COMMUN. MATH. SCI.
Vol. 21, No. 2, pp. 559–579

KAM PERSISTENCE FOR MULTISCALE GENERALIZED HAMILTONIAN SYSTEMS*

WEICHAO QIAN†, XUE YANG‡, AND YONG LI§

Abstract. This paper concerns the persistence of invariant tori for multiscale generalized Hamiltonian systems. A multiscale nondegenerate condition on Poisson manifold comparing Kolmogorov nondegenerate one on symplectic manifold and multiscale iso-energetically nondegenerate condition on Poisson manifold comparing iso-energetically nondegenerate one due to Arnold are introduced, hence some multiscale KAM theorems and multiscale iso-energetic KAM theorems on Poisson manifold are established. And we give three applications by a direct example, first order PDEs and steady Euler fluid path flow, respectively.

Keywords. Multiscale generalized Hamiltonian systems; multiscale nondegenerate condition; multiscale iso-energetically nondegenerate condition; KAM persistence.

AMS subject classifications. 37J40; 70H08.

1. Introduction

Consider a nearly integrable real analytic multiscale generalized Hamiltonian system of the following form:

$$H(x,y)=\varepsilon_0 h_0(y)+\varepsilon_1 h_1(y)+\cdots+\varepsilon_{m_0} h_{m_0}(y)+\varepsilon^2 P(x,y), \tag{1.1}$$

with the Poisson structure matrix $I(y)=\begin{pmatrix} 0 & B \\ -B^T & C \end{pmatrix}$ to be specialized below, where $y\in G$, $G\subset R^l$ is a bounded closed region (closure of a bounded, nonempty open set), $x\in T^n=R^n/2\pi Z^n$; $h_i(y)$, $0\le i\le m_0$, and $P(x,y)$ are real analytic functions; ε_{m_0}, $0\le i\le m_0$, and ε are parameters with $0<\varepsilon<\min\{\varepsilon_0,\varepsilon_1,\cdots,\varepsilon_{m_0}\}\ll 1$; l and n are positive integers and m_0 is nonnegative integer.

When $m_0=0$ and $I=J$ (the standard symplectic matrix), where $l=n$, system (1.1) is just the classical Hamiltonian system. The celebrated KAM theory duo to Kolmogorov [24], Arnold [1] and Moser [33] asserts the persistence of Lagrangian invariant tori, which answers certain stability questions of the planetary systems. And for the persistence of lower dimensional invariant tori, see [12,21–23,29,30,34,38–40,43,45,52], especially, for resonant invariant tori, see [11,14,26,27,41,48]. For a long time, one has been trying to establish the KAM type results for nonsymmetric Hamiltonian systems, i.e. generalized Hamiltonian ones under consideration. When $m_0=0$, $l<n$ and $l+n$ is even, the system is co-isotropic, for which we refer the reader to [10,16,17,35,37,51]. Actually, by some technical reasons the development of KAM theory for 'odd-dimensional' systems is a challenging problem, as pointed out in [28,32,46], and the relative theorem, where $l+n$ is odd, was given in [25] on the Poisson manifolds, which can be applied to

*Received: September 12, 2021; Accepted (in revised form): June 26, 2022. Communicated by Shi Jin.

†College of Mathematics, Jilin University, Changchun 130012, P.R. China (qian_wc@163.com).

‡College of Mathematics, Jilin University, Changchun 130012, P.R. China; School of Mathematics and Statistics & Center for Mathematics and Interdisciplinary Sciences, Northeast Normal University, Changchun 130024, P.R. China (xueyang@jlu.edu.cn).

§Corresponding author. College of Mathematics, Jilin University, Changchun 130012, P.R. China; School of Mathematics and Statistics & Center for Mathematics and Interdisciplinary Sciences, Northeast Normal University, Changchun 130024, P.R. China (liyongmath@163.com).

the perturbation of three-dimensional incompressible fluid flows [7,32,36]. Also see [13]. Furthermore, the KAM theory about atropic tori can be found in [18,19,46,47].

On the other hand, starting with Arnold's research [2] for $m_0=1$, the KAM stability for multiscale Hamiltonian systems has been paid high attention. When $m_0\geq 1$ and $I=J$, $l=n$, with the degeneracy-removing condition in many restricted $3-$body as well as $n-$body, there exists a family of invariant tori [4–6, 8, 20] and with high order degeneracy-removing condition [15]. This has been applied to spatial lunar problem by Meyer, Palacián and Yanguas [31]. And for further research, see [42, 44, 49, 50]. Naturally, one can ask whether there is a family of invariant tori when system (1.1) is one of multiscale with $m_0\geq 1$. Especially, when $m_0>1$, this will become very complex due to possible independence of these small parameters. In this paper we will study the KAM persistence for such a multiscale generalized Hamiltonian system (1.1).

Let the Poisson structure matrix $I=(I_{ij}):G\times T^n\to R^{(l+n)\times(l+n)}$ be a real analytic, antisymmetric, matrix-valued function with $rank\ I>0$ and satisfy the Jacobi identity:

$$\sum_{a=1}^{l+n}\Big(I_{ia}\frac{\partial I_{jk}}{\partial z_a}+I_{ja}\frac{\partial I_{ki}}{\partial z_a}+I_{ka}\frac{\partial I_{ij}}{\partial z_a}\Big)=0 \tag{1.2}$$

for all $z=(y,x)\in G\times T^n$ and $i,j,k=1,2,\cdots,l+n$. Such a structure matrix defines a 2-form ω^2 (Poisson structure): $\omega^2(\cdot,I\omega^1)=\omega^1(\cdot)$, for all 1-form ω^1 defined on $G\times T^n$, which can also be determined in the following way:

$$\{f_1,f_2\}=df_2(Idf_1)=\langle\nabla f_1,I\nabla f_2\rangle=\omega^2(Idf_1,Idf_2),$$

for all smooth functions f_1 and f_2 defined on $G\times T^n$, where $\{\cdot,\cdot\}$ denotes the Poisson bracket and ∇ denotes the standard Euclidean gradient on $R^l\times T^n$. To ensure the invariance 2-form ω^2 relative to T^n, the structure matrix I should be independent of $x\in T^n$, i.e. $I=I(y)$, $y\in G$. Then on the Poisson manifold $(G\times T^n,\omega^2)$ the motion equation of (1.1) associated to the 2-form ω^2 reads

$$\dot z=I(y)\nabla(N(y,\tilde\varepsilon)+\varepsilon^2P(y,x)), \tag{1.3}$$

where $\tilde\varepsilon=(\varepsilon_1,\cdots,\varepsilon_{m_0})$, $N(y,\tilde\varepsilon)=\varepsilon_0h_0(y)+\varepsilon_1h_1(y)+\cdots+\varepsilon_{m_0}h_{m_0}(y)$ and $z=(y,x)^T$. Moreover, we require that the unperturbed system associated to (1.3) is completely integrable, i.e. $y=(y_1,y_2,\cdots,y_l)^T\in G$ need to satisfy the involution conditions: $\{y_i,y_j\}=0,i,j=1,2,\cdots,l$. Thereby,

$$I(y)=\begin{pmatrix} O & B(y)\\ -B^T(y) & C(y)\end{pmatrix}, \tag{1.4}$$

where $O=O_{l,l}$ is a zero matrix, $B=B_{l,n}$, $C=C_{n,n}$ with $C^T=-C$.

Let $\varepsilon=0$ in (1.3). Then the motion equation for the unperturbed system, $N(y,\tilde\varepsilon)$, reads

$$\begin{cases}\dot y=0\\ \dot x=\omega(y),\end{cases}$$

where $\omega=(\bar\omega_1(y),\cdots,\bar\omega_n(y))=-B^T(y)\partial_yN(y,\tilde\varepsilon)$, $\tilde\varepsilon=(\varepsilon_1,\cdots,\varepsilon_{m_0})$. Hence, the phase space $G\times T^n$ is foliated into invariant $n-$tori $\{T_y:y\in G\}$ carrying parallel flows under the incommensurate condition: $\langle k,\omega\rangle\neq 0$ for $\forall k\in Z^n\setminus\{0\}$. What we will show is the persistence of invariant tori under small perturbation. To this aim, let

$g(\varsigma,y)=(\langle\varsigma,\omega_0\rangle,\cdots,\langle\varsigma,\omega_{m_0}\rangle)^T$, $\varsigma\in S^n$, where S^n is the n dimensional unit sphere and $\omega_i=-B^T(y)\partial_y h_i$, $0\le|i|\le m_0$. We introduce the following multiscale nondegenerate condition:

(A) There exists an $N>1$ such that

$$rank\{\partial_y^\alpha g:0\le|\alpha|\le N,\quad \forall y\in G\}=m_0+1,\qquad \forall\varsigma\in S^n.$$

REMARK 1.1. As is well-known, the weakest condition ensuring the persistence of invariant tori is Rüssmann nondegenerate condition [45]. Condition **(A)** is a multiscale nondegenerate one on Poisson manifold as compared to the Rüssmann nondegenerate condition on symplectic manifold. For multiscale nondegenerate condition on symplectic manifold, refer to [42].

REMARK 1.2. In fact, condition (**A**) is also equivalent to the following:

(**A′**) There is a positive integer N such that

$$rank\{\partial_y^\alpha\omega:0\le|\alpha|\le N,\ \forall y\in G\}=n\ for\ |\varepsilon_i|>0,$$

where $\omega=-B^T(y)\partial_y N(y,\tilde\varepsilon)$, $\tilde\varepsilon=(\varepsilon_1,\cdots,\varepsilon_{m_0})$, n is the dimension of the variable x.

(Kol) Assume

$$(\partial_y^2N)^T\partial_y^2N\ge\min\{\varepsilon_0^2,\cdots,\varepsilon_{m_0}^2\}I_{l\times l}.$$

REMARK 1.3. Assumption (**Kol**) is a multiscale nondegenerate condition on Poisson manifold as compared to the Kolmogorov nondegenerate one on symplectic manifold. If $m_0=0$, for a similar nondegenerate condition on Poisson manifold, refer to [25].

(Iso) Assume

$$\begin{pmatrix}\partial_y^2N & \Omega^T\\ \Omega & 0\end{pmatrix}^T\begin{pmatrix}\partial_y^2N & \Omega^T\\ \Omega & 0\end{pmatrix}\ge\min\{\varepsilon_0^2,\cdots,\varepsilon_{m_0}^2\}I_{(l+1)\times(l+1)}.$$

REMARK 1.4. Assumption (**Iso**) is a multiscale isoenergetically nondegenerate condition on Poisson manifold as compared to the isoenergetically nondegenerate one given by Arnold [2] on symplectic manifold. However, Arnold's condition does not involve multiscale. Hence (**Iso**) seems to be first multiscale isoenergetically nondegenerate condition.

Our main result can be stated as follows.

THEOREM 1.1. *Consider Hamiltonian* (1.1) *with the Poisson structure* $I(y)$, *i.e.* (1.4) *and the Jacobi identity.*

(1) *Assume multiscale nondegenerate condition* (**A**). *Then there exist a* $\Delta_0>0$ *and a family of Cantor sets* $G_\varepsilon\subset G$, $0<\varepsilon\le\Delta_0$, *such that for any* $y\in G_\varepsilon$ *the unperturbed torus* T_y *persists and gives rise to an analytic, Diophantine, invariant* $n-$*torus of the perturbed system with small perturbed frequency* $\omega_\varepsilon(y)$. *Moreover, the Lebesgue measure* $|G\setminus G_\varepsilon|\to0$ *as* $\varepsilon\to0$;

(2) *Assume* (**Kol**) *and* (**A**) *on* G. *Then there exist a* $\Delta_0>0$ *and a family of Cantor sets* $G_\varepsilon\subset G$, $0<\varepsilon\le\Delta_0$, *such that for any* $y\in G_\varepsilon$ *the unperturbed Diophantine tori will persist and give rise to perturbed tori preserving corresponding unperturbed toral frequencies.*

(3) *Let $\Sigma=\{y:N(y)=c\}$ be a given energy surface. Assume* (**Iso**) *and* (**A**) *on Σ. Then there exist a $\Delta_0>0$ and a family of Cantor sets $\Sigma_\varepsilon\subset\Sigma$, $0<\varepsilon\leq\Delta_0$, such that for any $y\in\Sigma_\varepsilon$ the unperturbed Diophantine tori will persist and give rise to perturbed tori keeping the same energy and maintaining the frequency ratio. Moreover, $|\Sigma\setminus\Sigma_\varepsilon|\to 0$ as $\varepsilon\to 0$.*

To have a good understanding of the main result, we show three applications first.

Example 1.1. Consider the following Hamiltonian system:

$$H(x,y,\tilde{\varepsilon})=\langle\omega,y\rangle+\frac{1}{2}\langle y,Ay\rangle+\varepsilon^2P(x,y), \tag{1.5}$$

defined on a Poisson manifold with Poisson structure $I(y)$ satisfying (1.4) and the Jacobi identity, where

$$\omega=\varepsilon_0\omega_0+\cdots+\varepsilon_{m_0}\omega_{m_0},\ A=\begin{pmatrix}\varepsilon_0 I_0 & 0 & \cdots & 0\\ 0 & \varepsilon_1 I_1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \varepsilon_{m_0}I_{m_0}\end{pmatrix},$$

$I_0,\ I_1,\ \cdots,\ I_{m_0}$ are identity matrices with dimension $\tilde{m}_0,\ \tilde{m}_1,\ \cdots,\ \tilde{m}_{m_0}$, respectively, and $\tilde{m}_0+\tilde{m}_1+\cdots+\tilde{m}_{m_0}=n$. It is easy to check

$$\begin{aligned}A^TA&=\begin{pmatrix}\varepsilon_0^2 I_0 & 0 & \cdots & 0\\ 0 & \varepsilon_1^2 I_1 & \cdots & 0\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \cdots & \varepsilon_{m_0}^2 I_{m_0}\end{pmatrix}\\ &\geq\min\{\varepsilon_0^2,\cdots,\varepsilon_{m_0}^2\}\begin{pmatrix}I_0 & 0 & \cdots & 0\\ 0 & I_1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & I_{m_0}\end{pmatrix}.\end{aligned}$$

Using Theorem 1.1, if ω satisfies condition (**A**), for Hamiltonian (1.5) there is a family of invariant tori, on which the frequency is $B^T\omega$.

Example 1.2. Consider the following first order PDEs:

$$\sum_{i=1}^{n}a_i(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l)\frac{\partial u}{\partial I_i}+\sum_{i=1}^{l}a_{n+i}(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l)\frac{\partial u}{\partial\theta_i}=0, \tag{1.6}$$

where $I=(I_1,\cdots,I_n)\in G\subset R^n$, $\theta=(\theta_1,\cdots,\theta_l)\in T^l$, G is a bounded closed region.

We give some basic definitions first. In (1.6), $(a_1,\cdots,a_{n+l})$ is called the **characteristic direction**. The following equations:

$$\begin{cases}\frac{dI_1}{dt}=a_1(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l),\\ \qquad\vdots\\ \frac{dI_n}{dt}=a_n(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l),\\ \qquad\vdots\\ \frac{d\theta_1}{dt}=a_{n+1}(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l),\\ \qquad\vdots\\ \frac{d\theta_l}{dt}=a_{n+l}(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l),\end{cases} \tag{1.7}$$

and

$$\begin{cases} \frac{dI_1}{dt} = a_1(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l), \\ \qquad\vdots \\ \frac{dI_n}{dt} = a_n(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l), \\ \qquad\vdots \\ \frac{d\theta_1}{dt} = a_{n+1}(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l), \\ \qquad\vdots \\ \frac{d\theta_l}{dt} = a_{n+l}(I_1,\cdots,I_n,\theta_1,\cdots,\theta_l), \\ \frac{du}{dt} = 0, \end{cases} \tag{1.8}$$

are called **characteristic equations and full characteristic equations, respectively.**

As is well-known, the picture of integral curves for (1.6) is determined by full characteristic equations. From full characteristic equations (1.8), we get $u=c$. Therefore, the characteristic equations are basic to show the picture of integral curves. Next, we give the definitions of integrable characteristic equations and nearly integrable characteristic equations. The characteristic equations are **integrable** if there is a Hamiltonian $H(I)$ and a matrix $B_{n\times l}(I)$ such that

(1) $a_i=0$, $1\le i\le n$,

(2) $\begin{pmatrix} a_{n+1} \\ \vdots \\ a_{n+l} \end{pmatrix} = -B^T\frac{\partial H}{\partial I}$.

For integrable characteristic equations, if $B^T\frac{\partial H}{\partial I}$ is incommensurate, the picture of integral curves on space (I,θ,u) is a torus for any $I\in G$ and $u=c$.

Question: What is the picture of the integral curves if the characteristic equations are not integrable?

The characteristic equations are **nearly integrable** if there are two Hamiltonians $H(I)$ and $P(I,\theta)$, two matrices $B(I)$ and $C(I)$ with $C=-C^T$ such that

(1) $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \varepsilon B\frac{\partial P}{\partial \theta}$,

(2) $\begin{pmatrix} a_{n+1} \\ \vdots \\ a_{n+l} \end{pmatrix} = -B^T\frac{\partial H}{\partial I} + \varepsilon C\frac{\partial P}{\partial \theta}$.

REMARK 1.5. Combining the definitions of integrable characteristic equations and nearly integrable characteristic equations, $u=c$ is a surface on space (I,θ,u) with a differential structure $J(I)$, usually to be a Poisson one, where $J(I)=\begin{pmatrix} 0 & B \\ -B^T & C \end{pmatrix}$, $C=-C^T$.

(A2) There is an $N>1$ such that

$$rank\{\partial_I^\alpha g, 0\le|\alpha|\le N\}=1, \forall\varsigma\in S^l,$$

where $g=\varsigma B^T\frac{\partial H}{\partial I}$.

Theorem 1.2. *Let J be a Poisson structure, i.e., J satisfy the Jacobi identity. The picture of integral curves with integrable characteristic equations is a torus T^l and denote it by $\mathcal{T}$.*

(1) *Assume* (**A2**) *on G. Then there exist a $\varepsilon_0>0$ and a family of Cantor sets $G_\varepsilon\subset G$, $0<\varepsilon\le\varepsilon_0$, such that for any $I\in G_\varepsilon$ and $u=c$ the picture of integral curve with nearly integrable characteristic equations on space (I,θ,u) keeps a torus $\mathcal{T}_\varepsilon$. Moreover, the Lebesgue measure $|G\setminus G_\varepsilon|\to 0$ as $\varepsilon\to 0$.*

(2) *Assume* (**A2**) *and rank $\frac{\partial^2 H}{\partial I^2}=n$ on G. Then there exist a $\varepsilon_0>0$ and a family of Cantor sets $G_\varepsilon\subset G$, $0<\varepsilon\le\varepsilon_0$, such that for any $I\in G_\varepsilon$ and $u=c$ the picture of integral curve with nearly integrable characteristic equations on space (I,θ,u) keeps a torus $\mathcal{T}_\varepsilon$ and the frequencies between $\mathcal{T}$ and $\mathcal{T}_\varepsilon$ are same. Moreover, the Lebesgue measure $|G\setminus G_\varepsilon|\to 0$ as $\varepsilon\to 0$.*

(3) *Let $\Sigma=\{I:H=c\}$ be a given energy surface. Assume* (**A2**) *and*

$$rank\begin{pmatrix} \frac{\partial^2 H}{\partial I^2} & \Omega^T \\ \Omega & 0 \end{pmatrix}=n+1$$

hold on Σ. Then there exist a $\varepsilon_0>0$ and a family of Cantor sets $\Sigma_\varepsilon\subset\Sigma$, $0<\varepsilon\le\varepsilon_0$, such that for any $I\in\Sigma_\varepsilon$ and $u=c$ the picture of integral curves with nearly integrable characteristic equations on space (I,θ,u) persists as a torus $\mathcal{T}_\varepsilon$, on which the frequency ratio is kept. Moreover, the Lebesgue measure $|\Sigma\setminus\Sigma_\varepsilon|\to 0$ as $\varepsilon\to 0$.

Remark 1.6. In this application we give the definitions of integrable characteristic equations and nearly integrable characteristic equations. And we prove that the picture of integral curves for semi-linear equation with nearly integrable characteristic equations is a family of tori. The results also hold for semi-linear equation with nearly multiscale integrable characteristic equations.

Example 1.3. For multiscale generalized Hamiltonians, an important and direct application is about the three-dimensional multiscale steady Euler fluid path flows. The persistence of invariant 2-tori or 1-tori (on the cylinder) after suitable perturbations is a significant way to understand the barrier of fluid transport and mixing, which brings KAM theory into play.

A fundamental result about three-dimensional volume-preserving flows given by Arnold [3] shows that the system admits either invariant tori with trajectories all closed or all dense, or invariant annuli with trajectories all closed, when the steady Euler velocity field is not everywhere collinear with its vorticity field in a domain, which uses crucially the fact that the vorticity associated with a steady Euler flow is an infinitesimal generator of a volume-preserving spatial symmetry group. Without the fact mentioned above, the persistence of invariant 2-tori under volume-preserving perturbations was shown in [32] by using the KAM theory developed in [9] for volume-preserving maps. And for the persistence to general perturbation, we refer the reader to [25]. The frequencies of the results mentioned above are only with one scale. Then a nature question is whether there is a family of invariant tori for multiscale three-dimensional steady Euler fluid path flows.

For example, consider the nearly planar flow reduced by a divergence-free multiscale system of ordinary differential equations of the following form:

$$\begin{cases} \dot{z_1} = \frac{\partial H(z_1,z_2)}{\partial z_2} \\ \dot{z_2} = -\frac{\partial H(z_1,z_2)}{\partial z_1} \\ \dot{z_3} = \varepsilon_1 h(z_1,z_2) \end{cases}, \tag{1.9}$$

where z_1, z_2 *and* $z_3 \in R^1$. We assume that the steady Euler flow admits a family of elliptic vortex lines, i.e.

(H) There is a region $\mathcal{D}$ of the $(z_1,z_2)-$plane in which the level sets $H(z_1,z_2)=c$ are closed curves,

which is generally satisfied for steady Euler flow.

Under assumption (**H**), (1.9) becomes the following form of action-angle:

$$\begin{cases} \dot{\mathcal{I}} = 0 \\ \dot{\theta} = \omega_1(\mathcal{I}) \\ \dot{z_3} = \varepsilon_1 h(\mathcal{I},\theta) \end{cases}. \tag{1.10}$$

Suppose $\omega_1 \neq 0$ in $\mathcal{D}$. Then with the volume-preserving transformation $\phi = z_3 + \frac{\theta}{2\pi}\int_0^{2\pi}\frac{\varepsilon_1 h(\mathcal{I},\theta)}{\omega_1(\mathcal{I})}d\theta - \int \frac{\varepsilon_1 h(\mathcal{I},\theta)}{\omega_1(\mathcal{I})}d\theta$, (1.10) arrives at

$$\begin{cases} \dot{\mathcal{I}} = 0 \\ \dot{\theta} = \omega_1(\mathcal{I}) \\ \dot{\phi} = \varepsilon_1 \omega_2(\mathcal{I}) \end{cases}, \tag{1.11}$$

where $\phi \in S^1$ or R^1 and $\omega_2(\mathcal{I}) = \frac{\omega_1(\mathcal{I})}{2\pi}\int_0^{2\pi}\frac{h(\mathcal{I},\theta)}{\omega_1(\mathcal{I})}d\theta$, which describes nearly planar flow.

Actually, system (1.11) is equivalent to an integral Hamiltonian $N(\mathcal{I})$ with the Poisson structure matrix

$$I = \begin{pmatrix} 0 & B(\mathcal{I}) \\ -B^T(\mathcal{I}) & C \end{pmatrix},$$

where $\mathcal{I} \in R^1$, $\begin{pmatrix} \omega_1 \\ \varepsilon_1\omega_2 \end{pmatrix} = -B^T(\mathcal{I})\partial_{\mathcal{I}}N(\mathcal{I})$, $\varepsilon \ll \varepsilon_1$. With the incommensurate condition the persistence of invariant tori is obvious. Then a direct and important problem is whether there is a family of invariant tori under small perturbation. In other words, consider a nearly integrable real analytic multiscale three-dimensional Hamiltonian of the following form:

$$H(\mathcal{I},\psi) = N(\mathcal{I}) + \varepsilon^2 P(\mathcal{I},\psi) \tag{1.12}$$

with the Poisson structure matrix

$$I = \begin{pmatrix} 0 & B(\mathcal{I}) \\ -B^T(\mathcal{I}) & C \end{pmatrix},$$

where $\mathcal{I} \in \Lambda \subset R^1$, $\psi = (\theta,\phi) \in T^2$, $\begin{pmatrix} \omega_1 \\ \varepsilon_1\omega_2 \end{pmatrix} = -B^T(\mathcal{I})\partial_{\mathcal{I}}N(\mathcal{I})$, $\varepsilon \ll \varepsilon_1$. We state our result about the persistence of invariant tori for Hamiltonian (1.12) as follows.

(A3) There is a positive integer N such that

$$rank\{\partial_{\mathcal{I}}^{\alpha}\omega : 0 \le |\alpha| \le N\} = 2 \ for \ |\varepsilon_i| > 0,$$

where $\omega = -B^T(\mathcal{I})\partial_{\mathcal{I}}N(\mathcal{I})$.

THEOREM 1.3. *Consider Hamiltonian* (1.12) *with the Poisson structure* $I(\mathcal{I})$ *satisfying the Jacobi identity.*

(1) *Assume* (**A3**) *on* Λ. *Then there exist a* $\Delta_0>0$ *and a family of Cantor sets* $\Lambda_\varepsilon\subset\Lambda$, $0<\varepsilon\leq\Delta_0$, *such that for any* $\mathcal{I}\in\Lambda_\varepsilon$ *the unperturbed torus* $T_{\mathcal{I}}$ *persists and gives rise to an analytic, Diophantine, invariant torus of the perturbed system with small perturbed frequency* $\omega_\varepsilon(\mathcal{I})$. *Moreover, the Lebesgue measure* $|\Lambda\setminus\Lambda_\varepsilon|\to 0$ *as* $\varepsilon\to 0$;

(2) *Assume* (**A3**) *and* $\partial^2_{\mathcal{I}}N^T\partial^2_{\mathcal{I}}N\geq\varepsilon_1^2$ *on* Λ. *Then there exist a* $\Delta_0>0$ *and a family of Cantor sets* $\Lambda_\varepsilon\subset\Lambda$, $0<\varepsilon\leq\Delta_0$, *such that for any* $\mathcal{I}\in\Lambda_\varepsilon$ *the unperturbed Diophantine tori will persist and give rise to perturbed tori which preserve corresponding unperturbed toral frequency. Moreover, the Lebesgue measure* $|\Lambda\setminus\Lambda_\varepsilon|\to 0$ *as* $\varepsilon\to 0$;

(3) *Let* $\Sigma=\{\mathcal{I}:N=c\}$ *be a given energy surface. Assume* (**A3**) *and*

$$\begin{pmatrix}\partial^2_{\mathcal{I}}N & \Omega^T\\ \Omega & 0\end{pmatrix}^T\begin{pmatrix}\partial^2_{\mathcal{I}}N & \Omega^T\\ \Omega & 0\end{pmatrix}\geq\varepsilon_1^2 I_{2\times 2}$$

on Σ. *Then there exist a* $\Delta_0>0$ *and a family of Cantor sets* $\Sigma_\varepsilon\subset\Sigma$, $0<\varepsilon\leq\Delta_0$, *such that for any* $\mathcal{I}\in\Sigma_\varepsilon$ *the unperturbed Diophantine tori on* Σ_ε *will persist and give rise to perturbed tori keeping the same energy and maintaining the frequency ratio. Moreover,* $|\Sigma\setminus\Sigma_\varepsilon|\to 0$ *as* $\varepsilon\to 0$.

REMARK 1.7. In applications, one does have the freedom to determine the Poisson matrix $I(\mathcal{I})$ according to either the form of a perturbation or the nature of a particular problem.

Actually, this paper is an extension of reference [25]. The extension contains not only the scales in integrable part from 1 to m_0 but also the results show the preservation of frequency and frequency ratio. Of course, this is not simple. It seems to be the first work to give multiscale nondegenerate condition on Poisson manifold as compared to the Kolmogorov nondegenerate one on symplectic manifold and multiscale iso-energetically nondegenerate condition on Poisson manifold as compared to the iso-energetically nondegenerate one on symplectic manifold. And we use an amendatory KAM iteration mentioned in [44] to reduce the harm caused by multiscale. The paper is organized as follows. In Section 2, we provide an amendatory KAM iteration for a parameterized multiscale generalized Hamiltonian. The proof of the main result is placed in Section 3.

2. Abstract Hamiltonian

Throughout the paper, unless specified otherwise, 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 denote by $|\cdot|_D$ the supremum norm of functions on a domain D. Also, for any two complex column vectors ξ,ζ of the same dimension, $\langle\xi,\zeta\rangle$ always means $\xi^T\zeta$, i.e. the transpose of ξ times ζ. For the sake of brevity, we shall not specify smoothness orders for functions having obvious orders of smoothness indicated by taking their derivatives.

Consider a parameterized Hamiltonian system of the following form:

$$\begin{aligned}\mathcal{H} &= \mathcal{N}(y,\xi,\tilde{\varepsilon})+\varepsilon\mathcal{P}(x,y,\xi),\\ \mathcal{N} &= e(\xi)+\langle\Omega(\xi),y\rangle+h(y,\xi),\\ \Omega &= \varepsilon_0\omega_0+\varepsilon_1\omega_1+\cdots+\varepsilon_{m_0}\omega_{m_0},\end{aligned}\tag{2.1}$$

$$h = \langle y, A(\xi)y\rangle,$$
$$A = \varepsilon_0 A_0 + \varepsilon_1 A_1 + \cdots + \varepsilon_{m_0} A_{m_0},$$

defined on $D(r,s)=\{(x,y):|Im\ x|<r,\ |y|<s\}$, a $(r,s)-$complex neighborhood of $T^n\times\{0\}\subset T^n\times R^l$, where $\mathcal{P}=\varepsilon P(x,y,\xi)$, $\xi\in\Lambda=\{\lambda:|\lambda|\le\delta_1\}\subset R^d$, $\tilde{\varepsilon}=(\varepsilon_1,\cdots,\varepsilon_{m_0})$ and ε defined as above. Denote $\bar{\Lambda}=\{\lambda:|\lambda|\le\delta_1-\bar{\eta}\}$.

(A1) There is an $N>1$ such that

$$rank\{\partial_\xi^\alpha\tilde{g}:0\le|\alpha|\le N\}=m_0+1,\forall\varsigma\in S^n,$$

where $\bar{g}=(\langle\varsigma,\bar{\omega}_0\rangle,\cdots,\langle\varsigma,\bar{\omega}_{m_0}\rangle)$, $\varsigma\in S^n$, $\bar{\omega}_i=-B^T\omega_i$, $0\le|i|\le m_0$.

THEOREM 2.1. *Consider Hamiltonian* (2.1) *with the Poisson structure* $I(y)$, *i.e.* (1.4) *and the Jacobi identity.*

(1) *Assume* (**A1**). *Then there exist a* $\Delta_0>0$ *and a family of Cantor sets* $\Lambda_\varepsilon\subset\Lambda$, $0<\varepsilon\le\Delta_0$, *such that for any* $y\in\Lambda_\varepsilon$ *the unperturbed torus* T_y *persists and gives rise to an analytic, Diophantine, invariant* $n-$*torus of the perturbed system with small perturbed frequency* $\omega_\varepsilon(y)$. *Moreover, the Lebesgue measure* $|\Lambda\setminus\Lambda_\varepsilon|\to 0$ *as* $\varepsilon\to 0$;

(2) *Assume* (**A1**) *and* $A^TA\ge\min\{\varepsilon_1^2,\cdots,\varepsilon_{m_0}^2\}I_{l\times l}$ *on* Λ. *Then there exist a* $\Delta_0>0$ *and a family of Cantor sets* $\Lambda_\varepsilon\subset\Lambda$, $0<\varepsilon\le\Delta_0$, *such that for any* $y\in\Lambda_\varepsilon$ *the unperturbed Diophantine tori will persist and give rise to perturbed tori which preserve corresponding unperturbed toral frequency. Moreover, the Lebesgue measure* $|\Lambda\setminus\Lambda_\varepsilon|\to 0$ *as* $\varepsilon\to 0$;

(3) *Let* $\Sigma=\{y:N=c\}$ *be a given energy surface. Assume* (**A1**) *and*

$$\begin{pmatrix} A & \Omega^T \\ \Omega & 0\end{pmatrix}^T\begin{pmatrix} A & \Omega^T \\ \Omega & 0\end{pmatrix}\ge\min\{\varepsilon_1^2,\cdots,\varepsilon_{m_0}^2\}I_{(l+1)\times(l+1)}$$

on Σ. *Then there exist a* $\Delta_0>0$ *and a family of Cantor sets* $\Sigma_\varepsilon\subset\Sigma$, $0<\varepsilon\le\Delta_0$, *such that for any* $y\in\Sigma_\varepsilon$ *the unperturbed Diophantine tori on* Σ_ε *will persist and give rise to perturbed tori keeping the same energy and maintaining the frequency ratio. Moreover,* $|\Sigma\setminus\Sigma_\varepsilon|\to 0$ *as* $\varepsilon\to 0$.

2.1. KAM steps. Let us begin with system (2.1) by regarding it as a Hamiltonian of $0-$step, and rewriting it as follows:

$$\begin{aligned}\mathcal{H}_0 &= \mathcal{N}_0(y,\xi,\tilde{\varepsilon})+\varepsilon\mathcal{P}_0(x,y,\xi),\\ \mathcal{N}_0 &= e_0(\xi)+\langle\Omega_0(\xi),y\rangle+h_0(y,\xi),\\ \Omega_0 &= \varepsilon_0\omega_0^0+\varepsilon_1\omega_1^0+\cdots+\varepsilon_{m_0}\omega_{m_0}^0,\\ h_0 &= \langle y, A_0(\xi)y\rangle,\\ A_0 &= \varepsilon_0A_0^0+\varepsilon_1A_1^0+\cdots+\varepsilon_{m_0}A_{m_0}^0,\end{aligned}\tag{2.2}$$

defined on $D(r_0,s_0)=\{(x,y):|Im\ x|<r_0,\ |y|<s_0\}$, a $(r_0,s_0)-$complex neighborhood of $T^n\times\{0\}\subset T^n\times R^l$, where $\mathcal{P}_0=\varepsilon P(x,y,\xi)$, $\xi\in\Lambda_0\subset R^d$. Moreover, let $\gamma_0=\varepsilon^{\frac{1}{12(7+N)}}$, $s_0=\varepsilon^{\frac{1}{3}}$, $\mu_0=\varepsilon^{\frac{1}{4}}$, $\bar{\eta}_0=\varepsilon^{\frac{3(1-b)-4(b+\sigma)}{14N}}$, where b and σ are constants to be determined next. Then by Cauchy estimate we have

$$|\partial_\xi^q\mathcal{P}_0|_{D(r_0,s_0)\times\bar{\Lambda}_0}<\frac{\gamma_0^{N+7}s_0^2\mu_0}{\bar{\eta}_0^N},\qquad |q|\le N.$$

Next, we will show the KAM iteration from $\nu-$step to $(\nu+1)-$step. For simplicity, 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. Suppose, at $\nu-$th step, we have obtained the following smooth family of real analytic Hamiltonians

$$\begin{aligned}\mathcal{H}(x,y,\xi) &= \mathcal{N}(y,\xi,\tilde{\varepsilon})+\varepsilon\mathcal{P}(x,y,\xi),\\ \mathcal{N}(y,\xi,\tilde{\varepsilon}) &= e(\xi,\tilde{\varepsilon})+\langle\Omega(\xi,\tilde{\varepsilon}),y\rangle+h(y,\xi,\tilde{\varepsilon}),\\ \Omega(\xi,\tilde{\varepsilon}) &= \varepsilon_0\omega_0+\varepsilon_1\omega_1+\cdots+\varepsilon_{m_0}\omega_{m_0},\\ h(y,\xi,\tilde{\varepsilon}) &= \langle y,A(\xi,\tilde{\varepsilon})y\rangle,\\ A(\xi,\tilde{\varepsilon}) &= \varepsilon_0A_0+\varepsilon_1A_1+\cdots+\varepsilon_{m_0}A_{m_0},\end{aligned} \tag{2.3}$$

where $(x,y)\in D(r,s)=\{(x,y):|Im\ x|<r,\ |y|<s\}$, a $(r,s)-$complex neighborhood of $T^n\times\{0\}\subset T^n\times R^l$, $\xi\in\Lambda\subset R^d$. Moreover,

$$|\partial_\xi^q\mathcal{P}|_{D(r,s)\times\bar{\Lambda}}\le\frac{\gamma^{N+7}s^2\mu}{\bar{\eta}^N},\qquad |q|\le N. \tag{2.4}$$

We need to construct a canonical transformation Φ_+, which, on a small phase domain $D(r_+,s_+)$ and a smaller parameter domain Λ_+, transforms (2.3) into a family of Hamiltonians with the following form

$$\mathcal{H}_+=\mathcal{H}\circ\Phi_+=\mathcal{N}_++\varepsilon\mathcal{P}_+$$

enjoying the similar properties to (2.3) but with a much smaller unintegrable perturbation $\mathcal{P}_+$.

All constants below, for simplicity, denoted by c, are positive and independent of the iteration process. Define $\tilde{\varepsilon}=(\varepsilon_0,\cdots,\varepsilon_{m_0})$ with $|\tilde{\varepsilon}|=\sum_{i=0}^{m_0}|\varepsilon_i|$, and let

$$\begin{aligned}r_+ &= \delta r-d(1-\frac{\delta^2}{2})r_0,\qquad s_+=s^{1+b+\sigma},\qquad \gamma_+=\frac{\gamma_0}{4}+\frac{\gamma}{2},\\ K_+ &= ([\log\frac{1}{s}]+1)^3,\quad D_+=D(s_+,r_+),\quad \tilde{D}=D(s_0,r_++\frac{5}{8}(r-r_+)),\\ D_i &= D(is_+,r_++\frac{i-1}{8}(r-r_+)),\ i=1,\cdots,8,\quad \bar{\eta}_+=\bar{\eta}-\frac{\bar{\eta}_0}{2^{\nu+1}},\end{aligned}$$

where a_0, b, σ, d are chosen so that $1<b\ll\sigma\ll1$, $0<d\ll1$, $2-m(b+\sigma)-\sigma>\frac{3}{2}$, $\delta(1+b+\sigma)>1$ and $\delta=1-d$. Hereafter, we let $\tau>\max\{0,n(n+1)-1,l(l+1)-1,(N+1)N-1\}$ be fixed.

2.1.1. Truncation. Consider the Taylor-Fourier series of $\mathcal{P}$

$$\mathcal{P}=\sum_{|k|\in Z^n,\ i\in Z_+^n}P_{ki}y^ie^{\sqrt{-1}\langle k,x\rangle},$$

and denote the truncation of $\mathcal{P}$ by $\mathcal{R}$ with the following form

$$\mathcal{R}=\sum_{|k|\le K_+,\ |i|\le2}P_{ki}y^ie^{\sqrt{-1}\langle k,x\rangle}.$$

With the following assumptions

$$s_+\le\frac{s}{16}, \tag{2.5}$$

$$\int_{K_+}^{\infty} \lambda^{n+N} e^{-\frac{\lambda(r-r_+)}{16}} d\lambda \le s^{3(1+b+\sigma)}, \tag{2.6}$$

and using *Lemma* 2.1 in [25] we get

$$|\partial_\xi^q (\mathcal{P}-\mathcal{R})|_{D_8 \times \bar{\Lambda}} \le \frac{c\gamma^{N+7}\mu(s^{3(1+b+\sigma)} + \frac{s_+^3}{s})}{\bar{\eta}^N}.$$

2.1.2. Homological equations. To average out all harmonic terms of $\mathcal{R}$, i.e. all terms $P_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle}$, $0<|k|\le K_+$, $|i|\le 2$, consider the following homological equations

$$\{\mathcal{N}, F\} + \varepsilon(\mathcal{R}-[\mathcal{R}]) - Q = 0, \tag{2.7}$$

where

$$\begin{aligned} F &= \sum_{0<|k|\le K_+,\ |i|\le 2} f_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle}, \\ [\mathcal{R}] &= \frac{1}{(2\pi)^n}\int_{T^n} \mathcal{R}(x,y)dx, \\ Q &= \sum_{\substack{0<|k|\le K_+,\\ |i|\le 2}} \sqrt{-1}\langle k, (B^T(y,\xi) - B^T(y_0,\xi))(\Omega(\xi)+\partial_y h(y,\xi,\tilde{\varepsilon}))\rangle f_{ki} y^i e^{\sqrt{-1}\langle k,x\rangle}. \end{aligned}$$

By comparing coefficients of (2.7), we set formally

$$\sqrt{-1}\langle k, B^T(y_0,\xi)(\Omega(\xi,\tilde{\varepsilon}) + \partial_y h(y,\xi,\tilde{\varepsilon}))\rangle f_{ki} = \varepsilon P_{ki}. \tag{2.8}$$

Put

$$B^T(\Omega + \partial_y h) = \varepsilon_0 \tilde{\omega}_0 + \varepsilon_1 \tilde{\omega}_1 + \cdots + \varepsilon_{m_0}\tilde{\omega}_{m_0} + O(\varepsilon_0 y + \cdots + \varepsilon_{m_0} y)$$

and denote

$$L_k = \langle k, B^T(y_0,\xi)(\Omega(\xi,\tilde{\varepsilon}) + \partial_y h(y,\xi,\tilde{\varepsilon}))\rangle.$$

Furthermore, on Λ_+, by the following assumption

$$s \cdot K_+^{\tau+1} = o(\gamma_0), \tag{2.9}$$

we have

$$\begin{aligned} |L_k| &= |(\varepsilon_0, \cdots, \varepsilon_{m_0})(\langle k, \tilde{\omega}_0\rangle, \cdots, \langle k, \tilde{\omega}_{m_0}\rangle) \\ &\qquad + (\varepsilon_0, \cdots, \varepsilon_{m_0})(\langle k, O(|y|)\rangle, \cdots, \langle k, O(|y|)\rangle)| \\ &\ge c|\tilde{\varepsilon}|\frac{\gamma}{|k|^\tau} - |\tilde{\varepsilon}||k|O(|y|) \\ &\ge c|\tilde{\varepsilon}|\frac{\gamma_0}{|k|^\tau}, \end{aligned}$$

where $\Lambda_+ = \{\xi \in \Lambda : |\langle k, \omega\rangle| > \frac{|\tilde{\varepsilon}|\gamma}{|k|^\tau},\ 0<|k|\le K_+\}$, $\omega = \varepsilon_0\tilde{\omega}_0 + \cdots + \varepsilon_{m_0}\tilde{\omega}_{m_0}$. Then (2.8) is solvable on Λ_+. Moreover, all solutions f_{ki}, $0<|k|\le K_+$, $|i|\le m$, are real analytic on

Λ_+. Thus we have found out the desired generalized Hamiltonian F, which is real analytic in both $\xi \in \Lambda_+$ and $(y,x) \in D_8$. Moreover,

$$|\partial_\xi^q \partial_y^j \partial_x^z F| \leq \frac{s^{2-|i|}\mu}{\bar{\eta}^N} \Gamma(r-r_+),$$

where $\Gamma(r-r_+) = \sum_{|j|\leq 2,\ 0<|k|\leq K_+} |k|^{\tau(|j|+|q|)+|j|+|q|-1+|z|} e^{-\frac{|k|(r-r_+)}{8}}$. We place the details in Appendix A.

2.1.3. Estimate for transformations. Let $\Phi_+ = \phi_F^1$ be the time-1 map of the motion equation associated to F, i.e.

$$\dot{z} = I(y,\xi)\nabla F(y,x,\tilde{\varepsilon}), \tag{2.10}$$

where $z = (y,x)^T$. Then Φ_+ is a canonical transformation and

$$\bar{\mathcal{H}}_+ = \mathcal{H}\circ\Phi_+ = (\mathcal{N}+\varepsilon\mathcal{R})\circ\phi_F^1 + \varepsilon(\mathcal{P}-\mathcal{R})\circ\phi_F^1 = \bar{\mathcal{N}}_+ + \bar{\mathcal{P}}_+,$$

where

$$\begin{aligned} \bar{\mathcal{P}}_+ &= \int_0^1 \{\mathcal{R}_t, F\}\circ\phi_F^t dt + \varepsilon(\mathcal{P}-\mathcal{R})\circ\phi_F^1 + Q, \\ \bar{\mathcal{N}}_+ &= \mathcal{N} + \varepsilon[\mathcal{R}], \\ \mathcal{R}_t &= (1-t)\{\mathcal{N}, F\} + \varepsilon\mathcal{R}. \end{aligned}$$

It should be pointed out that due to the Jacobi identity the structure matrix I is kept unchanged at each KAM step. Let $I(z)$, $z = (x,y)$, be a structure matrix on $G\times T^n$ and $\phi_F^t(z)$ be the flow generated by a vector field $I(z)\nabla F(z)$. Then by Jacobi identity,

$$\partial_z\phi_F^t(z)^T I(z)\partial_z\phi_F^t(z) = I(\phi_F^t(z)),$$

which implies the preservation of the Poisson structure on $G\times T^n$ under the transformation $z_1 = \phi_F^1(z)$. And with the following assumptions

$$\frac{c\mu\Gamma(r-r_+)}{\bar{\eta}^N} < \frac{1}{8}(r-r_+), \tag{2.11}$$

$$\frac{cs^2\mu\Gamma(r-r_+)}{\bar{\eta}^N} < 3s_+, \tag{2.12}$$

we have

$$|D\phi_F^t - Did| \leq \frac{c\mu\Gamma(r-r_+)}{\bar{\eta}^N}.$$

And we place the details in Appendix B.

Consider the transformation

$$\phi : x \to x,\ y \to y + y^*.$$

Then

$$\mathcal{H}_+ = \bar{\mathcal{H}}_+\circ\phi = e_+ + \langle\Omega_+, y\rangle + h_+(y) + \mathcal{P}_+,$$

where

$$\begin{aligned}
e_+ &= e+\langle\Omega(\xi),y^*\rangle+\frac{1}{2}\langle y^*,Ay^*\rangle+[\mathcal{R}](y^*),\\
\Omega_+ &= \Omega+Ay^*+\varepsilon P_{01},\\
A_+ &= A+\varepsilon\partial_y^2[\mathcal{R}](y^*),\\
h_+ &= \langle y,A_+y\rangle,\\
\mathcal{P}_+ &= \bar{\mathcal{P}}_+\circ\phi.
\end{aligned}$$

Moreover, by induction, we have $|\partial_\xi^q y^*|\le c\frac{\gamma^{N+7}s\mu}{\bar\eta^N}$. We place the details in Appendix C.

REMARK 2.1. Let

$$\begin{cases}\langle\Omega(\xi),y^*\rangle+\frac{1}{2}\langle y^*,Ay^*\rangle+[\mathcal{R}](y^*)=0,\\ Ay^*+\varepsilon P_{01}+t\Omega(\xi)=0.\end{cases}\tag{2.13}$$

Assume

$$\begin{pmatrix}A & \Omega^T\\ \Omega & 0\end{pmatrix}^T\begin{pmatrix}A & \Omega^T\\ \Omega & 0\end{pmatrix}\ge\min\{\varepsilon_0^2,\cdots,\varepsilon_{m_0}^2\}I_{(l+1)\times(l+1)}.$$

Then there is a solution $(y^*,t)^T$ for (2.13), which implies the preservation of frequency ratio on a given energy surface.

2.1.4. Estimate for new Hamiltonian. By the estimate of $|\partial_\xi^l y^*|$ together with definitions of e_+, Ω_+ and A_+, we have

$$\begin{aligned}
|\partial_\xi^q(e_+-e)| &\le \frac{c\gamma^{N+7}s\mu}{\bar\eta^N},\\
|\partial_\xi^q(\Omega_+-\Omega)| &\le \frac{c\gamma^{N+7}s\mu}{\bar\eta^N},\\
|\partial_\xi^q(A_+-A)| &\le \frac{c\gamma^{N+7}\mu}{\bar\eta^N}.
\end{aligned}$$

LEMMA 2.1. *Assume*

$$\Delta_+<\frac{\gamma_+^{N+7}s_+^2\mu_+}{\bar\eta_+^N},\tag{2.14}$$

where

$$\Delta_+=\frac{\mu(c\mu s^4\Gamma^2(r-r_+)+\gamma^{N+7}\mu s_+^3(1+\frac{1}{s})+s^3\mu\Gamma(r-r_+)+\gamma^{N+7}s^4)}{\bar\eta^N}.$$

Then there is a constant c such that

$$|\partial_\xi^q P_+|_{D_+\times\bar\Lambda_+}\le c\frac{\gamma_+^{N+7}s_+^2\mu_+}{\bar\eta_+^N}.$$

Proof. Using the definition of Q, by Cauchy estimate, we deduce

$$|\partial_\xi^q Q|_{D_+\times\bar\Lambda_+}\le\frac{cs^3\mu\Gamma(r-r_+)}{\bar\eta^N}.$$

With the definition of Poisson bracket and direct calculating, we have

$$|\partial_\xi^q \int_0^1 \{R_t, F\} \circ \phi_F^t dt \circ \phi| \leq \frac{c\mu^2 s^4 \Gamma^2(r - r_+)}{\bar{\eta}^N}.$$

Besides

$$|\partial_\xi^q (\mathcal{P} - \mathcal{R}) \circ \phi_F^1 \circ \phi| \leq \frac{c\gamma^{N+7}\mu}{\bar{\eta}^N}(s^{3(1+b+\sigma)} + \frac{s_+^3}{s}),$$

finally,

$$|\partial_\xi^q P_+| \leq \Delta_+ \leq \frac{\gamma_+^{N+7} s_+^2 \mu_+}{\bar{\eta}_+^N}.$$

$\square$

Now, we finish a KAM step.

2.2. Iteration lemma. In this section, we will prove an Iteration Lemma which guarantees the inductive construction of canonical transformations in all KAM steps. Let r_0, s_0, γ_0, μ_0, Λ_0, $\mathcal{H}_0$, $\mathcal{N}_0$, e_0, Ω_0, $\mathcal{P}_0$ be given as above and let $\tilde{D}_0 = D(r_0, \beta_0)$, $D_0 = D(r_0, s_0)$, $K_0 = 0$, $\Phi_0 = id$. For any $\nu = 0, 1, \cdots$, we label all index-free qualities in **Section 2** by ν and all '$+$'$-$indexed qualities in **Section 2** by $\nu + 1$. This defines sequences

$$r_\nu,\ s_\nu,\ \mu_\nu,\ K_\nu,\ \Lambda_\nu,\ D_\nu,\ \tilde{D}_\nu,\ \mathcal{H}_\nu,\ \mathcal{N}_\nu,\ e_\nu,\ \Omega_\nu,\ \omega_\nu,\ h_\nu,\ \mathcal{P}_\nu,\ \Phi_\nu,$$

for $\nu = 0, 1, \cdots$. In particular,

$$\begin{aligned} \mathcal{H}_\nu &= \mathcal{H}_\nu(x, y) = \mathcal{N}_\nu + \mathcal{P}_\nu, \\ \mathcal{N}_\nu &= e_v + \langle \Omega_\nu, y \rangle + h_\nu(y), \end{aligned}$$

where $(y, x) \in \tilde{D}_\nu$, $\xi \in \Lambda_\nu$, $e_\nu = e_\nu(y_0)$, $\omega_\nu = -B^T(y_0, \xi)\Omega_\nu(\xi, \tilde{\varepsilon})$, $\Omega_\nu = \Omega_\nu(\xi, \tilde{\varepsilon})$ is analytic on Λ_ν, and $h_\nu = h_\nu(y, \xi, \tilde{\varepsilon})$ and $\mathcal{P}_\nu = \mathcal{P}_\nu(y, x, \xi)$ are analytic in $\xi \in \Lambda_\nu$ and $(y, x) \in \tilde{D}_\nu$. Moreover, for $\nu = 1, 2, \cdots$,

$$\begin{aligned} s_\nu &= s_{\nu-1}^{1+b+\sigma}, \quad \mu_\nu = c_0 s_{\nu-1}^\sigma \mu_{\nu-1}, \quad \gamma_\nu = \gamma_0(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}), \quad K_\nu = ([\log \frac{1}{s_{\nu-1}}] + 1)^3, \\ \Delta_\nu &= \frac{\mu_{\nu-1}}{\bar{\eta}_{\nu-1}^N}(c\mu_{\nu-1} s_{\nu-1}^4 \Gamma^2(r_{\nu-1} - r_\nu) + \gamma_{\nu-1}^{N+7}\mu_{\nu-1} s_\nu^3 (1 + \frac{1}{s_{\nu-1}}) \\ &\qquad + s_{\nu-1}^3 \mu \Gamma(r_{\nu-1} - r_\nu) + \gamma_{\nu-1}^{N+7} s_{\nu-1}^4), \\ \Lambda_\nu &= \{\xi \in \Lambda_{\nu-1} : |\langle k, \omega_{\nu-1}(\xi) \rangle| > \frac{|\tilde{\varepsilon}|\gamma_{\nu-1}}{|k|^\tau}, 0 < |k| \leq K_\nu\}, \qquad \bar{\eta}_\nu = \bar{\eta}_{\nu-1} - \frac{\bar{\eta}_0}{2^{\nu+1}}, \\ D_\nu &= D(r_\nu, s_\nu), \quad \tilde{D}_\nu = D(r_\nu + \frac{7}{8}(r_{\nu-1} - r_\nu), s_0), \end{aligned}$$

where c_0 is the maximum among c mentioned above.

LEMMA 2.2. *If $\mu_0 = \mu_0(\varepsilon)$ is sufficiently small, then the following hold for all $\nu = 0, 1, \cdots$:*

(1)

$$|e_\nu - e_0|_{\Lambda_\nu},\ |\Omega_\nu - \Omega_0|_{\Lambda_\nu},\ |\omega_\nu - \omega_0|_{\Lambda_\nu},\ |h_\nu - h_0|_{\Lambda_\nu} \le 2\gamma_0^{N+7}\mu_*,$$
$$|e_{\nu+1} - e_\nu|_{\Lambda_\infty},\ |\Omega_{\nu+1} - \Omega_\nu|_{\Lambda_\infty},\ |\omega_{\nu+1} - \omega_\nu|_{\Lambda_\infty},\ |h_{\nu+1} - h_\nu|_{\Lambda_\infty} \le \frac{\gamma_0^{N+7}\mu_*}{2^{\nu+1}},$$
$$|\partial_\xi^q \mathcal{P}_\nu|_{D_\nu\times\bar\Lambda_\nu} \le \frac{\gamma_\nu^{N+7}s_\nu^2\mu_\nu}{\bar\eta_\nu^N};$$

(2) $\Phi_{\nu+1}: \tilde D_{\nu+1}\times\Lambda_{\nu+1}\to\tilde D_\nu$ *is canonical and real analytic with respect to* $(y,x)\in\tilde D_{\nu+1}$, $\xi\in\Lambda_{\nu+1}$. *Moreover,* $\mathcal{H}_{\nu+1}=\mathcal{H}_\nu\circ\Phi_{\nu+1}$, *and, on* $\tilde D_{\nu+1}\times\Lambda_{\nu+1}$,

$$|\Phi_{\nu+1}-id|,|D\Phi_{\nu+1}-Did|,|D^i\Phi_{\nu+1}|\le\frac{\mu_*}{2^{\nu+1}},\quad 2\le i\le m;$$

(3) $\Lambda_{\nu+1}=\{\xi\in\Lambda_\nu:|\langle k,\omega_\nu(\xi)\rangle|>\frac{|\tilde\varepsilon|\gamma_\nu}{|k|^\tau},K_\nu<|k|\le K_{\nu+1}\}$.

Proof. The lemma will be proved by performing the KAM steps inductively. What we should do is to verify conditions (2.5), (2.6), (2.9), (2.11), (2.12) and (2.14) for all $\nu=0,1,\cdots$.

Inductively,

$$\begin{aligned}\mu_\nu &= c_0^\nu\mu_0 s_0^{\frac{\sigma((1+b+\sigma)^\nu-1)}{b+\sigma}},\\ s_\nu &= s_0^{(1+b+\sigma)^\nu}.\end{aligned}$$

Then

$$s_{\nu+1}=s_\nu s_0^{(1+b+\sigma)^\nu(b+\sigma)}\le s_\nu s_0^{b+\sigma}\le\frac{s_\nu}{16},$$

i.e. (2.5) holds. Denote

$$E_\nu=\frac{r_\nu-r_{\nu+1}}{8}=\frac{\delta^{\nu+2}\gamma_0(1-\delta)}{16}$$

and with $\delta(1+b+\sigma)>1$ we have

$$\frac{E_\nu}{2}\log\frac{1}{s_\nu} = \frac{\delta^{\nu+2}\gamma_0(1-\delta)}{32}\log s_0^{-(1+b+\sigma)^\nu}\ge-\frac{\gamma_0\delta^2(1-\delta)}{32}\log s_0\ge1.$$

Therefore

$$\begin{aligned}&\log(n+N+1)!+3(n+N)\log([\log\frac{1}{s_\nu}]+1)-\frac{E_\nu}{2}([\log\frac{1}{s_\nu}]+1)^3\\ &\le-(m+1)(1+b+\sigma)\log\frac{1}{s_\nu}.\end{aligned}$$

Thus

$$\int_{K_{\nu+1}}^\infty\lambda^{n+N}e^{-\frac{\lambda E_\nu}{2}}d\lambda\le(n+N+1)!K_{\nu+1}^{n+N}e^{-\frac{K_{\nu+1}E_\nu}{2}}\le s_{\nu+1}^3,$$

i.e. (2.6) holds. Similarly,

$$s_\nu K_{\nu+1}^{\tau+1}=s_0^{(1+b+\sigma)^\nu}([\log\frac{1}{s_\nu}]+1)^{3(\tau+1)}\le s_0^{(1+b+\sigma)^\nu}(\log\frac{1}{s_0^{(1+b+\sigma)^\nu}}+2)^{3(\tau+1)}.$$

Since $x^{\beta}(\log\frac{1}{x}+c)^{\xi}\to 0$, as $x\to 0$, where $\beta>0$, $\xi>1$ and $c>1$ are constant, hypothesis (2.9), $s_\nu K_{\nu+1}^{\tau+1}=O(\gamma_0)$, is obvious.

Let $l_0=b$, $\eta=8+n+4[\tau]+4$, where $[\tau]$ is the integral part of τ. Without doubt

$$\Gamma_\nu=\sum_{|j|\le 2,\ 0<|k|\le K_{\nu+1}}|k|^{\tau(|j|+|q|+1)+|j|+|q|+|z|}e^{-\frac{|k|(r-r_+)}{8}}\le\int_1^\infty\lambda^\eta e^{-\lambda E_\nu}d\lambda\le\frac{\eta!}{E_\nu^\eta}.$$

Besides

$$\frac{\mu_\nu^{l_0}}{E_\nu^{\eta+1}}=(\frac{16}{\gamma_0(1-\delta)\delta^{\nu+2}})^{\eta+1}c_0^\nu\mu_0^{l_0}s_0^{\frac{\sigma}{b+\sigma}((1+b+\sigma)^\nu-1)}\le c_*\mu_0^{l_0}(\frac{s_0^\sigma c_0}{\delta^{\eta+1}})^\nu,$$

we have

$$\frac{c_0\mu_\nu\Gamma_\nu}{E_\nu\bar\eta_\nu^N}\le c_0\eta!\frac{\mu_\nu^{l_0}}{E_\nu^{\eta+1}}\le 1,$$

as ε_0 small enough, i.e. (2.11) holds. By $\frac{c_0s_\nu\mu_\nu\Gamma_\nu}{s_{\nu+1}\bar\eta_\nu^N}\le 3$, (2.12) is obvious.

Moreover, by making ε_0 small, we have $c_0\mu_\nu^{a_0}\Gamma_\nu^3\le\frac{1}{2^\nu}$. Next, for each $\nu\ge 1$

$$\begin{aligned}
|\partial_\xi^l\Delta_{\nu+1}| &\le \big(c\mu_\nu^2s_\nu^4\Gamma_\nu^2(r_\nu-r_{\nu+1})+\gamma_\nu^{N+7}\mu_\nu s_{\nu+1}^3(1+\frac{1}{s_\nu})\\
&\quad+s_\nu^4\mu_\nu\Gamma_\nu(r_\nu-r_{\nu+1})+c\gamma_\nu^{N+7}s_\nu^4\mu_\nu\big)/\bar\eta_\nu^N\\
&\le[2c\mu_{\nu+1}^2s_{\nu+1}^2s_\nu^{4-2\sigma-2(1+b+\sigma)}+2\gamma_\nu^{N+7}\mu_{\nu+1}s_{\nu+1}^ms_\nu^{1+\sigma}\\
&\quad+2\gamma_\nu^{N+7}\mu_{\nu+1}s_{\nu+1}^2s_\nu^b+3s_\nu^{4-\sigma-2(1+b+\sigma)}\mu_{\nu+1}s_{\nu+1}^2]\frac{\Gamma_\nu^2(r_\nu-r_{\nu+1})}{\bar\eta_\nu^N}\\
&\le\gamma_{\nu+1}^{N+7}s_{\nu+1}\mu_{\nu+1}[2s_\nu^{1+b}+2s_\nu^b+2\frac{s_\nu^{4-2\sigma-2(1+b+\sigma)}\mu_{\nu+1}}{\gamma_0^{N+7}}\\
&\quad+3\frac{s_\nu^{4-\sigma-2(1+b+\sigma)}}{\gamma_0^{N+7}}]\frac{\Gamma_\nu^2(r_\nu-r_{\nu+1})}{\bar\eta_\nu^N}\\
&\le c\gamma_{\nu+1}^{N+7}s_{\nu+1}^2\mu_{\nu+1}(s_\nu^b+\frac{s_\nu^{\frac{3}{2}}}{\gamma_0^{N+7}})\frac{\Gamma_\nu^2(r_\nu-r_{\nu+1})}{\bar\eta_\nu^N}\\
&\le C\gamma_{\nu+1}^{N+7}s_{\nu+1}^2\mu_{\nu+1}s_\nu^{l_0}\frac{\Gamma_\nu^2(r_\nu-r_{\nu+1})}{\bar\eta_\nu^N}\\
&\le\frac{C\gamma_{\nu+1}^{N+7}s_{\nu+1}^2\mu_{\nu+1}}{\bar\eta_{\nu+1}^N},
\end{aligned}$$

i.e. (2.14) holds. □

For brevity we omit the measure estimate of $|\Lambda_0\setminus\Lambda_*|$ and for details we refer the reader to [25] and [42].

3. Proof of main theorem

Actually, the main task of the persistence of invariant tori for multiscale generalized Hamiltonian (1.1) is to achieve the program of the KAM iteration consisting of infinite KAM steps by induction. Without loss of generality, we assume that there is a closed region $\Lambda\subset R^d$ and a C^{l_0} diffeomorphism $y:\Lambda\to M$ $(=y(\Lambda))$. Let $\xi\in\Lambda$ and consider the transformation: $y \mapsto y + y(\xi)$. Then (1.1) turns into a parameterized Hamiltonian system of the following form:

$$\begin{aligned} \mathcal{H} &= \mathcal{N}(y,\xi,\tilde{\varepsilon}) + \varepsilon^2 P(x,y,\xi), \\ \mathcal{N} &= e(\xi) + \langle \Omega(\xi), y\rangle + h(y,\xi) + O(|y|^3), \\ \Omega &= \varepsilon_0 \partial_y h_0 + \varepsilon_1 \partial_y h_1 + \cdots + \varepsilon_{m_0} \partial_y h_{m_0}, \\ h &= \langle y, A(\xi) y\rangle, \\ A &= \varepsilon_0 \partial_y^2 h_0 + \varepsilon_1 \partial_y^2 h_1 + \cdots + \varepsilon_{m_0} \partial_y^2 h_{m_0}, \end{aligned} \tag{3.1}$$

where $\xi \in \Lambda = \{\lambda : |\lambda| \le \delta_1\} \subset R^d$, $\tilde{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_{m_0})$ and ε defined as above. Consider the following symplectic transformation:

$$x \to x, \quad y \to \sqrt{\varepsilon} y, \quad H \to \frac{1}{\sqrt{\varepsilon}} H,$$

and denote $\sqrt{\varepsilon} = \varepsilon$, then the Hamiltonian (3.1) is changed to

$$\begin{aligned} \mathcal{H} &= \mathcal{N}(y,\xi,\tilde{\varepsilon}) + \varepsilon^2 \bar{P}(x,y,\xi), \\ \mathcal{N} &= \frac{e(\xi)}{\varepsilon} + \langle \Omega(\xi), y\rangle + \varepsilon h(y,\xi), \end{aligned} \tag{3.2}$$

where $\varepsilon^2 \bar{P} = \varepsilon^3 P + \varepsilon^2 O(|y|^3)$, Ω and h defined as above. Moreover,

$$|\bar{P}| \le c\gamma^{N+7} s^2 \mu,$$

if $\gamma = \varepsilon^{\frac{1}{12(7+N)}}$, $s = \varepsilon^{\frac{1}{3}}$, $\mu = \varepsilon^{\frac{1}{4}}$. Hence, with Theorem 2.1 we can get Theorem 1.1.

REMARK 3.1. Normal forms (3.2) and (2.1) seem a little different. In normal form (3.2), the perturbation is $\varepsilon^2 \bar{P}(x,y,\xi) = O(\varepsilon^3)$, which is small comparing the integrable part. The difficulty is the term $\varepsilon h(y,\xi)$, which is bad for the preservation of the frequency. In fact, in our case, this difficulty could be overcome since the coefficient of the perturbation is ε^2. We could achieve the proof of Theorem 1.1 step by step using the proof of Theorem 2.1.

Acknowledgements. We are grateful to anonymous referees for their careful reading and valuable suggestions. The first author was supposed by China Postdoctoral Science Foundation (2021M701396). The second author was supposed by National Basic Research Program of China (No. 2013CB834100), NSFC (No. 12071175) and Jilin-DRC (No. 2017C028-1). The third author was supposed by NSFC (No. 11571065, No. 11171132, No. 120701132), National Research Program of China Grant (No. 2013CB834100) and Natural Science Foundation of Jilin Province (No. 20200201253JC).

Appendix A. The norm on the derivative of F. By induction, we have

$$\begin{aligned} |\partial_\xi^q \partial_y^j L_k^{-1}| &\le c|k|^{|j|+|q|-1} |\tilde{\varepsilon}|^{|j|+|q|-1} \frac{|k|^{\tau(|j|+|q|)}}{|\tilde{\varepsilon}|^{|j|+|q|} \gamma^{|j|+|q|}} \\ &\le \frac{|k|^{\tau(|j|+|q|)+|j|+|q|-1}}{|\tilde{\varepsilon}| \gamma^{|j|+|q|}}. \end{aligned}$$

Then

$$|\partial_\xi^q \partial_y^j f_{ki}| \le \frac{c|k|^{\tau(|j|+|q|)+|j|+|q|-1}}{|\tilde{\varepsilon}| \gamma^{|j|+|q|}} \varepsilon \frac{\gamma^{N+7} s^{2-|i|} \mu}{\bar{\eta}^N} e^{-|k| r}$$

$$\leq c|k|^{\tau(|j|+|q|)+|j|+|q|-1}\frac{s^{2-|i|}\mu}{\bar{\eta}^N}e^{-|k|r}.$$

Finally,

$$\begin{aligned}|\partial_\xi^q\partial_y^j\partial_x^z F| &\leq \sum_{|j|\leq 2,\ 0<|k|\leq K_+}|k|^z|\partial_y^j(\partial_\xi^q f_{ki}y^i)|e^{|k|(r_+ + \frac{7}{8}(r-r_+))}\\ &\leq \frac{s^{2-|i|}\mu}{\bar{\eta}^N}\Gamma(r-r_+),\end{aligned}$$

where $\Gamma(r-r_+)=\sum_{|j|\leq 2,\ 0<|k|\leq K_+}|k|^{\tau(|j|+|q|)+|j|+|q|-1+|z|}e^{-\frac{|k|(r-r_+)}{8}}$.

Appendix B. The estimate on ϕ_F^t. Denote $\phi_{F_1}^t$, $\phi_{F_2}^t$ as the components of ϕ_F^t in y, x planes, respectively, and let X_F be the vector field on the right-hand side of (2.10), i.e.

$$\begin{pmatrix}\dot{y}\\ \dot{x}\end{pmatrix}=I(y,\xi)\nabla F(y,x)=\begin{pmatrix}0 & B\\ -B^T & C\end{pmatrix}\begin{pmatrix}\partial_y F\\ \partial_x F\end{pmatrix}.$$

Then $\phi_F^t=id+\int_0^t X_F\circ\phi_F^\lambda d\lambda$. For any $(y,x)\in D_3$, we let $t_*=\sup\{t\in[0,1]:\phi_F^t(y,x)\in D_4\}$. By (2.5), i.e. $s_+\leq\frac{s}{16}$, we have $D_4\subset D_*$, where $D_4=D(4s_+,r_++\frac{3}{8}(r-r_+))$, $D_*=D(\frac{s}{2},r_++\frac{6}{7}(r-r_+))$. Then

$$\begin{aligned}|\phi_{F_1}^t(y,x)| &= |y|+|\int_0^t B(\phi_{F_1}^\lambda+y_0)F_x\circ\phi_F^\lambda d\lambda|\leq s_++\frac{cs^2\mu\Gamma(r-r_+)}{\bar{\eta}^N}\leq 4s_+,\\ |\phi_{F_2}^t(y,x)| &= |x|+|\int_0^t -B(\phi_{F_1}^\lambda+y_0)F_y\circ\phi_F^\lambda+C(\phi_{F_1}^\lambda+y_0)F_x\circ\phi_F^\lambda d\lambda|\\ &= r_++\frac{7}{8}(r-r_+)+\frac{c\mu\Gamma(r-r_+)}{\bar{\eta}^N}\leq r_++\frac{3}{8}(r-r_+),\end{aligned}$$

where B, C are the matrices defined as above. This shows that $\phi_F^t(y,x)\in D_4$ for all $0\leq t\leq t_*$. Hence $t_*=1$ holds, i.e. $\phi_F^t:D_3\to D_4$ for all $0\leq t\leq 1$. Therefore, $\Phi_+:D_+\to D(s,r)$. It follows straightforward form the argument above that $|\phi_F^t-id|_{\tilde{D}}\leq\frac{c\mu\Gamma(r-r_+)}{\bar{\eta}^N}$.

With Gronwall Inequality and

$$D\phi_F^t = Did+\int_0^t((DI\cdot DF)\circ\phi_F^\lambda\cdot D\phi_F^\lambda+(I\cdot D^2F)\circ\phi_F^\lambda\cdot D\phi_F^\lambda)d\lambda,$$

we have

$$\begin{aligned}|D\phi_F^t-Did| &= |\int_0^t(DI\cdot DF-I\cdot D^2F)\circ\phi_F^\lambda\cdot(D\phi_F^\lambda-Did)\\ &\qquad+(DI\cdot DF-I\cdot D^2F)\circ\phi_F^\lambda d\lambda|\\ &\leq\int_0^t e^{\int_s^t|(DI\cdot DF-I\cdot D^2F)\circ\phi_F^r|dr}|(DI\cdot DF-I\cdot D^2F)\circ\phi_F^s|ds\\ &\leq c|DI\cdot DF-I\cdot D^2F|e^{c|DI\cdot DF-I\cdot D^2F|}\\ &\leq\frac{c\mu\Gamma(r-r_+)}{\bar{\eta}^N}.\end{aligned}$$

Appendix C. The estimate on ϕ. Let

$$Ay^* = -\varepsilon p_{01}.$$

Then

$$(y^*)^T A^T A y^* = -\varepsilon^2 p_{01}^T p_{01}.$$

Assume

$$A^T A \geq \min\{\varepsilon_1^2, \cdots, \varepsilon_m^2\} I.$$

Obviously,

$$\frac{(y^*)^T}{|y^*|}(A^T A - \min\{\varepsilon_1^2, \cdots, \varepsilon_m^2\} I)\frac{y^*}{|y^*|}|y^*|^2 \geq 0.$$

Therefore,

$$\begin{aligned} \min\{\varepsilon_1^2, \cdots, \varepsilon_m^2\}|y^*|^2 &= \frac{(y^*)^T}{|y|}\min\{\varepsilon_1^2, \cdots, \varepsilon_m^2\} I \frac{y^*}{|y^*|}|y^*|^2 \\ &\leq \frac{(y^*)^T}{|y^*|} A^T A \frac{y^*}{|y^*|}|y^*|^2 \\ &= \varepsilon^2 p_{01}^T p_{01}, \end{aligned}$$

i.e.

$$\sqrt{|y^*|^2} \leq \sqrt{\frac{\varepsilon^2}{\min\{\varepsilon_1^2, \cdots, \varepsilon_m^2\}} p_{01}^T p_{01}} \leq \sqrt{p_{01}^T p_{01}}.$$

Therefore, $|y^*| \leq c\gamma^{N+7} s\mu$. By induction, we have $|\partial_\xi^q y^*| \leq c\frac{\gamma^{N+7} s\mu}{\bar{\eta}^N}$.

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*, Russ. Math. Surv., 18(5):13–40, 1963. 1

[2] V.I. Arnold, *Small denominators and problems of stability of motion in classical and celestial mechanics*, Russ. Math. Surv., 18(6):91–192, 1963. 1, 1.4

[3] V.I. Arnold, *Sur la topologie des écoulements stationaries des fluides fluiddes parfaits*, C.R. Acad. Sci., 261:17–20, 1965. 1.3

[4] V.I. Arnold, *Dynamical Systems III*, Springer-Verlag, Berlin, 1988. 1

[5] L. Biasco, L. Chierchia, and E. Valdinoci, *N-dimensional elliptic invariant tori for the planar (N+1)-body problem*, SIAM J. Math. Anal., 37(5):2560–2588, 2006. 1

[6] A.D. Bruno, *Analytic form of differential equations*, Tr. Mosk. Mat. Obs., 25:131–288, 1971. 1

[7] J.R. Cary and R.G. Littlejohn, *Hamiltonian mechanics and its application to magnetic field line flow*, Ann. Phys., 151(1):1–34, 1982. 1

[8] A. Celletti and L. Chierchia, *KAM stability for a three-body problem of the solar system*, Z. Angew. Math. Phys., 57:33–41, 2006. 1

[9] C.Q. Cheng and Y.S. Sun, *Existence of invariant tori in three dimensional measure-preserving mapping*, Celestial Mech. Dyn. Astron., 47(3):275–292, 1990. 1.3

[10] F.Z. Cong and Y. Li, *Existence of higher-dimensional invariant tori for Hamiltonian systems*, J. Math. Anal. Appl., 222(1):255–267, 1998. 1

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

[12] L.H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 15(4):115–147, 1988. 1

[13] A. Fortunati and S. Wiggins, *A Kolmogorov theorem for nearly integrable Poisson systems with asymptotically decaying time-dependent perturbation*, Regul. Chaot. Dyn., 20(4):476–485, 2015. 1

[14] Y.C. Han, Y. Li, and Y. Yi, *Degenerate lower dimensional tori in Hamiltonian systems*, J. Differ. Equ., 227(2):670–691, 2006. 1

[15] Y.C. Han, Y. Li, and Y. Yi, *Invariant tori in Hamiltonian systems with high order proper degeneracy*, Ann. Henri Poincaré, 10:1419–1436, 2010. 1

[16] M.R. Herman, *Différéntiabilité optimae et contreexemples à la fermeture en topologie C^∞ des orbites récurrentes de flots hamiltoniens*, C.R. Acad. Sci. Paris. Sér. I Math., 313(1):49–51, 1991. 1

[17] M.R. Herman, *Exemples de flots Hamiltoniens dont aucune perturbation en topologie C^∞ n'a d'orbites périodiques sur un ouvert de surfaces dénergies*, C.R. Acad. Sci. Paris. Sér. I Math., 312:989–994, 1991. 1

[18] Q.D. Huang, F.Z. Cong, and Y. Li, *Persistence of elliptic invariant tori for Hamiltonian systems*, Nonlinear Anal., 45(2):241–260, 2001. 1

[19] Q.D. Huang, F.Z. Cong, and Y. Li, *Persistence of hyperbolic invariant tori for Hamiltonian systems*, J. Differ. Equ., 164(2):355–379, 2000. 1

[20] M. Inarrea, V. Lanchares, J.F. Palacián, A.I. Pascual, J.P. Salas, and P. Yanguas, *Reduction of some perturbed Keplerian problems*, Chaos Solitons Fractals, 27(2):527–536, 2006. 1

[21] S.B. Kuksin, *Nearly Integrable Infinite-Dimensional Hamiltonian Systems*, Lect. Notes Math., Springer-Verlag, 1556, 1993. 1

[22] S.B. Kuksin, *Hamiltonian perturbations of infinite dimensional linear systems with an imaginary spectrum*, Funct. Anal. Its Appl., 21(3):192–205, 1987. 1

[23] S.B. Kuksin, *Elements of a qualitative theory of Hamiltonian PDEs*, Doc. Math. J. DMV (Extra volume ICM) II, 819–829, 1998. 1

[24] A.N. Kolmogorov, *On conservation of conditionally periodic motions for a small change in Hamilton's function*, Dokl. Akad. Nauk SSSR, 98:527–530, 1954. 1

[25] Y. Li and Y. Yi, *Persistence of invariant tori in generalized Hamiltonian systems*, Ergod. Theor. Dyn. Sys., 22(4):1233–1261, 2002. 1, 1.3, 1.3, 1, 2.1.1, 2.2

[26] Y. Li and Y. Yi, *A quasi-periodic Poincaré's theorem*, Math. Ann., 326(4):649–690, 2003. 1

[27] Y. Li and Y. Yi, *Persistence of lower dimensional tori of general types in Hamiltonian systems*, Trans. Amer. Math. Soc., 357(4):1565–1600, 2005. 1

[28] R. de la Llave, *Recent progress in classical mechanics*, in K. Schmüdgen (eds.), Mathematical Physics X, Springer, Berlin, 3–19, 1992. 1

[29] V.K. Melnikov, *On certain cases of conservation of almost periodic motions with a small change of the Hamiltonian function*, Soviet Math. Dokl., 165:1592–1596, 1965. 1

[30] V.K. Melnikov, *A certain family of conditionally periodic solutions of a Hamiltonian system*, Soviet Math. Dokl., 181:882–886, 1968. 1

[31] K.R. Meyer, J.F. Palacián, and P. Yanguas, *Geometric averaging of Hamiltonian systems: periodic solutions, stability, and KAM tori*, SIAM J. Appl. Dyn. Syst., 10(3):817–856, 2011. 1

[32] I. Mezić and S. Wiggins, *On the integrability and perturbation of three-dimensional fluid flows with symmetry*, J. Nonlinear Sci., 4(2):157–194, 1994. 1, 1.3

[33] J. Moser, *On invariant curves of area preserving mappings of an annulus*, Nathr. Akad. Wiss. Gott. Math. Phys. K1., 2:1–20, 1962. 1

[34] J. Moser, *Convergent series expansions for quasi-periodic motions*, Math. Ann., 169:136–176, 1967. 1

[35] J. Moser, *Old and new applications of KAM theory*, in C. Simó (eds.), Hamiltonian Systems with Three or More Degrees of Freedom, NATO ASI Series, Vol. 533, Springer, 84–192, 1999. 1

[36] P.J. Olver, *Applications of Lie Groups to Differential Equations*, Springer, New York, 1986. 1

[37] I.O. Parasyuk, *Conservation of multidimensional invariant tori of Hamiltonian systems*, Ukr. Mat. J., 36(4):467–473, 1984. 1

[38] J. Pöschel, *Integrability of Hamiltonian systems on Cantor sets*, Commun. Pure Appl. Math., 35(5):653–696, 1982. 1

[39] J. Pöschel, *On elliptic lower-dimensional tori in Hamiltonian systems*, Math. Z., 202(4):559–608, 1989. 1

[40] J. Pöschel, *A KAM-theorem for some nonlinear partial differential equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. Ser. IV, 23(1):119–148, 1996. 1

[41] W.C. Qian, Y. Li, and X. Yang, *The isoenergetic KAM-type theorem at resonant case for nearly integrable Hamiltonian systems*, J. Appl. Anal. Comput., 5:1616–1638, 2019. 1

[42] W.C. Qian, Y. Li, and X. Yang, *Multiscale KAM theorem for Hamiltonian systems*, J. Differ. Equ., 266(1):70–86, 2019. 1, 1.1, 2.2
[43] W.C. Qian, Y. Li, and X. Yang, *Melnikov's conditions in matrices*, J. Dynam. Differ. Equ., 32(4):1779–1795, 2020. 1
[44] W.C. Qian, Y. Li, and X. Yang, *Persistence of Lagrange invariant tori at tangent degeneracy*, J. Differ. Equ., 268(9):5078–5112, 2020. 1, 1
[45] H. Rüssmann, *Talk held at the ETH Zurich, February 1986*. 1, 1.1
[46] M.B. Sevryuk, *Some problems of the KAM-theory: conditionally periodic motions in typical systems*, Russ. Math. Surv., 50(2):341–353, 1995. 1
[47] M.B. Sevryuk, *The classical KAM theory at the dawn of the twenty-first century. Dedicated to Vladimir Igorevich Arnold on the occasion of his 65th birthday*, Mosc. Math. J., 3(3):1113–1144, 2003. 1
[48] D.V. Treshchëv, *The mechanism of destruction of resonance tori of Hamiltonian systems*, Mat. USSR Sb., 68(1):181, 1991. 1
[49] L. Xu, Y. Li, and Y. Yi, *Lower-dimensional tori in multi-scale, nearly integrable Hamiltonian systems*, Ann. Henri Poincaré, 18(1):53–83, 2017. 1
[50] L. Xu, Y. Li, and Y. Yi, *Poincaré-Treshchev mechanism in multiscale, nearly integrable Hamiltonian systems*, J. Nonlinear Sci., 28(1):337–369, 2018. 1
[51] J. Yoccoz, *Travaux de Herman sur les tores invariants*, Asterisque, 206(4):311–344, 1992. 1
[52] J. You, *Perturbations of lower-dimensional tori for Hamiltonian systems*, J. Differ. Equ., 152(1):1–29, 1999. 1