Kolmogorov’s Theorem for Degenerate Hamiltonian Systems with Continuous Parameters

Jiayin Du\textsuperscript{a,}\textsuperscript{*}, Yong Li\textsuperscript{b,}\textsuperscript{2}, Hongkun Zhang\textsuperscript{c,3}

\textsuperscript{a}College of Mathematics and Statistics, and Center for Mathematics and Interdisciplinary Sciences, Northeast Normal University, Changchun, 130024, P. R. China.
\textsuperscript{b}College of Mathematics, Jilin University, Changchun, 130012, P. R. China.
\textsuperscript{c}Department of Mathematics and Statistics, University of Massachusetts, Amherst, 01003, USA.

Abstract

In this paper, we investigate Kolmogorov type theorems for small perturbations of degenerate Hamiltonian systems. These systems are index by a parameter $\xi$ as $H(y, x, \xi) = \langle \omega(\xi), y \rangle + \varepsilon P(y, x, \xi, \varepsilon)$ where $\varepsilon > 0$. We assume that the frequency map, $\omega$, is continuous with respect to $\xi$. Additionally, the perturbation function, $P(y, x, \cdot, \varepsilon)$, maintains Hölder continuity about $\xi$.

We prove that persistent invariant tori retain the same frequency as those of the unperturbed tori, under certain topological degree conditions and a weak convexity condition for the frequency mapping. Notably, this paper presents, to our understanding, pioneering results on the KAM theorem under such conditions—with only assumption of continuous dependence of frequency mapping $\omega$ on the parameter.

Keywords: Hamiltonian system, invariant tori, frequency-preserving, Kolmogorov’s theorem, degeneracy, continuous parameter.

2020 MSC: 37J40, 70H08, 70K43

Contents

1 Introduction ........................................... 2
  1.1 Degeneracy ........................................... 3
  1.2 Regularity ........................................... 3
  1.3 Our work ........................................... 4

2 Main results ........................................... 4

3 KAM step ........................................... 8
  3.1 Description of the 0-th KAM step ....................... 8
  3.2 Induction from $\nu$-th KAM step ....................... 9

\textsuperscript{*}Corresponding author

E-mail address : dujy668@nenu.edu.cn
\textsuperscript{2}E-mail address : liyong@jlu.edu.cn
\textsuperscript{3}E-mail address : hongkunz@umass.edu

Preprint submitted to a June 9, 2024
1. Introduction

As a conservation law of energy, Hamiltonian systems are frequently considered to describe models arose in celestial mechanics or the motion of charged particles in magnetic fields, see [9, 24, 38].

The classical KAM theory, as presented by Arnold, Kolmogorov, and Moser [1, 20, 26], posits that under the Kolmogorov non-degenerate condition, most invariant tori of an integrable Hamiltonian system can withstand small perturbations. While these tori might undergo minor deformations, they transform into other invariant tori that retain the original frequency.

Numerous methods have been explored to study the persistence of invariant tori and the preservation of toral frequency within Hamiltonian systems under certain non-degenerate conditions. For instance, the KAM approach was used in [2, 4, 12, 22, 30, 35]. The direct method using Lindstedt series can be found in references [8, 13, 15], while renormalization group techniques were discussed in [5, 16]. Notably, the study presented in [10] introduced the idea of partial preservation of unperturbed frequencies and delved into the persistence problem on a specified smooth sub-manifold for real analytic Hamiltonian systems, particularly under the Rüssmann-like non-degenerate condition. For insights under analogous conditions, see also [37].

Yet, in the context of persistence, two fundamental questions emerge that warrant attention:

\textbf{Q1}: In the event of a failure in the Kolmogorov non-degenerate condition, can the invariant tori with the same frequency still be preserved under small perturbations?

\textbf{Q2}: If the regularity of the frequency mapping diminishes to mere continuity, can the aforementioned result withstand small perturbations?

To shed light on these questions, we review previous findings and offer a more comprehensive overview.
1.1. Degeneracy

Consider the real analytic nearly integrable Hamiltonian system

$$H(y, x, \varepsilon) = h(y) + \varepsilon P(y, x, \varepsilon),$$

(1.1)

where $x$ is the angle variable in the standard torus $\mathbb{T}^n$, $n$ refers to the dimension; $y$ is the action variable in a bounded closed region $G \subset \mathbb{R}^n$, and $\varepsilon > 0$ is a small parameter.

A fundamental assumption in historical research is the Kolmogorov non-degenerate condition. However, if we assume that there exists a $y_0 \in G$ such that,

$$\det \frac{\partial^2 h(y_0)}{\partial y^2} = 0,$$

then the Kolmogorov condition is not satisfied. The spatial solar system serves as a prominent example of this situation, as detailed in [14]. Naturally, a question arises: does the persistence result still stand under these conditions? This question has been a primary motivation for this research.

In fact, even under weaker non-degenerate conditions, KAM tori might not preserve their frequencies. As demonstrated in [6, 34, 36], under the Brjuno non-degenerate condition and Rüssmann non-degenerate condition, the presumption of an unchanged frequency may not necessarily hold true. This is because the frequency of perturbed tori can undergo slight variations. Similar observations are noted in [3, 7, 10, 11, 18, 32, 39]. Consequently, deriving conditions that assure the persistence of frequencies for KAM tori in the context of a degenerate Hamiltonian becomes rather challenging. Furthermore, the issue of the perturbed invariant tori maintaining a consistent frequency has seldom been tackled for degenerate systems.

1.2. Regularity

On the matter of regularity, it’s worth noting the distinctions in the studies of various researchers. Kolmogorov [20] and Arnold [1] focused on real analytic Hamiltonian systems. In contrast, Moser [26] illustrated that Hamiltonian systems don’t necessarily need to be analytic; a high, albeit finite, level of regularity for the Hamiltonian suffices. This regularity requirement was later reduced to $C^3$ in work by [33]. Further important contributions on this topic can be found in [4, 21, 19, 35]. Moreover, the scenario where the frequency mapping has Lipschitz continuous parameters has been explored in [29]. A subsequent question of interest is: what are the implications when the regularity of the frequency mapping is merely continuous with respect to its parameters?

More precisely, we consider a family of Hamiltonian systems under small perturbations:

$$H(y, x, \xi, \varepsilon) = \langle \omega(\xi), y \rangle + \varepsilon P(y, x, \xi, \varepsilon),$$

(1.2)

where $(y, x) \in G \times \mathbb{T}^n$ and $\xi$ is a parameter in a bounded closed region $O \subset \mathbb{R}^n$. The function $\omega(\cdot)$ is continuous with respect to $\xi$ on $O$. The function $P(\cdot, \cdot, \xi, \varepsilon)$ is real analytic with respect to $y$ and $x$, and $P(y, x, \cdot, \varepsilon)$ is Hölder continuous with respect to the parameter $\xi$ with Hölder index $\beta$, for some $0 < \beta < 1$. Additionally, $\varepsilon > 0$ is a small parameter.

It’s important to note that in the conventional KAM iteration process, the regularity of the frequency mapping concerning the parameters must be at least Lipschitz continuous. This ensures that the parameter domain remains intact. However, when the regularity of the frequency mapping is less stringent than Lipschitz continuous, the traditional method of parameter excavation becomes infeasible. This necessitates the exploration of novel approaches to address the issue.
1.3. Our work

Regarding regularity, when the frequency mapping is continuous with respect to parameters, we prove that the perturbed invariant tori retain the same Diophantine frequency as their unperturbed counterparts for Hamiltonian systems as described in (1.2), see Theorem 1. For the degeneracy problem, persistence results under the highly degenerate Hamiltonian system (1.1) are proved in Theorem 2.

We establish sufficient conditions based on the topological degree condition (A0) and the weak convexity condition (A1) for frequency mapping. Detailed descriptions of these conditions are provided in Section 2. In deriving our primary results, we employ the quasi-linear KAM iteration procedure as in [10, 17, 23, 31]. Notably, we introduce a parameterized family Hamiltonian systems to counteract frequency drift. Specifically, we adjust the action variable to maintain constant frequency for the highly degenerate Hamiltonian system (1.1). It’s also noteworthy that the weak convexity condition proposed in this paper is necessary regardless of the smoothness level of the frequency mapping, as evidenced by Proposition 1.

It should be pointed out that the KAM-type theorems associated with parameter family are due to Moser [25], Pöschel [28]. However, our results are different from theirs: a Diophantine frequency can be given in advance, but Moser’s systems need to be modified in KAM iteration and hence cannot be given beforehand; in Pöschel’s approach, the frequency set need to be dug out in KAM process. Our method is to find a parameter in the family of systems by translating parameter. Of course, it does not work generally. As pointed out in our paper, the weak convexity condition (A1) is indispensable. To our knowledge, this setting seems to be first.

The rest of this paper is organized as follows. In Section 2, we state our main results (Theorems 1, 2, 3 and 4). We will describe the quasi-linear iterative scheme, show the detailed construction and estimates for one cycle of KAM steps in Section 3. In Section 4, we complete the proof of Theorem 1 by deriving an iteration lemma and showing the convergence of KAM iterations. In Section 5, we prove Theorem 2, which covers the analytic situation, and is also a special case of Theorem 1. We also prove Theorem 4 by directly computing. Finally, the proof of Theorem 3 can be found in Appendix B.

2. Main results

To state our main results we need first to introduce a few definitions and notations.

(1) Given a domain \( D \subset G \times \mathbb{T}^n \), we let \( \overline{D} \), \( \partial D \) denote the closure of \( D \) and the boundary of \( D \), respectively. \( D^r := \overline{D} \setminus \partial D \) refers to the interior.

(2) We shall use the same symbol \( |\cdot| \) to denote an equivalent vector norm and its induced matrix norm, absolute value of functions, etc, and use \( |\cdot|_D \) to denote the supremum norm of functions on a domain \( D \).

(3) For the perturbation function \( P(y, x, \xi) \), which is analytic about \( y \) and \( x \) and Hölder continuous about \( \xi \) with Hölder index \( \beta, 0 < \beta < 1 \), we define its norm as follows

\[
\| P \|_0 = \| P \|_0 + \| P \|_\beta
\]

where

\[
\| P \|_\beta \equiv \sup_{\xi \in \mathbb{C}, \xi \neq 0} \frac{|P(y, x, \xi) - P(y, x, \xi')|}{|\xi - \xi'|^\beta}, \quad \forall (y, x) \in D. \quad (2.3)
\]
(4) For any two complex column vectors $\xi, \eta$ in the same space, $\langle \xi, \eta \rangle$ always stands for $\xi^\top \eta$.

(5) $I_d$ is the unit map, and $I_d$ is the unit matrix.

(6) For a vector value function $f$, $Df$ denotes the Jacobian matrix of $f$, and $J_f = detDf$ its Jacobian determinant.

(7) All Hamiltonian in the sequel are endowed with the standard symplectic structure.

(8) As pointed out in [30], the real analyticity of the Hamiltonian $H(y, x)$ about $y$ and $x$ on $G \times \mathbb{T}^n$ implies that the analyticity extends to a complex neighbourhood $D(s, r)$ of $G \times \mathbb{T}^n$, where $D(s, r)$ is defined for some $0 < s, r < 1$, with

$$D(s, r) := \{(y, x) : dist(y, G) < s, |\Im x| < r\}.$$ 

(9) For $\forall \delta > 0, y_0 \in G$, let

$$B_\delta(y_0) := \{y \in G : |y - y_0| < \delta\},$$

$$\bar{B}_\delta(y_0) := \{y \in G : |y - y_0| \leq \delta\}.$$

Let $\Omega \subset \mathbb{R}^n$ be a bounded and open domain. We first give the definition of the degree for $f \in C^2(\bar{\Omega}, \mathbb{R}^n)$, see [27].

**Definition 2.1.** If $f \in C^2(\bar{\Omega}, \mathbb{R}^n)$, and $p \in \mathbb{R}^n \setminus f(\partial \Omega)$, let $\varsigma = \inf_{x \in \partial \Omega} ||f(x) - p||$.

(1) Denote $N_f = \{x | x \in \Omega, J_f(x) = 0\}$. If $p \notin f(N_f)$, then

$$\deg(f, \Omega, p) := \sum_{x \in f^{-1}(p)} \text{sign} \left( J_f(x) \right),$$

setting $\deg(f, \Omega, p) = 0$ if $f^{-1}(p) = \emptyset$.

(2) If $p \in f(N_f)$, by Sard theorem (see [27]), $\deg(f, \Omega, p)$ is defined as $\deg(f, \Omega, p) := \deg(f, \Omega, p_1)$, for some (and all) $p_1 \notin f(N_f)$ such that $||p_1 - p|| < \frac{\varsigma}{\varsigma}$. We next extend the definition of degree to the general continuous mapping, see [27].

**Definition 2.2.** (Brouwer’s Degree) Let $g \in C(\bar{\Omega}, \mathbb{R}^n)$, and $p \in \mathbb{R}^n \setminus g(\partial \Omega)$, i.e., $\varsigma_* := \inf_{x \in \partial \Omega} ||g(x) - p|| > 0$. Let

$$S := \left\{ f \in C^2(\bar{\Omega}, \mathbb{R}^n), \max_{x \in \Omega} \left(||g(x) - f(x)|| < \varsigma_*\right) \right\}.$$ 

Define the degree of $g$ as :

$$\deg(g, \Omega, p) := \deg(f, \Omega, p), \quad \forall f \in S.$$

We are now ready to state our assumptions. Mainly we consider (1.2), i.e., for any $\varepsilon > 0$ small enough, we consider the parameterized family of perturbed Hamiltonian equations

$$\begin{cases} H : G \times \mathbb{T}^n \times O \to \mathbb{R}^1, \\ H(y, x, \xi) = \langle \omega(\xi), y \rangle + \varepsilon P(y, x, \xi, \varepsilon). \end{cases}$$

First, we make the following assumptions:
(A0) Fix $\xi_0 \in O^o$ such that
\[
\deg (\omega(\cdot), O^o, \omega(\xi_0)) \neq 0.
\] (2.4)

(A1) There are $\sigma > 0$, $0 < L \leq \beta$, ($\beta$ was defined in (2.3)), such that
\[
|\omega(\xi) - \omega(\xi_0)| \geq \sigma |\xi - \xi_0|^\frac{1}{2}, \quad \forall \xi, \xi_0 \in O.
\] (2.5)

(A2) For the given $\xi_0 \in O^o$, $\omega(\xi_0)$ satisfies the Diophantine condition
\[
|\omega(\xi_0)| > \frac{\gamma}{|k|^\tau}, \quad k \in \mathbb{Z}^n \setminus \{0\},
\] (2.6)
where $k = (k_1, \cdots, k_n)$, $|k| = |k_1| + \cdots + |k_n|$, $\gamma > 0$ and $\tau > n - 1$.

Then, we have the following main results:

**Theorem 1.** Consider Hamiltonian system (1.2). Assume that (A0), (A1) and (A2) hold. Then there exists a sufficiently small $\varepsilon_0 > 0$, for any $0 < \varepsilon < \varepsilon_0$, there exist $\xi_0 \in O$ and a symplectic transformation $\Psi$, such that
\[
H(\Psi, (y, x, \xi_0, \varepsilon)) = e_\ast + \langle \omega(\xi_0), y \rangle + \tilde{h}_\ast(y, \xi_0) + P_\ast(y, x, \xi_0, \varepsilon),
\]
where $e_\ast$ is a constant, $\tilde{h}_\ast(y, \xi_0) = O(\|y\|^2)$, $P_\ast = O(\|y\|^2)$. Thus the perturbed Hamiltonian system $H(y, x, \xi_0, \varepsilon)$ admits an invariant torus with frequency $\omega(\xi_0)$.

**Remark 2.1.** It should be emphasized that we deal with the degenerate Hamiltonian system in which the frequency mapping is continuous about parameters and the perturbation is Hölder continuous about parameters in this theorem. It seems to be the first version in KAM theory.

In the following, we will give some examples to state that conditions (A0) and (A1) are indispensable, especially for condition (A1). See below for a counter example:

**Proposition 1.** Consider the Hamiltonian system (1.2), for $n = 2$, with
\[
\omega(\xi) = (\omega_1(\xi_1), \omega_2(\xi_2))^\top, \quad eP = P_0(\varepsilon)y_2,
\]
where
\[
\begin{align*}
\omega_1(\xi_1) &= \tilde{\omega}_1 + \xi_1, \quad \xi_1 \in (-1, 1), \\
\omega_2(\xi_2) &= \begin{cases} 
\tilde{\omega}_2 + \exp(-\frac{1}{\xi_2 + \frac{1}{2}}), & \xi_2 \in (-1, -\frac{1}{2}], \\
\tilde{\omega}_2, & \xi_2 \in [-\frac{1}{2}, \frac{1}{2}], \\
\tilde{\omega}_2 - \exp(-\frac{1}{\xi_2 + \frac{1}{2}}), & \xi_2 \in (\frac{1}{2}, 1),
\end{cases}
\end{align*}
\]
\[
\tilde{\omega} = (\tilde{\omega}_1, \tilde{\omega}_2)^\top
\]
satisfies Diophantine condition (2.6), and
\[
P_0(\varepsilon) = \begin{cases} 
0, & \varepsilon = 0, \\
\varepsilon^\ell \sin \frac{\ell}{\pi}, & \varepsilon \neq 0, \ell \in \mathbb{Z}^* \setminus \{0\}.
\end{cases}
\]
Then condition (A1) fails for any parameter $\xi \in (-1, 1)$. Moreover, Theorem 1 fails.
Remark 2.2. This counter example implies that (A1) is necessary no matter how smooth the frequency mapping \( \omega(\xi) \) is.

Nevertheless, one asks what happens to the frequency mapping in the analytic situation. As a special case of our Theorem 1, we also obtain the Kolmogorov’s theorem for analytic Hamiltonian systems under degenerate conditions. This is stated in the following theorem.

**Theorem 2.** Consider the real analytic Hamiltonian system (1.1). Fix \( \xi_0 \in G \) such that (A0), (A1) and (A2) hold for \( \omega(\xi) = \nabla h(\xi), \ O = G, \ ) and \( L > 0. \) Then there exist a sufficiently small positive constant \( \varepsilon' > 0 \) such that if \( 0 < \varepsilon < \varepsilon' \), there exists \( \gamma_\varepsilon \in G \) such that Hamiltonian system (1.1) at \( y = \gamma_\varepsilon \) admits an invariant torus with frequency \( \nabla h(\xi_0) \).

This theorem is proved in Section 5.1.

Next, we will give an example that satisfies conditions (A0)-(A1). For simplicity we use the action variable \( y \) as the parameter \( \xi \).

**Theorem 3.** Consider the Hamiltonian system (1.1) with
\[
h(y) = \langle \omega, y \rangle + \frac{1}{2l+2} |y|^{2l+2},
\]
where \( y \in G \subset \mathbb{R}^n, \ l \) is a positive integer, \( \omega \in \mathbb{R}^n \setminus \{0\} \) satisfies the Diophantine condition (2.6). Then the Hamiltonian system (1.1) admits an invariant torus with frequency \( \omega \) for any small enough perturbation.

The proof can be found in Appendix B.

**Proposition 2.** If \( h(y) = \langle \omega, y \rangle + \frac{1}{2l+2} |y|^{2l+1} \) in Hamiltonian system (1.1), \( \omega \in \mathbb{R}^n \setminus \{0\} \) satisfies the Diophantine condition (2.6), then the system may not admit torus with frequency \( \omega \).

The proof can be found in Appendix C.

Above results imply that condition (A0) is indispensable for \( n > 1 \) case. Furthermore, we also prove that for \( n = 1 \), the persistence results in Theorem 3 hold under some weaker conditions, provided that the frequency satisfies Diophantine condition (A2).

**Theorem 4.** Consider Hamiltonian (1.1) with
\[
h(y) = \omega y + g(y), \quad \varepsilon P(y, x, \varepsilon) = \varepsilon P(y),
\]
where \( y \in G = [-1, 1] \subset \mathbb{R}^1, \ \omega \) satisfies Diophantine condition (2.6).

1. If \( g(y) \in C^{2l+1}, \ g'(0) = \cdots = g^{2l}(0) = 0, \ g^{2l+1}(0) \neq 0, \) \( \ell \) is a positive integer, then the perturbed system admits at least two invariant tori with frequency \( \omega \) for the small enough perturbation satisfying \( \varepsilon P'(y) \) \( \text{sign}(g^{2l+1}(0)) < 0; \) conversely, if \( \varepsilon P'(y) \) \( \text{sign}(g^{2l+1}(0)) > 0, \) the unperturbed invariant torus with frequency \( \omega \) will be destroyed.

2. If \( g(y) \in C^{2l+2}, \ g'(0) = \cdots = g^{2l+1}(0) = 0, \ g^{2l+2}(0) \neq 0, \) \( \ell \) is a positive integer, then the perturbed system admits an invariant tori with frequency \( \omega \) for any small enough perturbation.

Remark 2.3. We don’t know whether the results in Theorem 4 can be extended to higher dimensions or not.
3. KAM step

In this section, we will describe the quasi-linear iterative scheme, show the detailed construction and estimates for one cycle of KAM steps, which is essential to study the KAM theory, see [10, 17, 22, 23, 30]. It should be pointed out that in our KAM iteration, we present a new way to move parameters; while in the usual KAM iteration, one has to dig out a decreasing series of parameter domains, see [10, 17, 23, 29, 30, 31, 32].

3.1. Description of the 0-th KAM step.

Given an integer \( m > L + 1 \), where \( L \) was defined as in (A1). Denote \( \rho = \frac{1}{\sqrt[4]{m+1}} \), and let \( \eta > 0 \) be an integer such that \( (1 + \rho)\eta > 2 \). We define

\[
y = e^{\frac{\eta}{\mu_0}},
\]

(3.7)

Consider the perturbed Hamiltonian (1.2). We first define the following 0-th KAM step parameters:

\[
r_0 = r, \quad \gamma_0 = \gamma, \quad e_0 = 0, \quad \tilde{h}_0 = 0, \quad \mu_0 = \frac{1}{\sqrt[4]{m+1}}, \quad (3.8)
\]

\[
s_0 = \frac{s_0}{16(M^* + 2)K^*}, \quad O_0 = \{ \xi \in O | [\xi - \xi_0] < \text{dist}(\xi_0, \partial O) \},
\]

\[
D(s_0, r_0) := \{ (y,x) : \text{dist}(y,G) < s_0, |\text{Im}x| < r_0 \},
\]

where \( 0 < s_0, \gamma_0, \mu_0 \leq 1 \), \( \tau > n-1 \), \( M^* > 0 \) is a constant defined as in Lemma 3.3, and

\[
K_1 = ((\log \frac{1}{\mu_0}) + 1)^{\eta}.
\]

Therefore, we can write

\[
H_0 =: H(y, x, \xi_0) = N_0 + P_0, \quad N_0 =: N_0(y, \xi_0, \varepsilon) = e_0 + \langle \omega(\xi_0), y \rangle + \tilde{h}_0, \quad P_0 =: \varepsilon P(y, x, \xi_0, \varepsilon).
\]

We first prove an important estimate.

**Lemma 3.1.**

\[
\|P_0\|_{D(s_0, r_0)} \leq \gamma_0^{\frac{\eta}{\mu_0}} s_0^\frac{1}{\mu_0} \mu_0.
\]

(3.9)

**Proof.** Using the fact \( \gamma_0^{\frac{\eta}{\mu_0}} = \varepsilon^\frac{\eta}{\mu_0} \) and \( \log \frac{1}{\mu_0} + 1 < \frac{1}{\mu_0} \), we have

\[
s_0^\frac{1}{\mu_0} = \frac{s^\frac{1}{\mu_0} e^{s^\frac{1}{\mu_0}}}{16^m(M^* + 2)^m K^*} > \frac{s^\frac{1}{\mu_0} e^{s^\frac{1}{\mu_0}}}{16^m(M^* + 2)^m K^*} \geq \frac{s^\frac{1}{\mu_0} e^{s^\frac{1}{\mu_0}}}{16^m(M^* + 2)^m}.
\]

Moreover, let \( \varepsilon_0 > 0 \) be small enough so that

\[
\varepsilon_0^{\frac{1}{\mu_0} - \frac{1}{\mu_0}} \|P_0\|_{D(s_0, r_0)} \frac{16^m(M^* + 2)^m}{s^m} \leq 1,
\]

(3.10)
using the fact that \( \mu_0 = \varepsilon_1 \), we get
\[
\gamma_{0}^{m+2} s_{0}^{m} \mu_0 \geq \frac{s_{0}^{m} e_{m}(\varepsilon_1, 1/2 \pm \delta)}{16^m(M^* + 2)^m} \geq \frac{s_{0}^{m} e_{m}(\varepsilon_1, 1/2 \pm \delta)}{16^m(M^* + 2)^m}
\]
and by \((3.10)\) and \(0 < \varepsilon < \varepsilon_0\),
\[
\varepsilon^{\frac{1}{2}} \frac{s_{0}^{m} e_{m}(\varepsilon_1, 1/2 \pm \delta)}{16^m(M^* + 2)^m} \leq 1,
\]
i.e.,
\[
\varepsilon^{\frac{1}{2}} \frac{s_{0}^{m} e_{m}(\varepsilon_1, 1/2 \pm \delta)}{16^m(M^* + 2)^m} \leq 1.
\]
Then by \((3.11)\) and \((3.12)\),
\[
\|P_0\|_{D(s_{0}, r_{0})} \leq \varepsilon^{\frac{1}{2}} \frac{s_{0}^{m} e_{m}(\varepsilon_1, 1/2 \pm \delta)}{16^m(M^* + 2)^m} \leq \gamma_{0}^{m+2} s_{0}^{m} \mu_0,
\]
which implies \((3.9)\).

The proof is complete.

3.2. Induction from \(\nu\)-th KAM step

3.2.1. Description of the \(\nu\)-th KAM step

We now define the \(\nu\)-th KAM step parameters:
\[
r_{\nu} = \frac{r_{\nu-1}}{2} + \frac{r_{0}}{4}, \quad s_{\nu} = \frac{1}{16^m(M^* + 2)^m}, \quad \mu_{\nu} = \frac{1}{8^m(M^* + 2)^m},
\]
where \(\rho = \frac{1}{2(\nu+1)}\).

Now, suppose that at \(\nu\)-th step, we have arrived at the following real analytic Hamiltonian:
\[
H_{\nu} = N_{\nu} + P_{\nu},
N_{\nu} = e_{\nu} + \langle \Omega(0), y \rangle + h_{\nu}(y, \xi),
\]
defined on \(D(s_{\nu}, r_{\nu})\) and
\[
\|P_{\nu}\|_{D(s_{\nu}, r_{\nu})} \leq \gamma_{0}^{m+2} s_{\nu}^{m} \mu_{\nu}.
\]
The equation of motion associated to \(H_{\nu}\) is
\[
\begin{align*}
\dot{y}_{\nu} &= -\partial_{v} H_{\nu}, \\
\dot{x}_{\nu} &= \partial_{v} H_{\nu}.
\end{align*}
\]

Except for additional instructions, we will omit the index for all quantities of the present KAM step (at \(\nu\)-th step) and use \(+\) to index all quantities (Hamiltonian, domains, normal form, perturbation, transformation, etc.) in the next KAM step (at \((\nu + 1)\)-th step). To simplify the notations, we will not specify the dependence of \(P, P_{\nu}\) etc. All the constants \(c_{1} - c_{6}\) below are
positive and independent of the iteration process, and we will also use $c$ to denote any intermediate positive constant which is independent of the iteration process.

Define

$$r_+ = \frac{r + r_0}{4},$$

$$s_+ = \frac{1}{8} \alpha s, \quad \alpha = \mu^{2\nu} = \mu^{\pi \nu},$$

$$\mu_+ = 8^n c_0 \mu^{1 + \nu}, \quad c_0 = \max\{1, c_1, c_2, \cdots, c_6\},$$

$$K_+ = ([\log \frac{1}{\mu}] + 1)^{\nu};$$

$$\dot{D} = D(s, r_+ + \frac{7}{8}(r - r_+)),$$

$$\dot{D} = D(s, r_+ + \frac{6}{8}(r - r_+)),$$

$$D(s) = \{y \in \mathbb{C}^n : |y| < s\},$$

$$D_{\alpha} = D(i^{\alpha} s, r_+ + \frac{i - 1}{8}(r - r_+)), \quad i = 1, 2, \cdots, 8,$$

$$D_\nu = D_{\alpha} = D(s, r_+),$$

$$O_\nu = \{\xi : dist(\xi, O) \leq \mu^\frac{1}{\nu}\},$$

$$\Gamma(r - r_+) = \sum_{0 \leq |\beta| \leq K} |\beta|^{r+s} e^{-|\beta|^{\frac{1}{2}}};$$

3.2.2. Construct a symplectic transformation

We will construct a symplectic coordinate transformation $\Phi_+$:

$$\Phi_+ : (y_+, x_+) \in D(s_+, r_+) \rightarrow \Phi_+(y_+, x_+) = (y, x) \in D(s, r)$$  \hspace{1cm} (3.16)

such that it transforms the Hamiltonian (3.13) into the Hamiltonian of the next KAM cycle (at $(\nu + 1)$-th step), i.e.,

$$H_+ = H \circ \Phi_+ = N_+ + P_+,$$  \hspace{1cm} (3.17)

where $N_+$ and $P_+$ have similar properties as $N$ and $P$ respectively on $D(s_+, r_+)$, and the equation of motion (3.15) is changed into

$$\begin{cases}
\dot{y}_+ = -\partial_{x_+} H_+,
\dot{x}_+ = \partial_{y_+} H_+.
\end{cases}$$  \hspace{1cm} (3.18)

In the following, we prove (3.18). Let $\Phi_+(y_+, x_+) := (\Phi_+^1(y_+, x_+), \Phi_+^2(y_+, x_+))$, by (3.16), we have

$$\begin{bmatrix}
\dot{y}_+ \\
\dot{x}_+
\end{bmatrix} = \begin{bmatrix}
(\partial_{y_+} \Phi_+^1) y_+ & (\partial_{x_+} \Phi_+^1) x_+ \\
(\partial_{y_+} \Phi_+^2) y_+ & (\partial_{x_+} \Phi_+^2) x_+
\end{bmatrix} D \Phi_+ \begin{bmatrix}
\dot{y}_+ \\
\dot{x}_+
\end{bmatrix},$$

by (3.16) and (3.17), we get

$$\begin{bmatrix}
\partial_{y_+} H \\
\partial_{x_+} H
\end{bmatrix} = \begin{bmatrix}
\partial_{y_+} H \partial_{y_+} y_+ & \partial_{x_+} H \partial_{y_+} y_+ \\
\partial_{y_+} H \partial_{x_+} y_+ & \partial_{x_+} H \partial_{x_+} x_+
\end{bmatrix} = \begin{bmatrix}
\partial_{y_+} \Phi_+^1 & \partial_{y_+} \Phi_+^2 \\
\partial_{x_+} \Phi_+^1 & \partial_{x_+} \Phi_+^2
\end{bmatrix} \begin{bmatrix}
\partial_{y_+} H \\
\partial_{x_+} H
\end{bmatrix}.$$
Then this together with (3.15) yields
\[
\begin{pmatrix}
y' \\
x'
\end{pmatrix} = \Phi_+^{-1} \begin{pmatrix}
y \\
x
\end{pmatrix} = \Phi_+^{-1} J \begin{pmatrix}
\frac{\partial_i H}{\partial y} \\
\frac{\partial_i H}{\partial x}
\end{pmatrix} = \Phi_+^{-1} J (\Phi_+^{-1})^\top \begin{pmatrix}
\frac{\partial_i H}{\partial y} \\
\frac{\partial_i H}{\partial x}
\end{pmatrix},
\]
where \( J \) is the standard symplectic matrix, i.e.,
\[
J = \begin{pmatrix}
0 & -I_d \\
I_d & 0
\end{pmatrix}.
\]
This finishes the proof of (3.18).

Next, we show the detailed construction of \( \Phi_+ \) and the estimates of \( P_+ \).

3.2.3. Truncation

Consider the Taylor-Fourier series of \( P \):
\[
P = \sum_{k \in \mathbb{Z}^n, \ell \in \mathbb{Z}^m} p_{k\ell} y^\ell e^{\sqrt{-1} \langle k, x \rangle},
\]
and let \( R \) be the truncation of \( P \) of the form
\[
R = \sum_{|k| > K, |\ell| > m} p_{k\ell} y^\ell e^{\sqrt{-1} \langle k, x \rangle}.
\]
Next, we will prove that the norm of \( P - R \) is much smaller than the norm of \( P \) by selecting truncation appropriately, see the below lemma.

**Lemma 3.2.** Assume that
\[
(H1) : \int_{\mathbb{R}^+} t^m e^{-t/\mu} dt \leq \mu.
\]
Then there is a constant \( c_1 \) such that
\[
\|P - R\|_{D_\alpha} \leq c_1 \gamma_{0}^{p+m+2} s^m \mu^2, \tag{3.19}
\]
\[
\|R\|_{D_\alpha} \leq c_1 \gamma_{0}^{p+m+2} s^m \mu. \tag{3.20}
\]

**Proof.** Denote
\[
I = \sum_{|k| > K, \ell \in \mathbb{Z}^n} p_{k\ell} y^\ell e^{\sqrt{-1} \langle k, x \rangle},
\]
\[
II = \sum_{|k| \leq K, |\ell| > m} p_{k\ell} y^\ell e^{\sqrt{-1} \langle k, x \rangle}.
\]
Then
\[
P - R = I + II.
\]
To estimate $I$, we note by (3.14) that
\[
\left\| \sum_{|\xi| \leq K_1} p_{k_1}e^{i|\xi|y} \right\| \leq |P|_{D_{(1,2)}} e^{-|\xi|y} \leq \gamma_0^{m+2} s^n \mu e^{-|\xi|y},
\] (3.21)
where the first inequality has been frequently used in [10, 11, 17, 22, 30, 31, 32, 35] and the detailed proof see [35]. This together with (H1) yields
\[
|I|_{D_1} \leq \sum_{|\xi| > K_1} \left| \sum_{|p| \leq K} p_{k_1}e^{i|\xi|y} \right| \leq \sum_{|\xi| > K_1} |P|_{D_{(1,1)}} e^{-|\xi|y} \leq \gamma_0^{m+2} s^n \mu \int_{|\xi| > K_1} r^\alpha e^{-i\xi \cdot \nu} dt,
\] (3.22)
It follows from (3.14) and (3.22) that
\[
|P - I|_{D_1} \leq |P|_{D_{(1,1)}} + |I|_{D_1} \leq 2 \gamma_0^{m+2} s^n \mu.
\]
For $|p| = m + 1$, let $\int$ be the obvious antiderivative of $\frac{\partial^p}{\partial y^p}$. Then the Cauchy estimate of $P - I$ on $D_n$ yields
\[
|I|_{D_n} = \left| \int \frac{\partial^p}{\partial y^p} \sum_{|\xi| \leq K_1, |p| = m} p_{k_1}e^{i|\xi|y} dy \right|_{D_n}
\]
\[
\leq \left| \int \frac{\partial^p}{\partial y^p} (P - I) dy \right|_{D_n}
\]
\[
\leq \frac{c}{s^{m+1}} \int |P - I| dy_{D_n}
\]
\[
\leq 2 \frac{c}{s^{m+1}} \gamma_0^{m+2} s^n \mu (\alpha s)^{m+1}
\]
\[
\leq c \gamma_0^{m+2} s^n \mu^2.
\]
Thus,
\[
|P - R|_{D_n} = |I + II|_{D_n} \leq c \gamma_0^{m+2} s^n \mu^2,
\] (3.23)
and therefore,
\[
|R|_{D_1} \leq |P - R|_{D_1} + |P|_{D_{(1,1)}} \leq c \gamma_0^{m+2} s^n \mu.
\] (3.24)
Next, we estimate $||P - R||_{C^\alpha}$. In view of the definition of $|| \cdot ||_{C^\alpha}$, for $y, x \in D_n$, we have
\[
||P - R||_{C^\alpha} = \sup_{\xi \subset \xi} \left| \int \frac{\partial^p}{\partial y^p} (P(x, y, \xi) - R(x, y, \xi) - (P(x, y, \xi) - R(x, y, \xi))) dy \right|
\]
\[
\leq \sup_{\xi \subset \xi} \left| \int \frac{\partial^p}{\partial y^p} (P(x, y, \xi) - R(x, y, \xi) - (P(x, y, \xi) - R(x, y, \xi))) dy \right|
\]
\[
\leq \sup_{\xi \subset \xi} \frac{e^{\alpha |\xi| y}}{|\xi|^{\alpha}} \leq \frac{e^{\alpha |\xi| y}}{|\xi|^{\alpha}}
\]
\[
\begin{align*}
&\leq \sup_{\xi \neq \zeta} \frac{\int |P(x, y, \xi) - P(x, y, \zeta)| dy}{|\xi - \zeta|^p} \\
&\leq \sup_{\xi \neq \zeta} \frac{c |P(x, y, \xi) - P(x, y, \zeta)|}{|\xi - \zeta|^p} \leq c_0 |\xi - \zeta|^{-\beta} \leq c_0 |\xi - \zeta|^{-\beta},
\end{align*}
\]  

where the third inequality follows from Cauchy estimate and the last inequality follows from (3.14).

Similarly, we get
\[
\|R\|_{C^\beta} < \|P - R\|_{C^\beta} + \|P\|_{C^\beta} \leq c_0^{n+m+2} s^m \mu.
\]  

It follows from (3.23), (3.24), (3.25) and (3.26) that (3.19) and (3.20) hold. The proof is complete.

3.2.4. Homological Equation

As usual, we shall construct a symplectic transformation as the time 1-map \(\phi^1_F\) of the flow generated by a Hamiltonian \(F\) to eliminate all resonant terms in \(R\), i.e., all terms \(p_k y^e e^{\sqrt{-1} \langle k, \omega(\xi_0) + \partial_j \tilde{h} \rangle f_{k_i} y^e e^{\sqrt{-1} (k, x)}\), \(0 < |k| \leq K_+, |i| \leq m\).

To do so, we first construct a Hamiltonian \(F\) of the form
\[
F = \sum_{0 < |k| \leq K_+, |i| \leq m} f_{k_i} y^e e^{\sqrt{-1} (k, x)},
\]  

satisfying the equation
\[
\{N, F\} + R - [R] = 0,
\]  

where \([R] = \frac{1}{2\gamma_0} \int_{\mathbb{R}} R(y, x) dx\) is the average of the truncation \(R\).

Substituting (3.27) into (3.28) yields that
\[
\begin{align*}
&- \sum_{0 < |k| \leq K_+, |i| \leq m} \sqrt{-1} \left\langle k, \omega(\xi_0) + \partial_j \tilde{h} \right\rangle f_{k_i} y^e e^{\sqrt{-1} (k, x)} \\
&+ \sum_{0 < |k| \leq K_+, |i| \leq m} p_k y^e e^{\sqrt{-1} (k, x)} = 0.
\end{align*}
\]

By comparing the coefficients above, we then obtain the following quasi-linear equations:
\[
\sqrt{-1} \left\langle k, \omega(\xi_0) + \partial_j \tilde{h} \right\rangle f_{k_i} = p_k, \quad |i| \leq m, \quad 0 < |k| \leq K_+.
\]  

We declare that the quasi-linear equations (3.29) is solvable under some suitable conditions. The details can be seen in the following lemma:

**Lemma 3.3.** Assume that

\[\textbf{(H2)}: \max_{|i| \leq 2} \left\|\partial_i^j \tilde{h} - \partial_i^j \tilde{h}_0\right\|_{D^{(j)}} \leq \mu_0,\]

\[\textbf{(H3)}: 2s < \frac{\gamma_0}{(M^* + 2)K_+^{n+1}},\]
where
\[
M^* = \max_{|\xi| \leq 2} |\partial_\xi^j h_0(\xi)|.
\]

Then the quasi-linear equations (3.29) can be uniquely solved on \(D(s)\) to obtain a family of functions \(f_\xi\) which are analytic in \(y\), and satisfy the following properties:
\[
\|\partial_\xi^j f_\xi\|_{D(s)} \leq c_2 |k|^{(|j|+1)r+|i|} \gamma_0^{m+|j|} s^{-m-|j|} \mu e^{-|k|r},
\]
(3.30)

for all \(|i| \leq m, 0 < |k| \leq K_+\), \(|i| \leq 2\), where \(c_2\) is a constant.

**Proof.** For all \(y \in D(s)\), by (H2), (H3), we have
\[
|\partial_\xi^j h|_{D(s)} = \left| (\partial_\xi^j h - \partial_\xi^j h_0) + \partial_\xi^j h_0 \right|_{D(s)} \leq (1 + M^*)|y| < (1 + M^*)s < \frac{\gamma_0}{2|k|^{r+1}}
\]
and
\[
\|\partial_\xi^j h\|_{C^r} = \sup_{\xi, y} \left| \partial_\xi^j h(y, \xi) - \partial_\xi^j h(y, \xi_0) \right| \leq \mu_0 |y| < \frac{\gamma_0}{2|k|^{r+1}},
\]
which imply that
\[
\|\partial_\xi^j h\|_{D(s)} < \frac{\gamma_0}{2|k|^{r+1}}.
\]
(3.31)

It follows from (3.31) and (A2) that
\[
\left\| \left( k, \omega(\xi_0) + \partial_\xi^j h(y) \right) \right\|_{D(s)} > \frac{\gamma_0}{|k|^2} - \frac{\gamma_0}{2|k|^2} = \frac{\gamma_0}{2|k|^2}.
\]
(3.32)

Hence
\[
L_k = \sqrt{-1} \left( k, \omega(\xi_0) + \partial_\xi^j h(y) \right)
\]
(3.33)
is invertible, and
\[
f_\xi = L_k^{-1} p_\xi,
\]
(3.34)
for all \(y \in D(s), 0 < |k| \leq K_+, |i| \leq m\). Let \(0 < |k| \leq K_+\). We note by the first inequality of (3.21) and Cauchy estimate that
\[
\|P_{L_k}\| \leq \|\partial_\xi^j P\|_{D(s)} e^{-|k|r} \leq \gamma_0^{m+2} s^{-m-|j|} \mu e^{-|k|r}, \quad |i| \leq m,
\]
(3.35)
and by (3.32) and (3.33) that
\[
\|\partial_\xi^j L_k^{-1}\|_{D(s)} \leq c_2 |k|^{(|j|+1)r+|i|} \gamma_0^{m+1} s^{-m-|j|} \mu e^{-|k|r}, \quad |i| \leq 2.
\]
(3.36)

So, by (3.34), (3.35) and (3.36), we get
\[
\|\partial_\xi^j f_\xi\|_{D(s)} \leq c_2 |k|^{(|j|+1)r+|i|} \gamma_0^{m+2} s^{-m-|j|} \mu e^{-|k|r}
\]
\[
= c_2 |k|^{(|j|+1)r+|i|} \gamma_0^{m+1} s^{-m-|j|} \mu e^{-|k|r}, \quad |i| \leq 2.
\]

The proof is complete.
Next, we apply the above transformation $\phi^1_F$ to Hamiltonian $H$, i.e.,

$$H \circ \phi^1_F = (N + R) \circ \phi^1_F + (P - R) \circ \phi^1_F =
$$

$$= (N + R) + [N, F] + \int_0^1 [(1 - t)[N, F] + R, F] \circ \phi^1_F dt
$$

$$+ (P - R) \circ \phi^1_F =
$$

$$= N + [R] + \int_0^1 [R, F] \circ \phi^1_F dt + (P - R) \circ \phi^1_F =: \tilde{N}_e + \tilde{P}_e,$n

where

$$\tilde{N}_e = N + [R] = e_+ + \langle \omega(\xi), y \rangle + \left( \sum_{j=0}^{y} p^1_{01}(\xi), y \right) + \tilde{h}_+(y, \xi),$$

$$e_+ = e + p^0_{00},$$

$$\tilde{h}_+ = \tilde{h}(y, \xi) + [R] - p^0_{00} - \langle p^0_{01}(\xi), y \rangle,$n

$$\tilde{P}_+ = \int_0^1 [R, F] \circ \phi^1_F dt + (P - R) \circ \phi^1_F,$n

$$R_t = (1 - t)[R] + tR.$n

3.2.5. Translation

In this subsection, we will construct a translation so as to keep the frequency unchanged. It should be pointed out that we present a new way to move parameters, but in the usual KAM iteration, one has to dig out a decreasing series of parameter domains, in which the Diophantine condition doesn’t hold, see [10, 17, 23, 29, 30, 31, 32].

Consider the translation

$$\phi: x \rightarrow x, \quad y \rightarrow y, \quad \tilde{\xi} \rightarrow \tilde{\xi} + \xi_+ - \xi,$n

where $\xi_+$ is to be determined. Let

$$\Phi_+ = \phi^1_F \circ \phi.$n

Then

$$H \circ \Phi_+ = N_e + P_e,$n

$$N_e = \tilde{N}_e \circ \phi = e_+ + \langle \omega(\xi_+), y \rangle + \left( \sum_{j=0}^{y} p^1_{01}(\xi_+), y \right) + \tilde{h}_+(y, \xi_+),$$

$$P_e = \tilde{P}_+ \circ \phi.$$n

3.2.6. Frequency-preserving

In this subsection, we will show that the frequency can be preserved in the iteration process. Recall the topological degree condition (A0) and the weak convexity condition (A1). The former ensures that the parameter $\xi_e$ can be found in the parameter set to keep the frequency unchanged at this KAM step. The later assures that the distance between $\xi_e$ and $\xi$ is smaller than the distance between $\xi$ and $\xi_{e-1}$, i.e., the sequence of parameters is convergent after infinite steps of iteration. The following lemma is crucial to our arguments.
Lemma 3.4. Assume that

\[(H4) : \left\| \sum_{j=0}^{\nu} p_{01}^j \right\|_{\Omega(L,Q)} < \mu_0.\]

There exists \(\xi_+ \in B_{c_{\mu_1}}(\xi) \subset O^\nu\) such that

\[\omega(\xi_+) + \sum_{j=0}^{\nu} p_{01}^j(\xi_+) = \omega(\xi_0).\]  \hspace{1cm} (3.43)

Proof. The proof will be completed by an induction on \(\nu\). We start with the case \(\nu = 0\). It is obvious that \(\omega(\xi_0) = \omega(\xi_0)\). Now assume that for some \(\nu > 0\) we have got

\[\omega(\xi_i) + \sum_{j=0}^{i-1} p_{01}^j(\xi_i) = \omega(\xi_0), \quad \xi_i \in B_{c_{\mu_1}}(\xi_{i-1}) \subset O^\nu, \quad i = 1, \cdots, \nu.\]  \hspace{1cm} (3.44)

We need to find \(\xi_+\) near \(\xi\) such that

\[\omega(\xi_+) + \sum_{j=0}^{\nu} p_{01}^j(\xi_+) = \omega(\xi_0).\]  \hspace{1cm} (3.45)

In view of the property of topological degree, \((H4)\) and \((A0)\), we have

\[\deg \left( \omega(\cdot) + \sum_{j=0}^{\nu} p_{01}^j(\cdot), O^\nu, \omega(\xi_0) \right) = \deg (\omega(\cdot), O^\nu, \omega(\xi_0)) \neq 0,\]

i.e., there exists at least a \(\xi_+ \in O^\nu\) such that (3.43) holds.

Next, we estimate \(|\xi_+ - \xi|\). (3.20) in Lemma 3.2 implies that

\[\left\| p_{01}^j \right\|_{C^p} < c\mu_j, \quad j = 0, 1, \cdots, \nu,\]

i.e.,

\[\left| p_{01}^j(\xi_+) - p_{01}^j(\xi) \right| < c\mu_j |\xi_+ - \xi|^p, \quad \forall \xi_+, \xi \in O.\]  \hspace{1cm} (3.46)

According to (3.44) and (3.45), we get

\[\omega(\xi_+) - \omega(\xi) + \sum_{j=0}^{\nu-1} \left( p_{01}^j(\xi_+) - p_{01}^j(\xi) \right) = -p_{01}^\nu(\xi_+).\]  \hspace{1cm} (3.47)

This together with \((A1)\) and (3.46) yields

\[\left| p_{01}^\nu(\xi_+) \right| = \left| \omega(\xi_+) - \omega(\xi) + \sum_{j=0}^{\nu-1} \left( p_{01}^j(\xi_+) - p_{01}^j(\xi) \right) \right| \geq |\omega(\xi_+) - \omega(\xi)| - \sum_{j=0}^{\nu-1} \left| p_{01}^j(\xi_+) - p_{01}^j(\xi) \right|\]
\[ \geq \sigma |\xi_+ - \xi|^l - c \sum_{j=0}^{n-1} \mu_j |\xi_+ - \xi|^l \]

\[ \geq (\sigma - c \sum_{j=0}^{n-1} \mu_j) |\xi_+ - \xi|^l, \quad (3.48) \]

where the last inequality follows from \(0 < L \leq \beta\) in (A1). Then, by (3.48) and (3.20) in Lemma 3.2, we have

\[ |\xi_+ - \xi|^l \leq \frac{p_0(|\xi_+|)}{\sigma} \frac{c \mu}{\sigma - c \sum_{j=0}^{n-1} \mu_j} < \frac{2c \mu}{\sigma}, \]

which implies \( \xi_+ \in B_{p_0}^{\sigma, \xi}(\xi_+)\). From \( \xi \in O'\) in (3.44) and the fact \( \epsilon \) is small enough (i.e., \( \mu \) is small enough), we have \( B_{p_0}^{\sigma, \xi}(\xi) \subset O'\).

The proof is complete. \( \square \)

3.2.7. Estimate on \( N_+ \)

Now, we give the estimate on the next step \( N_+ \).

**Lemma 3.5.** There is a constant \( c_3 \) such that the following holds:

\[ |\xi_+ - \xi|^l \leq c_3 |\xi|^l, \quad (3.49) \]

\[ |e_+ - e|^l \leq c_3 |\xi|^l. \quad (3.50) \]

\[ \|\hat{h}_+ - \hat{h}\|_{\mathcal{D}(0)} \leq c_3 |\xi|^l. \quad (3.51) \]

**Proof.** It is obvious by \( \xi_+ \in B_{p_0}^{\sigma, \xi}(\xi)\) in Lemma 3.4 that (3.49) holds. It follows from (3.38) and (3.39) that (3.50) and (3.51) hold. \( \square \)

3.2.8. Estimate on \( \Phi_+ \)

Recall that \( F \) is as in (3.27) with the coefficients and its estimate given by Lemma 3.3. Then, we have the following estimate on \( F \).

**Lemma 3.6.** There is a constant \( c_4 \) such that for all \( |j| + |i| \leq 2 \),

\[ \|\partial^j_{\xi} \partial^i_{y} F\|_{\mathcal{D}} \leq c_4 \gamma_0^{n+m+1} s^{-|i|} \mu \Gamma(r - r_+). \quad (3.52) \]

**Proof.** By (3.27) and (3.30), we have

\[ \|\partial^j_{\xi} \partial^i_{y} F\|_{\mathcal{D}} \leq \sum_{|j|+|i|=|k|} |k| \gamma_0^{n+m+1} s^{-|i|} \mu \Gamma(r - r_+). \]

The proof is complete. \( \square \)
Lemma 3.7. Assume that

\begin{align*}
\text{(H5)} & : c_4 s^{m-1} \mu \Gamma (r - r_s) < \frac{1}{8} (r - r_s), \\
\text{(H6)} & : c_4 s^m \mu \Gamma (r - r_s) < \frac{1}{8} \alpha s.
\end{align*}

Then the following holds.

(1) For all $0 \leq t \leq 1$, the mappings

\begin{align*}
\phi^t_F : D_{+0} & \to D_{+0}, \\
\phi : O & \to O,
\end{align*}

are well defined.

(2) $\Phi_+ : D_+ \to D(s, r)$.

(3) There is a constant $c_5$ such that

\begin{align*}
\left\| \phi^t_F - id \right\|_D & \leq c_5 \mu \Gamma (r - r_s), \\
\left\| D\phi^t_F - Id \right\|_D & \leq c_5 \mu \Gamma (r - r_s), \\
\left\| D^2 \phi^t_F \right\|_D & \leq c_5 \mu \Gamma (r - r_s).
\end{align*}

(4)

\begin{align*}
\left\| \Phi_+ - id \right\|_D & \leq c_5 \mu \Gamma (r - r_s), \\
\left\| D\Phi_+ - Id \right\|_D & \leq c_5 \mu \Gamma (r - r_s), \\
\left\| D^2 \Phi_+ \right\|_D & \leq c_5 \mu \Gamma (r - r_s).
\end{align*}

Proof. (1) (3.54) immediately follows from (3.49) and the definition of $O_+$. To verify (3.53), we denote $\phi^t_{F, y}, \phi^t_{F, x}$ as the components of $\phi^t_F$ in the $y, x$ planes, respectively. Let $X_F = (F_y, -F_x) \Gamma$ be the vector field generated by $F$. Then

$$
\phi^t_F = id + \int_0^t X_F \circ \phi^u_F \, du, \quad 0 \leq t \leq 1.
$$

For any $(y, x) \in D_{+0}$, we let $t_\ast = \sup \{ \tau \in [0, 1] : \phi^\tau_F (y, x) \in D_0 \}$. Then for any $0 \leq t \leq t_\ast$, by $(y, x) \in D_{+0}$, (3.52) in Lemma 3.6, (H5) and (H6), we can get the following estimates:

\begin{align*}
\left\| \phi^t_{F, y} (y, x) \right\|_{D_{+0}} & \leq |y| + \int_0^t \left\| F_y \circ \phi^u_F \right\|_{D_{+0}} \, du \\
& \leq \frac{1}{4} \alpha s + c_4 s^m \mu \Gamma (r - r_s) \\
& \leq \frac{3}{8} \alpha s,
\end{align*}

\begin{align*}
\left\| \phi^t_{F, x} (y, x) \right\|_{D_{+0}} & \leq |x| + \int_0^t \left\| F_x \circ \phi^u_F \right\|_{D_{+0}} \, du \\
& \leq \frac{1}{18} \alpha s.
\end{align*}
Thus, \( \phi_F' \in D_{\alpha}^{\Phi'} \subset D_{\alpha} \), i.e. \( t_1 = 1 \) and (1) holds.

(2) It follows from (1) that (2) holds.

(3) Using (3.52) in Lemma 3.6 and (3.55), we immediately have

\[
\|\phi_F' - id\|_D \leq c_5 \mu(r - r_0).
\]

By (3.52) in Lemma 3.6, (3.55) and Gronwall Inequality, we get

\[
\|D\phi_F' - Id\|_D \leq \left\| \int_0^T DX_F \circ \phi_F' D\phi_F' d\lambda \right\|_D \\
\leq \int_0^T \|DX_F \circ \phi_F'\|_D \|D\phi_F' - Id\|_D d\lambda + \int_0^T \|DX_F \circ \phi_F'\|_D d\lambda \\
\leq c_3 \mu (r - r_0).
\]

It follows from the induction and a similar argument that we have the estimates on the 2-order derivatives of \( \phi_F' \), i.e.,

\[
\|D^2\phi_F'\|_D \leq c_3 \mu (r - r_0).
\]

(4) now follows from (3).

The proof is complete. \( \square \)

3.2.9. Estimate on \( P_+ \)

In the following, we estimate the next step \( P_+ \).

**Lemma 3.8.** Assume (H1)-(H6). Then there is a constant \( c_6 \) such that,

\[
\|P_+\|_{D_0} \leq c_6 \gamma_0^{n+1} \mu^2 \Gamma^2 (r - r_0 + (r - r_0)).
\] (3.56)

Moreover, if

\[
(\text{H7}) : \mu^2 (\Gamma^2 (r - r_0) + \Gamma (r - r_0)) \leq 1
\]

then,

\[
\|P_+\|_{D_0} \leq \gamma_0^{n+1} \mu^2. \] (3.57)

**Proof.** By (3.19) and (3.20) in Lemma 3.2, (3.52) in Lemma 3.6 and Lemma 3.7 (3), we have that, for all \( 0 \leq t \leq 1 \),

\[
\|\{R_t, F\} \circ \phi_F'\|_{D_{\alpha}} \leq c_7 \gamma_0^{n+1} \mu^2 \Gamma^2 (r - r_0),
\]

\[
\|\{P - R\} \circ \phi_F'\|_{D_{\alpha}} \leq c_7 \gamma_0^{n+1} \mu^2 \Gamma (r - r_0).
\]

So, by (3.40),

\[
\|\tilde{P}_t\|_{D_{\alpha}} \leq c_7 \gamma_0^{n+1} \mu^2 \left( \Gamma^2 (r - r_0) + \Gamma (r - r_0) \right).
\]
By (H7), we see that,
\[
\|P\|_{D_0} \leq 8^m c_0 \mu_1^{1+p} \mu_1^{-2p} \pi \tau y_0^{m+2} \left( \mu^2 \left( \Gamma^2(r - r_+) + \Gamma(r - r_+) \right) \right) 
\leq y_0^{m+2} \mu_1^{2m+2} \mu_2
\]
which implies (3.57).

The proof is complete.

This completes one cycle of KAM steps.

4. Proof of Theorem 1

4.1. Iteration lemma

In this subsection, we will prove an iteration lemma which guarantees the inductive construction of the transformations in all KAM steps.

Let \( r_0, s_0, \gamma_0, \mu_0, H_0, N_0, e_0, \bar{h}_0, P_0 \) be given at the beginning of Section 3 and let \( D_0 = D(s_0, r_0) \), \( K_0 = 0, \Phi_0 = \text{id} \). We define the following sequence inductively for all \( \nu = 1, 2, \cdots \).

\[
\begin{align*}
    r_\nu &= r_0 \left( 1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}} \right), \\
    s_\nu &= \frac{1}{8} \alpha_{\nu-1} s_{\nu-1}, \\
    \alpha_\nu &= \frac{1}{8} \mu_\nu = \frac{1}{8} \mu_{\nu-1}, \\
    \mu_\nu &= 8^m c_0 \mu_1^{1+p}, \\
    K_\nu &= \left( \left\lfloor \log \left( \frac{1}{\mu_{\nu-1}} \right) \right\rfloor + 1 \right) \gamma_\nu, \\
    \tilde{D}_\nu &= D \left( \frac{1}{2} s_\nu, r_\nu + \frac{6}{8} (r_{\nu-1} - r_\nu) \right).
\end{align*}
\]

**Lemma 4.1.** Denote

\[
\mu_\ast = \frac{\mu_0}{(M_0 + 2)^{m-1} K_1^{(\nu+1)(m-1)}},
\]

If \( \varepsilon \) is small enough, then the KAM step described on the above is valid for all \( \nu = 0, 1, \cdots \), resulting the sequences

\[
H_\nu, N_\nu, e_\nu, \bar{h}_\nu, P_\nu, \Phi_\nu,
\]

\( \nu = 1, 2, \cdots \), with the following properties:

(1)

\[
\begin{align*}
    |e_{\nu+1} - e_\nu| &\leq \frac{\mu_\ast^2}{2^{\nu}}, \quad (4.58) \\
    |e_\nu - e_0| &\leq 2 \mu_\ast^2, \quad (4.59)
\end{align*}
\]
\[ \|\tilde{h}_{v+1} - \tilde{h}_v\|_{D(v,s)} \leq \frac{\mu_{v}^{\frac{1}{2}}}{2v}, \quad (4.60) \]
\[ \|\tilde{h}_v - \tilde{h}_0\|_{D(0,s)} \leq 2\mu_{v}, \quad (4.61) \]
\[ \|P_v\|_{D(v,s,v)} \leq \frac{\mu_{v}^{\frac{1}{2}}}{2v}, \quad (4.62) \]
\[ |\xi_{v+1} - \xi_v| \leq \left(\frac{\mu_{v}^{\frac{1}{2}}}{2v}\right)^{\frac{1}{2}}. \quad (4.63) \]

(2) **$\Phi_{v+1} : \tilde{D}_{v+1} \rightarrow \tilde{D}_v$ is symplectic,** and
\[ \|\Phi_{v+1} - id\|_{\tilde{D}_{v+1}} \leq \frac{\mu_{v}^{\frac{1}{2}}}{2v}. \quad (4.64) \]

Moreover, on $D_{v+1}$,
\[ H_{v+1} = H_v \circ \Phi_{v+1} = N_{v+1} + P_{v+1}. \]

**Proof.** The proof amounts to the verification of (H1)-(H7) for all $\nu$. For simplicity, we let $r_0 = 1$. It follows from $\varepsilon$ small enough that $\mu_0$ is small. So, we see that (H2), (H4)-(H7) hold for $\nu = 0$. From (3.8), (H3) holds for $\nu = 0$. According to the definition of $\mu_\nu$, we see that
\[ \mu_\nu = (8^m c_0)^{\frac{1}{1-p}} \frac{1}{\mu_0}. \quad (4.65) \]

Let $\zeta \gg 1$ be fixed and $\mu_0$ be small enough so that
\[ \mu_0 < \left(\frac{1}{8^m c_0 \zeta}\right)^{\frac{1}{2}} < 1. \quad (4.66) \]

Then
\[ \mu_1 = 8^m c_0 \mu_0^{1+p} < \frac{1}{\zeta} \mu_0 < 1, \]
\[ \mu_2 = 8^m c_0 \mu_1^{1+p} < \frac{1}{\zeta^2} \mu_1 < \frac{1}{\mu_0}, \]
\[ \vdots \]
\[ \mu_v = 8^m c_0 \mu_{v-1}^{1+p} < \cdots < \frac{1}{\zeta^v} \mu_0. \quad (4.67) \]

Denote
\[ \Gamma_v = \Gamma(r_v - r_{v+1}). \]

We notice that
\[ \frac{r_v - r_{v+1}}{r_0} = \frac{1}{2v^{r+2}}. \quad (4.68) \]

Since
\[ \Gamma_v \leq \int_1^{r_v} t^{3r+5} e^{-\frac{r_v}{21}} dt. \]
\[ \leq (3^r + 5)!2^{(v+5)(3^r+5)}, \]

it is obvious that if \( \zeta \) is large enough, then

\[ \mu_i^2 \Gamma_i^i \leq \mu_i^2 (\Gamma_i^i + \Gamma_i) \leq 1, \quad i = 1, 2, \]

which implies that \( (H7) \) holds for all \( v \geq 1 \), and

\[ \mu_i \Gamma_i \leq \mu_i^{1-p} \leq \frac{\mu_0^{1-p}}{\zeta^{(1-p)v}}. \quad (4.69) \]

By (4.68) and (4.69), it is easy to verify that \( (H5), (H6) \) hold for all \( v \geq 1 \) as \( \zeta \) is large enough and \( \epsilon \) is small enough.

By (3.20) in Lemma 3.2 and (4.67), we have

\[ \left\| \sum_{j=0}^{v} P_{0j} \right\|_{D(h, r)} < c \sum_{j=0}^{v} \frac{1}{j!}\mu_0 < c\mu_0, \]

which implies \( (H4) \).

To verify \( (H3) \), we observe by (4.65) and (4.67) that

\[ \frac{1}{4} (M^* + 2) \mu_v^{2r_{v+1}} < \frac{1}{2^{v+2}}, \]

as \( \zeta \) is large enough. Then

\[ 2 (M^* + 2) s_v K_{v+1} \leq \frac{s_{v-1}}{4} (M^* + 1) \mu_v^{2r_{v+1}} K_{v+1} \]
\[ \leq \frac{\gamma_0}{2v+2} < \gamma_0, \quad (4.70) \]

which verifies \( (H3) \) for all \( v \geq 1 \).

Let \( \zeta^{1+p} \geq 2 \) in (4.66), (4.67). We have that for all \( v \geq 1 \)

\[ c_0 \mu_v \leq \frac{\mu_0}{2^v} \leq \frac{\mu_1}{2^v}, \quad (4.72) \]

\[ c_0 \mu_v \Gamma_v \leq \frac{\mu_0}{2^v} \leq \frac{\mu_1}{2^v}, \quad (4.73) \]

\[ c_0 \mu_v^{m-1} \mu_v \leq \frac{\mu_0^{1+2r_{m-1}}}{2^{v+3}} \leq \frac{\mu_1}{2^v}. \quad (4.74) \]

The verification of \( (H2) \) follows from (4.72) and an induction application of (3.51) in Lemma 3.5 for all \( v = 0, 1, \ldots \).

Since \( (1 + \rho)^{\nu} > 2 \), we have

\[ \frac{1}{2^{v+6}} \left( \log \frac{1}{\mu} + 1 \right)^q \geq \frac{1}{2^{v+6}} (1 - (1 + \rho)^{\nu}) \log (8^n c_0) - (1 + \rho)^{\nu} \log \mu_0 \]
\[ \geq -\frac{1}{2^{v+6}} (1 + \rho)^{\nu} (\log \mu_0)^{\nu} \geq 1. \]
It follows from the above that
\[
\log (n+1)! + (\nu + 6) n \log 2 + 3n\eta \log \left( \log \frac{1}{\mu} \right) + 1 \leq \log (n+1)! + (\nu + 6) n \log 2 + 3n\eta \log \left( \log \frac{1}{\mu} \right) + 1 \leq -\log \frac{1}{\mu},
\]
as \mu is small, which is ensured by making \( \varepsilon \) small. Thus,
\[
\int_{K_{n+1}}^\infty e^{-\frac{t}{\mu}} dt \leq (n+1)! 2^{(\nu+6)n} K_{n+1} e^{-\frac{\varepsilon}{\mu^2}} \leq \mu,
\]
i.e. (H1) holds.

Above all, the KAM steps described in Section 3 are valid for all \( \nu \), which give the desired sequences stated in the lemma.

Now, (4.58) and (4.60) follow from Lemma 3.5, (4.72) and (4.74); by adding up (4.58), (4.60) for all \( \nu = 0, 1, \ldots \), we can get (4.59), (4.61); (4.62) follows from (3.57) in Lemma 3.8 and (4.72); (4.63) follows from (3.49) in Lemma 3.5 and (4.72); (2) follows from Lemma 3.7. \( \square \)

4.2. Convergence

The convergence is standard. For the sake of completeness, we briefly give the framework of proof. Let
\[
\Psi^\nu = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_\nu, \quad \nu = 1, 2, \cdots.
\]
By Lemma 4.1, we have
\[
D_{\nu+1} \subset D_n,
\]
\[
\Psi^\nu : D_n \to D_0,
\]
\[
H_0 \circ \Psi^\nu = H = N_\nu + P_\nu,
\]
\[
N_\nu = e_\nu + \left( \omega(\xi_\nu) + \sum_{j=0}^{\nu-1} p_{0j}^0(\xi_j), y \right) + \hat{h}_\nu(y, \xi_\nu),
\]
\( \nu = 0, 1, \cdots \), where \( \Psi_0 = id \). Using (4.64) and the identity
\[
\Psi^\nu = id + \sum_{i=1}^\nu \left( \Psi^i - \Psi^{i-1} \right),
\]
it is easy to verify that \( \Psi^\nu \) is uniformly convergent and denote the limitation by \( \Psi^\infty \).

In view of Lemma 4.1, it is obvious to see that \( e_\nu, \hat{h}_\nu, \xi_\nu \) converge uniformly about \( \nu \), and denote its limitation by \( e_\infty, \hat{h}_\infty, \xi_\infty \). By Lemma 3.4, we have
\[
\omega(\xi_1) + p_{01}^0(\xi_1) = \omega(\xi_0),
\]
23
\[ \omega(\xi_2) + p_{01}^0(\xi_2) + p_{01}^1(\xi_2) = \omega(\xi_0), \]
\[ \vdots \]
\[ \omega(\xi_\nu) + p_{01}^0(\xi_\nu) + \cdots + p_{01}^{\nu-1}(\xi_\nu) = \omega(\xi_0). \tag{4.75} \]

Taking limits at both sides of (4.75), we get
\[ \omega(\xi_\infty) + \sum_{j=0}^{\infty} p_{01}^j(\xi_\infty) = \omega(\xi_0). \]

Then, on \( D(\frac{m}{r}, \frac{n}{r}) \), \( N_\nu \) converge uniformly to
\[ N_\infty = e_\infty + \langle \omega(\xi_0), y \rangle + \bar{h}_\infty(y, \xi_\infty). \]

Hence, on \( D(\frac{m}{r}, \frac{n}{r}) \),
\[ P_\nu = H_0 \circ \Psi_\nu - N_\nu \]
converge uniformly to
\[ P_\infty = H_0 \circ \Psi_\infty - N_\infty. \]

Since
\[ \|P_\nu\|_{D_\nu} \leq \gamma_0^{\nu+m+2}\gamma_0^{\mu_\nu}, \]
by (4.72), we have that it converges to 0 as \( \nu \to \infty \). So, on \( D(0, \frac{m}{r}) \),
\[ J\nabla P_\infty = 0. \]

Thus, for the given \( \xi_0 \in O \), the Hamiltonian
\[ H_\infty = N_\infty + P_\infty \]
admits an analytic, quasi-periodic, invariant \( n \)-torus \( \mathbb{T}^n \times \{0\} \) with the Diophantine frequency \( \omega(\xi_0) \), which is the corresponding unperturbed toral frequency.

5. Proof of Theorems 2 and 4

First, we briefly give the proof framework of Theorem 2 because it can follow the KAM step in Section 3, where we mainly point out the two major differences from the proof of Theorem 1. The first one is that the homotopy invariance and excision of topological degree are used to keep the frequency unchanged in the iteration process not by picking parameters because we consider a Hamiltonian not a family of Hamiltonian. The other one is that the transformation defines on a smaller domain because we see the action-variable as parameter and the translation of parameter is equivalent to the action-variable's.
The following lemma is crucial to our arguments.

\[ \text{Lemma 5.1.} \] There exists \( \xi_0 \in B_{\mu_0}(\xi) \) such that

\[ \omega_0(\xi_0) = \cdots = \omega_0(\xi_0). \]
where \( i = 1, 2, \ldots, v \). Then, we need to find \( \xi_i \) near \( \xi \) such that \( \omega_i(\xi_i) = \omega(\xi) \). In view of (5.81), we observe that

\[
|\omega_i(y) - \omega(y)| = O(\mu).
\]  

(5.84)

We split

\[
\omega_v(y) - \omega(\xi) = (\omega(y) - \omega(\xi)) + (\omega_v(y) - \omega(y)).
\]  

(5.85)

Consider homotopy \( H_i(y) : [0, 1] \times O \rightarrow \mathbb{R}^n \),

\[
H_i(y) := (\omega(y) - \omega(\xi)) + t(\omega_v(y) - \omega(y)).
\]

For any \( y \in \partial O, t \in [0, 1] \), by (A1), we have that

\[
|H_i(y)| \geq |\omega(y) - \omega(\xi)| - |\omega_v(y) - \omega(y)| \geq |\omega_0(y) - \omega(\xi_0)| - \sum_{i=0}^v |\omega_{i+1}(y) - \omega_i(y)|
\]

\[
\geq \sigma|y - \xi_0|^L - \sum_{i=0}^v s_i^{n+m+2} s_i^m \mu_i > \frac{\sigma \delta^L}{2},
\]

where \( \delta := \min|y - \xi_0|, \forall y \in \partial O \).

So, it follows from homotopy invariance and (A0) that

\[
\deg(H_i(\cdot), O^0, 0) = \deg(H_0(\cdot), O^0, 0) \neq 0.
\]

We note by (A1), (5.84) and (5.85) that for any \( y \in O \setminus B_{y^0/\mu}(\xi) \),

\[
|H_i(y)| = |\omega_i(y) - \omega(\xi)| \geq |y - \xi|^L - c_1 y_0^{n+m+2} s_i^m \mu_{i+1}
\]

\[
\geq \delta^L \mu - c_1 y_0^{n+m+2} s_i^m \mu_{i+1} \geq \frac{\delta^L \mu}{2}.
\]

Hence, by excision, we have that

\[
\deg(H_i(\cdot), B_{y^0/\mu}(\xi), 0) = \deg(H_i(\cdot), O^0, 0) \neq 0,
\]

i.e., there exists at least a \( \xi_i \in B_{y^0/\mu}(\xi) \), such that \( H_i(\xi_i) = 0 \), i.e.,

\[
\omega_i(\xi_i) = \omega(\xi),
\]

which implies (5.83).

The proof is complete.
In the following, we prove

\[ \phi : D_{4\alpha} \to D_{5\alpha} \]  

which is different from (3.54) in Lemma 3.7. Recall that \( m > L + 1 \) and \( \alpha = \mu_{\frac{1}{4}} \), we have

\[ cs\mu_{\frac{1}{4}} < \frac{1}{8} \alpha s. \]  

For \( \forall (y, x) \in D_{4\alpha} \), we note by \( \xi_+ \in B_{\alpha s}(\xi) \) in Lemma 5.1 and (5.87) that

\[ |y + \xi_+ - \xi| < |y| + |\xi_+ - \xi| < \frac{1}{8} \alpha s + cs\mu_{\frac{1}{4}} < \frac{1}{4} \alpha s, \]  

which implies (5.86).

Next, we prove Theorem 4 by a direct method.

5.2. Proof of Theorem 4

(1) The unperturbed motion of (1.1) is described by the equation

\[ \begin{cases} \dot{y} = 0, \\ \dot{x} = h'(y). \end{cases} \]

The flow is \( x = h'(y)t + x_0, y \in G \), where \( x_0 \) is an initial condition. Notice that

\[ h''(0) = 0, \]

i.e., \( h(y) \) is degenerate at \( \xi_0 = 0 \). Obviously, by simple calculation, we get

\[ \text{deg} (h'(y) - h'(0), B_\delta(0), 0) = 0, \]

i.e., (A0) fails, then Theorem 3 is not applicable.

Note that the perturbed motion equation is

\[ \begin{cases} \dot{y} = 0, \\ \dot{x} = h'(y) + \epsilon P'(y). \end{cases} \]

The flow is \( x = (h'(y) + \epsilon P'(y))t + x_1, y \in G \), where \( x_1 \) is an initial condition. To ensure the frequency is equal to \( h'(0) \), we need to find a solution of the following equation in \( G \): 

\[ h'(y) + \epsilon P'(y) = h'(0), \]

i.e.,

\[ g'(y) = -\epsilon P'(y). \]  

Notice that the Taylor expansion of \( g'(y) \) at \( \xi_0 = 0 \) is

\[ g'(y) = g'(0) + g''(0)y + \cdots + g^{2\ell+1}(0)y^{2\ell} + o(y^{2\ell}), \]

then the equation (5.88) is equivalent to

\[ g^{2\ell+1}(0)y^{2\ell} + o(y^{2\ell}) = -\epsilon P'(y), \]

27
which is solvable provided that $\varepsilon P' (y) \text{sign} \left( g^{2L+1}(0) \right) < 0$. So the perturbed system admits at least two invariant tori with frequency $\omega = h'(0)$ for the small enough perturbation satisfying $\varepsilon P' (y) \text{sign} \left( g^{2L+1}(0) \right) < 0$. Conversely, if $\varepsilon P' (y) \text{sign} \left( g^{2L+1}(0) \right) > 0$, the unperturbed invariant torus with frequency $\omega = h'(0)$ will be destroyed.

(2) Note that $h(y)$ is degenerate in $\xi_0 = 0$. Obviously, by simple calculation, we get

$$\deg \left( h'(y) - h'(0), B_0(0), 0 \right) \neq 0.$$  

Then, by Theorem 3, the above persistence result hold. In addition, we can also directly prove this result. Similarly, we need to solve the following equation in $G$:

$$h'(y) + \varepsilon P'(y) = h'(0),$$

i.e.,

$$g'(y) = -\varepsilon P'(y). \quad (5.89)$$

Notice that the Taylor expansion of $g'(y)$ at $\xi_0 = 0$ is

$$g'(y) = g'(0) + g''(0)y + \cdots + g^{2L+2}(0)y^{2L+1} + o \left( y^{2L+1} \right),$$

then the equation (5.89) is equivalent to

$$g^{2L+2}(0)y^{2L+1} + o \left( y^{2L+1} \right) = -\varepsilon P'(y),$$

whose solution always exists in $G$ for any small enough perturbation. Hence, the perturbed system admits an invariant torus with frequency $\omega = h'(0)$ for any small enough perturbation.

6. Appendix A. Proof of Proposition 1

Proof. Obviously, for $\forall \xi \in (-1, 1) \times (-1, 1)$,

$$(\omega - \bar{\omega})(-\xi) = -(\omega - \bar{\omega})(\xi),$$

and for $\forall \xi \in \partial(-1, 1) \times (-1, 1)$,

$$(\omega - \bar{\omega})(\xi) \neq 0.$$  

Using Borsuk’s theorem in [27], we have

$$\deg \left( \omega(\cdot) - \bar{\omega}, (-1, 1) \times (-1, 1), 0 \right) \neq 0,$$

i.e.,

$$\deg \left( \omega(\cdot), (-1, 1) \times (-1, 1), \bar{\omega} \right) \neq 0,$$

i.e., (A0) holds. For $\xi, \xi_\ast \in \left[ -\frac{1}{2}, \frac{1}{2} \right]$, and $\xi \neq \xi_\ast$, we have

$$\omega(\xi) - \omega(\xi_\ast) = 0,$$

but

$$|\xi - \xi_\ast| > 0, \quad \forall L > 0.$$  

which shows that (A1) fails. Note that the flow of unperturbed motion equation is
\[ x = \omega(\xi)t + x_0, \quad \xi \in (-1, 1) \times (-1, 1), \]
where \( x_0 \) is an initial condition, and the flow of perturbed motion equation is
\[ x = \left( \omega(\xi) + (0, P_0(\varepsilon)) \right)t + x_0, \quad \xi \in (-1, 1) \times (-1, 1). \]
In order to keep the frequency \( \omega(0) = \bar{\omega} \) unchanged, we have to solve the following equation
\[ \omega(\xi) + (0, P_0(\varepsilon))^T = \bar{\omega}, \]
i.e.,
\[ \omega(\xi) - \bar{\omega}(\xi) = (\xi_1, \omega_2 - \bar{\omega}_2)^T = -(0, P_0(\varepsilon))^T, \]
which implies that the second component \( \xi_2 \) of solution \( \xi \) is discontinuous and alternately appears on \((-1, -\frac{1}{2})\) and \((\frac{1}{2}, 1)\) as \( \varepsilon \to 0^- \). So, this example shows that condition (A1) is necessary no matter how smooth the frequency mapping \( \omega(\xi) \) is.

7. Appendix B. Proof of Theorem 3

Proof. Notice that
\[ \nabla h(y) - \nabla h(0) = y|y|^{2l}. \]
For \( 0 < \delta < 1 \), \( B_\delta(0) \) denotes the open ball centered at the origin with radius \( \delta \). We have that \( \nabla h(y) - \nabla h(0) \) is odd and unequal to zero on \( \partial B_\delta(0) \), i.e.,
\[ \nabla h(-y) - \nabla h(0) = -y|y|^{2l} = -(\nabla h(y) - \nabla h(0)), \quad \nabla h(y) - \nabla h(0) \neq 0, \quad \forall y \in \partial B_\delta(0). \]
It follows from Borsuk’s theorem in [27] that,
\[ \deg (\nabla h(y) - \nabla h(0), B_\delta(0), 0) \neq 0. \]
Obviously, there exist \( \sigma = \frac{\min_{y \in \mathbb{R}^1} (\| y \|^{2l} - \| y \|^{2l})}{2^{2l+1}} \) and \( L = 2l + 1 \) such that
\[ |\nabla h(y) - \nabla h(y_*)| \geq \sigma |y - y_*|^{2l}, \quad y \in B_\delta(0), y \in B_\delta(0) \setminus B_{\delta}(y_*), \]
where \( \delta > 0, B_{\delta}(y_*) \subset B_\delta(0) \). So, by Theorem 2, the perturbed system admits an invariant torus with frequency \( \omega \) for any small enough perturbation.

8. Appendix C. Proof of Proposition 2

Proof. Let \( \varepsilon P = \varepsilon y, \varepsilon > 0 \). Notice that for \( y \in G \subset \mathbb{R}^1 \),
\[ h'(y) = \omega + y^{2l}, \quad h'(0) = \omega, \quad h''(y)|_{y=0} = 0, \]
which implies that the Hamiltonian \( H \) is degenerate at \( y = 0 \). By the definition of degree, we have for \( 0 < \delta < 1 \)
\[ \deg (\nabla h(y) - \nabla h(0), B_\delta(0), 0) = 0. \]
i.e., (A0) fails. Then, Theorem 2 cannot be used to prove the persistence result of keeping frequency unchanged.

Note that the flow of unperturbed motion equation at $y = 0$ is

$$x = \omega t + x_0,$$

where $x_0$ is an initial condition, and the flow of perturbed motion equation is

$$x = \left(\omega + y^2 + \epsilon\right)t + x_0, \quad y \in G.$$

In order to preserve frequency $\omega$, we need to solve $y^2 + \epsilon = 0$ in $G$, which has no real solution in $G$. Hence, the persistence result of keeping frequency unchanged fails.

Acknowledgments

The second author (Li Yong) is supported by National Basic Research Program of China (Grant number [2013CB8-34100]), National Natural Science Foundation of China (Grant numbers [11571065], [11171132], and [12071175]), and Natural Science Foundation of Jilin Province (Grant number [20200201253JC]).

References

[1] Arnold, V.I.: Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. Uspehi Mat. Nauk.18,5,13-40(1963). MR0163025

[2] Benettin, G., Galgant, L., Giorgilli, A., Strelcyn, J.: A proof of Kolmogorov’s theorem on invariant tori using canonical transformations defined by the Lie method. Nuovo Cimento.79B,2,201-223(1984). MR0743977

[3] Biasco, L., Chierchia, L., Treschev, D.: Stability of nearly integrable, degenerate Hamiltonian systems with two degrees of freedom. J. Nonlinear Sci.16,1,79-107(2006). MR2202903

[4] Bourennoura, A.: Positive measure of KAM tori for finitely differentiable Hamiltonians. J. Ec. polytech. Math.7,1113-1132(2020). MR4167789

[5] Bricmont, J., Gawedzki, K., Kupiainen, A.: KAM theorem and quantum field theory. Commun. Math. Phys.201,699-727(1999). MR1685894

[6] Bruno, A.D.: Nondegeneracy conditions in the Kolmogorov theorem. Soviet Math. Dokl.45,1,221-225(1992). MR1171798

[7] Cheng, C.Q., Sun, Y.S.: Existence of KAM tori in degenerate Hamiltonian systems. J. Differential Equations.114,1,288-335(1994). MR1302146

[8] Chierchia, L., Falcolini, C.: A direct proof of a theorem by Kolmogorov in Hamiltonian systems. Ann. Scuola Norm. Sup. Pisa Cl. Sci.21,4,541-593(1994). MR1318772

[9] Chierchia, L.: Periodic solutions of the planetary N-body problem. XVIIth International Congress on Mathematical Physics.269-280(2014). MR3204477

[10] Chow, S.N., Li, Y., Yi, Y.F.: Persistence of invariant tori on submanifolds in Hamiltonian systems. J. Nonlinear Sci.12,585-617(2002). MR1938331

[11] Cong, F.Z., Küpper, T., Li, Y., You, J.G.: KAM-type theorem on resonant surfaces for nearly integrable Hamiltonian systems. J. Nonlinear Sci.10,1,49-68(2000). MR1730569

[12] de la Llave, R., González, A., Jorba, À., Villanueva, J.: KAM theory without action-angle variables. Nonlinearity.18,2,855-895(2005). MR2122688

[13] Eliasson, L.H.: Absolutely convergent series expansions for quasi-periodic motions. Math. Phys. Elect. J.2,1-33(1996). MR1399458

[14] Féjoz, J.: Démonstration du ‘théorème d’Arnold’ sur la stabilité du système planétaire (d’après Herman). Ergodic Theory Dynam. Systems.24,5,1521-1582(2004). MR2104595

[15] Gallavotti, G.: Twistless KAM tori. Commun. Math Phys.164,145-156(1994). MR1288156

[16] Gallavotti, G., Gentile, G., Mastropietro, V.: Field theory and KAM tori. Math. Phys. Elect. J.1,1-13(1995). MR1359460
[17] Han, Y.C., Li, Y., Yi, Y.F.: Invariant tori in Hamiltonian systems with high order proper degeneracy. Ann. Henri Poincaré.10,1419-1436(2010). MR2639543
[18] Heinz, H.: Non-degeneracy conditions in KAM theory. Indag. Math. (N.S.).22,3-4,241-256(2011). MR2853608
[19] Herman, M.R.: Sur les courbes invariantes par les difféomorphismes de l’anneau. Vol. 1. (French) [On the curves invariant under diffeomorphisms of the annulus. Vol. 1] With an appendix by Albert Fathi. With an English summary. Astérisque, 103-104. Société Mathématique de France, Paris, 1983, pp. i+221. MR0728564
[20] Kolmogorov, A.N.: On conservation of conditionally periodic motions for a small change in Hamilton’s function. Dokl. Akad. Nauk SSSR.98,527-530(1954). MR0068687
[21] Koudjinan, C.E.: A KAM theorem for finitely differentiable Hamiltonian systems. J. Differential Equations.269,6,4720-4750(2020). MR4104457
[22] Li, Y., Yi, Y.F.: Persistence of invariant tori in generalized Hamiltonian systems. Ergodic Theory Dynam. Systems.22,4,1233-1261(2002). MR1926285
[23] Li, Y., Yi, Y.F.: A quasi-periodic Poincaré’s theorem. Math. Ann.326,649-690(2003). MR2003447
[24] Meyer, K.R.: Periodic solutions of the N-body problem. Springer-Verlag, Berlin,1999. MR1736548
[25] Moser, J.: Convergent series expansions for quasi-periodic motions. Math. Ann.169,136-176(1967). MR0208078
[26] Moser, J.: On invariant curves of area-preserving mapping of an annulus. Nachr. Akad. Wiss. Göttingen.II.1-20(1962). MR0147741
[27] Motreanu, D., Motreanu, V.V., Papageorgiou, N.: Topological and variational methods with applications to nonlinear boundary value problems. Springer Science+Business Media, LLC,2014. MR3136201
[28] Pöschel, J.: On elliptic lower-dimensional tori in Hamiltonian systems. Math. Z.202,4,559-608(1989). MR1022821
[29] Pöschel, J.: A KAM theorem for some nonlinear partial differential equations. Ann. Sc. Norm. Super. Pisa Cl. Sci.23,1,119-148(1996). MR1401420
[30] Pöschel, J.: A lecture on the classical KAM theorem. Proc. Symp. Pure Math.69,1-37(2004). MR1858551
[31] Qian, W.C., Li, Y., Yang, X.: Multiscale KAM theorem for Hamiltonian systems. J. Differential Equations.266,1,70-86(2019). MR3870557
[32] Qian, W.C., Li, Y., Yang, X.: Melnikov’s conditions in matrices. J. Dynam. Differential Equations.32,4,1779-1795(2020). MR4171876
[33] Rüssmann, H., Kleine Nenner I.: Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II, 67-105, 1970. MR0273156
[34] Rüssmann, H.: Invariant tori in non-degenerate nearly integrable Hamiltonian systems. Regul. Chaotic Dyn.6,2,119-204(2001). MR1843664
[35] Salamon, D.A.: The Kolmogorov-Arnold-Moser theorem. Math. Phys. Electron. J.10,1-37(2004). MR2111297
[36] Sevryuk, M.B.: KAM-stable Hamiltonians. J. Dynam. Control Systems.1,3,351-366(1995). MR1354540
[37] Sevryuk, M.B.: Partial perservation of frequencies in KAM theory. Nonlinearity.19,1099-1140(2006). MR2221801
[38] Wayne, C.E.: An introduction to KAM theory. Lectures in Appl. Math.31(1996). MR1363023
[39] Xu, J.X., You, J.G., Qu, Q.J.: Invariant tori for nearly integrable Hamiltonian systems with degeneracy. Math. Z.226,3,375-387(1997). MR1483538