*Research Article*

# On the Reducibility for a Class of Quasi-Periodic Hamiltonian Systems with Small Perturbation Parameter

## Jia Li and Junxiang Xu

*Department of Mathematics, Southeast University, Nanjing 210096, China*

Correspondence should be addressed to Jia Li, lijia831112@163.com

We consider the following real two-dimensional nonlinear analytic quasi-periodic Hamiltonian system $\dot{x} = J\nabla_x H$, where $H(x,t,\varepsilon) = (1/2)\beta(x_1^2 + x_2^2) + F(x,t,\varepsilon)$ with $\beta \neq 0$, $\partial_x F(0,t,\varepsilon) = O(\varepsilon)$ and $\partial_{xx} F(0,t,\varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$. Without any nondegeneracy condition with respect to $\varepsilon$, we prove that for most of the sufficiently small $\varepsilon$, by a quasi-periodic symplectic transformation, it can be reduced to a quasi-periodic Hamiltonian system with an equilibrium.

## 1. Introduction

We first give some definitions and notations for our problem. A function $f(t)$ is called a quasi-periodic function with frequencies $\omega = (\omega_1, \omega_2, \ldots, \omega_l)$ if $f(t) = F(\omega_1 t, \omega_2 t, \ldots, \omega_l t)$ with $\theta_i = \omega_i t$, where $F(\theta_1, \theta_2, \ldots, \theta_l)$ is $2\pi$ periodic in all the arguments $\theta_j$, $j = 1, 2, \ldots, l$. If $F(\theta)$ ($\theta = (\theta_1, \theta_2, \ldots, \theta_l)$) is analytic on $D_\rho = \{\theta \in C^l/2\pi Z^l \mid |\operatorname{Im}\theta_i| \leq \rho, \ i = 1, 2, \ldots, l\}$, we call $f(t)$ analytic quasi-periodic on $D_\rho$. If all $q_{ij}(t)$ ($i, j = 1, 2 \ldots, n$) are analytic quasi-periodic on $D_\rho$, then the matrix function $Q(t) = (q_{ij}(t))_{1 \leq i,j \leq n}$ is called analytic quasi-periodic on $D_\rho$.

If $f(t)$ is analytic quasi-periodic on $D_\rho$, we can write it as Fourier series:

$$f(t) = \sum_{k \in Z^l} f_k e^{\mathrm{i}\langle k, \omega \rangle t}. \tag{1.1}$$

Define a norm of $f$ by $\|f\|_\rho = \sum_{k \in Z^l} |f_k| e^{|k|\rho}$. It follows that $|f_k| \leq \|f\|_\rho e^{-|k|\rho}$. If the matrix function $Q(t)$ is analytic quasi-periodic on $D_\rho$, we define the norm of $Q$ by $\|Q\|_\rho = n \times \max_{1 \leq i,j \leq n} \|q_{ij}\|_\rho$. It is easy to verify $\|Q_1 Q_2\|_\rho \leq \|Q_1\|_\rho \|Q_2\|_\rho$. The average of $Q(t)$ is denoted

by $[Q] = ([q_{ij}])_{1 \le i,j \le n}$, where

$$[q_{ij}] = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} q_{ij}(t)dt. \tag{1.2}$$

For the existence of the above limit, see [1].

Denote

$$D(r, \rho, \varepsilon_0) = \left\{ (x, \theta, \varepsilon) \in C^n \times \left( \frac{C^l}{2\pi Z^l} \right) \times C \mid |x| \le r, \ \theta \in D_\rho, \ |\varepsilon| \le \varepsilon_0 \right\}, \tag{1.3}$$

where $x = (x_1, x_2, \ldots, x_n)$ and $|x| = |x_1| + |x_2| + \cdots + |x_n|$.

Let $f(x, t, \varepsilon)$ be analytic quasi-periodic of $t$ and analytic in $x$ and $\varepsilon$ on $D(r, \rho, \varepsilon_0)$. Then $f(x, t, \varepsilon)$ can be expanded as

$$f(x, t, \varepsilon) = \sum_{m=0}^{\infty} \sum_{k \in Z^l} f_{mk}(x) \varepsilon^m e^{i \langle k, \omega \rangle t}. \tag{1.4}$$

Define a norm by

$$\|f\|_{D(r, \rho, \varepsilon_0)} = \sum_{m=0}^{\infty} \sum_{k \in Z^l} |f_{mk}|_r \varepsilon_0^m e^{\rho|k|}, \tag{1.5}$$

where $|f_{mk}|_r = \sup_{|x| \le r} |f_{mk}(x)|$. Note that

$$\|f_1 \cdot f_2\|_{D(r, \rho, \varepsilon_0)} \le \|f_1\|_{D(r, \rho, \varepsilon_0)} \cdot \|f_2\|_{D(r, \rho, \varepsilon_0)}. \tag{1.6}$$

*Problems*

The reducibility on the linear differential system has been studied for a long time. The well-known Floquet theorem tells us that if $A(t)$ is a $T$-periodic matrix, then the linear system $\dot{x} = A(t)x$ is always reducible to the constant coefficient one by a $T$-periodic change of variables. However, this cannot be generalized to the quasi-periodic system. In [2], Johnson and Sell considered the quasi-periodic system $\dot{x} = A(t)x$, where $A(t)$ is a quasi-periodic matrix. Under some "full spectrum" conditions, they proved that $\dot{x} = A(t)x$ is reducible. That is, there exists a quasi-periodic nonsingular transformation $x = \phi(t)y$, where $\phi(t)$ and $\phi(t)^{-1}$ are quasi-periodic and bounded, such that $\dot{x} = A(t)x$ is transformed to $\dot{y} = By$, where $B$ is a constant matrix.

In [3], Jorba and Simó considered the reducibility of the following linear system:

$$\dot{x} = (A + \varepsilon Q(t))x, \quad x \in R^n, \tag{1.7}$$

where $A$ is an $n \times n$ constant matrix with $n$ different eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $Q(t)$ is analytic quasi-periodic with respect to $t$ with frequencies $\omega = (\omega_1, \omega_2, \ldots, \omega_l)$. Here $\varepsilon$ is a small perturbation parameter. Suppose that the following nonresonance conditions hold:

$$\left| \langle k, \omega \rangle \sqrt{-1} + \lambda_i - \lambda_j \right| \geq \frac{\alpha}{|k|^\tau}, \tag{1.8}$$

for all $k \in Z^l \setminus \{0\}$, where $\alpha > 0$ is a small constant and $\tau > l - 1$. Assume that $\lambda_j^0(\varepsilon)$ ($j = 1, 2, \ldots, n$) are eigenvalues of $A + \varepsilon[Q]$. If the following non-degeneracy conditions hold:

$$\frac{d}{d\varepsilon} \left( \lambda_i^0(\varepsilon) - \lambda_j^0(\varepsilon) \right) \Big|_{\varepsilon=0} \neq 0, \quad \forall i \neq j, \tag{1.9}$$

then authors proved that for sufficiently small $\varepsilon_0 > 0$, there exists a nonempty Cantor subset $E \subset (0, \varepsilon_0)$, such that for $\varepsilon \in E$, the system (1.7) is reducible. Moreover, $\operatorname{meas}((0, \varepsilon_0) \setminus E) = o(\varepsilon_0)$.

Some related problems were considered by Eliasson in [4, 5]. In the paper [4], to study one-dimensional linear Schrödinger equation

$$\frac{d^2 q}{dt^2} + Q(\omega t)q = Eq, \tag{1.10}$$

Eliasson considered the following equivalent two-dimensional quasi-periodic Hamiltonian system:

$$\dot{p} = (E - Q(\omega t))q, \quad \dot{q} = p, \tag{1.11}$$

where $Q$ is an analytic quasi-periodic function and $E$ is an energy parameter. The result in [4] implies that for almost every sufficiently large $E$, the quasi-periodic system (1.11) is reducible. Later, in [5] the author considered the almost reducibility of linear quasi-periodic systems. Recently, the similar problem was considered by Her and You [6]. Let $C^\omega(\Lambda, gl(m, C))$ be the set of $m \times m$ matrices $A(\lambda)$ depending analytically on a parameter $\lambda$ in a closed interval $\Lambda \subset R$. In [6], Her and You considered one-parameter families of quasi-periodic linear equations

$$\dot{x} = \left( A(\lambda) + g(\omega_1 t, \ldots, \omega_l t, \lambda) \right) x, \tag{1.12}$$

where $A \in C^\omega(\Lambda, gl(m, C))$, and $g$ is analytic and sufficiently small. They proved that under some nonresonance conditions and some non-degeneracy conditions, there exists an open and dense set $\mathcal{A}$ in $C^\omega(\Lambda, gl(m, C))$, such that for each $A \in \mathcal{A}$, the system (1.12) is reducible for almost all $\lambda \in \Lambda$.

In 1996, Jorba and Simó extended the conclusion of the linear system to the nonlinear case. In [7], Jorba and Simó considered the quasi-periodic system

$$\dot{x} = (A + \varepsilon Q(t))x + \varepsilon g(t) + h(x, t), \quad x \in R^n, \tag{1.13}$$

where $A$ has $n$ different nonzero eigenvalues $\lambda_i$. They proved that under some nonresonance conditions and some non-degeneracy conditions, there exists a nonempty Cantor subset $E \subset (0, \varepsilon_0)$, such that the system (1.13) is reducible for $\varepsilon \in E$.

In [8], the authors found that the non-degeneracy condition is not necessary for the two-dimensional quasi-periodic system. They considered the two-dimensional nonlinear quasi-periodic system:

$$\dot{x} = Ax + f(x, t, \varepsilon), \quad x \in R^2, \tag{1.14}$$

where $A$ has a pair of pure imaginary eigenvalues $\pm\sqrt{-1}\omega_0$ with $\omega_0 \neq 0$ satisfying the nonresonance conditions

$$|\langle k, \omega \rangle| \geq \frac{\alpha}{|k|^\tau}, \qquad |\langle k, \omega \rangle - 2\omega_0| \geq \frac{\alpha}{|k|^\tau} \tag{1.15}$$

for all $k \in Z^l \setminus \{0\}$, where $\alpha > 0$ is a small constant and $\tau > l - 1$. Assume that $f(0, t, \varepsilon) = O(\varepsilon)$ and $\partial_x f(0, t, \varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$. They proved that either of the following two results holds:

(1) for $\forall \varepsilon \in (0, \varepsilon_0)$, the system (1.14) is reducible to $\dot{y} = By + O(y)$ as $y \to 0$;

(2) there exists a nonempty Cantor subset $E \subset (0, \varepsilon_0)$, such that for $\varepsilon \in E$ the system (1.14) is reducible to $\dot{y} = By + O(y^2)$ as $y \to 0$.

Note that the result (1) happens when the eigenvalue of the perturbed matrix of $A$ in KAM steps has nonzero real part. But the authors were interested in the equilibrium of the transformed system and obtained a small quasi-periodic solution for the original system.

Motivated by [8], in this paper we consider the Hamiltonian system and we have a better result.

## 2. Main Results

**Theorem 2.1.** *Consider the following real two-dimensional Hamiltonian system*

$$\dot{x} = J\nabla_x H, \quad x \in R^2, \tag{2.1}$$

*where $H(x, t, \varepsilon) = (1/2)\beta(x_1^2 + x_2^2) + F(x, t, \varepsilon)$ with $\beta \neq 0$, $F(x, t, \varepsilon)$ is analytic quasi-periodic with respect to $t$ with frequencies $\omega = (\omega_1, \omega_2, \ldots, \omega_l)$ and real analytic with respect to $x$ and $\varepsilon$ on $D(r, \rho, \varepsilon_0)$, and*

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.2}$$

*Here $\varepsilon \in (0, \varepsilon_0)$ is a small parameter. Suppose that $\partial_x F(0, t, \varepsilon) = O(\varepsilon)$ and $\partial_{xx} F(0, t, \varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$. Moreover, assume that $\beta$ and $\omega$ satisfy*

$$|\langle k, \omega \rangle| \geq \frac{\alpha_0}{|k|^\tau}, \tag{2.3}$$

$$|\langle k, \omega \rangle - 2\beta| \geq \frac{\alpha_0}{|k|^\tau} \tag{2.4}$$

*for all $k \in Z^l \setminus \{0\}$, where $\alpha_0 > 0$ is a small constant and $\tau > l - 1$.*

*Then there exist a sufficiently small $\varepsilon_* \in (0, \varepsilon_0]$ and a nonempty Cantor subset $E_* \subset (0, \varepsilon_*)$, such that for $\varepsilon \in E_*$, there exists an analytic quasi-periodic symplectic transformation $x = \phi_*(t)y + \psi_*(t)$ on $D_{\rho/2}$ with the frequencies $\omega$, which changes (2.1) into the Hamiltonian system $\dot{y} = J\nabla_y H_*$, where $H_*(y, t, \varepsilon) = 1/2\beta_*(\varepsilon)(y_1^2 + y_2^2) + F_*(y, t, \varepsilon)$, where $F_*(y, t, \varepsilon) = O(y^3)$ as $y \to 0$. Moreover, $\text{meas}((0, \varepsilon_*) \setminus E_*) = o(\varepsilon_*)$ as $\varepsilon_* \to 0$. Furthermore, $\beta_*(\varepsilon) = \beta + O(\varepsilon)$ and $\|\phi_* - Id\|_{\rho/2} + \|\psi_*\|_{\rho/2} = O(\varepsilon)$, where $Id$ is the 2-order unit matrix.*

## 3. The Lemmas

The proof of Theorem 2.1 is based on KAM-iteration. The idea is the same as [7, 8]. When the non-degeneracy conditions do not happen, the small parameter $\varepsilon$ is not involved in the nonresonance conditions. So without deleting any parameter, the KAM step will be valid. Once the non-degeneracy conditions occur at some step, they will be kept for ever and we can apply the results with the non-degeneracy conditions. Thus, after infinite KAM steps, the transformed system is convergent to a desired form.

We first give some lemmas. Let $R = (r_{ij})_{1 \leq i, j \leq 2}$ be a Hamiltonian matrix. Then we have $r_{11} + r_{22} = 0$. Define a matrix $R_A = (1/2)dJ$ with $d = r_{12} - r_{21}$. Let

$$B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ \sqrt{-1} & -\sqrt{-1} \end{pmatrix}. \tag{3.1}$$

It is easy to verify

$$B^{-1}R_A B = \frac{1}{2}\text{diag}\left(\sqrt{-1}d, -\sqrt{-1}d\right),$$

$$B^{-1}(R - R_A)B = \frac{1}{2}\begin{pmatrix} 0 & \sigma' - \sqrt{-1}\kappa' \\ \sigma' + \sqrt{-1}\kappa' & 0 \end{pmatrix}, \tag{3.2}$$

where $\sigma' = 2r_{11}$ and $\kappa' = r_{21} + r_{12}$.

In the same way as in [7, 8], in KAM steps we need to solve linear homological equations. For this purpose we need the following lemma.

**Lemma 3.1.** *Consider the following equation of the matrix:*

$$\dot{P} = AP - PA + R(t), \tag{3.3}$$

*where $A = \beta(\varepsilon)J$ with $|\beta(\varepsilon)| > \mu$, $\mu > 0$ is a constant, and $R(t) = (r_{ij}(t))_{1 \le i, j \le 2}$ is a real analytic quasi-periodic Hamiltonian matrix on $D_\rho$ with frequencies $\omega$. Suppose $\beta(\varepsilon)$ and $R$ are smooth with respect to $\varepsilon$ and $|\varepsilon\beta'(\varepsilon)| \le c_0$ for $\varepsilon \in E \subset (0, \varepsilon_*)$, where $c_0$ is a constant. Note that here and below the dependence of $\varepsilon$ is usually implied and one does not write it explicitly for simplicity. Assume $[R]_A = 0$, where $[R]$ is the average of $R$. Suppose that for $\varepsilon \in E$, the small divisors conditions (2.3) and the following small divisors conditions hold:*

$$\left| \langle k, \omega \rangle - 2\beta(\varepsilon) \right| \ge \frac{\alpha}{|k|^{\tau'}}, \tag{3.4}$$

*where $\tau' > 2\tau + l$. Let $0 < s < \rho$ and $\rho_1 = \rho - s$. Then there exists a unique real analytic quasi-periodic Hamiltonian matrix $P(t)$ with frequencies $\omega$, which solves the homological linear equation (3.3) and satisfies*

$$\|P\|_{\rho_1} \le \frac{c}{\alpha s^v} \|R\|_\rho, \qquad \|\varepsilon\partial_\varepsilon P\|_{\rho_1} \le \frac{c}{\alpha^2 s^{v'}} \left( \|R\|_\rho + \|\varepsilon\partial_\varepsilon R\|_\rho \right), \tag{3.5}$$

*where $v = \tau' + l$, $v' = 2\tau' + l$ and $c > 0$ is a constant.*

*Remark 3.2.* The subset $E$ of $(0, \varepsilon_*)$ is usually a Cantor set and so the derivative with respect to $\varepsilon$ should be understood in the sense of Whitney [9].

*Proof.* Let $\overline{P} = B^{-1}PB$, where $B$ is defined by (3.1). Similarly, define $\overline{A}$, $\overline{R}$, $\overline{R}_A$. Then (3.3) becomes

$$\dot{\overline{P}} = \overline{A}\,\overline{P} - \overline{P}\,\overline{A} + \overline{R}(t), \tag{3.6}$$

where

$$\overline{A} = \mathrm{diag}\left( \sqrt{-1}\beta, -\sqrt{-1}\beta \right). \tag{3.7}$$

Moreover, $\overline{R}_A$ and $\overline{R} - \overline{R}_A$ have the same forms as (3.2) and (3.2), respectively
Noting that $[R]_A = 0$, we have $[\overline{R}]_A = 0$. Write $\overline{P} = (\overline{p}_{ij})_{i,j}$ and $\overline{R} = (\overline{r}_{ij})_{i,j}$. Obviously, we have $\overline{r}_{11} = -\overline{r}_{22}$ with $[\overline{r}_{ii}] = 0$.

Insert the Fourier series of $\overline{P}$ and $\overline{R}$ into (3.6). Then it follows that $\overline{p}_{ii}^0 = 0$, $\overline{p}_{ii}^k = \overline{r}_{ii}^k / (\langle k, \omega \rangle \sqrt{-1})$ for $k \ne 0$, and

$$\overline{p}_{ij}^k = \frac{\overline{r}_{ij}^k}{\sqrt{-1}(\langle k, \omega \rangle \pm 2\beta)} \quad \text{for } i \ne j. \tag{3.8}$$

Since $\overline{R}$ is analytic on $D_\rho$, we have $|\overline{R}_k| \le \|\overline{R}\|_\rho e^{-|k|\rho}$. So it follows

$$\left\| \overline{P} \right\|_{\rho-s} \le \sum_{k \in \mathbb{Z}^l} \left| \overline{P}_k \right| e^{|k|\,(\rho-s)} \le \frac{c}{\alpha s^v} \|R\|_\rho. \tag{3.9}$$

Note that here and below we always use $c$ to indicate constants, which are independent of KAM steps.

Since $A$ and $R(t)$ are real matrices, it is easy to obtain that $P(t)$ is also a real matrix. Obviously, it follows that $\overline{p}_{11} = -\overline{p}_{22}$ and the trace of the matrix $\overline{P}$ is zero. So is the trace of $P$. Thus, $P$ is a Hamiltonian matrix.

Now we estimate $\|\varepsilon \partial P/\partial \varepsilon\|_{\rho_1}$. We only consider $\overline{p}_{12}$ and $\overline{p}_{21}$ since $\overline{p}_{11}$ and $\overline{p}_{22}$ are easy.

For $i \neq j$ we have

$$\frac{d\overline{p}_{ij}^k(\varepsilon)}{d\varepsilon} = \frac{\pm 2\beta'(\varepsilon)\overline{r}_{ij}^k - (\langle k, \omega \rangle \pm 2\beta)\overline{r}_{ij}^k{}'(\varepsilon)}{-\sqrt{-1}(\langle k, \omega \rangle \pm 2\beta)^2}. \tag{3.10}$$

Then, in the same way as above we obtain the estimate for $\|\varepsilon(\partial P/\partial \varepsilon)\|_{\rho_1}$. $\qquad\square$

The following lemma will be used for the zero order term in KAM steps.

**Lemma 3.3.** *Consider the equation*

$$\dot{x} = Ax + g(t), \tag{3.11}$$

*where $A$ is the same as in Lemma 3.1, and $g$ is real analytic quasi-periodic in $t$ on $D_\rho$ with frequencies $\omega$ and smooth with respect to $\varepsilon$. Suppose that the small divisors conditions (3.4) hold. Then there exists a unique real analytic quasi-periodic solution $x(t)$ with frequencies $\omega$, which satisfies*

$$\|x\|_{\rho_1} \leq \frac{c}{\alpha s^v}\|g\|_\rho, \qquad \left\|\varepsilon \frac{\partial x}{\partial \varepsilon}\right\|_{\rho_1} \leq \frac{c}{\alpha^2 s^{v'}}\left(\|g\|_\rho + \left\|\varepsilon \frac{\partial g}{\partial \varepsilon}\right\|_\rho\right), \tag{3.12}$$

*where $s, \rho_1, v, v'$ are defined in Lemma 3.1.*

*Proof.* Similarly, let $\overline{x} = B^{-1}x, \overline{A} = B^{-1}AB$ and $\overline{g}(t) = B^{-1}g(t)$. Then (3.11) becomes

$$\dot{\overline{x}} = \overline{A}\,\overline{x} + \overline{g}(t), \tag{3.13}$$

where $\overline{A} = \mathrm{diag}(\sqrt{-1}\beta, -\sqrt{-1}\beta)$. Expanding $\overline{x} = (\overline{x}_1, \overline{x}_2)$ and $\overline{g} = (\overline{g}_1, \overline{g}_2)$ into Fourier series and using (3.13), we have

$$\overline{x}_i^k = \frac{\overline{g}_i^k}{\sqrt{-1}\left(\langle k, \omega \rangle + (-1)^i \beta\right)}. \tag{3.14}$$

Using $2k$ in place of $k$ in (3.4), we have

$$\left|\langle k, \omega\rangle - \beta(\varepsilon)\right| \geq \frac{\alpha}{2|k|^{\tau'}}. \tag{3.15}$$

Thus, in the same way as the proof of Lemma 3.1, we can estimate $\|x\|_{\rho_1}$ and $\|\varepsilon\partial_\varepsilon x\|_{\rho_1}$. We omit the details.                                                                                   $\square$

The following lemma is used in the estimate of Lebesgue measure for the parameter $\varepsilon$ in the case of non-degeneracy.

**Lemma 3.4.** *Let $\psi(\varepsilon) = \sigma\varepsilon^N + \varepsilon^N f(\varepsilon)$, where $N$ is a positive integer and $f$ satisfies that $f(\varepsilon) \to 0$ as $\varepsilon \to 0$ and $|f'(\varepsilon)| \leq c$ for $\varepsilon \in (0, \varepsilon_*)$. Let $\phi(\varepsilon) = \langle k, \omega\rangle - 2\beta - \psi(\varepsilon)$. Let*

$$O = \left\{ \varepsilon \in (0, \varepsilon_*) \mid |\phi(\varepsilon)| \geq \frac{\alpha}{|k|^{\tau'}}, \ \forall k \neq 0 \right\}, \tag{3.16}$$

*where $\tau' \geq 2\tau + l$, $\alpha \leq (1/2)\alpha_0, \sigma \neq 0$. Suppose that the small condition (2.4) holds. Then when $\varepsilon_*$ is sufficiently small, one has*

$$\mathrm{meas}(0, \varepsilon_*) \setminus O \leq c \frac{\alpha}{\alpha_0^2} \varepsilon_*^{N+1}, \tag{3.17}$$

*where $c$ is a constant independent of $\alpha_0, \alpha, \varepsilon_*$*

*Proof.* Let

$$O_k = \left\{ \varepsilon \in (0, \varepsilon_*) \mid |\phi(\varepsilon)| < \frac{\alpha}{|k|^{\tau'}} \right\}. \tag{3.18}$$

By assumption, if $\varepsilon_*$ is sufficient small, we have that $|\psi(\varepsilon)| \leq 2\sigma\varepsilon^N$ and $|\psi'(\varepsilon)| \geq (\sigma/2)\varepsilon^{N-1}$ for $\varepsilon \in (0, \varepsilon_*)$. If $\varepsilon^N \leq \alpha_0/(4\sigma|k|^\tau)$, by (2.4) we have

$$\left|\phi(\varepsilon)\right| \geq \left|\langle k, \omega\rangle - 2\beta\right| - \left|\psi(\varepsilon)\right| \geq \frac{\alpha}{|k|^{\tau'}}. \tag{3.19}$$

Thus, we only consider the case that $\varepsilon_*^N \geq \varepsilon^N \geq (\alpha_0/(4\sigma|k|^\tau))$. We have $|k| \geq (\alpha_0/(4\sigma\varepsilon_*^N))^{1/\tau} = K$. Since

$$\left|\phi'(\varepsilon)\right| = \left|\psi'(\varepsilon)\right| \geq \frac{\sigma}{2}\varepsilon^{N-1} \geq \frac{\alpha_0}{8|k|^\tau \varepsilon_*}, \tag{3.20}$$

we have $\text{meas}(O_k) \leq ((2\alpha)/|k|^{\tau'}) \times ((8|k|^{\tau}\varepsilon_*)/\alpha_0) = (16\alpha\varepsilon_*)/(|k|^{\tau'-\tau}\alpha_0)$. So

$$\begin{aligned}
\text{meas}((0,\varepsilon_*) \setminus 0) &\leq \sum_{|k| \geq K} \text{meas}(O_k) \leq \frac{16\alpha}{\alpha_0}\varepsilon_* \sum_{|k| \geq K} \frac{1}{|k|^{\tau'-\tau}} \\
&\leq \frac{c\alpha}{\alpha_0}\varepsilon_* K^{l-\tau'+\tau} \leq \frac{c\alpha}{\alpha_0^2}\varepsilon_*^{N+1},
\end{aligned} \tag{3.21}$$

where $c$ is a constant independent of $\alpha_0, \alpha$, and $\varepsilon_*$. □

Below we give a lemma with the non-degeneracy conditions.

**Lemma 3.5.** *Consider the real nonlinear Hamiltonian system $\dot{x} = J\nabla_x H$, where*

$$H(x,t,\varepsilon) = \frac{1}{2}\beta\left(x_1^2 + x_2^2\right) + F(x,t,\varepsilon) \quad \text{with } \beta \neq 0. \tag{3.22}$$

*Suppose that $F(x,t,\varepsilon)$ is analytic quasi-periodic with respect to $t$ with frequencies $\omega$ and real analytic with respect to $x$ and $\varepsilon$ on $D(r,\rho,\varepsilon_0)$. Let $f(x,t,\varepsilon) = J\nabla_x F(x,t,\varepsilon)$. Assume that $f(0,t,\varepsilon) = O(\varepsilon^{2m_0})$ and $\partial_x f(0,t,\varepsilon) = O(\varepsilon^{m_0})$ as $\varepsilon \to 0$, where $m_0$ is a positive integer. Let $Q(t,\varepsilon) = \partial_x f(0,t,\varepsilon) = \sum_{k \geq m_0} Q_k(t)\varepsilon^k$. Suppose there exists $m_0 \leq k \leq 2m_0 - 1$ such that $[Q_k]_A \neq 0$ and the nonresonance conditions (2.3) and (2.4) hold. Then, for sufficiently small $\varepsilon_* > 0$, there exists a nonempty Cantor subset $E_* \subset (0,\varepsilon_*)$, such that for $\varepsilon \in E_*$, there exists a quasi-periodic symplectic transformation $x = \phi_*(t)y + \psi_*(t)$ with the frequencies $\omega$, which changes the Hamiltonian system to $\dot{y} = J\nabla_y H_*$, where*

$$H_*(y,t,\varepsilon) = \frac{1}{2}\beta_*(\varepsilon)\left(y_1^2 + y_2^2\right) + F_*(y,t,\varepsilon), \tag{3.23}$$

*where $F_*(y,t,\varepsilon) = O(y^3)$ as $y \to 0$. Moreover, $\text{meas}((0,\varepsilon_*) \setminus E_*) = O(\varepsilon_*^{m_0+1})$ as $\varepsilon_* \to 0$. Furthermore, $\beta_*(\varepsilon) = \beta + O(\varepsilon^{m_0})$ and $\|\phi_* - Id\|_{\rho/2} + \|\psi_*\|_{\rho/2} = O(\varepsilon^{m_0})$.*

*Proof*

*KAM Step*

The proof is based on a modified KAM iteration. In spirit, it is very similar to [7, 8]. The important thing is to make symplectic transformations so that the Hamiltonian structure can be preserved. Note that $[Q_k]_A \neq 0$ for some $m_0 \leq k \leq 2m_0 - 1$ is a non-degeneracy condition.

Consider the following Hamiltonian system

$$\dot{x} = Ax + f(x,t,\varepsilon), \tag{3.24}$$

where $A = \beta(\varepsilon)J$ and $f$ is analytic quasi-periodic with respect to $t$ with frequencies $\omega$ and real analytic with respect to $x$ and $\varepsilon$ on $D = D(r,\rho,\varepsilon_*)$.

Let $\|f\|_D \leq \alpha r \widetilde{\varepsilon}$ and $\|\varepsilon \partial_\varepsilon f\|_D \leq \alpha r \widetilde{\varepsilon}$. Let $Q(t, \varepsilon) = \partial_x f(0, t, \varepsilon)$, $g(t, \varepsilon) = f(0, t, \varepsilon)$ and

$$h(x, t, \varepsilon) = f(x, t, \varepsilon) - g(t, \varepsilon) - Q(t, \varepsilon)x. \tag{3.25}$$

Then $h$ is the higher-order term of $f$. Moreover, the matrix $Q(t, \varepsilon)$ is Hamiltonian. Let $[Q]_A = \widehat{\beta}(\varepsilon)J$.

The system (3.24) is written as

$$\dot{x} = (A_+ + R(t, \varepsilon))x + g(t, \varepsilon) + h(x, t, \varepsilon), \tag{3.26}$$

where $A_+ = A + [Q]_A = \beta_+(\varepsilon)J$ and $R = Q - [Q]_A$. By assumption we have

$$\|g\|_\rho \leq \alpha r \widetilde{\varepsilon}, \qquad \|Q\|_\rho \leq \alpha \widetilde{\varepsilon}, \qquad \|h\|_D \leq 3\alpha r \widetilde{\varepsilon}. \tag{3.27}$$

Moreover, we have

$$\|\varepsilon \partial_\varepsilon g\|_\rho \leq \alpha r \widetilde{\varepsilon}, \qquad \|\varepsilon \partial_\varepsilon Q\|_\rho \leq \alpha \widetilde{\varepsilon}, \qquad \|\varepsilon \partial_\varepsilon h\|_D \leq 3\alpha r \widetilde{\varepsilon}. \tag{3.28}$$

Now we want to construct the symplectic change of variables $x = T'y = e^{P(t)}y$ to (3.26), where $P$ is a Hamiltonian matrix to be defined later. Then we have

$$\dot{y} = \left( e^{-P}(A_+ + R - \dot{P})e^P + e^{-P}\left( \dot{P}e^P - \frac{d}{dt}e^{P(t)} \right) \right)y$$
$$+ e^{-P}g(t, \varepsilon) + e^{-P}h\left(e^P y, t, \varepsilon\right). \tag{3.29}$$

Let $W = e^P - I - P$ and $\widetilde{W} = e^{-P} - I - P$. Then the system (3.29) becomes

$$\dot{y} = (A_+ + R - \dot{P} + A_+P - PA_+)y + Q'y + e^{-P}g(t, \varepsilon) + e^{-P}h\left(e^P y, t, \varepsilon\right), \tag{3.30}$$

where

$$Q' = -P(R - \dot{P}) + (R - \dot{P})P - P(A_+ + R - \dot{P})P$$
$$- P(A_+ + R - \dot{P})W + (A_+ + R - \dot{P})W$$
$$+ \widetilde{W}(A_+ + R - \dot{P})e^P + e^{-P}\left( \dot{P}e^P - \frac{d}{dt}e^P \right). \tag{3.31}$$

We would like to have

$$\dot{P} - A_+P + PA_+ = R, \tag{3.32}$$

where $R = Q - [Q]_A$. Suppose the small divisors conditions (2.3) hold. Let $E_+ \subset (0, \varepsilon_*)$ be a subset such that for $\varepsilon \in E_+$ the small divisors conditions hold:

$$\left| \langle k, \omega \rangle - 2\beta_+(\varepsilon) \right| \geq \frac{\alpha_+}{|k|^{\tau'}}, \quad \forall k \in Z^l \setminus \{0\}, \tag{3.33}$$

where $\tau' > 2\tau + l$. By Lemma 3.1, we have a quasi-periodic Hamiltonian matrix $P(t)$ with frequencies $\omega$ to solve the above equation with the following estimates:

$$\|P\|_{\rho-s} \leq \frac{c\|Q\|_\rho}{\alpha_+ s^v} \leq \frac{c\tilde{\varepsilon}}{s^v},$$

$$\left\| \varepsilon \frac{\partial P}{\partial \varepsilon} \right\|_{\rho-s} \leq \frac{c}{\alpha_+^2 s^{v'}} \left( \|Q\|_\rho + \left\| \varepsilon \frac{\partial Q}{\partial \varepsilon} \right\|_\rho \right) \leq \frac{c\tilde{\varepsilon}}{\alpha_+ s^{v'}}, \tag{3.34}$$

where $v = \tau' + l$, $v' = 2\tau' + l$ and $c > 0$ is a constant. Then the system (3.30) becomes

$$\dot{y} = A_+ y + f'(y, t, \varepsilon), \tag{3.35}$$

where $f' = Q'y + e^{-P}g(t, \varepsilon) + e^{-P}h(e^P y, t, \varepsilon)$.

By Lemma 3.3, let us denote by $\underline{x}$ the solution of $\dot{x} = A_+ x + g'(t, \varepsilon)$ on $D_{\rho-2s}$, where $g' = e^{-P}g(t, \varepsilon)$. Then, by Lemma 3.3 we have

$$\|\underline{x}\|_{\rho-2s} \leq \frac{c\|g\|_{\rho-s}}{\alpha_+ s^v} \leq \frac{cr\tilde{\varepsilon}}{s^v},$$

$$\left\| \varepsilon \frac{\partial \underline{x}}{\partial \varepsilon} \right\|_{\rho-2s} \leq \frac{c}{\alpha_+^2 s^{v'}} \left( \|g\|_{\rho-s} + \left\| \varepsilon \frac{\partial g}{\partial \varepsilon} \right\|_{\rho-s} \right) \leq \frac{cr\tilde{\varepsilon}}{\alpha_+ s^{v'}}. \tag{3.36}$$

Under the symplectic change of variables $y = T''x_+ = \underline{x} + x_+$, the Hamiltonian system (3.35) is changed to

$$\dot{x}_+ = A_+ x_+ + f_+(x_+, t, \varepsilon), \tag{3.37}$$

where $A_+ = \beta_+ J$ and

$$f_+ = Q' \cdot T'' + e^{-P}h \circ T' \circ T''. \tag{3.38}$$

Let the symplectic transformation $T = T' \circ T''$. Then $x = Tx_+ = \phi(t)x_+ + \psi(t)$, where $\phi(t) = e^{P(t)}$ and $\psi(t) = e^{P(t)}\underline{x}(t)$. It is easy to obtain that if $\|P\|_{\rho-2s} \leq 1/2$, then

$$
\|\phi - I\|_{\rho-2s} \leq \frac{c\widetilde{\varepsilon}}{s^v}, \qquad \|\varepsilon\partial_\varepsilon\phi\|_{\rho-2s} \leq \frac{c\widetilde{\varepsilon}}{\alpha_+ s^{v'}},
$$

$$
\|\psi\|_{\rho-2s} \leq \frac{cr\widetilde{\varepsilon}}{s^v}, \qquad \|\varepsilon\partial_\varepsilon\psi\|_{\rho-2s} \leq \frac{cr\widetilde{\varepsilon}}{\alpha_+ s^{v'}}. \tag{3.39}
$$

Under the symplectic change of variables $x = Tx_+$, the Hamiltonian system (3.24) becomes (3.37).

Below we give the estimates for $A_+$ and $f_+$. Obviously, it follows that $A_+(\varepsilon) - A = [Q]_A = \widehat{\beta}(\varepsilon)J$ and

$$
|\beta_+(\varepsilon) - \beta(\varepsilon)| = |\widehat{\beta}(\varepsilon)| \leq c\alpha\widetilde{\varepsilon}, \qquad |\varepsilon(\beta'_+(\varepsilon) - \beta'(\varepsilon))| = |\varepsilon\widehat{\beta}'(\varepsilon)| \leq c\alpha\widetilde{\varepsilon}. \tag{3.40}
$$

By (3.38) we have

$$
f_+(x_+,t,\varepsilon) = Q'(t)(x_+ + \underline{x}(t)) + e^{-P(t)}h\left(e^{P(t)}(x_+ + \underline{x}(t)),t,\varepsilon\right). \tag{3.41}
$$

Let $\rho_+ = \rho - 2s$, and $r_+ = \eta r$ with $\eta \leq 1/8$. If $c\widetilde{\varepsilon}/\alpha_+ s^{v+v'} \leq \eta$, it follows that $\|\underline{x}\|_{\rho-2s} \leq (1/8)r$. Let $D_+ = D(r_+, s_+, \varepsilon_*)$. Note that $Q'$ and $h$ only consist of high-order terms of $P$ and $x$, respectively. It is easy to see $|e^{P(t)}(x_+ + \underline{x}(t))| \leq 4\eta r \leq r$. By all the estimates (3.27), (3.28), (3.34), and (3.36), and using usual technique of KAM estimate, we have

$$
\|f_+\|_{D_+} \leq \frac{c\widetilde{\varepsilon}^2}{s^{2v}}\eta r + c\alpha r\widetilde{\varepsilon}\eta^2 \leq \left(\frac{c\widetilde{\varepsilon}}{s^{2v}} + c\alpha\eta\right)r_+\widetilde{\varepsilon},
$$

$$
\|\varepsilon\partial_\varepsilon f_+\|_{D_+} \leq \frac{c\widetilde{\varepsilon}^2}{\alpha_+ s^{v+v'}}\eta r + c\alpha r\widetilde{\varepsilon}\eta^2 \leq \left(\frac{c\widetilde{\varepsilon}}{\alpha s^{v+v'}} + c\alpha\eta\right)r_+\widetilde{\varepsilon}. \tag{3.42}
$$

Let $\alpha_+ = \alpha/2$ and $\eta = c\widetilde{\varepsilon}/(\alpha^2 s^{v+v'})$. Then we have

$$
\|f_+\|_{D_+} \leq c\alpha_+ r_+\eta\widetilde{\varepsilon} = \alpha_+ r_+\widetilde{\varepsilon}_+, \qquad \widetilde{\varepsilon}_+ = c\eta\widetilde{\varepsilon}. \tag{3.43}
$$

Similarly, we have

$$
\|\varepsilon\partial_\varepsilon f_+\|_{D_+} \leq \alpha_+ r_+\widetilde{\varepsilon}_+. \tag{3.44}
$$

Note that KAM steps only make sense for the small parameter $\varepsilon$ satisfying small divisors conditions. However, by Whitney's extension theorem, for convenience all the functions are supposed to be defined for $\varepsilon$ on $[0, \varepsilon_*]$.

*KAM Iteration*

Now we can give the iteration procedure in the same way as in [7] and prove its convergence.

At the initial step, let $f_0 = f$. Let $f(x, t, \varepsilon) = f(0, t, \varepsilon) + \partial_x f(0, t, \varepsilon)x + h(x, t, \varepsilon)$. By assumption, if $\varepsilon_*$ is sufficiently small, we have that for all $\varepsilon \in [0, \varepsilon_*]$

$$
\left| f(0, t, \varepsilon) \right| \leq c\varepsilon^{2m_0}, \qquad \left| \partial_x f(0, t, \varepsilon) \right| \leq c\varepsilon^{m_0},
$$
$$
\left| \varepsilon\partial_\varepsilon f(0, t, \varepsilon) \right| \leq c\varepsilon^{2m_0}, \qquad \left| \varepsilon\partial_\varepsilon\partial_x f(0, t, \varepsilon) \right| \leq c\varepsilon^{m_0}. \tag{3.45}
$$

Moreover,

$$
|h(x, t, \varepsilon)| \leq c|x|^2, \quad |\varepsilon\partial_\varepsilon h(x, t, \varepsilon)| \leq c|x|^2, \quad \forall |x| \leq \varepsilon^{m_0}, \quad \forall \varepsilon \in [0, \varepsilon_*]. \tag{3.46}
$$

Let $r_0 = \varepsilon^{m_0}$, $\rho_0 = \rho$, $s_0 = \rho_0/8$, $D_0 = D(r_0, \rho_0, \varepsilon_*)$, and $\widetilde{\varepsilon}_0 = c\varepsilon^{m_0}/\alpha_0$. Then we have

$$
\left| f_0 \right|_{D_0} \leq \alpha_0 r_0 \widetilde{\varepsilon}_0, \qquad \left| \varepsilon\partial_\varepsilon f_0 \right|_{D_0} \leq \alpha_0 r_0 \widetilde{\varepsilon}_0. \tag{3.47}
$$

For $n \geq 1$, let

$$
\alpha_n = \frac{\alpha_{n-1}}{2}, \qquad s_n = \frac{s_{n-1}}{2}, \qquad \rho_n = \rho_{n-1} - 2s_{n-1},
$$
$$
\eta_{n-1} = \frac{c\widetilde{\varepsilon}_{n-1}}{\alpha_{n-1}^2 s_{n-1}^{v+v'}}, \qquad r_n = \eta_{n-1} r_{n-1}, \qquad \widetilde{\varepsilon}_n = c\eta_{n-1}\widetilde{\varepsilon}_{n-1}. \tag{3.48}
$$

Then we have a sequence of quasi-periodic symplectic transformations $\{T_n\}$ satisfying $T_n x = \phi_n(t)x + \psi_n(t)$ with

$$
\left\| \phi_n - I \right\|_{\rho_{n+1}} \leq \frac{c\widetilde{\varepsilon}_n}{s_n^v}, \qquad \left\| \psi_n \right\|_{\rho_{n+1}} \leq \frac{cr_n\widetilde{\varepsilon}_n}{s_n^v}. \tag{3.49}
$$

Let $T^n = T_0 \circ T_1 \cdots \circ T_{n-1}$. Then under the transformation $x = T^n y$ the Hamiltonian system $\dot{x} = A_0 x + f_0(x, t, \varepsilon)$ is changed to $\dot{y} = A_n y + f_n(y, t, \varepsilon)$.

Moreover, $A_n(\varepsilon) = \beta_n(\varepsilon)J$ satisfies $A_{n+1} - A_n = [Q_n]_A$ and

$$
\left| \beta_{n+1}(\varepsilon) - \beta_n(\varepsilon) \right| \leq c\alpha_n\widetilde{\varepsilon}_n, \qquad \left| \varepsilon\left( \beta'_{n+1}(\varepsilon) - \beta'_n(\varepsilon) \right) \right| \leq c\alpha_n\widetilde{\varepsilon}_n, \tag{3.50}
$$

$$
\left\| f_n \right\|_{D_n} \leq \alpha_n r_n \widetilde{\varepsilon}_n. \tag{3.51}
$$

*Convergence*

By the above definitions we have $\eta_n/\eta_{n-1} = c\widetilde{\varepsilon}_n/\widetilde{\varepsilon}_{n-1} = c\eta_{n-1}$. Thus, we have $\eta_n \leq c\eta_{n-1}^2$ and so $c\eta_n \leq (c\eta_{n-1})^2 \leq (c\eta_0)^{2^n}$. Note that $\eta_0 = c\widetilde{\varepsilon}_0/(\alpha_0^2 s_0^{v+v'}) \leq c\varepsilon^{m_0}/(\alpha_0^2\rho_0^{v+v'})$. Suppose that $\varepsilon_*$ is sufficiently small such that for $0 < \varepsilon < \varepsilon_*$ we have $c\eta_0 \leq 1/2$. $T_n$ are affine, so are $T^n$

with $T^n x = \phi^n(t)x + \psi^n(t)$. By the estimates (3.49) it is easy to prove that $\phi^n(t)$ and $\psi^n(t)$ are convergent and so $T^n$ is actually convergent on the domain $D(r/2, \rho/2)$. Let $T^n \to T_*$ and $T_* x = \phi_*(t)x + \psi_*(t)$. It is easy to see that the estimates for $\phi_*$ and $\psi_*$ in Theorem 2.1 hold.

Using the estimate for $f_n$ and Cauchy's estimate, we have $|f_n(0, t, \varepsilon)| \le \alpha_n r_n \tilde{\varepsilon}_n \to 0$ and $|\partial_x f_n(0, t, \varepsilon)| \le \alpha_n \tilde{\varepsilon}_n \to 0$ as $n \to \infty$. Let $f_n \to f_*$. Then it follows that $f_*(x, t, \varepsilon) = O(x^2)$.

By the estimates (3.50) for $\beta_n$ we have $\beta_n \to \beta_*$. Thus, by the quasi-periodic symplectic transformation $x = T_* y$, the original system is changed to $\dot{y} = A_* y + f_*(y, t, \varepsilon)$ with $A_* = \beta_* J$.

*Estimate of Measure*

Let

$$E_n = \left\{ \varepsilon \in (0, \varepsilon_*) \mid |\langle \omega, k \rangle - 2\beta_n(\varepsilon)| \ge \frac{\alpha_n}{|k|^{\tau'}} \right\}. \tag{3.52}$$

Note that $\beta_n = \beta_1 + \psi$, where $\psi = \sum_{j=1}^{n-1} \beta_{j+1} - \beta_j, \beta_1 = \beta + \hat{\beta}$, and $\hat{\beta} J = [Q]_A$. Note that $\tilde{\varepsilon}_1 = c\tilde{\varepsilon}_0^2/(\alpha_0^2 s_0^{v+v'})$ and $\tilde{\varepsilon}_0 = c\varepsilon^{m_0}/\alpha_0$. By the estimates (3.50), we have $\psi(\varepsilon) = O(\varepsilon^{2m_0})$ and $\varepsilon\psi'(\varepsilon) = O(\varepsilon^{2m_0})$. By assumption, $[Q]_A$ is analytic with respect to $\varepsilon$ and there exists $m_0 \le N \le 2m_0 - 1$ such that $[Q]_A = \delta\varepsilon^N + O(\varepsilon^{N+1})$ with $\delta \ne 0$. Thus, $\beta_1(\varepsilon) = \beta + \delta\varepsilon^N + O(\varepsilon^{N+1})$. By Lemma 3.4, we have $\text{meas}((0, \varepsilon_*) - E_n) \le c(\alpha_n/\alpha_0^2)\varepsilon_*^{N+1}$. Let $E_* = \bigcap_{n \ge 1} E_n$. By $\alpha_n = \alpha_0/2^n$, it follows that $\text{meas}((0, \varepsilon_*) - E_*) \le c\varepsilon_*^{N+1}/\alpha_0$. Thus Lemma 3.5 is proved. $\qquad\square$

## 4. Proof of Theorem 2.1

As we pointed previously, once the non-degeneracy conditions are satisfied in some KAM step, the proof is complete by Lemma 3.5. If the non-degeneracy conditions never happen, the small parameter $\varepsilon$ does not involve into the small divisors and so the systems are analytic in $\varepsilon$. To prepare for KAM iteration, we need a preliminary step to change the original system to a suitable form.

*Preliminary Step*

We first give the preliminary KAM step. Let

$$\dot{x} = Ax + f(x, t, \varepsilon), \tag{4.1}$$

where $A = \beta J$ and $f = J\nabla_x F$. By Lemma 3.3, denote by $\underline{x}$ the solution of $\dot{x} = Ax + f(0, t, \varepsilon)$ on $D_{3\rho/4}$. Under the change of variables $x = T_0 x_+ = \underline{x} + x_+$, the Hamiltonian system (2.1) becomes

$$\dot{x}_+ = Ax_+ + f_1(x_+, t, \varepsilon), \tag{4.2}$$

where $f_1(x_+, t, \varepsilon) = f(\underline{x} + x_+, t, \varepsilon) - f(0, t, \varepsilon)$ satisfying $f_1(0, t, \varepsilon) = O(\varepsilon^2)$ and $\partial_{x_+} f_1(0, t, \varepsilon) = O(\varepsilon)$.

*KAM Step*

The next step is almost the same as the proof of **Lemma 3.5** and even more simple. In the KAM iteration, we only need consider the case that the non-degeneracy condition never happens. In this case, the normal frequency has no shift, which is equivalent to $A_n = A$ for all $n \geq 1$ in the KAM steps in the above nondegenerate case. Moreover, the small divisors conditions are always the initial ones as (2.3) and (2.4) and are independent of the small parameter $\varepsilon$. Thus, we need not delete any parameter. Moreover, the analyticity in $\varepsilon$ remains in the KAM steps, which makes the estimate easier. At the first step, we consider $\dot{x} = Ax + f_1(x, t, \varepsilon)$. In the same way as the case of nondegenerate case, let $r_1 = \varepsilon, \rho_1 = 3\rho/4, \varepsilon_1 = \varepsilon_0, D_1 = D(r_1, \rho_1, \varepsilon_1)$, and $\tilde{\varepsilon}_1 = c\varepsilon/\alpha_0$. Then we have $\|f_1\|_{D_1} \leq \alpha_0 r_1 \tilde{\varepsilon}_1$.

At $n$th step, we consider the Hamiltonian system

$$\dot{x} = Ax + f_n(x, t, \varepsilon), \tag{4.3}$$

where $f_n$ is analytic quasi-periodic with respect to $t$ with frequencies $\omega$ and real analytic with respect to $x$ and $\varepsilon$ on $D_n = D(r_n, \rho_n, \varepsilon_n)$. Moreover, $\|f_n\|_{D_n} \leq \alpha_0 r_n \tilde{\varepsilon}_n$. Suppose

$$Q_n(t, \varepsilon) = \partial_x f_n(0, t, \varepsilon) = O\left(\varepsilon^{2^{n-1}}\right), \quad f_n(0, t, \varepsilon) = O\left(\varepsilon^{2^n}\right). \tag{4.4}$$

Since $Q_n$ is analytic with respect to $\varepsilon$, it follows that

$$Q_n = \sum_{k=2^{n-1}}^{\infty} Q_n^k \varepsilon^k. \tag{4.5}$$

Truncating the above power series of $\varepsilon$, we let

$$R_n(t, \varepsilon) = \sum_{k=2^{n-1}}^{2^n-1} Q_n^k \varepsilon^k, \qquad \tilde{Q}_n = Q_n - R_n. \tag{4.6}$$

Because the non-degeneracy conditions do not happen in KAM steps, we must have $[R_n]_A = 0$. In the same way as the proof of **Lemma 3.5**, we have a quasi-periodic symplectic transformation $T_n$ with $T_n x = \phi_n(t)x + \psi_n(t)$ satisfying (3.49). Let $T^n = T_1 \circ T_2 \cdots \circ T_{n-1}$.

By the transformation $x = T^n y$, the system (4.3) is changed to

$$\dot{y} = Ay + f_{n+1}(y, t, \varepsilon), \tag{4.7}$$

where $f_{n+1} = \tilde{Q}_n \cdot T_n'' + Q_n' \cdot T_n'' + e^{-P_n} \cdot h_n \circ T_n = \tilde{Q}_n(\underline{x}_n + y) + Q_n'(\underline{x}_n + y) + e^{-P_n} h_n(e^{P_n}(\underline{x}_n + y))$.

The last two terms can be estimated similarly as those of (3.41). Note that

$$\tilde{Q}_n = Q_n - R_n = \sum_{k > 2^n} Q_n^k \varepsilon^k \tag{4.8}$$

only consists of the higher order terms of $\varepsilon$. So, in the same way as [8, 10], we use the technique of shriek of the domain interval of $\varepsilon$ to estimate the first term.

Let $r_1 = \varepsilon, \rho_1 = 3\rho/4, \varepsilon_1 = \varepsilon_0$ and $s_1 = \rho/16$.

Define $s_{n+1} = s_n/2, \rho_{n+1} = \rho_n - 2s_n, \eta_n = (1/8)e^{-(4/3)^n}$, $r_{n+1} = \eta_n r_n, \delta_n = 1 - (2/3)^n$ and $\varepsilon_{n+1} = \delta_n \varepsilon_n$. Let $D_{n+1} = D(r_{n+1}, \rho_{n+1}, \varepsilon_{n+1})$.

If $c\tilde{\varepsilon}_n/s_n^{2v} \leq \eta_n < (1/8)$, it follows that

$$\|f_{n+1}\|_{D_{n+1}} \leq \left(\alpha_0 \tilde{\varepsilon}_n e^{-(4/3)^n} + \left(\frac{c\tilde{\varepsilon}_n}{s_n^v}\right)^2\right)\eta_n r_n + c\alpha_0 r_n \tilde{\varepsilon}_n \eta_n^2 \leq \alpha_0 r_{n+1}\tilde{\varepsilon}_{n+1}, \tag{4.9}$$

where $\tilde{\varepsilon}_{n+1} = c\eta_n \tilde{\varepsilon}_n$. Moreover, it is easy to see

$$\partial_x f_{n+1}(0, t, \varepsilon) = O\left(\varepsilon^{2^n}\right), \quad f_{n+1}(0, t, \varepsilon) = O\left(\varepsilon^{2^{n+1}}\right). \tag{4.10}$$

Now we verify $c\tilde{\varepsilon}_n/s_n^{2v} \leq \eta_n < 1/8$. Let $G_n = c\tilde{\varepsilon}_n/s_n^{2v}$. By $G_n = ce^{-(4/3)^{n-1}}16^v G_{n-1}$, it follows that

$$G_n = (c16^v)^{n-1} e^{-[(4/3)^{n-1}+(4/3)^{n-2}+\cdots+(4/3)^1]} G_1 = (c16^v)^{n-1} e^4 e^{-4(4/3)^{n-1}} G_1. \tag{4.11}$$

Note that $G_1 = c\tilde{\varepsilon}_1/s_1^{2v}$. If $\tilde{\varepsilon}_1$ is sufficiently small, we have $c\tilde{\varepsilon}_n/s_n^{2v} = G_n \leq \eta_n$.

Note that $(cr_n\tilde{\varepsilon}_n/s_n^v) \to 0$ and $(c\tilde{\varepsilon}_n/(\eta_n s_n^v)) \to 0$ as $n \to \infty$, and $\tilde{\varepsilon}_n \leq cs_n^{2v}G_n$. Let $\varepsilon_* = \prod_{n\geq 1}(1-(2/3)^n)\varepsilon_0$. Thus, in the same way as before we can prove the convergence of the KAM iteration for all $\varepsilon \in (0, \varepsilon_*)$ and obtain the result of Theorem 2.1. We omit the details.

*Remark 4.1.* As suggested by the referee, we can also introduce an outer parameter to consider the Hamiltonian function $H(x, t, \varepsilon) = \langle \omega, I \rangle + (1/2)(\beta_* + \sigma(\varepsilon))(x_1^2 + x_2^2) + F(x, t, \varepsilon)$, where $(\theta, I)$ are the angle variable and the action variable and $x = (x_1, x_2)$ are a pair of normal variables. In the same way as in [11], $\sigma(\varepsilon)$ is the modified term of the normal frequency. Then by some technique as in [11–13], we can also prove Theorem 2.1.

## Acknowledgments

## References

[1] N. N. Bogoljubov, J. A. Mitropoliski, and A. M. Samoilenko, *Methods of Accelerated Convergence in Nonlinear Mechanics*, Springer, New York, NY, USA, 1976.

[2] R. A. Johnson and G. R. Sell, "Smoothness of spectral subbundles and reducibility of quasiperiodic linear differential systems," *Journal of Differential Equations*, vol. 41, no. 2, pp. 262–288, 1981.

[3] A. Jorba and C. Simó, "On the reducibility of linear differential equations with quasiperiodic coefficients," *Journal of Differential Equations*, vol. 98, no. 1, pp. 111–124, 1992.

[4] L. H. Eliasson, "Floquet solutions for the one-dimensional quasi-periodic Schrödinger equation," *Communications in Mathematical Physics*, vol. 146, no. 3, pp. 447–482, 1992.

[5] L. H. Eliasson, "Almost reducibility of linear quasi-periodic systems," in *Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999)*, vol. 69 of *Proceedings of Symposia in Pure Mathematics*, pp. 679–705, American Mathematical Society, Providence, RI, USA, 2001.

[6] H.-L. Her and J. You, "Full measure reducibility for generic one-parameter family of quasi-periodic linear systems," *Journal of Dynamics and Differential Equations*, vol. 20, no. 4, pp. 831–866, 2008.

[7] A. Jorba and C. Simó, "On quasi-periodic perturbations of elliptic equilibrium points," *SIAM Journal on Mathematical Analysis*, vol. 27, no. 6, pp. 1704–1737, 1996.

[8] X. Wang and J. Xu, "On the reducibility of a class of nonlinear quasi-periodic system with small perturbation parameter near zero equilibrium point," *Nonlinear Analysis*, vol. 69, no. 7, pp. 2318–2329, 2008.

[9] H. Whitney, "Analytic extensions of differentiable functions defined in closed sets," *Transactions of the American Mathematical Society*, vol. 36, no. 1, pp. 63–89, 1934.

[10] L. Zhang and J. Xu, "Persistence of invariant torus in Hamiltonian systems with two-degree of freedom," *Journal of Mathematical Analysis and Applications*, vol. 338, no. 2, pp. 793–802, 2008.

[11] J. Moser, "Convergent series expansions for quasi-periodic motions," *Mathematische Annalen*, vol. 169, pp. 136–176, 1967.

[12] M. B. Sevryuk, "KAM-stable Hamiltonians," *Journal of Dynamical and Control Systems*, vol. 1, no. 3, pp. 351–366, 1995.

[13] K. Soga, "A point-wise criterion for quasi-periodic motions in the KAM theory," *Nonlinear Analysis*, vol. 73, no. 10, pp. 3151–3161, 2010.