$$\frac{1}{r}\|\mathcal{R}^{10}\|^{\sup}_{q;D(s)} \le \|X_R\|^{\sup}_{r,q;D(s,r)},$$

where $D(s) = \{\theta : |\mathrm{Im}\theta| < s\}$. Writing $R^{10} = \langle \mathcal{R}^{10}, z\rangle = \sum_{j\in\mathbb{Z}_1} R_j(\theta, \xi)z_j$, and similarly F^{10}, the equation $\{F^{10}, N\} = R^{10}$ further decomposes into

$$2\pi(-\mathrm{i}\partial_\omega F_j + \Omega_j F_j) = R_j, \quad j \in \mathbb{Z}_1.$$

By the non-degeneracy condition (5.3) and the diophantine condition (7.1), we have uniformly on $\mathcal{O}$

$$|\Omega_j| \ge m|j|^d, \quad j \in \mathbb{Z}_1,$$

$$|\langle k, \omega(\xi)\rangle + \Omega_j(\xi)| \ge \frac{\alpha|j|^{1+\delta}}{|k|^\tau}, \quad k \ne 0,\ j \in \mathbb{Z}_1.$$

The unique solution F_j satisfies the estimate

$$|F_j|^{\sup}_{D(s-2\sigma)} \le \frac{C}{\alpha\sigma^{\tau+1}|j|^{1+\delta}}|R_j|^{\sup}_{D(s-\sigma,r)}, \quad j \in \mathbb{Z}_1.$$

Since $p - q \le \delta$, this and Lemma A.2 imply

$$\|\mathcal{F}^{10}\|^{\sup}_{p,D(s-2\sigma)} \le \frac{C}{\alpha\sigma^{\tau+1}}\|\mathcal{R}^{10}\|^{\sup}_{q,D(s)} \le \frac{C}{\alpha\sigma^{\tau+1}} r\|X_R\|^{\sup}_{r,q;D(s,r)}.$$

The same estimate holds for $\mathcal{F}^{01}$. Multiplying $\mathcal{F}^{10}$ with z and $\mathcal{F}^{01}$ with $\bar{z}$ and using $p > \frac{3}{2}$ this gives

$$\frac{1}{r^2}|F^1|^{\sup}_{D(s-2\sigma,r)} \le \frac{C}{\alpha\sigma^{\tau+1}}\|X_R\|^{\sup}_{r,q;D(s,r)},$$

finally with Cauchy's estimate

$$\|X_{F^1}\|^{\sup}_{r,p;D(s-3\sigma,r)} \le \frac{C}{\alpha\sigma^{\tau+1}}\|X_R\|^{\sup}_{r,q;D(s,r)}.$$

To obtain Lipschitz estimates, we study first the differences $\Delta F_j = F_j(\xi) - F_j(\zeta)$ for $\xi, \zeta \in \mathcal{O}$. We obtain

$$2\pi(-\mathrm{i}\partial_\omega \Delta F_j + \Omega_j \Delta F_j) = \Delta R_j + 2\pi(\mathrm{i}\partial_{\Delta\omega}F_j - F_j\Delta\Omega_j), \quad j \in \mathbb{Z}_1.$$

The right hand side is known, so ΔF_j uniquely solves the same kind of equation as F_j. So we obtain

$$\begin{aligned}|\Delta F_j|^{\sup}_{D(s-3\sigma)} &\le \frac{C}{\alpha\sigma^{\tau+1}|j|^{1+\delta}}\left(|\Delta R_j|^{\sup}_{D(s-\sigma)} + \frac{1}{\sigma}|F_j|^{\sup}_{D(s-2\sigma)}(|\Delta\omega| + |\Delta\Omega_j|^{\sup}_{D(s)})\right),\\ &\le \frac{C}{\alpha\sigma^{\tau+1}|j|^{1+\delta}}|\Delta R_j|^{\sup}_{D(s-\sigma)} + \frac{C}{\alpha^2\sigma^{2\tau+3}|j|^{2(1+\delta)}}|R_j|^{\sup}_{D(s-\sigma)}(|\Delta\omega| + |\Delta\Omega_j|^{\sup}_{D(s)}).\end{aligned}$$

Then

$$\|\Delta\mathcal{F}^{10}\|^{\sup}_{p;D(s-3\sigma)} \le \frac{C}{\alpha\sigma^{\tau+1}}\|\Delta\mathcal{R}^{10}\|^{\sup}_{q;D(s)} + \frac{C}{\alpha^2\sigma^{2\tau+3}}\|\mathcal{R}^{10}\|^{\sup}_{q;D(s)}(|\Delta\omega| + |\Delta\Omega|^{\sup}_{-\delta;D(s)}).$$

Dividing by $|\xi - \zeta| \ne 0$ and taking the supremum over $\mathcal{O}$,

$$\|\mathcal{F}^{10}\|^{\mathcal{L}}_{p;D(s-3\sigma)} \le \frac{C}{\alpha\sigma^{2\tau+3}}\left(\|\mathcal{R}^{10}\|^{\mathcal{L}}_{q;D(s)} + \frac{M}{\alpha}\|\mathcal{R}^{10}\|^{\sup}_{q;D(s)}\right).$$

The same estimate applies to $\mathcal{F}^{01}$. So for the vector field of F^1, we finally get

$$\|X_{F^1}\|^{\mathcal{L}}_{r,p;D(s-4\sigma)} \leq \frac{C}{\alpha\sigma^{2\tau+3}}(\|X_R\|^{\mathcal{L}}_{r,q;D(s,r)} + \frac{M}{\alpha}\|X_R\|^{\sup}_{r,q;D(s,r)}).$$

This concludes the discussion of F^1.

Now, consider the third equation in (7.8). Write $R^2 = R^{20} + R^{11} + R^{02}$ and similarly F^2 and N^2. This equation decomposes into

$$\{F^{ij}, N\} + \widehat{N}^{ij} = R^{ij}, \tag{7.9}$$

while $\widehat{N}^{ij} = 0$ for $i \neq j$.

Consider the equation for F^{11}, which is slightly more complicated than the ones for F^{20} and F^{02}. Writing $R^{11} = \langle \mathcal{R}^{11}z, \bar{z}\rangle$, we have $\mathcal{R}^{11} = R_{z\bar{z}}\mid_{z,\bar{z}=0}$. Thus, $\mathcal{R}^{11}$ is the Jacobian of R_z with respect to $\bar{z}$ at $\bar{z} = 0$. By Cauchy's inequality, we have

$$\|\mathcal{R}^{11}\|^{\sup}_{q,p;D(s)} \leq \frac{1}{r}\|R_z\|^{\sup}_{q;D(s,r)} \leq \|X_R\|^{\sup}_{r,q;D(s,r)},$$

where $\|\cdot\|_{q,p}$ denotes the operator norm included by $\|\cdot\|_p$ and $\|\cdot\|_q$ in the source and target spaces, respectively.

Note that P contains no terms of the form $z_j\bar{z}_{-j}$ and $e^{i\langle k,\theta\rangle}z_j\bar{z}_j$ with $k \neq 0$, write more explicitly

$$R^{11} = \sum_{\substack{i\neq j\\ i,j\in\mathbb{Z}_1}} R_{ij}(\theta,\xi)z_i\bar{z}_j + \sum_{j\in\mathbb{Z}_1} R_{jj}(\xi)z_j\bar{z}_j + \sum_{\substack{j\in\mathbb{Z}_1\\ |j|\leq\frac{1}{2}\max\{|n_1|,|n_2|\}|k|}} R_{-jj}(\theta,\xi)z_j\bar{z}_{-j}$$

and similarly F^{11}. The Eq. (7.9) decomposes into

$$2\pi(-\mathrm{i}\partial_\omega F_{ij} + (\Omega_i - \Omega_j)F_{ij}) = R_{ij}, \quad i \neq j,$$

and $F_{jj} = 0$ for $j \neq 0$, $F_{-jj} = 0$ for $|j| > \frac{1}{2}\max\{|n_1|, |n_2|\}|k|$.

Again, by the conditions (5.5) and (7.1), we have

$$|\Omega_i - \Omega_j| \geq m(|i|^{d-1} + |j|^{d-1}),$$

$$|\langle k, \omega(\xi)\rangle + \Omega_i(\xi) - \Omega_j(\xi)| \geq \frac{\alpha(|i|^\delta + |j|^\delta)}{|k|^\tau},$$

we obtain

$$(|i|^\delta + |j|^\delta)|F_{ij}|^{\sup}_{D(s-2\sigma)} \leq \frac{C}{\alpha\sigma^{\tau+1}}|R_{ij}|^{\sup}_{D(s-\sigma,r)}, \quad i \neq j.$$

With Lemma A.3, this yield

$$\|\mathcal{F}^{11}\|^{\sup}_{p,p;D(s-2\sigma)}, \|\mathcal{F}^{11}\|^{\sup}_{q,q;D(s-2\sigma)} \leq \frac{C}{\alpha\sigma^{\tau+1}}\|\mathcal{R}^{11}\|^{\sup}_{q,p;D(s)} \leq \frac{C}{\alpha\sigma^{\tau+1}}\|X_R\|^{\sup}_{r,q;D(s,r)}.$$

The same, and even better estimates hold for $\mathcal{F}^{20}$ and $\mathcal{F}^{02}$. Multiplying with $z, \bar{z}$ we then get

$$\frac{1}{r^2}|F^2|^{\sup}_{D(s-2\sigma,r)} \leq \frac{C}{\alpha\sigma^{\tau+1}}\|X_R\|^{\sup}_{r,q;D(s,r)},$$

finally with Cauchy's estimate

$$\|X_{F^2}\|^{\sup}_{r,p;D(s-3\sigma,r)} \leq \frac{C}{\alpha\sigma^{\tau+1}}\|X_R\|^{\sup}_{r,q;D(s,r)}.$$

The estimate for the Lipschitz semi-norm of X_{F^2} is obtained by the same arguments as the one for X_{F^1}, and the result is analogous. We therefore omit it.

The estimates of $X_{\widehat{N}}$ follow from the observation that

$$\widehat{N} = \sum_{|l|=1} P_{0l00}I^l + \sum_{j\in\mathbb{Z}_1} P_{00jj}(\xi)z_j\bar{z}_j.$$

The final estimates of the lemma are obtained by replacing σ by $\frac{\sigma}{4}$ throughout the proof. ∎

Remark 7.1. The crucial, somewhat hidden feature of the first two of these estimates is the "p" on their left hand sides and the "q" on their right hand sides. That means, the solution X_F is bounded in the stronger norm $\|\cdot\|_{r,p}$ rather than $\|\cdot\|_{r,q}$.

For later reference, the estimates of Lemma 7.1 may be condensed as follows. For $\lambda \geq 0$, define

$$\|X\|_r^\lambda = \|X\|_r^{\sup} + \lambda \|X\|_r^{\mathcal{L}}.$$

The symbol "λ" in $\|X\|_r^\lambda$ will always be used in this role and never has the meaning of exponentiation.

Lemma 7.2. The estimates of Lemma 7.1 imply that

$$\|X_F\|^\lambda_{r,p;D(s-\sigma,r)\times\mathcal{O}} \leq \frac{C}{\alpha\sigma^{2\tau+3}}\|X_R\|^\lambda_{r,q;D(s,r)\times\mathcal{O}},$$

$$\|X_{\widehat{N}}\|^\lambda_{r,q;D(s-\sigma,r)\times\mathcal{O}} \leq C\|X_R\|^\lambda_{r,q;D(s,r)\times\mathcal{O}},$$

for $0 < \sigma < s$ and $0 < \lambda \leq \frac{\alpha}{M}$ with another C of the same form as in Lemma 7.1.

The preceding lemma also gives us an estimate of $\|DX_F\|^\lambda_{r,p,p;D(s-2\sigma,r)\times\mathcal{O}}$ with the help of Cauchy's estimate.

Lemma 7.3. Under the assumptions of Lemma 7.1,

$$\|DX_F\|^\lambda_{r,p,p;D(s-\sigma,r)\times\mathcal{O}}, \|DX_F\|^\lambda_{r,q,q;D(s-\sigma,r)\times\mathcal{O}} \leq \frac{C}{\alpha\sigma^{2\tau+3}}\|X_R\|^\lambda_{r,q;D(s,r)\times\mathcal{O}}.$$

Proof: The proof can be found on page 160 in Ref. 7. ■

2. Approximation estimates

We recall some approximation results in Ref. 14, which show that the second order approximation of P can be controlled by P, and that $P - R$ is small when we contract the domain (this contraction is governed by the new parameter η).

Lemma 7.4. Let P satisfies Proposition 5.3 and consider its Taylor approximation R of the form (7.4). Then there exists $C > 0$ so that for all $\eta > 0$,

$$\|X_R\|^\lambda_{r,q;D(s,r)} \leq \|X_P\|^\lambda_{r,q;D(s,r)},$$

$$\|X_{(P-R)}\|^\lambda_{\eta r,q;D(s,4\eta r)} \leq C\eta\|X_P\|^\lambda_{r,q;D(s,r)}. \tag{7.10}$$

For the proof, see Ref. 14. ■

3. Estimation on the coordinate transformation

In the section, we give some estimates for X_F^t. The formulas (7.11) and (7.12) will be used to prove our coordinate transformation is well-defined. Inequality (7.13) and (7.14) will be used to check the convergence of the iteration.

Lemma 7.5. If $\varepsilon_0 \leq \frac{\alpha\sigma^{2\tau+4}\eta^2}{C}$, we then have

$$X_F^t : D(s-2\sigma, \frac{r}{2}) \to D(s-\sigma, r), \quad -1 \leq t \leq 1, \tag{7.11}$$

Similarly,

$$X_F^t : D(s-3\sigma, \frac{r}{4}) \to D(s-2\sigma, \frac{r}{2}), \quad -1 \leq t \leq 1, \tag{7.12}$$

Moreover,

$$\|X_F^t - Id\|^{\lambda}_{r,p;D(s-2\sigma,\frac{r}{2})} < C\|X_F\|^{\lambda}_{r,p;D(s-\sigma,r)}, \tag{7.13}$$

$$\|DX_F^t - Id\|^{\lambda}_{r,q,q;D(s-3\sigma,\frac{r}{4})} < C\|DX_F\|^{\lambda}_{r,q,q;D(s-\sigma,r)}, \tag{7.14}$$

for $0 \le \lambda \le \frac{\alpha}{M}$. *The latter estimate also holds in the* $\|\cdot\|_{r,p,p}$ *−norm.*

We can use Lemma 7.2 and Lemma A.4 to prove this lemma. ■

4. Estimation for new perturbation and the new normal form

The map $\Phi = X_F^1$ defined above transforms H into $H_+ = H\circ\Phi = N_+ + P_+$ on $D(s-\sigma, \frac{r}{2})$, where $N_+ = N + \widehat{N}$ and

$$P_+ = \int_0^1 \{R(t), F\}\circ X_F^t dt + (P - R)\circ X_F^1,$$

where $R(t) = (1-t)\widehat{N} + tR$. Hence,

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

From the paper,[14] we have known the following result:

$$\|(X_F^t) * Y\|_{\eta r,q;D(s-4\sigma,\eta r)} \le C\|Y\|_{\eta r,q;D(s-2\sigma,4\eta r)}, \quad 0 \le t \le 1. \tag{7.15}$$

We already have estimated $\|X_P - X_R\|^{\lambda}_{\eta r,q}$ in (7.10), so it remains to consider the commutator $\|X_{\{R(t),F\}}\|_{r,q}$.

First, we have

$$\begin{aligned}\|X_{R(t)}\|^{\lambda}_{r,q;D(s-\sigma,r)} &\le \|X_{\widehat{N}}\|^{\lambda}_{r,q;D(s-\sigma,r)} + \|X_R\|^{\lambda}_{r,q;D(s-\sigma,r)}\\ &\le C\|X_P\|^{\lambda}_{r,q;D(s,r)}.\end{aligned}$$

Moreover, we have the pointwise estimate

$$\begin{aligned}\|X_{\{R(t),F\}}\|_{r,q} &\le \|DX_{R(t)}\cdot X_F\|_{r,q} + \|X_{R(t)}\cdot DX_F\|_{r,q},\\ &\le \|DX_{R(t)}\|_{r,q,p}\|X_F\|_{r,p} + \|DX_F\|_{r,q,q}\|X_{R(t)}\|_{r,q}.\end{aligned}$$

By the product rule for Lipschitz-norms and Cauchy's estimate, we thus obtain

$$\begin{aligned}\|X_{\{R(t),F\}}\|^{\lambda}_{r,q;D(s-2\sigma,\frac{r}{2})} &\le \|DX_{R(t)}\|^{\lambda}_{r,q,p;D(s-2\sigma,\frac{r}{2})}\|X_F\|^{\lambda}_{r,p;D(s-2\sigma,\frac{r}{2})}\\ &\quad+\|DX_F\|^{\lambda}_{r,q,q;D(s-2\sigma,\frac{r}{2})}\|X_{R(t)}\|^{\lambda}_{r,q;D(s-2\sigma,\frac{r}{2})},\\ &\le \frac{C}{\alpha\sigma^{2\tau+3}}(\|X_P\|^{\lambda}_{r,q;D(s,r)})^2,\end{aligned}$$

for $0 \le \lambda \le \frac{\alpha}{M}$. Hence, also

$$\begin{aligned}\|X_{\{R(t),F\}}\|^{\lambda}_{\eta r,q;D(s-2\sigma,\frac{r}{2})} &\le \frac{1}{\eta^2}\|X_{\{R(t),F\}}\|^{\lambda}_{r,q;D(s-2\sigma,\frac{r}{2})},\\ &\le \frac{C}{\alpha\sigma^{2\tau+3}\eta^2}(\|X_P\|^{\lambda}_{r,q;D(s,r)})^2.\end{aligned}$$

Together with the estimate on $X_{(P-R)}$ in (7.10) and with that in (7.15), we finally arrive at the estimate

$$\|X_{P_+}\|^{\lambda}_{\eta r,q;D(s-2\sigma,\frac{r}{2})} \le C\eta\|X_P\|^{\lambda}_{r,q;D(s,r)} + \frac{C}{\alpha\sigma^{2\tau+3}\eta^2}(\|X_P\|^{\lambda}_{r,q;D(s,r)})^2,$$

for $0 \le \lambda \le \frac{\alpha}{M}$. This is the bound for the new perturbation.

Now turn to the new frequencies $\omega_+(\xi) = \omega(\xi) + \widehat{\omega}(\xi)$ and $\bar{\Omega}_+(\xi) = \bar{\Omega}(\xi) + \overline{\widehat{\Omega}}(\xi)$. For $\widehat{N}$, we have the estimate

$$\|X_{\widehat{N}}\|^{\lambda}_{r,q;D(s-\sigma,r)} \leq C\|X_R\|^{\lambda}_{r,q;D(s,r)},$$

for $0 \leq \lambda \leq \frac{\alpha}{M}$. The weighted norm implies that we have $|\widehat{\omega}(\xi)| \leq \|X_{\widehat{N}}\|^{\sup}_{r,q}$ and $\|\overline{\widehat{\Omega}}z\|_q \leq r\|X_{\widehat{N}}\|^{\sup}_{r,q}$ on $D(s, r)$, and consequently $|\overline{\widehat{\Omega}}|_{q-p} \leq \|X_{\widehat{N}}\|^{\sup}_{r,q}$. The same holds for the Lipschitz semi-norms. Since $p - q \leq \delta$, we obtain

$$|\widehat{\omega}|^{\lambda}_{\mathcal{O}} + |\widehat{\Omega}|^{\lambda}_{-\delta;D(s-\sigma)} \leq C\|X_P\|^{\lambda}_{r,q;D(s,r)},$$

where $\Omega = (\Omega_j)_{j\in\mathbb{Z}_1}$ and $\widehat{\Omega}_j = \sigma_j\overline{\widehat{\Omega}}_j$.

In order to control the assumptions of the KAM step for the iteration, we notice that the last estimate also implies

$$\begin{aligned}|\langle l, \widehat{\Omega}\rangle|_{D(s-\sigma)} &\leq |l|_{\delta}|\widehat{\Omega}|_{-\delta;D(s-\sigma)},\\ &\leq \langle l\rangle_{d-1}\|X_P\|^{\lambda}_{r,q;D(s,r)},\\ &\leq \varepsilon_0\langle l\rangle_{d-1}.\end{aligned}$$

Then $|\langle l, \Omega_+(\xi)\rangle|_{D(s-\sigma)} \geq m_+\langle l\rangle_{d-1}$ for $0 < |l| \leq 2$ provided that ε_0 is sufficiently small.

Lemma 7.6. $P_+ \in A$ with respect to n_1, n_2.

Proof: Note that

$$\begin{aligned}P_+ = P - R + \{P, F\} + \frac{1}{2!}\{\{N, F\}, F\} + \frac{1}{2!}\{\{P, F\}, F\}\\ + \cdots + \frac{1}{n!}\{\cdots\{N, \underbrace{F\}\cdots, F}_{n}\} + \frac{1}{n!}\{\cdots\{P, \underbrace{F\}\cdots, F}_{n}\} + \cdots.\end{aligned}$$

Since $P \in A$ with respect to n_1, n_2, then F, so do $P - R$, $\{N, F\}$ and $\{P, F\}$. The lemma follows from Lemma 2.2 and Lemma 2.3. ■

VIII. ITERATIVE LEMMA AND CONVERGENCE

To iterate the KAM step infinitely, we now choose sequences for all its parameters. Set $\alpha_0 = \alpha$, $m_0 = m$, $M_0 = M$, $r_0 = r$, $s_0 = s$. Moreover, for given ε_0 and all $\nu \geq 0$, we should define the sequences as follows:

$$\begin{aligned}&\sigma_0 = \frac{s}{8}, \quad \sigma_{\nu+1} = \frac{\sigma_\nu}{2}, \quad s_{\nu+1} = s_\nu - 2\sigma_\nu,\\ &\eta_\nu^3 = \frac{\varepsilon_\nu}{\alpha_\nu\sigma_\nu^{2\tau+3}}, \quad r_{\nu+1} = \eta_\nu r_\nu, \quad D_\nu = D(s_\nu, r_\nu),\\ &\alpha_0 = \varepsilon_0^{\frac{1}{3}}, \quad \alpha_\nu = \varepsilon_\nu^{\frac{1}{3}},\\ &M_\nu = M_0(2 - 2^{-\nu}), \quad \lambda_\nu = \frac{\alpha_\nu}{M_\nu},\\ &m_\nu = \frac{m_0}{2}(1 + 2^{-\nu}), \quad \varepsilon_{\nu+1} = C(\alpha_\nu\sigma_\nu^{2\tau+3})^{-\frac{1}{3}}\varepsilon_\nu^{\frac{4}{3}},\\ &\mathcal{O}_\nu = \{\xi \in \mathcal{O}_{\nu-1} : |\langle k, \omega_\nu(\xi)\rangle| \geq \frac{\alpha_\nu}{|k|^\tau}, \quad k \neq 0; \quad |\langle k, \omega_\nu(\xi)\rangle + \Omega_j^\nu(\xi)| \geq \frac{\alpha_\nu|j|^{1+\delta}}{|k|^\tau};\\ &\qquad |\langle k, \omega_\nu(\xi)\rangle + \Omega_i^\nu(\xi) + \Omega_j^\nu(\xi)| \geq \frac{\alpha_\nu(|i| + |j|)(|i|^\delta + |j|^\delta)}{|k|^\tau};\\ &\qquad |\langle k, \omega_\nu(\xi)\rangle + \Omega_i^\nu(\xi) - \Omega_j^\nu(\xi)| \geq \frac{\alpha_\nu(|i|^\delta + |j|^\delta)}{|k|^\tau}, \quad |k| + ||i| - |j|| \neq 0\}.\end{aligned} \tag{8.1}$$

A. Iterative lemma

The proceeding analysis can be summarized as follows:

Lemma 8.1. Let

$$\varepsilon_0 \le \frac{\alpha_0 \sigma_0^{2\tau+6}}{2C}, \quad \alpha_0 \le \frac{m_0}{2}.$$

Suppose, $H_\nu = N_\nu + P_\nu$ is given on $D(s_\nu, r_\nu) \times \mathcal{O}_\nu$ which is real analytic in $(\theta, I, z, \bar{z}) \in D_\nu$ and Whitney smooth in $\xi \in \mathcal{O}_\nu$, where

$$N_\nu = e_\nu + \langle \omega_\nu, I\rangle + \sum_{j\in\mathbb{Z}_1} \bar{\Omega}_j^\nu(\xi) z_j \bar{z}_j.$$

Its coefficients satisfy

$$|\omega_\nu|^L_{\mathcal{O}_\nu} + |\Omega^\nu|^L_{-\delta;\mathcal{O}_\nu} \le M_\nu,$$

$$|\omega_{\nu+1} - \omega_\nu|^{\lambda_\nu}_{\mathcal{O}_\nu} < \varepsilon_\nu,$$

$$|\Omega^{\nu+1} - \Omega^\nu|^{\lambda_\nu}_{-\delta;D_{\nu+1}\times\mathcal{O}_\nu} \le \varepsilon_\nu,$$

and

$$|\langle l, \Omega^\nu(\xi)\rangle| \ge m_\nu \langle l\rangle_{d-1}, \quad \forall 0 < |l| \le 2. \tag{8.2}$$

$P_\nu \in \mathcal{A}$ with respect to n_1, n_2, and

$$\|X_{P_\nu}\|^\lambda_{r_\nu, q; D(s_\nu, r_\nu)\times\mathcal{O}_\nu} \le \varepsilon_\nu.$$

Then there exists a family of symplectic coordinate transformation

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

and a closed subset

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu \setminus \bigcup_{k,l} \mathcal{R}^{\nu+1}_{k,l}(\alpha_{\nu+1}),$$

where

$$\begin{aligned}
\mathcal{R}^{\nu+1}_{k,l}(\alpha_{\nu+1}) = \{\xi \in \mathcal{O}_\nu : {} & |\langle k, \omega_{\nu+1}(\xi)\rangle| < \frac{\alpha_{\nu+1}}{|k|^\tau}, \quad k \ne 0;\\
& |\langle k, \omega_{\nu+1}(\xi)\rangle + \Omega^{\nu+1}_j(\xi)| < \frac{\alpha_{\nu+1}|j|^{1+\delta}}{|k|^\tau};\\
& |\langle k, \omega_{\nu+1}(\xi)\rangle + \Omega^{\nu+1}_i(\xi) + \Omega^{\nu+1}_j(\xi)| < \frac{\alpha_{\nu+1}(|i| + |i|)(|i|^\delta + |j|^\delta)}{|k|^\tau};\\
& |\langle k, \omega_{\nu+1}(\xi)\rangle + \Omega^{\nu+1}_i(\xi) - \Omega^{\nu+1}_j(\xi)| < \frac{\alpha_{\nu+1}(|i|^\delta + |j|^\delta)}{|k|^\tau},\\
& |k| + ||i| - |j|| \ne 0\},
\end{aligned} \tag{8.3}$$

such that for $H_{\nu+1} = H_\nu \circ \Phi_{\nu+1} = N_{\nu+1} + P_{\nu+1}$ the same assumptions as above are satisfied with "$\nu + 1$" in place of "ν."

B. Convergence

Suppose that the assumptions of Theorem 2 are satisfied. To apply the iterative lemma with $\nu > 0$, set $s_0 = s$, $r_0 = r$, $H_0 = H$, $\alpha_0 = \alpha$. Assume that the small divisor conditions are satisfied by setting $\mathcal{O}_0 = \mathcal{O}$. Thus, the iterative lemma can be applied inductively, then for $\nu \geq 1$, we obtain the following sequence:

$$\Psi^\nu = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_\nu : D_\nu \times \mathcal{O}_{\nu-1} \to D_0,$$

such that

$$H_\nu = H \circ \Psi^\nu = \langle \omega_\nu(\xi), I \rangle + \sum_{j \in \mathbb{Z}_1} \bar{\Omega}_j^\nu(\xi) z_j \bar{z}_j + P_\nu(\theta, I, z, \bar{z}, \xi).$$

Let $\widetilde{\mathcal{O}} = \bigcap_{\nu=0}^{\infty} \mathcal{O}_\nu$. As in Refs. 13 and 14, thanks to Lemma 7.5, it concludes that $N_\nu, \omega_\nu, \bar{\Omega}^\nu, \Psi^\nu$ and $D\Psi^\nu$ converge uniformly on $D(\frac{s}{2}, 0) \times \widetilde{\mathcal{O}}$ with

$$N_\infty = e_\infty + \langle \omega_\infty, I \rangle + \sum_{j \in \mathbb{Z}_1} \bar{\Omega}_j^\infty(\xi) z_j \bar{z}_j.$$

Since

$$\varepsilon_{\nu+1} = C(\alpha_\nu \sigma_\nu^{2\tau+3})^{-\frac{1}{3}} \varepsilon_\nu^{\frac{4}{3}},$$

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

Let X_H^t be the flow of X_H. Since $H \circ \Psi^\nu = H_\nu$, we have

$$X_H^t \circ \Psi^\nu = \Psi^\nu \circ X_{H_\nu}^t. \tag{8.4}$$

The uniform convergence of $\Psi^\nu, D\Psi^\nu, X_{H_\nu}$ implies that the limits can be taken on the both sides of (8.4). Hence, on $D(\frac{s}{2}, 0) \times \widetilde{\mathcal{O}}$, we get

$$X_H^t \circ \Psi^\infty = \Psi^\infty \circ X_{H_\infty}^t, \tag{8.5}$$

and

$$\Psi^\infty : D(\frac{s}{2}, 0) \times \widetilde{\mathcal{O}} \to D(s, r) \times \mathcal{O},$$

it follows from (8.5) that we get an invariant finite dimensional tori $\Psi^\infty(\mathbb{T}^2 \times \{\xi\})$ for the original perturbed Hamiltonian system at $\xi \in \widetilde{\mathcal{O}}$. We remark that the frequencies $\omega_*(\xi) = \omega_\infty(\xi)$ associated with $\Psi^\infty(\mathbb{T}^2 \times \{\xi\})$ are slightly deformed from the unperturbed ones $\omega(\xi)$. The normal behaviors of the invariant tori $\Psi^\infty(\mathbb{T}^2 \times \{\xi\})$ are governed by their respective normal frequencies $\bar{\Omega}_n^\infty$.

IX. MEASURE ESTIMATES

For convenience, we have set $\mathcal{O}_0 = \mathcal{O}$. Then at each step, we have to exclude the following resonant sets:

$$R^{\nu+1} = \bigcup_{|k|>0, l} R_{k,l}^{\nu+1}(\alpha_{\nu+1}), \tag{9.1}$$

where $\mathcal{R}_{k,l}^{\nu+1}(\alpha_{\nu+1})$ has been described in (8.3), then

$$\mathcal{O} \setminus \widetilde{\mathcal{O}} = \bigcup_{\nu \geq 0} \mathcal{R}^{\nu+1}.$$

Note that

$$\mathcal{R}_{k,l}^{\nu+1}(\alpha_{\nu+1}) \subset \widetilde{\mathcal{R}}_{k,l}^{\nu+1}(\alpha_{\nu+1}),$$

where

$$\widetilde{\mathcal{R}}_{k,l}^{\nu+1} = \{\xi \in \mathcal{O}_\nu : |\langle k, \omega_{\nu+1} \rangle + \langle l, \Omega^{\nu+1} \rangle| < \frac{\alpha_{\nu+1} \langle l \rangle_\delta}{|k|^\tau}\}.$$

Now we will prove that the measure of the set $\widetilde{\mathcal{R}}_{k,l}^{\nu+1}$ is small, so does $\mathcal{R}_{k,l}^{\nu+1}$.

Lemma 9.1. If $\widetilde{\mathcal{R}}^{\nu}_{k,l}(\alpha_\nu) \neq \emptyset$, then

$$\langle l \rangle_{d-1} \leq c|k|,$$

where $c = 4(1 + |\omega|^{\sup}_{\mathcal{O}})/m$ is independent of ν.

Proof: If there exists $\xi \in \widetilde{\mathcal{R}}^{\nu}_{k,l}$, then (8.2) implies that for $k \neq 0$,

$$\begin{aligned} |\langle k, \omega_\nu(\xi)\rangle| &\geq |\langle l, \Omega^\nu(\xi)\rangle| - \alpha_\nu \frac{\langle l \rangle_\delta}{|k|^\tau}, \\ &\geq m_\nu \langle l \rangle_{d-1} - \alpha_\nu \langle l \rangle_\delta, \\ &\geq \frac{m}{4} \langle l \rangle_{d-1}, \end{aligned}$$

since $\langle l \rangle_\delta \leq \langle l \rangle_{d-1}$ for $\delta \leq d - 1$ and $\alpha_\nu \leq \frac{m_\nu}{2}$, $m_\nu \geq \frac{m}{2}$ by construction. Hence,

$$\frac{m}{4} \langle l \rangle_{d-1} \leq |k||\omega_\nu(\xi)| \leq |k|(1 + |\omega|^{\sup}_{\mathcal{O}}).$$

■

Lemma 9.2. For fixed $\nu + 1$, k, l,

$$\mathrm{meas}(\widetilde{\mathcal{R}}^{\nu+1}_{kl}(\alpha_{\nu+1})) < C\rho_\nu \frac{\alpha_{\nu+1}}{|k|^{\tau+1}},$$

where ρ_ν is the diameter of $\mathcal{O}_\nu$.

Proof: Denote

$$f(\xi) = \langle k, \omega_{\nu+1}(\xi)\rangle + \langle l, \Omega^{\nu+1}(\xi)\rangle,$$

let vector v satisfy $\langle k, v\rangle = |k|$. It follows that

$$\frac{df(\xi + tv)}{dt} \geq C|k| > 0,$$

where C is some positive constant. Then the proof of this lemma is evident by using Lemma A.5, so we omit it here. ■

Lemma 9.3. For fixed $\nu + 1 \geq 0$,

$$\mathrm{meas}(\bigcup_{k,l} \widetilde{\mathcal{R}}^{\nu+1}_{k,l}(\alpha_{\nu+1})) \leq C\rho_\nu \alpha_{\nu+1},$$

where C is a constant.

Proof: For a fixed k, it suffices to consider l with $\langle l \rangle_{d-1} \leq c|k|$ according to Lemma 9.1. Taking into account that $|l|_{d-1} \leq 2\langle l \rangle_{d-1}$, we get

$$\mathrm{card}\{l : |l| \leq 2, \langle l \rangle_{d-1} \leq c|k|\} \leq c|k|^s, \quad s = \frac{2}{d-1}.$$

Hence, by Lemma 9.2,

$$\mathrm{meas}(\bigcup_{l} \widetilde{\mathcal{R}}^{\nu+1}_{kl}(\alpha_{\nu+1})) \leq C\rho_\nu \frac{\alpha_{\nu+1}}{|k|^{\tau+1-s}}.$$

If we choose $\tau \geq s + 2$, then

$$\mathrm{meas}(\bigcup_{k,l} \widetilde{\mathcal{R}}^{\nu+1}_{kl}(\alpha_{\nu+1})) \leq C\rho_\nu \alpha_{\nu+1}.$$

So, Lemma 9.3 follows. ■

By Lemma 9.3, we can obtain the following result about the finite dimension Lebesgue measure of $(\mathcal{O}_\nu \backslash \mathcal{O}_{\nu+1})$, i.e.,

$$\text{meas}(\mathcal{O}_\nu \setminus \mathcal{O}_{\nu+1}) = \text{meas}(\bigcup_{k,l} \mathcal{R}_{kl}^{\nu+1}(\alpha_{\nu+1})) \leq \text{meas}(\bigcup_{K,l} \widetilde{\mathcal{R}}_{k,l}^{\nu+1}(\alpha_{\nu+1})) = \mathcal{O}(\alpha_{\nu+1}) \to 0,$$

as $\nu \to \infty$. It follows that the total measure of all excluded parameters can be as small as we wish, and we will finally get a Cantor-like parameter set $\widetilde{\mathcal{O}} = \bigcap_{\nu=0}^{\infty} \mathcal{O}_\nu$.

ACKNOWLEDGMENTS

The authors are grateful to Jiangong You for his valuable discussions. This work is in part supported by the National Science Foundation of China (NSFC) (Grant No. 11271180). This work is also in part supported by the Scientific Research and Innovation Project for college postgraduates in Jiangsu, China (CXZZ12_0029) and a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

APPENDIX: SOME TECHNICALITIES

Lemma A.1: (Generalized Cauchy inequalities)

$$\|F_\theta\|_{D(s-\sigma,r)} \leq \frac{c}{\sigma}\|F\|_{D(s,r)},$$

$$\|F_I\|_{D(s,\frac{1}{2}r)} \leq \frac{c}{r^2}\|F\|_{D(s,r)},$$

and

$$\|F_z\|_{D(s,\frac{1}{2}r)} \leq \frac{c}{r}\|F\|_{D(s,r)},$$

$$\|F_{\bar{z}}\|_{D(s,\frac{1}{2}r)} \leq \frac{c}{r}\|F\|_{D(s,r)}.$$

Lemma A.2. Let u_j, $j \geq 1$, be complex functions on $\mathbb{T}^n$ that are real analytic on $D(s) = \{\theta : |\text{Im}\theta| < s\}$. Then

$$(\sum_{j\geq 1} \sup_{\theta\in D(s-\sigma)} |u_j(\theta)|^2)^{\frac{1}{2}} \leq \frac{4^n}{\sigma^n} \sup_{\theta\in D(s)} (\sum_{j\geq 1} |u_j(\theta)|^2)^{\frac{1}{2}},$$

for $0 < \sigma < s \leq 1$.

Proof: The proof can be found on page 262–263 in Ref. 7. ■

Lemma A.3. Let $A = (A_{ij})_{i,j\neq 0}$ be a bounded operator on l^2, which depends on $\theta \in \mathbb{T}^n$ such that all coefficients are analytic on $D(s) = \{\theta : |\text{Im}\theta| < s\}$. Suppose $B = (B_{ij})_{i,j\neq 0}$ is another operator on l^2 depending on θ whose coefficients satisfy

$$\sup_{D(s)} |B_{ij}(\theta)| \leq \frac{1}{\big||i| - |j|\big|} \sup_{D(s)} |A_{ij}(\theta)|, \quad i \neq |j|,$$

and $B_{jj} = 0$, $B_{-jj} = 0$ for $j \neq 0$. Then B is a bounded operator on l^2 for every $\theta \in D(s)$, and

$$\sup_{\theta\in D(s-\sigma)} \|B(\theta)\| \leq \frac{4^n}{\sigma^n} \sup_{\theta\in D(s)} \|A(\theta)\|,$$

for $0 < \sigma < s \leq 1$.

Proof: The proof can be found on page 263–264 in Ref. 7. ■

Let V be an open domain in a real Banach space E with norm$\|\cdot\|$, B a subset of another real Banach space, and $X : V \times B \to E$ a parameter dependent vector field on V, which is C^1 on V and Lipschitz on B. Let ϕ^t be its flow. Suppose there is a subdomain $U \subset V$ such that $\phi^t : U \times B \to V$ for $-1 \le t \le 1$.

Lemma A.4. Under the preceding assumptions,

$$\|\phi^t - id\|_U \le \|X\|_V,$$

$$\|\phi^t - id\|_U^{\mathcal{L}} \le \exp(\|DX_V\|)\|X\|_V^{\mathcal{L}}.$$

For the proof, see Ref. 14. ■

Lemma A.5. Suppose that $g(u)$ is a C^N function on the closure $\bar{I}$, where $I \subset R^1$ is an interval. Let $I_h = \{u : |g(u)| \le h\}$, $h > 0$. If for some constant $d > 0$, $|g^N(u)| \ge d$ for $\forall u \in I$, then $|I_h| \le ch^{\frac{1}{N}}$, where $|I_h|$ denotes the Lebesgue measure of I_h and the constant $c = 2(2 + 3 + \ldots + N + d^{-1})$.

For the proof, see Ref. 16. ■

[1] Bambusi, D. and Graffi, S., "Time quasi-periodic unbounded perturbations of Schrödinger operators and KAM methods," Commun. Math. Phys. **219**, 465–480 (2001).

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

[3] Bourgain, J., "Recent progress on quasi-periodic lattic Schrödinger operators and Hamiltonian PDEs," Russ. Math. Surveys **59**(2), 231–246 (2004).

[4] Geng, J. and You, J., "KAM tori of Hamiltonian perturbations of 1D linear beam equations," J. Math. Anal. Appl. **277**, 104–121 (2003).

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

[6] Geng, J. and Yi, Y., "Quasi-periodic solutions in a nonlinear Schrödinger equation," J. Differ. Equations **233**, 512–542 (2007).

[7] Kappeler, T. and Pöschel, J., *KdV and KAM* (Springer-Verlag, Berlin/Heidelberg, 2003).

[8] Klainerman, S., "Long-time behaviour of solutions to nonlinear wave equations," Arch. Ration. Mech. An. **78**(1), 73–98 (1982).

[9] Kuksin, S. B., *Analysis of Hamiltonian PDEs* (Oxford University Press, Oxford, 2000).

[10] Kuksin, S. B., "Fifteen years of KAM for PDE," *Translations-American Mathematical Society* (American Mathematical Society, 2004).

[11] Lax, P. D., "Development of singularities of solutions of nonlinear hyperbolic partial differential equations," J. Math. Phys. **5**, 611–613 (1964).

[12] Liang, Z. and You, J., "Quasi-periodic solutions for 1D Schrödinger equation with higher order nonlinearity," SIAM J. Math. Anal. **36**, 1965–1990 (2005).

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

[14] Pöschel, J., "A KAM theorem for some nonlinear partial differential equations," Ann. Sc. Norm. Sup. Pisa. **23**, 119–148 (1996).

[15] Pöschel, J. and Trubowitz, E., *Inverse Spectral Theory* (Academic, New York, 1987).

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

[17] Liu, J. and Yuan, X., "Spectrum for quantum Duffing oscillator and small-divisor equation with large variable coefficient," Commun. Pur. Appl. Math. **63**, 1145–1172 (2010).

[18] Liu, J. and Yuan, X., "A KAM theorem for Hamiltonian particle differential equations with unbounded perturbations," Commun. Math. Phys. **307**, 629–673 (2011).

[19] Zhang, J., Gao, M., and Yuan, X., "KAM tori for reversible partial differential equations," Nonlinearity **24**, 1189–1228 (2011).