DE GRUYTER Adv. Nonlinear Stud. 2016; 16 (4):719–735

Research Article

Dongfeng Zhang*, Junxiang Xu and Hao Wu

On Invariant Tori with Prescribed Frequency in Hamiltonian Systems

DOI: 10.1515/ans-2015-5051
Received March 11, 2016; accepted May 4, 2016

Abstract: In this paper we are mainly concerned with the persistence of invariant tori with prescribed frequency for analytic nearly integrable Hamiltonian systems under the Brjuno–Rüssmann non-resonant condition, when the Kolmogorov non-degeneracy condition is violated. As it is well known, the frequency of the persisting invariant tori may undergo some drifts, when the Kolmogorov non-degeneracy condition is violated. By the method of introducing external parameters and rational approximations, we prove that if the Brouwer topological degree of the frequency mapping is nonzero at some Brjuno–Rüssmann frequency, then the invariant torus with this frequency persists under small perturbation.

Keywords: KAM Theory, Hamiltonian Systems, Invariant Tori, Non-Degeneracy Condition, Non-Resonant Condition

MSC 2010: 70H08, 37J40, 34C27

Communicated by: Russell Johnson

1 Introduction and Main Results

Consider the following Hamiltonian dynamical system:

$$\begin{cases} \dot{x} = H_y(x,y) = N_y(y) + P_y(x,y), \\ \dot{y} = -H_x(x,y) = -P_x(x,y), \end{cases} \tag{1.1}$$

where $H(x,y) = N(y) + P(x,y)$ is the Hamiltonian function, and $(x,y) \in \mathbb{T}^n \times D$, with $\mathbb{T}^n$ being the usual n-dimensional torus and D a bounded connected open domain of $\mathbb{R}^n$. Suppose $N(y)$ and $P(x,y)$ are real analytic on the complex neighborhood of D and $\mathbb{T}^n \times D$.

If $P = 0$, the Hamiltonian system (1.1) is integrable and has invariant tori $\mathbb{T}^n \times \{y_0\}$ for all $y_0 \in D$, on which there exists a linear flow $x(t) = x_0 + \omega(y_0)t$, $y(t) = y_0$ with the frequency $\omega(y_0) = N_y(y_0)$.

If $P \neq 0$, the Hamiltonian system is in general no longer integrable, and we are interested to know whether these invariant tori persist under small perturbation.

The classical KAM theorem asserts that if the frequency $\omega(y)$ satisfies the Kolmogorov non-degeneracy

$$\det\Big(\frac{\partial\omega}{\partial y}\Big) = \det\Big(\frac{\partial^2 N}{\partial y^2}\Big) \neq 0 \tag{1.2}$$

and the Diophantine condition

$$|\langle k, \omega(y)\rangle| \geq \frac{\alpha}{|k|^\tau} \quad \text{for all } 0 \neq k \in \mathbb{Z}^n, \tag{1.3}$$

***Corresponding author: Dongfeng Zhang:** Department of Mathematics, Southeast University, Nanjing 210096, P. R. China, e-mail: zhdf@seu.edu.cn
Junxiang Xu, Hao Wu: Department of Mathematics, Southeast University, Nanjing 210096, P. R. China, e-mail: xujun@seu.edu.cn, haow@seu.edu.cn

where $\alpha > 0$, $\tau > n - 1$, then most of the invariant tori can persist, being only slightly deformed, when the perturbation P is sufficiently small.

The natural question to ask is whether these conditions (1.2) and (1.3) are necessary for the persistence of invariant tori in Hamiltonian systems.

During the development of Hamiltonian systems and KAM theory, a lot of scholars are dedicated to weakening the non-degeneracy condition and non-resonant condition. First, there are various well-known results on the degenerate Hamiltonian systems, i.e. $\det(\partial\omega/\partial y) = 0$. Brjuno [5] proved that the majority of the invariant tori of unperturbed Hamiltonian systems are preserved if $\operatorname{rank}(\omega, \partial\omega/\partial y) = n$. Cheng and Sun [10] obtained the persistence of invariant tori under the following assumptions:

(i) $\operatorname{rank}(\partial\omega/\partial y) = m$ for all $y \in D$, where $0 < m < n$,
(ii) there exists a twist curve on the range of any neighborhood of y_0 for all $y_0 \in D$, where twist curve means that its every curvature component is not zero.

Rüssmann [22] announced the following result: system (1.1) possesses many invariant tori, if $\omega(y)$ do not lie in any hyperplane passing through the origin, i.e.

$$a_1\omega_1(y) + a_2\omega_2(y) + \cdots + a_n\omega_n(y) \not\equiv 0 \quad \text{on } D \text{ for all } (a_1, \ldots, a_n) \in \mathbb{R}^n \setminus \{0\}. \tag{1.4}$$

Condition (1.4) is usually called the Rüssmann non-degeneracy condition and it is the sharpest one for the KAM theorem. Its detailed proof was given in [23, 25]. Later, Xu, You and Qiu [30] demonstrated that in the real analytic case, the Rüssmann non-degeneracy condition is equivalent to the fact that there exists a positive integer m depending on $N(y)$ and D such that

$$\operatorname{rank}\left\{\omega(y), \frac{\partial^\beta \omega(y)}{\partial y^\beta} : |\beta| \le m\right\} = n$$

for all $y \in D$. Especially, for the case $m = 1$, condition (1.4) is equivalent to the Kolmogorov non-degeneracy condition. Moreover, from [30] we have that the Rüssmann non-degeneracy condition is also equivalent to the fact that there exists a point $y_0 \in D$ such that

$$\operatorname{rank}\left\{\omega(y_0), \frac{\partial^\beta \omega(y_0)}{\partial y^\beta} : |\beta| \le n-1\right\} = n.$$

On the other hand, the Diophantine condition (1.3) can also be weakened. For the case of dimension two, Hanßmann and Si [14] investigated the existence of quasi-periodic solutions of non-autonomous two-dimensional reversible and Hamiltonian systems under the Brjuno condition defined as

$$\sum_{k=0}^{\infty} \frac{\log q_{k+1}}{q_k} < \infty, \tag{1.5}$$

where $\{q_k\}$ are the denominators of the continued fraction expansion of the rotation number ω. Gentile [13] showed that the invariant curves of analytic exact symplectic twist maps of the cylinder with Brjuno rotation number are preserved under small perturbation. For the high-dimensional case, Rüssmann [24] and Pöschel [21] proved the persistence of invariant tori of vector fields and Hamiltonian systems under the Brjuno–Rüssmann non-resonant condition

$$|\langle k, \omega(y)\rangle| \ge \frac{\alpha}{\Delta(|k|)} \quad \text{for all } k \in \mathbb{Z}^n \setminus \{0\}, \tag{1.6}$$

where $\alpha > 0$, and Δ is called Brjuno–Rüssmann approximation function. These are continuous, increasing, unbounded functions $\Delta : [1, \infty) \to [1, \infty)$ such that $\Delta(1) = 1$ and

$$\int_1^{\infty} \frac{\ln \Delta(t)}{t^2}\, dt < \infty.$$

Recently, Bounemoura and Fischler [3, 4] used rational approximations to prove the same results.

Remark. The non-resonant condition (1.6) is equivalent to the Brjuno condition (1.5) in the case of the plane. Moreover, if we choose $\Delta(t) = t^\tau$, the Brjuno–Rüssmann non-resonant condition (1.6) corresponds to the Diophantine condition (1.3).

Recently, the Liouvillean frequency received a lot of attention in the literature on small divisor problems arising in the context of analytic dynamical systems. Let $\mu \in (0,1)$ be irrational and denote by p_n/q_n the n-th convergence of μ. Define

$$L(\mu) = \limsup_{n\to\infty} \frac{\ln q_{n+1}}{q_n}. \tag{1.7}$$

Then $L(\mu)$ measures how Liouvillean μ is. Notice that $L(\mu) = 0$, if μ satisfies the Diophantine condition (1.3) or Brjuno conditions (1.5) and (1.6). Moreover, $L(\mu)$ has an equivalent definition:

$$\limsup_{|k|\to\infty} \frac{1}{|k|} \ln \frac{1}{|e^{2\pi i k\mu} - 1|} = L(\mu).$$

For the well-known results on Liouvillean frequency, Avila, Fayad and Krikorian [2] proved the rotation reducibility of $SL(2,\mathbb{R})$ cocycles with one frequency μ, irrespective of any Diophantine condition on the base dynamics. Hou and You [16] proved that a quasi-periodic linear differential equation in $sl(2,\mathbb{R})$ with two frequencies $(1,\mu)$, where μ is irrational, is almost reducible provided that the coefficients are analytic and close to constant.

In all the above results in [3, 4, 13, 14, 21, 24], the frequency satisfies the Kolmogorov non-degeneracy condition. Under this condition, the frequency can be regarded as parameters. However, under the Rüssmann non-degeneracy condition, the frequency can not be regarded as independent parameters, so the previous methods in [3, 4, 13, 14, 21, 24] are not valid. By an improved KAM iteration with parameters, we proved that the results in [21, 24] also hold under the Rüssmann non-degeneracy condition [34, 35], but the frequency of invariant tori may undergo some drifts.

In fact, the difference between the two kinds of non-degeneracy conditions is that under the former condition the invariant tori with prescribed frequency can persist, but under the latter condition one can only get the existence of a family of invariant tori, while there is no information on the persistence of any torus with given frequency.

In this paper, without assuming the Kolmogorov non-degeneracy condition, we are mainly concerned with the persistence of invariant tori with prescribed frequency, which satisfies the Brjuno–Rüssmann non-resonant condition. The method of rational approximations in [3, 4] is closely related to the fact that the frequency of each KAM step remains the same. But under the Rüssmann non-degeneracy condition, the frequency of invariant tori may undergo some drifts. So the methods in [3, 4] can not be directly applied. By the method of introducing external parameter and rational approximations, we will prove that if the Brouwer topological degree of the frequency mapping is nonzero at some Brjuno–Rüssmann frequency, then the invariant torus with this frequency persists under small perturbation.

First, with any approximation function Δ we define two other functions

$$\Lambda(Q) = Q\Delta(Q), \quad \text{and} \quad \Lambda^{-1}(t) = \sup\{Q \ge 1 : \Lambda(Q) \le t\} \quad \text{for } t \ge \Lambda(1).$$

The following theorem is the main result of this paper.

Theorem 1.1. *Let $H(x,y) = N(y) + P(x,y)$ be real analytic on the complex neighborhood of $\mathbb{T}^n \times D$, with the frequency $\omega(y) = N_y(y)$. Suppose that $\omega_0 = \omega(y_0)$, $y_0 \in D$, is the prescribed frequency, satisfying the Brjuno–Rüssmann non-resonant condition*

$$|\langle k, \omega_0\rangle| \ge \frac{\alpha}{\Delta(|k|)} \quad \text{for all } k \in \mathbb{Z}^n \setminus \{0\},$$

where the function Δ satisfies

$$\frac{C}{\ln 2} \int_Q^\infty \frac{\ln \Lambda(t)}{t^2} dt < \frac{s}{2},$$

and the Brouwer degree of the frequency mapping $\omega(y)$ at ω_0 on D is not zero, i.e. $\deg(\omega(y), D, \omega_0) \neq 0$. *Then, if the perturbation P is sufficiently small, the Hamiltonian system* (1.1) *has an invariant torus with ω_0 as its frequency.*

Remark. The power function t^λ, $\lambda > nm - 1$, and exponent function e^{t^λ}, $0 < \lambda < 1$ can be chosen as a Brjuno–Rüssmann approximation function.

Remark. Recently, the authors [2, 16] obtained a series of important results on the reducibility of quasi-periodic cocycles in the case of Liouvillean frequency. We conjecture that when $L(\mu) = 0$, the persistence of invariant tori with Liouvillean frequency $\omega = (1, \mu)$, μ is an irrational number, in Hamiltonian systems can also be obtained.

At first we introduce some parameters and change the Hamiltonian system (1.1) to a parameterized system, and at the end we take some suitable parameters to guarantee the convergence of KAM iteration. For any $\xi \in D$, we expand $N(y)$ in a small neighborhood of ξ: writing $y = \xi + I$ for I close to zero, we get

$$N(y) = N(\xi) + \langle \nabla N(\xi), I \rangle + \int_0^1 (1-t)\langle \nabla^2 N(\xi + tI)I, I \rangle.$$

Let $x = \theta$. We can eventually define

$$H(\theta, I; \xi) = e(\xi) + \langle \omega(\xi), I \rangle + P(\theta, I; \xi),$$

where $e(\xi) = N(\xi)$, $\omega(\xi) = \nabla N(\xi)$, and

$$P(\theta, I; \xi) = \int_0^1 (1-t)\langle \nabla^2 N(\xi + tI)I, I \rangle + P(\theta, \xi + I).$$

Note that $\xi \in D$ are regarded as parameters, $e(\xi)$ is an energy constant, which is usually omitted, and $\omega : \xi \to \omega(\xi)$ is called frequency mapping.

As usual, we denote by $\mathbb{Z}$ and $\mathbb{Z}_+$ the set of integers and positive integers. We define the variables and parameters domain by

$$D(s, r) = \{(\theta, I) \in \mathbb{C}^n/\mathbb{Z}^n \times \mathbb{C}^n : |\operatorname{Im}\theta|_\infty \le s, |I|_1 \le r\},$$

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

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

where $\sigma > r > 0$ is a small constant.

Define Π_σ as the complex neighborhood of Π in $\mathbb{C}^n$ with the radius σ, that is

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

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

$$\begin{cases} \dot\theta = H_I = \omega(\xi) + P_I(\theta, I; \xi), \\ \dot I = -H_\theta = -P_\theta(\theta, I; \xi), \end{cases} \tag{1.8}$$

where

$$|H|_{D(s,r)\times\Pi_\sigma} = \sup_{(\theta,I;\xi)\in D(s,r)\times\Pi_\sigma} |H(\theta, I; \xi)|.$$

Thus the persistence of invariant tori for Hamiltonian system (1.1) is reduced to that of the family of invariant tori for Hamiltonian system (1.8) depending on parameters $\xi \in \Pi$.

Theorem 1.2. *Let $H(\theta, I; \xi) = \langle \omega(\xi), I\rangle + P(\theta, I; \xi)$ be real analytic on $D(s, r) \times \Pi_\sigma$, where $\Pi \subset \mathbb{R}^n$ is a bounded simply connected domain. Let $\omega_0 = \omega(\xi_0)$, $\xi_0 \in \Pi$. Suppose that ω_0 satisfies the Brjuno–Rüssmann non-resonant condition*

$$|\langle k, \omega_0\rangle| \geq \frac{\alpha}{\Delta(|k|)} \quad \textit{for all } k \in \mathbb{Z}^n \setminus \{0\},$$

where the function Δ satisfies

$$\frac{C}{\ln 2}\int_Q^\infty \frac{\ln \Lambda(t)}{t^2}\,dt < \frac{s}{2}.$$

Moreover, the Brouwer degree of the frequency mapping $\omega(\xi)$ at ω_0 on Π is not zero, i.e.

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

Then there exists a sufficiently small $\varepsilon > 0$ such that if $|P|_{D(s,r)\times\Pi_\sigma} \leq \varepsilon$, there exists $\xi_* \in \Pi$ such that the Hamiltonian system (1.8) *at $\xi = \xi_*$ has an invariant torus with ω_0 as its frequency.*

Remark. Another direction is to weaken the analyticity of Hamiltonian systems. Popov [18] obtained the persistence and Gevrey smoothness of invariant tori for Gevrey Hamiltonian systems. Wagener [26] gave a more general conclusion. We proved that the results in [18, 26] also hold under the Rüssmann non-degeneracy condition [28, 32, 33]. For the persistence or destruction of invariant tori for finitely differential Hamiltonian systems, we refer to [9, 11, 12, 19].

Remark. In [27], we consider the following reversible system with normal degenerate equilibrium point:

$$\begin{cases} \dot{x} = \omega_0 + Q(x)y + P_1(x, y, u, v), \\ \dot{y} = P_2(x, y, u, v), \\ \dot{u} = y_m^{2n_0-1} + v^{2n_0} + P_3(x, y, u, v), \\ \dot{v} = u + P_4(x, y, u, v), \end{cases} \tag{1.9}$$

where $(x, y, u, v) \in \mathbb{T}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}$, $(m \geq n + 1)$, $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$, $n_0 > 0$ is a positive integer, $Q(x)$ is an $n \times m$-matrix, and P_i $(i = 1, 2, 3, 4)$ are small perturbations. The corresponding involution G is

$$(x, y, u, v) \mapsto (-x, y, -u, v).$$

By the KAM method and the special structure of unperturbed nonlinear terms, we proved that the reversible system (1.9) has at least one n-dimensional invariant torus with ω_0 as its frequency, when ω_0 satisfies the Diophantine condition (1.3). By the methods in this paper, the above results can be generalized to the Brjuno–Rüssmann non-resonant condition (1.6).

2 Proof of the Main Results

It is effective to introduce an artificial external parameter γ and consider the following Hamiltonian system:

$$\begin{cases} \dot{\theta} = H_I = \omega(\xi) + \gamma + P_I(\theta, I; \xi), \\ \dot{I} = -H_\theta = -P_\theta(\theta, I; \xi), \end{cases} \tag{2.1}$$

where $H(\theta, I; \xi, \gamma) = \langle \omega(\xi) + \gamma, I\rangle + P(\theta, I; \xi)$. The Hamiltonian system (2.1) with $\gamma = 0$ returns to the original Hamiltonian system (1.8).

The idea of introducing parameters was proposed by Herman [15] and heavily employed later on by others in [6–8, 17, 20, 29, 31]. We will first give a KAM theorem for the Hamiltonian system (2.1) with parameters (ξ, γ) and then prove Theorem 1.2.

Let $d = \max_{\xi_1,\xi_2\in\Pi_\sigma} |\omega(\xi_1) - \omega(\xi_2)|$, and define

$$B(\omega, d) = \{\gamma \in \mathbb{C}^n : \operatorname{dist}(\gamma, \omega) < d\}.$$

Let $M = \Pi_\sigma \times B(0, 2d + 1)$. The Hamiltonian $H(\theta, I; \xi, \gamma)$ is real analytic on $D(s, r) \times M$.

Suppose that $\omega_0 = \omega(\xi_0)$ satisfies the Brjuno–Rüssmann non-resonant condition

$$|\langle k, \omega_0\rangle| \geq \frac{\alpha}{\Delta(|k|)} \quad \text{for all } k \in \mathbb{Z}^n \setminus \{0\},$$

where the function Δ satisfies

$$\int_1^\infty \frac{\ln \Delta(t)}{t^2} dt < \infty.$$

Let $h = \frac{\alpha}{2Q\Delta(Q)}$. Then for all $\omega \in B(\omega_0, h)$ it follows that

$$|\langle k, \omega\rangle| \geq \frac{\alpha}{2\Delta(|k|)} \quad \text{for all } 0 < |k| \leq Q. \tag{2.2}$$

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

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

then we have an analytic curve

$$\{\Gamma_{\omega_0} : y = y(\xi),\ \xi \in \Pi\} \subset M,$$

which is determined implicitly by the equation

$$\omega(\xi) + y + \hat{N}_*(\xi, y) = \omega_0,$$

where $|\hat{N}_*(\xi, y)| \leq 2\varepsilon/r$, $|\hat{N}_{*\xi}(\xi, y)| + |\hat{N}_{*y}(\xi, y)| \leq 1/2$, and a parameterized symplectic mapping

$$\Phi(\cdot, \cdot; \xi, y) : D(\frac{s}{2}, \frac{r}{2}) \to D(s, r), (\xi, y) \in \Gamma_{\omega_0},$$

where Φ is C^∞-smooth in (ξ, y) on Γ_{ω_0} in the sense of Whitney and analytic in (θ, I) on $D(s/2, r/2)$, such that for each $(\xi, y) \in \Gamma_{\omega_0}$, we have

$$H(\Phi(\theta, I; \xi, y); \xi, y) = \langle \omega_0, I\rangle + P_*(\theta, I; \xi, y),$$

where P_* satisfies $P_*(\theta, 0) = \partial_I P_*(\theta, 0) = 0$. Therefore, the Hamiltonian systems (2.1) *have an invariant torus $\Phi(\mathbb{T}^n, 0; \xi, y)$ with ω_0 as its frequency.*

Now we first use Theorem 2.1 to prove Theorem 1.2. In fact, we only need to prove that the external parameter y has at least one zero point, so that the Hamiltonian system (2.1) returns to the original Hamiltonian system (1.8). By the estimates of $\hat{N}_*(\xi, y)$ and using the implicit function theorem, the equation

$$\omega(\xi) + y + \hat{N}_*(\xi, y) = \omega_0$$

determines an analytic curve

$$y(\xi) = \omega_0 - \omega(\xi) + \hat{y}(\xi),$$

satisfying

$$|\hat{y}(\xi)| < \frac{2\varepsilon}{r}, \quad |\hat{y}_\xi(\xi)| \leq \frac{4\varepsilon}{r}.$$

Assuming

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

if we choose Q_0 sufficiently large such that ε is sufficiently small, we have

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

Then we have at least one $\xi_* \in \Pi$ such that $y(\xi_*) = 0$. Therefore, the Hamiltonian system (1.8) with

$$H(\theta, I; \xi_*) = H(\theta, I; \xi_*, y(\xi_*))$$

has an invariant torus with ω_0 as its frequency.

Now we begin to prove Theorem 2.1; its detailed proof consists of a KAM step, setting the parameters and iteration, and the convergence of iteration.

The idea is to use the method of introducing an external parameter to have a good control of frequency drift, so that we can obtain a Cantor-like family of analytic curves in KAM iteration, on which the frequency remains the same and satisfies the Brjuno–Rüssmann non-resonant condition. Every KAM iteration is carried out in the neighborhood of one curve, the radius of the neighborhood gradually tends to zero. When the radius of the neighborhood shrinks to zero, the family of curves can converge the curve, on which the frequency is prescribed and satisfies the Brjuno–Rüssmann non-resonant condition.

2.1 The KAM Step

In this section we describe our linear iterative scheme with respect to the Hamiltonian system (2.1) for one KAM step. Suppose we are now in the n-th step, and in what follows the quantities without subscripts refer to those at the n-th step, while the quantities with subscripts "+" denote the corresponding ones at the $(n+1)$-th step. We will use the same notation c to indicate different constants, which are independent of the iterative process.

Suppose that at the n-th step, the Hamiltonian system is written as

$$\begin{cases} \dot{\theta} = H_I = \omega(\xi) + y + \hat{N}(\xi, y) + P_I(\theta, I; \xi), \\ \dot{I} = -H_\theta = -P_\theta(\theta, I; \xi), \end{cases} \tag{2.3}$$

where $H = \langle \Omega(\xi, y), I \rangle + P(\theta, I; \xi, y)$, the frequency $\Omega(\xi, y) = \omega(\xi) + y + \hat{N}(\xi, y)$. We summarize one KAM step in the following lemma.

Lemma 2.2 (KAM Step). *Consider the real analytic Hamiltonian system* (2.3), *with*

$$\|P\|_{D(s,r)\times M} \le \varepsilon,$$

and

$$\frac{c\varepsilon}{r} < h, \quad h = \frac{\alpha}{2Q\Delta(Q)}, \quad Q\rho > 1, \tag{2.4}$$

where $0 < \rho < s$, $Q > 1$ is a sufficiently large constant, which will be chosen below. Let $\omega(\xi_0) = \omega_0 = (1, \bar{\omega}_0)$ be the prescribed Brjuno–Rüssmann non-resonant frequency (2.2). *Suppose that the function $\hat{N}(\xi, y)$ satisfies*

$$|\hat{N}_\xi(\xi, y)| + |\hat{N}_y(\xi, y)| \le \frac{1}{2} \quad \text{for all } (\xi, y) \in M \tag{2.5}$$

such that the equation

$$\omega(\xi) + y + \hat{N}(\xi, y) = \omega_0$$

defines implicitly an analytic curve

$$\Gamma : y = y(\xi), \quad \xi \in \Pi_\sigma \to y(\xi) \in B(0, 2d+1)$$

such that

$$\Gamma = \{(\xi, y(\xi)) : \xi \in \Pi_\sigma\} \subset M.$$

Moreover, for $h = \frac{\alpha}{2Q\Delta(Q)}$, $L = 2 + \max_{\xi\in\Pi_\sigma} |\omega_\xi(\xi)|$, we define $\delta = \frac{h}{L}$ such that

$$\mathrm{B}(\Gamma, \delta) = \{(\xi', y') \in \Pi_\sigma \times \mathbb{C}^n : |\xi' - \xi| + |y' - y(\xi)| \le \delta, (\xi, y) \in \Gamma\} \subset M.$$

Next, we define

$$\begin{gathered} s_+ = s - \rho, \quad r_+ = \eta r, \quad h_+ = \frac{1}{4}h, \quad \sigma_+ = \sigma - \frac{\delta}{2}, \\ \Lambda_+ = 2\Lambda(Q_0), \quad Q_+ = \Lambda^{-1}(\Lambda_+), \quad \rho_+ Q_+ = C \ge 1, \end{gathered}$$

where η is a fixed small constant. Then there exists

$$M_+ = \{(\xi', y') \in \Pi_{\sigma_+} \times \mathbb{C}^n : |\xi' - \xi| + |y' - y(\xi)| \le \frac{\delta}{2}, (\xi, y) \in \Gamma\} \subset M \tag{2.6}$$

such that for any $(\xi, y) \in M_+$, there exists a symplectic mapping $\Phi(\cdot, \cdot; \xi, y) : D(s_+, r_+) \to D(s, r)$, which transforms the Hamiltonian system (2.3) to

$$H_+ = \langle \Omega_+(\xi, y), I\rangle + P_+(\theta, I; \xi, y),$$

where the new frequency $\Omega_+(\xi, y) = \omega(\xi) + y + \hat{N}_+(\xi, y)$, $\hat{N}_+(\xi, y) = \hat{N}(\xi, y) + \Delta\hat{N}(\xi, y)$, the new perturbation P_+ and the drift term $\Delta\hat{N}(\xi, y)$ satisfy

$$|P_+|_{D(s_+,r_+)\times M_+} \le \varepsilon_+ = \frac{\eta}{8}\varepsilon$$

and

$$|\Delta\hat{N}(\xi, y)| \le \frac{\varepsilon}{r} \quad \text{for all } (\xi, y) \in M, \tag{2.7}$$

$$|\Delta\hat{N}_\xi(\xi, y)| + |\Delta\hat{N}_y(\xi, y)| \le \frac{2\varepsilon}{r\delta} \quad \text{for all } (\xi, y) \in M_+. \tag{2.8}$$

Moreover, the symplectic mapping Φ has the estimates

$$|W(\Phi - id)|_{D(s_+,r_+)\times M_+} \le \frac{c\Delta(Q)\varepsilon}{r\rho}, \quad |W(D\Phi - Id)W^{-1}|_{D(s_+,r_+)\times M_+} \le \frac{c\Delta(Q)\varepsilon}{r\rho},$$

where $W = \operatorname{diag}(\frac{1}{r}I_n, \frac{1}{\rho}I_n)$.

Thus if

$$\frac{2\varepsilon}{r\delta} \le \frac{1}{4}, \tag{2.9}$$

the equation

$$\omega(\xi) + y + \hat{N}(\xi, y) + \Delta\hat{N}(\xi, y) = \omega_0$$

defines implicitly an analytic curve

$$\Gamma_+ : y_+ = y_+(\xi) : \xi \in \Pi_{\sigma_+} \to y_+(\xi) \in \mathrm{B}(0, 2d + 1)$$

with $\sigma_+ = \sigma - \frac{1}{2}\delta$, satisfying

$$|y_+(\xi) - y(\xi)| \le \frac{2\varepsilon}{r} \le \frac{1}{4}\delta \tag{2.10}$$

and

$$\Gamma_+ = \{(\xi, y_+(\xi)) : \xi \in \Pi_{\sigma_+}\} \subset M_+. \tag{2.11}$$

If

$$\delta_+ \le \frac{1}{4}\delta, \tag{2.12}$$

then we have $\mathrm{B}(\Gamma_+, \delta_+) \subset M_+$.

Proof. We divide the proof of the KAM step into several steps. First notice that condition (2.4) implies the following inequalities:

$$h < \frac{c}{Q\Delta(Q)}, \quad \varepsilon < \frac{c\rho r}{\Delta(Q)}, \quad \varepsilon < \frac{cr}{Q\Delta(Q)}, \quad 1 < Q\rho. \tag{2.13}$$

A. Linear Approximation. Let R be the linearization of P in I at $I = 0$, i.e.

$$R(\theta, I) = P(\theta, 0) + \langle \partial_I P(\theta, 0), I\rangle.$$

Using the Cauchy estimate, it is easy to see that $|R|_{s,r} < c\varepsilon$. Moreover, using Lemma A.3, we have

$$|P - R|_{s,2\eta r} \le \frac{(2\eta)^2}{1 - 2\eta}\varepsilon \le \frac{\eta\varepsilon}{16}, \tag{2.14}$$

where η is small enough.

B. Rational Approximations. Since ω_0 is non-resonant, by Lemma A.1, given $Q_0 > 1$, there exist n rational vectors $v_1, \dots, v_n$ with denominators $q_1, \dots, q_n$, such that $q_1 v_1, \dots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$, and satisfy

$$|\omega_0 - v_j| < \frac{c}{q_j Q}, \quad 1 \le q_j \le c\Delta(Q), \quad 1 \le j \le n.$$

C. Extension of rational approximations. First note that the mapping

$$\Omega : (\xi, y) \in \mathrm{B}(\Gamma, \delta) \to \Omega(\xi, y) \in \mathrm{B}(\omega_0, h)$$

is well defined. In fact, for any $(\xi', y') \in \mathrm{B}(\Gamma, \delta)$, there exists $(\xi, y(\xi)) \in \Gamma$ such that

$$|\xi' - \xi| + |y' - y(\xi)| < \delta,$$

so that

$$\begin{aligned} |\Omega(\xi', y') - \omega_0| &= |\omega(\xi') + y' + \hat{N}(\xi', y') - (\omega(\xi) + y + \hat{N}(\xi, y))| \\ &\le (2 + \max_{\xi \in \Pi_\sigma} |\omega_\xi(\xi)|)\delta \le h. \end{aligned}$$

Thus, for any $\Omega(\xi, y) \in \mathrm{B}(\omega_0, h)$, using the conditions $h = \frac{\alpha}{2Q\Delta(Q)}$ and $q_j < c\Delta(Q)$, we have

$$\begin{aligned} |\Omega(\xi, y) - v_j| &\le |\Omega(\xi, y) - \omega_0| + |\omega_0 - v_j| < h + \frac{c}{q_j Q} \\ &< \frac{\alpha}{2Q\Delta(Q)} + \frac{c}{q_j Q} < \frac{c}{q_j Q}. \end{aligned}$$

Remark. As the classical KAM iteration in [10, 20, 21, 23, 25, 35], this step is equivalent to the following step.

C. Extension of the Small Divisor Estimates. We can extend the small divisor conditions to the neighborhood $\mathrm{B}(\Gamma, \delta)$. For any $(\xi', y') \in \mathrm{B}(\Gamma, \delta)$, $0 < |k| < Q$, we have

$$|\langle k, \Omega(\xi', y')\rangle| \ge \frac{\alpha}{2\Delta(|k|)}. \tag{2.15}$$

In fact, for any $(\xi', y') \in \mathrm{B}(\Gamma, \delta)$, there exists $(\xi, y(\xi)) \in \Gamma$ such that $|\xi' - \xi| + |y' - y(\xi)| \le \delta$. So it follows that

$$\begin{aligned} |\langle k, \Omega(\xi', y') - \omega_0\rangle| &\le |k||\omega(\xi') + y' + \hat{N}(\xi', y') - (\omega(\xi) + y + \hat{N}(\xi, y))| \\ &\le |k|h \le \frac{\alpha}{2\Delta(Q)}. \end{aligned}$$

Note that ω_0 satisfies the Brjuno–Rüssmann non-resonant condition. This proves claim (2.15).

D. Construction of the Symplectic Mappings. Let $P_1 = R$, and define inductively

$$P_{j+1} = [P_j]_{v_j} = [\cdots [R]_{v_1} \cdots]_{v_j}, \quad 1 \le j \le n,$$

where $[\cdot]_{v_j}$ denotes the average in the direction of v_j. The symplectic mapping Φ is the composition of the time-1 mapping of the Hamiltonian flow F_j, i.e.

$$\Phi = X^1_{F_1} \circ \cdots \circ X^1_{F_n}.$$

In the following we will construct the symplectic mapping by induction. For $j = 1$, write $\bar{H} = N_1 + S_1 + P_1$, where $N_1 = \langle v_i, I\rangle$, $S_1 = \langle \Omega(\xi, y), I\rangle - \langle v_1, I\rangle$, $P_1 = R$. Next we expand $\bar{H} \circ X^1_{F_1}$ with respect to t at zero, and write

$$\begin{aligned}\bar{H}\circ X^1_{F_1} &= (N_1+S_1+P_1)\circ X^1_{F_1}\\ &= N_1+\{N_1,F_1\}+\int_0^1(1-t)\{\{N_1,F_1\},F_1\}\circ X^t_{F_1}\,dt\\ &\quad + S_1+\int_0^1\{S_1,F_1\}\circ X^t_{F_1}\,dt+P_1+\int_0^1\{P_1,F_1\}\circ X^t_{F_1}\\ &= N+\{N_1,F_1\}+P_1+\int_0^1\{(1-t)\{N_1,F_1\}+P_1+S_1,F_1\}\circ X^t_{F_1}\,dt.\end{aligned}$$

The point is to find F_1 such that

$$\{N_1,F_1\}+P_1=[R]_{v_1}=P_2. \tag{2.16}$$

Equation (2.16) can be solved without Fourier expansions by the following integral formula:

$$F_1(\theta,I)=q_1\int_0^1(P_1-P_2)(\theta+tq_1v_1,I)t\,dt.$$

Then, $(1-t)\{N_1,F_1\}+P_1+S_1=(1-t)P_2+tP_1+S_1$, and altogether we obtain

$$\bar{H}\circ\Phi=N+P_2+P_1^+,$$

where

$$P_1^+=\tilde{P}_1=\int_0^1\{(1-t)P_2+tP_1+S_1,F_1\}\circ X^t_{F_1}\,dt$$

will be put in the new perturbation.

Moreover, we have the estimates $|P_1|_{s,r}\le|R|_{s,r}<c\varepsilon$ and

$$|F_1|_{s,r}\le q_1|P_1|_{s,r}<c\Delta(Q)\varepsilon. \tag{2.17}$$

Next, for any $0\le j\le n$, define $s_j^+=s-\frac{j\rho}{2n}$ and $r_j^+=r-\frac{jr}{4n}$. Using (2.17) and the Cauchy estimate, we have

$$|\partial_\theta F_1|_{s_1^+,r_1^+}\le\frac{c}{\rho}|F_1|_{s,r}\le\frac{c\Delta(Q)\varepsilon}{\rho}$$

and

$$|\partial_I F_1|_{s_1^+,r_1^+}\le\frac{c}{r}|F_1|_{s,r}\le\frac{c\Delta(Q)\varepsilon}{r}.$$

Let $X^1_{F_1}=(U_1,V_1)$. As R is affine in I, therefore U_1 is affine in I, V_1 is independent of I, and we have the estimates

$$|U_1-\mathrm{id}\,|_{s_1^+,r_1^+}\le|\partial_\theta F_1|_{s_1^+,r_1^+}\le\frac{c\Delta(Q)\varepsilon}{\rho},\quad |V_1-\mathrm{id}\,|_{s_1^+,r_1^+}\le|\partial_I F_1|_{s_1^+,r_1^+}\le\frac{c\Delta(Q)\varepsilon}{r}. \tag{2.18}$$

The Jacobian of $X^1_{F_1}$ is given by the matrix

$$DX^1_{F_1}=\begin{pmatrix}\partial_I U_1 & \partial_\theta U_1\\ 0 & \partial_\theta V_1\end{pmatrix}.$$

If we define $s_j=s-\frac{j\rho}{n}$ and $r_j=r-\frac{jr}{2n}$, then (2.18) and the Cauchy estimates imply that

$$|\partial_I U_1-\mathrm{Id}\,|_{s_1,r_1}\le\frac{c\Delta(Q)\varepsilon}{r\rho},\quad|\partial_\theta U_1|_{s_1,r_1}\le\frac{c\Delta(Q)\varepsilon}{\rho^2},\quad|\partial_\theta V_1-\mathrm{Id}\,|_{s_1,r_1}\le\frac{c\Delta(Q)\varepsilon}{r\rho}. \tag{2.19}$$

Together with the inequalities (2.4) and (2.18), we can ensure that

$$|U_1 - \mathrm{id}\,|_{s_1,r_1} \le \frac{r}{2n}, \quad |V_1 - \mathrm{id}\,|_{s_1,r_1} \le \frac{\rho}{n},$$

which implies that the time-1 mapping $X^1_{F_1}$ of the Hamiltonian flow of F_1 defines a symplectic real analytic mapping

$$X^1_{F_1} = (U_1, V_1) : D(s_1, r_1) \to D(s_2, r_2).$$

The estimates (2.18) and (2.19) can be conveniently written as

$$|W(X^1_{F_1} - \mathrm{id})|_{s_1,r_1} \le \frac{c\Delta(Q)\varepsilon}{r\rho}, \quad |W(DX^1_{F_1} - \mathrm{Id})W^{-1}|_{s_1,r_1} \le \frac{c\Delta(Q)\varepsilon}{r\rho},$$

where $W = \mathrm{diag}(r^{-1}\,\mathrm{Id}, \rho^{-1}\,\mathrm{Id})$.

Let $N_j = \langle v_j, I\rangle$, $S_j = \langle \Omega(\xi, y), I\rangle - \langle v_j, I\rangle$. By induction, we have

$$\begin{aligned}
\bar{H} \circ X^1_{F_1} \circ \cdots \circ X^1_{F_{j-1}} \circ X^1_{F_j} &= (N + P_j + P^+_{j-1}) \circ X^1_{F_j} = (N_j + S_j + P_j + P^+_{j-1}) \circ X_{F^1_j} \\
&= N_j + \{N_j, F_j\} + \int_0^1 (1-t)\{\{N_j, F_j\}, F_j\} \circ X^t_{F_j}\, dt \\
&\quad + S_j + \int_0^1 \{S_j, F_j\} \circ X^1_{F_j}\, dt + P_j + \int_0^1 \{P_j, F_j\} \circ X^t_{F_j}\, dt + P^+_{j-1} \circ X^1_{F_j} \\
&= N + \{N_j, F_j\} + P_j + \int_0^1 \{(1-t)\{N_j, F_j\} + P_j + S_j, F_j\} \circ X^t_{F_j}\, dt + P^+_{j-1} \circ X^1_{F_j}.
\end{aligned}$$

Similarly, the equation

$$\{N_j, F_j\} + P_j = P_{j+1} = [\cdots [R]_{v_1} \cdots]_{v_j}$$

can be solved by the integral formula

$$F_j(\theta, I) = q_j \int_0^1 (P_j - P_{j+1})(\theta + t q_j v_j, I) t\, dt,$$

which yields that

$$\bar{H} \circ X^1_{F_1} \circ \cdots \circ X^1_{F_j} = N + P_{j+1} + P^+_j,$$

where $P^+_j = \tilde{P}_j + P^+_{j-1} \circ X^1_{F_j}$ will be put in the new perturbation, and

$$\tilde{P}_j = \int_0^1 \{(1-t)P_{j+1} + tP_j + S_j, F_j\} \circ X^t_{F_j}\, dt$$

satisfies

$$|\tilde{P}_j|_{s_j,r_j} < \frac{c\Delta(Q)\varepsilon^2}{r\rho} + \frac{\varepsilon}{Q\rho} < \frac{c\varepsilon}{Q\rho}. \tag{2.20}$$

The symplectic mapping

$$X^1_{F_j} = (U_j, V_j) : D(s_j, r_j) \to D(s_{j+1}, r_{j+1})$$

is well defined, and

$$|W(X^1_{F_j} - \mathrm{id})|_{s_j,r_j} \le \frac{c\Delta(Q)\varepsilon}{r\rho}, \quad |W(DX^1_{F_j} - \mathrm{Id})W^{-1}|_{s_j,r_j} \le \frac{c\Delta(Q)\varepsilon}{r\rho},$$

where $W = \mathrm{diag}(r^{-1}\,\mathrm{Id}, \rho^{-1}\,\mathrm{Id})$.

Altogether, let $\Phi = X^1_{F_1} \circ \cdots \circ X^1_{F_n}$. Note that $\Phi : D(s-\rho, \frac{r}{2}) \to D(s, r)$, so $\Phi : D(s-\rho, r\eta) \to D(s, r)$ when η is small enough. Moreover, we have the estimates

$$|W(\Phi - \mathrm{id})|_{s-\rho,\eta r} < \frac{c\Delta(Q)\varepsilon}{r\rho}$$

and

$$|W(D\Phi - \mathrm{Id})W^{-1}|_{s-\rho,\eta r} < \frac{c\Delta(Q)\varepsilon}{r\rho}.$$

E. Construction of Parameters Domains. Let M_+ be defined by (2.6). It follows that M_+ is closed and

$$M_+ \subset \mathrm{B}(\Gamma, \delta) \subset M, \quad \mathrm{dist}(M_+, \partial M) \geq \frac{\delta}{2},$$

where ∂M is the boundary of M.

Let $\Delta\hat{N}(\xi, y) = [\partial_I P(\theta, 0)]$. By the Cauchy estimates, the estimates (2.7) and (2.8) hold. Set

$$\hat{N}_+(\xi, y) = \hat{N}(\xi, \eta) + \Delta\hat{N}(\xi, y).$$

By the implicit function theorem, if

$$\left|\frac{\partial \hat{N}_+}{\partial y}\right| \leq \frac{1}{2} \quad \text{for all } (\xi, y) \in M,$$

which will be verified in (2.22), the equation

$$\omega(\xi) + y + \hat{N}_+(\xi, y) = \omega_0$$

defines implicitly an analytic curve

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

Note that y_+ and y satisfy

$$\omega(\xi) + y_+ + \hat{N}_+(\xi, y) = \omega(\xi) + y + \hat{N}(\xi, y) = \omega_0.$$

Then it follows that

$$\begin{aligned} |y_+(\xi) - y(\xi)| &\leq |\hat{N}_+(\xi, y_+(\xi)) - \hat{N}(\xi, y(\xi))| \\ &\leq |\hat{N}(\xi, y_+(\xi)) - \hat{N}(\xi, y(\xi))| + |\Delta\hat{N}(\xi, y_+(\xi))| \\ &\leq \frac{1}{2}|y_+(\xi) - y(\xi)| + \frac{\varepsilon}{r}. \end{aligned}$$

Hence the conclusions (2.10) and (2.11) hold. By (2.12), we have $\mathrm{B}(\Gamma_+, \delta_+) \subset M_+$.

F. Estimates of the New Perturbations. As $H = N + P = N + R + (P - R) = \bar{H} + P - R$, this gives

$$H \circ \Phi = \bar{H} \circ \Phi + (P - R) \circ \Phi = N + P_{n+1} + P^+_n + (P - R) \circ \Phi.$$

By Lemma A.2, $P_{n+1} = [\cdots [R]_{v_1} \cdots]_{v_n} = [R]$, we arrive at

$$H \circ \Phi = N + [R] + P_+ = \langle \Omega_+(\xi, y), I \rangle + P_+,$$

where $\Omega_+(\xi, y) = \omega(\xi) + y + \hat{N}_+(\xi, y)$, $P^+ = P^+_n + (P - R) \circ \Phi$.

Concerning P^+_n, using (2.20) and the inequalities (2.13), we can ensure that

$$|P^+_n|_{s-\rho,\eta r} < \frac{c\varepsilon}{Q\rho} < \frac{\eta\varepsilon}{16}.$$

Using the inequalities (2.13), we may assume that the image of Φ actually sends $D(s-\rho, \eta r)$ into $D(s, 2\eta r)$, and together with (2.14), we obtain the estimates

$$\begin{aligned} |P_+|_{s-\rho,\eta r} &\leq |P^+_n|_{s-\rho,\eta r} + |(P - R) \circ \Phi|_{s-\rho,\eta r} \\ &\leq |P^+_n|_{s-\rho,\eta r} + |(P - R)|_{s,2\eta r} \leq \frac{\eta\varepsilon}{16} + \frac{\eta\varepsilon}{16} \leq \frac{\eta\varepsilon}{8}. \end{aligned}$$

Thus this ends the proof of the KAM step. □

2.2 Setting the Parameters and Iteration

In this section we choose some suitable parameters and verify some assumptions in Lemma 2.2 so that the above iteration can go on infinitely.

At the initial step, let

$$s_0 = s, \quad r_0 = r, \quad \rho_0 = \rho < s, \quad h_0 = \frac{\alpha}{2Q_0\Delta(Q_0)}, \quad \varepsilon_0 = \varepsilon,$$
$$\sigma_0 = \sigma, \quad Q_0\rho_0 = C \geq 1, \quad \delta_0 = \frac{h_0}{L} = \frac{\alpha}{2LQ_0\Delta(Q_0)},$$

where $L = 2 + \max_{\xi\in\Pi_\sigma} |\omega_\xi(\xi)|$, Q_0 is a sufficiently large constant, which will be chosen below.

Recall that η is a fixed small constant, and now we define the following sequences of parameters:

$$\varepsilon_i = \left(\frac{\eta}{16}\right)^i \varepsilon, \quad r_i = \eta^i r, \quad \Lambda_i = 4^i\Lambda(Q_0),$$
$$Q_i = \Lambda^{-1}(\Lambda_i) = \sup\{Q \geq 1 : \Lambda(Q) \leq \Lambda_i\},$$
$$h_i = \frac{\alpha}{2Q_i\Delta(Q_i)} \Leftrightarrow h_i = \left(\frac{1}{4}\right)^i h, \quad \delta_i = \frac{h_i}{L},$$
$$\sigma_{i+1} = \sigma_i - \frac{1}{2}\delta_i, \quad \rho_i = CQ_i^{-1}, \quad s_{i+1} = s_i - \rho_i,$$

where $L = 2 + \max_{\xi\in\Pi_\sigma} |\omega_\xi(\xi)|$, $C \geq 1$ is a sufficiently large constant, depending only on n.

In the following we will check all the assumptions in Lemma 2.2. By induction and the definition of ε_i, r_i and δ_i, we have

$$\frac{c\varepsilon_i}{r_i} < h_i, \quad \frac{\delta_{i+1}}{\delta_i} \leq \frac{1}{2}$$

and

$$\frac{2\varepsilon_i}{r_i\delta_i} = \frac{2L\varepsilon}{rh}\left(\frac{1}{4}\right)^i \leq \frac{1}{4} \quad \text{for all } i \geq 0.$$

Thus the assumptions (2.4), (2.9) and (2.12) hold.

Lemma 2.3 (Iterative Lemma). *Let $H_0 = \langle \omega(\xi) + y, I\rangle + P_0(\theta, I; \xi)$ be real analytic on $D(s, r) \times M$, satisfying*

$$|P_0|_{D(s,r)\times M} \leq \varepsilon_0$$

and

$$\frac{c\varepsilon_0}{r_0} < h_0, \quad h_0 = \frac{\alpha}{2Q_0\Delta(Q_0)},$$

where Q_0 is sufficiently large such that (2.21) *is satisfied. Then, for each $i \in \mathbb{N}$, there exists a real analytic symplectic mapping*

$$\Phi^i : D(s_i, r_i) \times M_i \to D(s, r) \times M$$

such that $H \circ \Phi^i = N_i + P_i$ with $N_i = \langle \Omega_i(\xi, y), I\rangle$, $\Omega_i(\xi, y) = \omega(\xi) + y + \hat{N}_i(\xi, y)$, and

$$|P_i|_{D(s_i,r_i)\times M_i} \leq \varepsilon_i = \left(\frac{\eta}{8}\right)^i \varepsilon.$$

Moreover, we have the estimates

$$|W_0(\Phi^{i+1} - \Phi^i)|_{D(s_{i+1},r_{i+1})\times M_{i+1}} < \frac{c\varepsilon_i}{r_ih_i},$$

where $W_0 = \operatorname{diag}(\frac{1}{r_0}\operatorname{Id}, \frac{1}{\rho_0}\operatorname{Id})$.

Proof. The proof follows by applying Lemma 2.2 repeatedly. So we only check the assumptions in Lemma 2.2 such that the KAM step is valid for any $i \geq 0$. By induction and the definition of the parameters, we obtain the following relations:

$$\frac{c\varepsilon_i}{r_i} < h_i, \quad \frac{2\varepsilon_i}{r_i\delta_i} \leq \frac{1}{4}, \quad \frac{\delta_{i+1}}{\delta_i} \leq \frac{1}{2} \quad \text{for all } i \geq 0.$$

□

2.3 Convergence of Iteration

We first verify that the sequence of s_i tends to a positive limit, if Q_0 is sufficiently large. Indeed, let $x = \Lambda^{-1}(2^y \Lambda(Q_0))$, we have

$$\sum_{i\geq 0} \frac{1}{Q_i} \leq \int_0^{+\infty} \frac{dy}{\Lambda^{-1}(2^y \Lambda(Q_0))} \leq \frac{1}{\ln 2} \int_0^{+\infty} \frac{d\Lambda(x)}{x\Lambda(x)}.$$

Integrating by parts and requiring $\Lambda(Q_0) \geq Q_0 > 1$, we get

$$\sum_{i\geq 0} \frac{1}{Q_i} \leq \frac{1}{\ln 2} \int_{Q_0}^{+\infty} \frac{\ln \Lambda(x)}{x^2} dx.$$

Now as $\rho_i = CQ_i^{-1}$, it follows that

$$\sum_{i\geq 0} \rho_i = C \sum_{i\geq 0} \frac{1}{Q_i} \leq \frac{C}{\ln 2} \int_{Q_0}^{+\infty} \frac{\ln \Lambda(x)}{x^2} dx.$$

Hence, by choosing Q_0 sufficiently large , we can achieve that

$$\sum_{i\geq 0} \rho_i \leq \frac{C}{\ln 2} \int_{Q_0}^{+\infty} \frac{\ln \Lambda(x)}{x^2} dx < \frac{s}{2}, \tag{2.21}$$

and thus

$$s_* = \lim_{i\to+\infty} s_i \geq \frac{s}{2}.$$

By the Iterative Lemma 2.3, we have

$$|W_0(\Phi^{i+1} - \Phi^i)|_{D_{i+1}\times M_{i+1}} \leq \frac{c\varepsilon_i}{r_i h_i}$$

and

$$|W_0 D(\Phi^{i+1} - \Phi^i)|_{D_{i+1}\times M_{i+1}} \leq \frac{c\varepsilon_i}{r_i h_i}.$$

Note that $D_* = D(s_*, 0)$, $M_* = \bigcap_{i\geq 0} M_i$, and $\Phi_* = \lim_{i\to\infty} \Phi^i$. Thus the mappings Φ^i converge uniformly on

$$\bigcap_{i\geq 0} D_i \times M_i = D(s_*, 0) \times M_*$$

to a mapping Φ_*, which is real analytic on $D(s_*, 0)$ and uniformly continuous on M_*. Moreover, we have

$$|W_0(\Phi_* - \mathrm{id})|_{D_*\times M_*} \leq \frac{c\varepsilon}{rh}$$

and

$$|W_0 D(\Phi_* - \mathrm{id})|_{D_*\times M_*} \leq \frac{c\varepsilon}{rh}.$$

Since Φ^i is affine in I, the symplectic mappings Φ^i converge to Φ_* on $D(\frac{s}{2}, \frac{r}{2})$, and satisfy

$$|W_0(\Phi_* - \mathrm{id})|_{D(\frac{s}{2}, \frac{r}{2})\times M_*} \leq \frac{c\varepsilon}{rh}.$$

Let $\sigma_* = \sigma - \frac{1}{2} \sum_{i=0}^{\infty} \delta_i$. It follows that $\sigma_* \geq \sigma - \frac{2}{3}\delta_0$. By the definition of δ_0, we can choose Q_0 sufficiently large such that $\delta_0 \leq \sigma$. Therefore $\sigma_* \geq \frac{1}{3}\sigma$, and $\Pi_{\sigma_*} \subset \bigcap_{i\geq 0} \Pi_{\sigma_i}$.

By iteration, we have

$$\hat{N}_i(\xi, y) = \sum_{k=0}^{i-1} \Delta \hat{N}_k(\xi, y).$$

Now we prove the convergence of $\hat{N}_i(\xi, y)$. Combining with estimates for $\Delta\hat{N}_i$, we have that for all $(\xi, y) \in M_i$,

$$|\hat{N}_i(\xi, y)| \le \sum_{k=0}^{i-1} \frac{\varepsilon_k}{r_k} = \sum_{k=0}^{i-1} \Big(\frac{1}{16}\Big)^k \frac{\varepsilon}{r} \le \frac{2\varepsilon}{r}.$$

Similarly, for all $(\xi, y) \in M_i$,

$$|\hat{N}_{i\xi}(\xi, y)| + |\hat{N}_{iy}(\xi, y)| \le \sum_{k=0}^{i-1} \frac{2\varepsilon_k}{r_k\delta_k} = \sum_{k=0}^{i-1} \frac{2L\varepsilon}{rh}\Big(\frac{1}{4}\Big)^k \le \frac{4L\varepsilon}{rh}.$$

Thus if Q_0 is sufficiently large such that ε is sufficiently small, we have

$$|\hat{N}_{i\xi}(\xi, y)| + |\hat{N}_{iy}(\xi, y)| \le \frac{1}{2} \quad \text{for all } (\xi, y) \in M_i, \tag{2.22}$$

and assumption (2.5) holds.

Let $\hat{N}_* = \lim_{i\to\infty} \hat{N}_i$. Then for all $(\xi, y) \in M_*$ we have

$$|\hat{N}_*(\xi, y)| \le \frac{2\varepsilon}{r}, \quad |\hat{N}_{*\xi}(\xi, y)| + |\hat{N}_{*y}(\xi, y)| \le \frac{1}{2}.$$

Similarly, we can prove the convergence of $y_i(\xi)$ on Π_{σ_*}. In fact, we can choose Q_0 sufficiently large such that $\frac{2\varepsilon_i}{r_i\delta_i} \le \frac{1}{4}$ for all $i \ge 0$. Then for $j \ge i$, it follows that

$$|y_j(\xi) - y_i(\xi)| \le \sum_{k=i}^{j-1} \frac{\varepsilon_k}{r_k} \le \frac{\delta_i}{2}.$$

Let $y_*(\xi) = \lim_{i\to\infty} y_i(\xi)$. Then we have

$$|y_*(\xi) - y_i(\xi)| \le \frac{\delta_i}{2},$$

which implies $\Gamma^* = \{(\xi, y_*(\xi)) : \xi \in \Pi_{\sigma_*}\} \subset M_i$, and therefore $\Gamma^* \subset M_* = \bigcap_{i\ge 0} M_i$. Moreover, for $(\xi, y) \in \Gamma^*$ we have

$$\omega(\xi) + y + \hat{N}_*(\xi, y) = \omega_0.$$

Therefore, for any $(\xi, y) \in \Gamma^*$ with ω_0 satisfying the Brjuno–Rüssmann non-resonant condition, the symplectic mapping $\Phi_*(\cdot, \cdot; \xi, y)$ transforms the Hamiltonian system (2.1) into

$$H_* = \langle \omega_0, I\rangle + P_*(\theta, I; \xi),$$

where P_* satisfies $P_*(\theta, 0) = \partial_I P_*(\theta, 0) = 0$. This completes the proof of Theorem 2.1.

A Appendix

In this section we formulate some technical lemmas which have been used in the previous section.

Lemma A.1. *Let $\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n$ be the Brjuno–Rüssmann non-resonant frequency, with $\bar{\omega}_0 \in [-1, 1]^{n-1}$. For any $Q > 1$, there exist n rational vectors $v_1, \dots, v_n$ with denominators $q_1, \dots, q_n$, such that $q_1 v_1, \dots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$, and satisfy*

$$|\omega_0 - v_j| \le \frac{c}{q_j Q}, \quad 1 \le q_j \le c\Delta(Q), \quad 1 \le j \le n.$$

Now given a q-rational vector v and a function P defined on $D(s, r) \times M$, we define

$$[P]_v(\theta, I; \xi, y) = \int_0^1 P(\theta + tqv, I; \xi, y)\, dt.$$

Similarly, given n rational vectors $v_1, \dots, v_n$, we define $[P]_{v_1,\dots,v_n} = [\cdots [P]_{v_1} \cdots]_{v_n}$, and

$$[P](I;\xi,\gamma) = \int_{\mathbb{T}^n} P(\theta, I; \xi, \gamma)\, d\theta.$$

The fact that the vectors $q_1 v_1, \dots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$ yields the following lemma.

Lemma A.2. *Let $v_1, \dots, v_n$ be rational vectors with denominators $q_1, \dots, q_n$, such that $q_1 v_1, \dots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$, and let P be a function defined on $D(s,r) \times M$. Then*

$$[\cdots [P]_{v_1} \cdots]_{v_n} = [P].$$

For the detailed proof of the above two lemmas, we refer to [3].

Lemma A.3. *Let P be an analytic function defined on $D(s,r) \times M$, and R be the linearization of P in I at $I = 0$, i.e.*

$$R(\theta, I) = P(\theta, 0) + \langle \partial_I P(\theta, 0), I \rangle.$$

Then, for any $0 < \eta < 1$, we have the estimate

$$|P - R|_{s,\eta r} \leq \frac{\eta^2}{1-\eta} |P|_{s,r}.$$

This lemma deals with the estimates of the remainder of the Taylor expansion at order one of an analytic function. A proof of this lemma can be found in [1].

Funding: The work was supported by the National Natural Science Foundation of China (11001048 and 11371090) and the Natural Science Foundation of Jiangsu Province of China (SBK201321532).

References

[1] J. Albrecht, An analytic KAM-theorem, preprint (2007), http://arxiv.org/abs/0705.0275.

[2] A. Avila, B. Fayad and R. Krikorian, A KAM scheme for SL(2, R) cocycles with Liouvillean frequencies, *Geom. Funct. Anal.* **21** (2011), 1001–1019.

[3] A. Bounemoura and S. Fischler, A Diophantine duality applied to the KAM and Nekhoroshev theorems, *Math. Z.* **275** (2013), 1135–1167.

[4] A. Bounemoura and S. Fischler, The classical KAM theorem for Hamiltonian systems via rational approximations, *Regul. Chaotic Dyn.* **19** (2014), 251–265.

[5] A. D. Brjuno, Analytic form of differential equations, *Trans. Moscow Math. Soc.* **25** (1971), 131–288.

[6] H. W. Broer and G. B. Huitema, Unfoldings of quasi-periodic tori in reversible systems, *J. Dynam. Differential Equations* **7** (1995), no. 1, 191–212.

[7] H. W. Broer, G. B. Huitema and M. B. Sevryuk, Families of quasi-periodic motions in dynamical systems depending on parameters, in: *Nonlinear Dynamical Systems and Chaos* (Groningen 1995), Progr. Nonlinear Differential Equations Appl. 19, Birkhäuser, Basel (1996), 171–211.

[8] H. W. Broer, G. B. Huitema, F. Takens and B. L. J. Braaksma, Unfoldings of quasi-periodic tori, *Mem. Amer. Math. Soc.* **83** (1990), no. 421, 1–175.

[9] C. Cheng, Non-existence of KAM torus, *Acta Math. Sin. (Engl. Ser.)* **27** (2011), 397–404.

[10] C. Cheng and Y. Sun, Existence of KAM tori in degenerate Hamiltonian systems, *J. Differential Equations* **114** (1994), 288–335.

[11] C. Cheng and L. Wang, Destruction of Lagrangian torus for positive definite Hamiltonian systems, *Geom. Funct. Anal.* **23** (2013), 848–866.

[12] L. Chierchia and D. Qian, Moser's theorem for lower dimensional tori, *J. Differential Equations* **206** (2004), 55–93.

[13] G. Gentile, Invariant curves for exact symplectic twist maps of the cylinder with Bryuno rotation numbers, *Nonlinearity* **28** (2015), 2555–2585.

[14] H. Hanßmann and J. Si, Quasi-periodic solutions and stability of the equilibrium for quasi-periodically forced planar reversible and Hamiltonian systems under the Bruno condition, *Nonlinearity* **23** (2010), 555–577.

[15] M. Herman, Dynamics connected with indefinite normal torsion, in: *Twist Mappings and Their Applications* (Minneapolis 1989/90), IMA Vol. Math. Appl. 44, Springer, New York (1992), 153–182.

[16] X. Hou and J. You, Almost reducibility and non-perturbative reducibility of quasi-periodic linear systems, *Invent. Math.* **190** (2012), 209–260.

[17] Y. V. Loveikin and I. O. Parasyuk, Invariant tori of locally Hamiltonian systems that are close to conditionally integrable systems, *Ukrainian Math. J.* **59** (2007), 70–99.

[18] G. Popov, KAM theorem for Gevrey Hamiltonians, *Ergodic Theory Dynam. Systems* **24** (2004), 1753–1786.

[19] J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, *Comm. Pure Appl. Math.* **35** (1982), 653–696.

[20] J. Pöschel, A Lecture on the classical KAM theorem, *Proc. Sympos. Pure Math.* **69** (2001), 707–732.

[21] J. Pöschel, KAM à la R, *Regul. Chaotic Dyn.* **16** (2011), 17–23.

[22] H. Rüssmann, Nondegeneracy in the perturbation theory of integrable dynamical systems, in: *Stochastics, Algebra and Analysis in Classical and Quantum Dynamics* (Marseille 1988), Math. Appl. 59, Kluwer, Dordrecht (1990), 211–223.

[23] H. Rüssmann, Invariant tori in non-degenerate nearly integrable Hamiltonian systems, *Regul. Chaotic Dyn.* **6** (2001), 119–204.

[24] H. Rüssmann, KAM iteration with nearly infinitely small steps in dynamical systems of polynomial character, *Discrete Contin. Dyn. Syst. Ser. S* **3** (2010), 683–718.

[25] M. B. Sevryuk, KAM stable Hamiltonians, *J. Dyn. Control Syst.* **1** (1995), 351–366.

[26] F. Wagener, A note on Gevrey regular KAM theory and the inverse approximation lemma, *Dyn. Syst.* **18** (2003), 159–163.

[27] X. Wang, J. Xu and D. Zhang, On the persistence of degenerate lower-dimensional tori in reversible systems, *Ergodic Theory Dynam. Systems,* **35** (2015), 2311–2333.

[28] J. Xu and J. You, Gevrey smoothness of invariant tori for analytic nearly integrable Hamiltonian systems under Russmann's non-degeneracy condition, *J. Differential Equations* **235** (2007), 609–622.

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

[30] J. Xu, J. You and Q. Qiu, Invariant tori of nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* **226** (1997), 375–386.

[31] J.-C. Yoccoz, Travaux de Herman sur les tores invariants, in: *Séminaire Bourbaki, Vol. 1991/92, Exposés 745–759*, Astérisque 206, Société Mathématique de France, Paris (1992), Exposé 754, 311–344.

[32] D. Zhang and J. Xu, Gevrey-smoothness of elliptic lower-dimensional invariant tori in Hamiltonian systems under Rüssmann's non-degeneracy condition, *J. Math. Anal. Appl.* **322** (2006), 293–312.

[33] D. Zhang and J. Xu, On elliptic lower dimensional tori for Gevrey-smooth Hamiltonian systems under Rüssmann's non-degeneracy condition, *Discrete Contin. Dynam. Systems Ser. A* **16** (2006), 635–655.

[34] D. Zhang and J. Xu, Invariant curves of analytic reversible mappings under Brjuno–Rüssmann's non-resonant condition, *J. Dynam. Differential Equations* **26** (2014), no. 4, 989–1005.

[35] D. Zhang and J. Xu, On invariant tori of vector field under weaker non-degeneracy condition, *NoDEA Nonlinear Differential Equations Appl.* **22** (2015), 1381–1394.