Journal of Applied Analysis and Computation
Volume 9, Number 5, October 2019, 1616–1638
Website:http://jaac.ijournal.cn
DOI:10.11948/20180008

THE ISOENERGETIC KAM-TYPE THEOREM AT RESONANT CASE FOR NEARLY INTEGRABLE HAMILTONIAN SYSTEMS*

Weichao Qian[1], Yong Li[2,1,†] and Xue Yang[2,1]

Abstract In this paper, we study the persistence of resonant invariant tori on energy surfaces for nearly integrable Hamiltonian systems under the usual R*ü*ssmann nondegenerate condition. By a quasilinear iterative scheme, we prove the following things: (1) The majority of resonant tori on a given energy surface will be persisted under R*ü*ssmann nondegenerate condition. (2) The maximal number about the preserved frequency components on a perturbed torus is characterized by the smaller of the maximal rank of the Hessian matrices of the unperturbed system and the nondegeneracy of resonance. (3) If unperturbed systems admit subisoenergetic nondegeneracy on an energy surface, then the majority of the unperturbed resonant tori on the energy surface will be persisted and give rise to a family of perturbed tori with the same energy, whose frequency ratios among respective "nondegenerate" components are preserved.

Keywords Isoenergetic KAM-type theorem, resonant case, nearly integrable Hamiltonian systems.

MSC(2010) 37J40.

1. Introduction

Consider a family of real analytic nearly integrable Hamiltonian systems with the following form:

$$H(x, y) = H_0(y) + \varepsilon P(x, y), \tag{1.1}$$

where $y \in G \subset R^d$, $x \in T^d$; $H_0(y)$ is a real analytic function; $P(x, y)$, a perturbation, is a real analytic function and $\varepsilon > 0$ is a small parameter.

As well known, the celebrated KAM theory due to Kolmogorov [16], Arnold [1] and Moser [20] asserts that, if integrable part is nondegenerate, then nearly integrable systems still exhibit a large extent quasiperiodic motions on invariant tori. Furthermore, these tori $\{T_y : y \in G\}$, which contain resonant tori of all

†the corresponding author. Email address:liyongmath@163.com(Y. Li)
1School of Mathematics and Statistics & Center for Mathematics and Interdisciplinary Sciences, Northeast Normal University, 5268 People Street, 130024, China
2School of Mathematics, Jilin University, 2699 Qianjin Street, 130012, China
*The second author was supposed by National Basic Research Program of China (grant No. 2013CB834100), NSFC (grant No. 11571065), JilinDRC (grant No. 2017C028-1). The third author was supposed by Science and Technology Developing Plan of Jilin Province (grant No. 20180101220JC).

type of resonances, tend to be destroyed under arbitrary generic perturbation and born out a resonance zone consisting of both stochastic trajectories and regular orbits [5, 27]. To characterize regular orbits in the resonance region, an essential problem is to study mechanism of destruction for resonant tori and the persistence of certain lower dimensional tori which are split from the resonant ones.

Such a persistence problem was first considered by Poincaré [21] within a class of maximal resonances, which is foliated into periodic orbits. With respect to (1.1), the Poincaré theorem states that any periodic orbit associated to a nondegenerate relative equilibrium will be persisted. After a long time, Treshchëv [29] proved the persistence of hyperbolic tori in a case of general resonance. Eliasson [11], Chierchia etc [6], Rudnev etc [24] and Medvedev etc [19] also obtained similar results for the case of multiplicity one resonant.

To characterize general nondegenerate perturbation, Cong etc [9] proved that in general resonance case, under $g-$nondegeneracy condition (see **Definition 1.4**), (1.1) can be reduced to the following form:

$$H(x,y,z) = e + \langle \omega, y\rangle + \frac{\delta}{2}\langle z, Mz\rangle + \varepsilon P(x,y,z),$$

where M is a nonsingular matrix, and gives rise to a family of lower dimensional invariant tori. Without the $g-$nondegeneracy condition, Li etc [18] also obtained the same conclusion. And for some recent researches about resonance tori with respect to multiplicity r, $r > 1$, we refer readers to [10, 15, 32, 33].

Of course, there is a fair amount of attentions given to the fixed Diophantine torus about preservation of toral frequency. See Benettin etc [3] for KAM approach, Gallavotti [13], Chierchia etc [6] and Eliasson [12] for a direct method using Lindstedt series, Gallavotti etc [14] and Bricmont etc [4] for renormalization groups techniques, Chow etc [8] for the partial preservation of frequencies on a submainfold, Sevryuk for the degenerate systems [25] and the partial preservation not only frequencies but also Floquet exponents [26], and Xu etc [31] for topological degree techniques. Then, on a given energy surface $\mathcal{M}$, is there a family of resonant invariant tori which preserve the toral frequencies? Furthermore, on the energy surface $\mathcal{M}$, is there a family of resonant invariant tori in which the frequency ratios are preserved?

To deal with above problem, it is indispensable to set up a subgroup g of Z^d on energy surface $\mathcal{M}$. Before setting up the subgroup g, we are going to introduce some definitions.

Definition 1.1. For $H_0(y)$, $\omega = \frac{\partial H_0}{\partial y}(y)$ is called nonresonant, if $\langle k, \omega\rangle \neq 0$ for any $k \in Z^d \setminus \{0\}$. Otherwise, it is resonant.

Definition 1.2. If there is a subgroup g of Z^d, generated by independent integer vectors $\tau_1, \ldots, \tau_{m_0}$, such that $\langle k, \omega\rangle = 0$ for all $k \in g$ and $\langle k, \omega\rangle \neq 0$ for all $k \in Z^d/g$, then ω is called a frequency with multiplicity m_0 resonance.

Definition 1.3. For any given subgroup g, the m $(= d - m_0)$ dimensional surface $\widetilde{\Lambda}(g, G) = \{y \in G : \langle k, \omega\rangle = 0, k \in g\}$ is called $g-$resonant surface.

Remark 1.1. Under the nondegenerate condition, locally, a $g-$resonant surface is diffeomorphic to R^m. In a typical way [1], by passing to finite coverings which will also lead to the global result on G, we may assume that $\widetilde{\Lambda}(g, G)$ is globally differomorphic to a subdomain in R^m without loss of generality.

Similar to [29], by group theory, on a given energy surface $\mathcal{M}$ there are integer vectors $\tau_1', \cdots, \tau_m' \in Z^d$ such that Z^d is generated by $\tau_1, \cdots, \tau_{m_0}, \tau_1', \cdots, \tau_m'$ and $\det K_0 = 1$, where $K_0 = (K_*, K^{'})$, $K_* = (\tau_1', \cdots, \tau_m')$, $K^{'} = (\tau_1, \cdots, \tau_{m_0})$ are $d \times d$, $d \times m$, $d \times m_0$, respectively, and K_* generates the quotient group Z^d/g, while $K^{'}$ generates the group g.

Definition 1.4. If H_0 is nondegenerate and $\det K^{'T} \frac{\partial^2 H_0}{\partial y^2}(y) K^{'} \neq 0$ for $y \in \widetilde{\Lambda}(g, G)$, then H_0 is $g-$nondegenerate.

Without loss of generality, we suppose that $\mathcal{M}$ admits a global coordinate, i.e., there is a bounded closed region $\Lambda \in R^{d_0}$ and a C^{l_0} diffeomorphism $y : \Lambda \to \mathcal{M}$ such that $\mathcal{M} = y(\Lambda)$. For $\lambda \in \Lambda$ under the transformation

$$y \mapsto y + y(\lambda),$$

Hamiltonian (1.1) reads

$$H(x, y, \lambda; \varepsilon) = N(y, \lambda) + P(x, y, \lambda; \varepsilon),$$

where

$$\begin{aligned} N(y, \lambda) &= H_0(y(\lambda)) + \langle \frac{\partial H_0}{\partial y}(y(\lambda)), y - y(\lambda) \rangle + h(y, \lambda), \\ h(y, \lambda) &= \frac{1}{2} \langle y - y(\lambda), \frac{\partial^2 H_0}{\partial y^2}(y(\lambda))(y - y(\lambda)) \rangle + O(|y - y(\lambda)|^3), \\ P(x, y, \lambda; \varepsilon) &= \varepsilon P(x, y, \lambda). \end{aligned}$$

Let

$$\Gamma = K_0^T \frac{\partial^2 H_0}{\partial y^2}(\lambda) K_0 = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix},$$

where Γ_{11}, Γ_{12}, Γ_{21}, Γ_{22} are $m \times m$, $m \times m_0$, $m_0 \times m$, $m_0 \times m_0$ matrices, respectively, and $\Gamma_{12} = \Gamma_{21}^T$, $\Gamma_{22} = K^{'T} \frac{\partial^2 H_0}{\partial y^2}(\lambda) K^{'}$.

Then on energy surface $\mathcal{M}$ with the symplectic coordinate transformation $y - y(\lambda) = K_0 p$, $q = K_0^T x$, Hamiltonian (1.1) is changed to

$$H(q, p) = e_0 + \langle \omega^*, p^{'} \rangle + \frac{1}{2} \langle p^{''}, \Gamma_{22}(\lambda) p^{''} \rangle + O(|p|^3) + O(|p^{'}|^2) + O(|p^{'}| \cdot |p^{''}|) + \varepsilon \bar{P}(p, q, \omega^*),$$

where

$$\begin{aligned} \omega^* &= K_*^T \omega(\lambda) \in \Lambda(g, G), \\ \Lambda(g, G) &= \{\omega^* \in R^m : y \in \widetilde{\Lambda}(g, G)\}, \\ p^{'} &= (p_1, \cdots, p_m)^T, \\ p^{''} &= (p_{m+1}, \cdots, p_d)^T, \\ \bar{P}(q, p) &= P((K_0^T)^{-1} q, y(\lambda) + K_0 p). \end{aligned}$$

Here we used the fact that $\Lambda(g, G)$ is diffeomorphic to the $m-$dimensional surface $\widetilde{\Lambda}(g, G)$.

Let

$$\Lambda^{'}(g,G)=\{\lambda\in\Lambda(g,G):|\langle k,\omega\rangle|>\frac{\gamma_0}{|k|^{\tau}},\forall k\in Z^m\setminus\{0\}\},$$

where $|k|=\sum_{i=1}^{m}|k_i|$; γ_0 and τ are fixed positive constants. For any $\omega\in\Lambda^{'}(g,G)$, we separate the first-order resonant terms from the perturbation by a canonical transformation of coordinates,

$$(p,q\ mod\ 2\pi)\to(Y,X\ mod\ 2\pi):p=\tfrac{\partial S(q,Y)}{\partial q},X=\tfrac{\partial S(q,Y)}{\partial Y},$$

where

$$\begin{aligned}S&=\langle Y,q\rangle+\varepsilon\sum_{k\in Z^m\setminus\{0\}}\frac{\sqrt{-1}h_k}{\langle k,\omega\rangle}(q^{''},\omega)e^{\sqrt{-1}\langle k,q^{'}\rangle},\\ h_k&=\int_0^{2\pi}\bar{P}(q,0)e^{\sqrt{-1}\langle k,q^{'}\rangle}dq^{'}.\end{aligned}$$

Then

$$\begin{aligned}p^{'}&=Y^{'}+\sqrt{-1}\varepsilon\sum_{k\in Z^m\setminus\{0\}}kS_ke^{\sqrt{-1}\langle k,q^{'}\rangle},\\ p^{''}&=Y^{''}+O(\varepsilon),\\ X&=q,\\ S_k&=\frac{\sqrt{-1}h_k}{\langle k,\omega\rangle}.\end{aligned}$$

The new Hamiltonian reads as

$$\begin{aligned}H(X,Y)=&\langle\omega,Y^{'}\rangle+\frac{1}{2}\langle Y^{''},\Gamma_{22}(\lambda)Y^{''}\rangle+\varepsilon h_0(X^{''},\omega)+O(\varepsilon Y)+O(\varepsilon^2)+O(|Y|^3)\\ &+O(|Y^{'}|^2)+O(|Y^{'}||Y^{''}|),\end{aligned}$$

up to an irrelevant constant. We assume that $h_0(X^{''},\omega)$ has a nondegenerate critical point, say $X_0^{''}$. Taking $X=X$, $Y=\varepsilon^{\frac{1}{2}}\bar{Y}$, $\bar{H}=\varepsilon^{-\frac{1}{2}}H$ and using Taylor expansion at the equilibrium point, the Hamiltonian arrives at

$$\begin{aligned}\bar{H}(X,Y)&=H(X,\varepsilon^{\frac{1}{2}}\bar{Y})/\varepsilon^{\frac{1}{2}}\\ &=\langle\omega,y\rangle+\varepsilon O(|y|^2)+\frac{\varepsilon}{2}(\langle u,V_0u\rangle+\langle v,U_0v\rangle)+\varepsilon O(|u|^3)+O(\varepsilon^2)+\varepsilon O(|y||v|),\end{aligned}$$

where we replace $\bar{H}$, $X^{'}$, $\bar{Y}^{'}$, $X^{''}$, $\bar{Y}^{''}$, $\varepsilon^{\frac{1}{2}}$, Γ_{22} and $\frac{\partial^2h_0}{\partial X^{''2}}(0,\omega)$ by H, x, y, u, v, ε, U_0 and V_0, respectively.

Denote $z=(u,v)$, $M=\begin{pmatrix}V_0&0\\0&U_0\end{pmatrix}$, $\varepsilon O(|y|^2)=\frac{\varepsilon}{2}\langle y,Ay\rangle+\varepsilon\hat{y}(y)=\frac{\varepsilon}{2}\langle y,Ay\rangle+\varepsilon O(|y|^3)$ and $\varepsilon O(|u|^3)=\varepsilon\hat{z}(z)$.

Finally, we get

$$H(x,y,u,v)=N+P,\tag{1.2}$$

where

$$\begin{aligned} N &= \langle \omega(\lambda), y\rangle + h, \\ h &= \frac{\varepsilon}{2}\langle y, A(\lambda)y\rangle + \varepsilon\hat{y}(y) + \frac{\varepsilon}{2}\langle z, M(\lambda)z\rangle + \varepsilon\hat{z}(z), \\ \hat{y}(y) &= O(|y|^3), \\ \hat{z}(z) &= O(|u|^3), \\ P &= O(\varepsilon^2) + \varepsilon O(|y||v|), \end{aligned}$$

with $z = (u, v)^T$, and $x \in T^m$, $y \in R^m$, $u, v \in R^{m_0}$, $\lambda \in \Lambda$, which is a bounded closed region on R^n. In the above, all $\lambda-$dependences are of class C^{l_0} for some $l_0 \geq d$. Hereinafter, rewrite $m = d$.

For any small ε, let $s = \varepsilon^{\frac{1}{3}}$. Consider Hamiltonian (1.1) in

$$D(r, s) = \{(x, y, z) : |Im\ x| < r,\ |y| < s^2,\ |z| < s\}.$$

It is easy to see that on $D(r, s) \times \Lambda$

$$|\partial_\lambda^l P| \leq C\varepsilon^2.$$

Set $\delta = \varepsilon$, $\gamma = \varepsilon^{\frac{\frac{1}{3}-\iota}{d+9}}$, $\mu = \varepsilon^\iota$, $\iota \in (0, \frac{1}{3})$. Hence

$$|\partial_\lambda^l P| \leq \delta\gamma^{d+9}\mu s^2, \qquad |l| \leq d.$$

Since $P(x, y, z)$ is a real analytic function defined on some complex neighborhood of $T^d \times \{0\} \times \{0\}$, with the Taylor-Fourier series, we have

$$P = \sum_{|k| \in Z^d,\ |j|, |q| \in Z_+^d} P_{kjq} y^j z^q e^{\sqrt{-1}\langle k, x\rangle}.$$

We assume the following conditions hold:

(A1) $rank\ \{\frac{\partial^\alpha \omega}{\partial \lambda^\alpha} : |\alpha| \leq d-1\} = d$ for all $\lambda \in \Lambda$.

(A2) For given n, $0 < n \leq m$, $rank\ A(\lambda) = n$ and $rank\ M(\lambda) = 2m_0$ on Λ.

$(A1^{'})$ For given n, $0 < n \leq m$, there is a smoothly varying $n \times n$ minor $\mathcal{A}$ of $A(\lambda)$ on $\mathcal{M}$, a given energy surface, such that

$$\det \begin{pmatrix} \mathcal{A}(\lambda) & \omega^*(\lambda) \\ \omega^*(\lambda)^T & 0 \end{pmatrix} \neq 0,$$

where $\omega^*(\lambda) = \frac{\partial N}{\partial y_*}$, $y_* = (y_{i_1}, \cdots, y_{i_n})^T$ and i_1, i_2, $\cdots$, i_n denote the row indices of $\mathcal{A}(\lambda)$ in $A(\lambda)$; $rank M(\lambda) = 2m_0$.

Remark 1.2. If ω is real analytic, Λ is connected, and $\max_{\lambda\in\Lambda}\ rank\ \{\frac{\partial^\alpha \omega}{\partial \lambda^\alpha} : |\alpha| \leq d-1\} = d$, condition $(A1)$ is Rüssmann condition.

Remark 1.3. Actually, condition $(A1^{'})$ is a condition called subisoenergetic nondegenerate condition, where $\mathcal{A}$ is a minor of $\frac{\partial^2 H_0}{\partial y^2}$, hence weaker than the following Arnold's isoenergetic nondegenerate condition [2]:

(AI) On a given level set: $H_0 = c$,

$$\det \begin{pmatrix} \frac{\partial^2 H_0}{\partial y^2} & \frac{\partial H_0}{\partial y} \\ \frac{\partial H_0}{\partial y}^T & 0 \end{pmatrix} \neq 0.$$

By using quasilinear iterative scheme introduced in [18], we have the following theorem.

Theorem 1.1. *Consider* (1.1)*. Suppose H_0 is g-nondegenerate for a given g, and $h_0(X'', Y)$ has an analytic family of nondegenerate critical points for all $y \in \tilde{\Lambda}(g, G)$. Let $\mathcal{M}$ be an energy surface given by $\{H_0(y) = E\}$ and $\tau > l_1(d-1) - 1$ be fixed, where l_1 is a constant.*

(i) Assume $(A1)$ *and* $(A2)$ *on $\mathcal{M}$. Then there exists a $\varepsilon_0 > 0$ and a family of Cantor sets $\Lambda_\varepsilon(g, G) \subset \tilde{\Lambda}(g, G)$, $0 < \varepsilon < \varepsilon_0$, with $|\tilde{\Lambda}(g, G) \setminus \Lambda_\varepsilon(g, G)| = O(\gamma^{\frac{1}{l_1}})$, such that each $y \in \Lambda_\varepsilon(g, G)$, the unperturbed tours T_y persists and gives rise to a family of real analytic, quasiperiodic invariant tori $T_{\varepsilon,y}$ which preserves the frequency components $\omega_{i_1}(y)$, $\omega_{i_2}(y)$, $\cdots$, $\omega_{i_n}(y)$ of the unperturbed toral frequency $\omega(y)$. Moreover, these perturbed tori form a C^{l_0-1} Whitney smooth family;*

(ii) If $(A1)$ *and* $(A1^{'})$ *hold on $\mathcal{M}$, then each perturbed torus $T_{\varepsilon,y}$ preserves the ratios of the i_1, $\cdots$, i_n components of its toral frequency $\omega_\varepsilon(y)$, i.e.,*

$$[\omega_{\varepsilon,i_1}(y) : \cdots : \omega_{\varepsilon,i_n}(y)] = [\omega_{i_1}(y) : \cdots : \omega_{i_n}(y)],$$

where $\omega_{\varepsilon,i_j}(y)$ and $\omega_{i_j}(y)$ are the i_j-th components of $\omega_\varepsilon(y)$ and $\omega(y)$, respectively, for $j = 1, 2, \cdots, n$.

Remark 1.4. Notice that d_0 is the dimension of parameter and d is the dimension of the space of x, thus there are two cases: (1) $d_0 > d$, then nondegenerate condition $(A1)$ fails with respect to the original parameters of Hamiltonian systems and extra deformation parameters need to be added so that a joint nondegenerate condition of type $(A1)$ can be hold with respect to the extended parameters; (2) $d_0 \leq d$, *Theorem* 1.1 has a direct application to nearly integrable Hamiltonian systems (1.1) with respect to the persistence of invariant tori on a submanifold of G.

The article is organized as follows. In Section 2 we show the quasilinear iterative scheme for one KAM cycle. And in Section 3 we complete the proof of our results by deriving an iteration lemma and giving measure estimates.

2. The KAM Step

Throughout, for any two complex column vectors ξ, ζ with the same dimension, $\langle \xi, \zeta \rangle$ always stands for $\xi^T \zeta$, i.e., the transpose of ξ times ζ. Below, unless specified explanation, we shall use the same symbol $|\cdot|$ to denote an equivalent (finite dimensional) vector norm and its induced matrix norm, absolute value of functions, and measure of sets, etc., and use $|\cdot|_D$ to denote the supremum norm of functions on a domain D. For sake of brevity, we shall not specify smoothness orders for functions having obvious orders of smoothness indicated by their derivatives taken.

Moreover, all Hamiltonians in the sequel are associated to the standard symplectic structure.

The KAM iteration process consists of infinitely many KAM steps. From each cycle of KAM steps, one can find constructions and estimations of desired symplectic transformations and their domains, perturbed frequencies and new perturbations. For the sake of convenience, we shall omit the index for all quantities of the $\nu-$th KAM step and use $'+'$ to index all quantities in the $(\nu+1)-$th KAM step. All constants below are positive and independent of the iteration process. To simplify the nations, we shall suspend the $\lambda-$dependence in most terms in this section.

We start with the Hamiltonian systems

$$H = N + \varepsilon P, \tag{2.1}$$

defined on

$$D(r,s) = \{(x,y,z) : |Im\ x| < r,\ |y| < s^2,\ |z| < s\},$$

where

$$\begin{aligned} &N = e + \langle \omega, y\rangle + \frac{\delta}{2}\langle y, A(\lambda)y\rangle + \delta\hat{y}(y) + \frac{\delta}{2}\langle z, M(\lambda)z\rangle + \delta\hat{z}(z),\\ &P = O(\varepsilon^2) + \varepsilon O(|y||v|),\\ &|\partial_\lambda^l P|_{D(r,s)} \le \delta\gamma^{d+9}s^2\mu, \qquad |l| \le d. \end{aligned} \tag{2.2}$$

Consider (2.1) and define

$$\begin{aligned} &r_0 = r, \qquad \gamma_0 = 4\gamma, \qquad \beta_0 = s, \qquad \Lambda_0 = \Lambda, \qquad H_0 = H,\\ &e_0 = e, \qquad A_0 = A, \qquad \mathcal{A} = \mathcal{A}, \qquad M_0 = M, \qquad P_0 = P,\\ &N_0 = e_0(\lambda) + \langle \omega_0(\lambda), y\rangle + h_0, \quad \hat{y}_0(y) = \hat{y}(y), \quad \hat{z}_0(z) = \hat{z}(z),\\ &h_0 = \frac{\delta}{2}\langle y, A_0(\lambda)y\rangle + \delta\hat{y}_0(y) + \frac{\delta}{2}\langle z, M_0(\lambda)z\rangle + \delta\hat{z}_0(z). \end{aligned}$$

For sake of simplicity, we assume that $0 < r_0, \beta_0, \gamma_0 \le 1$ and $\mathcal{A}_0$ is the $n\times n$ minor of A_0 with $\det\mathcal{A}_0 \ne 0$. By monotonicity, we define μ_0, s_0 implicitly through the following equations:

$$\begin{aligned} \mu &= \frac{4^{d+5}\mu_0}{(M^*+1)^2K_1^{2\tau}},\\ s_0 &= \frac{\beta_0\mu_0}{16(M^*+1)}K_1^{\tau}, \end{aligned} \tag{2.3}$$

where

$$\begin{aligned} M^* &= \max_{|l|\le d,\ |j|+|i|\le 7, \lambda\in\Lambda_0} |\partial_\lambda^l\partial_y^j\partial_z^i h_0(y,z;\lambda)|,\\ K_1 &= ([\log\frac{1}{\mu_0}]+1)^{3\eta},\\ \eta &= \frac{\log 2}{\log 7 - \log 6}. \end{aligned}$$

It is clear that μ_0 is small if and only if μ is, and

$$\mu_0 = o(\mu^{1-\varepsilon}), \quad \forall\ 0 < \varepsilon \ll 1. \tag{2.4}$$

If μ is small enough, it is simple to get

$$16(M^*+1)K_1^\tau > 1.$$

Hence, $s_0 < \min\{\beta_0, \gamma_0\}$. Due to (2.3),

$$\frac{\mu}{\mu_0} = 4^{d+7}(\frac{s_0}{\beta_0\gamma})^2. \tag{2.5}$$

Besides (2.2), we have

$$|\partial_\lambda^l P_0|_{D(r_0,s_0)} \le \delta\gamma_0^{d+7}s_0^2\mu_0, \quad |l| \le d. \tag{2.6}$$

In the following, we characterize the quasilinear iteration scheme for Hamiltonian (2.1) in one KAM cycle, say, from $\nu-$th KAM step to the $(\nu+1)-$th. Now, suppose that after ν steps, we have arrived at the following real analytic Hamiltonian systems:

$$H = N + P,$$

where

$$\begin{aligned} N &= e + \langle \omega, y\rangle + h(y,z), \\ h(y,z) &= \frac{\delta}{2}\langle y, Ay\rangle + \delta\hat{y}(y) + \frac{\delta}{2}\langle z, Mz\rangle + \delta\hat{z}(z), \\ P &= O(\varepsilon^2) + \varepsilon O(|y||v|), \end{aligned}$$

defined on a phase domain

$$D(r,s) = \{(x,y,z) : |Im\ x| < r,\ |y| < s^2,\ |z| < s\},$$

and $\lambda \in \Lambda \subset \Lambda_0$. In addition, suppose that $\mathcal{A}$ is an $n\times n$ nonsingular minor of A on Λ, and $P = P(x,y,z;\lambda)$ satisfies, for some $0<\mu<\mu_0$, $0<\gamma\le\gamma_0$,

$$|\partial_\lambda^l P(x,y,z;\lambda)|_{D(r,s)} \le \delta\gamma^{d+7}s^2\mu,\ |l|\le d.$$

By considering both averaging and translation, we shall find a symplectic transformation Φ_+, which, on a small phase domain $D(r_+,s_+)$ and a smaller parameter domain Λ_+, transforms the Hamiltonian into the Hamiltonian of the next KAM step, i.e.

$$H_+ = H\circ\Phi_+ = N_+ + P_+,$$

where N_+, P_+ enjoy similar properties as N, P respectively on $D(r_+,s_+)\times\Lambda_+$. Define

$$\begin{aligned} r_+ &= \frac{r}{2} + \frac{r_0}{4}; \\ s_+ &= \frac{1}{8}\alpha s; \\ \alpha &= \mu^{2\sigma} = \mu^{\frac{1}{3}}; \\ \beta_+ &= \frac{\beta}{2} + \frac{\beta_0}{4}; \\ \gamma_+ &= \frac{\gamma}{2} + \frac{\gamma_0}{4}; \\ K_+ &= ([\log\frac{1}{\mu}]+1)^{3\eta}; \end{aligned}$$

$$D_{\frac{i}{8}\alpha} = D(r_+ + \frac{i-1}{8}(r-r_+), \frac{i}{8}\alpha s),\ i = 1, 2, \cdots, 8;$$
$$D(\xi) = \{y \in C^d : |y| < \xi^2,\ |z| < \xi\},\ \xi > 0;$$
$$\hat{D}(\xi) = D(r_+ + \frac{7}{8}(r-r_+), \xi),\ \xi > 0;$$
$$D_+ = D_{\frac{1}{8}\alpha} = D(r_+, s_+);$$
$$\tilde{D}_+ = D(r_+ + \frac{3}{4}(r-r_+), \beta_+);$$
$$\Gamma(r-r_+) = \sum_{0<|k|\le K_+} |k|^{4m_0^2(l+d+8)\tau+d+l+8} e^{-|k|\frac{r-r_+}{8}}.$$

2.1. Truncation

Consider the Taylor-Fourier series of P,

$$P = \sum_{|k|\in Z^d,\ |j|,|q|\in Z_+^d} P_{kjq} y^j z^q e^{\sqrt{-1}\langle k,x\rangle}, \tag{2.7}$$

and let R be the truncation of P with the following form:

$$R = \sum_{|k|\le K_+} (P_{k00} + \langle P_{k10}, y\rangle + \langle P_{k01}, z\rangle + \langle P_{k02} z, z\rangle) e^{\sqrt{-1}\langle k,x\rangle}.$$

Then

$$P - R = (\sum_{|k|>K_+} + \sum_{|k|\le K_+, 2|j|+|q|\ge 3})\ P_{kjq}\ y^j\ z^q\ e^{\sqrt{-1}\langle k,x\rangle}.$$

Thus, on a smaller domain $D_{\frac{7}{8}\alpha}$, under assumption

$$\int_{K_+}^{\infty} t^{d+m} e^{-t\frac{r-r_+}{8}} \le \mu, \tag{2.8}$$

where K_+ is defined as above, we have, for $|l| \le d$,

$$|\partial_\lambda^l (P-R)|_{D_{\frac{7}{8}\alpha}} \le C\delta\gamma^{d+7} s^2 \mu^2,$$
$$|\partial_\lambda^l R|_{D_{\frac{7}{8}\alpha}} \le C\delta\gamma^{d+7} s^2 \mu.$$

The details can be obtained with the same techniques as ones in [8].

2.2. Averaging and Quasilinear Equations

As usual, we shall construct the averaging transformation as the time 1−map Φ_F^1 of the flow generated by a Hamiltonian F. Let F has the following form:

$$F = \sum_{0\ne|k|\le K_+} (F_{k00} + \langle F_{k10}, y\rangle + \langle F_{k01}, z\rangle + \langle F_{k02} z, z\rangle) e^{\sqrt{-1}\langle k,x\rangle},$$

where F_{ijk}=$F_{ijk}(y, z)$, differing from the usual linear iteration. Let $[R]$=$\int_{T^d} R(x, \cdot)dx$ be the average of truncation R. Substituting F into the equation

$$\{N, F\} + R - [R] = 0, \tag{2.9}$$

where $\{\cdot,\cdot\}$ represents Poisson brackets, then

$$\begin{aligned}\bar{H} &= H\circ\Phi_F^1\\ &= (N+R)\circ\Phi_F^1+(P-R)\circ\Phi_F^1\\ &= N+[R]+\bar{P}_+(x,y,z;\lambda)\\ &= \bar{e}_+ + \langle\bar{\omega}_+,y\rangle+\frac{\delta}{2}\langle y,Ay\rangle+\delta\hat{y}(y)+\frac{\delta}{2}\langle z,M_+z\rangle+\delta\hat{z}(z)+\langle P_{001},z\rangle+\bar{P}_+,\end{aligned}$$

where

$$\begin{aligned}\bar{e}_+ &= e+P_{000},\\ \bar{\omega}_+ &= \omega+P_{010},\\ M_+ &= M+P_{002},\\ \bar{P}_+(x,y,z) &= \int_0^1\{R_t,F\}\circ\Phi_F^t dt+(P-R)\circ\Phi_F^1,\\ R_t &= (1-t)[R]+tR.\end{aligned}$$

Let Y, Z, p_{010}, p_{001} be the vectors formed by the n components (maybe not the first n components) of y, z, P_{010}, P_{001}, respectively, and denote

$$\hat{y}(y)=\hat{y}(\begin{pmatrix}Y\\0\end{pmatrix}),$$

$$\hat{z}(z)=\hat{z}(\begin{pmatrix}Z\\0\end{pmatrix}).$$

Then by implicit function theorem, the equation

$$\delta\mathcal{A}Y+\delta\partial_Y\hat{y}(Y)=-p_{010}, \tag{2.10}$$

$$\delta M_+Z+\delta\partial_Z\hat{z}(Z)=-p_{001} \tag{2.11}$$

admits a unique solution Y^*, Z^*, respectively, on $D(s)$, which also smoothly depends on λ, where $\mathcal{A}$ is an $n\times n$ minors of A. Define

$$y^*=\begin{pmatrix}Y^*\\0\end{pmatrix},$$

$$z^*=\begin{pmatrix}Z^*\\0\end{pmatrix}.$$

By (2.10) and (2.11), we clearly have

$$\delta Ay^*+\delta\partial_y\hat{y}(y^*)=-\begin{pmatrix}p_{010}\\0\end{pmatrix},$$

$$\delta M_+z^*+\delta\partial_z\hat{z}(z^*)=-\begin{pmatrix}p_{001}\\0\end{pmatrix}.$$

Consider the following translation:

$$\Phi : x \to x, y \to y + y^*, z \to z + z^*,$$

then

$$\begin{aligned} H_+ &= \bar{H}_+ \circ \phi \\ &= e_+ + \langle \omega_+, y \rangle + h_+(y,z) + P_+(x,y,z), \end{aligned}$$

where

$$\begin{aligned} e_+ &= \bar{e}_+ + \langle \bar{\omega}_+, y^* \rangle + \frac{\delta}{2}\langle y^*, Ay^* \rangle + \delta\hat{y}(y^*) + \frac{\delta}{2}\langle z^*, M_+ z^* \rangle + \delta\hat{z}(z^*) + \langle P_{001}, z^* \rangle, \\ \omega_+ &= \omega + P_{010} - \begin{pmatrix} p_{010} \\ 0 \end{pmatrix}, \\ h_+(y,z) &= \delta\hat{y}(y+y^*) - \delta\hat{y}(y^*) - \delta\langle \partial_y \hat{y}(y^*), y \rangle - \frac{\delta}{2}\langle \partial_y^2 \hat{y}(y^*) y,\ y \rangle + \delta\hat{z}(z+z^*) \\ &\quad - \delta\hat{z}(z^*) - \delta\langle \partial_z \hat{z}(z^*), z \rangle - \frac{\delta}{2}\langle \partial_z^2 \hat{z}(z^*) z,\ z \rangle + \frac{\delta}{2}\langle y, Ay \rangle + \frac{\delta}{2}\langle \partial_y^2 \hat{y}(y^*) y,\ y \rangle \\ &\quad + \frac{\delta}{2}\langle z, M_+ z \rangle + \frac{\delta}{2}\langle \partial_z^2 \hat{z}(z^*) z,\ z \rangle, \\ P_+ &= \bar{P}_+(x,y,z;\lambda) \circ \phi. \end{aligned} \tag{2.12}$$

Now, we will solve the homological equations.

Lemma 2.1 (p.256, [17]). *Let A, B, C be $r \times r$, $s \times s$, $r \times s$ matrices, respectively, and let X be an $r \times s$ unknown matrix. Then the matrix equation*

$$AX + XB = C,$$

is solvable if and only if the vector equation

$$(E_s \otimes A + B^T \otimes E_r)X' = C'$$

is solvable, where $X' = (X_1^T, \cdots, X_s^T)^T$ and $C' = (C_1^T, \cdots, C_s^T)^T$ with $X = (X_1, \cdots, X_s)$ and $C = (C_1, \cdots, C_s)$. Moreover,

$$||X'|| \le ||(E_s \otimes A + B^T \otimes E_r)^{-1}|| \cdot |C'|.$$

In view of **Lemma2.1** and (2.9), by comparing coefficients, we have

$$\sqrt{-1}\langle k, \omega + \partial_y h(y) \rangle F_{k00} = P_{k00}, \tag{2.13}$$

$$\sqrt{-1}\langle k, \omega + \partial_y h(y) \rangle F_{k10} = P_{k10}, \tag{2.14}$$

$$[-\sqrt{-1}\langle k, \omega + \partial_y h(y) \rangle I_{2m_0} + \partial_z h_1(z) J] F_{k01} = -P_{k01}, \tag{2.15}$$

$$\begin{aligned} &[-\sqrt{-1}\langle k, \omega + \partial_y h(y) \rangle I_{4m_0^2} + (\partial_z h_1(z) J) \otimes I_{2m_0} \\ &\quad + I_{2m_0} \otimes (\partial_z h_1(z) J)] F_{k02} = -P_{k02}, \end{aligned} \tag{2.16}$$

where

$$\begin{aligned} h(y) &= \frac{\delta}{2}\langle y, Ay \rangle + \delta\hat{y}(y), \\ h(z) &= \frac{\delta}{2}\langle z, Mz \rangle + \delta\hat{z}(z) \end{aligned}$$

$$= zh_1(z).$$

To control the norm of F, we solve homological equations on the set

$$\Lambda_+(g,G) = \{\lambda \in \Lambda(g,G) : |\langle k,\omega\rangle| > \frac{\gamma}{|k|^\tau},\ |\det A_1| > \frac{\gamma^{2m_0}}{|k|^{2\tau m_0}},$$
$$|\det A_2| > \frac{\gamma^{4m_0^2}}{|k|^{4m_0^2\tau}} \quad \text{for all} \quad 0 < |k| \le K_+\},$$

where

$$A_1 = -\sqrt{-1}\langle k,\omega\rangle I_{2m_0} + \partial_z h_1(z)J, \tag{2.17}$$
$$A_2 = -\sqrt{-1}\langle k,\omega\rangle I_{4m_0^2} + (\partial_z h_1(z)J)\otimes I_{2m_0} + I_{2m_0}\otimes(\partial_z h_1(z)J). \tag{2.18}$$

In the following, we also will use the similar notations.

2.3. Estimate on N_+

Theorem 2.1. *The following facts hold.*

(i) F_{k00}, F_{k10}, F_{k01}, F_{k02} *satisfy the following properties:*

$$|\partial_\lambda^l F_{k00}| \le C\delta\Gamma(r-r_+)s^2\mu, \tag{2.19}$$
$$|\partial_\lambda^l F_{k10}| \le C\delta\Gamma(r-r_+)\mu, \tag{2.20}$$
$$|\partial_\lambda^l F_{k01}| \le C\delta\Gamma(r-r_+)s\mu, \tag{2.21}$$
$$|\partial_\lambda^l F_{k02}| \le C\delta\Gamma(r-r_+)\mu, \tag{2.22}$$

where C is a constant;

(ii) F *can be extended to functions of Hölder class* $C^{5,d-1+\sigma_0}(\hat{D}(\beta_0)\times\Lambda_0)$, *where* $0<\sigma_0<1$ *is fixed. Moreover, there is a constant C such that*

$$\|F\|_{C^{5,d-1+\sigma_0}}(\hat{D}(\beta_0)\times\Lambda_0) \le C\delta\mu s^2\Gamma(r-r_+).$$

Proof. The proofs of (2.19)-(2.22) are analogic. We are going to primarily testify (2.19). Let $(y,z,\lambda)\in D(s)\times\Lambda_+$. By the definition of M^*,

$$|\partial_y h(y)| \le (M^*+1)|y| < (M^*+1)s < \tfrac{\gamma}{2|k|^\tau},$$

provided

$$2s < \frac{r-r_+}{(M^*+1)K_+^\tau}. \tag{2.23}$$

It follows that

$$L_k = |\langle k,\omega(\lambda)+\partial_y h(y)\rangle| > \tfrac{\gamma}{2|k|^\tau}.$$

Hence, L_k is nonvanishing on Λ_+. Therefore,

$$F_{k00} = L_k^{-1}P_{k00}, \quad \forall\ (y,z,\lambda)\in D(s)\times\Lambda_+,\ 0<|k|\le K_+. \tag{2.24}$$

By Cauchy estimate, we know that

$$|\partial_\lambda^l P_{k00}| \le |\partial_\lambda^l P|e^{-|k|r} \le C\delta\gamma^{d+7}s^2\mu e^{-|k|r}, \quad |l| \le d.$$

Retrospecting back to differential and integral calculus as well as using (2.23), inductively, we deduce that

$$|\partial^q L_k^{-1}| \le C|k|^{|q|}|L_k^{-1}|^{|q|+1}. \tag{2.25}$$

Using (2.24) and (2.25), we get

$$\begin{aligned}
|\partial_\lambda^l F_{k00}| &= |\partial_\lambda^l (L_k^{-1} P_{k00})| \\
&= |\sum_{|l'|=1}^{|l|} \begin{pmatrix} l \\ l' \end{pmatrix} (\partial_\lambda^{l-l'} L_k^{-1})\ (\partial_\lambda^{l'} P_{k00})| \\
&\le \sum_{|i|+|j|=|l|} \frac{|k|^{\tau(|i|+|j|+|l|)+|i|+|j|+|l|}}{\gamma^{|i|+|j|+|l|+1}} C\delta\gamma^{d+7}s^2\mu e^{-|k|r} \\
&\le C\delta s^2\mu\Gamma(r-r_+).
\end{aligned}$$

With the same method,

$$|\partial_\lambda^l F_{k10}| = |\partial_\lambda^l (L_k^{-1} P_{k10})| \le C\delta\Gamma(r-r_+)\mu.$$

Let

$$|L_k^{(2)}| = |\det(A_1 + \langle k, \partial_y h(y)\rangle I_{2m_0})|.$$

Thus,

$$|L_k^{(2)}| \ge \tfrac{\gamma^{2m_0}}{|k|^{2m_0\tau}} - C_0 s K_+^{2m_0} \ge \tfrac{\gamma^{2m_0}}{2|k|^{2m_0\tau}},$$

provided

$$2sC_0 K_+^{4m_0^2\mu+4m_0^2+1} < \gamma^{4m_0^2}. \tag{2.26}$$

Therefore,

$$|\partial_\lambda^l F_{k01}| = |\sum_{|l'|=1}^{|l|} \begin{pmatrix} l \\ l' \end{pmatrix} (\partial_\lambda^{l-l'} (L_k^{(2)})^{-1})\ (\partial_\lambda^{l'} P_{k01})| \le C\delta s\mu\Gamma(r-r_+).$$

Similarly,

$$|\partial_\lambda^l F_{k02}| \le C\delta\Gamma(r-r_+)\mu.$$

(ii) follows from the standard Whitney extension theorem [22, 28].

□

Theorem 2.2. *The following hold:*

(i) There is a constant C such that the following hold for all $|l| \le d$

$$|\partial_\lambda^l y^*|_{\Lambda_+} \le C\gamma^{d+7}\mu, \tag{2.27}$$

$$|\partial_\lambda^l z^*|_{\Lambda_+} \le C\gamma^{d+7}\mu, \tag{2.28}$$

$$|\partial_\lambda^l e_+ - \partial_\lambda^l e|_{\Lambda_+} \le C\delta\gamma^{d+7}\mu, \tag{2.29}$$

$$|\partial_\lambda^l \omega_+ - \partial_\lambda^l \omega|_{\Lambda_+} \le C\delta\gamma^{d+7}\mu, \tag{2.30}$$

$$|\partial_y^i\partial_z^j\partial_\lambda^l h_+(y,z) - \partial_y^i\partial_z^j\partial_\lambda^l h(y,z)|_{\Lambda_+} \le C\delta\gamma^{d+7}\mu, \tag{2.31}$$

where $|i| + |j| \le 7$;

(ii) y^**,* z^* *can be extended to functions of Hölder class* $C^{5,d-1+\sigma_0}(\hat{D}(\beta_0)\times\Lambda_0)$*, respectively, where* $0<\sigma_0<1$ *is fixed. Moreover, there is a constant* C *such that*

$$\|y^*\|_{C^{d-1+\sigma_0}(\Lambda_0)} \le C\mu\Gamma(r-r_+),$$
$$\|z^*\|_{C^{d-1+\sigma_0}(\Lambda_0)} \le C\mu\Gamma(r-r_+).$$

Proof. Since the proofs of (2.27)-(2.31) are similar, we will mainly prove (2.27). Let $\lambda\in\Lambda_+$. Denote

$$M_* = \max_{\lambda\in\Lambda_0}|\mathcal{A}_0^{-1}(\lambda)|+1,$$
$$B(y,\lambda) = \mathcal{A} + (\int_0^1 \partial_y^2\hat{y}(\theta y)d\theta)y.$$

Then by (2.10),

$$B(Y^*)Y^* = -p_{010}. \tag{2.32}$$

Supposing

$$\max_{|l|\le d,2|i|+|j|\le 7}|\partial_\lambda^l\partial_y^i\partial_z^j h - \partial_\lambda^l\partial_y^i\partial_z^j h_0|_{D(s)\times\Lambda_+} \le \mu_0^{\frac{1}{2}}\delta, \tag{2.33}$$

using the same method as one in [18], we can get that $B(Y^*)$ is nonsingular and

$$|B^{-1}(Y^*)| \le \frac{|\mathcal{A}_0^{-1}|}{1-|\mathcal{A}_0-B(Y^*)||\mathcal{A}_0^{-1}|} \le \frac{2M^*}{\delta}.$$

Hence,

$$|y^*| = |Y^*| \le 2\frac{M_*}{\delta}|\partial_y P|_{D(s)} \le 2M_*\gamma^{d+7}\mu.$$

Differentiating (2.32) with respect to λ, under assumption

$$4\delta M_*(M^*+1)\gamma^{d+7}\mu < \frac{1}{2} \tag{2.34}$$

and induction, we have

$$|\partial_\lambda^l y^*| < CM_*\gamma^{d+7}\mu.$$

In the same way, we can easily get

$$|\partial_\lambda^l z^*| < 2M_*\gamma^{d+7}\mu.$$

We place the maximal value of M_+ as M_*, too. Similarly,

$$|\partial_\lambda^l e_+ - \partial_\lambda^l e|_{\Lambda_+} \le C\delta\gamma^{d+7}\mu,$$
$$|\partial_\lambda^l \omega_+ - \partial_\lambda^l \omega|_{\Lambda_+} \le C\delta\gamma^{d+7}\mu,$$
$$|\partial_y^i\partial_z^j\partial_\lambda^l h_+(y,z) - \partial_y^i\partial_z^j\partial_\lambda^l h(y,z)|_{\Lambda_+} \le C\delta\gamma^{d+7}\mu,$$

where $|i|+|j|\le 7$.

For details of (ii), refer to [22, 28]. □

2.4. Estimate on Φ_+

By estimates of $\partial_\lambda^l F_{k00}$, $\partial_\lambda^l F_{k10}$, $\partial_\lambda^l F_{k01}$, $\partial_\lambda^l F_{k02}$ and the form of F,

$$|\partial_\lambda^l F| \leq C\delta s^2 \mu\Gamma(r - r_+).$$

With Cauchy estimate, we have

$$(r - r_+)|\partial_\lambda \partial_y F|, \quad s^2|\partial_\lambda \partial_x F|, \quad s|\partial_\lambda \partial_z F| \quad \leq \quad C\delta s^2 \mu\Gamma(r - r_+). \tag{2.35}$$

Inductively,

$$|D^n \partial_\lambda^l F| \leq C\delta\mu\Gamma(r - r_+), \ |n| \leq 4.$$

Denote $\Phi_F^t = (\phi_1^t, \phi_2^t, \phi_3^t)^T$, where ϕ_1^t, ϕ_2^t and ϕ_3^t are components of Φ_F^t in directions of x, y, z, respectively. Let (x, y, z) be any point in $D_{\frac{1}{4}\alpha}$ and let $t_* = \sup\{t \in [0, 1] : \Phi_F^t(x, y, z) \in D_\alpha\}$. We note that $D_\alpha \subset \hat{D}(s)$. Using the identity

$$\Phi_F^t = id + \textstyle\int_0^t X_F \circ \Phi_F^u du,$$

where $X_F = (F_y, -F_x, JF_z)^T$ denotes the vector field generated by F, we have

$$|\Phi_1^t(x, y, z) - x| \leq \textstyle\int_0^t |F_y \circ \Phi_F^u|_{D_\alpha} du \leq |F_y|_{\hat{D}(s)} < C\delta\Gamma(r - r_+)\mu < \frac{1}{8}(r - r_+),$$

with assumption

$$C\delta\Gamma(r - r_+)\mu < \frac{1}{8}(r - r_+); \tag{2.36}$$

$$|\Phi_2^t(x, y, z) - y| \leq \textstyle\int_0^t |F_x \circ \Phi_F^u|_{D_\alpha} du \leq |F_x|_{\hat{D}(s)} < C\delta s\mu\Gamma(r - r_+) \leq \frac{1}{8}\alpha,$$

with hypothesis

$$C\delta s\mu\Gamma(r - r_+) \leq \frac{1}{8}\alpha; \tag{2.37}$$

$$|\Phi_3^t(x, y, z) - z| \leq \textstyle\int_0^t |JF_z \circ \Phi_F^u|_{D_\alpha} du \leq |F_z|_{\hat{D}(s)} < C\delta s\mu\Gamma(r - r_+) < \frac{1}{8}\alpha s,$$

supposing

$$C\delta s\mu\Gamma(r - r_+) < \frac{1}{8}\alpha s. \tag{2.38}$$

Therefore, $\Phi_F^t : D_{\frac{1}{4}\alpha} \to D_{\frac{1}{2}\alpha}$. By estimations of $|\partial_\lambda^l y^*|_{\Lambda_+}$ and $|\partial_\lambda^l z^*|_{\Lambda_+}$, it is easy to see $\phi : D_{\frac{1}{8}\alpha} \to D_{\frac{1}{4}\alpha}$.

The above imply that $\Phi_+ = \Phi_F^t \circ \phi : D_+ \to D_{\frac{1}{2}\alpha}$ is well defined, symplectic and real analytic for all $\lambda \in \Lambda_+$. We, now, consider Φ_+ on the domain $\tilde{D}_+$.

Theorem 2.3. *Let F, y^* and z^* be the extended functions defined as above. Then*

$$\Phi_+ = \Phi_F^1 \circ \phi : \hat{D}_+ \to D(r, \beta)$$

is of classes C^4 and also depends $C^{d-1+\sigma_0}$ smoothly on $\lambda \in \Lambda_0$, where σ_0 define as above. Moreover, there is a constant C such that:

$$\|\Phi_+ - id\|_{C^{4,d-1+\sigma_0}(\tilde{D}_+ \times \Lambda_0)} \leq C\mu\Gamma(r - r_+).$$

Proof. It is easy to see that Φ_+ maps $\hat{D}_+$ into $D(r,\beta)$ for all $\lambda \in \Lambda_0$. We note that

$$\Phi_F^t = id + \int_0^t X_F \circ \Phi_F^u du, 0 < t < 1,$$
$$\|X_F\|_{C^{4,d-1+\sigma_0}(\hat{D}(\beta_0)\times\Lambda_0)} \le C\|F\|_{C^{5,d-1+\sigma_0}(\hat{D}(\beta_0)\times\Lambda_0)},$$

where $X_F = (F_y, -F_x, JF_z)^T$ is the vector field generated by F. Supposing

$$C\mu\delta\Gamma(r-r_+) < \frac{1}{8}(r-r_+), \tag{2.39}$$
$$C\mu\delta\Gamma(r-r_+) + C\delta\mu < \beta - \beta_+, \tag{2.40}$$

by applying Gronwall inequality and the definition of Φ_F^t, inductively, we have that on $\tilde{D}_+ \times \Lambda_0$,

$$|\Phi_F^t - id|, |\partial_y \Phi_F^t - I_{2n}|, |\partial_y^j \Phi_F^t| \le C\delta\mu\Gamma(r-r_+). \tag{2.41}$$

This theorem holds with the help of the identity

$$\Phi_+ - id = (\Phi_F^1 - id)\circ\phi + \begin{pmatrix} 0 \\ y^* \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z^* \end{pmatrix}.$$

□

2.5. Frequency and Ratios

As for the preservation of the energy on resonant tori, y^* and z^* are defined so that $e_+ = e = E$. Therefore, we consider equations

$$\langle \bar{\omega}_+, y^*\rangle + \frac{\delta}{2}\langle y^*, Ay^*\rangle + \delta\hat{y}(y^*) + \frac{\delta}{2}\langle z^*, M_+ z^*\rangle + \delta\hat{z}(z^*) + \langle P_{001}, z^*\rangle = 0,$$
$$\delta M_+ z^* + \delta\partial_z\hat{z}(z^*) = -p_{001},$$

which, by implicit function theorem, clearly admits a local smooth solution y^*, z^* on $\mathcal{M}$, respectively.

As for the preservation of ratios for the toral frequency on the resonant tori, y^* and z^* need to be choose such that $e_+ = e = E$, and $[\omega_{+,i_1} : \cdots : \omega_{+,i_n}] = [\omega_{i_1} : \cdots : \omega_{i_n}]$. Therefore, we consider equations

$$(\mathcal{A} + \frac{\partial\hat{h}}{\partial(y_{i_1},\cdots,y_{i_n})}(y^*))(y^*_{i_1},\cdots,y^*_{i_n})^T - t^*(\omega_{i_1} : \cdots : \omega_{i_n})^T = -(p_{010,i_1},\cdots,p_{010,i_n})^T,$$
$$\langle(\bar{\omega}_{i_1},\cdots,\bar{\omega}_{i_n})^T, (y^*_{i_1},\cdots,y^*_{i_n})\rangle + \frac{\delta}{2}\langle y^*, Ay^*\rangle + \delta\hat{y}(y^*)$$
$$+ \frac{\delta}{2}\langle z^*, M_+ z^*\rangle + \delta\hat{z}(z^*) + \langle P_{001}, z^*\rangle = 0,$$
$$\delta M_+ z^* + \delta\partial_z\hat{z}(z^*) = -p_{001},$$

which, by subisoenergetic nondegenerate condition $(A1')$ and implicit function theorem, admit a local smooth solution (y^*, z^*, t^*) such that $y^*_j = 0$ and $z^*_j = 0$ if $j \overline{\in} \{i_1, \cdots, i_n\}$.

Under the symplectic transformation

$$\Phi_+ = \Phi_F^1 \circ \phi,$$

the new Hamiltonian reads

$$H \circ \Phi_+ = N_+ + P_+,$$

where

$$\begin{aligned}
N_+ &= E + \langle \omega_+, y \rangle + h_+(y, z), \\
P_+ &= \bar{P}_+ \circ \phi, \\
\omega_+ &= \omega + P_{010} + Ay^* + \partial_y \hat{h}(y^*),
\end{aligned}$$

and $h_+(y,z)$, A_+ and $\hat{h}(y)$ have the same forms as above. Thus, the new normal form is reduced to the desired case.

Theorem 2.4. *Assume* (2.26). *Then for all* $0 < |k| \le K_+$, $\lambda \in \Lambda_+$,

$$|\langle k, \omega_+ \rangle| > \frac{\gamma_+}{|k|^\tau}, \ |\det A_{1,+}| > \frac{\gamma_+^{2m_0}}{|k|^{2\tau m_0}}, \ |\det A_{2,+}| > \frac{\gamma_+^{4m_0^2}}{|k|^{4m_0^2 \tau}},$$

where

$$\begin{aligned}
A_{1,+} &= -\sqrt{-1}\langle k, \omega_+ \rangle I_{2m_0} + \partial_z h_{1,+}(z) J, \\
A_{2,+} &= -\sqrt{-1}\langle k, \omega_+ \rangle I_{4m_0^2} + (\partial_z h_{1,+}(z) J) \otimes I_{2m_0} + I_{2m_0} \otimes (\partial_z h_{1,+}(z) J).
\end{aligned}$$

Proof. By the definition of Λ_+ and (2.26), this theorem can be proved easily. □

2.6. Estimate on P_+

Theorem 2.5. *There is a constant* C, *such that, on* $D_+ \times \Lambda_+$,

$$|\partial_\lambda^l P_+| \le C\delta\gamma^{d+7} s^2 \mu^2 (\Gamma(r - r_+) + 1), \quad |l| \le d.$$

Proof. We know

$$\begin{aligned}
P_+ &= \bar{P}_+ \circ \phi \\
&= \left(\int_0^1 \{R_t, F\} \circ \Phi_F^t dt + (P - R) \circ \Phi_F^1\right) \circ \phi.
\end{aligned}$$

By above estimates, we see that, for all $|l| \le d$, $0 \le t \le 1$,

$$\begin{aligned}
|\partial_\lambda^l \{R_t, F\} \circ \Phi_F^t|_{D_{\frac{1}{4}\alpha} \times \Lambda_+} &\le C\delta\gamma^{d+7} s^2 \mu \Gamma(r - r_+), \\
|\partial_\lambda^l (P - R) \circ \Phi_F^t|_{D_{\frac{1}{4}\alpha} \times \Lambda_+} &\le C\delta\gamma^{d+7} s^2 \mu.
\end{aligned}$$

Hence, by the definition of P_+,

$$|\partial_\lambda^l P_+| \le C\delta\gamma^{d+7} s^2 \mu^2 (\Gamma(r - r_+) + 1).$$

Let C_0 be the maximal one among $C's$ we mentioned above and define

$$\mu_+ = 64 C_0 \mu^{1+\sigma}.$$

Assume

$$\mu^{\sigma}(\Gamma(r-r_+)+1) \le \frac{\gamma_+^{d+7}}{\gamma^{d+7}}, \quad \text{on } D_+ \times \Lambda_+, \tag{2.42}$$

then

$$\begin{aligned} |\partial_\lambda^l P_+| &\le 64C_0 \delta s_+^2 \mu^{1+\sigma} \mu^{\frac{1}{3}-2\sigma} (\mu^\sigma \gamma^{d+7}(\Gamma(r-r_+)+1)) \\ &\le \delta \gamma_+^{d+7} s_+^2 \mu_+, \qquad |l| \le d. \end{aligned}$$

□

This completes one KAM step.

3. Proof of Main Results

3.1. Iteration Lemma

Consider (2.1) and let r_0, s_0, γ_0, β_0, μ_0, Λ_0, H_0, N_0, e_0, ω_0, h_0, A_0, P_0 be given as above. And let $\hat{D}_0 = D(r_0, \beta_0)$. We define the following sequences, inductively, for all $\nu = 1, 2, \cdots$:

$$r_\nu = r_0\Big(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}\Big), \qquad s_\nu = \frac{1}{8}\alpha_{\nu-1} s_{\nu-1}, \qquad \alpha_\nu = \mu_\nu^{2\sigma} = \mu_\nu^{\frac{1}{3}},$$

$$\beta_\nu = \beta_0\Big(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}\Big), \qquad \mu_\nu = 64C_0 \mu_{\nu-1}^{1+\sigma}, \qquad \gamma_\nu = \gamma_0\Big(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}\Big),$$

$$K_\nu = \Big(\Big[\log \frac{1}{\mu_{\nu-1}}\Big]+1\Big)^{3\eta}, \ \tilde{D}_\nu = D\Big(r_\nu + \frac{3}{4}(r_{\nu-1} - r_\nu), \beta_\nu\Big), \ D_\nu = D(r_\nu, s_\nu),$$

$$\Lambda_\nu(g,G) = \Big\{\lambda \in \Lambda_{\nu-1}(g,G) : |\langle k, \omega_{\nu-1}\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, |\det A_{1,\nu-1}| > \frac{\gamma_{\nu-1}^{2m_0}}{|k|^{2m_0\tau}},$$
$$|\det A_{2,\nu-1}| > \frac{\gamma_{\nu-1}^{4m_0^2}}{|k|^{4\tau m_0^2}} \quad \textit{for all} \quad 0 < k \le K_\nu\Big\}.$$

Theorem 3.1. *If* (2.6) *holds for a sufficiently small* $\mu_0 = \mu_0(r_0, \beta_0, m, d, \tau)$*, then the KAM step described as one in Section 2 is valid for all* $\nu = 0, 1, \cdots$*, resulting in sequences* $\Lambda_\nu, H_\nu, N_\nu, e_\nu, \omega_\nu, h_\nu, A_\nu, P_\nu, \Phi_\nu$*,* $\nu = 1, 2, \cdots$*, with the following properities:*

(i) $\Phi_\nu : \hat{D} \times \Lambda_\nu \to \hat{D}_{\nu-1}$ *is symplectic for each* $\lambda \in \Lambda_\nu$*, and is of class* $C^{4,d-1+\sigma_0}$*, where* $0 < \sigma_0 < 1$ *is fixed, and*

$$\|\Phi_\nu - id\|_{C^{4,d-1+\sigma_0}(\hat{D}_\nu \times \Lambda_\nu)} \le \frac{\mu^{\frac{1}{2}}}{2^\nu}. \tag{3.1}$$

Moreover, on $\hat{D}_\nu \times \Lambda_\nu$*,*

$$H_\nu = H_{\nu-1} \circ \Phi_\nu = N_\nu + P_\nu,$$

where

$$H_\nu = N_\nu + P_\nu,$$

$$N_\nu = e_\nu + \langle \omega_\nu, y\rangle + h_\nu,$$
$$h_\nu = \frac{\delta}{2}\langle y, A_\nu y\rangle + \delta\hat{y}(y) + \frac{\delta}{2}\langle z, M_\nu z\rangle + \delta\hat{z}(z),$$

A_ν has $n\times n$ real nonsingular minor $\mathcal{A}_\nu$, $\varepsilon\hat{y}(y) = O(|y|^3)$, $\varepsilon\hat{z}(z) = O(|z|^3)$;

(ii) $(\omega_\nu(\lambda))_q = (\omega_{\nu-1}(\lambda))_q$ for all $q = 1, 2, \cdots n$ and $\lambda \in \Lambda_\nu$;

(iii) For all $|l| \le d$,

$$|\partial_\lambda^l e_\nu - \partial_\lambda^l e_{\nu-1}|_{\Lambda_\nu} \le \delta\gamma_0^{d+6}\frac{\mu}{2^\nu}; \tag{3.2}$$
$$|\partial_\lambda^l e_\nu - \partial_\lambda^l e_0|_{\Lambda_\nu} \le \delta\gamma_0^{d+6}\mu; \tag{3.3}$$
$$|\partial_\lambda^l \omega_\nu - \partial_\lambda^l \omega_{\nu-1}|_{\Lambda_\nu} \le \delta\gamma_0^{d+6}\frac{\mu}{2^\nu}; \tag{3.4}$$
$$|\partial_\lambda^l \omega_\nu - \partial_\lambda^l \omega_0|_{\Lambda_\nu} \le \delta\gamma_0^{d+6}\mu; \tag{3.5}$$
$$|\partial_\lambda^l\partial_y^i\partial_z^j h_\nu - \partial_\lambda^l\partial_y^i\partial_z^j h_{\nu-1}|_{D_\nu\times\Lambda_\nu} \le \delta\gamma_0^{d+6}\frac{\mu^{\frac{1}{2}}}{2^\nu}, \quad |i|+|j|\le 7; \tag{3.6}$$
$$|\partial_\lambda^l\partial_y^i\partial_z^j h_\nu - \partial_\lambda^l\partial_y^i\partial_z^j h_0|_{D_\nu\times\Lambda_\nu} \le \delta\gamma_0^{d+6}\mu^{\frac{1}{2}}, \quad |i|+|j|\le 7; \tag{3.7}$$
$$|\partial_\lambda^l P_\nu|_{D_\nu\times\Lambda_\nu} \le \delta\gamma_\nu^{d+7}s_\nu^2\mu_\nu; \tag{3.8}$$

(iv)

$$\Lambda_\nu(g,G) = \{\lambda\in\Lambda_{\nu-1}(g,G) : |\langle k, \omega_{\nu-1}\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, |\det A_{1,\nu-1}| > \frac{\gamma_{\nu-1}^{2m_0}}{|k|^{2m_0\tau}},$$
$$|\det A_{2,\nu-1}| > \frac{\gamma_{\nu-1}^{4m_0^2}}{|k|^{4\tau m_0^2}} \textit{ for all } 0<|k|\le K_\nu\}.$$

Proof. Actually, it suffices to verify above assumptions that we put forward for all ν. The method of verifying hypothesises mentioned above is standard. For sake of brevity we here only show the proof of (2.42). As for details of other hypothesises', we refer readers to [8]. For simplicity, we let $r_0 = \beta_0 = 1$. By choosing μ_0 small, we also see that other assumptions are hold for $\nu = 0$. By the definition of μ_ν, we have that

$$\mu_\nu = (64C_0)^{(1+\sigma)^\nu - 1}\mu_0{}^{(1+\sigma)^\nu}. \tag{3.9}$$

Therefore,

$$\mu_\nu = 64C_0\mu_{\nu-1}^{1+\sigma} < \cdots < \frac{1}{\zeta^\nu}\mu_0, \tag{3.10}$$

where $\zeta > 1$ and

$$\mu_0 < (\frac{1}{64C_0\zeta})^\sigma < 1. \tag{3.11}$$

Denote

$$\Gamma_\nu = \Gamma(r_\nu - r_{\nu-1}),$$

where

$$r_\nu - r_{\nu-1} = \frac{1}{2^{\nu+2}} = \frac{\beta_\nu - \beta_{\nu+1}}{\beta_0}. \tag{3.12}$$

Since

$$\begin{aligned}\Gamma_\nu &\le \int_1^\infty \lambda^{4m_0^2(|l|+d+8)\tau+d+|l|+8} e^{-\frac{\lambda}{2^{\nu+6}}} d\lambda \\ &\le 4m_0^2(|l|+d+8)\tau+d+|l|+8)!2^{(\nu+6)(4m_0^2(|l|+d+8)\tau+d+|l|+8)},\end{aligned}$$

it is clear that ζ is sufficiently large, then

$$\mu_\nu^\sigma \Gamma_\nu < \mu_\nu^\sigma(\Gamma_\nu+1) < \frac{\gamma_{\nu+1}^{d+7}}{\gamma_\nu^{d+7}}.$$

Therefore, (2.42) holds for all $\nu \ge 1$.

Above all, KAM step described in Section 2 are valid for all ν, which gives the desired sequences stated as this theorem. Now, we accomplish proofs of (i), (ii) and (iii). The proof of (iv) is standard. Details could be found in [8]. □

3.2. Convergence and Measure Estimates

Let

$$\Psi^\nu = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_\nu, \quad \nu = 1, 2, \cdots.$$

Then $\Psi^\nu : \tilde{D}_\nu \times \Lambda_0 \to \tilde{D}_0$, and

$$\begin{aligned} H_0 \circ \Psi^\nu = H_\nu = N_\nu + P_\nu, \\ N_\nu = e_\nu + \langle \omega_\nu(\lambda), y\rangle + h_\nu(y,\lambda), \quad \nu = 0, 1, \cdots,\end{aligned}$$

where $\Psi_0 = id$.

Simply, N_ν converges uniformly to N_∞, P_ν converges uniformly to P_∞ and $\partial_y^j P_\infty = 0$, $|j| \le 2$. Details refer to [8].

Hence for each $\lambda \in \Lambda_*$, $T^d \times \{0\}$ is an analytic invariant torus of H_∞ with the toral frequency $\omega_\infty(\lambda)$, which, by the definition of Λ_ν, satisfies

$$\begin{aligned} |\langle k, \omega_\infty(\lambda)\rangle| > \frac{\gamma}{2|k|^\tau} \quad \text{for all} \quad k \in Z^d/\{0\}, \\ (\omega_\infty(\lambda))_q \equiv (\omega_0(\lambda))_q \quad \text{for all} \quad 1 \le q \le n.\end{aligned}$$

Following the Whitney extention of Ψ^ν, all e_ν, ω_ν, h_ν, P_ν, ($\nu = 0, 1, \cdots$) admit uniformly $C^{d-1+\sigma_0}$ extensions in $\lambda \in \Lambda_0$ with derivatives in λ up to order $d-1$ satisfying same estimates (3.2)-(3.11). Thus, e_∞, ω_∞, h_∞, P_∞ are C^{d-1} Whitney smooth in $\lambda \in \Lambda_*$, and derivatives for $e_\infty - e_0$, $\omega_\infty - \omega_0$, $h_\infty - h_0$ satisfy similar estimates as ones in (3.3), (3.5), (3.7). Consequently, the perturbed tori form a C^{d-1} Whitney smooth family on Λ_*.

The measure estimate is the same as one in [8]. For sake of simplicity, we only show the idea. Denote

$$\Lambda_{\nu+1} = \{\lambda \in \Lambda_\nu : |\langle k, \omega_\nu\rangle| \le \frac{\gamma_\nu}{|k|^\tau}, |\det A_{1,\nu}| \le \frac{\gamma_\nu^{2m_0}}{|k|^{2m_0\tau}}, |\det A_{2,\nu}| \le \frac{\gamma_\nu^{4m_0^2}}{|k|^{4m_0^2\tau}},$$

$$
\begin{aligned}
&\qquad K_{\nu-1} < |k| \le K_\nu\} \\
&= S_1 \bigcup S_2 \bigcup S_3,
\end{aligned}
$$

where

$$
\begin{aligned}
S_1 &= \{\lambda \in \Lambda_\nu : |\langle k, \omega_\nu\rangle| \le \frac{\gamma_\nu}{|k|^\tau}, K_{\nu-1} < |k| \le K_\nu\}, \\
S_2 &= \{\lambda \in \Lambda_\nu : |\det A_{1,\nu}| \le \frac{\gamma_\nu^{2m_0}}{|k|^{2m_0\tau}}, K_{\nu-1} < |k| \le K_\nu\}, \\
S_3 &= \{\lambda \in \Lambda_\nu : |\det A_{2,\nu}| \le \frac{\gamma_\nu^{4m_0^2}}{|k|^{4m_0^2\tau}}, K_{\nu-1} < |k| \le K_\nu\}.
\end{aligned}
$$

There are three cases, $d_0 = d$, $d_0 > d$ and $d_0 < d$, to be considered. When $d_0 = d$, using the same method as one in [23], $|S_1| \le c(\frac{\gamma}{|k|^{\tau+1}})^{\frac{1}{l}}$, where l is a constant. Recall

$$
\begin{aligned}
|\det A_{2,\nu}| &= |\det(-\sqrt{-1}\langle k, \omega\rangle I_{4m_0^2} + (\partial_z h_1(z)J) \otimes I_{2m_0} + I_{2m_0} \otimes (\partial_z h_1(z)J))| \\
&= ||k|^{4m_0^2} \det(-\sqrt{-1}\langle \varsigma, \omega\rangle I_{4m_0^2} + \frac{\delta}{|k|} M)|,
\end{aligned}
$$

where δ is small enough, $\varsigma = \frac{k}{|k|} \in S^d$ and $\delta M = (\partial_z h_1(z)J) \otimes I_{2m_0} + I_{2m_0} \otimes (\partial_z h_1(z)J)$. When $\langle \varsigma, \omega\rangle \neq 0$, with the continuity of determinant the leading part of $\det(-\sqrt{-1}\langle \varsigma, \omega\rangle I_{4m_0^2} + \frac{\delta}{|k|} M)$ is $\det(-\sqrt{-1}\langle \varsigma, \omega\rangle I_{4m_0^2})$. Then, $|S_3| \le c(\frac{\gamma}{|k|^{\tau+1}})^{\frac{1}{l}}$. When $\langle \varsigma, \omega\rangle = 0$, under condition $(A1)$ $rank\,\{\frac{\partial^\alpha \langle \varsigma, \omega\rangle}{\partial \lambda^\alpha} : |\alpha| \le N_1\} = 1$ for any $\varsigma \in S^d$. Since $(4m_0^2)!\{\frac{\partial^\alpha \langle \varsigma, \omega\rangle}{\partial \lambda^\alpha} : |\alpha| \le N_1\}$ is a minor of $\{\frac{\partial^\alpha (\langle \varsigma, \omega\rangle)^{4m_0^2}}{\partial \lambda^\alpha} : |\alpha| \le 4m_0^2 - 1 + N_1\}$, $rank\{\frac{\partial^\alpha (\langle \varsigma, \omega\rangle)^{4m_0^2}}{\partial \lambda^\alpha} : |\alpha| \le 4m_0^2 - 1 + N_1\} = 1$. Then, with sufficient small δ, $rank\{\frac{\partial^\alpha \det A_{2,\nu}}{\partial \lambda^\alpha} : |\alpha| \le N_2\} = 1$. Therefore, using the same step as one in [23], $|S_3| \le c(\frac{\gamma}{|k|^{\tau+1}})^{\frac{1}{l}}$. Similarly, $|S_2| \le c(\frac{\gamma}{|k|^{\tau+1}})^{\frac{1}{l}}$. Then $\Lambda_{\nu+1} \le c(\frac{\gamma}{|k|^{\tau+1}})^{\frac{1}{l}}$. Hence

$$
|\Lambda_0 \setminus \Lambda_*| \le \sum_{\nu=0}^{\infty} \sum_{K_\nu \le |k| \le K_{\nu+1}} |\Lambda_{\nu+1}| \le O(\gamma^{\frac{1}{l}}).
$$

For the other two cases, $d_0 < d$ and $d_0 > d$, with the same techniques as in [8,23] $|S_1| < c(\frac{\gamma}{|k|^{\tau+1}})^{\frac{1}{l}}$. Then, using the step mentioned above, we have $|\Lambda_0 \setminus \Lambda_*| = O(\gamma^{\frac{1}{l}})$.

Now, we complete the proof of the Theorem 1.1.

Acknowledgements. The authors are grateful to the anonymous referees for their useful suggestions which improve the contents of this article.

References

[1] V. I. Arnold, *Proof of a theorem by A. N. Kolmogorov on the preservation of quasi-periodic motions under small perturbations of the Hamiltonian*, Usp. Mat. Nauk., 1963, 113(5), 13–40.

[2] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, Berlin, 1991.

[3] G. Benettin, L. Galgani, A. Giorgilli and J. Strelcyn, *A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method*, Nuovo Cimento B, 1984, 79(2), 201–223.

[4] J. Bricmont, K. Gawedzki and A. Kupiainen, *KAM theorem and quantum field theory*, Commun. Math. Phys., 1999, 201(2), 699–727.

[5] H. Broer, G. Huitema and M. Sevryuk, *Quasi-Periodic Motions in Families of Dynamical Systems*, Springer-Verlag, Berlin, 1996.

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

[7] L. Chierchia and G. Gallavotti, *Drift and diffusion in phase space*, Ann. Inst. H. Poincaré Phys. Theor., 1994, 64(1), 1–144.

[8] S. -N. Chow, Y. Li and Y. Yi, *Persistence of invariant tori on submanifolds in Hamiltonian system*, J. Nonl. Sci., 2002, 12(6), 585–617.

[9] F. Z. Cong, T. Küpper, Y. Li and J. G. You, *KAM-type theorem on resonant surfaces for nearly integrable Hamiltonian systems*, J. Nonl. Sci., 2000, 10(1), 49–68.

[10] L. Corsi and G. Gentile, *Resonant tori of arbitrary codimension for quasi-periodically forced systems*, NoDEA Nonlinear Differential Equations Appl., 2017, 24(1), 3–24.

[11] L. H. Eliasson, *Biasymptotic solutions of perturbed integrable Hamiltonian systems*, Bol. Soc. Mat., 1994, 25(1), 57–76.

[12] L. H. Eliasson, *Absolutely convergent series expansions for quasiperiodic motions*, Math. Phys. Elect. J., 1996, 2, 1–33.

[13] G. Gallavotti, *Twistless KAM tori*, Commun. Math. Phys., 1994, 164(1), 145–156.

[14] G. Gallavotti, G. Gentile and V. Mastropietro, *Field theory and KAM tori*, Math. Phys. Elect. J., 1995, 1, 1–13.

[15] A. González-Enríquez, A. Haro and R. de la Llave, *Singularity theory for non-twist KAM tori*, Mem. Amer. Math. Soc., 2014, 1607(227), 1–115.

[16] A. N. Kolmogorov, *On conservation of conditionally periodic motions for a small change in Hamilton's function*, Dokl. Akad. Nauk. SSSR, 1954, 98, 527–530.

[17] P. Lancaster, *Theory of matrices*, Academic Press, New York, 1969.

[18] Y. Li and Y. Yi, *A quasiperiodic Poincaré's theorem*, Math. Ann., 2003, 326(4), 649–690.

[19] A. G. Medvedev, A. I. Neishtadt and D. V. Treshchëv, *Lagrangian tori near resonances of near-integrable Hamiltonian systems*, Nonlinearity, 2015, 28(7), 2105–2130.

[20] J. Moser, *On invariant curves of area preserving mappings of an annulus*, Nachr. Akad. Wiss. Gött. Math. Phys. K1, 1962, 1–20.

[21] H. Poincaré, *Les Méthodes Nouvelles de la Mécaniques Céleste*, I-III, Gauthier-Villars,1892, 1893, 1899. (The English translation: New methods of celestial mechanics, AIP Press,Williston, 1992.)

[22] J. Pöschel, *Integrability of Hamiltonian systems on cantor sets*, Commun. Pure Appl. Math., 1982, 35(5), 653–696.

[23] W. C. Qian, Y. Li and X. Yang, *Multiscale KAM theorem for Hamiltonian systems*, J. Differ. Equ., 2019, 266(1), 70–86.

[24] M. Rudnev and S. Wiggins, *KAM theory near multiplicity one resonant surfaces in perturbations of A-priori stable Hamiltonian systems*, J. Nonl. Sci., 1997, 7(2), 177–209.

[25] M. B. Sevryuk, *Partial preservation of frequencies in KAM theory*, Nonlinearity, 2006, 19(5), 1099–1140.

[26] M. B. Sevryuk, *Partial preservation of frequencies and Floquet exponents in KAM theory (Rüssian)*, translated from Tr. Mat. Inst. Steklova, 2007, 259(2), 174–202.

[27] C. L. Siegel and J. K. Moser, *Lectures on celestial mechanics*, Springer, Berlin, 1971.

[28] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, New Jersey, 1970.

[29] D. V. Treshchëv, *Mechanism for destroying resonance tori of Hamiltonian systems*, Mat. USSR Sb., 1989, 1439(10), 1325–1346.

[30] J. X. Xu, J. G. You and Q. J. Qiu, *Invariant tori for nearly integrable Hamiltonian systems with degeneracy*, Math. Z., 1997, 226(3), 375–387.

[31] J. X. Xu and J. G. You, *Persistence of the non-twist torus in nearly integrable Hamiltonian systems*, Proc. Amer. Math. Soc., 2010, 138(7), 2385–2395.

[32] L. Xu, Y. Li and Y. Yi, *Lower-dimensional tori in multi-scale, nearly integrable Hamiltonian systems*, Ann. Henri Poincaré, 2017, 18(1), 53–83.

[33] L. Xu, Y. Li and Y. Yi, *Poincaré-Treshchëv mechanism in multi-scale, nearly integrable Hamiltonian systems*, J. Nonl. Sci., 2018, 28(1), 337–369.