A Note on KAM Theorem for Symplectic Mappings

Zai-jiu Shang

Institute of Mathematics
Academia Sinica
Beijing 100080

China

Max-Planck-Institut
für Mathematik
Gottfried-Claren-Str. 26
53225 Bonn

Germany

MPI 96-116

A Note on KAM Theorem for Symplectic Mappings

Zai-jiu Shang

Institute of Mathematics, Academia Sinica, Beijing 100080, China

E-mail: zaijiu@math03.math.ac.cn

1. Introduction and main result

a. In this note we prove the existence of differentiable foliation structures of invariant tori for nearly integrable symplectic mappings in the sense of Whitney as J. Pöschel did in the case of Hamiltonian systems([10]). The formulation of the main theorem is almost a copy of that given in [10] and the proof also follows from essentially the same arguments only some different technical details should be pointed out particularly. The main aim of the note is to give two relevant estimates: one is for perturbations under which a majority of invariant tori of the integrable mapping persists and another is for invariant Liouville measure of the set filled by the invariant tori of the perturbed mappings in phase space. The estimates are both given explicitly in terms of three parameters: the first one is γ, which appears in the well-known diophantine condition satisfied by the frequencies, say $\omega = (\omega_1, \cdots, \omega_n)$, of invariant tori

$$\left|e^{i\langle k,\omega\rangle} - 1\right| \geq \frac{\gamma}{|k|^\tau}, \quad \text{for } 0 \neq k = (k_1, \cdots, k_n) \in \mathbf{Z}^n \tag{1.1}$$

with another fixed constant $\tau > 0$, where $\langle k, \omega\rangle = \sum_{j=1}^{n} k_j\omega^{(j)}$ and $|k| = \sum_{j=1}^{n} |k_j|$ for integers $k \in \mathbf{Z}^n$. The other two, say θ and Θ, are used to describe the nondegeneracy of the frequency map $\omega : I \to \Omega$, of the unperturbed integrable mapping, and its inverse, for example, of the form

$$\theta\,|p_1 - p_2| \leq |\omega(p_1) - \omega(p_2)| \leq \Theta\,|p_1 - p_2|\,. \tag{1.2}$$

Here I and Ω are the domains of action variables and the corresponding frequency values respectively. We always assume that they are open in $\mathbf{R}^n$. In fact, in our case we require $\omega : I \to \Omega$ to be analytic so ω is assumed to be defined and analytic in some complex extension, say $I + r$, with radius r of real domain $I \subset \mathbf{R}^n$ of action variables (cf. [10]) and (1.2) is assumed to be true for p_1, $p_2 \in I + r$ with $|p_1 - p_2| \leq r$. Note that this nondegeneracy condition implies that the frequency map ω is invertible in any ball with radius r and centered in I and so it is stronger than the standard Kolmogorov's nondegeneracy assumption but equivalent to the Pöschel's assumption ([10]).

It should be noted that there are already some optimal results on the estimates in terms of γ. For example, H. Rüssmann([11], $n = 2$) and N. V. Svanidze([14]) in mapping case with $\omega(p) = p$ (and so $\theta = \Theta = 1$) and J. Pöschel([10]) and A. I. Neishtadt([8]) in Hamiltonian system case proved that if a perturbation has a norm, say ε, in some relevant function space, proportional to γ^2 with coefficient, say Δ, not dependent of γ and small enough, then the perturbed mapping has invariant tori, deformed slightly from those of unperturbed one, whose complement in the phase space has invariant Liouville measure proportional to γ with coefficient, say C, not dependent of γ. Therefore, the invariant tori of the perturbed mapping form a set with large measure in the phase space if γ is sufficiently small (in fact, γ may be chosen reasonably to be proportional to $\sqrt{\varepsilon}$ from the fact that ε is admitted to be proportional to γ^2 (cf. [10], [8])). In this note we also obtain the same results.

We note that in most general cases, the bounds of both Δ and C in various versions of KAM theorem depends on the "nondegenerate " behavior, in some sense, of the frequency map $\omega : I \to \Omega$ (cf. [1,2,6,7] and [3,12]). Here we work simply under the "ordinary" nondegeneracy assumption (1.2) (this is the strongest nondegeneracy assumption) and investigate the dependence of the bounds of Δ and C, say Δ_0 and C_0 respectively, on the nondegeneracy (or twist) parameters θ and Θ. We obtain the dependence to be $\Delta_0 = \delta_0 \theta \Theta^{-2}$ and $C_0 = c_0(\theta\Theta^{-1})^{-n}$, where δ_0 and c_0 are constants not dependent of θ, Θ or γ. If ω only satisfies the Kolmogorov's nondegeneracy condition, say,

$$\theta|dp| \le |d\omega(p)| \le \Theta|dp|, \quad \text{for } p \in I + r \tag{1.2$'$}$$

with constants $0 < \theta \le \Theta$, then we may get the estimate of the form $\Delta_0 = \delta_0\theta^2\Theta^{-3}$ for perturbation part. We remark that this estimate is better than that given in [5], in the case of analytic perturbations, by directly estimating the convergence of the Lindstedt series for individual KAM tori, which is essentially of the form $\varepsilon = \delta_0\gamma^4\theta^2\Theta^{-4}$ – this estimate can not be applied to small twist problem.

We are interested in such estimates because they relate to some significant problems such as the small twist problem and therefore also relate to the stability of symplectic integrators in solving integrable and nearly integrable Hamiltonian systems. In the small twist problem, a small parameter, say s, enters into the frequency map and so the parameters θ, Θ and even γ are all s-dependent, say, they turn into $s\theta$, $s\Theta$ and $s\gamma$ respectively. For the case of $n = 1$, Moser obtained the existence of invariant curves ([7]). But for $n > 1$, I have not seen an available reference. Our estimates naturally lead to a proper answer to the problem.

b. We consider an exact symplectic mapping $S : (p, q) \to (\hat{p}, \hat{q})$ to be defined in phase space $I \times \mathbf{T}^n$ by its generating function $H : I \times \mathbf{T}^n \to \mathbf{R}$ through relation

$$\begin{cases} \hat{p} = p - \frac{\partial H}{\partial q}(\hat{p}, q) \\ \hat{q} = q + \frac{\partial H}{\partial p}(\hat{p}, q), \end{cases} \tag{1.3}$$

where I is an open and usually bounded set of $\mathbf{R}^n$ in which the action variables p vary and $\mathbf{T}^n$ is the standard n-torus which is identified as $\mathbf{R}^n/2\pi\mathbf{Z}^n$ and the angle variables q vary on. When $H(p, q) = H_0(p)$ does not depend on angle variables, then (1.3) represents an integrable mapping $S = S_0 : (p, q) \to (\hat{p}, \hat{q}) = (p, q + \omega(p))$, which is well defined on $I \times \mathbf{T}^n$, with frequency map

$$\omega(p) = \frac{\partial H_0}{\partial p}(p),\ p \in I. \tag{1.4}$$

Under S_0, the phase space $I \times \mathbf{T}^n$ is completely foliated into invariant n-tori $\{p\} \times \mathbf{T}^n$, $p \in I$, on each of which the iterations of S_0 are linear with frequencies $\omega = \omega(p)$. When a perturbation, say $h(p, q)$, is added to H_0, i.e., when $H(p, q) = H_0(p) + h(p, q)$, then (1.3) does not define an integrable mapping generally. However, KAM theorem shows that the perturbed mapping S, defined by the perturbed generating function H according to (1.3), still exibits in a large extent the integrable behavior even not in any open domain but at least in some Cantor set of the phase space if the frequency map ω is nondegenerate in some sense (see [1,2,6,7] for strong (ordinary) nondegeneracy and [3,12] for weak nondegeneracy) and if the perturbation h belongs to the class $C^a(I \times \mathbf{T}^n)$ for a suitably large number $a > 0$ with a sufficiently small associated norm (cf. [10]). We note that for h small in $C^2(I \times \mathbf{T}^n)$, (1.3) really defines an exact symplectic mapping on $I' \times \mathbf{T}^n$ with some open subset I', of I, whose boundary is a little away from the boundary of I with distance equal to the C^2-norm of h.

As a Banach space, the class $C^a(I \times \mathbf{T}^n)$, with the corresponding norm denoted as $\|\cdot\|_{a, I\times\mathbf{T}^n}$, is always understood according to [10]. In this note, we also get the anisotropic differentiability of the foliations of invariant tori. So we also need the class $C^{\nu_1,\nu_2}(I \times \mathbf{T}^n)$ of anisotropic differentiable functions, whose exact definition was also given in [10]. The norm of a function in $C^{\nu_1,\nu_2}(I \times \mathbf{T}^n)$ is denoted by $\|\cdot\|_{\nu_1,\nu_2;I\times\mathbf{T}^n}$. We also use another norm $\|\cdot\|_{\nu_1,\nu_2;I\times\mathbf{T}^n,\rho}$ for $\rho > 0$ defined by

$$\|u\|_{\nu_1,\nu_2;I\times\mathbf{T}^n,\rho} = \|u \circ \sigma_\rho\|_{\nu_1,\nu_2;\sigma_\rho^{-1}(I\times\mathbf{T}^n)} \tag{1.5}$$

for $u \in C^{\nu_1,\nu_2}(I \times \mathbf{T}^n)$, where σ_ρ denotes the partial stretching $(x, y) \to (\rho x, y)$ for $(x, y) \in I \times \mathbf{T}^n$. Note that the following relation between these two norms is valid for $0 < \rho \leq 1$:

$$\|u\|_{\nu_1,\nu_2;\rho} \leq \|u\|_{\nu_1,\nu_2} \leq \rho^{-\nu_1} \|u\|_{\nu_1,\nu_2;\rho}, \tag{1.6}$$

where we dropped the domains to simplify the notations.

Write $\Omega = \omega(I)$ and denote by Ω_γ the set of those frequencies, in Ω, which satisfy the diophantine condition (1.1) for the given γ and whose distance to the boundary of Ω is at least equal to 2γ. The difference $\Omega \setminus \bigcup_{\gamma>0} \Omega_\gamma$ is a zero set if $\tau > n + 1$ and so Ω_γ is large for small γ.

Now we formulate our main result as follows.

Theorem 1.1. *With given positive integer n and given $\tau > n + 1$, consider mapping S to be defined in phase space $I \times \mathbf{T}^n$ by generating function $H(p, q) = H_0(p) + h(p, q)$ in the form of (1.3), where $H_0(p)$ is analytic in $p \in I + r$ with $r > 0$ and $h(p, q)$ belongs to the class $C^{\alpha\lambda+\lambda+\tau}(I \times \mathbf{T}^n)$ with fixed $\lambda > \tau + 1$ and $\alpha > 1$ not in the discrete set*

$$\Lambda = \{i/\lambda + j : i, j \geq 0 \, integer\}.$$

Suppose that the frequency map $\omega : I \to \Omega$, defined by (1.4) from H_0, satisfies the nondegeneracy condition (1.2) for p_1, $p_2 \in I + r$ with $|p_1 - p_2| \leq r$ where the constants θ and Θ satisfy $0 < \theta \leq \Theta$. Then there exists a positive constant δ_0, depending only on n, τ, λ and α, such that for any $0 < \gamma \leq \min(1, \frac{1}{2}r\Theta)$, if

$$\|h\|_{\alpha\lambda+\lambda+\tau,I\times\mathbf{T}^n;\gamma\Theta^{-1}} \leq \delta_0 \gamma^2 \theta \Theta^{-2}, \tag{1.7}$$

then there exist a closed set $I_\gamma \subset I$, a surjective map $\omega_\gamma : I_\gamma \to \Omega_\gamma$ of $C^{\alpha+1}$ class and an injection $\Phi : I_\gamma \times \mathbf{T}^n \to \mathbf{R}^n \times \mathbf{T}^n$ of $C^{\alpha,\alpha\lambda}$ class, in the Whitney's sense, such that

(1) $S \circ \Phi = \Phi \circ R$. Moreover, this equation may be differentiated as often as Φ allows, where R is the integrable rotation on $I_\gamma \times \mathbf{T}^n$ with frequency map ω_γ.

(2) If Ω is a bounded open set of type D in the Arnold's sense [1], then we have measure estimate

$$m\mathcal{E}_\gamma \geq \left(1 - c_4\left(\theta\Theta^{-1}\right)^{-n}\gamma\right)m\mathcal{E}, \tag{1.9}$$

where m denotes the invariant Liouville measure of the phase space $\mathcal{E} = I \times \mathbf{T}^n$ and $\mathcal{E}_\gamma = \Phi(I_\gamma \times \mathbf{T}^n)$; c_4 is a positive constant depending on n, τ, a and the geometry of the domain Ω.

(3) If h is of $C^{\beta\lambda+\lambda+\tau}$ class with $\alpha \leq \beta$ not in Λ, then we have further that $\omega_\gamma \in C^{\beta+1}(I_\gamma)$ and $\Phi \in C^{\beta,\beta\lambda}(I_\gamma \times \mathbf{T}^n)$. Moreover,

$$\left\|\sigma_{\gamma\Theta^{-1}}^{-1} \circ (\Phi - I)\right\|_{\beta,\beta\lambda;\gamma\Theta^{-1}}, \gamma^{-1}\left\|\omega_\gamma - \omega\right\|_{\beta+1;\gamma\Theta^{-1}} \leq c_5\gamma^{-2}\Theta\left\|h\right\|_{\beta\lambda+\lambda+\tau;\gamma\Theta^{-1}} \tag{1.10}$$

with constant c_5 depending on n, τ, λ and β, here we have droped the domains in the notation of norms.

Theorem 1.2. *In Theorem 1.1, if the frequency map ω satisfies the nondegeneracy condition* (1.2)′, *then the required smallness condition for h is, instead of* (1.7),

$$\|h\|_{\alpha\lambda+\lambda+\tau,I\times\mathbf{T}^n;\gamma\Theta^{-1}} \leq \delta_0\gamma^2\theta^2\Theta^{-3} \tag{1.11}$$

with sufficiently small δ_0, depending only on n, τ, λ and α, under which the conclusions of Theorem 1.1 are also true with the same estimates.

Remark: We have following two further conclusions in the above theorems:

1. If $h \in C^\infty(I \times \mathbf{T}^n)$, then $\omega_\gamma \in C^\infty(I_\gamma)$ and $\Phi \in C^\infty(I_\gamma \times \mathbf{T}^n)$ with the estimates (1.10) for any $\beta \geq \alpha$.

2. If $h \in C^\omega(I \times \mathbf{T}^n)$, then we have further $\Phi \in C^{\infty,\omega}(I_\gamma, \mathbf{T}^n)$ under a further smallness condition for δ which also depends on the radius of analyticity of h with respect to angle variables.

2. Outline of the proof of Theorem 1.1

In this section, we outline the proof of Theorem 1.1. The detailed arguments will be given in the latter sections.

a. First we transform the mapping S by the partial coordinates stretching $\sigma_\rho : (x, y) \to (p, q) = (\rho x, y)$ of phase space $I \times \mathbf{T}^n$ and obtain a new mapping

$T = \sigma_\rho^{-1} \circ S \circ \sigma_\rho : (x, y) \to (\hat{x}, \hat{y})$ to be defined in the new phase space $I_\rho \times \mathbf{T}^n$ by

$$\begin{cases} \hat{x} = x - \frac{\partial F}{\partial y}(\hat{x}, y) \\ \hat{y} = y + \frac{\partial F}{\partial x}(\hat{x}, y), \end{cases} \tag{2.1}$$

where

$$F(x, y) = F_0(x) + f(x, y) \tag{2.2}$$

is well defined on $I_\rho \times \mathbf{T}^n$ with

$$F_0(x) = \rho^{-1} H_0(\rho x), \quad f(x, y) = \rho^{-1} h(\rho x, y) \tag{2.3}$$

and

$$I_\rho = \rho^{-1} I = \{x \in \mathbf{R}^n | \rho x \in I\}. \tag{2.4}$$

For the time being, ρ is considered as a free parameter. $F_0(x)$ is real analytic for $x \in I_\rho + r_\rho$ with $r_\rho = \rho^{-1} r$ and f belongs to the class $C^a(I_\rho \times \mathbf{T}^n)$ with $a = \alpha\lambda + \lambda + \tau$. So the new mapping T satisfies the same assumptions of Theorem 1.1 in which only I, r, H, H_0, h are replaced by I_ρ, r_ρ, F, F_0, f respectively. Correspondingly, the frequency map of integrable mapping associated to the generating function F_0 turns into

$$\tilde{\omega}(x) = \frac{\partial F_0}{\partial x}(x), \quad x \in I_\rho$$

and the nondegenerate condition for the mapping turns out to be

$$\rho\theta |x_1 - x_2| \le |\tilde{\omega}(x_1) - \tilde{\omega}(x_2)| \le \rho\Theta |x_1 - x_2|, \quad x_1, x_2 \in I_\rho + r_\rho, |x_1 - x_2| \le r_\rho. \tag{2.5}$$

In addition, from (2.3), we have

$$\|f\|_{a, I_\rho \times \mathbf{T}^n} = \rho^{-1} \|h\|_{a, I \times \mathbf{T}^n; \rho}.$$

From now on we fix $\rho = \gamma\Theta^{-1}$. Then the assumption $0 < \gamma \le \frac{1}{2} r\Theta$ in Theorem 1.1 implies that $0 < \rho \le \frac{1}{2} r$ and so $r_\rho \ge 2$. Let I_ρ^* be the set of points in I_ρ with distance at least one to its boundary and let

$$I_{\rho;\gamma} = \tilde{\omega}^{-1}(\Omega_\gamma) \cap I_\rho. \tag{2.6}$$

Then, from (2.5) and the definition of Ω_γ, we have

$$(I_{\rho;\gamma} + 1) \cap \mathbf{R}^n \subset I_\rho^* \subset (I_\rho^* + 1) \cap \mathbf{R}^n \subset I_\rho, \tag{2.7}$$

and

$$\gamma\mu\,|x_1 - x_2| \le |\tilde{\omega}(x_1) - \tilde{\omega}(x_2)| \le \gamma\,|x_1 - x_2|\,, \quad x_1,\, x_2 \in I_\rho + 2,\ |x_1 - x_2| \le 2 \tag{2.8}$$

with $\mu = \theta\Theta^{-1}$.

b. We approximate f by real analytic functions. Let

$$s_j = s_0 4^{-j}, \quad r_j = s_j^\lambda, \quad j = 0, 1, 2, \cdots \tag{2.9}$$

with fixed $\lambda > \tau + 1$ and $s_0 > 0$ to be determined later on and let

$$\mathcal{U}_j = I_\rho \times \mathbf{T}^n + (4s_j, 4s_j).$$

Then by Approximation Lemma of [10], p. 665, there exist real analytic functions f_j defined on $\mathcal{U}_0$ with $f_0 = 0$ such that, for $f \in C^b(I \times \mathbf{T}^n)$ with $b \ge a$, we have

$$\begin{aligned} &|f_j - f_{j-1}|_{\mathcal{U}_j} \le s_j^b c_b \,\|f\|_{b;I_\rho\times\mathbf{T}^n} \quad j = 1, 2, \cdots, \\ &\|f - f_j\|_{b',I_\rho^*\times\mathbf{T}^n} \to 0 \quad (j \to \infty) \quad \text{for } 0 < b' < b, \end{aligned} \tag{2.10}$$

where c_b is a positive constant only depending on b, n and s_0 but not depending on the domain I_ρ and hence not depending on the parameter ρ. Moreover, we may require f_j to be 2π-periodic in the last n variables. In (2.10), $|\cdot|_{\mathcal{U}_j}$ denotes the maximum norm of analytic functions on the complex domains $\mathcal{U}_j$. Almost all the notations in this note come from [10], which we will not specify particularly.

c. We give the KAM iteration process which is essentially the same as that given by Pöschel in Hamiltonian system case ([10]). Associating to each f_j, we will define a mapping $T_j : (x, y) \to (\hat{x}, \hat{y})$ by

$$\begin{cases} \hat{x} = x - \frac{\partial F_j}{\partial y}(\hat{x}, y) \\ \hat{y} = y + \frac{\partial F_j}{\partial x}(\hat{x}, y) \end{cases} \tag{2.11}$$

with $F_j(x, y) = F_0(x) + f_j(x, y)$, which is well-defined and real analytic on $\mathcal{U}_j$ if $4s_j \le r_\rho = \rho^{-1}r$ which is always true for $j = 0, 1, \cdots$ if we choose $0 < s_0 \le 4^{-1}$(noting that $0 < \gamma < \frac{1}{2}r\Theta$). We will show that each of T_j for $j \ge 0$ is really well-defined on a complex domain, of the phase space $I_\rho \times \mathbf{T}^n$, which is appropriate for the KAM iteration if h is bounded by (1.7) with δ_0 chosen to be small enough, independently of γ, θ and Θ. In fact, as j approaches infinity, T_j

will converge to the T in $C^{a'}$ norm with $0 < a' < a - 1$ on some sub-domain $I'_\beta \times \mathbf{T}^n$ of the phase space $I_\rho \times \mathbf{T}^n$ with I'_β being a sufficiently large open subset of I_ρ. The central problem is to find transformations Φ_j and integrable rotations R_j with frequency maps, say $\omega^{(j)}$, such that, as $j \to \infty$,

$$C_j = R_j^{-1} \circ \Phi_j^{-1} \circ T_j \circ \Phi_j \to \text{identity}, \; \Phi_j \to \tilde{\Phi}, \; R_j \to \tilde{R}, \; \omega^{(j)} \to \tilde{\omega}_\gamma, \qquad (2.12)$$

in $C^{\tilde{a}}$ norm, with a suitable positive number $\tilde{a}$, on $\tilde{I}_{\rho;\gamma} \times \mathbf{T}^n$ with some closed set $\tilde{I}_{\rho;\gamma} \subset I_\rho$ such that $\tilde{\omega}_\gamma(\tilde{I}_{\rho;\gamma}) = \Omega_\gamma$; here $\tilde{\Phi}$ and $\tilde{R}$ are well-defined on $\tilde{I}_{\rho;\gamma} \times \mathbf{T}^n$ and $\tilde{\omega}_\gamma$, as the frequency map of the rotation $\tilde{R}$, is well-defined on $\tilde{I}_{\rho;\gamma}$. So in the limit we have

$$T \circ \tilde{\Phi} = \tilde{\Phi} \circ \tilde{R} \quad \text{on } \tilde{I}_{\rho;\gamma} \times \mathbf{T}^n. \qquad (2.13)$$

Reversing the mapping T to S by the partial coordinates stretching σ_ρ and, at the same time, reversing $\tilde{\Phi}$ to Φ and $\tilde{R}$ to R with the frequency map $\omega_\gamma(p) = \tilde{\omega}_\gamma(\rho^{-1}p)$, then we have

$$S \circ \Phi = \Phi \circ R \quad \text{on } I_\gamma \times \mathbf{T}^n$$

with

$$I_\gamma = \rho \tilde{I}_{\rho;\gamma} = \{p \in \mathbf{R}^n | \, \rho^{-1}p \in \tilde{I}_{\rho;\gamma}\},$$

a closed subset of I and such that $\omega_\gamma(I_\gamma) = \Omega_\gamma$. This is just the conclusion (1) of Theorem 1.1.

The construction of Φ_j and R_j uses the KAM iteration which we describe as follows.

Assume

$$|f_j - f_{j-1}| \sim \varepsilon_j, \quad j = 1, 2, \cdots \qquad (2.14)$$

with a decreasing sequence of positive numbers $\{\varepsilon_j\}_1^\infty$ and suppose we have already found a transformation Φ_j and a rotation R_j with frequency map $\omega^{(j)}$ such that

$$C_j = R_j^{-1} \circ \Phi_j^{-1} \circ T_j \circ \Phi_j \qquad (2.15)$$

satisfies

$$|C_j - I| \sim \varepsilon_{j+1}. \qquad (2.16)$$

We then construct a transformation Ψ_j and a new rotation R_{j+1} with frequency map $\omega^{(j+1)}$ such that

$$\Phi_{j+1} = \Phi_j \circ \Psi_j \qquad (2.17)$$

and (2.16) is also true for the next index $j+1$ with C_{j+1} defined by (2.15) where j is replaced by $j+1$. As was remarked in [10], "for this procedure to be successful it is essential to have precise control over the various domains of definition".

We define the transformation $\Psi_j : (\xi, \eta) \to (x, y)$ implicitly with the help of a generating function ψ_j by

$$x = \xi + \frac{\partial \psi_j}{\partial y}(\xi, y), \quad y = \eta - \frac{\partial \psi_j}{\partial x}(\xi, y). \tag{2.18}$$

To define the function ψ_j, we consider the mapping

$$B_j = R_j^{-1} \circ \Phi_j^{-1} \circ T_{j+1} \circ \Phi_j \tag{2.19}$$

which is also near identity and assumed to be given implicitly from its generating function b_j by

$$\hat{x} = x - \frac{\partial b_j}{\partial y}(\hat{x}, y), \quad \hat{y} = y + \frac{\partial b_j}{\partial x}(\hat{x}, y). \tag{2.20}$$

The function ψ_j is then determined from b_j by the following holomorphic equation

$$\psi_j(x, y + \omega^{(j)}(x)) - \psi_j(x, y) + \widetilde{b}_j(x, y) = 0, \tag{2.21}$$

where $\widetilde{b}_j(x, y) = b_j(x, y) - [b_j](x)$ with $[b_j]$ being the mean value of b_j over $\mathbf{T}^n$. Define

$$\omega^{(j+1)}(x) = \omega^{(j)}(x) + \frac{\partial [b_j]}{\partial x}(x), \tag{2.22}$$

then $R_{j+1} : (x, y) \to (\hat{x}, \hat{y})$ is just given by

$$\hat{x} = x, \quad \hat{y} = y + \omega^{(j+1)}(x). \tag{2.23}$$

With just defined Ψ_j and R_{j+1}, we easily show that, formally,

$$\Psi_j^{-1} \circ R_j \circ B_j \circ \Psi_j = R_{j+1} \circ C_{j+1}.$$

Similar formal calculation to that in [10] shows that (2.16) is valid if we replace j by $j+1$.

We will not solve the equation (2.21) exactly indeed. Instead, we will solve an equation with a replacement of $\widetilde{b}_j$ by a finite truncation of its fourier expansion with respect to angle variables so that "only finitely many resonances remain, and we obtain a real analytic solution ψ_j defined on an open set"([10], pp. 677). The idea goes back to Arnol'd [1,2].

In the following sections, we will make the above arguments more precise by carefully controlling the domains of definition of functions and mappings and giving the relevant estimates. For the convenience of the later statements, we follow [10] and abbreviate the differentiability orders by

$$a = \alpha\lambda + \lambda + \tau \le b = \beta\lambda + \lambda + \tau, \tag{2.24}$$

and set, for $f \in C^b(I_\rho \times \mathbf{T}^n)$,

$$\|f\|_{b, I_\rho \times \mathbf{T}^n} = \gamma\mu\delta_b. \tag{2.25}$$

In addition, we denote by π_k the translation $(x, y) \to (x, y + 2\pi e_k)$ of $\mathbf{R}^n$ for $k = 1, \cdots, n$, where e_k is the vector of $\mathbf{R}^n$ with the k-th entry equal to one and others equal to zero. With this notation, a well-defined transformation on $I \times \mathbf{T}^n$ may be viewed as a transformation on $I \times \mathbf{R}^n$ which commutes with π_k, $k = 1, \cdots, n$.

3. Iterative lemma and proof of Theorem 1.1

Now we state a so-called iterative lemma which is a quantitative formulation of the KAM iteration process outlined in the above section.

Lemma 3.1. *Assume that F_0 is real analytic on $I_\rho + 2$ with the gradient map $\tilde{\omega} = \frac{\partial F_0}{\partial x}$ satisfying the condition (2.8) in which $\gamma > 0$ and $0 < \mu \le 1$ are given in advance. For fixed $\lambda > \tau + 1$ and $\alpha > 1$, let $f \in C^a(I_\rho \times \mathbf{T}^n)$ with $a = \alpha\lambda + \lambda + \tau$ and let f_j be real analytic approximants to the f as given in the preceeding section. Assume $\delta_a \le \delta$. If δ is small enough, then for each $j \ge 0$, there exist a closed set $I_{\rho;\gamma}^{(j)} \subset I_\rho$, a real analytic function $F_0^{(j)}$ on $V_j = I_{\rho;\gamma}^{(j)} \times \mathbf{T}^n + (r_j, s_j)$ which is independent of the angle variables, and a real analytic exact symplectic map $\Phi_j : V_j \to \mathcal{U}_j$ which commutes with π_k, $k = 1, \cdots, n$ such that, for $f \in C^b(I_\rho \times \mathbf{T}^n)$ with $b = \beta\lambda + \lambda + \tau \ge a$, the following hold.*

(i) For $j \ge 1$, $I_{\rho;\gamma}^{(j)} + r_j \subset I_{\rho;\gamma}^{(j-1)} + \frac{5}{4}r_j \subset I_{\rho;\gamma}^{(j-1)} + r_{j-1}$.

(ii) $\omega^{(j)} = \frac{\partial F_0^{(j)}}{\partial x}$ maps $I_{\rho;\gamma}^{(j)}$ onto Ω_γ and satisfies the following nondegeneracy condition: for x_1, $x_2 \in I_{\rho;\gamma}^{(j)} + r_j$ with $|x_1 - x_2| \le r_j$,

$$\left(1 - \sum_{k=1}^{j} s_k\right) \gamma\mu \,|x_1 - x_2| \le \left|\omega^{(j)}(x_1) - \omega^{(j)}(x_2)\right| \le \left(1 + \sum_{k=1}^{j} s_k\right) \gamma \,|x_1 - x_2| \,. \tag{3.1}$$

(iii) Φ_0 *is the identity and for* $j \geq 1$, Φ_j *is well-defined and real analytic on*

$$V_{j-1}^2 = I_{\rho;\gamma}^{(j-1)} \times \mathbf{T}^n + (2r_j, 2s_j)$$

and commutes with π_k, $k = 1, \cdots, n$ *and*

$$|\Phi_j - \Phi_{j-1}|_{V_{j-1}^2} \leq r_j^\beta \cdot c_1 \mu \delta_b. \tag{3.2}$$

(iv) $C_j = R_j^{-1} \circ \Phi_j^{-1} \circ T_j \circ \Phi_j$ *is well-defined and real analytic on* V_j *and commutes with* π_k, $k = 1, \cdots, n$ *and*

$$|C_j - I|_{V_j} \leq s_{j+1}^{b-1} \cdot c_2 \gamma \mu \delta_b, \quad \textit{with } c_2 \geq 16 c_b, \tag{3.3}$$

where R_j *is the rotation map* $(x, y) \to (x, y + \omega^{(j)}(x))$.

(v) $\omega^{(0)} = \tilde{\omega}$ *and for* $j \geq 1$,

$$\left|\omega^{(j)} - \omega^{(j-1)}\right|_{V_j} \leq r_j^{\beta+1} \cdot c_3 \gamma \mu \delta_b. \tag{3.4}$$

The proof of Lemma 3.1 is postponed to the next section. Now we prove Theorem 1.1, taking this lemma for granted. We let

$$\widetilde{I}_{\rho;\gamma} = \bigcap_{j=0}^{\infty} \overline{I_{\rho;\gamma}^{(j)} + r_j}.$$

By (i) of Lemma 3.1, $\widetilde{I}_{\rho;\gamma}$ is a nonempty closed subset of I_ρ. Moreover, $\widetilde{I}_{\rho;\gamma} \subset I_{\rho;\gamma}^{(j+1)} + r_{j+1}$ and hence $\widetilde{I}_{\rho;\gamma} + r_{j+1} \subset I_{\rho;\gamma}^{(j)} + r_j$ for each j. Therefore, $\omega^{(j)}$, Φ_j and C_j are well defined on $\widetilde{V}_{j+1} = \widetilde{I}_{\rho;\gamma} \times \mathbf{T}^n + (r_{j+1}, s_{j+1})$. From the estimates (3.2) – (3.4) and the Inverse Approximation Lemma of [10], p. 665, the sequences $(\omega^{(j)} - \tilde{\omega})$ and $(\Phi_j - I)$ have real limits

$$\tilde{\omega}_\gamma - \tilde{\omega} \in C^{\beta+1}(\widetilde{I}_{\rho;\gamma}), \quad \|\tilde{\omega}_\gamma - \tilde{\omega}\|_{\beta+1;\widetilde{I}_{\rho;\gamma}} \leq c_5 \gamma \mu \delta_b \tag{3.5}$$

and

$$\widetilde{\Phi} - I \in C^{\beta,\beta\lambda}(\widetilde{I}_{\rho;\gamma} \times \mathbf{T}^n), \quad \left\|\widetilde{\Phi} - I\right\|_{\beta,\beta\lambda;\widetilde{I}_{\rho;\gamma} \times \mathbf{T}^n} \leq c_5 \mu \delta_b, \tag{3.6}$$

respectively, such that

$$\left\|\omega^{(j)} - \tilde{\omega}_\gamma\right\|_{\beta+1-\kappa,\widetilde{I}_{\rho;\gamma}} \to 0, \quad \left\|\Phi_j - \widetilde{\Phi}\right\|_{\beta-\iota,(\beta-\iota)\lambda;\widetilde{I}_{\rho;\gamma} \times \mathbf{T}^n} \to 0 \tag{3.7}$$

for any $\kappa > 0$ and $\iota > 0$, here c_5 is a constant depending only on n, τ, λ, β. From (3.3), it follows that

$$|T_j \circ \Phi_j - \Phi_j \circ R_j|_{\widetilde{V}_{j+1}} \leq s_{j+1}^{b-1} \cdot c_6 \gamma \mu \delta_b \tag{3.8}$$

since R_j and Φ_j are uniformly bounded with their Jocobian if δ is small as requied — this fact follows from (2.8), (3.2) and (3.4) where we take $b = a$ and we will point it out again in next section. As argued in [10] and noting that $\|T_j - T\|_{b-1-\kappa; I_\rho^* \times \mathbf{T}^n} \to 0$ for any $\kappa > 0$ with δ small enough, we see that (3.8) implies the following equation

$$T \circ \widetilde{\Phi} = \widetilde{\Phi} \circ \widetilde{R} \tag{3.9}$$

on $\widetilde{I}_{\rho;\gamma} \times \mathbf{T}^n$, which may be differentiated as often as $\widetilde{\Phi}$ allows, where $\widetilde{R}$, with $\widetilde{\omega}_\gamma$ as its frequency map, is the limit of R_j. Transforming T back to S by the coordinates stretching σ_ρ with $\rho = \gamma\Theta^{-1}$, and accordingly, transforming $\widetilde{I}_{\rho;\gamma}$ to I_γ, $\widetilde{\Phi}$ to Φ, $\widetilde{R}$ to R, $\widetilde{\omega}_\gamma$ to ω_γ, we then get the conclusions (1) and (3) of Theorem 1.1 and the estimates (1.10). From the fact that $\omega^{(j)}(I_{\rho;\gamma}^{(j)}) = \Omega_\gamma$ for any j we easily show that ω_γ maps I_γ onto Ω_γ. The measure estimate (1.9) follows, by Arnold's argument (cf. [1]), from the uniform boundedness of the Jacobian determinant of ω_γ from above and below by $(2\Theta)^n$ and $\left(\frac{1}{2}\theta\right)^n$ respectively, which is easily observed from (3.1) and the fact that $d\omega_\gamma = \left(\gamma\Theta^{-1}\right)^{-1} d\widetilde{\omega}_\gamma$. The proof of Theorem 1.1 is completed.

4. Proof of the iterative lemma

We set $I_{\rho;\gamma}^{(0)} = \widetilde{\omega}^{-1}(\Omega_\gamma)$. Then Lemma 3.1 is valid for $j = 0$. Assume the lemma is proven for indices $0, \cdots, j$. Then $C_j = R_j^{-1} \circ \Phi_j^{-1} \circ T_j \circ \Phi_j$ is well-defined and real analytic on V_j and commutes with π_k, $k = 1, \cdots, n$ and

$$|C_j - I|_{V_j} \leq \gamma \varepsilon_b^j = s_{j+1}^{b-1} \cdot c_2 \gamma \mu \delta_b \tag{4.1}$$

by (iv). To prove Lemma 3.1 for $j+1$, we consider the map $B_j = R_j^{-1} \circ \Phi_j^{-1} \circ T_{j+1} \circ \Phi_j$, as suggested in Section 2. First, we have

Lemma 4.1. *If δ is small enough, then B_j is well-defined and real analytic on*

$$V_j^3 = I_{\rho;\gamma}^{(j)} \times \mathbf{T}^n + (3r_{j+1}, 3s_{j+1})$$

and commutes with π_k, $k = 1, \cdots, n$. Moreover, we have

$$|B_j - I|_{V_j^3} \leq 2\gamma\varepsilon_b^j. \tag{4.2}$$

To prove this lemma, we may write, formally, $B_j = C_j \circ D_j$, where

$$D_j = \Phi_j^{-1} \circ T_j^{-1} \circ T_{j+1} \circ \Phi_j.$$

So far the mappings T_{j+1} and T_j^{-1} are defined only implicitly. First we need to determine the domains of definition of them.

Lemma 4.2. *If δ is small, then T_{j+1} is well defined on*

$$U_{j+1}^* = I_\rho^* \times \mathbf{T}^n + \left(\frac{13}{4}s_{j+1}, \frac{13}{4}s_{j+1}\right)$$

and maps this domain into

$$U_j^{**} = I_\rho^{**} \times \mathbf{T}^n + \left(\frac{14}{4}s_{j+1}, 7s_{j+1}\right)$$

on which T_j^{-1} is well-defined, where

$$I_\rho^{**} = \left(I_\rho^* + \frac{1}{4}\right) \cap \mathbf{R}^n.$$

Moreover, T_{j+1} and T_j^{-1} commute with π_k, $k = 1, \cdots, n$ and,

$$\left|T_j^{-1} \circ T_{j+1} - I\right|_{U_{j+1}^*} \leq s_{j+1}^{b-1} \cdot 8c_b\gamma\mu\delta_b. \tag{4.3}$$

Proof of Lemma 4.2. It is observed that Lemma 4.2 of [10] does not apply to the proof of this lemma. But the contraction argument in an appropriate Banach space still works in our case. To express T_{j+1} in explicit form, let us first solve $\hat{x}$ in terms of x and y from

$$\hat{x} = x - \frac{\partial f_{j+1}}{\partial y}(\hat{x}, y), \tag{4.4}$$

the first equation of (2.11) where j is replaced by $j+1$. Let $\mathcal{M}_{j+1}$ be the set of all real analytic n-valued functions $\eta(x, y)$, on U_{j+1}^*, which are 2π-periodic in the last n variables and satisfy

$$|\eta|_{U_{j+1}^*} \leq \frac{1}{4}, \quad |\Im\eta|_{U_{j+1}^*} \leq \frac{1}{4}s_{j+1}. \tag{4.5}$$

Then $\mathcal{M}_{j+1}$ is a Banach space. Now consider the map

$$\mathcal{F}(\eta)(x, y) = -\frac{\partial f_{j+1}}{\partial y}(x + \eta(x, y), y) \tag{4.6}$$

which is well-defined for $\eta \in \mathcal{M}_{j+1}$ and maps the space into itself because f_{j+1} is real analytic on $\mathcal{U}_{j+1}$ and 2π-periodic in the last n variables and, from (2.10) and Cauchy's estimates, for $\eta \in \mathcal{M}_{j+1}$ and $(x, y) \in U^*_{j+1}$,

$$
\begin{aligned}
|\mathcal{F}(\eta)(x,y)| &= \left|\frac{\partial f_{j+1}}{\partial y}(x+\eta(x,y),y)\right| \leq \sum_{k=0}^{j}\left|\frac{\partial}{\partial y}(f_{k+1}-f_k)\right|_{I^{**}_\rho \times \mathbf{T}^n + \left(\frac{14}{4}s_{k+1}, \frac{14}{4}s_{k+1}\right)} \\
&\leq \sum_{k=0}^{j}\left(\frac{1}{2}s_{k+1}\right)^{-1}|f_{k+1}-f_k|_{\mathcal{U}_{k+1}} \leq \Big(\sum_{k=0}^{j} s_{k+1}^{a-1}\Big)\cdot 2c_a\gamma\mu\delta_a \\
&\leq 2^{-(a-2)}s_0^{a-1}c_a\delta \leq \frac{1}{4}, \quad \text{if } \delta \leq \delta_1 = 4^{a-3}c_a^{-1}
\end{aligned} \tag{4.7}
$$

and

$$
\begin{aligned}
|\Im\mathcal{F}(\eta)(x,y)| &\leq \left|\frac{\partial f_{j+1}}{\partial y}(x+\eta(x,y),y) - \frac{\partial f_{j+1}}{\partial y}(\Re x + \Re\eta(x,y), \Re y)\right| \\
&\leq \left|\frac{\partial^2 f_{j+1}}{\partial y \partial x}(\bar{x},\bar{y})(\Im x + \Im\eta(x,y)) + \frac{\partial^2 f_{j+1}}{\partial y\partial y}(\bar{x},\bar{y})(\Im y)\right| \\
&\leq 4^{-(a-2)}nc_a\delta\cdot 8s_{j+1} \leq \frac{1}{4}s_{j+1}, \quad \text{if } \delta \leq \delta_2 = 4^{a-5}(nc_a)^{-1}.
\end{aligned} \tag{4.8}
$$

In the above, $(\bar{x},\bar{y}) \in I^{**}_\rho \times \mathbf{T}^n + \left(\frac{14}{4}s_{j+1}, \frac{14}{4}s_{j+1}\right)$ and, we have used the estimates, for example,

$$
\begin{aligned}
\left|\frac{\partial^2 f_{j+1}}{\partial y\partial x}(\bar{x},\bar{y})\right| &\leq \sum_{k=0}^{j}\left|\frac{\partial^2}{\partial y \partial x}(f_{k+1}-f_k)\right|_{I^{**}_\rho \times \mathbf{T}^n + \left(\frac{14}{4}s_{k+1}, \frac{14}{4}s_{k+1}\right)} \\
&\leq \sum_{k=0}^{j}\frac{2!}{\left(\frac{1}{2}s_{k+1}\right)^2}|f_{k+1}-f_k|_{\mathcal{U}_{k+1}} \leq \sum_{k=0}^{j} s_{k+1}^{a-2}\cdot 8c_a\delta_a \\
&\leq 2^{-(a-5)}s_0^{a-2}c_a\delta \leq 4^{-(a-2)}c_a\delta, \quad \text{if } 0 < s_0 \leq \tfrac{1}{4},
\end{aligned} \tag{4.9}
$$

with the notice that $a = \alpha\lambda + \lambda + \tau > 3\tau + 2 > 3n + 2 \geq 5$. By (4.9), we see that, for η_1, $\eta_2 \in \mathcal{M}_{j+1}$, $(x, y) \in U^*_{j+1}$,

$$
\begin{aligned}
|\mathcal{F}(\eta_1)(x,y) - \mathcal{F}(\eta_2)(x,y)| &= \left|\frac{\partial f_{j+1}}{\partial y}(x+\eta_1(x,y),y) - \frac{\partial f_{j+1}}{\partial y}(x+\eta_2(x,y),y)\right| \\
&\leq \left|\sum_{k=0}^{j}\frac{\partial^2 f_{k+1}}{\partial y\partial x}(x^*_k,y)(\eta_1(x,y)-\eta_2(x,y))\right| \\
&\leq \frac{1}{2}|\eta_1-\eta_2|_{U^*_{j+1}}, \quad \text{if } \delta \leq \delta_2.
\end{aligned}
$$

In the above equation, for each k, x_k^* is a point in $I_\rho^{**} + \frac{14}{4}s_{j+1}$. This shows that the map $\mathcal{F} : \mathcal{M}_{j+1} \to \mathcal{M}_{j+1}$ is contractive. Therefore, there exists a unique $y^* \in \mathcal{M}_{j+1}$ such that

$$\mathcal{F}(y^*) = y^*, \tag{4.10}$$

and $T_{j+1} : (x, y) \to (\hat{x}, \hat{y})$ is expressed explicitly in the form

$$\begin{cases} \hat{x} = x + y^*(x, y) \\ \hat{y} = y + \tilde{\omega}(x + y^*(x, y)) + \frac{\partial f_{j+1}}{\partial x}(x + y^*(x, y), y), \end{cases} \tag{4.11}$$

which is clearly well-defined and real analytic for $(x, y) \in U_{j+1}^*$ and commutes with π_k, $k = 1, \cdots, n$. It is easy to check that T_{j+1} maps U_{j+1}^* into U_j^{**} by using the condition (2.8) for $\tilde{\omega}$ and the similar arguments to the above. Also, the contraction arguments may be applied to show that T_j^{-1} is well-defined and real analytic on U_j^{**} and commutes with π_k, $k = 1, \cdots, n$ only but, instead of (4.6), here we need to consider the following map

$$\mathcal{G}(\eta)(x, y) = -\frac{\partial f_j}{\partial x}(x, y - \tilde{\omega}(x) + \eta(x, y)) \tag{4.12}$$

which maps the Banach space, say $\mathcal{M}_j^*$, of real analytic n-valued functions $\eta(x, y)$ with 2π-periodic in the last n variables y for (x, y) in U_j^{**} and with

$$|\eta|_{U_j^{**}} \le \frac{1}{4}, \quad |\Im\eta|_{U_j^{**}} \le \frac{1}{4}s_j, \tag{4.13}$$

into itself contractively if δ is small enough (say, $\delta \le \delta_2$). Therefore we have a unique fixed point $y^{**} \in \mathcal{M}_j^*$ of the map $\mathcal{G}$ and T_j^{-1} is explicitly given by

$$\begin{cases} \hat{x} = x + \frac{\partial f_j}{\partial y}(x, y - \tilde{\omega}(x) + y^{**}(x, y)) \\ \hat{y} = y - \tilde{\omega}(x) + y^{**}(x, y). \end{cases} \tag{4.14}$$

Simple calculations will show that T_j^{-1} maps U_j^{**} into

$$U_j^{***} = I_\rho^{***} \times \mathbf{T}^n + (2s_j, 3s_j)$$

with $I_\rho^{***} = (I_\rho^* + \frac{1}{2}) \cap \mathbf{R}^n$. It remains to verify (4.3). Note that $T_j^{-1} \circ T_{j+1} : (x, y) \to (\hat{x}, \hat{y})$ can be written in the form

$$\begin{cases} \hat{x} = x - \left(\frac{\partial f_{j+1}}{\partial y}(\tilde{x}, y) - \frac{\partial f_j}{\partial y}(\tilde{x}, \hat{y})\right) \\ \hat{y} = y + \left(\frac{\partial f_{j+1}}{\partial x}(\tilde{x}, y) - \frac{\partial f_j}{\partial x}(\tilde{x}, \hat{y})\right), \end{cases} \tag{4.15}$$

where $(\tilde{x}, \tilde{y}) = T_{j+1}(x, y)$. Solve $\hat{y}$ in terms of x and y from the second equation of (4.15), we get

$$\hat{y} = y + y^{***}(\tilde{x}, y), \tag{4.16}$$

where $y^{***}(\tilde{x}, y)$ is the fixed point of the contractive map

$$\mathcal{H}(\eta)(\tilde{x}, y) = \frac{\partial f_{j+1}}{\partial x}(\tilde{x}, y) - \frac{\partial f_j}{\partial x}(\tilde{x}, y + \eta(\tilde{x}, y)), \tag{4.17}$$

on the Banach space, say $\mathcal{M}_j^{**}$, of real analytic n-valued functions $\eta(\tilde{x}, y)$ with 2π-periodic in variables y for $(\tilde{x}, y) \in \widetilde{U}_j^{**} = I_\rho^{**} \times \mathbf{T}^n + \left(\frac{14}{4}s_{j+1}, \frac{13}{4}s_{j+1}\right)$ and with

$$|\eta|_{\widetilde{U}_j^{**}} \leq \frac{1}{4}s_{j+1}. \tag{4.18}$$

The well-definedness and the contractivity of the map $\mathcal{H}$ on $\mathcal{M}_j^{**}$ are easily proved by the previous arguments with the notice of the inequality (2.10) for $f_{j+1} - f_j$. With the well-defined $y^{***}(\tilde{x}, y)$ and $(\tilde{x}, \tilde{y}) = T_{j+1}(x, y)$, the map $T_j^{-1} \circ T_{j+1}$: $(x, y) \to (\hat{x}, \hat{y})$ is given explicitly by (4.15) the second equation of which is in fact equivalent to (4.16). Direct verification shows that, if $\delta \leq \delta_2$,

$$\begin{aligned}
|y^{***}|_{\widetilde{U}_j^{**}} &= \sup_{(\tilde{x},y)\in\widetilde{U}_j^{**}} \left| \frac{\partial f_{j+1}}{\partial x}(\tilde{x}, y) - \frac{\partial f_j}{\partial x}(\tilde{x}, y + y^{***}(\tilde{x}, y)) \right| \\
&\leq \sup_{(\tilde{x},y)\in\widetilde{U}_j^{**}} \left| \frac{\partial f_{j+1}}{\partial x}(\tilde{x}, y) - \frac{\partial f_j}{\partial x}(\tilde{x}, y) \right| \\
&+ \sup_{(\tilde{x},y)\in\widetilde{U}_j^{**}} \left| \frac{\partial f_j}{\partial x}(\tilde{x}, y) - \frac{\partial f_j}{\partial x}(\tilde{x}, y + y^{***}(\tilde{x}, y)) \right| \\
&\leq s_{j+1}^{b-1} \cdot 4c_b\gamma\mu\delta_b + \frac{1}{2}\, |y^{***}|_{\widetilde{U}_j^{**}},
\end{aligned}$$

and therefore

$$|y^{***}|_{\widetilde{U}_j^{**}} \leq s_{j+1}^{b-1} \cdot 8c_b\gamma\mu\delta_b. \tag{4.19}$$

In a similar way, we get

$$\sup_{(\tilde{x},y)\in\widetilde{U}_j^{**}} \left| \frac{\partial f_{j+1}}{\partial y}(\tilde{x}, y) - \frac{\partial f_j}{\partial y}(\tilde{x}, y + y^{***}(\tilde{x}, y)) \right| \leq s_{j+1}^{b-1} \cdot 8c_b\gamma\mu\delta_b. \tag{4.20}$$

This verifies (4.3) and completes the proof of Lemma 4.2.

Now we return to the proof of Lemma 4.1. By inequality (3.2) for indices $0, 1, \cdots, j$ and Cauchy's estimates, we easily prove that, if $\delta \leq \delta_3$ with

$$\delta_3 = \left(8nc_{1,a} \sum_{k=1}^{\infty} r_k^{\alpha-1}\right)^{-1} = \frac{4^{(\alpha-1)\lambda} - 1}{8nc_{1,a} s_0^{(\alpha-1)\lambda}},$$

then

$$\frac{1}{2} |z_1 - z_2| \leq |\Phi_j(z_1) - \Phi_j(z_2)| \leq 2 |z_1 - z_2| \tag{4.21}$$

for $z_1 \in V_j$ and $|z_1 - z_2| \leq \frac{1}{4} r_j$, which implies that $|D\Phi_j|_{V_j}, |D\Phi_j^{-1}|_{V_j} \leq 2$. This, together with (3.2), implies that $\Phi_j(V_j) \subset U_{j+1}^*$. From (4.21), we also have

$$\Phi_j(V_j^3) + \frac{1}{8} r_j \subset \Phi_j(V_j). \tag{4.22}$$

By (4.3) and the fact that $s_{j+1}^{a-1} \cdot 8c_a \gamma \mu \delta \leq \frac{1}{8} r_j$ if $\delta \leq \delta_4 = 4^{\lambda-3} c_a^{-1}$, we see that D_j is well-defined on V_j^3 and maps this domain into V_j. Therefore, B_j is well-defined on V_j^3. On the other hand, from (4.3) and the fact that $|D\Phi_j^{-1}|_{V_j} \leq 2$ and $c_2 \geq 16c_b$, we get

$$|D_j - 1|_{V_j^3} \leq 2 \left|T_j^{-1} \circ T_{j+1} \circ \Phi_j - \Phi_j\right|_{V_j^3} \leq 2 \left|T_j^{-1} \circ T_{j+1} - I\right|_{U_{j+1}^*} \leq \gamma \varepsilon_b^j.$$

Therefore,

$$|B_j - I|_{V_j^3} \leq |C_j \circ D_j - D_j|_{V_j^3} + |D_j - I|_{V_j^3} \leq |C_j - I|_{V_j} + |D_j - I|_{V_j^3} \leq 2\gamma\varepsilon_b^j.$$

It is clear that B_j commutes with π_k, $k = 1, \cdots, n$. Lemma 4.1 is then proved.

B_j is an exact symplectic map. So there exists a generating function, say b_j, such that $B_j : (x, y) \to (\hat{x}, \hat{y})$ is generated from b_j by (2.20). By the exact symplecticity of B_j and the previous contraction arguments, we have

Lemma 4.3. *If* $\delta \leq \delta_5 = (8nc_{2,a})^{-1}$, *then* b_j *is well-defined and real analytic on*

$$\tilde{V}_j^* = I_{\rho;\gamma}^{(j)} \times \mathbf{T}^n + \left(\frac{5}{2} r_{j+1}, \frac{5}{2} s_{j+1}\right)$$

and is 2π*-periodic in the last* n *variables. Moreover, we have*

$$\left|\frac{\partial b_j}{\partial x}\right|_{\tilde{V}_j^*}, \left|\frac{\partial b_j}{\partial y}\right|_{\tilde{V}_j^*} \leq 2\gamma\varepsilon_b^j. \tag{4.23}$$

Now we may define

$$F_0^{(j+1)}(x) = F_0^{(j)}(x) + [b_j](x), \quad \omega^{(j+1)}(x) = \frac{\partial F_0^{(j+1)}}{\partial x}(x), \tag{4.24}$$

where $[b_j](x) = \frac{1}{(2\pi)^n} \int\limits_{\mathbf{T}^n} b_j(x,y) d^n y$. Then $F_0^{(j+1)}$ and $\omega^{(j+1)}$ are well-defined and real analytic on $\widetilde{V}_j^*$ and independent of the angle variables. Below we first show that there exists a closed subset $I_{\rho;\gamma}^{(j+1)}$ of I_ρ such that (i), (ii) anf (v) of Lemma 3.1 are true for index $j+1$. For this, we state the following lemma cited from [9] in which the proof can be found.

Lemma 4.4. *Assume that $F : D \to \mathbf{R}^n, G : D \to \mathbf{R}^n$ are two continuously differentiable mappings where D is an open subset of $\mathbf{R}^n$. Let C be a bounded open set such that $\bar{C} \subset D$ where $\bar{C}$ denotes the closure of C. Let $y \in \mathbf{R}^n$ satisfy*

$$L = \min_{x \in \partial C} \{ \|F(x) - y\|_2 \} > 0.$$

If

$$\sup_{x \in \bar{C}} \{ \|F(x) - G(x)\|_2 \} \le \frac{L}{8},$$

then $deg(F, C, y) = deg(G, C, y)$. Where $deg(F, C, y)$ denotes the degree of the mapping F associated to the point y and the region C, and $\|\cdot\|_2$ denotes the usual Euclidean metric.

We continue our proof. To apply Lemma 4.4, we fix a point $x^* \in I_{\rho;\gamma}^{(j)}$ and let $F = \omega^{(j)}$, $G = \omega^{(j+1)}$, $D = \{x^*\} + \frac{1}{2} r_j$, $C = \{x^*\} + \frac{1}{4} r_j$, $y = \omega^* = \omega^{(j)}(x^*)$. Since $0 < s_0 \le 1$, we have from (3.1)

$$\frac{2}{3} \gamma \mu \, |x_1 - x_2| \le |F(x_1) - F(x_2)| \le \frac{4}{3} \gamma \, |x_1 - x_2| \,, \quad x_1, \, x_2 \in D. \tag{4.25}$$

Therefore,

$$L = \min_{x \in \partial C} \{ \|F(x) - y\|_2 \} \ge \min_{x \in \partial C} \{ \left| \omega^{(j)}(x) - \omega^{(j)}(x^*) \right| \} \ge \frac{1}{6} \gamma \mu r_j$$

and

$$\sup_{x \in \bar{C}} \{ \|F(p) - G(p)\|_2 \} \le \sqrt{n} \left| \frac{\partial b_j}{\partial x} \right|_{\widetilde{V}_j^*} \le \frac{L}{8}, \quad \text{if } \delta \le \delta_5.$$

By Lemma 4.4, we have $\deg(\omega^{(j+1)}, C, \omega^*) = \deg(\omega^{(j)}, C, \omega^*) = 1$. From the theorem of topology degree of a mapping, there exists a point $x_1^* \in C$ such that $\omega^{(j+1)}(x_1^*) = \omega^*$. Moreover, we have

$$\begin{aligned} |x_1^* - x^*| &\le \left(\frac{2}{3} \gamma \mu \right)^{-1} \left| \omega^{(j)}(x_1^*) - \omega^{(j)}(x^*) \right| \\ &\le \left(\frac{2}{3} \gamma \mu \right)^{-1} \left| \frac{\partial [b_j]}{\partial x}(x_1^*) \right| \le \frac{1}{4} r_{j+1}. \end{aligned}$$

Let $I_{\rho;\gamma}^{(j+1)}$ be the set of points $x_1^* \in I_\rho$ such that $\omega^{(j+1)}(x_1^*) = \omega^{(j)}(x^*)$ for $x^* \in I_{\rho;\gamma}^{(j)}$. Then the conclusion *(i)* of Lemma 3.1 is obviously true for $j+1$ and $\omega^{(j+1)}$ maps $I_{\rho;\gamma}^{(j+1)}$ onto Ω_γ. So $\omega^{(j+1)}$ is well-defined on V_{j+1} and (3.4) is also valid for $j+1$ with $c_3 = 2c_2$. Seeing that, for $x_1,\ x_2 \in I_{\rho;\gamma}^{(j+1)} + r_{j+1}$ with $|x_1 - x_2| \le r_{j+1}$,

$$\begin{aligned}\left|\frac{\partial[b_j]}{\partial x}(x_1) - \frac{\partial[b_j]}{\partial x}(x_2)\right| &\le \frac{1}{(2\pi)^n}\int_{\mathbf{T}^n}\left|\frac{\partial^2 b_j}{\partial x \partial x}(\tilde{x}, y)(x_1 - x_2)\right| d^n y \\ &\le n \cdot r_{j+1}^{-1} \cdot \left|\frac{\partial b_j}{\partial x}\right|_{\tilde{V}_j^*} \cdot |x_1 - x_2| \le s_{j+1} \cdot \gamma\mu\, |x_1 - x_2| \quad (4.26)\end{aligned}$$

if $\delta \le \delta_5$ with $\tilde{x} \in \overline{x_1 x_2}$, the straight line connecting x_1 and x_2 and completely contained in $I_{\rho;\gamma}^{(j)} + \frac{5}{2} r_{j+1}$, the equation (3.1) is then verified for $j+1$. In (4.26) we have applied the Cauchy's estimate to analytic function $\dfrac{\partial b_j}{\partial x}$. It remains to verify *(iii)* and *(iv)* for $j+1$. For this, we define Ψ_j.

Note that the Fourier series expansion

$$b_j(x, y) = [b_j](x) + \sum_{0 \ne k \in \mathbf{Z}^n} b_{j;k}(x) e^{i\langle k,, y\rangle}$$

is defined on $\tilde{V}_j^*$. Let $\tilde{b}_j(x, y) = \sum_{0 \ne k \in \mathbf{Z}^n} b_{j;k}(x) e^{i\langle k,, y\rangle}$. We take generating function $\psi_j(x, y)$, of the symplectic transformation Ψ_j, as the solution of the equation

$$\psi_j(x, y + \omega^{(j)}(x)) - \psi_j(x, y) + T_{m_j}\tilde{b}_j(x, y) = 0, \tag{4.27}$$

with $[\psi_j](x) = 0$, where

$$T_{m_j}\tilde{b}_j(x, y) = \sum_{0 < |k| \le m_j} b_{j;k}(x) e^{i\langle k,, y\rangle}$$

with m_j given by

$$m_j^{\tau+1} = \frac{1}{r_j}. \tag{4.28}$$

We find

$$\psi_j(x, y) = -\sum_{0 < |k| \le m_j} \frac{b_{j;k}(x)}{e^{i\langle k, \omega^{(j)}(x)\rangle} - 1} e^{i\langle k,, y\rangle}. \tag{4.29}$$

Lemma 4.5. *$\psi_j(x, y)$ is well-defined and real analytic on $\tilde{V}_j^*$ and is 2π-periodic in each variables of y. Moreover, if $f \in C^b(I_\rho \times \mathbf{T}^n)$, then we have the following estimates for its gradient:*

$$\left|\frac{\partial \psi_j}{\partial x}(x, y)\right| \le \sigma_1 s_{j+1}^{-2\tau}\varepsilon_b^j, \quad \left|\frac{\partial \psi_j}{\partial y}(x, y)\right| \le \sigma_1 s_{j+1}^{-\tau}\varepsilon_b^j \tag{4.30}$$

for (x, y) *in*

$$V_j^* = I_{\rho;\gamma}^{(j)} + \left(\frac{9}{4}r_{j+1}, \frac{9}{4}s_{j+1}\right),$$

where σ_1 *is a positive constant depending only on* n *and* τ.

Proof: The well-definedness of ψ_j is proved by simply verifying the nonvanishing of $\left(e^{i\langle k,\omega^{(j)}(x)\rangle} - 1\right)$ for $x \in I_{\rho;\gamma}^{(j)} + \frac{5}{2}r_{j+1}$ and $0 < |k| \le m_j$. By the definitions of $I_{\rho;\gamma}^{(j)}$ and m_j with the notice that $r_{j+1} = 4^{-\lambda}r_j$ and that $\lambda \ge 2$, we get for $x^* \in I_{\rho;\gamma}^{(j)}$, $|x - x^*| \le \frac{5}{2}r_{j+1}$ and $0 < |k| \le m_j$,

$$\begin{aligned}\left|e^{i\langle k,\omega^{(j)}(x)\rangle} - 1\right| &\ge \left|e^{i\langle k,\omega^{(j)}(x^*)\rangle} - 1\right| - \left|e^{i\langle k,\omega^{(j)}(x)\rangle} - e^{i\langle k,\omega^{(j)}(x^*)\rangle}\right| \\ &\ge \frac{\gamma}{|k|^\tau} - 2\left|\langle k, \omega^{(j)}(x) - \omega^{(j)}(x^*)\rangle\right| \\ &\ge \frac{\gamma}{|k|^\tau} - 2m_j \cdot 2\gamma \cdot \frac{5}{2}r_{j+1} \ge \frac{1}{4} \cdot \frac{\gamma}{|k|^\tau},\end{aligned}$$

which does not vanish. The analyticity and the periodicity in the last n variables of this function are clear. The remainder of this lemma can be proved by differentiating the equation (4.27) with respect to x and y, and then estimating $\left|\frac{\partial\psi_j}{\partial x}\right|$ and $\left|\frac{\partial\psi_j}{\partial y}\right|$ over V_j^* in terms of the norms $\left|\frac{\partial b_j}{\partial x}\right|_{\tilde{V}_j^*}$ and $\left|\frac{\partial b_j}{\partial y}\right|_{\tilde{V}_j^*}$ respectively, by using the standard arguments. The details are referred to [10] and [13] for a similar problem.

Therefore, $\Psi_j : (\xi, \eta) \to (x, y)$ is well-defined from ψ_j by (2.18). More precisely, we have the following result which is easily proved by the standard contraction argument.

Lemma 4.6. *If* $\delta \le \delta_6 = (8n\sigma_1 c_{2,a})^{-1}$, *then* Ψ_j *and its inverse* Ψ_j^{-1} *are well-defined on* V_j^2 *and map* V_j^2 *into* V_j^* *with*

$$|\Psi_j - I|_{V_j^2} \le s_{j+1}^{-2\tau} \cdot \sigma_1 \varepsilon_b^j; \tag{4.31}$$

and $\Psi_j^{-1} \circ R_j \circ B_j \circ \Psi_j$ *is well defined and real analytic on* V_{j+1} *and maps* V_{j+1} *into* $\tilde{V}_j^*$.

Let $\Phi_{j+1} = \Phi_j \circ \Psi_j$. Then, Φ_{j+1} is well-defined and real analytic on V_j^2 and commutes with π_k, $k = 1, \cdots, n$ and, from (4.21),

$$|\Phi_{j+1} - \Phi_j|_{V_j^2} = |\Phi_j \circ \Psi_j - \Phi_j|_{V_j^2} \le 2\,|\Psi_j - I|_{V_j^2} \le r_{j+1}^\beta \cdot c_1\mu\delta_b$$

with $c_1 = 2\sigma_1 c_2$. (3.2) is then verified. Next we verify (3.3) for $j+1$. Note that $C_{j+1} = R_{j+1}^{-1} \circ \Psi_j^{-1} \circ R_j \circ B_j \circ \Psi_j$, which is analytically defined on V_{j+1} and commutes with π_k, $k = 1, \cdots, n$ and, moreover, it may be written as $(\xi, \eta) \to (\hat{\xi}, \hat{\eta})$ with

$$\hat{\xi} = \xi - \frac{\partial \psi_j}{\partial y}(\hat{\xi}, \hat{y}) + \frac{\partial \psi_j}{\partial y}(\xi, y) - \frac{\partial b_j}{\partial y}(\hat{x}, y)$$

$$\hat{\eta} = \eta + \frac{\partial \psi_j}{\partial \xi}(\hat{\xi}, \hat{y}) - \frac{\partial \psi_j}{\partial \xi}(\xi, y) + \frac{\partial b_j}{\partial x}(\hat{x}, y) - \omega^{(j+1)}(\hat{\xi}) + \omega^{(j)}(\hat{x}),$$

where $(x, y) = \Psi_j(\xi, \eta) \in T_{\beta;\gamma}^{(j)} \times \mathbf{T}^n + \left(\frac{6}{4} r_{j+1}, \frac{6}{4} s_{j+1}\right)$, $(\hat{x}, \hat{y}) = R_j \circ B_j(x, y) \in I_{\rho;\gamma}^{(j)} \times \mathbf{T}^n + \left(\frac{7}{4} r_{j+1}, 2 s_{j+1}\right)$, $(\hat{\xi}, \hat{\eta}) = R_{j+1}^{-1} \circ \Psi_j^{-1}(\hat{x}, \hat{y}) \in I_{\rho;\gamma}^{(j)} \times \mathbf{T}^n + \left(2 r_{j+1}, \frac{10}{4} s_{j+1}\right)$ for $(\xi, \eta) \in V_{j+1}$ if $\delta \leq \delta_6$. So,

$$\begin{aligned}
|\hat{\xi} - \xi| &= \left| \frac{\partial \psi_j}{\partial y}(\hat{\xi}, \hat{y}) - \frac{\partial \psi_j}{\partial y}(\xi, y) + \frac{\partial b_j}{\partial y}(\hat{x}, y) \right| \\
&\leq I_1 + I_2 + I_3 + I_4, \\
|\hat{\eta} - \eta| &= \left| \frac{\partial \psi_j}{\partial \xi}(\hat{\xi}, \hat{y}) - \frac{\partial \psi_j}{\partial \xi}(\xi, y) + \frac{\partial b_j}{\partial x}(\hat{x}, y) - \omega^{(j+1}(\hat{\xi}) + \omega^{(j)}(\hat{x}) \right| \\
&\leq J_1 + J_2 + J_3 + J_4 + J_5 + J_6 + J_7,
\end{aligned}$$

where

$$I_1 = \left| \frac{\partial \psi_j}{\partial y}(\hat{\xi}, \hat{y}) - \frac{\partial \psi_j}{\partial y}(\xi, \hat{y}) \right|,$$

$$I_2 = \left| \frac{\partial \psi_j}{\partial y}(\xi, \hat{y}) - \frac{\partial \psi_j}{\partial y}(\xi, y + \omega^{(j)}(\xi)) \right|,$$

$$I_3 = \left| \frac{\partial \psi_j}{\partial y}(\xi, y + \omega^{(j)}(\xi)) - \frac{\partial \psi_j}{\partial y}(\xi, y) + \frac{\partial b_j}{\partial y}(\xi, y) \right|,$$

$$I_4 = \left| \frac{\partial b_j}{\partial y}(\hat{x}, y) - \frac{\partial b_j}{\partial y}(\xi, y) \right|,$$

$$J_1 = \left| \frac{\partial \psi_j}{\partial \xi}(\hat{\xi}, \hat{y}) - \frac{\partial \psi_j}{\partial \xi}(\xi, \hat{y}) \right|,$$

$$J_2 = \left| \frac{\partial \psi_j}{\partial \xi}(\xi, \hat{y}) - \frac{\partial \psi_j}{\partial \xi}(\xi, y + \omega^{(j)}(\xi)) \right|,$$

$$J_3 = \left| \frac{\partial \psi_j}{\partial \xi}(\xi, y + \omega^{(j)}(\xi)) - \frac{\partial \psi_j}{\partial \xi}(\xi, y) - \frac{\partial \omega^{(j)}}{\partial \xi}(\xi) \frac{\partial \psi_j}{\partial y}(\xi, y + \omega^{(j)}(\xi)) + \frac{\partial \tilde{b}_j}{\partial x}(\xi, y) \right|,$$

$$J_4 = \left|\frac{\partial b_j}{\partial x}(\hat{x}, y) - \frac{\partial b_j}{\partial x}(\xi, y)\right|,$$

$$J_5 = \left|\omega^{(j)}(\hat{x}) - \omega^{(j)}(\xi) - \frac{\partial \omega^{(j)}}{\partial \xi}(\xi)(\hat{x} - \xi)\right|,$$

$$J_6 = \left|\frac{\partial \omega^{(j)}}{\partial \xi}(\xi)\Big(\frac{\partial \psi_j}{\partial y}(\xi, y + \omega^{(j)}(\xi)) - \frac{\partial \psi_j}{\partial y}(\xi, y) + \frac{\partial \tilde{b}_j}{\partial y}(\hat{x}, y)\Big)\right|,$$

$$J_7 = \left|\frac{\partial [b_j]}{\partial x}(\hat{\xi}) - \frac{\partial [b_j]}{\partial x}(\xi)\right| + \left|\omega^{(j)}(\hat{\xi}) - \omega^{(j)}(\xi)\right|.$$

By Taylor formula and Cauchy's estimates with the notice that all the concerned variables are in the corresponding shrunken domains with shrunken width, say, $(\frac{1}{4}r_{j+1}, \frac{1}{4}s_{j+1})$, we obtain

$$I_1 \le n \cdot \left(\frac{1}{4}r_{j+1}\right)^{-1} \left|\frac{\partial \psi_j}{\partial y}\right|_{V_j^*} \cdot \left|\hat{\xi} - \xi\right| \le s_{j+1}^{-\lambda-\tau} \cdot 4n\sigma_1 \varepsilon_a^j \cdot \left|\hat{\xi} - \xi\right|,$$

$$I_2 \le n \cdot \left(\frac{1}{4}s_{j+1}\right)^{-1} \left|\frac{\partial \psi_j}{\partial y}\right|_{V_j^*} \cdot \left|\hat{y} - y - \omega^{(j)}(\xi)\right| \le s_{j+1}^{-2\tau-1} \cdot \sigma_2 \varepsilon_a^j \cdot \gamma \varepsilon_b^j$$

with $\sigma_2 = 8n\sigma_1(\sigma_1 + 3)$, here the following estimate has been used

$$\begin{aligned}
\left|\hat{y} - y - \omega^{(j)}(\xi)\right| &= \left|\omega^{(j)}(\hat{x}) - \omega^{(j)}(\xi) + \frac{\partial b_j}{\partial x}(\hat{x}, y)\right| \\
&\le 2\gamma|\hat{x} - \xi| + \left|\frac{\partial b_j}{\partial x}\right|_{\tilde{V}_j^*} \\
&\le 2\gamma\left|\frac{\partial b_j}{\partial y}\right|_{\tilde{V}_j^*} + 2\gamma\left|\frac{\partial \psi_j}{\partial y}\right|_{V_j^*} + \left|\frac{\partial b_j}{\partial x}\right|_{\tilde{V}_j^*} \\
&\le s_{j+1}^{-\tau} \cdot 2(\sigma_1 + 3) \cdot \gamma \varepsilon_b^j. \qquad (4.32)
\end{aligned}$$

I_3 is bounded by (cf. [10], p. 684)

$$I_3 \le \left|\frac{\partial b_j}{\partial y} - T_{m_j}\frac{\partial b_j}{\partial y}\right|_{I_{\rho;\gamma}^{(j)} \times \mathbf{T}^n + \left(\frac{6}{4}r_{j+1}, \frac{6}{4}s_{j+1}\right)} \le s_{j+1}^{\alpha-\lambda-\tau-1} \cdot \gamma \varepsilon_b^j,$$

if s_0 is small enough but depends only on n, τ, λ, α. From (4.32), we get the estimate $|\hat{x} - \xi| \le s_{j+1}^{-\tau} \cdot (\sigma_1 + 2) \cdot \varepsilon_b^j$, which implies

$$I_4 \le n \cdot \left(\frac{1}{4}r_{j+1}\right)^{-1} \left|\frac{\partial b_j}{\partial y}\right|_{\tilde{V}_j^*} \cdot |\hat{x} - \xi| \le s_{j+1}^{-\lambda-\tau} \cdot 8n(\sigma_1 + 2)\varepsilon_a^j \cdot \gamma \varepsilon_b^j.$$

Noting that $\varepsilon_a^j = s_{j+1}^{a-1} \cdot c_{2,a}\mu\delta_a \leq s_{j+1}^{a-1} \cdot \sigma_2^{-1}$ if δ is small enough, say,

$$\delta \leq \delta_7 = (\sigma_2 c_{2,a})^{-1},$$

we get, by combining the above estimates,

$$\left|\hat{\xi} - \xi\right| \leq s_{j+1}^{a-\lambda-\tau-1} \cdot 6\gamma\varepsilon_b^j. \tag{4.33}$$

By estimating J_k, $k = 1, \cdots, 7$ in a similar way to the above with making use of the previous estimates, we obtain

$$|\hat{\eta} - \eta| \leq s_{j+1}^{a-\lambda-2\tau-1} \cdot 6\gamma\varepsilon_b^j, \tag{4.34}$$

if δ is smaller, say,

$$\delta \leq \delta_8 = \left(\sigma_2^2 c_{2,a}\right)^{-1}$$

and s_0 is chosen to be small as before. To summarize, if $\delta \leq \delta_0 = \min_{1\leq l\leq 8} \delta_l$ with s_0 sufficiently small and depending on n, τ, λ, α, then

$$|C_{j+1} - I|_{V_{j+1}} \leq 6s_{j+1}^{(\alpha-1)\lambda} \cdot \gamma\varepsilon_b^j \tag{4.35}$$

due to the fact that $(\alpha-1)\lambda < a - \lambda - 2\tau - 1 < a - \lambda - \tau - 1$. As argued in [10], p. 688, for $b = a$, (4.35) is bounded by $\gamma\varepsilon_a^{j+1}$ with an appropriate choice of s_0, say, $s_0^{(\alpha-1)\lambda} \leq 4^{-(a-1)}/6$. And for any finite $b \geq a$, we can also bound (4.35) by $\gamma\varepsilon_b^{j+1}$ in only finitely many iteration steps — the number of steps needed is, say, $N_{\beta,\alpha} = (\beta-1)/(\alpha-1)$ and, of course, the constant $c_{2,b}$ involved in ε_b^j has to be adjusted to a larger one, say, $4^{(\beta-\alpha)\lambda N_{\beta,\alpha}} c_{2,b}$, and the other related constants $c_{1,b}$, $c_{3,b}$ also change accordingly, but all of these constants do not change for $b = a$, which implies that the smallness of δ required by the induction does not change and therefore, does not depend on β. This shows that the iteration from j-th step to $(j+1)$-th step may be carried out and therefore, Lemma 3.1 is proved.

5. Proof of Theorem 1.2

To prove Theorem 1.2, we only need to reexamine Lemma 3.1 and its proof. Under the assumption (1.2)′ for ω, the corresponding nondegeneracy condition for $\tilde{\omega}$ in Lemma 3.1 is

$$\gamma\mu\,|dx| \leq |d\tilde{\omega}(x)| \leq \gamma\,|dx|, \quad x \in I_\rho + 2 \tag{5.1}$$

with $\mu = \theta\Theta^{-1}$ and $\rho = \gamma\Theta^{-1}$. Accordingly, the equation (3.1) of Lemma 3.1 turns out to be

$$\left(1 - \sum_{k=1}^{j} s_k\right) \gamma\mu\,|dx| \le |d\tilde{\omega}(x)| \le \left(1 + \sum_{k=1}^{j} s_k\right) \gamma\,|dx|\,, \quad x \in I_{\rho;\gamma}^{(j)} + r_j \qquad (5.2)$$

and all other assumptions and conclusions in Lemma 3.1 remain unchanged. The smallness condition for h, of the form (1.11), and therefore, the smallness condition $0 < \delta_a \le \delta$ with the corresponding notation

$$\|f\|_{\beta\lambda+\lambda+\tau;I_\rho\times\mathbf{T}^n} = \gamma\mu^2\delta_b, \quad b = \beta\lambda + \lambda + \tau \ge a = \alpha\lambda + \lambda + \tau, \qquad (5.3)$$

instead of (2.25), and with f defined from h by (2.3), is needed only for proving the existence of $I_{\rho;\gamma}^{(j+1)}$ such that $\omega^{(j+1)}\big(I_{\rho;\gamma}^{(j+1)}\big) = \Omega_\gamma$ from j-th step to $(j+1)$-th step in induction. But this is no problem because in this case, to apply Lemma 4.4, we need only to let $C = \{x^*\} + \frac{1}{8n^2}\mu r_j$ without any other change. From (5.2) and the induction assumption for the first j steps, we have

$$\frac{2}{3}\gamma\mu\,|dx| \le |dF(x)| \le \frac{4}{3}\gamma\,|dx|\,, \quad x \in D \qquad (5.4)$$

which implies that $\left|\frac{\partial F}{\partial x}\right|_D \le \frac{4}{3}\gamma$. By Taylor's formula for $F(x_1) - F(x_2)$ up to second order and Cauchy's estimate with the notice that $C + 4^{-1}r_j \subset D$ and $\left|\frac{\partial^2 F}{\partial x \partial x}\right|_C \le \left(4^{-1}r_j\right)^{-1} \cdot \frac{4}{3}\gamma$, we easily show that, for $x_1, x_2 \in C$,

$$|F(x_1) - F(x_2)| \ge \frac{1}{3}\gamma\mu\,|x_1 - x_2|\,. \qquad (5.5)$$

As a result, we get $L = \min_{x\in\partial C}\{\|F(x) - y\|_2\} \ge \frac{1}{24n^2}\gamma\mu^2 r_j$ with $y = \omega^{(j)}(x^*)$ as assumed tacitly. Therefore, $\sup_{x\in\bar{C}}\{\|F(x) - G(x)\|_2\} \le \sqrt{n}\left|\frac{\partial b_j}{\partial x}\right|_{\tilde{V}_j^*} \le \frac{1}{8}\cdot\frac{1}{24n^2}\gamma\mu^2 r_j$, if δ is small enough but only depending on n, τ, λ and α because the assumption (5.3) will certainly lead to such an estimate for b_j, which guarantees the existence of $I_{\rho;\gamma}^{(j+1)}$ as required. The proof of the remainder is standard.

6. Application to small twist problem

A direct application of the above theorems gives the existence of invariant tori with a smooth foliation structure (differentiable, C^∞-smooth or analytic according to the corresponding smoothness of the considered mapping respectively) of a nearly integrable symplectic mapping with a small twist. The result may be formulated as follows.

Theorem 6.1. *Under the assumptions of Theorem 1.1, consider one parameter family of mappings $S_t : (p,q) \to (\hat{p},\hat{q})$ with $S_0 = I$ and $S_1 = S$, to be defined in phase space $I \times \mathbf{T}^n$ by*

$$\begin{cases} \hat{p} = p - t\frac{\partial H}{\partial q}(\hat{p},q) = p - t\frac{\partial h}{\partial q}(\hat{p},q) \\ \hat{q} = q + t\frac{\partial H}{\partial p}(\hat{p},q) = q + t\omega(\hat{p}) + t\frac{\partial h}{\partial p}(\hat{p},q). \end{cases} \tag{6.1}$$

Under the smallness conditions for h of Theorems 1.1 and 1.2 (in the case when only nondegeneracy condition $(1.2)'$ is satisfied by ω), the corresponding conclusions of them are still valid for S_t, $0 < t \leq 1$, only with the following remarks:

1. Ω_γ is replaced by

$$\Omega_{t,\gamma} = \left\{ \omega \in \Omega_* : \left| e^{i\langle k, t\omega\rangle} - 1 \right| \geq \frac{t\gamma}{|k|^\tau} \quad \text{for } k \in \mathbf{Z}^n \setminus \{0\} \right\}, \tag{6.2}$$

which means to depend on both γ and t, where Ω_* denotes the set of points in Ω with distance to its boundary at least equal to 2γ; and accordingly, I_γ is replaced by $I_{t,\gamma}$, a closed subset of I; ω_γ replaced by $\omega_{t,\gamma} : I_{t,\gamma} \to \Omega_{t,\gamma}$, an onto map; Φ replaced by $\Phi_t : I_{t,\gamma} \times \mathbf{T}^n \to \mathbf{R}^n \times \mathbf{T}^n$ and R replaced by $R_t : (\xi,\eta) \to (\xi, \eta + t\omega_{t,\gamma}(\xi))$.

2. If Ω is a bounded open set of type D in Arnold's sense [1], then we have the following Lebesgue measure estimate

$$m(\Omega \setminus \Omega_{t,\gamma}) \leq D\gamma m\Omega \tag{6.3}$$

for $t \in (0,1]$, with constant D only depending on n, τ and the geometry of Ω. So in this case, $\Omega_{t,\gamma}$ is still a large Cantor set in Ω if γ is small enough.

We conclude the note by remarking that Theorem 6.1 implies the existence of invariant tori with smooth foliation structure and therefore, also implies the existence of n independent smooth invariant functions which are in involution and well-defined on the set filled by the invariant tori in the Whitney's sense, of a symplectic numerical integrator applied to an integrable or a nearly integrable Hamiltonian system if the system is nondegenerate and the time-step size of the integrator is small enough. The invariant tori are just those level sets of the n invariant functions. The nondegeneracy of a nearly integrable system means the nondegeneracy of the integrable part of the system. Symplectic integrators or, in other words, symplectic algorithms with their computer performance are referred to [4].

Acknowledgments. The essential part of this work was carried out during the author's visit to Max-Planck-Institut für Mathematik in Bonn. He thanks MPIM for their hospitality and Max-Planck-Gesellschaft for their financial support under the exchange program with Academia Sinica. The author was partially supported by the National Natural Science Foundation of China.

References

1. Arnold V. I., Proof of A. N. Kolmogorov's theorem on the preservation of quasi-periodic motions under small perturbations of the Hamiltonian, *Russ. Math. Surv.* **18:5** (1963), 9-36.
2. Arnold V. I. and A. Avez, Ergodic Problems of Classical Mechanics, New York, 1968.
3. Cheng C. Q. and Sun Y. S., Existence of KAM tori in degenerate Hamiltonian systems, *J. Differential Equations,* **114:1**(1994), 288-335.
4. Feng Kang, The Hamiltonian way for computing Hamiltonian dynamics, in "Applied and Industrial Mathematics", Ed. R. Spigler, Netherlands: Kluwer, 1991, 17-35.
5. Gentile G. and Mastropietro V., Methods for the analysis of the Lindstedt series for KAM tori and renormalizatibility in classical mechanics – A review with some applications, *Reviews in Math. Phy.* **8:3**(1996), 393-444.
6. Kolmogorov A. N., On conservation of conditionally periodic motions under small perturbations of the Hamiltonian, *Dokl. Akad. Nauk. SSSR* **98** (1954), 527–530.
7. Moser J., On invariant curves of area–preserving mappings of an annulus, *Nachr. Akad. Wiss. Gott. Math. Phys. Kl.,* 1962, 1–20.
8. Neishtadt A. I., Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions, *J. Appl. Math. Mech.* **45:6**(1982), 766-772.
9. Ortega J. M. and Rheinboldz W. C., Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.
10. Pöschel J., Integrability of Hamiltonian systems on Cantor sets, *Commun.*

Pure Appl. Math. **35** (1982), 653-695.

11. Rüssmann H., On the existence of invariant curves of twist mappins of an anulus, in "Geometric Dynamics", Proceedings, Rio de Janeiro 1981, Ed. J. Palis, Lecture Notes in Math. **1007,** Springer-Verlag, 677–718.

12. Rüssmann H., On twist Hamiltonian, in "Colloque International: Mécanique céleste et systèmes hamiltoniens, Marseille, 1990.

13. Rüssmann H., On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus, Lect. Notes Phys. **38,** 1975, 598-624.

14. Svanidze N. V., Small perturbations of an integrable dynamical system with an integral invariant, Proceedings of the Steklov Institute of Mathematics. 1981, Issue 2, 127-151.

PERMANENT ADDRESS: INSTITUTE OF MATHEMATICS, ACADEMIA SINICA, BEIJING, 100080, P. R. CHINA

E-MAIL: zaijiu@math03.math.ac.cn

CURRENT ADDRESS: MAX-PLANCK-INSTITUT FÜR MATHEMATIK, GOTTFRIED-CLAREN-STRASSE 26, 53225, BONN, GERMANY

E-MAIL: shang@mpim-bonn.mpg.de