J. Nonlinear Sci. Vol. 10: pp. 49–68 (2000)
DOI: 10.1007/s003329910003

Journal of Nonlinear Science

KAM-Type Theorem on Resonant Surfaces for Nearly Integrable Hamiltonian Systems

F. Cong,[1] T. Küpper,[2] Y. Li,[1,*] and J. You[3]

[1] Department of Mathematics, Jilin University, Changchun 130023, People's Republic of China

[2] Mathematisches Institut, Universität Köln, D-50931 Köln, Germany

[3] Department of Mathematics, Nanjing University, Nanjing 210093, People's Republic of China

Received December 18, 1997; revised December 30, 1998; accepted June 21, 1999
Communicated by Stephen Wiggins

Summary. In this paper, we consider analytic perturbations of an integrable Hamiltonian system in a given resonant surface. It is proved that, for most frequencies on the resonant surface, the resonant torus foliated by nonresonant lower dimensional tori is not destroyed completely and that there are some lower dimensional tori which survive the perturbation if the Hamiltonian satisfies a certain nondegenerate condition. The surviving tori might be elliptic, hyperbolic, or of mixed type. This shows that there are many orbits in the resonant zone which are regular as in the case of integrable systems. This behavior might serve as an obstacle to Arnold diffusion. The persistence of hyperbolic lower dimensional tori has been considered by many authors [5], [6], [15], [16], mainly for multiplicity one resonant case. To deal with the mechanisms of the destruction of the resonant tori of higher multiplicity into nonhyperbolic lower dimensional tori, we have to deal with some small coefficient matrices that are the generalization of small divisors.

Key words. Hamiltonian systems, resonant invariant tori, KAM-type theorem

MSC numbers. 58F05, 58F27, 58F30

1. Introduction

Consider a Hamiltonian system

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

* Supported partially by the Heinrich-Hertz-Stiftung in Germany and the grant of NSF-China.

where $y \in G \subset R^n$, $x \in T^n$, H_0, and P are real analytic functions defined on a complex neighborhood of the closed bounded region G and the torus $T^n(= R^n/2\pi Z^n)$, respectively; H_0 satisfies the standard nondegeneracy condition det $\frac{\partial^2 H_0}{\partial y^2}(y) \neq 0$ in G. P is a perturbation and $\varepsilon > 0$ is a small parameter.

For the unperturbed Hamiltonian $H_0(y)$, $\omega = \frac{\partial H_0}{\partial y}(y)$ is called nonresonant if it satisfies $\langle k, \omega\rangle \neq 0$ for any $k \in Z^n\backslash 0$. Otherwise it is resonant. ω is called a multiplicity m_0 resonant frequency if there is a rank m_0 subgroup g of Z^n generated by independent integer vectors $\tau_1, \dots, \tau_{m_0}$ such that $\langle k, \omega\rangle = 0$ for all $k \in g$ and $\langle k, \omega\rangle \neq 0$ for all $k \in Z^n/g$.

For any given subgroup g,

$$O(g, G) = \{y \in G\colon \langle k, \omega(y)\rangle = 0, k \in g\}$$

is an $m = n - m_0$ dimensional surface, which is called a g-resonant surface. Locally it is diffeomorphic to R^{n-m_0}. In a typical way (see [1]), by passing to a finite covering which will also lead to the global result on G, we may assume that $O(g, G)$ is globally diffeomorphic to a subdomain in R^{n-m_0} without loss of generality.

For the trivial subgroup $g = 0$, according to the celebrated KAM (Kolmogorov-Arnold-Moser) theory (see [1], [8], [10]), most of the nonresonant tori of the integrable system persist under a small perturbation. This paper deals with the perturbation of the resonant tori. More precisely, for a given $O(g, G)$ of multiplicity $m_0 > 0$, we will investigate what happens to the resonant torus of the unperturbed system with frequency $\frac{\partial H_0}{\partial y}(y)$ for $y \in O(g, G)$ under a small perturbation. The perturbation of the resonant tori is more complicated. In general, it will be destroyed [11] by the perturbation. Note that if $y \in O(g, G)$, the invariant torus of H_0 with frequency $\frac{\partial H_0}{\partial y}(y)$ is foliated by $m = n - m_0$ dimensional tori. We will prove that, *for most of $y \in O(g, G)$ in the measure sense*, there are some lower dimensional tori on the resonant torus which survive *general* perturbations.

We will work out a KAM-type theorem in a resonant surface $O(g, G)$ for any given subgroup g of Z^n. We first set up the problem. Similar to [16], by group theory, there are integer vectors $\tau_1', \dots, \tau_m' \in Z^n$ such that Z^n is generated by $\tau_1, \dots, \tau_{m_0}, \tau_1', \dots, \tau_m'$ and $\det(K_0) = 1$, where $K_0 = (K_*, K')$, $K_* = (\tau_1', \dots, \tau_m')$, $K' = (\tau_1, \dots, \tau_{m_0})$ are $n \times n$, $n \times m$, and $n \times m_0$, respectively.

We say H_0 is g-nondegenerate if H_0 is nondegenerate and det $K'^T \frac{\partial^2 H_0}{\partial y^2}(y) K' \neq 0$ for $y \in O(g, G)$.

Since $P(x, y)$ is a real analytic function defined on some complex neighborhood of $T^n \times G$, using Fourier's expansion yields

$$P(x, y) = \sum_{k\in Z^n} P_k e^{\sqrt{-1}\langle k, x\rangle}.$$

For the subgroup g of Z^n, let

$$h_0(\varphi, y) = \sum_{k\in g} P_k e^{\sqrt{-1}\langle k, x\rangle} = \sum_{l\in Z^{m_0}} P_{K'l} e^{\sqrt{-1}\langle l, \varphi\rangle}, \tag{1.2}$$

where $\varphi = K'^T x$. Clearly, h_0 has at least $m_0 + 1$ critical points on T^{m_0}. Moreover, there are at least 2^{m_0} critical points if all of them are nondegenerate (see [9]).

Let φ_0 be a nondegenerated critical point of $h_0(\varphi, y)$, i.e., $\frac{\partial h_0}{\partial \varphi}(\varphi_0, y) = 0$, and $\frac{\partial^2 h_0}{\partial \varphi^2}(\varphi_0, y)$ is nonsingular.

Treshchev [16] proved that for any $y_0 \in O(g, G)$, there is an m-dimensional torus on the resonant torus which persists and only undergoes a small deformation if $\omega^* = K_*^T \omega(y_0)$ as an m-vector satisfies the Diophantine condition and no eigenvalue of $\Pi K'^T \frac{\partial^2 H_0}{\partial y^2}(y_0) K'$ is positive or zero (for simplicity, the later condition will be called Treshchev's hyperbolicity condition), where $\Pi = \frac{\partial^2 h_0}{\partial \varphi^2}(\varphi_0, y_0)$. Eliasson [6], Chierchia and Gallavotti [5], and Rudnev and Wiggins [15] also obtained a similar result for the multiplicity one resonant case. Actually, Treshchev's condition implies that the Hamiltonian equations of motion with Hamiltonian $H_0(y) + \varepsilon h_0(y, \varphi)$ have a hyperbolic m-dimensional torus for any $y \in O(g, G)$ close to y_0. By a series of symplectic changes of variables, he reduced (1.1) to the following Graff's form [7], with a small parameter:

$$H = \langle \omega^*, y\rangle + \frac{\varepsilon}{2}\langle y, \Gamma y\rangle + \varepsilon\langle z_-, \Omega_0(x, y) z_+\rangle + \varepsilon O(|z|^3), \qquad z_+, z_- \in C^{m_0},$$

where $x \in T^m$, $y \in R^m$, Γ is a nonsingular symmetric matrix, and $\mathrm{Re}\langle \gamma, \Omega\bar{\gamma}\rangle \geq \sigma|\gamma|^2$, for some $\sigma > 0$ and for all $\gamma \in C^{m_0}$. By a modified Graff's iteration scheme, he proved that there exists a canonical change of variables Φ such that

$$H \circ \Phi(X, Y, Z) = \langle \omega^*, Y\rangle + \frac{\varepsilon}{2}\langle Y, \tilde{\Gamma} Y\rangle + \varepsilon\langle Z_-, \tilde{\Omega}(X, Y) Z_+\rangle + \varepsilon O((|Y| + |Z|)^3),$$

where

$$\tilde{\Gamma} = \Gamma + O(\varepsilon), \qquad \tilde{\Omega} = \omega_0 + O(\varepsilon), \qquad Z = (Z_-, Z_+).$$

Hence the perturbed system under consideration admits a hyperbolic invariant torus with the same frequency ω^*.

Eliasson [6] considered the following case:

(C1) $g = \{l k_0 \colon l \in \mathbf{Z}\}$ with $\langle k_0, \omega\rangle = 0$, for some $k_0 \in Z^n \backslash 0$,
(C2) $|\langle k, \omega\rangle| \geq \frac{\gamma}{|k|^\tau}$, for $k \in Z^n \backslash g$,
(C3) $k_0^\top \frac{\partial^2 H_0}{\partial y^2}(y_0) k_0 > 0$,

where $\gamma, \tau > 0$. (C1) corresponds to the multiplicity one resonance; the frequency ω with (C2) is called the relative Diophantine to the group g; (C3) represents the g-nondegeneracy of H_0 at y_0. Indeed, at this situation, $K' = k_0$, and hence

$$\det K'^\top \frac{\partial^2 H_0}{\partial y^2}(y_0) K' = K'^\top \frac{\partial^2 H_0}{\partial y^2}(y_0) K' = k_0^\top \frac{\partial^2 H_0}{\partial y^2}(y_0) k_0 > 0.$$

Under his hyperbolicity assumption of the perturbation $P(x, 0)$, Eliasson deduced the persistence problem to the perturbation of the following integrable system:

$$\langle K_*^\top \omega(y_0), y'\rangle + \frac{1}{2}\beta y_1^2 - \frac{1}{2}\varepsilon a x_1^2 + O(y_1^3) + \varepsilon O(x_1^3) + \frac{1}{2}\langle y', M y'\rangle,$$

where $a > 0, \beta = k_0^\top \frac{\partial^2 H_0}{\partial y^2}(y_0) k_0$. Note $\Pi = -\varepsilon a$, and hence $\Pi K' \frac{\partial^2 H_0}{\partial y^2}(y_0) K' = -\varepsilon a \beta < 0$. This shows that, locally, Eliasson's result coincides with Treshchev's.

These approaches do not work when the torus is not hyperbolic. Here we consider the system on a whole resonant surface and prove the persistence result in a measure sense.

We will prove that, for general nondegenerate perturbations, there is an m-dimensional torus born from the resonant torus no matter if Treshchev's hyperbolicity condition holds or not. Certainly, in this case, the obtained torus might be elliptic, hyperbolic, or of mixed type.

The main result of this paper is the following:

Theorem 1. *Suppose that H is analytic. Moreover, H_0 is g-nondegenerate for a given g, and $h_0(\varphi, y)$ has an analytic family of nondegenerate critical points for all $y \in O(g, G)$. Then there is an $\varepsilon_0 > 0$ (depending on H_0, g, h_0) and a Cantor set $\Lambda_* \subset O(g, G)$ such that for $0 < \varepsilon < \varepsilon_0$, the system (1.1) admits a smooth family (in Whitney's sense) of m-dimensional invariant tori I_{y_0} parametrized by $y_0 \in \Lambda_*$. Moreover, the measure of Λ_* relative to $O(g, G)$ tends to 1 as $\varepsilon \to 0$.*

Here a map defined on a Cantor set is said to be smooth in Whitney's sense if its Whitney extension is smooth. For details, see [13].

Remark 1. It is well-known that for a nearly integrable Hamiltonian system with many degrees of freedom, the dynamical behavior (for example Arnold diffusion) of nearly integrable systems in the resonant zone is very complicated if the stability of orbits is destroyed. Hence it is important to study the mechanisms which lead to the destruction of resonant tori under perturbations. Theorem 1 provides some further description of those mechanisms, which shows that many orbits in the resonant zone are still quite as regular as in the integrable case for nondegenerate (typical) small perturbations. Those orbits might be hyperbolic, linearly stable, or of mixed type, which will influence the Arnold diffusion.

Since Treshchev's hyperbolicity condition is dropped in Theorem 1, the following stronger result is an immediate consequence, which can be compared with Poincaré's famous theorem [12] (See also [3], page 105).

Theorem 2. *Under the assumptions of Theorem 1, if all critical points of $h_0(\varphi, y)$ are nondegenerate, there is an $\varepsilon_0 > 0$ (depending on H_0, g, h_0) and a Cantor set $\Lambda'_* \subset O(g, G)$ such that for $0 < \varepsilon < \varepsilon_0$, the system (1.1) admits 2^{m_0} smooth families of m-dimensional invariant tori parametrized by $y_0 \in \Lambda'_*$. Moreover, the measure of Λ'_* relative to $O(g, G)$ tends to 1 as $\varepsilon \to 0$.*

Remark 2. Recall Poincaré's theorem on the resonant torus foliated by periodic solutions, i.e., *rank* $g = n - 1$, $m_0 = 1$ case. There it is proved that the perturbed system has at least two periodic solutions if all critical points of $h_0(\varphi, y)$ are nondegenerate. Theorem 2 can be regarded as a generalization of Poincaré's theorem to the invariant tori case on a Cantor set. Actually, Theorem 1 provides a positive answer to a conjecture about a higher dimensional version of Poincaré's theorem in [3].

Remark 3. Cheng [4] considered the multiplicity one resonant case and proved that there is an $n-1$ dimensional torus born from each resonant torus under a convexity condition of the unperturbed system without imposing any restriction on the perturbation. It seems that his approach does not work for higher resonant cases since it strongly depends on the restriction of the dimension. It is believed that for the resonant case of higher multiplicity, Cheng's result is also true, i.e., the nondegeneracy condition of the critical points is not essential although it has not been proved so far. Recently, Wang [17] discussed general cases and gave a persistence theorem similar to Theorem 1 under some additional assumptions.

After reduction to a suitable normal form at the nondegenerate critical point, Theorem 1 is obtained as a consequence of the following theorem for the special Hamiltonian system,

$$H = \langle \omega, y \rangle + \frac{\delta}{2}\langle z, M(\omega)z \rangle + P(x, y, z), \tag{1.3}$$

defined on the complex neighborhood $D(r,s) = \{(x, y, z, \omega) \mid |\operatorname{Im} x| < r, |y| < s^2, |z| < s, \omega \in \mathcal{O}\}$ of $T^m \times \{0\} \times \{0\} \times \mathcal{O} \subset T^m \times R^m \times R^{2m_0} \times R^m$; where $\mathcal{O} \subset R^m$ is a bounded closed region with positive measure, M is a symmetric matrix smoothly depending on ω. For system (1.3) we prove

Theorem 3. *Suppose that P is real analytic on some complex neighborhood $D(r,s)$ of the phase space $T^m \times \{0\} \times \{0\}$ and $\mathcal{O}$ with $|\det M(\omega)| \geq d > 0$. Then for a given parameter γ, there are sufficiently small μ_0, δ_0 such that if $\delta \leq \delta_0$, $\mu \leq \mu_0$, and*

$$|P| < s^2 \gamma^{4m_0^2} \delta\mu,$$

there exist a Cantor set $\mathcal{O}_* \subset \mathcal{O}$, a (Whitney) smooth family of symplectic changes

$$\Phi\colon D\left(r, \frac{s}{2}\right) \to D(r,s),$$

and a smooth map $\omega_\infty\colon \mathcal{O}_* \to R^m$, such that

$$H \circ \Phi = \langle \omega_\infty(\omega), y \rangle + \frac{\delta}{2}\langle z, M_*(\omega)z \rangle + P_*(x, y, z, \omega),$$

with

$$\partial_y^l \partial_z^p P_*|_{(y,z)=(0.0)} = 0,$$

for $|l| + |p| \leq 2$, where $\omega_\infty - id = O(\mu)$, $M_* - M = O(\mu)$. Thus for each $\omega \in \mathcal{O}_*$, the perturbed system (1.3) admits an invariant torus with frequency $\omega_\infty(\omega)$. Moreover, $|\mathcal{O} \backslash \mathcal{O}_*| = O(\gamma)$.

Remark 4. In Treshchev's and Eliasson's cases,

$$M = \begin{pmatrix} 0 & \Omega_0 \\ \Omega_0^\top & 0 \end{pmatrix}.$$

Hence

$$JM = \begin{pmatrix} \Omega_0^\top & 0 \\ 0 & -\Omega_0 \end{pmatrix},$$

where $J = \left(\begin{smallmatrix} 0 & -I_{m_0} \\ I_{m_0} & 0 \end{smallmatrix}\right)$, I_s denotes the unity matrix of order s. Then for each eigenvalue λ of JM,

$$|\mathrm{Re}\,\lambda| \geq \sigma.$$

However, in Theorem 3, we only require that M is nonsingular, and hence some eigenvalues of JM might be purely imaginary and multiple. This shows that Theorem 1 can conclude the existence of some nonhyperbolic invariant tori. Hence, Theorem 1 generalizes Treshchev's and Eliasson's results.

On the other hand, for a general system without small parameter δ,

$$N = \langle \omega, y \rangle + \frac{1}{2}\langle z, M_1 z \rangle, \tag{1.4}$$

if by a smooth symplectic change of variables

$$z = \begin{pmatrix} u \\ v \end{pmatrix} = \Phi \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \Phi(z_1),$$

one has

$$N \circ \Phi = \langle \omega, y \rangle + \frac{\delta}{2}\langle z_1, M_2 z_1 \rangle + O(|z|^3),$$

where $\delta M_2 = \partial\Phi(0)^\top M_1 \partial\Phi(0)$, then

$$\det M_2 = \frac{1}{\delta^{2m_0}} \det(\partial\Phi(0)^\top \partial\Phi(0)) \det M_1 = \frac{1}{\delta^{2m_0}} \det M_1, \tag{1.5}$$

since $\partial\Phi(0)^\top J \partial\Phi(0) = J$ implies $\det(\partial\Phi(0)^\top \partial\Phi(0)) = 1$. (1.5) shows that, generally, one cannot deduce (1.4) to the problem considered in Theorem 3. It is well-known that if δ is not small, say $\delta = 1$, one has to require more restrictions on the frequencies such as

$$|\langle k, \omega \rangle + \langle l, \Omega \rangle| \neq 0, \qquad |l| \leq 2,$$

for $\langle z, M(\omega) z \rangle = \sum_{i=1}^{m_0} \Omega_i(\omega) z_i z_{-i}$ (see Pöschel [14]), and

$$|\det(\sqrt{-1}\langle k, \omega \rangle I_{2m_0} + MJ)| \neq 0,$$
$$|\det(\sqrt{-1}\langle k, \omega \rangle I_{4m_0^2} + (MJ) \otimes I_{2m_0} - I_{2m_0} \otimes (JM))| \neq 0,$$

for general M (see You [19]). For sufficiently small δ, Theorem 3 does not need the above conditions, and thus it has a special advantage when $\frac{\delta}{2}\langle z, M(\omega) z \rangle$ in (1.3) arises from the perturbation. Since M may be nondiagonal and in general, since there is no symplectic transformation depending smoothly on ω which turns $M(\omega)$ in (1.3) into a diagonal one unless JM has distinct eigenvalues, Bourgain's theorem [2] also cannot be directly applied here.

As is well known, the set of all $2m_0 \times 2m_0$ symmetric matrices with distinct eigenvalues is open and dense in the space of all $2m_0 \times 2m_0$ symmetric matrices. However, in our problem, we do not know if the set $\mathcal{M} = \{\omega \in \mathcal{O}\colon\ M(\omega)$ has distinct eigenvalues$\}$ is nonempty. It is possible that $\mathcal{M}$ has a zero measure. Even if $\mathcal{M}$ has a positive measure, one cannot use perturbation technique to study such systems in the KAM theory. Actually, many systems in physical science, such as decoupled oscillators, admit such multiplicity, and this kind of multiplicity has been one of the difficult problems in the perturbation theory (see [2]).

Remark 5. To obtain the persistence of lower dimensional tori on the resonant torus, we will separate the resonant frequency into two parts: the relative diophantine one and the complementary one. This process is similar to Treshchev's. In that way, we can reduce (1.1) to the form (1.3). Note that the part related to the normal variables is not of the usual diagonal form, and that there is no symplectic change that transforms the nondiagonal form into the diagonal one. For that reason, some new linearized equations and small divisor conditions similar to [19] have to be taken into account.

Remark 6. We give an example to illustrate our result in an explicit way. Consider an analytic Hamiltonian

$$H = \frac{1}{2}\sum_{i=1}^{n} y_i^2 + \varepsilon F(x),$$

in $B(1,0) \times T^n$ where $B(1,0)$ is a unit ball in R^n centered at the origin with radius 1. Let g be the subgroup generated by $\tau_1' = (0, 0, \ldots, 0, 1), \ldots, \tau_{m_0}' = (0, \ldots, 0, 1, 0, \ldots, 0)$. Then

$$O(g, G) = \left\{(y, 0) \in B(1, 0), y \in R^{n-m_0}, |y| \le 1\right\},$$

which can be regarded as a unit ball in R^{n-m_0}. Suppose that $\bar{F}(x_{n-m_0+1}, \ldots, x_n) = \int_0^{2\pi} F(x)dx_1 \ldots dx_{n-m_0}$ has l nondegenerate critical points. By applying Theorem 1, we know H has at least l Cantor families of invariant tori of dimension $n - m_0$ parametrized by y' in a Cantor set. Moreover, treating the Cantor set as a set in the unit ball of R^{n-m_0}, it has positive measure which tends to the full measure as $\varepsilon \to 0$.

We note that, by Treshchev's result, one can obtain the persistence of a lower dimensional torus only if $\frac{1}{2}\sum_{i=n-m_0+1}^{n} y_i^2 + \bar{F}$ has a nondegenerate *hyperbolic* critical point. Generally there are some other types of critical points (see [9]). By our Theorem 2, we can get at least 2^{m_0} tori if all critical points of $\bar{F}$ are nondegenerate.

Let us outline the proof of the theorems. In Section 2, we reduce (1.1) to the normal form. This process is similar to [16]. After the reduction, Theorem 1 is a consequence of Theorem 3. A detailed proof of Theorem 3 is given in Section 3 by a KAM-type iteration. In Section 3, we describe one cycle of KAM steps. Section 4 provides an iteration lemma, which shows the validity of each step. Finally in Section 5, we give the proof of Theorem 3. Then we focus our attention on the measure estimate of the Cantor set.

2. Reduction to Normal Form

For the given subgroup g in Theorem 1, following Treshchev's transformation technique [16], we reduce (1.1), near a nondegenerate critical point of h_0, into the following normal form:

$$H = N_0 + P_0,$$

where

$$N_0 = \langle \omega, y\rangle + \frac{\varepsilon}{2}(\langle u, V_0 u\rangle + \langle v, U_0 v\rangle), \qquad |P_0| = O(\varepsilon^2),$$

and $x \in T^m$, $y \in R^m$, $u, v \in R^{m_0}$, ω varies in some m-dimensional subset with positive measure, V_0, U_0 are nonsingular matrices depending smoothly on ω. Then we will show that Theorem 3 implies Theorem 1.

Let

$$\Gamma = K_0^T \frac{\partial^2 H_0}{\partial y^2}(y_0) K_0 = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix},$$

where $\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22}$ are $m \times m$, $m \times m_0$, $m_0 \times m$, $m_0 \times m_0$ matrices, respectively, and $\Gamma_{12} = \Gamma_{21}^T$, $\Gamma_{22} = K'^T \frac{\partial^2 H_0}{\partial y^2}(y_0) K' (\equiv \hat{\Gamma})$. For any $y_0 \in O(g, G)$, by Taylor's formula, we expand the Hamiltonian (1.1) into the following form:

$$H(x, y) = \langle \omega(y_0), y - y_0\rangle + \frac{1}{2}\left\langle \frac{\partial^2 H_0}{\partial y^2}(y_0)(y - y_0), y - y_0\right\rangle + \varepsilon P(x, y) + O(|y - y_0|^3),$$

up to a constant. By the symplectic coordinate transformation $y - y_0 = K_0 p$, $q = K_0^T x$, the above Hamiltonian is changed to

$$\begin{aligned} H(q, p) &= \langle \omega^*, p'\rangle + \frac{1}{2}\langle p, \Gamma(y_0) p\rangle + \varepsilon \bar{P}(q, p) + O(|p|^3) \\ &= \langle \omega^*, p'\rangle + \frac{1}{2}\langle p'', \Gamma_{22} p''\rangle + \varepsilon \bar{P}(q, p) + O(|p|^3) + O(|p'|^2) \\ &\quad + O(|p'| \cdot |p''|), \end{aligned} \tag{2.1}$$

where $\omega^* = K_*^T \omega(y_0)$, $p' = (p_1, \ldots, p_m)^T$, $p'' = (p_{m+1}, \ldots, p_n)^T$,

$$\bar{P}(q, p) = P((K_0^T)^{-1} q, y_0 + K_0 p). \tag{2.2}$$

Denote by $\mathcal{O}(g, G) = \{\omega^* \in R^m : y \in O(g, G)\}$. We know that $\mathcal{O}(g, G)$ is a bounded region in R^m. Since $\mathcal{O}(g, G)$ is diffeomorphic to the m-dimensional surface $O(g, G)$, we will use ω^* as a parameter instead of y_0 in the following. This approach has been used by many authors [1], [14], [15], to simplify the proofs. In the following, we work with

$$\begin{aligned} H(q, p) &= \langle \omega^*, p'\rangle + \frac{1}{2}\langle p'', \Gamma_{22}(\omega^*) p''\rangle + \varepsilon \bar{P}(q, p, \omega^*) \\ &\quad + O(|p|^3) + O(|p'|^2) + O(|p'| \cdot |p''|), \end{aligned} \tag{2.3}$$

where $\omega^* \in \mathcal{O}(g, G)$ serves as a parameter.

Choose $\omega \in \mathcal{O}(g, G)$ such that

$$|\langle k, \omega\rangle| > \gamma_0 |k|^{-\tau}, \qquad \forall 0 \neq k \in Z^m, \tag{2.4}$$

where $|k| = \sum_{i=1}^{m} |k_i|$, and γ_0, τ are fixed positive constants. Denote by $\mathcal{O}'$ the set of ω satisfying (2.4).

For $\omega \in \mathcal{O}'$, we separate the first-order resonant terms from the perturbation by a canonical transformation of coordinates

$$(p, q \bmod 2\pi) \longrightarrow (Y, X \bmod 2\pi)\colon\ p = \frac{\partial S(q, Y)}{\partial q}, X = \frac{\partial S(q, Y)}{\partial Y},$$

where

$$S = \langle Y, q\rangle + \varepsilon \sum_{k \in Z^m \setminus 0} \frac{\sqrt{-1} h_k}{\langle \omega, k\rangle}(q'', \omega) e^{\sqrt{-1}\langle k, q'\rangle},$$

with $h_k = \int_0^{2\pi} \bar{P}(q, 0) e^{\sqrt{-1}\langle k, q'\rangle} dq'$. Then

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

From (2.4) it follows that S is real analytic. The new Hamiltonian function reads as

$$\begin{aligned} H(X, Y) &= \langle \omega, Y'\rangle + \frac{1}{2}\langle Y'', \Gamma_{22}(\omega) Y''\rangle + \varepsilon h_0(X'', \omega) \\ &\quad + O(\varepsilon Y) + O(\varepsilon^2) + O(|Y|^3) + O(|Y'|^2) + O(|Y'||Y''|). \end{aligned} \tag{2.5}$$

We have assumed that $h_0(X'', \omega)$ has a nondegenerate critical point, say X_0''. Without loss of generality, we assume $X_0'' = 0$ up to a linear coordinate transformation. (2.5) is then equivalent to the following:

$$\begin{aligned} H(X, Y) &= \langle \omega, Y'\rangle + \frac{1}{2}\langle Y'', \Gamma_{22}(\omega) Y''\rangle + \frac{1}{2}\varepsilon \left\langle \frac{\partial^2 h_0}{\partial \varphi^2}(0, \omega) X'', X''\right\rangle \\ &\quad + O(\varepsilon Y) + O(\varepsilon^2) + O(|Y|^3) + O(|Y'|^2) \\ &\quad + O(|Y'| \cdot |Y''|) + \varepsilon O(|X''|^3), \end{aligned} \tag{2.6}$$

up to an irrelevant constant $\varepsilon h_0(0)$.

In the next step, we scale the variable Y to reduce some less significant terms to a new perturbation. Take $Y = \varepsilon^{\frac{1}{2}} \bar{Y}$; it arrives at

$$\begin{aligned} H(X, \bar{Y}) &= H(X, \varepsilon^{\frac{1}{2}} \bar{Y})/\varepsilon^{\frac{1}{2}} \\ &= \langle \omega, \bar{Y}'\rangle + \frac{\varepsilon^{\frac{1}{2}}}{2}\left(\left\langle \bar{Y}'', \Gamma_{22}(\omega)\bar{Y}''\right\rangle + \left\langle X'', \frac{\partial^2 h_0}{\partial \varphi^2}(0, \omega) X''\right\rangle\right) \\ &\quad + \varepsilon^{\frac{1}{2}}\Big(\varepsilon^{\frac{1}{2}} O(\bar{Y}) + O(\varepsilon) + \varepsilon^{\frac{1}{2}} O(|\bar{Y}|^3) \\ &\quad + O(|\bar{Y}'|^2) + O(|\bar{Y}'| \cdot |\bar{Y}''|) + O(|X''|^3)\Big). \end{aligned} \tag{2.7}$$

We replace $X', \bar{Y}', X'', \bar{Y}'', \varepsilon^{\frac{1}{2}}, \Gamma_{22}$, and $\frac{\partial^2 h_0}{\partial \varphi^2}(0, \omega)$ by $x, y, u, v, \varepsilon, U_0$, and V_0, respectively. Then

$$H(x, y, u, v) = N_0 + P_0, \tag{2.8}$$

where

$$N_0 = \langle \omega, y \rangle + \frac{\varepsilon}{2}(\langle u, V_0(\omega)u \rangle + \langle v, U_0(\omega)v \rangle), \tag{2.9}$$

and

$$P_0 = O(\varepsilon^2) + \varepsilon O(|y|^2) + \varepsilon O(|y| \cdot |v|) + \varepsilon O(|u|^3). \tag{2.10}$$

This is the desired normal form.

By Whitney's extension theorem, we can assume that P_0 is a smooth function of ω in $\mathcal{O}$, and it coincides with our Hamiltonian only in the Cantor set $\mathcal{O}'(g, G)$.

The Proof of Theorem 1. Now we are in the position to prove that Theorem 3 implies Theorem 1.

For any small ε, let $s = \varepsilon^{\frac{1}{3}}$. We consider the Hamiltonian (2.8) in

$$D(r, s) = \{(x, y, u, v) \colon |\mathrm{Im}\, x| < r, |y| < s^2, |u| + |v| < s\}.$$

It is a special case of (1.3) if we set $z = (u, v)$, $M = \begin{pmatrix} U & \\ & V \end{pmatrix}$ and $\mathcal{O} = \mathcal{O}(g, G)$.

It is easy to see that in (2.8)

$$|P_0| \le C\varepsilon^2, \tag{2.11}$$

on $D(r, s) \times \mathcal{O}$ if ε is sufficiently small. Set $\delta = \varepsilon$, $\gamma = \varepsilon^{\frac{1}{4m_0^2}(\frac{1}{3}-\sigma)}$, $\mu = \varepsilon^\sigma$, $\sigma \in (0, \frac{1}{3})$. Then

$$|P_0| \le \delta \gamma^{4m_0^2} s^2 \mu,$$

on $D(r, s) \times \mathcal{O}$ if ε is sufficiently small. By Theorem 3, the Hamiltonian system with Hamiltonian (2.8) has a family of m-dimensional tori parametrized by $y' \in \mathcal{O}_*(g, G) \subset \mathcal{O}$. Note that (2.8) coincides with our Hamiltonian (2.3) only in the Cantor set $\mathcal{O}'(g, G)$. It follows that the Hamiltonian system with Hamiltonian (2.3) at $\omega \in \mathcal{O}_*(g, G) \cap \mathcal{O}'(g, G) \subset \mathcal{O}(g, G)$ has a rotational torus. It is well-known that $|\mathcal{O} - \mathcal{O}'| \to 0$ as $\gamma \to 0$. Since $\gamma = \varepsilon^{\frac{1}{4m_0^2}(\frac{1}{3}-\sigma)}$, it follows that $|\mathcal{O} - (\mathcal{O}_* \cap \mathcal{O}')| \to 0$ as $\varepsilon \to 0$. Note that $\mathcal{O}(g, G)$ is diffeomorphic to $O(g, G)$. Going back to $O(g, G)$, we have the conclusion of Theorem 1.

Here and later we use $|\cdot|$ to denote norms of vector and functions (sometimes with subscripts), and the measure of sets, and c_i's always denote the constants independent of the iteration process.

In the following, we will give a detailed proof for Theorem 3.

3. The KAM Step

The KAM iteration process consists of infinitely many KAM steps. From each cycle of KAM steps, one can find the constructions and estimates of the desired symplectic changes and their domains, perturbed frequencies, and new perturbations.

We start from Hamiltonian (1.3),

$$H = N + P,$$

defined in $D(r, s)$, where

$$N = e + \langle \omega, y \rangle + \frac{\delta}{2}\langle z, Mz \rangle,$$

with

$$\frac{1}{\delta\gamma^{4m_0^2}s^2}|P| \leq \mu \ll 1.$$

Since there is a detailed description of the KAM steps in [14], we only outline the proof from the ν-step to the $\nu+1$-step. We also refer to [19] for a detailed proof of the case $\delta = 1$ with an additional nondegenerate condition. For simplicity, we omit the index of the ν-step and denote by "$+$" and "$-$" the $\nu+1$-step and $\nu-1$-step respectively. Here we also would like to point out some differences between the usual approach and ours. The linearized equations are not the usual ones. This framework seems to be necessary for the problem under consideration, because the normal form we obtain is not of the norm $\sum \Omega_i z_i z_{-i}$, and in general, it is also impossible to turn it into such a form by a symplectic change of variables. This leads to the use of another nonresonant condition.

3.1. Approximating the Perturbation

Due to the presence of small divisors, one cannot remove all angle-variable-dependent terms in one step. Following the main idea of KAM theory, we will find a symplectic coordinate transformation such that the x-dependent term of the transformed system is much smaller at each KAM step. First, we truncate the perturbation P and keep the higher order terms to the next KAM step since they are already small enough.

Let R be the truncation of P of the form

$$R = \sum_{|k|\leq K_+} (P_{k00} + \langle P_{k10}, y\rangle + \langle P_{k01}, z\rangle + \langle z, P_{k02}z\rangle)e^{\sqrt{-1}\langle k,x\rangle}; \tag{3.1}$$

K_+ will be specified below. Then

$$\begin{aligned} P - R &= \left(\sum_{|k|>K_+} + \sum_{|k|\leq K_+, 2|j|+|q|\geq 3}\right) P_{kjq}e^{\sqrt{-1}\langle k,x\rangle}y^j z^q \\ &= I + II. \end{aligned} \tag{3.2}$$

We estimate $P - R$ on a smaller domain $D(r_+, \alpha s)$, $\alpha = \mu^{\frac{1}{3}} \in (0, 1)$, $r_+ < r$. First,

$$\begin{aligned} |I|_{D(r,s)} &\leq \sum_{|k|>K_+} |P|_{D(r,s)}e^{-|k|(r-r_+)} \leq \delta\gamma^{4m_0^2}s^2\mu \sum_{l>K} l^m e^{-l(r-r_+)} \\ &\leq \delta\gamma^{4m_0^2}s^2\mu \int_{K_+}^{\infty} \lambda^m e^{-\lambda(r-r_+)}\,\mathrm{d}\lambda \leq \delta\gamma^{4m_0^2}s^2\mu^2, \end{aligned} \tag{3.3}$$

provided

$$\int_K^{\infty} \lambda^m e^{-\lambda(r-r_+)}\,\mathrm{d}\lambda \le \mu. \tag{3.4}$$

Hence by (3.3)

$$|P-I|_{D(r,s)} \le |P|_{D(r,s)} + |I|_{D(r,s)} \le 2\delta\gamma^{4m_0^2} s^2 \mu.$$

Second,

$$\begin{aligned}
|II|_{D(r_+,\alpha s)} &= \left| \int \frac{\partial^{|i|+|p|}}{\partial y^i \partial z^p} \sum_{|k|\le K_+, 2|j|+|q|\ge 3} P_{kjq} e^{\sqrt{-1}\langle k,x\rangle} y^j z^q \,\mathrm{d}y\,\mathrm{d}z \right|_{D(r_+,\alpha s)} \\
&= \left| \int \frac{\partial^{|i|+|p|}}{\partial y^i \partial z^p} (P-I)\,\mathrm{d}y\,\mathrm{d}z \right|_{D(r_+,\alpha s)} \\
&\le \left| \int \left| \frac{\partial^{|i|+|p|}}{\partial y^i \partial z^p} (P-I) \right|_{D(r,\frac{s}{2})} \mathrm{d}y\,\mathrm{d}z \right|_{D(r_+,\alpha s)} \\
&\le c_1 \delta \gamma^{4m_0^2} \alpha^3 s^2 \mu = c_1 \delta \gamma^{4m_0^2} s^2 \mu^2,
\end{aligned} \tag{3.5}$$

where $2|i|+|p| = 3$ and $\int = \int_0^y \cdots \int_0^y \int_0^z \cdots \int_0^z$ with $2|i|+|p|$-times. Thus on $D(r_+,\alpha s)$,

$$|P-R| \le c_2 \delta \gamma^{4m_0^2} s^2 \mu^2. \tag{3.6}$$

Thus

$$|R|_{D(r,s)} \le c_3 \delta \gamma^{4m_0^2} s^2 \mu. \tag{3.7}$$

3.2. Linearized Equations

We have to find a Hamiltonian F such that the time 1-map ϕ_F^1 generated by X_F carries H into a new normal form with a smaller perturbation.

Formally we assume F is of the form,

$$F = \sum_{0\ne|k|\le K_+} (F_{k00} + \langle F_{k10}, y\rangle + \langle F_{k01}, z\rangle + \langle z, F_{k02} z\rangle) e^{\sqrt{-1}\langle k,x\rangle} + \langle F_{001}, z\rangle. \tag{3.8}$$

If

$$\{N, F\} + \tilde{R} + \langle P_{001}, z\rangle = 0, \tag{3.9}$$

then

$$\begin{aligned}
H \circ \phi_F^1 &= (N+R)\circ\phi_F^1 + (P-R)\circ\phi_F^1 \\
&= N + [R] - \langle P_{001}, z\rangle + \int_0^1 \{R_t, F\}\circ\phi_F^t\,\mathrm{d}t + (P-R)\circ\phi_F^1 \\
&= N_+ + P_+,
\end{aligned} \tag{3.10}$$

where

$$
\begin{aligned}
[R] &= \int_{T^m} R(x,\cdot)\,\mathrm{d}x, \\
\tilde{R} &= R - [R],\ R_t = (1-t)([R] - R - \langle P_{001}, z\rangle) + R, \\
N_+ &= N + [R] - \langle P_{001}, z\rangle,\ P_+ = \int_0^1 \{R_t, F\}\circ \phi_F^t\,\mathrm{d}t + (P-R)\circ\phi_F^1.
\end{aligned} \tag{3.11}
$$

Putting (3.1) and (3.8) into (3.9) yields

$$
\begin{aligned}
&-\sum_{k\neq 0}\sqrt{-1}\langle k,\omega\rangle(F_{k00} + \langle F_{k10}, y\rangle + \langle F_{k01}, z\rangle + \langle z, F_{k02}z\rangle)e^{\sqrt{-1}\langle k,x\rangle} \\
&\quad + \delta\sum_{k\neq 0}(\langle Mz, JF_{k01}\rangle + 2\langle Mz, JF_{k02}z\rangle)e^{\sqrt{-1}\langle k,x\rangle} + \delta\langle Mz, JF_{001}\rangle \\
&= -\sum_{0\neq|k|\le K_+}(P_{k00} + \langle P_{k10}, y\rangle + \langle P_{k01}, z\rangle + \langle z, P_{k02}z\rangle)e^{\sqrt{-1}\langle k,x\rangle} \\
&\quad - \langle P_{001}, z\rangle.
\end{aligned} \tag{3.12}
$$

Note that we have to solve F_{k02} with $F_{k02} = F_{k02}^{\top}$. Hence comparing coefficients we have

$$\sqrt{-1}\langle k,\omega\rangle F_{k00} = P_{k00}, \tag{3.13}$$

$$\sqrt{-1}\langle k,\omega\rangle F_{k10} = P_{k10}, \tag{3.14}$$

$$-\sqrt{-1}\langle k,\omega\rangle F_{k01} + \delta MJF_{k01} = -P_{k01}, \tag{3.15}$$

$$-\sqrt{-1}\langle k,\omega\rangle F_{k02} + \delta(MJ)F_{k02} - \delta F_{k02}(JM) = -P_{k02}, \tag{3.16}$$

$$\delta MF_{001} = -P_{001}. \tag{3.17}$$

(3.15) is equivalent to

$$\left[-\sqrt{-1}\langle k,\omega\rangle I_{2m_0} + \delta MJ\right]F_{k01} = -P_{k01}, \tag{3.18}$$

and (3.16) is equivalent to

$$\left[-\sqrt{-1}\langle k,\omega\rangle I_{4m_0^2} + \delta(MJ)\otimes I_{2m_0} + \delta I_{2m_0}\otimes(MJ)\right]F_{k02} = -P_{k02}. \tag{3.19}$$

The above linear systems (3.13)–(3.16) are solvable if the coefficient matrices are nonsingular. To control the norm of F, we solve them on the set

$$
\begin{aligned}
\mathcal{O}_+ = \Bigg\{&\omega\in\mathcal{O};\ |\langle k,\omega\rangle| > \frac{\gamma}{|k|^{\tau}},\qquad |\det A_1| > \frac{\gamma^{2m_0}}{|k|^{2\tau m_0}},\ \ |\det A_2| > \frac{\gamma^{4m_0^2}}{|k|^{4\tau m_0^2}}, \\
&\text{for } k\in Z^m \text{ with } 0<|k|\le K_+\Bigg\},
\end{aligned} \tag{3.20}
$$

where A_1 and A_2 denote the coefficient matrices of (3.18) and (3.19) respectively. In the following, we also will use similar notations.

3.3. Coordinate Changes

We will give some estimates of F and its derivatives, which are vital in proving the convergence of the transformation sequence and in estimating the new perturbation at each step. Set

$$D_i = D\left(r_+ + \frac{3}{4}(r - r_+), \frac{i}{4}s\right), \quad i = 1, 2, 3, 4.$$

By (3.13), (3.14), and Cauchy's estimate, we have

$$\begin{aligned} |F_{k00}| &\le \frac{|k|^\tau}{\gamma}\delta\gamma^{4m_0^2}s^2\mu e^{-|k|r} \le |k|^\tau\delta s^2\mu e^{-|k|r}, \\ |F_{k10}| &\le \frac{|k|^\tau}{\gamma}\delta\gamma^{4m_0^2}\mu e^{-|k|r} \le |k|^\tau\delta\mu e^{-|k|r}. \end{aligned}$$

From (3.18) and (3.19), it follows that on $\mathcal{O}(K_+)$,

$$|F_{k01}| \le c_4|k|^{2\tau m_0}\delta s\mu e^{-|k|r}, \qquad \|F_{k02}\| \le c_4|k|^{4\tau m_0^2}\delta\mu e^{-|k|r},$$

where $\|F_{k02}\|$ is defined to be the maximum of the norm of the entries. By (3.17),

$$|F_{001}| \le c_5 s\mu.$$

From (3.8) and the above estimates together with $\mu, \delta, \gamma \ll 1$, we obtain

$$\frac{1}{s^2}|F|_{D_3} \le c_6\mu\Gamma(r - r_+) + c_6\mu, \tag{3.21}$$

where

$$\Gamma(r - r_+) = \sum_{k\neq 0}|k|^{4\tau m_0^2}e^{-|k|\frac{r-r_+}{4}}.$$

By Cauchy's estimate, on D_2,

$$\frac{1}{r - r_+}|F_x|, s^2|F_y|, s|F_z| \le c_6 s^2\mu\Gamma(r - r_+) + c_6 s^2\mu. \tag{3.22}$$

Since F is a polynomial of y and z with orders 1 and 2 respectively, by (3.22) we obtain

$$|D^l F|_{D_1} \le c_6\mu\Gamma(r - r_+) + c_6\mu, \qquad |l| \le 4. \tag{3.23}$$

3.4. Estimates of the New Perturbation

From the last section we have

$$H_+ = H \circ \phi_F^1 = N_+ + P_+,$$

where

$$N_+ = N + [R] - \langle P_{001}, z\rangle = e_+ + \langle \omega_+, y\rangle + \frac{\delta}{2}\langle z, M_+ z\rangle,$$
$$P_+ = \int_0^1 \{R_t, F\} \circ \phi_F^t \, \mathrm{d}t + (P - R) \circ \phi_F^1,$$

with

$$e_+ = e + P_{000},$$
$$\omega_+ = \omega + P_{010}, |P_{010}| \le \delta\mu, \tag{3.24}$$
$$M_+ = M + P_{002}, |P_{002}| \le \delta\mu. \tag{3.25}$$

Let

$$D_{\frac{i}{2}\alpha} = D\left(r_+ + \frac{i-1}{2}(r - r_+), \frac{i}{2}\alpha s\right), \qquad i = 1, 2.$$

For one single KAM step, everything has been done but the estimate of the new perturbation P_+ on a smaller domain.

Note

$$\phi_F^t = \mathrm{id} + \int_0^t X_F \circ \phi_F^\lambda \, \mathrm{d}\lambda,$$

and

$$D\phi_F^t = I_{2n} + \int_0^t (DX_F) D\phi_F^\lambda \, \mathrm{d}\lambda = I_{2n} + \int_0^t J(D^2F) D\phi_F^\lambda \, \mathrm{d}\lambda. \tag{3.26}$$

From (3.22) we have

$$\phi_F^t\colon\ D_{\frac{1}{2}\alpha} \longrightarrow D_\alpha, \qquad 0 \le t \le 1,$$

provided

$$c_7\left(\frac{\mu}{\gamma^{4m_0^2}}\Gamma(r - r_+) + \mu\right) < \frac{1}{2}(r - r_+),$$
$$c_7\left(s^2\mu\Gamma(r - r_+) + s^2\mu\right) < \frac{1}{4}\alpha^2 s^2,$$
$$c_7(s\mu\Gamma(r - r_+) + s\mu) < \frac{1}{2}\alpha s. \tag{3.27}$$

It follows that $\phi_F^1\colon\ D_{\frac{1}{2}\alpha} \longrightarrow D_\alpha$.

It follows from (3.23) that

$$|D\phi_F^t - I_{2n}|_{D_{\frac{1}{2}\alpha}} \le 2|D^2F|_{D_{\frac{1}{2}\alpha}} \le c_7(\mu\Gamma(r - r_+) + \mu),$$
$$|D^2\phi_F^t|_{D_{\frac{1}{2}\alpha}} \le c_7(\mu\Gamma(r - r_+) + \mu). \tag{3.28}$$

Now we estimate P_+. By (3.6) and (3.22) we have that on $D_{\frac{1}{2}\alpha}$,

$$|P_+| \le c_8(\delta\gamma^{4m_0^2}s^2\mu^2\Gamma(r-r_+) + \delta\gamma^{4m_0^2}s^2\mu^2). \tag{3.29}$$

Thus, on $D(r_+, s_+) = D_{\frac{1}{2}\alpha}$,

$$\frac{1}{s_+^2}|P_+| \le c_8\delta\gamma^{4m_0^2}(\Gamma(r-r_+)+1)\mu^{\frac{4}{3}}.$$

Hence one cycle of KAM steps is completed.

4. Iteration Lemma

The following Iteration Lemma checks the validity of the KAM iteration. For given r_0, s_0, and μ_0, we define some sequences inductively:

$$\begin{aligned}
r_\nu &= r_0\left(1-\sum_{i=1}^{\nu}\frac{1}{2^{i+1}}\right),\\
\gamma_\nu &= \gamma_0\left(1-\sum_{i=1}^{\nu}\frac{1}{2^{i+1}}\right),\\
s_\nu &= \frac{1}{2}\alpha_{\nu-1}s_{\nu-1},\\
\alpha_\nu &= \mu_\nu^{\frac{1}{3}},\\
\mu_\nu &= (16c\alpha_{\nu-1})^{\frac{1}{6}}\mu_{\nu-1},\\
K_{\nu+1} &= \left[\frac{1}{\mu_\nu}\right]+1, \qquad K_0 = 0,\\
D_\nu &= D(r_\nu, s_\nu), \qquad \mathcal{O}_\nu = \mathcal{O}_{\gamma_\nu}(K_\nu), \qquad \mathcal{O}_0 = \mathcal{O}_{\gamma_0},
\end{aligned}$$

where c is $4^{m_0^2}$ times the biggest one of all c_i's, $[\cdot]$ denotes the integral part.

Lemma 4.1 (Iteration Lemma). *There is a sufficiently small μ_0, δ_0 depending only on r_0 and γ_0, so that the following hold for all ν. Let*

$$H_\nu = N_\nu + P_\nu, \quad N_\nu = e_\nu + \langle\omega, y\rangle + \frac{\delta}{2}\langle z, M_\nu z\rangle,$$

such that

$$\frac{1}{s_\nu^2}|P_\nu| \le \delta\gamma_\nu^{4m_0^2}\mu_\nu,$$

on $D_\nu \times \mathcal{O}_\nu$. Then there is a subset $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu$,

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu - \bigcup_{K_\nu<|k|\le K_{\nu+1}} R_k^{\nu+1}(\gamma_\nu),$$

where

$$R_k^{\nu+1}(\gamma_\nu) = \Bigg\{ \omega \in \mathcal{O}_\nu \colon\ |\langle k, \omega\rangle| \le \frac{\gamma_\nu}{|k|^\tau}, \quad or\ \ |\det A_{1,\nu}| \le \frac{\gamma_\nu^{2m_0}}{|k|^{2m_0\tau}}, \ \ or\ \ |\det A_{2,\nu}| \le \frac{\gamma_{\nu+1}^{4m_0^2}}{|k|^{4m_0^2\tau}} \Bigg\},$$

and a symplectic change

$$\Phi_\nu \colon\ D_{\nu+1} \times \mathcal{O}_{\nu+1} \longrightarrow D_\nu,$$

such that

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

and on $D_{\nu+1} \times \mathcal{O}_{\nu+1}$,

$$\frac{1}{\delta \gamma_{\nu+1}^{4m_0^2} s_{\nu+1}^2} |P_{\nu+1}| \le \mu_{\nu+1}.$$

Proof. By induction, one verifies that $c_8 \gamma_\nu^{4m_0^2}(\Gamma(r_\nu - r_{\nu+1}) + 1)\mu_\nu^{\frac{4}{3}} \le \gamma_{\nu+1}^{4m_0^2}\mu_{\nu+1}$ for all $\nu \ge 0$ as well as (3.4), (3.27). For simplicity, let $r_0 = 1$. Since the proof is standard—see for example [14]—here we only verify the first inequality. By the inductive assumptions, it suffices to prove

$$c\left(\Gamma\left(\frac{1}{2^{\nu+2}}\right) + 1\right)\mu_\nu^{\frac{4}{3}} < \mu_{\nu+1}. \tag{4.1}$$

It is equivalent to

$$\mu_\nu^{\frac{5}{18}} \Gamma\left(\frac{1}{2^{\nu+2}}\right) < \frac{1}{16}(16c)^{\frac{1}{6}}. \tag{4.2}$$

Note

$$\begin{aligned} \Gamma\left(\frac{1}{2^{\nu+2}}\right) &\le \int_1^\infty \lambda^{m+4\tau m_0^2} e^{-\lambda \frac{1}{2^{\nu+5}}}\, \mathrm{d}\lambda \\ &\le (m + 4\tau m_0^2)!\, 2^{(\nu+5)(m+4\tau m_0^2)}. \end{aligned} \tag{4.3}$$

Here we assume that $\tau > 2m + 2$ is an integer.

It is sufficient to prove

$$\mu_\nu^{\frac{5}{18}} (m + 4\tau m_0^2)!\, 2^{(\nu+5)(m+4\tau m_0^2)} < \frac{1}{16}(16c)^{\frac{1}{6}}. \tag{4.4}$$

It is true if μ_0 is sufficiently small. In fact, taking $\lambda \gg 1$, such that

$$\mu_0 < \frac{1}{(16c\lambda^{\frac{6\times 18}{5}})^3} < 1,$$

then

$$\begin{aligned}
\mu_1 &= (16c\mu_0^{\frac{1}{3}})^{\frac{1}{6}}\mu_0 < \frac{1}{\lambda^{\frac{18}{5}}}\mu_0 < 1,\\
\mu_2 &= (16c\mu_1^{\frac{1}{3}})^{\frac{1}{6}}\mu_1 < \frac{1}{\lambda^{\frac{18}{5}}}\mu_1 < \frac{1}{\lambda^{\frac{2\times 18}{5}}}\mu_0,\\
&\cdots\cdots\\
\mu_\nu &= (16c\mu_{\nu-1}^{\frac{1}{3}})^{\frac{1}{6}}\mu_{\nu-1} < \cdots < \frac{1}{\lambda^{\frac{18\nu}{5}}}\mu_0,
\end{aligned} \tag{4.5}$$

which implies (4.4). □

5. Proof of Theorem 3

Clearly, the Iteration Lemma can be applied to (1.3) for $\nu = 0$ by assumptions. Inductively we have the following sequences:

$$\begin{aligned}
&D_\nu \times \mathcal{O}_\nu \subset D_{\nu-1} \times \mathcal{O}_{\nu-1},\\
&\Psi^\nu = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_\nu\colon\ D_{\nu+1} \times \mathcal{O}_{\nu+1} \to D_0,\ \nu \geq 1,\\
&H \circ \Psi^\nu = H_\nu = N_\nu + P_\nu.
\end{aligned}$$

Let $\mathcal{O}_* = \bigcap_{\nu=0}^{\infty} \mathcal{O}_\nu$. By (3.28) and a typical argument similar to [14], N_ν, Ψ_ν, and $D\Psi^\nu$ converge uniformly on $D(r, \frac{s}{2}) \times \mathcal{O}_*$ with

$$\begin{aligned}
&N_\infty = \varepsilon_\infty + \langle \omega, y\rangle + \frac{\delta}{2}\langle z, M_\infty z\rangle = H_\infty,\\
&|M_0 - M_\infty| \leq c_9\mu_0.
\end{aligned} \tag{5.1}$$

Let ϕ_H^t be the flow of X_H. From $H \circ \Psi^\nu = H_\nu$ we obtain

$$\phi_H^t \circ \Psi^\nu = \Psi^\nu \circ \phi_{H_\nu}^t.$$

Taking the limit yields

$$\phi_H^t \circ \Psi^\infty = \Psi^\infty \circ \phi_{H_\infty}^t$$

on $D(\frac{1}{2}r_0, 0, 0) \times \mathcal{O}_*$. That means that for any $\omega \in \mathcal{O}_*$, (1.3) has an invariant torus.

We focus our attention on proving

$$|\mathcal{O}_0 \backslash \mathcal{O}_*| = O(\gamma_0).$$

Fix $k \in Z^m \backslash 0$. Let us estimate the measure of $R_k^{\nu+1}(\gamma_\nu)$. Let $g(s) = \det(A_{2,\nu}(s))$ and $g_1(s) = (\operatorname{Re} g(s))^2 + (\operatorname{Im} g(s))^2$.

Without loss of generality, we assume that $k_1 = \max\{|k_i|\}$. Let

$$\begin{aligned}
S_1 &= \left\{\omega_1\colon\ \omega = (\omega_1, \ldots, \omega_m),\ |g(\langle k, \omega\rangle)| \leq \frac{\gamma_\nu^{4m_0^2}}{|k|^{4m_0^2\tau}}\right\}\\
&= \left\{\omega_1\colon\ \omega = (\omega_1, \ldots, \omega_m),\ |g_1(\langle k, \omega\rangle)| \leq \frac{\gamma_\nu^{8m_0^2}}{|k|^{8m_0^2\tau}}\right\}.
\end{aligned}$$

Note that

$$\left|\frac{\partial^{8m_0^2}}{\partial\omega_1^{8m_0^2}}g_1(\langle k,\omega\rangle)\right| = |k_1|^{8m_0^2}(1+O(\delta)) = A,$$

where $O(\delta)$ is independent of k. Hence for small δ,

$$A \geq \frac{1}{2}|k_1|^{8m_0^2} > \frac{1}{2}.$$

By [18], Lemma 2.1,

$$\left|\left\{\omega_1\colon\ |g_1(\langle k,\omega\rangle)| \leq \frac{\gamma_\nu^{8m_0^2}}{|k|^{8m_0^2\tau}}\right\}\right| \leq c\frac{\gamma_\nu}{|k|^\tau},$$

where $c = (2(1+2+\cdots+8m_0^2)+2$. Therefore,

$$|S_1| \leq c\frac{\gamma_\nu}{|k|^\tau}.$$

Clearly, by Fubini's theorem,

$$\left|\left\{\omega\colon\ |g(\langle k,\omega\rangle)| \leq \frac{\gamma_\nu^{4m_0^2}}{|k|^{4m_0^2\tau}}\right\}\right| \leq c_9|S_1| \leq c_{10}\frac{\gamma_\nu}{|k|^\tau}.$$

Similarly,

$$\left|\left\{\omega\colon\ |\langle k,\omega\rangle| \leq \frac{\gamma_\nu}{|k|^\tau}\right\}\right| \leq c_{10}\frac{\gamma_\nu}{|k|^\tau},$$

$$\left|\left\{\omega\colon\ |\det A_{1,\nu}(\langle k,\omega\rangle)| \leq \frac{\gamma_\nu^{2m_0}}{|k|^{2m_0\tau}}\right\}\right| \leq c_{10}\frac{\gamma_\nu}{|k|^\tau}. \tag{5.2}$$

Hence

$$|R_k^\nu(\gamma_\nu)| \leq 3c_{10}\frac{\gamma_\nu}{|k|^\tau} \leq 3c_{10}\frac{\gamma_0}{|k|^\tau}. \tag{5.3}$$

Thus, by (5.3),

$$\left|\bigcup_{K_\nu<|k|\leq K_{\nu+1}} R_k^{\nu+1}(\gamma_\nu)\right| \leq 3c_{10}\gamma_0\sum_{K_\nu<|k|\leq K_{\nu-1}}\frac{1}{|k|^\tau} \leq 3c_{10}\gamma_0\sum_{i=K_\nu}^{K_{\nu+1}}\frac{1}{i^2}.$$

Note

$$\mathcal{O}\backslash\mathcal{O}_\gamma \subset \bigcup_{i=0}^{\nu}\bigcup_{K_i<|k|\leq K_{i+1}} R_k^{i+1}(\gamma_i).$$

Therefore

$$|\mathcal{O}\backslash\mathcal{O}_*| \leq 3c_{10}\gamma_0\sum_1^\infty\frac{1}{i^2} = O(\gamma_0). \tag{5.4}$$

This shows that $\mathcal{O}_*$ is nonempty if γ_0 is sufficiently small. The proof of Theorem 3 is completed. ☐

Acknowledgments

The authors express sincere thanks to the referees for their valuable advice and suggestions, which helped us improve the manuscript greatly, and to Professor S. Wiggins for his reprints and preprints, which inspired some further consideration.

References

[1] V. I. Arnold, Proof of A. N. Kolmogorov's theorem on the preservation of quasi periodic motions under small perturbations of the Hamiltonian, *Usp. Math. USSR* **18** (1963), 13–40.
[2] J. Bourgain, Construction of quasi-periodic solution for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, *Int. Math. Res. Not.* **11** (1994), 475–497.
[3] H. Broer, G. Huitema, and M. Sevryuk, *Quasi-Periodic Motions in Families of Dynamical Systems*, Lecture Notes in Math., **1645**, Springer-Verlag, New York, 1996.
[4] C.-Q. Cheng, Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems, *Commun. Math. Phys.* **177** (1996), 529–559.
[5] L. Chierchia and G. Gallavotti, Drift and diffusion in phase space, *Ann. Inst. H. Poincaré Phys. Theor.*, **69** (1994), 1–144.
[6] L. H. Eliasson, Biasymptotic solutions of perturbed integrable Hamiltonian systems, *Bol. Soc. Mat.* **25** (1994), 57–76.
[7] S. M. Graff, On the continuation of hyperbolic invariant tori for Hamiltonian systems, *J. Diff. Eq.* **15** (1974), 1–60.
[8] A. N. Kolmogorov, On quasi-periodic motions under small perturbations of the Hamiltonian, *Dokl. Akad. Nauk USSR* **98** (1954), 527–530.
[9] J. Milnor, *Morse Theory*, Princeton University Press, Princeton, NJ, 1963.
[10] J. Moser, On invariant curves of area preserving mappings of an annulus, *Nachr. Akad. Wiss. Gott. Math. Phys. K1*, **2** (1962), 1–20.
[11] C. L. Siegel and J. Moser, *Lectures on Celestial Mechanics*, Grundlehren **187**, Springer-Verlag, Berlin, 1971.
[12] H. Poincaré, *Les Méthodes Nouvelles de la Mécaniques Céleste*, I–III, Dover Publications, New York, 1957; English translation: *New Methods of Celestial Mechanics*, AIP Press, Williston, MD, 1992.
[13] J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, *Commun. Pure Appl. Math.* **35** (1982), 653–696.
[14] J. Pöschel, On the elliptic lower dimensional tori in Hamiltonian systems, *Math. Z.* **202** (1989), 559–608.
[15] M. Rudnev and S. Wiggins, KAM theory near multiplicity one resonant surfaces in perturbations of A-priori stable Hamiltonian systems, *J. Nonlin. Sci.* **7** (1997), 177–209.
[16] D. V. Treshchev, Mechanism for destroying resonance tori of Hamiltonian systems, *Mat. USSR Sb.* **180** (1989), 1325–1346.
[17] S.-L. Wang, The existence of lower dimensional tori of mixed type for generic Hamiltonian systems, Preprint, 1997.
[18] J. Xu, J. You, and Q. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* **226** (1997), 375–387.
[19] J. You, Perturbations of elliptic lower dimensional tori for Hamiltonian systems, Preprint, 1997.