PERSISTENCE OF FLOQUET INVARIANT TORI FOR A CLASS OF NON-CONSERVATIVE DYNAMICAL SYSTEMS

JUNXIANG XU

(Communicated by Carmen C. Chicone)

ABSTRACT. In this paper we consider a class of non-conservative dynamical system with small perturbation. By the KAM method we prove existence of Floquet invariant tori under the weakest non-resonant conditions.

1. Introduction and main results

With the development of KAM theory, there are many results (called KAM theorems) about persistence of invariant tori of integrable systems under small perturbation for some conservative systems such as hamiltonian systems and reversible systems. More recently, the classical KAM theorems have been improved greatly in weaker non-degenerate conditions and fewer non-resonant conditions. The weakest non-degenerate condition is Rüssmann’s non-degenerate condition (see [7, 11, 12]). Regarding weaker non-resonant conditions, in [13] we proved a KAM theorem for hamiltonian systems only under the first Melnikov’s non-resonant condition (1.2). The result implies that the second Melnikov’s non-resonant condition (1.3) is not necessary for the classical KAM theorems. We also obtained a similar result for reversible systems [14]. In this paper we want to generalize these results to a class of non-conservative systems.

Consider the following dynamical system:

\[
\begin{align*}
\dot{x} &= \xi + P_1(x, y; \xi), \\
\dot{y} &= \Omega(\xi)y + P_2(x, y; \xi),
\end{align*}
\]

where \((x, y) \in \mathbb{R}^n/2\pi Z^m \times \mathbb{R}^m\), \(\xi \in O \subset \mathbb{R}^n\) is a frequency parameter, \(\Omega\) is a matrix depending on \(\xi\), and \(P_1\) and \(P_2\) are small perturbations.

If \(P_1 = 0, P_2 = 0\), then the system (1.1) is integrable and for each \(\xi \in O\) has an invariant torus \(T^n \times \{0\}\), on which there exists a linear flow \(x = \xi t + x_0\), \(y = 0\). We want to know whether the system (1.1) still has many invariant tori if \(P_1\) and \(P_2\) are sufficiently small.

Moser [11] studied a system of the form (1.1) by the rapidly convergent method, and the result can be applied to both hamiltonian systems and reversible systems. By the usual KAM method, Broer, Huitema and Sevryuk [2] proved that if \(\xi\) and...
the eigenvalues of $\Omega$ satisfy the following non-resonant conditions:

\begin{align}
&(1.2) \quad \Omega_j(\xi) \neq 0, \quad \forall \, k \in \mathbb{Z}^n, |k| \leq M + 1, \quad 1 \leq j \leq m, \nonumber \\
&(1.3) \quad \Omega_j(\xi) - \Omega_j(\xi) \neq 0, \quad \forall \, |k| + |i - j| \neq 0, |k| \leq 2M + 1, \nonumber
\end{align}

where $i = \sqrt{-1}$, $\Omega_j(\xi), \Omega_2(\xi), \ldots, \Omega_m(\xi)$ are the eigenvalues of $\Omega$ and $M = \max_{1 \leq j \leq m} \sup_{\xi \in O} |\partial \Omega_j / \partial \xi|$, then, if $P_1$ and $P_2$ are sufficiently small, most of the invariant tori can persist. In the recent paper [3], Bambusi considered non-

By motivation of these works, we want to consider the persistence of invariant tori under a weaker non-resonant condition. We will show that if the first Melnikov’s non-resonance condition \[1.2\] holds, then the system \[1.1\] persists many invariant tori under a weaker non-resonant condition. We will show that if the first Melnikov’s

Before stating our results, we first give some notation and assumptions. Denote a complex neighborhood of $T^n \times \{0\}$ by

\[ D(s, r) = \{(x, y) \in C^n / 2\pi Z^n \times C^m \mid |\text{Im} \, x| \leq s, |y| \leq r\}, \]

where $|\text{Im} \, x| = \max_{1 \leq \xi \leq n} |\text{Im} \, x|$, $|y| = \max_{1 \leq \xi \leq m} |y|$.

Let $O$ be a bounded connected closed domain of $R^n$. Let $C^L(O)$ be the space of the $L$-th continuously differentiable function on $O$. For $f(\xi) \in C^L(O)$, define $\|f\|_L = \max_{|\xi| \leq L} \sup_{\xi \in O} |\partial \xi^n f(\xi)|$.

Let $f(x, y; \xi)$ be analytic in $(x, y)$ on $D(s, r)$ and let it belong to $C^L(O)$ in $\xi$. Then we have

\[ f = \sum_{k, l} f_{k, l}(\xi) e^{i(k, x)} y^l. \]

Define

\[ \|f\|_{D(s, r)}^L = \sup_{|y| \leq r} \sum_{k, l} \|f_{k, l}(\xi)\| e^{s|k|} y^l. \]

**Theorem 1.1.** Suppose $\Omega = \text{diag}(\Omega_1(\xi), \Omega_2(\xi), \ldots, \Omega_m(\xi))$ is analytic on $O$ and the non-resonant condition \[1.2\] holds. Also suppose that $P_1$ and $P_2$ are analytic in $(x, y) \in D(s, r)$ and belong to $C^L(O)$ in $\xi$, where $L = m^2$. For sufficiently small $\alpha > 0$, there exists an $\epsilon > 0$ depending on $\Omega, \alpha, O, n, m$ such that if

\[ (1.4) \quad \|P_1\|_{D(s, r)}^L \leq \epsilon, \quad \frac{1}{r} \|P_2\|_{D(s, r)}^L \leq \epsilon, \]

then there exists a non-empty subset $O_\ast$ of $O$, and for each $\xi \in O_\ast$ there exists an analytic transformation\[1\]

\[ \Phi_\ast(x, y; \xi) : D_{s/2, r/2} \to D_{s, r} \]

such that the system \[1.1\] is changed to

\[ (1.5) \quad \dot{x} = \omega_\ast + P_1, \quad \dot{y} = (\Omega + A_\ast)y + P_2, \]

where $A_\ast$ is a constant matrix, and $P_1$ and $P_2$ satisfy

\[ P_1(x, 0; \xi) = P_2(x, 0; \xi) = 0, \quad \partial_y P_2(x, y; \xi)|_{y=0} = 0. \]

Hence, $\forall \xi \in O_\ast$, $\Phi_\ast(T^n \times \{0\}; \xi)$ is an invariant torus of the system \[1.1\] with the frequency $\omega_\ast$, $\omega_\ast$ and $A_\ast$ satisfy $\|\omega_\ast(\xi) - \xi\|_L \leq 2\epsilon$ and $\|A_\ast\|_L \leq 2\epsilon$. Moreover, we have $\text{mes}(O \setminus O_\ast) \to 0$ as $\alpha \to 0$.

\[ ^1\Phi_\ast \text{ is analytic in } (x, y) \text{ and Whitney smooth in } \xi. \]
The proof of Theorem 1.1 is based on the following theorem.

**Theorem 1.2.** Let $\Omega(\xi) = \text{diag}(\Omega_1 I_{d_1}, \Omega_2 I_{d_2}, \ldots, \Omega_m I_{d_m})$, where $I_{d_j}$ is a $d_j$-unit matrix for $j = 1, 2, \cdots, m$. Let $L = (\max_{1 \leq i \leq m} d_i)^2$. Suppose $\Omega_j(\xi) (1 \leq j \leq m)$ are analytic on $O$ and satisfy the non-resonant conditions (1.2) and (1.3). Then the result of Theorem 1.1 holds true.

**Remark 1.3.** In [1] Moser considered a modified system of (1.1) so that the frequency can be fixed in KAM iteration. In this case, the transformation is analytic with respect to the frequency parameter, and the invariant tori belong to the modified system. Here we directly consider the system (1.1). So we must adjust the frequency parameter for small divisor conditions in KAM iteration. Thus, we cannot obtain the persisting invariant tori are for the original system.

**Remark 1.4.** In Theorem 1.1 if all $\Omega_j$ have non-zero real parts, this corresponds to the case of hyperbolic invariant tori and can be dealt in the same way as in hamiltonian system (see [3, 4]). So, we are really interested in the case that $\Omega_j$ are pure imaginary numbers.

## 2. Proof of Theorems

The difference between Theorem 1.1 and Theorem 1.2 is in normal frequencies. In Theorem 1.1 only condition (1.2) is required, while in Theorem 1.2 both (1.2) and (1.3) are needed. We will first prove that the case of Theorem 1.1 can be reduced to the special case of Theorem 1.2. Thus, the proof of Theorem 1.1 is reduced to that of Theorem 1.2. Then, by KAM method we prove Theorem 1.2.

**Proof of Theorem 1.1.** Below we prove that under the assumptions of Theorem 1.1 the normal frequencies can be changed by a transformation such that the new normal frequencies satisfy both the non-resonant conditions (1.2) and (1.3). As in [13, 14], we divide $\Omega_1, \Omega_2, \cdots, \Omega_m$ into some groups by a resonant relation. If there exists $k_{ij} \in \mathbb{Z}^n$ such that $\Omega_j(\xi) - \Omega_i(\xi) = \sqrt{-1}(\xi, k_{ij}), \forall \xi \in O$, we say that $\Omega_j$ is resonant with $\Omega_i$ on $O$. Since the normal frequencies are analytic on the connected domain $O$, if $\Omega_i$ and $\Omega_j$ are not resonant at one point, then they are not resonant for almost every point on $O$.

We re-arrange $\Omega_1, \Omega_2, \cdots, \Omega_m$ to put together the resonant normal frequencies on $O$ and denote them by

$$
\Omega_1, \cdots, \Omega_{d_1}, \Omega_1^2, \cdots, \Omega_{d_2}^2, \cdots, \Omega_{d_m}^m
$$

such that

$$
\Omega_j^i - \Omega_1^i = \sqrt{-1}(\xi, k_{ij}), \forall \xi \in O, \forall 1 \leq j \leq d_i, \forall 1 \leq i \leq m,
$$

and

$$
\Omega_j^i - \Omega_1^j \neq \sqrt{-1}(\xi, k), \forall i \neq j, \forall k \in \mathbb{Z}^n.
$$

We also re-arrange $y_1, y_2, \cdots, y_m$ correspondingly and denote them by $y_j^i$. Let $y^i = (y_1^i, y_2^i, \cdots, y_m^i)$, $1 \leq i \leq \tilde{m}$, and $y = (y_1, y_2, \cdots, y_{\tilde{m}})$.

Now define a transformation by

$$
x_+ = x, \quad y_{+j} = e^{\sqrt{-1}(\xi, k_j)} y_j^i, \quad 1 \leq i \leq \tilde{m}, \quad 1 \leq j \leq d_i.
$$

Under this transformation, the system (1.1) is changed to

$$
\dot{x}_+ = \xi + P_{+1}, \quad \dot{y}_+ = \Omega_+ (\xi) y_+ + P_{+2},
$$
where
\[ \Omega_+ (\xi) = \text{diag}(\Omega_1 I_{d_1}, \Omega_2 I_{d_2}, \ldots, \Omega_m I_{d_m}). \]

From the above we know that the resonant normal frequencies are changed to be multiple ones by the transformation, so that the non-resonant condition \((1.3)\) holds. Obviously, the non-resonance condition \((1.2)\) still holds after transformation. Thus, the non-resonant condition \((1.2)\) and \((1.3)\) both hold for the new normal frequencies \(\Omega_+\).

**Proof of Theorem 1.2** The proof of Theorem 1.2 is based on a KAM iteration, which divides into several steps.

**A. Outline of the KAM step.** At each step we consider the following system:
\[
\begin{cases}
\dot{x} = \omega (\xi) + P_1 (x, y; \xi), \\
\dot{y} = (\Omega (\xi) + A (\xi)) y + P_2 (x, y; \xi),
\end{cases}
\]
where \(A = \text{diag}(A_1, A_2, \ldots, A_m)\) with \(A_j\) being a \(d_j \times d_j\) matrix. Suppose \(P_1\) and \(P_2\) are analytic on \(D(s, r)\) in \((x, y)\) and belong to \(C^L (O)\) in \(\xi\). Moreover,
\[
\| \omega - \xi \| \leq 2e_0, \quad \| A \| \leq 2e_0, \quad \| P_1 \|_{D(s, r)} \leq \epsilon, \quad \| P_2 \|_{D(s, r)} \leq \epsilon.
\]

We want to find a transformation near the identity map and change the system \((2.1)\) to
\[
\begin{cases}
\dot{x} = \omega_+ (\xi) + P_{+1} (x, y; \xi), \\
\dot{y} = (\Omega (\xi) + A_+ (\xi)) y + P_{+2} (x, y; \xi),
\end{cases}
\]
where \(P_{+1}\) and \(P_{+2}\) are much smaller perturbations, and \(\omega_+\) and \(A_+\) are the corrections of \(\omega\) and \(A\), respectively.

**B. Construction of a transformation.** Define a transformation \(\Phi : (x_+, y_+) \to (x, y)\) by
\[
x = x_+ + h(x_+), \quad y = (I_{\tilde{m}} + a(x_+)) y_+ + b(x_+),
\]
where \(h(x), b(x)\) are \(n\)-, \(\tilde{m}\)-dimensional vector functions, respectively, and \(a(x)\) is an \(\tilde{m} \times \tilde{m}\)-matrix function with \(\tilde{m} = d_1 + d_2 + \cdots + d_m\). Let \(P_0^1 = P_1 |_{y=0}, \omega_+ = \omega + [P_0^1]\), \(P_2 = P_2 |_{y=0}, P_2 = \partial_y P_2 |_{y=0}\), where \([P_0^1]\) is the average of \(P_0^1\) with respect to \(x\) on \(T^n\). Define an operator \(\partial_{u} \omega\) by \(\partial_u f = \sum k i(k, \omega) f_k e^{i(k, x)}\), where \(f = \sum k f_k e^{i(k, x)}\). If \(f = (f_{ij})\) is a matrix, define \(\partial_u f = (\partial_u f_{ij})\). Let \([P_0^2]\) be the average of \(P_0^2\), where \(p_{ij}\) is a \(d_i \times d_j\) matrix. Let \(A = \text{diag}(p_{11}, p_{22}, \ldots, p_{mm})\) and \(A_+ = A + A\). Under the transformation \(\Phi\), the system \((2.1)\) is changed to
\[
\begin{cases}
\dot{x} = \omega_+ + (I_n + D_x h)^{-1}(P_0^1 - [P_0^1] - \partial_{u} h) + P_{+1}, \\
\dot{y} = (\Omega + A_+) y + (I_{\tilde{m}} + a)^{-1}[(\Omega + A)b - \partial_{u} b + P_2^0] \\
\quad + (\Omega + A) a y - a(\Omega + A) y - \partial_{u} a y + (P_2 - A) y + P_{+2},
\end{cases}
\]
where
\[
P_{+1} = (I_n + D_x h)^{-1} \left( P_1 - P_0^1 - D_x h [P_0^1] + \int_0^1 \langle \nabla_z P_4 (z + t \delta z), \delta z \rangle dt \right)
\]
\[ P_{\tau}^{2} = -(I_{m} + a)^{-1}(\partial P_{\tau} + [P_{0}] ay + \Omega P_{\tau}) + (I_{m} + a)^{-1}(P_{2} - P_{0}^{2} y - aAy) + \int_{0}^{1} \langle \nabla_{z} P_{2}(z + t\delta z), \delta z \rangle dt, \]

where \( D_{z} h \) is the Jacobian matrix of \( h \) with respect to \( x, z = (x, y) \), and \( \delta z = (h, ay + b) \). Note that in (2.4) we have already used the variables \((x, y)\) in place of \((x_{+}, y_{+})\) for simplifying notations.

If we find \( h, a, b \) such that
\[ \begin{align*}
\tag{2.7} -\partial_{\omega} h + P_{1}^{0} - [P_{1}^{0}] &= 0, \\
\tag{2.8} (\Omega + A)b - \partial_{\omega} b + P_{2}^{0} &= 0, \\
\tag{2.9} (\Omega + A)a - a(\Omega + A) - \partial_{\omega} a + P_{1}^{1} - \hat{A} &= 0,
\end{align*} \]

then the system (2.4) becomes (2.3).

C. Solving linear equations. Now we solve the linear equations (2.7), (2.8) and (2.9). Let \( h = \sum_{k} h_{k} e^{i(k,x)} \) and \( P_{1} = \sum_{k} P_{1}^{0} e^{i(k,x)} \). By (2.7) we have
\[ h_{k} = P_{1}^{0}/i(k, \omega), \ k \neq 0. \]

It is easy to see that \( \| P_{1}^{0} \|_{D(s,r)}^{\infty} \leq \epsilon \). By Cauchy’s estimates we have \( \| P_{1}^{0} \|_{L} \leq \epsilon e^{-s|k|} \). There exists a subset \( O_{\tau}^{0} \) of \( O \) such that for \( \xi \in O_{\tau}^{0} \) we have
\[ |\langle \omega(\xi), k \rangle| \geq \frac{\alpha}{|k|^{\tau}}, \ \forall k \neq 0. \]

Thus, for \( \xi \in O_{\tau}^{0} \), we have \( \| h_{k} \|_{L} \leq |\alpha|^{-L-1}|k|^{(1+L)\tau + L + 1} \epsilon \), and so
\[ \| h \|_{D(s, -\rho, r)} \leq \epsilon \alpha^{-L-1} \rho^{-\kappa} \quad \text{with } \kappa = (1 + L)\tau + L + n + 1. \]

Note that \( h \) should be defined on \( O_{\tau}^{0} \). By the Whitney’s extension theorem, we can extend it to the initial domain \( O \). So, for simplicity we do not concern ourselves with the domain of the parameter in all estimates of the KAM steps. However, in our problem it only makes sense for \( \xi \in O_{\tau}^{0} \).

For the equation (2.8), write \( b = (b_{1}, \ldots, b_{m}) \), where \( b_{j} \) is a \( d_{j} \)-dimensional vector. Let \( f = P_{2}(x, 0; \xi) = P_{0}^{2} \). It follows easily that
\[ \| f \|_{D(s, r)}^{L} \leq \| P_{2} \|_{D(s, r)}^{L} \leq r\epsilon. \]

In the same way, write \( f = (f_{1}, \ldots, f_{m}) \). The equation (2.8) is equivalent to
\[ \partial_{\omega} b_{j} - (\Omega_{j} I_{d_{j}} + A_{j})b_{j} = f_{j}, \ j = 1, 2, \ldots, m. \]

Let \( b_{j} = \sum_{k \in \mathbb{Z}^{n}} b_{j,k} e^{i(k,x)} \) and \( f_{j} = \sum_{k \in \mathbb{Z}^{n}} f_{j,k} e^{i(k,x)} \). Then we have
\[ \begin{align*}
\tag{2.11} \left[(i(k, \omega) - \Omega_{j})I_{d_{j}} - A_{j}\right]b_{j,k} &= f_{j,k}.
\end{align*} \]

Denote
\[ E = (i(k, \omega) - \Omega_{j})I_{d_{j}} - A_{j}. \]

By assumption (2.2) and Lemmas [3.1] and [3.2] there is \( K > 0 \) such that for sufficiently small \( \epsilon_{0} \), there exists a subset \( O_{\tau}^{1} \) of \( O \) such that, for all \( \xi \in O_{\tau}^{1} \) and
$|k| \geq K$, the matrix $E$ is non-singular. Moreover, we have

$$
\|E^{-1}\| \leq c(|k| + 1)^{\tau'} \alpha^{-L-1},
$$

where $\tau' = (1 + L)\tau + L$ and $\tau > (n + 1)L$.

Let

$$
O^1 = \{ \xi \mid |i(k, \omega(\xi)) + \Omega_j(\xi)| \geq \alpha_0, \forall |k| \leq K, \forall j = 1, 2, \ldots, m \},
$$

where $\alpha_0 \geq \alpha$. So if $\epsilon_0$ is sufficiently small, for $\xi \in O^1$, $E^{-1}$ exists for all $|k| \leq K$, and the above estimate for $E^{-1}$ also holds.

Thus, $\forall \xi \in O^1_+ \cap O^1$, we can solve the equation (2.8) for $b_{j,k}$ with

$$
\|b_{j,k}\| \leq c(|k| + 1)^{\tau'} \alpha^{-L-1} \|f_{j,k}\|^L.
$$

Therefore, we have

$$
(2.12) \quad \frac{1}{r} \|b\|_{D(s-r)} \leq \frac{c\epsilon}{\alpha \kappa_1^r \rho^r},
$$

where $\kappa = \tau' + n - 1$.

Now we consider the last equation, (2.9). Write $a = (a_{ij})_{1 \leq i, j \leq m}$ and $g = P_2 - \tilde{A} = (g_{ij})_{1 \leq i, j \leq m}$, where $a_{ij}$ and $g_{ij}$ are $d_i \times d_j$-matrices.

Then (2.9) becomes

$$
\partial_\omega a_{ij} + a_{ij}(\Omega_j I_{d_j} + A_j) - (\Omega_i I_{d_i} + A_i)a_{ij} = g_{ij}.
$$

In the same way as above we have

$$
i(k, \omega)a_{ij,k} - (\Omega_i + A_i)a_{ij,k} + g_{ij,k}(\Omega_j + A_j) = g_{ij,k},
$$

where $a_{ij,k}$ and $g_{ij,k}$ are the Fourier coefficients of $a_{ij}$ and $g_{ij}$, respectively. The above matrix equations for $a_{ij,k}$ are equivalent to the following linear equations for the elements of the matrix $a_{ij,k}$:

$$
[(i(k, \omega) + \Omega_i - \Omega_j)I_{d_i,d_j} + O(\epsilon_0)]\{a_{ij,k}\} = \{g_{ij,k}\},
$$

where $\{a_{ij,k}\}$ and $\{g_{ij,k}\}$ are vectors whose components are the elements of the matrices $a_{ij,k}$ and $g_{ij,k}$, respectively, and $O(\epsilon_0)$ is a $d_i d_j$-order matrix depending on $A_i$ and $A_j$. By the assumption (2.2), we have $\|O(\epsilon_0)\| \leq c \epsilon_0$.

Let $E = (i(k, \omega) + \Omega_i - \Omega_j)I_{d_i,d_j} + O(\epsilon_0)$. By the assumption (2.2) and Lemmas 3.1 and 3.2 there is a $K > 0$ such that for $|k| \geq K$, there exists a subset $O^2_+$ of $O$ such that, for $\xi \in O^2_+$, the matrix $E$ is non-singular. Moreover, we have

$$
\|E^{-1}\|^L \leq c(|k| + 1)^{\tau'} \alpha^{-L-1},
$$

where $\tau' = (1 + L)\tau + L$ and $\tau > (n + 1)L$.

Let

$$
O^2 = \{ \xi \mid |i(k, \xi) + \Omega_i(\xi) - \Omega_j(\xi)| \geq \alpha_0, \ i \neq j, \forall |k| \leq K \}.
$$

By analyticity of $\Omega_j(\xi)$ and condition (12.2), if $\alpha_0$ is sufficiently small, the measure of the set $O \setminus O^2$ is small. So for sufficiently small $\epsilon_0$, $E^{-1}$ exists, and the above estimate of $E^{-1}$ still holds.

Thus, for $i \neq j$ or $k \neq 0$ and $\forall \xi \in O^2_+ \cap O^2$, the coefficient matrix of the linear equation system for $\{a_{ij,k}\}$ is non-singular, and so we find $a_{ij,k}$ with

$$
\|a_{ij,k}\| \leq c(|k| + 1)^{\tau'} \alpha^{-L-1} \|g_{ij,k}\|^L.
$$
If \( i = j \) and \( k = 0 \), by the choice of \( \hat{A} \), it follows that \( g_{j,i,0} = 0 \) and so we let \( a_{j,j,0} = 0 \). It follows that
\[
\|a(x)\|_{\mathcal{D}(s-\rho, r)}^L \leq \frac{c\varepsilon}{\alpha L + 1} r^\rho.
\]

Let \( O_+ = O_0^0 \cap O_0^1 \cap O_1^2 \cap O_1 \cap O_2 \). For \( \xi \in O_+ \), we can obtain solutions of the equations (2.7), (2.8) and (2.9), \( h, b \) and \( a \) with the estimates (2.11), (2.12) and (2.13) holding.

By Cauchy’s estimate it follows that
\[
\|D_x h\|_{\mathcal{D}(s-2\rho, r)^L}, \quad \frac{1}{r} \|D_x b\|_{\mathcal{D}(s-2\rho, r)^L}, \quad \|D_x a\|_{\mathcal{D}(s-2\rho, r)^L} \leq \frac{c\varepsilon}{\alpha L + 1} r^{\rho + 1}.
\]

Moreover, by Lemma 3.1, for \( \tau \geq (n + 1)L \), we have
\[
\text{mes}(O \setminus (O_0^0 \cup O_0^1 \cup O_1^2)) \leq c\alpha \frac{1}{\tau} \sum_{k \appa 0} (|k| + 1) \frac{1}{\tau} \leq c\alpha \frac{1}{\tau}.
\]

By analyticity of the normal frequencies (see [8]), it easily follows that
\[
\text{mes}(O \setminus O^1 \cup O^2) \rightarrow 0, \quad \text{as} \quad \alpha \to 0.
\]

By the above discussion, if
\[
\xi \in O_+ \quad \text{and} \quad \frac{c\varepsilon}{\alpha L + 1} r^\rho \leq \eta < \frac{1}{8},
\]
we have a transformation
\[
\Phi : (x, y) \in D(s - \rho, \eta r) \rightarrow (x, y) \in D(s, \rho r) \subset D(s, r)
\]
such that
\[
\begin{align*}
\|x - x_+\|_{\mathcal{D}(s-\rho, r)^L}, & \quad \frac{1}{r} \|y - y_+\|_{\mathcal{D}(s-\rho, r)^L} \leq \frac{c\varepsilon}{\alpha L + 1} r^{\rho + 1}, \\
\|D_x x - I_\alpha\|_{\mathcal{D}(s-2\rho, r)^L}, & \quad \frac{1}{r} \|D_x y\|_{\mathcal{D}(s-2\rho, r)^L} \leq \frac{c\varepsilon}{\alpha L + 1} r^{\rho + 1}, \\
\|D_y y - I_\alpha\|_{\mathcal{D}(s-2\rho, r)^L}. & \quad \leq \frac{c\varepsilon}{\alpha L + 1} r^{\rho + 1}.
\end{align*}
\]

Moreover, it follows easily that
\[
\|\omega_+ - \omega\|_L \leq \epsilon, \quad \|A_+ - A\|_L \leq \epsilon.
\]

D. Estimates for new error terms. Let
\[
\eta = \left(\frac{\epsilon}{\alpha L + 1} r^{\rho + 1}\right)^{\frac{1}{r + 1}}, \quad r_+ = r\eta, \quad s_+ = s - 3\rho.
\]
We will prove
\[
\|P_+\|_{\mathcal{D}(s, r_+)^L} \leq \epsilon_+, \quad \|P_{+2}\|_{\mathcal{D}(s, r_+)^L} \leq \epsilon_+ + r_+
\]
with \( \epsilon_+ = c\varepsilon\eta \). Since \( P_{+1} \) is simpler, we only consider \( P_{+2} \).

By (2.13), we suppose
\[
\|a\|_{\mathcal{D}(s-\rho, r)^L} \leq c\varepsilon / \alpha L^1 + 1 \rho^\rho \leq \frac{1}{2}.
\]

By the Neumann series, it follows that \( (I_\alpha + a)^{-1} \) exists and we have
\[
\|((I_\alpha + a)^{-1})^L\|_{\mathcal{D}(s-\rho, r)^L} \leq 1 + \|a\|_{\mathcal{D}(s-\rho, r)^L} + (\|a\|_{\mathcal{D}(s-\rho, r)^L})^2 + \cdots \leq 2.
\]

Since \( \|ay + b\|_{\mathcal{D}(s-\rho, r)^L} \leq c\varepsilon \rho \), so
\[
\|\partial_{P_{+1} + [P_{+}]}(ay + b)\|_{\mathcal{D}(s-3\rho, r)^L} \leq c\varepsilon^2 \rho / \alpha L^1 + 1 \rho^\rho + 1.
\]
Again,
\[ \| P_2 - P_0^0 - P_2^1 y \|_{D(s, \eta r)}^2 \leq \eta^2 \| P_2 - P_0^0 \|_{D(s, r)}^2 \leq \eta^2 c \epsilon \]
and
\[ \| \Delta y \|_{D(s, \eta r)}^2 \leq \| \Delta y \|_{D(s, \eta r)}^2 \leq c \epsilon^2 \eta r / \alpha \rho^k. \]
Moreover, \( \forall t \in [0, 1] \) we have
\[ \| \langle \nabla z P_2(z + t \delta z), \delta z \rangle \|_{D(s-3\rho, \eta r)}^2 \leq c \epsilon r / \alpha \rho^k. \]
Thus, by (2.20) and combining all the above estimates we have
\[ \| P_{\nu+2} \|_{D(s+\epsilon r)}^2 \leq c \epsilon^2 \epsilon r = \epsilon_+ r_+. \]
Thus, the second inequality of (2.18) holds.

E. Convergence of iteration. Below we make KAM iteration. For given \( s, \epsilon, r \) in the theorem, we define some sequences inductively:
\[ \epsilon_0 = \epsilon,\quad r_0 = r,\quad s_0 = s,\quad \alpha_0 = \alpha,\quad \rho_0 = \frac{s_0}{12}, \quad \eta_0 = \frac{\epsilon_0}{\alpha \rho^{k+1}}, \]
\[ \epsilon_{\nu+1} = c \eta_\nu \epsilon_\nu,\quad r_{\nu+1} = \eta_\nu r_\nu,\quad s_{\nu+1} = s_\nu - 3 \rho_\nu,\quad \rho_{\nu+1} = \frac{1}{2} \rho_\nu, \]
\[ \alpha_{\nu+1} = \frac{\alpha_\nu}{2},\quad \eta_\nu = \left( \frac{c \rho^{k+1}}{\alpha \rho^{k+1}} \right)^{1/2},\quad D_{\nu} = D(s_\nu, r_\nu),\quad \nu = 0, 1, \ldots. \]

At the \( \nu \)-th step, we have a subset \( O_\nu \subset O \). For \( \xi \in O_\nu \), there exist transformations \( \Phi_\nu : (x_{\nu+1}, y_{\nu+1}) \rightarrow (x_{\nu}, y_{\nu}) \), satisfying
\[ \| x_{\nu+1} - x_\nu \|_{D_\nu}^2 \leq c \epsilon_\nu^2,\quad \| y_{\nu+1} - x_\nu \|_{D_\nu}^2 \leq c \epsilon_\nu^2. \]
Let \( \Phi_{\nu+1} = \Phi_\nu \circ \Phi_{\nu+1} \) with \( \Phi_0 = Id \). Let \( P_0^0 = P_1, P_0^1 = P_2, \omega^0 = \xi, A^0 = 0. \) Then by the transformations \( \Phi^\nu \) the system (1.1) is changed to
\[ \dot{x} = \omega^\nu(\xi) + P_1^\nu(x, y; \xi),\quad \dot{y} = \Omega(\xi) + A^\nu(\xi) y + P_2^\nu(x, y; \xi). \]
Moreover, we have
\[ \| \omega^{\nu+1} - \omega^\nu \|_{L^2} \leq \epsilon_\nu,\quad \| A^{\nu+1} - A^\nu \|_{L^2} \leq \epsilon_\nu, \]
\[ \| P_1^{\nu+1} \|_{D_\nu} \leq \epsilon_\nu,\quad \| P_2^{\nu+1} \|_{D_\nu} \leq r_\nu \epsilon_\nu. \]

By the definitions of \( \eta_\nu, \epsilon_\nu, \rho_\nu \), we have \( \eta_{\nu+1} \leq c \eta_\nu \), and so \( c \eta_{\nu+1} \leq c \eta_\nu \). If \( c \eta_0 \leq 2^{-1} \), then \( \eta_{\nu+1} \leq c^{-1} 2^{-1} \). Thus, we have \( \epsilon_{\nu+1} \leq 2^{-1} \epsilon_\nu \), and so \( \epsilon_\nu \leq 2^{-\nu} \epsilon_0 \). It is easy to see that the assumption (2.2) holds by (2.22).

Let \( O_\nu = \bigcap_{\nu \geq 1} O_\nu \). By the above estimates we can prove that \( \Phi^\nu \) is convergent to a transformation \( \Phi_\ast \) for \( (x, y; \xi) \in D_s(s/2, r/2) \times O_\ast \). The proof is the same as in the case of hamiltonian systems and reversible systems, so we omit the details and refer the reader to [9, 10]. Thus, it is easy to see that for \( \nu \rightarrow \infty \) we get \( \omega^\ast = \omega^\ast, A^\ast = A^\ast, P_1^\ast = P_1^\ast, P_2^\ast = P_2^\ast \). It follows easily that \( \| \omega^\ast - \xi \|_{L^2} \leq 2 \epsilon_0,\quad \| A^\ast - \xi \|_{L^2} \leq 2 \epsilon_0 \). By (2.22) we have
\[ P_1^\ast(x, y; \xi) |_{y=0} = 0,\quad P_2^\ast(x, y; \xi) |_{y=0} = 0,\quad \text{and} \quad \partial_y P_2^\ast(x, y; \xi) |_{y=0} = 0. \]

Thus, for \( \xi \in O_\ast \), by the transformation \( \Phi_\ast \) the system (1.1) is changed to the form of (1.5).
F. Measure estimate. Now we estimate the measure of the set $O_*$ to prove it is non-empty. By the above discussion, $O_\nu = O_0^\nu \cap O_1^\nu \cap O_2^\nu \cap O_1 \cap O^2$. Let $O_* = \bigcap_{\nu \geq 1} O_\nu$. Then $O_* = O_0^* \cap O_1^* \cap O_2^* \cap O_1 \cap O^2$. By (2.11), we have $	ext{mes}(O \setminus (O_0^0 \cup O_1^1 \cup O_2^2)) \leq c(\alpha_0)^{\frac{1}{2}}$. It follows that

$$
\text{mes}(O \setminus O_*) \leq \text{mes}(O \setminus O_1) + \text{mes}(O \setminus O_2) + c\alpha \to 0 \quad \text{(as } \alpha \to 0) \).
$$

3. Appendix

In this section, we give some lemmas, which are used to control the measure of parameters such that the equations (2.7), (2.8) and (2.9) are solvable. These results are actually Lemmas A.1 and A.2 in [13], so we refer to the reference for their proofs.

Lemma 3.1. Suppose that $\|\omega(\xi) - \xi\|^L \leq \epsilon$ and $f_j(\xi)$ are $L$-th continuously differentiable functions with $\|f_j(\xi)\|^L \leq M$ ($1 \leq j \leq L$). Let $P(\xi) = (p_{ij}(\xi))$ be an $L \times L$-matrix, and $L$-th continuously differentiable with respect to $\xi \in O$ with $\|P\|^L = \max_{1 \leq \xi \leq L} \sup_{\xi \in O} \left| \frac{\partial^L p_{ij}}{\partial \xi^L} \right| \leq \epsilon$. Let $D = (\omega, k) I + \text{diag}(f_1(\xi), f_2(\xi), \ldots, f_L(\xi)) + P(\xi)$ with $I$ being the unit matrix, and let $R_k(\alpha)$ be the subset of $O$ such that $\|D^{-1}\| > \frac{\alpha^L}{\epsilon}$, $\tau > nL$. Then, there is a sufficiently small $\epsilon_0 > 0$ and a sufficiently large positive integer $K$ only depending on $M$ such that, if $\epsilon \leq \epsilon_0$, for $\alpha > 0$ and $|k| \geq K$, we have

$$
\text{mes}(R_k(\alpha)) \leq \left( \frac{c\alpha}{|k|^\tau-L} \right)^{\frac{1}{2}},
$$

where $c$ is a constant depending on $\epsilon_0$ and $M$.

Lemma 3.2. Let $D$ be a matrix depending on $\xi$ and $\|D\|^L \leq M$. If $D$ is invertible with $\|D^{-1}\| \leq N$, then $\|D^{-1}\|^L \leq cM^L N^{L+1}$, where $c$ is a constant depending on $L$.

References


Department of Mathematics, Southeast University, Nanjing 210096, People’s Republic of China
E-mail address: xujun@seu.edu.cn