A KAM Theorem with Large Twist and Finite Smooth Large Perturbation

Lu Chen

Received: 29 November 2022 / Accepted: 20 April 2023 / Published online: 5 May 2023
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

Abstract

We study non-degenerate Hamiltonian systems of the form

\[ H(\theta, t, I) = \frac{H_0(I)}{e^a} + \frac{P(\theta, t, I)}{e^b}, \]

where \((\theta, t, I) \in T^{d+1} \times [1, 2]^d (T := \mathbb{R}/2\pi \mathbb{Z}), a, b\) are given positive constants with \(a > b\), \(H_0 : [1, 2]^d \to \mathbb{R}\) is real analytic and \(P : T^{d+1} \times [1, 2]^d \to \mathbb{R}\) is \(C^\ell\) with \(\ell = \frac{2(d+1)(5a-b+2ad)}{a-b} + \mu, 0 < \mu \ll 1\). We prove that, for the above Hamiltonian system, if \(\epsilon\) is sufficiently small, there is an invariant torus with given Diophantine frequency vector which obeys conditions (1.7) and (1.8). As for application, a finite network of Duffing oscillators with periodic external forces possesses Lagrange stability for almost all initial data.

Keywords  KAM theorem · Hamiltonian system · Duffing oscillator · Invariant torus · Lagrange stability

Mathematics Subject Classification  34L15 · 34B10 · 47E05

1 Introduction and Main Results

Consider a second order differential equation

\[ \ddot{x} + f(x, t)\dot{x} + g(x, t) = 0, \quad (1.1) \]

This work is supported by the Doctoral Starting Foundation of Quzhou University (No. BSYJ202115).

Lu Chen
chenlu@zju.edu.cn

1 College of Teacher Education, Quzhou University, Quzhou 324000, Zhejiang, People’s Republic of China
where \( f(x, t) = f(x, t + 1) \) and \( g(x, t) = g(x, t + 1) \). The boundedness problem of solutions of Eq. (1.1) has been widely investigated by many authors since the 1940s. In order to show the boundedness of solutions of Eq. (1.1), one began with constructing an absorbing compact domain in the phase space such that all solutions of Eq. (1.1) always enter this domain whenever \( t \geq t_0 \) (see [4, 9, 18]).

When \( f(x, t) \equiv 0 \), Eq. (1.1) is a Hamiltonian system

\[
\dot{x} = \frac{\partial H}{\partial y}(x, y, t), \quad \dot{y} = -\frac{\partial H}{\partial x}(x, y, t),
\]

with the Hamiltonian function

\[
H(x, y, t) = \frac{1}{2} y^2 + G(x, t),
\]

where \( G(x, t) = \int_0^x g(z, t)dz \). Then one cannot construct the above-mentioned absorbing compact domain in the phase space in order to show the boundedness of the solutions. One could try to apply KAM theory in order to find conditions which guarantee the boundedness of all the solutions. In 1976, Morris [13] obtained the first boundedness result by Moser’s twist theorem [14, 15], he showed that all solutions of the equation

\[
\ddot{x} + 2x^3 = p(t), \quad p(t) \in C^0
\]

are bounded. Subsequently, Morris’s boundedness result was, by Dieckerhoff–Zehnder [3] in 1987, extended to a wider class of systems

\[
\ddot{x} + x^{2n+1} + \sum_{i=0}^{2n} x^i p_i(t) = 0, \quad n \geq 1,
\]

(1.2)

where \( p_i(t) \in C^\nu \) \((i = 0, 1, \ldots, 2n)\) are 1-periodic functions, and

\[
\nu > 1 + \frac{4}{n} + [\log_2 n] \rightarrow \infty \quad \text{as} \quad n \rightarrow \infty.
\]

they [3] remarked that: “It is not clear whether the boundedness phenomenon is related to the smoothness in the \( t \)-variable or whether this requirement is a shortcoming of our proof.”

In 1989 and 1992, Liu [11, 12] proved the boundedness of solutions for

\[
\ddot{x} + x^{2n+1} + a(t)x = p(t), \quad n \geq 1,
\]

where \( a(t), p(t) \in C^0 \) are 1-periodic functions. In 1991, Laederich–Levi [8] relaxed the smoothness requirement of \( p_i(t) \) \((i = 0, 1, \ldots, 2n)\) for (1.2) to \( p_i(t) \in C^{5+\varepsilon} \), \( \varepsilon > 0 \). In 1998 and 2000, Yuan [25, 26] relaxed the smoothness requirement of
\[ \pi(t) = 0, 1, \ldots, 2n \) for (1.2) to \( p_i(t) \in C^2 \). One may see papers [5, 16, 24, 27] for more details.

In many research fields such as physics, mechanics and mathematical biology and so on arise networks of coupled Duffing oscillators of various forms (see [1, 6, 7, 22, 23] for details). What about the boundedness of the solutions of these coupled Duffing oscillators? Recently, Yuan et al. [28] studied the Lagrange stability for coupled Hamiltonian systems of \( m \) Duffing oscillators,

\[
\ddot{x}_i + x_i^{2n+1} + \frac{\partial}{\partial x_i} F(x,t) = 0, \quad i = 1, 2, \ldots, m, \tag{1.3}
\]

where the polynomial potential \( F = F(x,t) = \sum_{\alpha \in \mathbb{N}^m, |\alpha| \leq 2n+1} p_{\alpha}(t) x^\alpha, \ x \in \mathbb{R}^m \) with \( p_{\alpha}(t) \) are of period \( 2\pi \), and \( n \) is a given natural number. They proved Eq. (1.3) is almost Lagrange stable if \( p_{\alpha}(t) \) are real analytic.

KAM theory for finite smooth Hamiltonian systems is of independent interest and application, considering that there are many physical, mechanical, and mathematical biological models which are finite smooth. In the present paper, we study the Lagrange stability of Eq. (1.3) with \( p_{\alpha}(t) \in C^\ell \ (\ell = 2(m + 1)(4n + 2nm + 1) + \mu \) with \( 0 < \mu \ll 1 \).

We denote different positive constants by \( C \) in different positions. Then, we give the following KAM theorem.

**Theorem 1.1** Consider a Hamiltonian

\[
H(\theta, t, I) = \frac{H_0(I)}{e^a} + \frac{P(\theta, t, I)}{e^b}, \tag{1.4}
\]

where \( a, b \) are given positive constants with \( a > b \), and \( H_0 \) and \( P \) obey the following conditions:

1. \( H_0 : [1, 2]^d \to \mathbb{R} \) is real analytic and \( P : \mathbb{T}^{d+1} \times [1, 2]^d \to \mathbb{R} \) is \( C^\ell \), and

\[
||H_0|| := \sup_{I \in [1,2]^d} |H_0(I)| \leq c_1, \ |P|_{C^\ell(\mathbb{T}^{d+1} \times [1,2]^d)} \leq c_2, \tag{1.5}
\]

where \( \ell = \frac{2(d+1)(5a-b+2ad)}{a-b} + \mu \) with \( 0 < \mu \ll 1 \);

2. \( H_0 \) is non-degenerate in Kolmogorov’s sense:

\[
\det \left( \frac{\partial^2 H_0(I)}{\partial I^2} \right) \geq c_3 > 0, \ \forall \ I \in [1, 2]^d. \tag{1.6}
\]

Then there exists \( 0 < \epsilon^* \ll 1 \) such that for any \( \epsilon \) with \( 0 < \epsilon < \epsilon^* \), the Hamiltonian system

\[
\dot{\theta} = \frac{\partial H(\theta, t, I)}{\partial I}, \quad \dot{I} = -\frac{\partial H(\theta, t, I)}{\partial \theta}
\]
possesses a \((d+1)\)-dimensional invariant torus of rotational frequency vector \(\left(\frac{\alpha(I_0)}{a}, 2\pi\right)\) with \(\omega(I) := \frac{\partial H_0(I)}{\partial I}\), for any \(I_0 \in [1, 2]^d\) and \(\omega(I_0)\) obeying Diophantine conditions (we let \(B = 5a - b + 2ad\)):

(i)

\[
\left|\frac{k}{\varepsilon^a} + l\right| \geq \frac{\varepsilon^{-a + \frac{B}{2}}}{|k|^\tau_1} > \frac{\varepsilon^a}{|k|^\tau_2}, \quad k \in \mathbb{Z}^d \setminus \{0\}, \quad l \in \mathbb{Z}, \quad |k| + |l|
\]

\[\leq \varepsilon^{-\frac{B}{2}} \left(\log \frac{1}{\varepsilon}\right)^2 , \tag{1.7}\]

(ii)

\[
\left|\frac{k}{\varepsilon^a} + l\right| \geq \frac{\varepsilon^2}{|k|^\tau_1}, \quad k \in \mathbb{Z}^d \setminus \{0\},
\]

\[l \in \mathbb{Z}, \quad |k| + |l| > \varepsilon^{-\frac{B}{2}} \left(\log \frac{1}{\varepsilon}\right)^2 , \tag{1.8}\]

where \(\gamma = (\log \frac{1}{\varepsilon})^{-4}, \quad \tau_1 = d - 1 + \frac{(a-b)^2\mu}{1000(a+b+1)(d+3)(5a-b+2ad)}, \quad \tau_2 = d + \frac{(a-b)^2\mu}{1000(a+b+1)(d+3)(5a-b+2ad)}\).

**Remark 1.2** Theorem 1.1 is trivial if the perturbation \(P\) is independent of time \(t\). However, Theorem 1.1 is not at all trivial when \(P\) involves time \(t\). This can be seen in the following way. By time rescaling \(t = \varepsilon^a \tau\), and introducing new angle-action variable \((\phi, J) = (\varepsilon^a \tau, J)\), we get an autonomous Hamiltonian

\[
H_{\text{new}} = \tilde{H}_0(J, I) + \varepsilon^{a-b} P(\theta, \phi, I), \text{ where } \tilde{H}_0(J, I)
\]

\[= \varepsilon^a J + H_0(I) \text{ and new time } = \tau.
\]

It seems that the invariant tori could be obtained by a direct application of Moser’s classical twist theorem. Unfortunately, the unperturbed Hamiltonian \(\tilde{H}_0(J, I)\) does not satisfy the twist condition of Moser’s theorem (or Kolmogorov’s non-degenerate condition). In fact, the small divisors for \(H_{\text{new}}\) are

\[(*) := \left(\frac{k}{\varepsilon^a} + l\right), \quad (k, l) \in \mathbb{Z}^d \times \mathbb{Z} \setminus \{0, 0\}.
\]

At this time, the divisor \((*)\) is too small for Newton iteration to work in the proof of the KAM theory. For example, taking \(k = 0, l = 1\), we have

\[(*) = \varepsilon^a \ll \text{size of the perturbation } \varepsilon^{a-b} P(\theta, \phi, I).
\]

This implies that the solution of the homological equation is so large that the symplectic transformation is not well defined in Newton iteration.
Remark 1.3 Let $\Theta = \{I_0 \in [1, 2]^d : \omega(I_0) \text{ obeys the Diophantine conditions}\}$. We claim that the Lebesgue measure of $\Theta$ approaches to 1:

$$\text{Leb}\Theta \geq 1 - C \left( \log \frac{1}{\varepsilon} \right)^{-2} \to 1, \text{ as } \varepsilon \to 0.$$ 

Let

$$\tilde{\Theta}_{k,l} = \left\{ \xi \in \omega([1, 2]^d) : [\ast] \leq \frac{\varepsilon^B}{\gamma |k|^\tau_2}, |k| + |l| \leq \left( \frac{\log \frac{1}{\varepsilon}}{\varepsilon^B} \right)^2 \right\},$$

where $[\ast] := |\langle k, \xi \rangle + l|, k \in Z^d \setminus \{0\}, l \in Z$.

Let $f(\xi) = \frac{\varepsilon^B}{\gamma |a|} + l$. Since $k \neq 0$, there exists an unit vector $v \in Z^d$ such that

$$\frac{df(\xi)}{dv} \geq \frac{C|k|}{\varepsilon^a}. \quad (1.9)$$

Then, if, $k \in Z^d \setminus \{0\}, l \in Z, |k| + |l| \leq \varepsilon^{-B} (\log \frac{1}{\varepsilon})^2$, by (1.9), we have

$$\text{Leb}\tilde{\Theta}_{k,l} \leq C \frac{\gamma \cdot \varepsilon^B}{|k|^\tau_1 + 1}. \quad (1.10)$$

Thus,

$$\text{Leb} \left( \bigcup_{k \in Z^d \setminus \{0\}, l \in Z, |k| + |l| \leq \varepsilon^{-B} (\log \frac{1}{\varepsilon})^2} \tilde{\Theta}_{k,l} \right) \leq \sum_{|l| \leq \varepsilon^{-B} (\log \frac{1}{\varepsilon})^2} C \gamma \cdot \varepsilon^B \leq C \left( \log \frac{1}{\varepsilon} \right)^{-2}. \quad (1.11)$$

If $k \in Z^d \setminus \{0\}$, $l \in Z$, $|k| + |l| > \varepsilon^{-B} (\log \frac{1}{\varepsilon})^2$, we can let $c_5 = \max\{||\omega([1, 2]^d)||\} := \max\{\sum_{i=1}^d |\omega_i([1, 2]^d)|\}$. Noting that $|\langle k, \xi \rangle| \leq c_5|k|$. Thus if $|l| > \frac{c_5|k|}{\varepsilon^a} + 1$, then

$$\left| \frac{\langle k, \xi \rangle}{\varepsilon^a} + l \right| \geq |l| - \left| \frac{\langle k, \xi \rangle}{\varepsilon^a} \right| > \frac{c_5|k|}{\varepsilon^a} + 1 - \frac{c_5|k|}{\varepsilon^a} \geq 1 > \frac{\gamma}{|k|^\tau_2}.$$ 

It follows that $\tilde{\Theta}_{k,l} = \phi$. Now we assume $|l| \leq \frac{c_5|k|}{\varepsilon^a} + 1$, then by (1.9), we have

$$\text{Leb}\tilde{\Theta}_{k,l} \leq \frac{C \gamma \varepsilon^a}{|k|^\tau_1 + 1}. \quad (1.12)$$
Thus,
\[
\text{Leb} \left( \bigcup_{k \in \mathbb{Z}^d \setminus \{0\}, l \in \mathbb{Z}, |k| + |l| > \varepsilon - B} \Theta_{k,l} \right) \leq \sum_{k \neq 0} \sum_{|l| \leq \left\lfloor \frac{|k|}{\varepsilon} \right\rfloor + 1} \frac{C \gamma \varepsilon^d}{|k|^{\tau_2 + 1}} \leq \sum_{k \neq 0} \frac{C \gamma}{|k|^{\tau_2}} \leq C \left( \log \frac{1}{\varepsilon} \right)^{-4}.
\]
(1.13)

Let \( \Theta = [1, 2]^d \setminus \left( \bigcup_{k \in \mathbb{Z}^d \setminus \{0\}, l \in \mathbb{Z}} \omega^{-1}(\Theta_{k,l}) \right) \). By Kolmogorov’s non-degenerate condition, the map \( \omega : [1, 2]^d \to \omega([1, 2]^d) \) is a diffeomorphism in both direction. Then by (1.11) and (1.13), the proof of the claim is completed by letting \( \Theta = [1, 2]^d \setminus \left( \bigcup_{k \in \mathbb{Z}^d \setminus \{0\}, l \in \mathbb{Z}} \omega^{-1}(\Theta_{k,l}) \right) \).

Applying Theorem 1.1 to Eq. (1.3) we have the following theorem.

**Theorem 1.4** If \( p_\alpha(t) \in C^\ell \), where \( \ell = 2(m + 1)(4n + 2nm + 1) + \mu \) with \( 0 < \mu \ll 1 \), Eq. (1.3) is almost Lagrange stable, i.e., for almost all initial data \( (x_1(0), \dot{x}_1(0); \ldots; x_m(0), \dot{x}_m(0)) \), the solution \( (x_1(t), \dot{x}_1(t); \ldots; x_m(t), \dot{x}_m(t)) \) of Eq. (1.3) exists globally and obeys
\[
\sup_{t \in \mathbb{R}} \sum_{i=1}^m (|x_i(t)| + |\dot{x}_i(t)|) < C,
\]
where \( C \) depends on the initial data \( (x_1(0), \dot{x}_1(0); \ldots; x_m(0), \dot{x}_m(0)) \).

The paper is outlined as follows. In Sect. 2, we introduce an approximation lemma. In Sect. 3, first, we use the approximation lemma in Sect. 2 to analytize the perturbation \( P \), then, we apply a finite number of symplectic changes of coordinates by reducing the Hamiltonian to another with the integrable part plus another that depends only on the action variables and the time, and the last part that is much smaller. In Sect. 4, with a symplectic transformation that is not close to the identity, we reduce the Hamiltonian to an integrable part plus an error term. In Sect. 5, we prove an iterative lemma by a classical way. In Sect. 6, we prove Theorem 1.1 by the iterative lemma. Applying Theorem 1.1 to Eq. (1.3), we prove Theorem 1.4.

**2 Approximation Lemma**

First we denote by \( |\cdot| \) the norm of any finite dimensional Euclidean space. Let \( C^{\tilde{\mu}}(\mathbb{R}^m) \) for \( 0 < \tilde{\mu} < 1 \) denote the space of bounded Hölder continuous functions.
$f : \mathbb{R}^m \to \mathbb{R}^n$ with the norm

$$|f|_{C^{\tilde{\mu}}} = \sup_{0 < |x - y| < 1} \frac{|f(x) - f(y)|}{|x - y|^{\tilde{\mu}}} + \sup_{x \in \mathbb{R}^n} |f(x)|.$$  

If $\tilde{\mu} = 0$ the $|f|_{C^{\tilde{\mu}}}$ denotes the sup-norm. For $\tilde{\ell} = k + \tilde{\mu}$ with $k \in \mathbb{N}$ and $0 \leq \tilde{\mu} < 1$ we denote by $C^{\tilde{\ell}}(\mathbb{R}^m)$ the space of functions $f : \mathbb{R}^m \to \mathbb{R}^n$ with Hölder continuous partial derivatives $\partial^\alpha f \in C^{\tilde{\mu}}(\mathbb{R}^m)$ for all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}^m$ with the assumption that $|\alpha| := |\alpha_1| + \cdots + |\alpha_m| \leq k$. We define the norm

$$|f|_{C^{\tilde{\ell}}} := \sum_{|\alpha| \leq \tilde{\ell}} |\partial^\alpha f|_{C^{\tilde{\mu}}}$$

for $\tilde{\mu} = \tilde{\ell} - [\tilde{\ell}] < 1$. In order to give an approximate lemma, we define the kernel function

$$K(x) = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} \hat{K}(\xi) e^{i\langle x, \xi \rangle} d\xi, \ x \in \mathbb{C}^m,$$

where $\hat{K}(\xi)$ is a $C^\infty$ function with compact support, contained in the ball $|\xi| \leq a_1$ with a constant $a_1 > 0$, that satisfies

$$\partial^\alpha \hat{K}(0) = \begin{cases} 1, & \alpha = 0, \\ 0, & \alpha \neq 0. \end{cases}$$

Then $K : \mathbb{C}^m \to \mathbb{R}^n$ is a real analytic function with the property that for every $j > 0$ and every $p > 0$, there exists a constant $c = c(j, p) > 0$ such that for all $\beta \in \mathbb{N}^m$ with $|\beta| \leq j$,

$$|\partial^\beta K(x + iy)| \leq c(1 + |x|)^{-p} e^{|a_1| y}, \ x, y \in \mathbb{R}^m. \quad (2.1)$$

**Lemma 2.1** (Jackson–Moser–Zehnder) There is a family of convolution operators

$$(S_s F)(x) = s^{-m} \int_{\mathbb{R}^m} K(s^{-1}(x - y)) F(y) dy, \ 0 < s \leq 1, \ \forall F \in C^0(\mathbb{R}^m)$$

from $C^0(\mathbb{R}^m)$ into the linear space of entire functions on $\mathbb{C}^m$ such that for every $\tilde{\ell} > 0$ there exists a constant $c = c(\tilde{\ell}) > 0$ with the following properties: if $F \in C^{\tilde{\ell}}(\mathbb{R}^m)$, then for $|\alpha| \leq \tilde{\ell}$ and $|\text{Im } x| \leq s$,

$$|\partial^\alpha (S_s F)(x) - \sum_{|\beta| \leq \tilde{\ell} - |\alpha|} \partial^{\alpha + \beta} F(\text{Re } x)(i \text{ Im } x)^{\beta}/\beta!| \leq c|F|_{C^{\tilde{\ell}}(\mathbb{R}^m)} (\tilde{\ell} - |\alpha|)} \quad (2.3)$$
and in particular for $\rho \leq s$

$$|\partial^{\alpha} S_{s} F - \partial^{\alpha} S_{\rho} F|_{\rho} := \sup_{|\text{Im} x| \leq \rho} |\partial^{\alpha} (S_{s} F)(x) - \partial^{\alpha} (S_{\rho} F)(x)| \leq c |F|_{C^{\bar{\ell}} s^{\bar{\ell} - |\alpha|}}. \quad (2.4)$$

Moreover, in the real case

$$|S_{s} F - F|_{C^{p}} \leq c |F|_{C^{\bar{\ell}} s^{\bar{\ell} - p}}, \quad p \leq \bar{\ell}, \quad (2.5)$$

$$|S_{s} F|_{C^{p}} \leq c |F|_{C^{\bar{\ell}} s^{\bar{\ell} - p}}, \quad p \leq \bar{\ell}. \quad (2.6)$$

Finally, if $F$ is periodic in some variables then so are the approximating functions $S_{s} F$ in the same variables.

**Remark 2.2** Moreover we point out that from (2.6) one can easily deduce the following well-known convexity estimates

$$|f|_{C^{l-k}} \leq c |f|_{C^{k}} |f|_{C^{q-k}}, \quad k \leq q \leq l, \quad (2.7)$$

$$|f \cdot g|_{C^{s}} \leq c (|f|_{C^{s}} |f|_{C^{0}} + |f|_{C^{0}} |g|_{C^{s}}), \quad s \geq 0. \quad (2.8)$$

See [21, 29] for the proofs of Lemma 2.1 and the inequalities (2.7) and (2.8).

**Remark 2.3** From the definition of the operator $S_{s}$, we clearly have

$$\sup_{x, y \in \mathbb{R}^m, |y| \leq s} |S_{s} F(x + iy)| \leq C |F|_{C^{0}}. \quad (2.9)$$

In fact, for any $x, y \in \mathbb{R}^m$ with $|y| \leq s$, we have that

$$|S_{s} F(x + iy)| = |s^{-m} \int_{\mathbb{R}^m} K(s^{-1}(x + iy - z)) F(z) dz|$$

$$= | \int_{\mathbb{R}^m} K(is^{-1}y + \xi) F(x - s\xi) d\xi|$$

$$\leq |F|_{C^{0}} \int_{\mathbb{R}^m} |K(is^{-1}y + \xi)| d\xi$$

$$\leq C |F|_{C^{0}}, \quad (2.10)$$

where we used (2.1) in the last inequality.

Consider a function $F(\theta, t, I)$, where $F: T^{d+1} \times [1, 2]^d \to \mathbb{R}$, and $F$ satisfies

$$|F|_{C^{\bar{\ell}}(T^{d+1} \times [1, 2]^d)} \leq C.$$  

By Whitney’s extension theorem, we can find a function $\tilde{F}: T^{d+1} \times \mathbb{R}^d \to \mathbb{R}$ such that $\tilde{F}|_{T^{d+1} \times [1, 2]^d} = F$ (i.e. $\tilde{F}$ is the extension of $F$) and

$$|\tilde{F}|_{C^{|\alpha|}(T^{d+1} \times \mathbb{R}^d)} \leq C_{\alpha} |F|_{C^{|\alpha|}(T^{d+1} \times [1, 2]^d)}, \quad \forall \alpha \in \mathbb{Z}^{2d+1}_{+}, |\alpha| \leq \bar{\ell},$$
where $C_\alpha$ is a constant which depends only $\tilde{\ell}$ and $d$.

Let $z = (\theta, t, I)$ for brevity, for $\forall s > 0$,

\[
(S_s \tilde{F})(z) = s^{-(2d+1)} \int_{T^{d+1} \times \mathbb{R}^d} K(s^{-1}(z - \tilde{z})) \tilde{F}(\tilde{z}) d\tilde{z}.
\]

For any positive integer $p$, let $T^p_s = \{ x \in C^p / (2\pi \mathbb{Z})^p : |\text{Im} x| \leq s \}$, $R^p_s = \{ x \in C^p : |\text{Im} x| \leq s \}$. Fix a sequence of fast decreasing numbers $s_\nu \downarrow 0$, $\nu \in \mathbb{Z}^+ \cup \{0\}$.

Let

\[
F(\nu)(z) = (S_{2s_\nu} \tilde{F})(z), \quad \nu \geq 0.
\]

Then $F(\nu)$'s ($\nu \geq 0$) are entire functions in $C^{2d+1}$, in particular, which obey the following properties:

1. $F(\nu)$'s ($\nu \geq 0$) are real analytic on the complex domain $T^{d+1}_{2s_\nu} \times R^d_{2s_\nu}$;
2. The sequence of functions $F(\nu)$'s satisfies the estimates

\[
\sup_{z \in T^{d+1}_{2s_\nu} \times R^d_{2s_\nu}} |F^{(\nu)}(z) - \tilde{F}(z)| \leq C |F|_{C^{i}(T^{d+1}_{2s_\nu} \times [1,2]^d)} S_{\tilde{\ell}}, \quad (2.11)
\]

\[
\sup_{z \in T^{d+1}_{2s_{\nu+1}} \times R^d_{2s_{\nu+1}}} |F^{(\nu+1)}(z) - F^{(\nu)}(z)| \leq C |F|_{C^{i}(T^{d+1}_{2s_{\nu+1}} \times [1,2]^d)} S_{\tilde{\ell}}, \quad (2.12)
\]

where constants $C = C(d, \tilde{\ell})$ depend on only $d$ and $\tilde{\ell}$;
3. The first approximate $F^{(0)}(z) = (S_{2s_0} \tilde{F})(z)$ is “small” with respect to $F$. Precisely,

\[
|F^{(0)}(z)| \leq C |F|_{C^{i}(T^{d+1}_{2s_0} \times [1,2]^d)}, \quad \forall z \in T^{d+1}_{2s_0} \times R^d_{2s_0}, \quad (2.13)
\]

where the constant $C = C(d, \tilde{\ell})$ is independent of $s_0$;
4. From Lemma 2.1, we have that

\[
F(z) = F^{(0)}(z) + \sum_{\nu=0}^{\infty} (F^{(\nu+1)}(z) - F^{(\nu)}(z)), \quad z \in T^{d+1} \times [1,2]^d. \quad (2.14)
\]

Let

\[
F_0(z) = F^{(0)}(z), \quad F_{\nu+1}(z) = F^{(\nu+1)}(z) - F^{(\nu)}(z). \quad (2.15)
\]

Then

\[
F(z) = \sum_{\nu=0}^{\infty} F_{\nu}(z), \quad z \in T^{d+1} \times [1,2]^d. \quad (2.16)
\]
3 Normal Form

Let \( I_0 \in [1, 2]^d \) such that \( \omega(I_0) = \frac{\partial H_0}{\partial T}(I_0) \) obeys Diophantine conditions (1.7) and (1.8). Let

\[
\mu_1 = \frac{(a-b)^2 \mu}{1000(a+b+1)(d+3)(5a-b+2ad)}, \quad \mu_2 := 2\mu_1, \quad m_0 = 10 + \left[ E \right],
\]

where \( E = \max\{\frac{4B}{a-b-2(\tau_1+2\tau_2)\mu-2\mu_1}, \frac{2(\tau_1+3)(\tau_2+1)B}{B-2a-2(\tau_2+1)b-2(\tau_1+3)(\tau_2+1)B-8\mu_1(\tau_2+1)-2\mu_2}\} \), and \( a, b, \tau_1, \tau_2, B, \ell \) are the same as those in Theorem 1.1, and \( \left[ \cdot \right] \) is the integer part of a positive number. Define sequences

- \( \varepsilon_j = \varepsilon^{m_0}_{-1}, \ j = 0, 1, 2, \ldots, m_0, \ \varepsilon_j = \varepsilon_{j+1}^{1+\mu_3} \) with \( \mu_3 = \frac{(a-b)\mu}{10B}, \ j = m_0 + 1, m_0 + 2, \ldots; \)
- \( s_j = \varepsilon_j^{-\frac{1}{\mu}}, \ s^{(l)}_j = s_j - \frac{l}{10}(s_j - s_{j+1}), \ l = 0, 1, \ldots, 10, \ j = 0, 1, 2, \ldots; \)
- \( r_j = \varepsilon^{m_0}_{-\varepsilon_j+1}, \ j = 0, 1, 2, \ldots, m_0, \ r_j = r_{j-1}^{1+\mu_3}, \ j = m_0 + 1, m_0 + 2, \ldots; \)
- \( r^{(l)}_j = r_j - \frac{l}{10}(r_j - r_{j+1}), \ l = 0, 1, \ldots, 10, \ j = 0, 1, 2, \ldots; \)
- \( K_j = \frac{2B}{s_j} \log \frac{1}{\varepsilon_j}, \ j = 0, 1, 2, \ldots; \)
- \( B(r_j) = \{z \in \mathbb{C}^d : |z - I_0| \leq r_j\}, \ j = 0, 1, 2, \ldots\)

With the preparation of Sect. 2, we can rewrite the Hamiltonian (1.4) as follows:

\[
H(\theta, t, I) = \frac{H_0(I)}{\varepsilon^a} + \frac{1}{\varepsilon^b} \sum_{v=0}^{\infty} P_v(\theta, t, I), \tag{3.1}
\]

where

\[
P_v : T^{d+1}_{2s_v} \times \mathbb{R}^d_{2s_v} \to \mathbb{C}, \tag{3.2}
\]

is real analytic, and

\[
\sup_{(\theta, t, I) \in T^{d+1}_{2s_v} \times \mathbb{R}^d_{2s_v}} |P_v| \leq C\varepsilon_v. \tag{3.3}
\]

Let

\[
h^{(0)}(t, I) \equiv 0, \ P^{(0)} = P_0. \tag{3.4}
\]

Then we can rewrite the Hamiltonian (3.1) as follows:

\[
H^{(0)}(\theta, t, I) = \frac{H_0(I)}{\varepsilon^a} + \frac{h^{(0)}(t, I)}{\varepsilon^b} + \frac{\varepsilon_0 P^{(0)}(\theta, t, I)}{\varepsilon^b} + \sum_{v=1}^{\infty} \frac{P_v(\theta, t, I)}{\varepsilon^b}. \tag{3.5}
\]
Define
\[ D(s, r) = T^{d+1}_s \times B(r), \quad D(s, 0) = T^{d+1}_s, \quad D(0, r) = B(r). \]

For a function \( f \) defined in \( D(s, r) \), define
\[ \| f \|_{D(s, r)} = \sup_{(\theta, t, I) \in D(s, r)} |f(\theta, t, I)|. \]

Similarly, we can define \( \| f \|_{D(0, r)} \) and \( \| f \|_{D(s, 0)} \).

Clearly, (3.5) fulfills (3.9)–(3.11) with \( m = 0 \). Then we have the following lemma.

**Lemma 3.1** Suppose that we have had \( m + 1 \) \((m = 0, 1, 2, \ldots, m_0 - 1)\) symplectic transformations \( \Phi_0 = id, \Phi_1, \ldots, \Phi_m \) with
\[ \Phi_j : D(s_j, r_j) \rightarrow D(s_{j-1}, r_{j-1}), \quad j = 1, 2, \ldots, m \]
and
\[ \| \partial(\Phi_j - id) \|_{D(s_j, r_j)} \leq \frac{1}{2^{j+1}}, \quad j = 1, 2, \ldots, m \]
such that system (3.5) is changed by \( \Phi^{(m)} = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_m \) into
\[ H^{(m)} = H^{(0)} \circ \Phi^{(m)} = \frac{H_0(I)}{\varepsilon^a} + \frac{h^{(m)}(t, I)}{\varepsilon^b} + \frac{\varepsilon_m P^{(m)}(\theta, t, I)}{\varepsilon^b} + \sum_{\nu = m + 1}^{\infty} \frac{P_{\nu} \circ \Phi^{(m)}(\theta, t, I)}{\varepsilon^b}. \]

where
\[ \| h^{(m)}(t, I) \|_{D(s_m, r_m)} \leq C, \]  
\[ \| P^{(m)}(\theta, t, I) \|_{D(s_m, r_m)} \leq C, \]  
\[ \| P_{\nu} \circ \Phi^{(m)}(\theta, t, I) \|_{D(s_{\nu}, r_{\nu})} \leq C\varepsilon_{\nu}, \quad \nu = m + 1, m + 2, \ldots \]

Then there is a symplectic transformation \( \Phi_{m+1} \) with
\[ \Phi_{m+1} : D(s_{m+1}, r_{m+1}) \rightarrow D(s_m, r_m) \]
and
\[ \| \partial(\Phi_{m+1} - id) \|_{D(s_{m+1}, r_{m+1})} \leq \frac{1}{2^{m+2}}. \]
such that system (3.8) is changed by \( \Phi_{m+1} \) into \((\Phi^{(m+1)} = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{m+1}) \)

\[
H^{(m+1)} = H^{(m)} \circ \Phi_{m+1} = H^{(0)} \circ \Phi^{(m+1)}
= \frac{H_0(I)}{\varepsilon^a} + \frac{h^{(m+1)}(t, I)}{\varepsilon^b} + \frac{\varepsilon_{m+1} P^{(m+1)}(\theta, t, I)}{\varepsilon^b} + \sum_{v=m+2}^{\infty} \frac{P_v \circ \Phi^{(m+1)}(\theta, t, I)}{\varepsilon^b},
\]

where \( H^{(m+1)} \) satisfies (3.9)–(3.11) by replacing \( m \) by \( m + 1 \).

**Proof** Assume that the change \( \Phi_{m+1} \) is implicitly defined by

\[
\Phi_{m+1} : \begin{cases} 
I = \rho + \frac{\partial S}{\partial \theta}, \\
\phi = \theta + \frac{\partial S}{\partial \rho}, \\
t = t,
\end{cases} \tag{3.12}
\]

where \( S = S(\theta, t, \rho) \) is the generating function, which will be proved to be analytic in a smaller domain \( D(s_{m+1}, r_{m+1}) \). By a simple computation, we have

\[
dI \wedge d\theta = d\rho \wedge d\theta + \sum_{i, j=1}^{d} \frac{\partial^2 S}{\partial \rho_i \partial \theta_j} d\rho_i \wedge d\theta_j = d\rho \wedge d\phi.
\]

Thus, the coordinates change \( \Phi_{m+1} \) is symplectic if it exists. Moreover, we get the changed Hamiltonian

\[
H^{(m+1)} = \frac{H_0(\rho + \frac{\partial S}{\partial \theta})}{\varepsilon^a} + \frac{h^{(m)}(t, \rho + \frac{\partial S}{\partial \theta})}{\varepsilon^b} + \frac{\varepsilon_{m} P^{(m)}(\theta, t, \rho + \frac{\partial S}{\partial \theta})}{\varepsilon^b} + \frac{\partial S}{\partial t} + \sum_{v=m+2}^{\infty} \frac{P_v \circ \Phi^{(m+1)}(\theta, t, \rho)}{\varepsilon^b}, \tag{3.13}
\]

where \( \theta = \theta(\phi, t, \rho) \) is implicitly defined by (3.12). By Taylor formula, we have

\[
H^{(m+1)} = \frac{H_0(\rho)}{\varepsilon^a} + \frac{h^{(m)}(t, \rho)}{\varepsilon^b} + \left( \frac{\omega(\rho)}{\varepsilon^a} \right) + \frac{\partial S}{\partial t} + \frac{\varepsilon_{m} P^{(m)}(\theta, t, \rho)}{\varepsilon^b} + \sum_{v=m+2}^{\infty} \frac{P_v \circ \Phi^{(m+1)}(\theta, t, \rho)}{\varepsilon^b} + R_1, \tag{3.14}
\]
where \( \omega(\rho) = \frac{\partial H_0}{\partial T} (\rho) \) and

\[
R_1 = \frac{1}{\varepsilon^a} \int_0^1 (1 - \tau) \frac{\partial^2 H_0}{\partial I^2} (\rho + \tau \frac{\partial S}{\partial \theta}) \left( \frac{\partial S}{\partial \theta} \right)^2 d\tau + \frac{\varepsilon^m}{\varepsilon^b} \int_0^1 \frac{\partial P^{(m)}}{\partial I} (\theta, t, \rho + \tau \frac{\partial S}{\partial \theta}) \frac{\partial S}{\partial \theta} d\tau + \frac{1}{\varepsilon^b} \int_0^1 \frac{\partial h^{(m)}}{\partial I} (t, \rho + \tau \frac{\partial S}{\partial \theta}) \frac{\partial S}{\partial \theta} d\tau.
\]

(3.15)

Expanding \( P^{(m)}(\theta, t, \rho) \) into a Fourier series,

\[
P^{(m)}(\theta, t, \rho) = \sum_{(k,l) \in \mathbb{Z}^d \times \mathbb{Z}} \hat{P}^{(m)}(k, l, \rho) e^{i\langle k, \theta \rangle + lt} := P^{(m)}_1(\theta, t, \rho) + P^{(m)}_2(\theta, t, \rho),
\]

(3.16)

where \( P^{(m)}_1 = \sum_{|k| + |l| \leq K_m} \hat{P}^{(m)}(k, l, \rho) e^{i\langle k, \theta \rangle + lt}, \) \( P^{(m)}_2 = \sum_{|k| + |l| > K_m} \hat{P}^{(m)}(k, l, \rho) e^{i\langle k, \theta \rangle + lt} \). Then, we derive the homological equation:

\[
\frac{\partial S}{\partial t} + \left( \frac{\omega(\rho)}{\varepsilon^a}, \frac{\partial S}{\partial \theta} \right) + \frac{\varepsilon^m P^{(m)}_1(\theta, t, \rho)}{\varepsilon^b} - \frac{\varepsilon^m P^{(m)}_1(0, t, \rho)}{\varepsilon^b} = 0,
\]

(3.17)

where \( \hat{P}^{(m)}_1(0, t, \rho) \) is the 0-Fourier coefficient of \( P^{(m)}_1(\theta, t, \rho) \) as a function of \( \theta \). Let

\[
S(\theta, t, \rho) = \sum_{|k| + |l| \leq K_m, k \neq 0} \hat{S}(k, l, \rho) e^{i\langle k, \theta \rangle + lt}.
\]

(3.18)

By passing to Fourier coefficients, we have

\[
\hat{S}(k, l, \rho) = \frac{\varepsilon^m}{\varepsilon^b} \frac{i}{\varepsilon^{-a} \langle k, \omega(\rho) \rangle + l} \hat{P}^{(m)}(k, l, \rho), \quad |k| + |l| \leq K_m, k \in \mathbb{Z}^d \setminus \{0\}, l \in \mathbb{Z}.
\]

(3.19)

Then we can solve homological Eq. (3.17) by setting

\[
S(\theta, t, \rho) = \sum_{|k| + |l| \leq K_m, k \neq 0} \frac{\varepsilon^m}{\varepsilon^b} \frac{i}{\varepsilon^{-a} \langle k, \omega(\rho) \rangle + l} \hat{P}^{(m)}(k, l, \rho) e^{i\langle k, \theta \rangle + lt}.
\]

(3.20)

By (1.6) and (1.7), for \( \forall \rho \in B(r_m), |k| + |l| \leq K_m, k \neq 0 \), we have

\[
|\varepsilon^{-a} \langle k, \omega(\rho) \rangle + l| \geq |\varepsilon^{-a} \langle k, \omega(I_0) \rangle + l| - |\varepsilon^{-a} \langle k, \omega(I_0) - \omega(\rho) \rangle|.
\]
\[ \geq \frac{e^{-a + \frac{B}{\tau}} \gamma}{|k|^{t_1}} - C e^{-a} |k| r_m \]
\[ \geq \frac{e^{-a + \frac{B}{\tau}} \gamma}{2|k|^{t_1}}. \] (3.21)

Then, by (3.10) and (3.19)–(3.21), using Rüssmann subtle arguments [19, 20] to give optimal estimates of small divisor series (also see Lemma 5.1 in [10]), we get

\[ \| S(\theta, t, \rho) \|_{D(s_m^{(1)}, r_m^{(1)})} \leq C e^{a - b - \frac{B}{\tau}} e_m \| P^{(m)}(\theta, t, \rho) \|_{D(s_m, r_m)} \leq \frac{C e^{a - b - \frac{B}{\tau}} e_m}{\gamma s_m^{t_1}}. \] (3.22)

Then by the Cauchy’s estimate, we have

\[ \left\| \frac{\partial S}{\partial \theta} \right\|_{D(s_m^{(2)}, r_m^{(2)})} \leq \frac{C e^{a - b - \frac{B}{\tau}} e_m}{\gamma s_m^{t_1+1}} \ll r_m - r_{m+1}, \]
\[ \left\| \frac{\partial S}{\partial \rho} \right\|_{D(s_m^{(1)}, r_m^{(1)})} \leq \frac{C e^{a - b - \frac{B}{\tau}} e_m}{\gamma s_m r_m} \ll s_m - s_{m+1}. \] (3.23)

By (3.12) and (3.23) and the implicit function theorem, we get that there are analytic functions \( u = u(\phi, t, \rho), v = v(\phi, t, \rho) \) defined on the domain \( D(s_m^{(3)}, r_m^{(3)}) \) with

\[ \frac{\partial S(\theta, t, \rho)}{\partial \theta} = u(\phi, t, \rho), \quad \frac{\partial S(\theta, t, \rho)}{\partial \rho} = -v(\phi, t, \rho) \] (3.24)

and

\[ \| u \|_{D(s_m^{(3)}, r_m^{(3)})} \leq \frac{C e^{a - b - \frac{B}{\tau}} e_m}{\gamma s_m^{t_1+1}} \ll r_m - r_{m+1}, \]
\[ \| v \|_{D(s_m^{(3)}, r_m^{(3)})} \leq \frac{C e^{a - b - \frac{B}{\tau}} e_m}{\gamma s_m r_m} \ll s_m - s_{m+1} \] (3.25)

such that

\[ \Phi_{m+1}: \begin{cases} I = \rho + u(\phi, t, \rho), \\ \theta = \phi + v(\phi, t, \rho), \\ t = t. \end{cases} \] (3.26)

Then, we have

\[ \Phi_{m+1}(D(s_{m+1}, r_{m+1})) \subseteq \Phi_{m+1}(D(s_m^{(3)}, r_m^{(3)})) \subseteq D(s_m, r_m). \] (3.27)
Let
\[
\begin{align*}
    h^{(m+1)}(t, \rho) &= h^{(m)}(t, \rho) + \varepsilon_m P_1^{(m)}(0, t, \rho), \\
    \varepsilon_{m+1} P^{(m+1)}(\phi, t, \rho) &= \varepsilon_m P_2^{(m)}(\theta, t, \rho) + P_{m+1} \circ \Phi^{(m+1)}(\phi, t, \rho) + \sum_{v=m+2}^{\infty} P_v \circ \Phi^{(m+1)}(\phi, t, \rho) + R_1.
\end{align*}
\] (3.28)

Then by (3.14), (3.16), (3.17), (3.28) and (3.29), we have
\[
H^{(m+1)}(\phi, t, \rho) = \frac{H_0(\rho)}{\varepsilon^a} + \frac{h^{(m+1)}(t, \rho)}{\varepsilon^b} + \varepsilon_{m+1} P^{(m+1)}(\phi, t, \rho) + \sum_{v=m+2}^{\infty} P_v \circ \Phi^{(m+1)}(\phi, t, \rho).
\] (3.30)

By (3.10) and (3.16), it is not difficult to show that (see Lemma A.2 in [17]),
\[
\left\| \frac{\varepsilon_m P_2^{(m)}(\theta, t, \rho)}{\varepsilon^b} \right\|_{D(s_m^{(9)}, r_m^{(9)})} \leq \frac{C \varepsilon_m}{\varepsilon^b} K_{d+1}^{m} e^{-\frac{\varepsilon_m}{\varepsilon^b} \gamma} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b}.
\] (3.31)

By (3.10), (3.16), (3.28) and (3.31), we have
\[
\| h^{(m+1)} \|_{D(s_m^{(9)}, r_m^{(9)})} \leq \| h^{(m)} \|_{D(s_m^{(9)}, r_m^{(9)})} + \| \varepsilon_m P_1^{(m)}(0, t, \rho) \|_{D(s_m^{(9)}, r_m^{(9)})} \leq C.
\] (3.32)

By (3.10), (3.16), (3.22)–(3.25), we have
\[
\left\| \frac{1}{\varepsilon^a} \int_0^1 (1 - \tau) \frac{\partial^2 H_0}{\partial \tau^2} \left( \rho + \tau \frac{\partial S}{\partial \theta} \right) \frac{\partial S}{\partial \theta} \right\|_{D(s_m^{(9)}, r_m^{(9)})} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b} \tau_{m+1}^{d+1} e^{-\frac{\varepsilon_{m+1}}{\varepsilon^b} \gamma} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b},
\] (3.33)

\[
\left\| \frac{\varepsilon_m}{\varepsilon^b} \int_0^1 \frac{\partial P^{(m)}}{\partial \tau} \left( \theta, t, \rho + \tau \frac{\partial S}{\partial \theta} \right) \frac{\partial S}{\partial \theta} d\tau \right\|_{D(s_m^{(9)}, r_m^{(9)})} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b} \tau_{m+1}^{d+1} e^{-\frac{\varepsilon_{m+1}}{\varepsilon^b} \gamma} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b},
\] (3.34)

\[
\left\| \frac{1}{\varepsilon^b} \int_0^1 \frac{\partial h^{(m)}}{\partial \tau} \left( t, \rho + \tau \frac{\partial S}{\partial \theta} \right) \frac{\partial S}{\partial \theta} d\tau \right\|_{D(s_m^{(9)}, r_m^{(9)})} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b} \tau_{m+1}^{d+1} e^{-\frac{\varepsilon_{m+1}}{\varepsilon^b} \gamma} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b}.
\]
\begin{align}
\leq \frac{C}{\varepsilon^b r_m} \cdot \varepsilon^{a-b-\gamma \varepsilon_m} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b}.
\end{align}
(3.35)

By (3.25) and (3.26), we have
\begin{align}
\Phi_{m+1}(\phi, t, \rho) = (\theta, t, I), \quad (\phi, t, \rho) \in D(s_m^{(3)}, r_m^{(3)}).
\end{align}
(3.36)

By (3.25), (3.26) and (3.36), we have
\begin{align}
\| I - \rho \|_{D(s_m^{(3)}, r_m^{(3)})} \leq \frac{C \varepsilon_{m+1}}{\gamma s_m^{\tau_1+1}}, \quad \| \theta - \phi \|_{D(s_m^{(3)}, r_m^{(3)})} \leq \frac{C \varepsilon^{a-b-\gamma \varepsilon_m}}{\gamma s_m^{\tau_1} r_m}.
\end{align}
(3.37)

By (3.26), (3.37) and Cauchy’s estimate, we have
\begin{align}
\| \partial (\Phi_{m+1} - id) \|_{D(s_m^{(4)}, r_m^{(4)})} \leq \frac{C \varepsilon^{a-b-\gamma \varepsilon_m}}{\gamma s_m^{\tau_1+1} r_m}.
\end{align}
(3.38)

It follows that
\begin{align}
\| \partial (\Phi_{m+1} - id) \|_{D(s_{m+1}, r_{m+1})} \leq \frac{1}{2^{m+2}}.
\end{align}
(3.39)

By (3.6), (3.7), (3.27) and (3.39), we have
\begin{align}
\| \partial \Phi_{m+1}(\phi, t, \rho) \|_{D(s_{m+1}, r_{m+1})} &= \| (\partial \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_{m+1}) \\
&\quad (\partial \Phi_2 \circ \Phi_3 \circ \cdots \circ \Phi_{m+1}) \cdots (\partial \Phi_{m+1}) \|_{D(s_{m+1}, r_{m+1})} \\
&\leq \prod_{j=0}^{m} \left( 1 + \frac{1}{2^{j+2}} \right) \leq 2.
\end{align}
(3.40)

It follows that
\begin{align}
\Phi^{(m+1)}(D(s_v, r_v)) \subset T^{d+1}_{2s_v} \times \mathbb{R}^d_{2s_v}, \quad v = m + 1, m + 2, \cdots.
\end{align}
(3.41)

In fact, suppose that \( w = \Phi^{(m+1)}(z) \) with \( z = (\phi, t, \rho) \in D(s_v, r_v) \). Since \( \Phi^{(m+1)} \) is real for real argument and \( r_v < s_v \), we have
\begin{align}
|\text{Im} w| = |\text{Im} \Phi^{(m+1)}(z)| &= |\text{Im} \Phi^{(m+1)}(z) - \text{Im} \Phi^{(m+1)}(\text{Re} z)| \\
&\leq |\Phi^{(m+1)}(z) - \Phi^{(m+1)}(\text{Re} z)|
\end{align}
\[ \leq \| \partial \Phi^{(m+1)}(\phi, t, \rho) \|_{D(s_{m+1}, r_{m+1})} |\text{Im} z| \]
\[ \leq 2 |\text{Im} z| \leq 2s_{\nu}. \quad (3.42) \]

By (3.3) and (3.41), we have
\[ \| P_{m+1} \circ \Phi^{(m+1)}(\phi, t, \rho) \|_{D(s_{m+1}, r_{m+1})} \leq \frac{C \varepsilon_{m+1}}{\varepsilon^b}, \quad (3.43) \]
\[ \| P_v \circ \Phi^{(m+1)}(\phi, t, \rho) \|_{D(s_v, r_v)} \leq C \varepsilon_v, \quad v = m + 2, m + 3, \ldots \quad (3.44) \]

By (3.15), (3.25), (3.29), (3.31), (3.33)–(3.35) and (3.43), we have
\[ \| P^{(m+1)}(\phi, t, \rho) \|_{D(s_m, r_m)} \leq C. \quad (3.45) \]

The proof is finished by (3.27), (3.30), (3.32), (3.39), (3.44) and (3.45). \(\square\)

By Lemma 3.1, there is a symplectic transformation \(\Phi^{(m_0)} = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{m_0}\)
with
\[ \Phi^{(m_0)} : D(s_{m_0}, r_{m_0}) \to D(s_0, r_0) \]
such that system (3.5) is changed by \(\Phi^{(m_0)}\) into
\[ H^{(m_0)} = \frac{H_0(I)}{\varepsilon^a} + \frac{h^{(m_0)}(t, I)}{\varepsilon^b} + \frac{\varepsilon^B P^{(m_0)}(\theta, t, I)}{\varepsilon^b} + \sum_{v=m_0+1}^{\infty} \frac{P_v \circ \Phi^{(m_0)}(\theta, t, I)}{\varepsilon^b} \quad (3.46) \]

where
\[ \| h^{(m_0)}(t, I) \|_{D(s_{m_0}, r_{m_0})} \leq C, \quad (3.47) \]
\[ \| P^{(m_0)}(\theta, t, I) \|_{D(s_{m_0}, r_{m_0})} \leq C, \quad (3.48) \]
\[ \| P_v \circ \Phi^{(m_0)}(\theta, t, I) \|_{D(s_v, r_v)} \leq C \varepsilon_v, \quad v = m_0 + 1, m_0 + 2, \ldots \quad (3.49) \]

4 A Symplectic Transformation

Let \(h^{(m_0)}(I) = \hat{h}^{(m_0)}(0, I)\) be the 0-Fourier coefficient of \(h^{(m_0)}(t, I)\) as a function of \(t\). In order to eliminate the dependence of \(h^{(m_0)}(t, I)\) on the time-variable \(t\), we introduce the following transformation
\[ \Psi : \rho = I, \phi = \theta + \frac{\partial \tilde{S}(t, I)}{\partial I}, \quad (4.1) \]
where \( \tilde{S}(t, I) = \frac{1}{\varepsilon^b} \int_0^t (\{h^{(m_0)}\}(I) - h^{(m_0)}(\xi, I)) \, d\xi \). It is symplectic by easy verification \( \, d\rho \wedge d\phi = dI \wedge d\theta \). Noting that the transformation is not small. So \( \Psi \) is not close to the Identity. Let

\[
\tilde{s}_0 = \varepsilon^{b+\frac{(m_0+1)(2\tau_1+3)B}{m_0}} + 4\mu_1 + \frac{2B}{1}.
\]

\[
\tilde{r}_0 = \varepsilon^{a+(\tau_2+1)b+\frac{(m_0+1)(2\tau_1+3)(\tau_2+1)B}{m_0}} + 4\mu_1(\tau_2+1) + \mu_2 + \frac{2B(\tau_2+1)}{1},
\]

where \( \mu_1 = \frac{(a-b)^2\mu}{1000(a+b+1)(d+3)(5a-b+2ad)} \), \( \mu_2 = 2\mu_1 \). We introduce a domain

\[
\mathcal{D} := \left\{ t = t_1 + t_2i \in \mathbb{T}_{s_{m_0}} : |t_2| \leq \tilde{s}_0 \right\} \times \left\{ I = I_1 + I_2i \in B(r_{m_0}) : |I_2| \leq \tilde{r}_0 \right\},
\]

where \( t_1, t_2, I_1, I_2 \) are real numbers. Noting that \( h^{(m_0)}(t, I) \) is real for real arguments. Thus, for \( (t, I) \in \mathcal{D} \), we have

\[
\| \text{Im} \frac{\partial \tilde{S}(t, I)}{\partial I} \|_{\mathcal{D}} = \| \text{Im} \frac{\partial \tilde{S}}{\partial I}(t_1 + t_2i, I_1 + I_2i) - \text{Im} \frac{\partial \tilde{S}}{\partial I}(t_1, I_1) \|_{\mathcal{D}} \]
\[
\leq \left\| \frac{\partial \tilde{S}}{\partial I}(t_1 + t_2i, I_1 + I_2i) - \frac{\partial \tilde{S}}{\partial I}(t_1, I_1) \right\|_{\mathcal{D}} \]
\[
\leq \left\| \frac{\partial^2 \tilde{S}(t, I)}{\partial I \partial t} \right\|_{\mathcal{D}} \| t_2i \|_{\mathcal{D}} + \left\| \frac{\partial^2 \tilde{S}(t, I)}{\partial^2 I} \right\|_{\mathcal{D}} \| I_2i \|_{\mathcal{D}} \]
\[
\leq \frac{C\tilde{s}_0}{\varepsilon^b r_{m_0} s_{m_0}} + \frac{C\tilde{r}_0}{\varepsilon^b r_{m_0}^2}
\]
\[
\leq \frac{1}{2} s_{m_0}. \tag{4.2}
\]

By (3.46), (4.1) and (4.2), we have

\[
\Psi(T^d_{s_{m_0}} / 2 \times \mathcal{D}) \subset D(s_{m_0}, r_{m_0}) \tag{4.3}
\]

and

\[
\tilde{H}(\phi, t, \rho) = H^{(m_0)} \circ \Psi
\]
\[
= \frac{H_0(\rho)}{\varepsilon^{a}} + \frac{[h^{(m_0)}](\rho)}{\varepsilon^{b}} + \frac{\varepsilon^{B} \tilde{P}^{(m_0)}(\phi, t, \rho)}{\varepsilon^{b}}
\]
\[
+ \sum_{v=m_0+1}^{\infty} \frac{P_v \circ \Phi^{(m_0)} \circ \Psi(\phi, t, \rho)}{\varepsilon^{b}}, \tag{4.4}
\]

where \( \tilde{P}^{(m_0)}(\phi, t, \rho) = P^{(m_0)}(\phi - \frac{\partial}{\partial T} \tilde{S}(t, \rho), t, \rho) \) and \( \| \tilde{P}^{(m_0)} \|_{T^d_{s_{m_0}} / 2 \times \mathcal{D}} \leq C \).
5 Iterative Lemma

By (3.47), we have
\[ \varepsilon^{a-b} \left\| \frac{\partial^2 [h^{(m_0)}]}{\partial \rho^2} \right\|_{D(0, \frac{r_{m_0}}{2})} \leq \frac{C \varepsilon^{a-b}}{r_{m_0}^2} \ll 1. \] (5.1)

Then by (1.6), (3.47) and (5.1), solving the equation
\[ \frac{\partial H_0}{\partial \rho} (\tilde{I}_0) + \varepsilon^{a-b} \frac{\partial [h^{(m_0)}]}{\partial \rho} (\tilde{I}_0) = \omega(I_0), \] (5.2)
where \( \omega(I_0) = \frac{\partial H_0}{\partial \rho} (I_0). \) For any \( c > 0 \) and any \( y_0 \in \mathbb{R}^d, \) let
\[ B(y_0, c) = \{ z \in \mathbb{C}^d : |z - y_0| \leq c \}. \]

Define
\[ \tilde{D}(s, r(I)) = T_{s}^{d+1} \times B(I, r), \quad \tilde{D}(s, 0) = T_{s}^{d+1}, \quad \tilde{D}(0, r(I)) = B(I, r). \]

Let \( \tilde{\varepsilon}_0 = \varepsilon_{m_0} = \varepsilon^B. \) Noting that \( |\tilde{I}_0 - I_0| \ll r_{m_0}, \) and by (4.3), we have
\[ \Psi(\tilde{D}(\tilde{s}_0, \tilde{r}_0(\tilde{I}_0))) \subset D(s_{m_0}, r_{m_0}). \] (5.3)

Then we can rewrite the Hamiltonian (4.4) as follows:
\[ \tilde{H}^{(0)}(\theta, t, I) = \frac{H_0^{(0)}(I)}{\varepsilon^a} + \frac{\tilde{P}^{(0)}(\theta, t, I)}{\varepsilon^b} + \sum_{v=m_0+1}^{\infty} \frac{P_v \circ \Phi^{(m_0)} \circ \Psi(\theta, t, I)}{\varepsilon^b}, \] (5.4)

where \((\theta, t, I) \in \tilde{D}(\tilde{s}_0, \tilde{r}_0(\tilde{I}_0)), H_0^{(0)}(I) = H_0(I) + \varepsilon^{a-b} [h^{(m_0)}](I), \tilde{P}^{(0)} = \varepsilon^B \tilde{P}^{(m_0)}\) and
\[ \frac{\partial H_0^{(0)}}{\partial I} (\tilde{I}_0) = \omega(I_0), \] (5.5)
\[ \|\tilde{P}^{(0)}\|_{\tilde{D}(\tilde{s}_0, \tilde{r}_0(\tilde{I}_0))} \leq C \tilde{\varepsilon}_0. \] (5.6)
By (1.6) and (5.1), we get that there exist constants $M_0 > 0$, $h_0 > 0$ such that
\[
\det \left( \frac{\partial^2 H_0^{(0)}(I)}{\partial I^2} \right), \ det \left( \frac{\partial^2 H_0^{(0)}(I)}{\partial I^2} \right)^{-1} \leq M_0, \ \forall I \in \tilde{D}(0, \tilde{r}_0(\tilde{I}_0))
\]
(5.7)
and
\[
\left\| \frac{\partial^2 H_0^{(0)}(I)}{\partial I^2} \right\|_{\tilde{D}(0, \tilde{r}_0(\tilde{I}_0))} \leq h_0.
\]
(5.8)

Define sequences

- $\tilde{e}_0 = e^{m_0} = e^B$, $\tilde{e}_{j+1} = e^{1+\mu_3} = e^{m_0+1+j}$ with $\mu_3 = \frac{(a-b)\mu}{10B}$, $j = 0, 1, \ldots$;
- $\tilde{s}_0 = \exp \left( \frac{b+\frac{(m_0+1)(2\tau_1+3)B}{2}}{m_0} \right) + 4\mu_1 + \frac{2B}{\varepsilon}$ with $\mu_1 = \frac{1000(a+b+1)(d+3)(5a-b+2m_1)}{1000(a+b+1)(d+3)(5a-b+2m_1)}$, $\tilde{s}_{j+1} = \tilde{s}_j + \frac{1}{10} (\tilde{s}_j - \tilde{s}_{j+1})$, $l = 0, 1, \ldots, 10$, $j = 0, 1, 2, \ldots$;
- $\tilde{r}_0 = \exp \left( \frac{a+(\tau_2+1)b+\frac{(m_0+1)(2\tau_2+3)(\tau_2+1)B}{2}}{m_0} \right) + 4\mu_1 (\tau_2+1) + \mu_2 + \frac{2B}{\varepsilon}$ with $\mu_2 = \frac{2\mu_1}{\varepsilon}$, $\tilde{r}_{j+1} = \tilde{r}_j + \frac{1}{10} (\tilde{r}_j - \tilde{r}_{j+1})$, $l = 0, 1, \ldots, 10$, $j = 0, 1, 2, \ldots$;
- $\tilde{K}_j = \frac{2}{\tilde{s}_j} \log \frac{1}{\tilde{r}_j}$, $j = 0, 1, 2, \ldots$;
- $h_j = h_0 (2 - \frac{1}{2l})$, $j = 0, 1, 2, \ldots$;
- $M_j = M_0 (2 - \frac{1}{2l})$, $j = 0, 1, 2, \ldots$.

We claim that
\[
\| P_v \circ \Phi^{(m_0)} \circ \Psi(\theta, t, I) \|_{\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_0))} \leq C \varepsilon_v = C \tilde{e}_{v-m_0},
\]
(5.9)

In fact, for $(t, I) = (t_1 + t_2i, I_1 + I_2i) \in \tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_0))$, where $t_1$, $t_2$, $I_1$, $I_2$ are real numbers, we have
\[
\left\| \text{Im} \left( \frac{\partial \tilde{S}(t, I)}{\partial I} \right) \right\|_{\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_0))} \leq \left\| \text{Im} \left( \frac{\partial \tilde{S}(t_1 + t_2i, I_1 + I_2i)}{\partial I} \right) - \text{Im} \left( \frac{\partial \tilde{S}(t_1, I_1)}{\partial I} \right) \right\|_{\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_0))} + \left\| \frac{\partial^2 \tilde{S}(t, I)}{\partial I^2} \right\|_{\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_0))} \| I_2 i \|_{\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_0))}
\]
\[
\leq C \tilde{e}_{v-m_0} + \frac{C \tilde{r}_{v-m_0}^2}{\varepsilon \tilde{r}_{m_0} s_{m_0}} + \frac{C \tilde{r}_{v-m_0}^2}{\varepsilon \tilde{r}_{m_0}^2 s_{m_0}}
\]
\[ \leq \frac{1}{2} s_v. \]  (5.10)

It follows that
\[ \Psi(\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_0))) \subset \tilde{D}(s_v, \tilde{r}_{v-m_0}(\tilde{I}_0)). \]  (5.11)

Suppose that
\[ w = \Phi^{(m_0)}(z) \text{ with } z = (\theta, t, I) \in \tilde{D}(s_v, \tilde{r}_{v-m_0}(\tilde{I}_0)) \subset D(s_{m_0}, r_{m_0}). \]
Since \( \Phi^{(m_0)} \) is real for real argument and \( \tilde{r}_{v-m_0} < r_v < s_v \), then by (3.40) with \( m = m_0 - 1 \), we have
\[ |\text{Im} w| = |\text{Im} \Phi^{(m_0)}(z) - \text{Im} \Phi^{(m_0)}(\text{Re} z)| \]
\[ \leq |\Phi^{(m_0)}(z) - \Phi^{(m_0)}(\text{Re} z)| \]
\[ \leq \| \partial \Phi^{(m_0)}(\theta, t, I) \|_{D(s_{m_0}, r_{m_0})} |\text{Im} z| \]
\[ \leq 2 |\text{Im} z| \leq 2s_v. \]  (5.12)

Then by (5.11) and (5.12), we have
\[ \Phi^{(m_0)} \circ \Psi(\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_0))) \subset T^{d+1}_{2s_v} \times \mathbb{R}^d, \nu = m_0 + 1, m_0 + 2, \ldots \]
(5.13)

By (3.3) and (5.13), the proof of (5.9) is completed. Clearly, by (5.5)–(5.9), (5.4) fulfills (5.17)–(5.21) with \( m = 0 \). Then we have the following lemma.

**Lemma 5.1 (Iterative Lemma)** Suppose that we have had \( m + 1 \) (\( m = 0, 1, 2, \ldots \)) symplectic transformations \( \Phi_0 = \text{id}, \Phi_1, \ldots, \Phi_m \) with
\[ \tilde{\Phi}_j : \tilde{D}(\tilde{s}_j, \tilde{r}_j(\tilde{I}_j)) \to \tilde{D}(\tilde{s}_{j-1}, \tilde{r}_{j-1}(\tilde{I}_{j-1})), \quad j = 1, 2, \ldots, m \]  (5.14)
and
\[ \| \partial(\tilde{\Phi}_j - \text{id}) \|_{\tilde{D}(\tilde{s}_j, \tilde{r}_j(\tilde{I}_j))} \leq \frac{1}{2^{j+1}}, \quad j = 1, 2, \ldots, m \]  (5.15)

where \( \tilde{I}_j \in \mathbb{R}^d, \quad j = 0, 1, 2, \ldots, m \) such that system (5.4) is changed by \( \tilde{\Phi}^{(m)} = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_m \) into
\[ \tilde{H}^{(m)} = \tilde{H}^{(0)} \circ \tilde{\Phi}^{(m)} = \frac{H_0^{(m)}(I)}{\varepsilon^a} + \tilde{P}^{(m)}(\theta, t, I) + \sum_{\nu=m_0+m+1}^{\infty} \frac{P_{\nu} \circ \Phi^{(m_0)} \circ \Psi \circ \tilde{\Phi}^{(m)}(\theta, t, I)}{\varepsilon^b}, \]  (5.16)

where
\[ \frac{\partial H_0^{(m)}}{\partial I}(\tilde{I}_m) = \omega(I_0), \]  (5.17)
\[ \| \tilde{P}^{(m)} \| \tilde{D}(\tilde{s}_m, \tilde{r}_m(I_m)) \leq C \tilde{\epsilon}_m, \quad (5.18) \]
\[ \det \left( \frac{\partial^2 H_0^{(m)}(I)}{\partial I^2} \right), \det \left( \frac{\partial^2 H_0^{(m)}(I)}{\partial I^2} \right)^{-1} \leq M_m, \forall I \in \tilde{D}(0, \tilde{r}_m(I_m)), \quad (5.19) \]
\[ \left\| \frac{\partial^2 H_0^{(m)}(I)}{\partial I^2} \right\|_{\tilde{D}(0, \tilde{r}_m(I_m))} \leq h_m, \quad (5.20) \]
\[ \| P_v \circ \Phi^{(m_0)} \circ \Psi \circ \Phi^{(m)}(\theta, t, I) \|_{\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(I_m))} \leq C \tilde{\epsilon}_{v-m_0}, v = m_0 + m + 1, m_0 + m + 2, \ldots \quad (5.21) \]

Then there is a symplectic transformation \( \tilde{\Phi}_{m+1} \) with
\[ \tilde{\Phi}_{m+1} : \tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(I_{m+1})) \rightarrow \tilde{D}(\tilde{s}_m, \tilde{r}_m(I_m)) \quad (5.22) \]

and
\[ \| \partial(\tilde{\Phi}_{m+1} - \text{id})\|_{\tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(I_{m+1}))} \leq \frac{1}{2^{m+2}} \]

where \( \tilde{I}_{m+1} \in \mathbb{R}^d \) such that system (5.16) is changed by \( \tilde{\Phi}_{m+1} \) into (\( \tilde{\Phi}^{(m+1)} = \Phi_0 \circ \tilde{\Phi}_1 \circ \cdots \circ \tilde{\Phi}_{m+1} \))
\[ \tilde{H}^{(m+1)} = \tilde{H}^{(m)} \circ \Phi_{m+1} = \tilde{H}^{(0)} \circ \Phi^{(m+1)} = \frac{H_0^{(m+1)}(I)}{\varepsilon^a} + \frac{\bar{P}^{(m+1)}(\theta, t, I)}{\varepsilon^b} + \sum_{v=m_0+m+2}^{\infty} \frac{P_v \circ \Phi^{(m_0)} \circ \Psi \circ \Phi^{(m+1)}(\theta, t, I)}{\varepsilon^b}, \]

where \( \tilde{H}^{(m+1)} \) satisfies (5.17)–(5.21) by replacing \( m \) by \( m + 1 \).

**Proof** Assume that the change \( \tilde{\Phi}_{m+1} \) is implicitly defined by
\[ \tilde{\Phi}_{m+1} : \begin{cases} I = \rho + \frac{\partial S}{\partial \rho}, \\ \phi = \theta + \frac{\partial S}{\partial \rho}, \\ t = t, \end{cases} \quad (5.23) \]

where \( S = S(\theta, t, \rho) \) is the generating function, which will be proved to be analytic in a smaller domain \( \tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(I_{m+1})) \). By a simple computation, we have
\[ dI \wedge d\theta = d\rho \wedge d\theta + \sum_{i,j=1}^{d} \frac{\partial^2 S}{\partial \rho_i \partial \theta_j} d\rho_i \wedge d\theta_j = d\rho \wedge d\phi. \]
Thus, the coordinates change $\Phi_{m+1}$ is symplectic if it exists. Moreover, we get the changed Hamiltonian

\[ \tilde{H}^{(m+1)} = \tilde{H}^{(m)} \circ \tilde{\Phi}_{m+1} \]

\[ = \frac{H^{(m)}_0 (\rho + \frac{\partial S}{\partial \theta})}{\epsilon^a} + \frac{\tilde{P}^{(m)} (\theta, t, \rho + \frac{\partial S}{\partial \theta})}{\epsilon^b} \]

\[ + \frac{P_m + 1 \circ \Phi^{(m)} \circ \tilde{\Phi}^{(m+1)} (\phi, t, \rho)}{\epsilon^b} \]

\[ + \frac{\partial S}{\partial t} + \sum_{v=m_0+m+2}^{\infty} \frac{P_v \circ \Phi^{(m)} \circ \tilde{\Phi}^{(m+1)} (\phi, t, \rho)}{\epsilon^b}, \quad (5.24) \]

where $\theta = \theta (\phi, t, \rho)$ is implicitly defined by (5.23). By Taylor formula, we have

\[ \tilde{H}^{(m+1)} = \frac{H^{(m)}_0 (\rho)}{\epsilon^a} + \left( \frac{\omega^{(m)} (\rho)}{\epsilon^a} - \frac{\partial S}{\partial \theta} \right) + \frac{\frac{\partial S}{\partial t} + \tilde{P}^{(m)} (\theta, t, \rho)}{\epsilon^b} \]

\[ + \frac{P_m + 1 \circ \Phi^{(m)} \circ \tilde{\Phi}^{(m+1)} (\phi, t, \rho)}{\epsilon^b} + \sum_{v=m_0+m+2}^{\infty} \frac{P_v \circ \Phi^{(m)} \circ \tilde{\Phi}^{(m+1)} (\phi, t, \rho)}{\epsilon^b}, \quad (5.25) \]

where $\omega^{(m)} (\rho) = \frac{\partial H^{(m)}_0}{\partial t} (\rho)$ and

\[ R_1 = \frac{1}{\epsilon^a} \int_0^1 (1 - \tau) \frac{\partial^2 H^{(m)}_0}{\partial t^2} (\rho + \tau \frac{\partial S}{\partial \theta}) \left( \frac{\partial S}{\partial \theta} \right)^2 d\tau \]

\[ + \frac{1}{\epsilon^b} \int_0^1 \frac{\partial \tilde{P}^{(m)}}{\partial t} (\theta, t, \rho + \tau \frac{\partial S}{\partial \theta}) \frac{\partial S}{\partial \theta} d\tau. \quad (5.26) \]

Expanding $\tilde{P}^{(m)} (\theta, t, \rho)$ into a Fourier series,

\[ \tilde{P}^{(m)} (\theta, t, \rho) = \sum_{(k,l) \in \mathbb{Z}^d \times \mathbb{Z}} \tilde{P}^{(m)} (k, l, \rho) e^{i(k, \theta) + il} := \tilde{P}^{(m)}_1 (\theta, t, \rho) + \tilde{P}^{(m)}_2 (\theta, t, \rho), \]

\[ (5.27) \]

where $\tilde{P}^{(m)}_1 = \sum_{|k| + |l| \leq K_m} \tilde{P}^{(m)} (k, l, \rho) e^{i(k, \theta) + il}, \tilde{P}^{(m)}_2 = \sum_{|k| + |l| > K_m} \tilde{P}^{(m)} (k, l, \rho) e^{i(k, \theta) + il}$. Then, we derive the homological equation:

\[ \frac{\partial S}{\partial t} + \left( \frac{\omega^{(m)} (\rho)}{\epsilon^a} - \frac{\partial S}{\partial \theta} \right) + \frac{\tilde{P}^{(m)} (\theta, t, \rho)}{\epsilon^b} - \frac{\tilde{P}^{(m)} (0, 0, \rho)}{\epsilon^b} = 0, \quad (5.28) \]
where $\tilde{P}^{(m)}(0,0,\rho)$ is the 0-Fourier coefficient of $\tilde{P}^{(m)}(\theta, t, \rho)$ as a function of $(\theta, t)$. Let

$$S(\theta, t, \rho) = \sum_{|k|+|l| \leq \tilde{K}_m, (k,l) \neq (0,0)} \hat{S}(k, l, \rho) e^{i(k, \theta) + l t}. \quad (5.29)$$

By passing to Fourier coefficients, we have

$$\hat{S}(k, l, \rho) = \frac{i}{\epsilon^b} \cdot \frac{\tilde{P}^{(m)}(k, l, \rho)}{\epsilon^{-a} \langle k, \omega^{(m)}(\rho) \rangle + l}, \ |k| + |l| \leq \tilde{K}_m, (k, l) \in \mathbb{Z}^d \times \mathbb{Z} \setminus \{(0,0)\}. \quad (5.30)$$

Then we can solve homological Eq. (5.28) by setting

$$S(\theta, t, \rho) = \sum_{|k|+|l| \leq \tilde{K}_m, (k,l) \neq (0,0)} \frac{i}{\epsilon^b} \cdot \frac{\tilde{P}^{(m)}(k, l, \rho)}{\epsilon^{-a} \langle k, \omega^{(m)}(\rho) \rangle + l}e^{i(k, \theta) + l t}. \quad (5.31)$$

By (1.8), (5.17) and (5.19), for $\forall \rho \in \tilde{D}(\tilde{s}_m, \tilde{r}_m(\tilde{I}_m))$, $|k| + |l| \leq \tilde{K}_m, (k, l) \neq (0, 0)$, we have

$$|e^{-a} \langle k, \omega^{(m)}(\rho) \rangle + l| \geq |e^{-a} \langle k, \omega^{(m)}(\tilde{I}_m) \rangle + l - |e^{-a} \langle k, \omega^{(m)}(\tilde{I}_m) - \omega^{(m)}(\rho) \rangle| \geq \frac{\gamma}{|k|^2} - C \epsilon^{-a} |k| \tilde{r}_m$$

$$\geq \frac{\gamma}{2|k|^2}. \quad (5.32)$$

Then, by (5.18) and (5.30)–(5.32), using Rüschmann subtle arguments [19, 20] to give optimal estimates of small divisor series (also see Lemma 5.1 in [10]), we get

$$\|S(\theta, t, \rho)\|_{\tilde{D}(\tilde{s}_m^{(1)}, \tilde{r}_m(\tilde{I}_m))} \leq \frac{C \|\tilde{P}^{(m)}\|_{\tilde{D}(\tilde{s}_m, \tilde{r}_m(\tilde{I}_m))}}{\gamma \epsilon^{b \tilde{r}_m^2 \tilde{s}_m^2}} \leq \frac{C \tilde{e}_m}{\gamma \epsilon^{b \tilde{r}_m^2 \tilde{s}_m^2}}. \quad (5.33)$$

Then by Cauchy’s estimate, we have

$$\left\| \frac{\partial S}{\partial \theta} \right\|_{\tilde{D}(\tilde{s}_m^{(2)}, \tilde{r}_m(\tilde{I}_m))} \leq \frac{C \tilde{e}_m}{\gamma \epsilon^{b \tilde{r}_m^2 \tilde{s}_m^2 + 1}} \ll \tilde{r}_m - \tilde{r}_{m+1},$$

$$\left\| \frac{\partial S}{\partial \rho} \right\|_{\tilde{D}(\tilde{s}_m^{(2)} \tilde{r}_m^{(1)}(\tilde{I}_m))} \leq \frac{C \tilde{e}_m}{\gamma \epsilon^{b \tilde{r}_m^2 \tilde{s}_m^2 \tilde{r}_m}} \ll \tilde{s}_m - \tilde{s}_{m+1}. \quad (5.34)$$

By (5.23) and (5.34) and the implicit function theorem, we get that there are analytic functions $u = u(\phi, t, \rho), v = v(\phi, t, \rho)$ defined on the domain $\tilde{D}(\tilde{s}_m^{(3)}, \tilde{r}_m^{(3)}(\tilde{I}_m))$ with

$$\frac{\partial S(\theta, t, \rho)}{\partial \theta} = u(\phi, t, \rho), \frac{\partial S(\theta, t, \rho)}{\partial \rho} = -v(\phi, t, \rho) \quad (5.35)$$
and

\[ \|u\|_{\tilde{D}((\tilde{s}_m^{(3)}, \tilde{r}_m^{(3)}(I_m)))} \leq \frac{C\tilde{\varepsilon}_m}{\sqrt[2]{\varepsilon b^2 \tilde{s}_m^{t_2 + 1}}} \ll \tilde{r}_m - \tilde{r}_{m+1}, \]

\[ \|v\|_{\tilde{D}((\tilde{s}_m^{(3)}, \tilde{r}_m^{(3)}(I_m)))} \leq \frac{C\tilde{\varepsilon}_m}{\sqrt[2]{\varepsilon b^2 \tilde{s}_m^{t_2}}} \ll \tilde{s}_m - \tilde{s}_{m+1} \]  

(5.36)

such that

\[ \Phi_{m+1} : \begin{align*}
I &= \rho + u(\phi, t, \rho), \\
\theta &= \phi + v(\phi, t, \rho), \\
t &= t.
\end{align*} \]  

(5.37)

Then, we have

\[ \Phi_{m+1}(\tilde{D}(\tilde{s}_m^{(3)}, \tilde{r}_m^{(3)}(I_m))) \subseteq \tilde{D}(\tilde{s}_m, \tilde{r}_m(I_m)). \]  

(5.38)

Let

\[ H_0^{(m+1)}(\rho) = H_0^{(m)}(\rho) + \varepsilon^{a-b} \widehat{P}^{(m)}(0, 0, \rho). \]  

(5.39)

By Cauchy’s estimate and (5.18), we have

\[ \left\| \frac{\partial^p \widehat{P}^{(m)}(0, 0, \rho)}{\partial \rho^p} \right\|_{\tilde{D}(0, \tilde{r}_m^{(2)}(I_m))} \leq \frac{C\tilde{\varepsilon}_m}{\tilde{r}_m^p}, \quad p = 1, 2. \]  

(5.40)

By (5.19), (5.20), (5.39) and (5.40), we have

\[ \det \left( \frac{\partial^2 H_0^{(m+1)}(\rho)}{\partial \rho^2} \right), \det \left( \frac{\partial^2 H_0^{(m+1)}(\rho)}{\partial \rho^2} \right)^{-1} \leq M_{m+1}, \quad \forall \rho \in \tilde{D}(0, \tilde{r}_m^{(2)}(I_m)) \]  

(5.41)

and

\[ \left\| \frac{\partial^2 H_0^{(m+1)}(\rho)}{\partial \rho^2} \right\|_{\tilde{D}(0, \tilde{r}_m^{(2)}(I_m))} \leq h_{m+1}. \]  

(5.42)

By (5.39), we have

\[ \frac{\partial H_0^{(m+1)}(\rho)}{\partial \rho} = \frac{\partial H_0^{(m)}(\rho)}{\partial \rho} + \varepsilon^{a-b} \frac{\partial \widehat{P}^{(m)}(0, 0, \rho)}{\partial \rho}. \]  

(5.43)
Noting that $H^{(m+1)}_0(\rho)$, $\tilde{H}^{(m)}_0(\rho)$ and $\hat{P}^{(m)}(0, 0, \rho)$ are real analytic on $\tilde{D}(0, \tilde{r}_m^{(2)}(\tilde{I}_m))$ and that $\tilde{I}_m \in \mathbb{R}^d$. Then by (5.17), (5.39)–(5.41) and (5.43), it is not difficult to see that (see Appendix A “The Classical Implicit Function Theorem” in [2]) there exists a unique point $\tilde{I}_{m+1} \in \mathbb{R}^d$ so that

$$\frac{\partial H^{(m+1)}_0}{\partial \rho}(\tilde{I}_{m+1}) = \omega(I_0),$$

(5.44)

$$|\tilde{I}_{m+1} - \tilde{I}_m| \leq \frac{C \varepsilon a - b \tilde{\varepsilon}_m}{\tilde{r}_m} \ll \tilde{r}_m.$$

(5.45)

By (5.38) and (5.45), we have

$$\Phi_{m+1}(\tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(\tilde{I}_{m+1}))) \leq \Phi_{m+1}(\tilde{D}(\tilde{s}_m^{(9)}, \tilde{r}_m^{(9)}(\tilde{I}_{m+1})))$$

$$\leq \Phi_{m+1}(\tilde{D}(\tilde{s}_m^{(3)}, \tilde{r}_m^{(3)}(\tilde{I}_{m}))) \subseteq \tilde{D}(\tilde{s}_m, \tilde{r}_m(\tilde{I}_{m})).$$

(5.46)

Let

$$\tilde{P}^{(m+1)}(\phi, t, \rho) \varepsilon b = \tilde{P}_2^{(m)}(\theta, t, \rho) \varepsilon b + P_{m_0+m+1} \circ \Phi^{(m_0)} \circ \Psi \circ \tilde{\Phi}^{(m+1)}(\phi, t, \rho) \varepsilon b + R_1.$$  

(5.47)

Then by (5.25), (5.27), (5.28), (5.39) and (5.47), we have

$$\tilde{H}^{(m+1)}(\phi, t, \rho) = \frac{H^{(m+1)}_0(\rho)}{\varepsilon a} + \frac{\tilde{P}^{(m+1)}(\phi, t, \rho)}{\varepsilon b}$$

$$+ \sum_{v=m_0+m+2}^{\infty} P_v \circ \Phi^{(m_0)} \circ \Psi \circ \tilde{\Phi}^{(m+1)}(\phi, t, \rho) \varepsilon b. $$

(5.48)

By (5.18), (5.27) and (5.45), it is not difficult to show that (see Lemma A.2 in [17]), we have

$$\left\| \frac{\tilde{P}_2^{(m)}(\theta, t, \rho)}{\varepsilon b} \right\|_{\tilde{D}(\tilde{s}_m^{(9)}, \tilde{r}_m^{(9)}(\tilde{I}_{m+1})))} \leq \frac{C \tilde{\varepsilon}_m}{\varepsilon b} K_d^{d+1} \tilde{e}_m \varepsilon b \leq \frac{C \tilde{\varepsilon}_m^{m+1}}{\varepsilon b}. $$

(5.49)

By (5.18), (5.20), (5.33)–(5.36) and (5.45), we have

$$\left\| \frac{1}{\varepsilon a} \int_0^1 (1 - \tau) \frac{\partial^2 H^{(m)}_0}{\partial I^2} \left( \rho + \tau \frac{\partial S}{\partial \theta} \right) \left( \frac{\partial S}{\partial \theta} \right)^2 d\tau \right\|_{\tilde{D}(\tilde{s}_m^{(9)}, \tilde{r}_m^{(9)}(\tilde{I}_{m+1})))} \leq \frac{C \tilde{\varepsilon}_m^{m+1}}{\varepsilon b}, $$

(5.50)
By (5.36) and (5.37), we have

$$\tilde{\Phi}_{m+1}(\phi, t, \rho) = (\theta, t, I), \ (\phi, t, \rho) \in \tilde{D}(\tilde{s}^{(3)}_m, \tilde{r}^{(3)}_m(\tilde{I}_m)).$$

(5.52)

By (5.36), (5.37) and (5.52), we have

$$\|I - \rho\|_{\tilde{D}(\tilde{s}^{(3)}_m, \tilde{r}^{(3)}_m(\tilde{I}_m))} \leq \frac{C\tilde{\varepsilon}_m}{\gamma \varepsilon b^{\tilde{s}_m^{1/2}}}, \ \|\theta - \phi\|_{\tilde{D}(\tilde{s}^{(3)}_m, \tilde{r}^{(3)}_m(\tilde{I}_m))} \leq \frac{C\tilde{\varepsilon}_m}{\gamma \varepsilon b^{\tilde{s}_m^{1/2}}}. \quad (5.53)$$

By (5.37), (5.53) and Cauchy’s estimate, we have

$$\|\partial(\tilde{\Phi}_{m+1} - id)\|_{\tilde{D}(\tilde{s}^{(3)}_m, \tilde{r}^{(3)}_m(\tilde{I}_m))} \leq \frac{C\tilde{\varepsilon}_m}{\gamma \varepsilon b^{\tilde{s}_m^{1/2}}}. \quad (5.54)$$

By (5.45) and (5.54), we have

$$\|\partial(\tilde{\Phi}_{m+1} - id)\|_{\tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(\tilde{I}_{m+1}))} \leq \frac{1}{2^{m+2}}. \quad (5.55)$$

By (5.14), (5.15), (5.46) and (5.55), we have

$$\|\partial(\tilde{\Phi}^{(m+1)}(\phi, t, \rho))\|_{\tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(\tilde{I}_{m+1}))}$$

$$= \|\partial \tilde{\Phi}_1 \circ \tilde{\Phi}_2 \circ \cdots \circ \tilde{\Phi}_{m+1}(\tilde{I}_{m+1})(\tilde{I}_{m+1})\|_{\tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(\tilde{I}_{m+1}))}$$

$$\leq \prod_{j=0}^{m+1} \left(1 + \frac{1}{2^{j+2}}\right) \leq 2. \quad (5.56)$$

We claim that

$$\|P_v \circ \tilde{\Phi}^{(m_0)} \circ \Psi \circ \tilde{\Phi}^{(m+1)}(\phi, t, \rho)\|_{\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_m))}$$

$$\leq C\tilde{\varepsilon}_{v-m_0}, \ v = m_0 + m + 1, m_0 + m + 2, \ldots \quad (5.57)$$

In fact, suppose that $w = \tilde{\Phi}^{(m+1)}(z)$ with $z = (\phi, t, \rho) \in \tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_m))$. Since $\tilde{\Phi}^{(m+1)}$ is real for real argument and $\tilde{r}_{v-m_0} < \tilde{s}_{v-m_0}$, we have

$$|\text{Im} w| = |\text{Im} \tilde{\Phi}^{(m+1)}(z)| = |\text{Im} \tilde{\Phi}^{(m+1)}(z) - \text{Im} \tilde{\Phi}^{(m+1)}(\text{Re} z)|$$

$$\leq |\tilde{\Phi}^{(m+1)}(z) - \tilde{\Phi}^{(m+1)}(\text{Re} z)|$$

$$\leq \|\partial \tilde{\Phi}^{(m+1)}(\phi, t, \rho)\|_{\tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(\tilde{I}_{m+1}))} |\text{Im} z|$$

$$\leq 2 |\text{Im} z| \leq 2\tilde{s}_{v-m_0}. \quad (5.58)$$
By (5.14), (5.46) and (5.58), we have

\[ \tilde{\Phi}^{(m+1)}(\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_{m+1}))) \subseteq D_v := (T^{d+1}_{2\tilde{s}_{v-m_0}} \times \mathbb{R}^d_{2\tilde{s}_{v-m_0}}) \cap \tilde{D}(\tilde{s}_0, \tilde{r}_0(\tilde{I}_0)). \] (5.59)

For \((t, I) = (t_1 + t_2, I_1 + I_2) \in D_v\), where \(t_1, t_2, I_1, I_2\) are real numbers, we have

\[
\begin{align*}
\| \text{Im} \frac{\partial \tilde{S}(t, I)}{\partial I} \|_{D_v} &= \| \text{Im} \frac{\partial \tilde{S}}{\partial I} (t_1 + t_2, I_1 + I_2) - \text{Im} \frac{\partial \tilde{S}}{\partial I} (t_1, I_1) \|_{D_v} \\
&\leq \| \frac{\partial \tilde{S}}{\partial I} (t_1 + t_2, I_1 + I_2) - \frac{\partial \tilde{S}}{\partial I} (t_1, I_1) \|_{D_v} \\
&\leq \| \frac{\partial^2 \tilde{S}(t, I)}{\partial I \partial t} \|_D \| t_2 I \|_{D_v} + \| \frac{\partial^2 \tilde{S}(t, I)}{\partial^2 I} \|_D \| I_2 \|_{D_v} \\
&\leq \frac{C \tilde{s}_{v-m_0}}{\varepsilon^br_{m_0}^s_{m_0}} + \frac{C \tilde{s}_{v-m_0}}{\varepsilon^b r_m^2} \\
&\leq \frac{1}{2} \tilde{s}_v.
\end{align*}
\] (5.60)

By (5.3), (5.59) and (5.60), we have

\[
\begin{align*}
\Psi \circ \tilde{\Phi}^{(m+1)}(\tilde{D}(\tilde{s}_{v-m_0}, \tilde{r}_{v-m_0}(\tilde{I}_{m+1}))) \\
\subseteq \tilde{D}_v := \left(T^{d+1}_{s_v} \times \mathbb{R}^d_{s_{v-m_0}}\right) \cap \tilde{D}(s_{m_0}, r_{m_0}).
\end{align*}
\] (5.61)

Suppose that \(w = \Phi^{(m_0)}(z)\) with \(z = (\theta, t, I) \in \tilde{D}_v\). Since \(\Phi^{(m_0)}\) is real for real argument and \(2\tilde{s}_{v-m_0} < r_v < s_v\), then by (3.40) with \(m = m_0 - 1\), we have

\[
\begin{align*}
|\text{Im} w| &= |\text{Im} \Phi^{(m_0)}(z)| = |\text{Im} \Phi^{(m_0)}(z) - \text{Im} \Phi^{(m_0)}(\text{Re} z)| \\
&\leq |\Phi^{(m_0)}(z) - \Phi^{(m_0)}(\text{Re} z)| \\
&\leq \| \partial \Phi^{(m_0)}(\theta, t, I) \|_{D(s_{m_0}, r_{m_0})}|\text{Im} z| \\
&\leq 2|\text{Im} z| \leq 2s_v.
\end{align*}
\] (5.62)

Then by (5.61) and (5.62), we have

\[
\Phi^{(m_0)} \circ \Psi \circ \tilde{\Phi}^{(m+1)}(\tilde{D}(\tilde{s}_{v-m_0}) \\
\tilde{r}_{v-m_0}(\tilde{I}_{m+1})) \subseteq T^{d+1}_{2s_v} \times \mathbb{R}^d_{2s_v}, \quad v = m_0 + m + 1, \quad m_0 + m + 2, \ldots
\] (5.63)

By (3.3) and (5.63), the proof of (5.57) is completed. By (5.26), (5.45), (5.47), (5.49)–(5.51) and (5.57), we have

\[
\| \tilde{P}^{(m+1)} \|_{\tilde{D}(\tilde{s}_{m+1}, \tilde{r}_{m+1}(\tilde{I}_{m+1}))} \leq C \tilde{\epsilon}_{m+1}.
\] (5.64)
Then the proof is completed by (5.41), (5.42), (5.44), (5.46), (5.48), (5.55), (5.57) and (5.64).

\[ \square \]

6 Proof of Theorems 1.1 and 1.4

In Lemma 5.1, letting \( m \to \infty \) we get the following lemma:

**Lemma 6.1** There exists a symplectic transformation \( \tilde{\Phi}^{(\infty)} := \lim_{m \to \infty} \tilde{\Phi}_0 \circ \tilde{\Phi}_1 \circ \cdots \circ \tilde{\Phi}_m \) with

\[
\tilde{\Phi}^{(\infty)} : T^{d+1} \times \{ \tilde{I}_\infty \} \to D(\tilde{s}_0, \tilde{r}_0(\tilde{I}_0)), \tag{6.1}
\]

where \( \tilde{I}_\infty \in \mathbb{R}^d \) such that system (5.4) is changed by \( \tilde{\Phi}^{(\infty)} \) into

\[
\tilde{H}^{(\infty)}(\theta, t, I) = \tilde{H}^{(0)} \circ \tilde{\Phi}^{(\infty)} = \frac{H^{(\infty)}_0(I)}{\varepsilon^a}, \tag{6.2}
\]

where

\[
\frac{\partial H^{(\infty)}_0(I)}{\partial I}(\tilde{I}_\infty) = \omega(I_0), \tag{6.3}
\]

\[
\| \tilde{\Phi}^{(\infty)} - id \|_{T^{d+1} \times \tilde{I}_\infty} \leq \tilde{\varepsilon}_0 \frac{\pi}{\ell}. \tag{6.4}
\]

**Proof** By (5.36) and (5.56), for \( z = (\theta, t, I) \in T^{d+1} \times \tilde{I}_\infty \) and \( m = 0, 1, 2, \ldots \), we have

\[
\| \tilde{\Phi}_{m+1}(z) - \tilde{\Phi}_{m}(z) \|_{T^{d+1} \times \tilde{I}_\infty} = \| \tilde{\Phi}_m(\tilde{\Phi}_{m+1}(z)) - \tilde{\Phi}_m(z) \|_{T^{d+1} \times \tilde{I}_\infty} \\
\leq \| \tilde{\Phi}_m(\tilde{\Phi}_{m+1}(z)) \|_{T^{d+1} \times \tilde{I}_\infty} \| \tilde{\Phi}_{m+1}(z) - z \|_{T^{d+1} \times \tilde{I}_\infty} \\
\leq 2\tilde{\varepsilon}_m^{\frac{1}{2}}, \tag{6.5}
\]

where \( \tilde{\Phi}^{(0)} := id. \) Then, we have

\[
\| \tilde{\Phi}^{(\infty)}(z) - z \|_{T^{d+1} \times \tilde{I}_\infty} \leq \sum_{m=0}^{\infty} \| \tilde{\Phi}_{m+1}(z) - \tilde{\Phi}_m(z) \|_{T^{d+1} \times \tilde{I}_\infty} \leq \sum_{m=0}^{\infty} 2\tilde{\varepsilon}_m^{\frac{1}{2}} \leq \tilde{\varepsilon}_0^{\frac{1}{2}}. 
\]

This completes the proof of Lemma 6.1. \( \square \)

Then the proof of Theorem 1.1 is completed by (3.1), (3.5), (3.46), (4.4), (5.4) and Lemma 6.1. Applying Theorem 1.1 to equation (1.3) we have Theorem 1.4 (The proof is similar to Section 5 in [28], we omit it.).
References

1. Chatterjee, A.K., Kundu, A., Kulkarni, M.: Spatiotemporal spread of perturbations in a driven dissipative Duffing chain: an out-of-time-ordered correlator approach. Phys. Rev. E 102(5), 052103 (2020)

2. Chierchia, L.: Kolmogorov-Arnold-Moser (KAM) Theory. Mathematics of Complexity and Dynamical Systems, vol. 1–3, pp. 810–836. Springer, New York (2012)

3. Dieckerhoff, R., Zehnder, E.: Boundedness of solutions via the twist theorem. Ann. Sc. Norm. Super. Pisa 14(1), 79–95 (1987)

4. Graef, J.R.: On the generalized Liénard equation with negative damping. J. Differ. Equ. 12, 34–62 (1972)

5. Jiao, L., Piao, D., Wang, Y.: Boundedness for the general semilinear Duffing equations via the twist theorem. J. Differ. Equ. 252(1), 91–113 (2012)

6. Kiss, I.Z., Zhai, Y.M., Hudson, J.L.: Resonance clustering in globally coupled electrochemical oscillators with external forcing. Phys. Rev. E 77(4), 046204 (2008)

7. Kovaleva, A.: Capture into resonance of coupled Duffing oscillators. Phys. Rev. E 92(2), 022909 (2015)

8. Laederich, S., Levi, M.: Invariant curves and time-dependent potential. Ergod. Theory Dyn. Syst. 11(2), 365–378 (1991)

9. Levinson, N.: On the existence of periodic solutions for second order differential equations with a forcing term. J. Math. Phys. 32, 41–48 (1943)

10. Li, J., Qi, J., Yuan, X.: KAM theorem for reversible mapping of low smoothness with application (2019). arXiv:1910.08214 v1 [math.DS]

11. Liu, B.: Boundedness for solutions of nonlinear Hill’s equations with periodic forcing terms via Moser’s twist theorem. J. Differ. Equ. 79(2), 304–315 (1989)

12. Liu, B.: Boundedness for solutions of nonlinear periodic differential equations via Moser’s twist theorem. Acta Math. Sin. (N.S.) 8(1), 91–98 (1992)

13. Morris, G.R.: A case of boundedness of Littlewood’s problem on oscillatory differential equations. Bull. Aust. Math. Soc. 14(1), 71–93 (1976)

14. Moser, J.: On invariant curves of area-preserving mapping of annulus. Nachr. Akad. Wiss. Gottingen Math. Phys. 2, 1–20 (1962)

15. Moser, J.: Stable and Random Motion in Dynamical Systems. Annals of Mathematics Studies. Princeton University Press, Princeton (1973)

16. Peng, Y., Piao, D., Wang, Y.: Longtime closeness estimates for bounded and unbounded solutions of non-recurrent Duffing equations with polynomial potentials. J. Differ. Equ. 268(2), 513–540 (2020)

17. Pöschel, J.: A lecture on the classical KAM theorem. Proc. Symp. Pure Math. 69, 707–732 (2001)

18. Reuter, G.E.H.: A boundedness theorem for nonlinear differential equations of the second order. Proc. Camb. Philos. Soc. 47, 49–54 (1951)

19. Rüssmann, H.: On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In: Dynamical Systems, Theory and Applications, pp. 598–624. Springer (1975)

20. Rüssmann, H.: On the existence of invariant curves of twist mappings of an annulus. In: Geometric Dynamics, pp. 677–718. Springer (1983)

21. Salamon, D.A.: The Kolmogorov-Arnold-Moser theorem. Math. Phys. Electron. J. 10(3), 1–37 (2004)

22. Sarkar, P., Ray, D.S.: Vibrational antiresonance in nonlinear coupled systems. Phys. Rev. E 99(5), 052221 (2019)

23. Shen, J., Lazarides, N., Hizanidis, J.: Multi-branched resonances, chaos through quasiperiodicity, and asymmetric states in a superconducting dimer. Chaos 30(12), 123127 (2020)

24. Wang, Y.: Unboundedness in a Duffing equation with polynomial potentials. J. Differ. Equ. 160(2), 467–479 (2000)

25. Yuan, X.: Invariant tori of Duffing-type equations. J. Differ. Equ. 142(2), 231–262 (1998)

26. Yuan, X.: Lagrange stability for Duffing-type equations. J. Differ. Equ. 160(1), 94–117 (2000)

27. Yuan, X.: Boundedness of solutions for Duffing equation with low regularity in time. Chin. Ann. Math. Ser. B 38(5), 1037–1046 (2017)

28. Yuan, X., Chen, L., Li, J.: The KAM theorem with a large perturbation and application to the network of Duffing oscillators. Sci. China Math. 66(3), 457–474 (2023)

29. Zehnder, E.: Generalized implicit function theorems with applications to some small divisor problems I and II. Comm. Pure Appl. Math. 28, 91–140 (1975); 29(1), 49–111 (1976)
Publisher's Note  Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.