PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 138, Number 7, July 2010, Pages 2385–2395
S 0002-9939(10)10151-8
Article electronically published on February 18, 2010

PERSISTENCE OF THE NON-TWIST TORUS IN NEARLY INTEGRABLE HAMILTONIAN SYSTEMS

JUNXIANG XU AND JIANGONG YOU

(Communicated by Yingfei Yi)

Abstract. In this paper we consider analytic nearly integrable hamiltonian systems, and prove that if the frequency mapping has nonzero Brouwer topological degree at some Diophantine frequency, then the invariant torus with this frequency persists under small perturbations.

1. Introduction

Consider an analytic hamiltonian $H(q,p) = h(p) + f(q,p)$, where $(q,p) \in T^n \times D$, with T^n being the usual n-dimensional torus and D a bounded simply connected open domain of $\mathbb{R}^n$. $h(p)$ and $f(q,p)$ are real analytic on $\bar{D}$ and $\bar{D} \times T^n$, respectively. The corresponding hamiltonian system reads as

$$\begin{cases} \dot{q} &= H_p(q,p) = h_p(p) + f_p(q,p) \\ \dot{p} &= -H_q(q,p) = \qquad -f_q(q,p) \end{cases}. \tag{1.1}$$

If $f = 0$, the system (1.1) is integrable and possesses a family of invariant tori $T^n \times \{p_0\}$ for all $p_0 \in D$, with $\omega(p_0) = h_p(p_0)$ as its frequency. The whole phase space is occupied by the invariant tori. Under Kolmogorov's non-degeneracy condition, that is,

$$\det(\partial\omega/\partial p) = \det(h_{pp}) \neq 0,$$

the classical KAM theorem asserts that most of the tori will survive small perturbations [6, 1, 2, 7, 5, 8, 9]. What's more, for a fixed Diophantine frequency in the image of the frequency map, the perturbed system still has an invariant torus with this frequency (in this case, we say that the torus *persists* under small perturbations).

The classical KAM theorem can be extended to the case of Rüssmann's non-degeneracy condition

$$a_1\omega_1(p) + a_2\omega_2(p) + \cdots + a_n\omega_n(p) \not\equiv 0 \text{ on } \bar{D} \tag{1.2}$$

Received by the editors February 19, 2009, and, in revised form, August 3, 2009.
2010 *Mathematics Subject Classification.* Primary 34D10, 34D23; Secondary 34C27.
Key words and phrases. Hamiltonian system, KAM iteration, invariant tori, non-degeneracy condition.
The first author was supported by the National Natural Science Foundation of China (10571027).
The second author was partially supported by the National Basic Research Program of China (973 Program, 2007CB814800) and by the NNSF of China (Grant No. 10531050).

for all $(a_1, a_2, \cdots a_n) \in \mathbb{R}^n \setminus \{0\}$ [10, 12, 11, 4, 16, 15, 13] in the sense that the perturbed system (1.1) still has a family of invariant tori of positive measure.

However, under Rüssmann's non-degeneracy condition one can only get the existence of a family of invariant tori, but there is no information on the persistence or not of any individual torus. In fact, one cannot expect persistence when the image of the frequency map is a sub-manifold.

In this paper, we will investigate the persistence of tori without assuming Kolmogorov's non-degeneracy condition. Consider an unperturbed torus with frequency $\omega_0 = \omega(p_0) = h_p(p_0)$. If $h_{pp}(p_0) = 0$, we call this unperturbed torus *non-twist*. We will prove that if ω_0 is a Diophantine frequency and the topological degree $\deg(\omega, D, \omega_0) \neq 0$, then the perturbed system still has an invariant torus with ω_0 as its frequency, i.e., the torus persists under small perturbations. The following theorem is the main result of this paper.

Theorem 1.1. *Suppose that $h(p)$ and $f(q,p)$ are real analytic on $\bar{D}$ and $T^n \times \bar{D}$, respectively. Let $\omega(p) = h_p(p)$ and $\omega_0 = \omega(p_0)$, with $p_0 \in D$. Suppose that ω_0 satisfies the Diophantine condition*

$$\big|\langle \omega_0, k\rangle\big| \geq \frac{\alpha}{|k|^\tau}, \forall\, 0 \neq k \in \mathbb{Z}^n, \tag{1.3}$$

and that the Brouwer degree of the frequency mapping ω at ω_0 on D is not zero, i.e.

$$\deg(\omega, D, \omega_0) \neq 0.$$

Then there exists a sufficiently small positive constant $\epsilon > 0$ such that the system (1.1) has an invariant torus with ω_0 as its frequency if $\|f\| = \sup_{T^n \times D} |f(q,p)| \leq \epsilon$.

Remark. By a property of the topological degree, it follows easily that for a Diophantine frequency ω_* sufficiently close to ω_0, the invariant torus with ω_* as its frequency can also persist under small perturbations.

Remark. An example to which the above theorem can be applied is $\omega(p) = \omega_0 + (p_1^3, p_2^3 \cdots, p_n^3)$. At $p = 0$, ω is degenerate in the Kolmogorov sense and so the classic KAM theorem cannot be applied. Although ω satisfies Rüssmann's non-degeneracy condition, the previous KAM theorems cannot tell us whether the perturbed system has an invariant torus with ω_0 as the frequency.

Remark. Our result can be easily generalized to lower dimensional hyperbolic invariant tori. However, this is not true for the elliptic case.

We follow the paper [8] in the standard part of KAM iteration. First we linearize the hamiltonian system (1.1) at the invariant tori of the integrable system, and then we will consider instead a parameterized hamiltonian system. For any $\xi \in D$, let $p = \xi + I$ and $q = \theta$. Then,

$$\begin{aligned} H(q,p) &= h(\xi) + \langle h_p(\xi), I\rangle + f_h(I;\xi) + f(\theta, \xi + I) \\ &= e + \langle \omega(\xi), I\rangle + P(\xi, \theta, I), \end{aligned}$$

where $e = h(\xi)$, $\omega(\xi) = h_p(\xi)$, $P(\theta, I; \xi) = f_h(I, \xi) + f(\theta, \xi + I)$, and $\xi \in D$ is regarded as a parameter. Here e is an energy constant, which is usually omitted, $\omega : \xi \to \omega(\xi)$ is called the frequency mapping, and P is a small perturbation term.

Let

$$D(s,r) = \big\{(\theta, I) \in \mathbb{C}^n \times \mathbb{C}^n \big|\, |\mathrm{Im}\ \theta|_\infty \leq s,\ |I|_1 \leq r\big\},$$

where $|\mathrm{Im}\ \theta|_\infty = \max_{1\le i\le n} |\mathrm{Im}\theta_i|$ and $|I|_1 = \sum_{1\le i\le n} |I_i|$. Let

$$\Pi = \{\xi \in D \mid \mathrm{dist}\,(\xi, \partial D) \ge \sigma\},$$

where $\sigma > r > 0$ is a small constant. Let Π_σ be the complex closed neighborhood of Π in $\mathbb{C}^n$ with radius σ, that is,

$$\Pi_\sigma = \{\xi \in \mathbb{C}^n \mid \mathrm{dist}\,(\xi, \Pi) \le \sigma\}.$$

Now the hamiltonian function $H(\xi;\theta,I)$ is real analytic in $(\xi;\theta,I)$ on $\Pi_\sigma \times D(s,r)$. The corresponding hamiltonian system becomes

$$\left\{\begin{array}{lll} \dot\theta & = & H_I = \omega(\xi) + P_I(\xi;\theta,I) \\ \dot I & = -H_\theta = & \qquad\quad -P_\theta(\xi;\theta,I) \end{array}\right. . \tag{1.4}$$

Thus, persistence of invariant tori for the nearly integrable system (1.1) is reduced to that of invariant tori for the family of hamiltonian systems (1.4) indexed by the parameter $\xi \in \Pi$.

We expand $P(\xi;\theta,I)$ as a Fourier series with respect to θ:

$$P(\xi;\theta,I) = \sum_{k\in\mathbb{Z}^n} P_k(\xi;I)\, e^{\mathrm{i}\langle k,\theta\rangle}.$$

Define

$$\|P\|_{\Pi_\sigma\times D(s,\,r)} = \sum_{k,\,l} \|P_k\|_{\sigma;r}\, e^{s|k|},$$

where $\|P_k\|_{\sigma;r} = \sup_{\xi\,\in\Pi_\sigma, |I|_1\le r} |P_k(\xi;I)|$.

Theorem 1.2. *Let $H(\xi;\theta,I) = \langle\omega(\xi), I\rangle + P(\xi;\theta,I)$ be real analytic on $\Pi_\sigma\times D(s,r)$, where $\Pi \subset \mathbb{R}^n$ is a bounded simply connected domain. Let $\omega_0 = \omega(\xi_0)$ with $\xi_0 \in \Pi$. Suppose that ω_0 satisfies* (1.3) *and that* $\deg(\omega,\Pi,\omega_0) \ne 0$. *Then there exists a sufficiently small positive constant $\epsilon > 0$ such that if $\|P\|_{\Pi_\sigma\times D(s,\,r)} \le \epsilon$, there exists $\xi_* \in \Pi$ such that the hamiltonian system* (1.4) *at $\xi = \xi_*$ has an invariant torus with ω_0 as its frequency.*

Remark. Theorem 1.2 also holds true if the hamiltonian system (1.4) is finitely smooth with respect to the parameter. For some related results we refer to [3] for details.

2. Proof of the theorems

Our key idea is to introduce an artificial external parameter λ and consider the following hamiltonian system:

$$\left\{\begin{array}{lll} \dot\theta & = & H_I = \omega(\xi) + \lambda + P_I(\xi;\theta,I) \\ \dot I & = -H_\theta = & \qquad\quad -P_\theta(\xi;\theta,I) \end{array}\right. , \tag{2.1}$$

where $H = H(\xi,\lambda;\theta,I) = \langle\omega(\xi)+\lambda, I\rangle + P(\xi;\theta,I)$. The hamiltonian system (1.4) corresponds to the hamiltonian system (2.1) with $\lambda = 0$. The method of introducing a parameter was used in [13] to give a simple proof of the KAM theorem under Rüssmann's non-degeneracy condition. We will first give a KAM theorem for the hamiltonian system (2.1) with parameters (ξ,λ) and then prove Theorem 1.2.

Let

$$d = \max_{\xi,\eta\in\Pi_\sigma} |\omega(\xi) - \omega(\eta)|$$

and define

$$B(\omega,d) = \{\lambda \in \mathbb{C}^n \mid \mathrm{dist}(\lambda,\omega) < d\}.$$

Let $O = (\bigcup_{\xi\in\Pi} B(\omega(\xi), d)) \cap \mathbb{R}^n$. It follows that

$$\omega(\Pi) = \big\{\omega(\xi) \mid \xi \in \Pi\big\} \subset O.$$

Let

$$O_\alpha = \{\Omega \in O \mid |\langle k, \Omega\rangle| \geq \frac{\alpha}{|k|^\tau},\ \forall k \in \mathbb{Z}^n \setminus \{0\}\big\}$$

and $O_{\alpha,\tilde\delta} = B(O_\alpha, \tilde\delta)$. Let $K > 0$ and $\tilde\delta = \frac{\alpha}{2K^{\tau+1}}$. Then, for all $\Omega \in O_{\alpha,\tilde\delta}$ it follows that

$$|\langle k, \Omega\rangle| \geq \frac{\alpha}{2|k|^\tau}, \quad 0 < |k| \leq K.$$

Let $M = \Pi_\sigma \times B(0, 2d+1)$. The hamiltonian $H(\xi, \lambda; \theta, I)$ is real analytic on $M \times D(s, r)$. Without loss of generality, write $P(\xi, \lambda; \theta, I) = P(\xi; \theta, I)$.

Theorem 2.1. *There exists a small $\epsilon > 0$ such that if*

$$\|P\|_{M\times D(s,r)} \leq \epsilon,$$

then we have a Cantor-like family of analytic curves in M,

$$\{\Gamma_\Omega : \lambda = \lambda(\xi), \xi \in \Pi \mid \Omega \in O_\alpha\},$$

which are determined implicitly by the equation

$$\lambda + \omega(\xi) + h(\xi, \lambda) = \Omega,$$

where $h(\xi, \lambda)$ is a C^∞-smooth function on M with $|h(\xi, \lambda)| \leq 2\epsilon/r$ and $|h_\lambda(\xi, \lambda)| + |h_\xi(\xi, \lambda)| \leq \frac{1}{2}$, and a parameterized family of symplectic mappings

$$\Phi(\xi, \lambda; \cdot, \cdot) : D(s/2, r/2) \to D(s, r),\ (\xi, \lambda) \in \Gamma = \bigcup_{\Omega\in O_\alpha} \Gamma_\Omega,$$

where Φ is C^∞-smooth in (ξ, λ) on Γ in the sense of Whitney and analytic in (θ, I) on $D(s/2, r/2)$, such that for each $(\xi, \lambda) \in \Gamma_\Omega$,

$$H(\xi, \lambda; \Phi(\xi, \lambda; \theta, I)) = \langle \Omega, I\rangle + P_*(\xi, \lambda; \theta, I),$$

where $P_*(\xi, \lambda; \theta, I) = O(I^2)$ at $I = 0$. Therefore, the hamiltonian system (1.4) *has invariant tori $\Phi(\xi, \lambda; T^n, 0)$ with Ω as their frequencies.*

Now we first use Theorem 2.1 to prove Theorem 1.2 and delay the proof of Theorem 2.1 until later. In fact, let $\Omega = \omega_0$ and then we have an analytic curve $\Gamma_{\omega_0} : \xi \in \Pi \to \lambda(\xi)$, implicitly determined by the equation $\lambda + \omega(\xi) + h(\xi, \lambda) = \omega_0$. By the implicit function theorem we have

$$\lambda(\xi) = \omega_0 - \omega(\xi) + \hat\lambda(\xi),\ \forall \xi \in \Pi.$$

Moreover, if ϵ is sufficiently small, we have $|\hat\lambda(\xi)| \leq 2\epsilon/r$ and $|\hat\lambda_\xi(\xi)| \leq 4\epsilon/r$. From the assumption it follows that

$$\deg(\omega_0 - \omega, \Pi, 0) \neq 0.$$

So if ϵ is sufficiently small we have

$$\deg(\lambda, \Pi, 0) = \deg(\omega_0 - \omega, \Pi, 0) \neq 0.$$

Then we have $\xi_* \in \Pi$ such that $\lambda(\xi_*) = 0$. Therefore, the hamiltonian system (1.4) with $H(\xi_*; \theta, I) = H(\xi_*, \lambda(\xi_*); \theta, I)$ has an invariant torus $\Phi(\xi_*, \lambda(\xi_*); T^n, 0)$ with ω_0 as the frequency.

Now it remains to prove Theorem 2.1. Our method is the standard KAM iteration. We should note that it is very important to keep the parameters ξ and λ in

the KAM iteration so that one can see clearly the dependence of KAM tori on the parameters.

KAM step. We summarize our KAM step in the following iteration lemma.

Lemma 2.1 (Iteration Lemma)**.** *Consider the hamiltonian*

$$H(\xi,\lambda;\theta,I)=N(\xi,\lambda;I)+P(\xi,\lambda;\theta,I),$$

where $N(\xi,\lambda;I)=\langle\Omega(\xi,\lambda),I\rangle$ *with* $\Omega(\xi,\lambda)=\omega(\xi)+\lambda+h(\xi,\lambda)$. *Assume that the following hold:*

A1. N *and* P *are analytic on* M *and* $M\times D(s,r)$, *respectively. With* $0<E<1$ *and* $0<\rho<s/5$, P *satisfies*

$$\|P\|_{M\times D(s,r)}\le\epsilon=\alpha r\rho^{\tau+n+1}E.$$

A2. *The function* h *satisfies*

$$|h_\lambda(\xi,\lambda)|+|h_\xi(\xi,\lambda)|<\frac{1}{2},\quad\forall(\xi,\lambda)\in M, \tag{2.2}$$

and for each $\Omega\in O_\alpha$ *the equation*

$$\Omega(\xi,\lambda)=\omega(\xi)+\lambda+h(\xi,\lambda)=\Omega$$

defines implicitly an analytic mapping

$$\lambda:\xi\in\Pi_\sigma\to\lambda(\xi)\in B(0,2d+1)$$

such that $\Gamma_\Omega=\{(\xi,\lambda(\xi))\mid\xi\in\Pi_\sigma\}\subset M$. *Moreover, for* $K>0$ *satisfying* $e^{-K\rho}=E$, $\tilde\delta=\frac{\alpha}{2K^{\tau+1}}$ *and* $\delta=\frac{2}{3}\tilde\delta$, *we have*

$$U(\Gamma_\Omega,\delta)=\{(\xi,\lambda')\in\Pi_\sigma\times\mathbb{C}^n\mid|\lambda'-\lambda(\xi)|\le\delta\}\subset M.$$

Then, there exist $M_+\subset M$ *and* $D(s_+,r_+)\subset D(s,r)$ *such that for any* $(\xi,\lambda)\in M_+$ *there exists a symplectic mapping*

$$\Phi(\xi,\lambda;\cdot,\cdot):D(s_+,\,r_+)\to D(s,r),$$

with Φ *real analytic on* $M_+\times D(s_+,r_+)$, *such that*

$$H_+(\xi,\lambda;\theta,I)=H(\xi,\lambda;\Phi(\xi,\lambda;\theta,I))=N_+(\xi,\lambda;I)+P_+(\xi,\lambda;\theta,I),$$

where $N_+(\xi,\lambda;I)=\langle\Omega_+(\xi,\lambda),I\rangle$ *with* $\Omega_+(\xi,\lambda)=\omega(\xi)+\lambda+h(\xi,\lambda)+\hat h(\xi,\lambda)$, *and the following conclusions hold:*

(**i**) *The new perturbation term* P_+ *satisfies*

$$\|P_+\|_{M_+\times D(s_+,r_+)}\le\epsilon_+=\alpha_+r_+\rho_+^{\tau+n+1}E_+$$

with

$$s_+=s-5\rho,\ \eta=\sqrt{E},\ \rho_+=\frac{1}{2}\rho,\ r_+=\eta r,\ E_+=cE^{\frac{3}{2}},$$

and

$$M_+=\big\{(\xi,\lambda')\in\mathbb{C}^n\times\mathbb{C}^n\mid\xi\in\Pi_{\sigma-\frac{1}{2}\delta},(\xi,\lambda)\in\Gamma,|\lambda'-\lambda|\le\frac{1}{2}\delta\big\},$$

where $\Gamma=\bigcup_{\Omega\in O_\alpha}\Gamma_\Omega$. *Moreover, for the mapping* Φ *we have the following estimates:*

$$\|W(\Phi-\mathrm{id})\|_{M_+\times D(s_+,r_+)}\le cE,$$

and

$$\|W(\mathcal{D}\Phi-\mathrm{Id})W^{-1}\|_{M_+\times D(s_+,r_+)}\le cE,$$

where $\mathcal{D}$ *is the differentiation operator with respect to* (θ,I) *and* W *is the matrix* $\mathrm{diag}(\rho^{-1}I_n,r^{-1}I_n)$ *with* I_n *being the* n*-th unit matrix.*

(ii) *$\hat{h}$ satisfies*

$$|\hat{h}(\xi,\lambda)| \le \frac{\epsilon}{r} = \alpha\rho^{\tau+n+1}E,\ \forall(\xi,\lambda)\in M$$

and

$$|\hat{h}_\lambda(\xi,\lambda)| + |\hat{h}_\xi(\xi,\lambda)| \le \frac{2\epsilon}{r\delta},\ \forall(\xi,\lambda)\in M_+.$$

Thus, if

$$2\alpha\rho^{\tau+n+1}E \le \frac{1}{4}\delta, \tag{2.3}$$

then the equation

$$\Omega_+(\xi,\lambda) = \omega(\xi) + \lambda + h_+(\xi,\lambda) = \Omega$$

determines implicitly an analytic mapping

$$\lambda_+ : \xi \in \Pi_{\sigma_+} \to \lambda_+(\xi) \in B(0, 2d+1) \text{ with } \sigma_+ = \sigma - \frac{1}{2}\delta$$

that satisfies

$$|\lambda_+(\xi) - \lambda(\xi)| \le \frac{2\epsilon}{r} = 2\alpha\rho^{\tau+n+1}E \le \frac{1}{4}\delta \tag{2.4}$$

and

$$\Gamma_\Omega^+ = \{(\xi, \lambda_+(\xi)) \mid \xi \in \Pi_{\sigma_+}\} \subset M_+. \tag{2.5}$$

Let $\delta_+ = \frac{\alpha}{2K_+^{\tau+1}}$ with K_+ satisfying $e^{\rho_+K_+} = E_+$. If

$$\delta_+ < \frac{1}{4}\delta, \tag{2.6}$$

then for all $\Omega \in O_\alpha$ we have $U(\Gamma_\Omega^+, \delta_+) \subset M_+$.

Remark. The above lemma is actually one step in our KAM iteration. If (2.3) and (2.6) hold and h_+ satisfies (2.2), then the assumptions **A1** and **A2** hold for H_+ and so the KAM step can be iterated.

Proof of the iteration lemma. Our KAM step is standard and we divide it into several parts.

A. Truncation. Let $R = P(\xi,\lambda;\theta,0) + \langle P_I(\xi,\lambda;\theta,0), I\rangle$. It follows easily that $\|R\|_{M\times D(s,r)} \le 2\|P\|_{M\times D(s,r)} \le 2\epsilon$. Let

$$R = \sum_{k\in\mathbb{Z}^n} R_k(\xi,\lambda;I)e^{\mathrm{i}\langle k,\theta\rangle}$$

and

$$R^K = \sum_{|k|\le K} R_k(\xi,\lambda;I)e^{\mathrm{i}\langle k,\theta\rangle}.$$

By definition, we have

$$\|R - R^K\|_{M\times D(s-\rho,r)} \le 2\epsilon e^{-K\rho}.$$

B. Construction of the symplectic mapping. The symplectic mapping is generated by a hamiltonian flow mapping at time 1, that is, $\Phi = X_F^t|_{t=1}$, where F is the generation function. It follows that

$$H\circ\Phi = N_+ + \{N,F\} + R^K - [R] + P_+,$$

where $[R]$ denotes the average of R on T^n, $N_+ = N + [R] = \langle I, \Omega_+(\xi,\lambda)\rangle$, $\{\cdot,\cdot\}$ is the Poisson bracket, and

$$P_+ = (R - R^K) + \int_0^1 \{(1-t)\{N,F\} + R, F\} \circ X_F^t \, dt + (P - R) \circ \Phi.$$

We choose F such that

$$\{N, F\} + R^K - [R] = 0. \tag{2.7}$$

It follows that

$$\Omega : (\xi,\lambda) \in U(\Gamma,\delta) \to \Omega(\xi,\lambda) \in O_{\alpha,\tilde{\delta}}.$$

Thus, we have

$$|\langle \Omega(\xi,\lambda), k\rangle| \geq \frac{1}{2}\frac{\alpha}{|k|^\tau}, \ \forall\, (\xi,\lambda) \in U(\Gamma,\delta), \ \forall\, 0 < |k| \leq K.$$

Let $\{F_k\}$ and $\{R_k\}$ be the Fourier coefficients of F and R with respect to θ. Thus

$$F_k = \frac{1}{\mathrm{i}\langle \Omega(\xi,\lambda), k\rangle} R_k, \ 0 < |k| \leq K$$

and $F_k = 0$ with $k = 0$ or $|k| > K$.

We have

$$\|F\|_{U(\Gamma,\delta)\times D(r,s-\rho)} \leq \frac{c\epsilon}{\alpha\rho^{\tau+n}}.$$

C. Estimates for the symplectic mapping. It follows that

$$\|WX_F\|_{U(\Gamma,\delta)\times D(r,s-2\rho)} \leq \frac{c\epsilon}{\alpha r \rho^{\tau+n+1}} = cE.$$

Thus, if $0 < \eta \leq \frac{1}{8}$ and $cE \leq \frac{1}{8}$, then for all $(\xi,\lambda) \in U(\Gamma,\delta)$ we have

$$\Phi(\xi,\lambda;\cdot,\cdot) = X_F^1 : D(r\eta, s-3\rho) \to D(2r\eta, s-2\rho).$$

Combining this with Cauchy's estimate, we have

$$\|W(\Phi - \mathrm{id})\|_{U(\Gamma,\delta)\times D(s-5\rho,\eta r)} \leq cE$$

and

$$\|W(\mathcal{D}\Phi - \mathrm{Id})W^{-1}\|_{U(\Gamma,\delta)\times D(s-5\rho,\eta r)} \leq cE.$$

Thus the estimates for Φ hold.

D. Estimates of the error terms. Since $\hat{h}(\xi,\lambda) = [P_I(\xi,\lambda;\cdot,0)]$, by the assumptions the estimate for $\hat{h}$ obviously holds. Let M_+ be defined as in conclusion **(i)**. Since the set O_α is closed, it follows easily that M_+ is also closed. Obviously, we have $\mathrm{dist}(M_+, \partial M) \geq \frac{1}{2}\delta$. By Cauchy's estimate, the estimates for $\hat{h}_\xi$ and $\hat{h}_\lambda$ follow easily.

Moreover, by (2.3) and the implicit function theorem, if

$$|h_{+\lambda}(\xi,\lambda)| \leq \frac{1}{2}, \forall (\xi,\lambda) \in M,$$

the equation

$$\Omega_+(\xi,\lambda) = \omega(\xi) + \lambda + h_+(\xi,\lambda) = \Omega$$

determines an analytic mapping

$$\lambda_+ : \xi \in \Pi_{\sigma_+} \to \lambda_+(\xi) \in B(0, 2d+1).$$

It is easy to see that statements (2.4) and (2.5) hold. By (2.6) we have that $U(\Gamma_\Omega^+, \delta_+) \subset M_+$. Thus, conclusion **(ii)** holds.

As usual, it follows that

$$\|P_+\|_{M_+\times D(s_+,r_+)} < c\left[\frac{\epsilon^2}{\alpha r\rho^{\tau+n+1}} + (\eta^2 + e^{-K\rho})\epsilon\right].$$

By the choice of the parameters, we have

$$\|P_+\|_{M_+\times D(s_+,r_+)} \le c\epsilon E \le c\alpha r_+\rho_+^{\tau+n+1}E^{\frac{3}{2}} = \alpha\ r_+\rho_+^{\tau+n+1}\ E_+,$$

where $E_+ = cE^{\frac{3}{2}}$ with c being a constant that depends only on n and τ. This implies conclusion **(i)**. □

Iteration. Now we choose some suitable parameters so that the above step can be iterated infinitely.

At the initial step, let $\rho_0 = s/20$, $r_0 = r$, $\epsilon_0 = \alpha r_0\rho_0^{\tau+n+1}E_0 = \epsilon$, and $\eta_0 = E_0{}^{\frac{1}{2}}$. Let K_0 satisfy $e^{-K_0\rho_0} = E_0$.

Assume the above parameters are all well defined for j. Then, we define $\rho_{j+1} = \frac{1}{2}\rho_j$, $\eta_j = E_j^{\frac{1}{2}}$, $r_{j+1} = \eta_j r_j$ and $E_{j+1} = cE_j^{\frac{3}{2}}$; ϵ_{j+1}, η_{j+1}, K_{j+1} are defined similarly.

Let $M_0 = \Pi_\sigma \times B(0, 2d+1)$ and $D_0 = D(s_0, r_0)$. Let $H_0 = H$. By the iteration lemma, we have a sequence of closed sets $\{M_j\}$ with $M_{j+1} \subset M_j$ and a sequence of symplectic mappings $\{\Phi_j\}$ such that for each $(\xi,\lambda) \in M_{j+1}$, $\Phi_j(\xi,\lambda;\cdot,\cdot) : D_{j+1} \to D_j$, where $D_j = D(s_j, r_j)$. Moreover, we have the following estimates:

$$\|W_j(\Phi_j - \mathrm{id})\|_{M_{j+1}\times D_{j+1}} \le cE_j$$

and

$$\|W_j(\mathcal{D}\Phi_j - \mathrm{Id})W_j^{-1}\|_{M_{j+1}\times D_{j+1}} \le cE_j.$$

Let $\Phi^j = \Phi_0 \circ \Phi_1 \circ \cdots \Phi_{j-1}$ with $\Phi^0 = \mathrm{id}$ and

$$H_j = H \circ \Phi^j = N_j + P_j,$$

where $N_j(\xi,\lambda;I) = \langle \Omega_j(\xi,\lambda), I\rangle$ with $\Omega_j(\xi,\lambda) = \omega(\xi)+\lambda+h_j(\xi,\lambda)$. Let $\tilde{\delta}_j = \frac{\alpha}{2K_j^{\tau+1}}$, $\delta_j = \frac{2}{3}\tilde{\delta}_j$ and $\sigma_j = \sigma_{j-1} - \frac{1}{2}\delta_{j-1}$ with $\sigma_0 = \sigma$. From the iteration lemma we know that for $\Omega \in O_\alpha$ the equation

$$\Omega_j(\xi,\lambda) = \omega(\xi) + \lambda + h_j(\xi,\lambda) = \Omega$$

on M_j defines implicitly an analytic mapping $\lambda = \lambda_j(\xi)$, $\xi \in \Pi_{\sigma_j}$, whose graph in M_j forms an analytic curve Γ_Ω^j.

Let $\Gamma_j = \bigcup_{\Omega\in O_\alpha}\Gamma_\Omega^j$. We have

$$M_{j+1} = \big\{(\xi,\lambda') \in \mathbb{C}^n\times\mathbb{C}^n \mid \xi \in \Pi_{\sigma_{j+1}}, (\xi,\lambda) \in \Gamma^j, |\lambda'-\lambda| \le \frac{1}{2}\delta_j\big\}.$$

Obviously, it follows that $M_{j+1} \subset M_j$ and $\mathrm{dist}(M_{j+1}, \partial M_j) \ge \frac{1}{2}\delta_j$.

Let

$$\hat{h}_j(\xi,\lambda) = \Omega_{j+1}(\xi,\lambda) - \Omega_j(\xi,\lambda).$$

Then we have

$$|\hat{h}_j(\xi,\lambda)| \le \frac{\epsilon_j}{r_j},\quad \forall\ (\xi,\lambda) \in M_j$$

and

$$|\hat{h}_{j\xi}(\xi,\lambda)| + |\hat{h}_{j\lambda}(\xi,\lambda)| \le \frac{2\epsilon_j}{\delta_j r_j},\quad \forall\ (\xi,\lambda) \in M_{j+1}.$$

Furthermore, we have

$$|\lambda_{j+1}(\xi) - \lambda_j(\xi)| \leq \frac{2\epsilon_j}{r_j}, \ \forall(\xi,\lambda) \in M_{j+1}.$$

Obviously it follows that

$$\|P_j\|_{M_j \times D_j} \leq \epsilon_j = \alpha \rho_j^{\tau+n+1} r_j E_j.$$

Convergence of the iteration. Now we prove convergence of the KAM iteration. In the same way as in [8, 15], it follows that if $c^{\frac{1}{2}} E_0 \leq \frac{1}{2}$, then

$$\|W_0 \mathcal{D}\Phi^j W_j^{-1}\|_{M_j \times D_j} \leq \prod_{i=1}^{j} (1 + cE_j) < 2.$$

So, we have

$$\|W_0(\Phi^j - \Phi^{j-1})\|_{M_j \times D_j} \leq cE_j$$

and

$$\|W_0 \mathcal{D}(\Phi^j - \Phi^{j-1})\|_{M_j \times D_j} \leq cE_j.$$

Let $D_* = D(0, \frac{1}{2}s)$, $M_* = \bigcap_{j \geq 0} M_j$ and $\Phi = \lim_{j\to\infty} \Phi^j$. Thus, we have

$$\|W_0(\Phi - \mathrm{id})\|_{M_* \times D_*} \leq cE_0$$

and

$$\|W_0(D\Phi - \mathrm{Id})\|_{M_* \times D_*} \leq cE_0.$$

Since Φ^j is affine in I, we have convergence of Φ^j to Φ on $D(r/2, s/2)$ and

$$\|W_0(\Phi - \mathrm{id})\|_{M_* \times D(s/2, r/2)} \leq cE_0.$$

Now we consider convergence of $\{h_j\}$. Let $F_j = \frac{2\epsilon_j}{\delta_j r_j}$. It follows that

$$\frac{F_{j+1}}{F_j} = (\frac{1}{2})^n \frac{(K_{j+1}\rho_{j+1})^{\tau+1}}{(K_j\rho_j)^{\tau+1}} \frac{E_{j+1}}{E_j} = (\frac{1}{2})^n \frac{x_{j+1}^{\tau+1} e^{-x_{j+1}}}{x_j^{\tau+1} e^{-x_j}},$$

where $x_j = K_j \rho_j$. By the iteration $E_{j+1} = cE_j^{\frac{3}{2}}$, if E_0 is sufficiently small, E_j are all sufficiently small and so $K_j\rho_j$ are sufficiently large. Since the function $x^{\tau+1}e^{-x}$ is decreasing for $x > \tau + 1$, we can choose a small $E_0 > 0$ such that $\frac{F_{j+1}}{F_j} \leq \frac{1}{4}$ and $F_j \leq \frac{1}{4}, \forall j \geq 0$. Thus the assumption (2.3) holds. Moreover, $\frac{\delta_{j+1}}{\delta_j} = (\frac{1}{2})^{\tau+1} \frac{x_j}{x_{j+1}} \leq \frac{1}{4}$. Thus, the condition (2.6) holds.

Let $\sigma_* = \sigma - \frac{1}{2}\sum_{j=0}^{\infty} \delta_j$. It follows that $\sigma_* \geq \sigma - \frac{2}{3}\delta_0$. If E_0 is sufficiently small such that $\delta_0 \leq \sigma$, then we have $\sigma_* \geq \frac{1}{3}\sigma$. Thus $\Pi_{\sigma_*} \subset \bigcap_{j\geq 0} \Pi_{\sigma_j}$.

By iteration we have $h_j = \sum_{i=0}^{j-1} \hat{h}_i$. Combining this with the estimates for $\hat{h}_j$, we have that for all $(\xi,\lambda) \in M_j$,

$$|h_j(\xi,\lambda)| \leq \sum_{i=0}^{j-1} \frac{1}{2}\delta_i F_i \leq \frac{1}{2}\delta_0 \sum_{i=0}^{j-1} F_i \leq \delta_0 F_0 \leq 2\epsilon/r.$$

Similarly, it follows that for all $(\xi,\lambda) \in M_j$,

$$|h_{j\xi}(\xi,\lambda)| + |h_{j\lambda}(\xi,\lambda)| \leq \sum_{i=0}^{j-1} F_i \leq 2F_0 \leq 12(-\ln E_0)^{\tau+n+1} E_0.$$

So if E_0 is sufficiently small we have

$$|h_{j\xi}(\xi,\lambda)| + |h_{j\lambda}(\xi,\lambda)| \le \frac{1}{2}, \quad \forall(\xi,\lambda) \in M_j,$$

and so the assumption (2.2) holds for all j.

Let $h = \lim_{j\to\infty} h_j$. Then for $(\xi,\lambda) \in M_*$ we have

$$|h(\xi,\lambda)| \le \frac{2\epsilon}{r}$$

and

$$|h_\xi(\xi,\lambda)| + |h_\lambda(\xi,\lambda)| \le 12(-\ln E_0)^{\tau+n+1} E_0 \le \frac{1}{2}.$$

In the same way it is easy to show that $\{\lambda_j\}$ is also convergent on Π_{σ_*}. In fact, we can choose E_0 sufficiently small such that $F_j \le \frac{1}{4}$ for all $j \ge 0$. Then for $i > j$ it follows that

$$|\lambda_i(\xi) - \lambda_j(\xi)| \le \sum_{l=j}^{i-1} F_l \delta_l \le 2F_j\delta_j \le \frac{\delta_j}{2}.$$

Let $\lambda_j(\xi) \to \lambda(\xi), \xi \in \Pi_{\sigma_*}$. Since $\Gamma_\Omega^j = \{(\xi,\lambda_j(\xi)) \mid \xi \in \Pi_{\sigma_j}\} \subset M_j$ and λ_j are all analytic on Π_{σ_*}, so is the limit λ. Let $i \to \infty$, and then we have

$$|\lambda(\xi) - \lambda_j(\xi)| \le \frac{\delta_j}{2}.$$

This implies that $\Gamma_\Omega^* = \{(\xi,\lambda(\xi)) \mid \xi \in \Pi_{\sigma_*}\} \subset M_j$ and so $\Gamma^* = \bigcup_{\Omega\in O_\alpha} \Gamma_\Omega^* \subset M_j$. Hence, $\Gamma^* \subset M_* = \bigcap_{j\ge 0} M_j$. Obviously, for $(\xi,\lambda) \in \Gamma_\Omega^*$ we have

$$\lambda + \omega(\xi) + h(\xi,\lambda) = \Omega.$$

In the same way as in [15] we can prove that h and P_* are C^∞-smooth with respect to (ξ,λ) on M_* in the Whitney sense. By Whitney's extension theorem [14], we can extend h and P_* to be C^∞-smooth on M, but this makes sense only on M_* for our problem.

Acknowledgment

We would like to thank the referee for some helpful comments that led to this revised version.

References

[1] V. I. Arnold. Proof of a theorem of A. N. Kolmogorov on the persistence of quasi-periodic motions under small perturbations of the Hamiltonian, *Russ. Math. Surv.* **18:5**(1963), 9-36. MR0163025 (29:328)

[2] V. I. Arnold. *Dynamical systems*, Encyclopaedia of Mathematical Sciences, 3, Springer-Verlag, 1988. MR0923953 (88m:58043)

[3] H. W. Broer, G. B. Huitema, F. Takens. Unfoldings of quasi-perodic tori. *Mem. Amer. Math. Soc.* **83:421**(1990), 1-81. MR1041003

[4] C.-Q. Cheng. Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems. *Comm. Math. Phys.* **177**(1996), 529-559. MR1385075 (97a:58161)

[5] L.H. Eliasson. Perturbations of stable invariant tori for Hamiltonian systems. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (4)* **15**(1988), 115-147. MR1001032 (91b:58060)

[6] A. N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR* **98:4**(1954), 527-530. MR0068687 (16:924c)

[7] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.* **169**(1967), 136-176. MR0208078 (34:7888)

[8] J. Pöschel. A lecture on the classical KAM theorem. School on dynamical systems. May 1992.

[9] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.* **35**(1982), 653-696. MR668410 (84d:58039)

[10] H. Rüssmann. On twist hamiltonian. Talk in the Colloque International: Mécanique Céleste et Systémes hamiltoniens, Marseille, 1990.

[11] H. Rüssmann. Nondegeneracy in the perturbation theory of integrable dynamical systems, *Stochastics, Algebra and Analysis in Classical and Quantum Dynamics*, Math. and Its Appl., 59, 211-223, Kluwer Academic, 1990. MR1052709 (91j:58140)

[12] H. Rüssmann. Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *Regular and Chaotic Dynamics* **6:2**(2001), 119-204. MR1843664 (2002g:37083)

[13] M. B. Sevryuk. KAM-stable Hamiltonians. *J. Dynamics Control Systems* **1**(1995), 351-366. MR1354540 (96m:58222)

[14] H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.* **36**(1934), 63-89. MR1501735

[15] Junxiang Xu, Jiangong You. Gevrey-smoothness of invariant tori for analytic nearly integrable Hamiltonian systems under Rüssmann's non-degeneracy condition. *Journal of Differential Equations* **235**(2007), 609-622. MR2317497 (2008b:37100)

[16] J. Xu, J. You, Q. Qiu. Invariant tori for nearly integrable Hamiltonian systems with degeneracy. *Mathematische Zeitschrift* **226**(1997), 375-387. MR1483538 (98k:58196)

Department of Mathematics, Southeast University, Nanjing 210096, People's Republic of China

E-mail address: xujun@seu.edu.cn

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

E-mail address: jyou@nju.edu.cn