A KAM theorem of degenerate infinite dimensional Hamiltonian systems (I)*

XU Junxiang (徐君祥), QIU Qingjiu (仇庆久) and YOU Jiangong (尤建功)

(Department of Mathematics, Nanjing University, Nanjing 210093, China)

Received September 15, 1995

Abstract A class of weaker nondegeneracy conditions is given and an existence theorem of invariant tori is proven for small perturbations of degenerate integrable infinite dimensional Hamiltonian systems under the weaker nondegeneracy conditions. The measure estimates of the parameter set are also given for which invariant tori exist. It is valuable to point out that by the motivation of finite dimensional situation the nondegeneracy conditions may be the weakest. Mainly KAM machine is used to prove the existence of invariant tori. The measure estimates for small divisor conditions, on which the measure estimates of the parameter set are based, will be given in the second paper.

Keywords: Hamiltonian systems, perturbation, invariant tori, small divisor conditions, KAM iteration.

1 Introduction and main theorem

This paper considers the problem of infinite dimensional Hamiltonian systems arising from some partial differential equations (see refs. [1—3]). In refs. [2, 3] the authors gave a KAM theorem under the nondegeneracy conditions of frequency with respect to parameters. In this paper we study the degenerate situation and prove a similar KAM theorem but under some weaker nondegeneracy conditions. This KAM theorem can also be applied to studying the existence of quasiperiodic solutions to one-dimensional nonlinear Schrödinger equations and wave equations, which will be discussed in another paper. Moreover, because our proofs are very long, we will finish them in two papers, (I) and (II).

Consider the following infinite dimensional Hamiltonian:

\[ \tilde{H} = \sum_{j \neq k \alpha} \omega_j(\xi) y_j + \frac{1}{2} \sum_{j \neq 1} \Omega_j(\xi)(u_j^2 + v_j^2). \]

On the phase space \((x, y, u, v) \in \Gamma^a = T^n \times \mathbb{R}^n \times \mathcal{S}^a \times \mathcal{S}^{a, p}, \) where \(T^n(1 \leq n < + \infty)\) is the usual \(n\)-torus, \(\mathbb{R}^n\) is \(n\)-dimensional Euclidean space and \(\mathcal{S}^{a, p}\) is the Hilbert space of all real (later complex) sequences \(u = (u_1, u_2, \ldots)\) with the norm defined by \(\|u\|_{a, p}^2 = \sum_{j \geq 1} |u_j|^2 j^{2p} e^{2|j|} < \infty. \)

* Project supported by the National Natural Science Foundation of China.
The motion equations of Hamiltonian $\tilde{H}$ are
\[ \dot{x} = \omega(\xi), \quad \dot{y} = 0, \quad \dot{u} = \Omega(\xi) v, \quad \dot{v} = -\Omega(\xi) u, \]
where the dot indicates the differentiation with respect to time $t$, $(\Omega u)_t = \dot{\Omega} u$. Obviously, for each $\xi \in \Pi$, $\tau = T^\pi \times \{0, 0, 0\} \subset T^\pi$ is an invariant torus for the Hamiltonian systems (1.1) with the frequencies $\omega(\xi) = (\omega_1(\xi), \cdots, \omega_r(\xi))$, so (1.1) are integrable.

Now we consider the small perturbations of $\tilde{H}$: $H = \tilde{H} + P$, where $P$ is a small perturbation. The corresponding Hamiltonian motion equations are
\[ \dot{x} = \omega(\xi) + P_x, \quad \dot{y} = -P_y, \quad \dot{u} = \Omega(\xi) v + P_v, \quad \dot{v} = -\Omega(\xi) u - P_v \]
(1.2)

Our main problem is under which conditions the Hamiltonian systems (1.2) keep many invariant tori. For this we first give some notations and assumptions.

Let $C^{N,1}(\Pi)$ be the $N$-order Lipschitz continuously differentiable function space. If $\Pi$ is a closed set, the derivatives of function on $\Pi$ are understood in the sense of Whitney (see definition A1 in Appendix, for some related contents see refs. [4, 5]), so the space $C^{N,1}(\Pi)$ is also understood in the sense of Whitney, where the integer $N$ will be decided below.

**Assumption A (Nondegeneracy Conditions).** Suppose for $\forall \xi \in \Pi$
\[
\begin{cases} 
\text{rank} \left\{ \frac{\partial \omega}{\partial \xi} \right\} = r, \\
\text{rank} \left\{ \frac{\partial^\beta \omega}{\partial \xi^\beta} \left| \forall \beta, 1 \leq |\beta| \leq n-r+1 \right\} = n,
\end{cases}
\]
where $\frac{\partial \omega}{\partial \xi}$ is a function vector group of all 1-order partial derivatives of $\omega$, and
\[
\frac{\partial^\beta \omega}{\partial \xi^\beta} = \left( \frac{\partial^\beta \omega_1}{\partial \xi^\beta}, \cdots, \frac{\partial^\beta \omega_r}{\partial \xi^\beta} \right).\] Moreover, for some $N > n-r+1$, $\omega$ belongs to $C^{N,1}(\Pi)$ with
\[
\|\omega\|_{C^{N,1}(\Pi)} = \max_{1 \leq j \leq N} \|\omega_j\|_{C^{N,1}(\Pi)} \leq M.
\]

**Assumption B (Spectral Asymptotics).** There exist $d \geq 1$ and $\delta < d - 1$ such that
\[
\Omega(\xi) = b^d j^d + b^d j^d + \cdots + O(j^d), \quad b > 0, \quad d < d',
\]
where the dots stand for finite lower order terms of $j$ and $b^d j^d + b^d j^d + \cdots$ are independent.
of the parameter $\xi$. Moreover, $\Omega_j$ satisfies

$$\|\Omega_j - b^j d - b' j^k - \cdots\|_{C^{N_1(N)}} \leq M_j j^k, \quad \forall j \geq 1.$$ 

Assumption C (Regularity of perturbations). The perturbation $P$ is real analytic in $(x, y, u, v)$ and belongs to $C^{N_1'(II)}$ in $\xi$, and for $\forall \xi \in \Pi$, the Hamiltonian vector field of $P$, $X_p = (P_x, -P_y, P_u, -P_v)^T$: $\Gamma^{\alpha, \beta} \rightarrow \Gamma^{\alpha, \beta}$ is real analytic with $p \leq \bar{p}$ and $p - \bar{p} < d - 1$. Without loss of generality, suppose $p - \bar{p} < \delta$.

Denote a complex neighbourhood of $\tau^*$ in $\Gamma^{\alpha, \beta}$ by

$$D(s, r) = \{(x, y, u, v) | | \text{Im} x | < s, |y| < r^2, \|u\|_{a, \bar{p}} + \|v\|_{a, \bar{p}} < r\},$$

where Imx is the image part of $x$ and $|\cdot|$ is the supnorm for $n$-dimensional vector, $r > 0$ is a radius of neighbourhood (notice that it has sense different from that in Assumption A). Let $\|\cdot\|_*=\|\cdot\|_{C_{N_1}(I)}$ for simplicity. Suppose $W(\xi) = (X(\xi), Y(\xi), U(\xi), V(\xi)) \in \Gamma^{\alpha, \beta}$. Define weighted norms by

$$\|W\|_* = \|W\|_{r, \bar{p}} = |X|^* + \frac{1}{r^2} |Y|^* + \frac{1}{r} \|U\|_{a, \bar{p}}^* + \frac{1}{r} \|V\|_{a, \bar{p}}^*,$$

where

$$|X|^* = \sum ||X_i||, |Y|^* = \sum ||Y_i||, \|U\|_{a, \bar{p}}^* = \sum ||U_i||^*, \|V\|_{a, \bar{p}}^* = \sum ||V_i||^* \text{ in } a_r.$$ 

If $f(\eta, \xi)$ is a function defined on $D \times \Pi$, denote by

$$\|f\|_D = \sup_{\eta \in D} \|f(\eta, \cdot)\|_{C^{N_1}(I)}.$$

Similarly define

$$\|W\|_{r, \bar{p}, \bar{d}} = \|W\|_{r, \bar{d}}, d = |X|^d + \frac{1}{r^2} |Y|^d + \frac{1}{r} \|U\|_{a, \bar{p}, \bar{d}}^d + \frac{1}{r} \|V\|_{a, \bar{p}, \bar{d}}^d.$$ 

Let $k \in \mathbb{Z}^n$ and $l \in \mathbb{Z}^{+\infty}$, and denote by $|k| = \sum |k_i|$, $|l| = \sum |l_j|$, $|l|_d = \sum |l_j| j^d$ and $|l|_a = \max \{1, \sum |l_j| j^d \}$. Let $A_k = (1 + |k|)^t$, where $t \geq \tau_0$ with

$$\tau_0 = \begin{cases} \left( n + \frac{2}{d - 1} \right)(n - r + 1), & \text{if } d > 1, \\ \left( n + 1 + \frac{\kappa + n - r + 1}{\kappa} \right)(n - r + 1), & \text{if } d = 1, \end{cases}$$

where $\kappa = \min\{d, d - 1 - \delta, d - d'\}$.

**Theorem A.** Suppose the Hamiltonian $H = \tilde{H} + P$ satisfies Assumptions A, B and C. Then there exist $K$ and $L$ such that if for $k$, $l$, $0 \neq |k| \leq K$, $|l| \leq 2$ and $|l|_a \leq L$, the following inequality
\[ \langle k, \omega(\xi) \rangle + \langle l, \Omega(\xi) \rangle \geq \alpha_0 > 0 \]

holds for all \( \xi \in \Pi \), then for sufficiently small \( \alpha > 0 \) (\( \alpha \leq \alpha_0 \)), there exists sufficiently small \( \epsilon = \epsilon_0 > 0 \) such that if \( \epsilon = \|X_p\|_{\omega, \Omega, n} < \epsilon_0 \), then there exist a nonempty Cantorian subset \( \Pi_k \) of \( \Pi \) and embeddings \( \Phi_k(\cdot, \cdot) : T^n \times \Pi \rightarrow T^n \times \overline{\Pi}_k \) satisfying

\[ \|\Phi_k - \Phi_0\|_r \leq \epsilon, \]

where \( \Phi_0 \) is the trivial embedding \( T^n \times \Pi \rightarrow T^n \times \{0, 0, \cdots, 0\} \subset T^n \times \overline{\Pi}_k \), and for \( \xi \in \Pi_k \), \( \Phi_k(\cdot, \cdot) \) is a real analytic embedding of a rotational torus for Hamiltonian systems (1.2) at \( \xi \) with its frequencies \( \omega(\xi) \) satisfying \( |\omega_1(\xi) - \omega(\xi)| \leq \epsilon \). Moreover, \( \text{mes} (\Pi - \Pi_k) \leq c (\text{diam} \Pi)^{n+1} \alpha^n \), where \( \mu = \frac{1}{n-r+1} \) if \( d > 1 \) and \( \mu = \frac{\kappa}{(\kappa + n-r+1)(n-r+1)} \) if \( d = 1 \).

The Cantorian subset \( \Pi_k \) in Theorem A is determined by the small divisor conditions at KAM step,

\[ \langle k, \omega(\xi) \rangle + \langle l, \Omega(\xi) \rangle \geq \alpha \frac{[l]}{A_k}, \quad \forall k \in \mathbb{Z}^n, \forall l \in \mathbb{Z}^{n+1}, \forall \|l\| \leq 2, \|k\| + \|l\| \neq 0. \quad (1.3) \]

That is, KAM iteration can only be done for the parameters that (1.3) hold, so we must exclude some points where (1.3) do not hold. Theorem B below guarantees that if \( \alpha \) is sufficiently small, \( \Pi_k \) is nonempty.

**Theorem B.** Let

\[ R_{k, l}(\alpha) = \left\{ \xi \middle| \langle k, \omega(\xi) \rangle + \langle l, \Omega(\xi) \rangle < \alpha \frac{[l]}{A_k} \right\}, \]

where \( k \in \mathbb{Z}^n, l \in \mathbb{Z}^{n+1}, \|l\| \leq 2, \|k\| + \|l\| \neq 0 \). If Assumptions A and B hold, then there exist \( K \) and \( L \) such that for given arbitrary small \( \alpha > 0 \),

\[ \text{mes} \left( \bigcup_{\|k\| > K \text{ or } \|l\| > L} R_{k, l}(\alpha) \right) \leq c (\text{diam} \Pi)^{n+1} \alpha^n, \]

where \( \mu \) is defined in Theorem A.

**Remark 1.** If rank \( \left\{ \frac{\partial \omega}{\partial \xi} \right\} = n \), Assumption A is the nondegeneracy case of frequency \( \omega \) with respect to parameter \( \xi \), which is studied in refs. [2, 3, 6]. In this paper we are interested in the case rank \( \left\{ \frac{\partial \omega}{\partial \xi} \right\} < n \), that is, the corresponding Hamiltonian systems are degenerate. If \( \omega(\xi) \) is analytic in \( \xi \), our nondegeneracy conditions are almost the sharpest (see [7]). If \( \omega \) is not analytic in \( \xi \), but \( \omega \in C^N'(\Pi) \), then the nondegeneracy conditions in Assumption A can be replaced by

\[ \text{rank} \left\{ \frac{\partial k \omega}{\partial \xi^\beta} \middle| \forall \beta, 1 \leq \|\beta\| \leq N \right\} = n, \forall \xi \in \Pi. \]
Thus Theorem A still holds with a little modification of the estimates of \( \text{mes}(\Pi - \Pi) \), that is, all \( n-r+1 \) in the estimates should be replaced by \( N \), where \( N \) is some fixed positive integer.

**Remark 2.** \( \varepsilon_0 \) in Theorem A depends not only on \( \alpha \) but also on the constants in all assumptions. Moreover, \( \varepsilon_0 \to 0 \) as \( \alpha \to 0 \), so \( \text{mes}(\Pi - \Pi) \to 0 \) as \( \varepsilon \to 0 \).

In the following KAM iteration we always use the same \( \varepsilon \) to denote the constants of estimates, which are independent of KAM step.

### 2 Proof of Theorem A

Since Theorem A and Theorem B hold separately, we first prove Theorem A under the assumption that Theorem B holds. The proof of Theorem B and some results related to derivatives of Whitney needed at KAM step will be given in the proof of Theorem B and Appendix in the second paper.

The proof of Theorem A is actually concrete application of KAM theory, so we only state the main step of KAM iteration, for some details of KAM technique, see refs. [2, 3, 6, 8], etc.

First do the transformation \( z = \frac{u - iv}{\sqrt{2}}, \bar{z} = \frac{u + iv}{\sqrt{2}} \). Thus we can consider the Hamiltonian \( H(x, y, z, \bar{z}) = \tilde{H} + P \) in the complex conjugate coordinates, where \( \tilde{H} = \langle \omega(\xi), y \rangle + \langle \Omega(\xi) z, \bar{\xi} \rangle \) and \( P = \sum_{k<q} P_{kmq} z^m \bar{z}^q e^{i(k \cdot x)} \), \( (x, y, z, \bar{z}) \in T^n \times \mathbb{R}^n \times L^n \times L^n \), where \( L^n \) is the Hilbert space consisting of complex sequences, which is defined as the real case. Moreover, \( P_{kq} \) depends on the parameter \( \xi \in \Pi \).

At each step of KAM iteration, the symplectic coordinate change \( \Phi \) is obtained as the time 1-map \( X^n_{\Phi} \) of the flow of Hamiltonian vector field \( X_\varphi \). Its generating function \( F \) and some normal correction \( \hat{H} \) to the given normal form \( N \) are solutions of the linear equation

\[
\{ F, \tilde{H} \} + \hat{H} = R,
\]

(2.1)

where \( \{ \cdot, \cdot \} \) is the Poisson product, \( R \) is the lower order approximation of perturbation \( P \) in \( y, z, \bar{z} \), that is,

\[
R = \sum_{2|m|+|q|+|\bar{q}| \leq 2} R_{kmq\bar{q}} z^m \bar{z}^q \bar{q}^\bar{q} e^{i(k \cdot x)}, \quad R_{kmq\bar{q}} = P_{kmq\bar{q}}, \quad 2|m|+|q+\bar{q}| \leq 2.
\]

Choose the normal correction

\[
\hat{H} = [R] = \sum_{2|m|+|q|+|\bar{q}| \leq 2, \bar{q} = \bar{q}} R_{kmq\bar{q}} z^m \bar{z}^q \bar{q}.
\]
Thus (2.1) is a linear equation of $F$. Below we solve this linear equation and estimate the generating function $F$.

Lemma 2.1. Suppose that uniformly on $\Pi \subset \Pi$

$$\langle k, \omega(\xi) \rangle + \langle l, \Omega(\xi) \rangle \geq \alpha \frac{|l|}{A_k}$$

for $\forall k \in \mathbb{Z}^n$, $\forall l \in \mathbb{Z}^n$ with $|l| \leq 2$, $|k| + |l| \neq 0$. Then the linear equation (2.1) has solutions $F$ and $\hat{H}$, which satisfy $[F] = 0$, $[\hat{H}] = \hat{H}$. Moreover,

$$||X_\xi||_{\mathcal{L}(\mathcal{D}(\zeta, r))} \leq ||X_R||_{\mathcal{L}(\mathcal{D}(\zeta, r))} \quad \text{and} \quad ||X_F||_{\mathcal{L}(\mathcal{D}(\zeta, r))} \leq \frac{cM}{\alpha^{N+2}}$$

where $v = (N+2)\tau + N \tau + n + \frac{n}{2}$, $M = (M_1 + M_2)^{N+2}$ and $|| \cdot ||^* = || \cdot ||_{C^{n,1}((\Pi_+)}$.

Proof. Since the first inequality holds obviously, we only prove the second inequality. Expand $F = \sum_{2m \mid |q|+1 \leq 2} F_{k=m} y^m \mathcal{D}(\zeta, \mathcal{D}(\zeta, r))$ and substitute it into (2.1). By comparing the coefficients, we have

$$i F_{k=m} = \frac{R_{k=m} y^m}{\langle k, \omega \rangle + \langle q - q, \Omega \rangle}$$

for $|k| + |q - q| \neq 0$,

otherwise.

We decompose $R = R^0 + R^1 + R^2$, where $R^i$ comprises all terms with $|q + q| = i$, $i = 0, 1, 2$. Thus we write them as $R^0 = \langle R^0, \xi \rangle + \langle R^0, \zeta \rangle$, $R^1 = \langle R^1, \xi \rangle + \langle R^1, \zeta \rangle$, $R^2 = \langle R^2, \xi \rangle + \langle R^2, \zeta \rangle$, where $R^0$ depend on $x$ and $\xi$, and $R^0$ depends in addition on $y$. Similarly we decompose $F$ and $\hat{H}$. Below we can consider the above terms individually, but without loss of generality we only consider $\hat{R} = R^0$ and $\hat{R} = R^1$. For simplicity, let $|| \cdot ||^* = || \cdot ||_{\mathcal{L}(\mathcal{D}(\zeta, r))}$. Then, $|| \cdot ||$, and $|| \cdot ||_{\mathcal{L}(\mathcal{D}(\zeta, r))}$ are correspondingly defined.

Since $\hat{R} = R_k |_{z = z = 0}$, from the definition of $|| \cdot ||_{\mathcal{L}(\mathcal{D}(\zeta, r))}$ we have

$$||\hat{R}||_{\mathcal{L}(\mathcal{D}(\zeta, r))} \leq r ||X_k||_{\mathcal{L}(\mathcal{D}(\zeta, r))} r ||X_k||_{\mathcal{L}(\mathcal{D}(\zeta, r))},$$

where $D(s) = \{ x \mid |\text{Im} x| < s \}$. Since $\hat{R}$ is an analytic mapping from $D(s)$ into $\mathcal{L}(\mathcal{D}(\zeta, r), \mathcal{L}(\mathcal{D}(\zeta, r))$, its Fourier coefficients $\hat{R}_k$ satisfy $L^2$-estimates

$$\sum_k ||\hat{R}_k||^2 e^{2\pi k} \leq 2^2 ||\hat{R}||^2_{\mathcal{L}(\mathcal{D}(\zeta, r))}.$$
\[ \tilde{\omega}(\xi) = (\omega(\xi), \Omega_\xi(\xi)) \in \Omega_+, \quad \text{if } i \neq j, \quad \text{and} \quad \frac{1}{\langle \omega, \Omega_\xi \rangle + \Omega_j(\xi)} = G_\xi[\tilde{\omega}(\xi)], \quad \text{where } \bar{k} = (k, 1) \in \mathbb{Z}^{d+1}. \]

By deriving the composite function \( G_\xi[\tilde{\omega}(\xi)] \), we have
\[
\| G_\xi[\tilde{\omega}(\xi)] \|^* \leq \frac{c A^{N+2}_k |k|^{N+1}}{\alpha^{N+2} j^d} \left( \| \omega \| + \max_j \| \Omega_j - b_j \| + \cdots \| \omega^j \| \right)^{N+1}.
\]

By Lemma A5 in Appendix of the second paper, it follows
\[
\| \hat{\Phi}_i \| = \frac{c A^{N+2}_k |k|^{N+1} M}{\alpha^{N+2} j^d} \| \hat{\Phi}_j \|.
\]

and
\[
\| \hat{\Phi}_i \| \leq \frac{c A^{N+2}_k |k|^{N+1} M}{\alpha^{N+2} j^d} \| \hat{\Phi}_j \|, \quad \text{where } M = (M_1 + M_2)^{N+1}. \]

Thus
\[
\| \hat{\Phi}_i \| \leq \frac{c M}{\alpha^{N+2} j^d} \| \hat{\Phi}_j \|.
\]

Since \( B_{v} = \left( \sum_k A_k^{2(N+2)} |k|^{2(N+1)} e^{-2\|k\|} \right)^{1/2} \leq \frac{c}{\sigma^r} \) with \( v = (N+2) \tau + N + 1 + \frac{n}{2} \), \( \frac{1}{r} \| \hat{\Phi}_i \| \leq \frac{c M}{\alpha^{N+2} j^d} \| X_{\Phi_i} \| \),
\[
\text{where } \| \hat{\Phi}_j \| \text{ denotes the operator norm of } \hat{\Phi} \text{ from } L^a \text{ to } L^q.
\]

Now we consider \( \hat{\Phi} = \hat{\Phi}_i \) and \( \hat{\Phi} = \hat{\Phi}_j \). Similar to the estimates of \( R_{10} \), it follows that \( \| R_{10} \| \leq r \| X_{\Phi_i} \| \), \( D(k, r) \), and hence
\[
\| \hat{\Phi}_i \| \leq \frac{1}{r} \| \hat{\Phi}_j \| \leq \| X_{\Phi_i} \| \leq \| X_{\Phi_j} \| \leq \| X_{\Phi_i} \| \),
\]

where \( \| X_{\Phi_i} \| \) denotes the operator norm of \( \Phi \) from \( L^a \) to \( L^q \).

Since this operator \( \Phi \) is equivalent to the operator \( \Phi = (V_i \Phi, W_j) \) from \( L^2 = L^{0,0} \) to \( L^2 \) and \( \| \hat{\Phi} \| \leq \| X_{\Phi_i} \| \), \( \| \hat{\Phi} \| \leq \| X_{\Phi_i} \| \), where \( V_i \) and \( W_j \) are some weights, \( \| \| \| \| \) is the operator norm from \( L^2 = L^{0,0} \) to \( L^2 \). Thus we need only consider the norm of \( L^2 \)-operator \( \Phi \).

Let \( \Phi = \sum_k \Phi_k e^{\xi(k, x)} \), \( \Phi = \sum_k \Phi_k e^{\xi(k, x)} \). From eq. (2.1) we know
\[
\langle k, \omega \rangle + \Omega_\xi - \Omega_j \rangle, \quad |k| + |i-j| \neq 0,
\]
where \( \{ \Phi_k | k \in \mathbb{Z}^d \} \) are the Fourier coefficients of \( (i,j) \)th components of \( \Phi \).

Again by Lemma A4 of the second paper and taking \( \tilde{\omega} = (\omega, \Omega_\xi - \Omega_j) \) and \( \bar{k} = (k, 1) \), in the same way as the above and combining \( 2|i^d - j^d| \geq |i-j||i^d-1 + j^d-1| \), we have
\[
\left\| \frac{1}{\langle k, \omega(\xi) \rangle + \Omega(\xi) - \Omega_{i}(\xi)} \right\|^* \leq \frac{cA^{N+2}|k|^{N+1}M}{\alpha^{N+2}|i-j|} \text{ for } i \neq j.
\]

Again by Lemma A5 in Appendix of the second paper

\[
\| \bar{F}_{ij} \|^* \leq \frac{cA^{N+2}|k|^{N+1}M}{\alpha^{N+2}|i-j|} \| \bar{R}_{ij} \|^*, \ i \neq j.
\]

Hence, by Lemma A2 in Appendix of the second paper, \( \|\| \bar{F}_{ij} \|_{L(0,s)} \| \leq \frac{cA^{N+2}|k|^{N+1}M}{\alpha^{N+2}} \|\| \bar{R}_{ij} \|_{L(0,s)} \| \).

Summing up for \( k \) as the above, we have

\[
\|\| \bar{F} \|_{L(0,s)} \| \leq \frac{cB_{ij} M}{\alpha^{N+2}} \|\| \bar{R} \|_{L(0,s)} \|.
\]

Going back to the original operator norm we have

\[
\frac{1}{r} \|\| \bar{F} \|_{D(0,s, r)} \| \leq \frac{cM}{\alpha^{N+2}} \|\| X_{R} \|_{L(0,s, r)} \|.
\]

Thus, this lemma is proved.

From Lemma 2.1, \( X_{F} \) is defined on \( D(s-\sigma, r) \times \Pi_+ \). Denote by \( \| \cdot \| \| \) the norm of \( C^N(\Pi_+) \). By Lemma A7 in Appendix of the second paper and Cauchy's inequality, we have

**Lemma 2.2.** If \( \| X_{F} \|_{L(0,s, r)} \| \leq \sigma \), then for \( \forall \xi \in \Pi_+ \), the flow \( X_{F} \) of the Hamiltonian system \( H = H(s, y, z) \) exists on \( D(s-2\sigma, r/2) \) for \( |t| \leq 1 \) and maps \( D(s-2\sigma, r/2) \) into \( D(s, r) \). Moreover, for \( |t| \leq 1 \)

\[
\| X_{F} - Id \|_{L(0,s-2\sigma, r/2), \Pi_+} \leq c \| X_{F} \|_{L(0,s-2\sigma, r/2), \Pi_+} \| X_{F} \|_{L(0,s, r)} \|,
\]

where \( \mathcal{D} \) is the differentiation operator with respect to \( (x, y, z, \bar{z}) \), \( id \) and \( Id \) are identity mapping and unit matrix, and the operator norm \( \| \cdot \|_{L^1} \) is defined by \( \| L \|_{L^1} = \sup_{x \neq 0} \| LW \|_{L^1} \).

Below we consider the new perturbation under the symplectic transformation \( \Phi = X_{F}^{-1} \). Let \( \varepsilon = \| X_{F} \|_{L(0,s, r)} \| \). From the above we have

\[
R = \sum_{2|n| + \overline{|q|} \leq 2} \sum_{2|q| + \overline{|q|} \leq 2} P_{k, m, q, r} e^{i\langle k, x \rangle}.
\]

Thus we have \( \| X_{F} \|_{L(0,s, r)} \| \leq 2 \| X_{F} \|_{L(0,s, r)} \| \leq 2 \varepsilon \), and for \( \eta \leq \frac{1}{8} \),

\[
\| X_{F} - Id \|_{L(0,s, r)} \| \leq 2 \eta \| X_{F} \|_{L(0,s, r)} \| \leq 2 \eta \varepsilon.
\]

Since \( \hat{H} = \sum_{2|n| + \overline{|q|} \leq 2} P_{k, m, q, r} e^{i\langle k, x \rangle} \), the new normal form is

\[
\bar{H} = \hat{H} + \hat{H} = \langle \omega_{+}, y \rangle + \langle \Omega_{+}, z, \bar{z} \rangle.
\]

By Lemma 2.1 and noticing that \( p - \bar{p} < \delta \), we have \( \| X_{\bar{H}}^* \|_{r, \delta(\bar{p}, \bar{r})} \leq 2 \varepsilon \), so,

\[
|\omega_+ - \omega_+^*|, \| \Omega_+ - \Omega_+^* \|_{\delta} \leq 2 \varepsilon,
\]

where \( \| \Omega \|_{j^*} = \max_{j \geq 1} \| \Omega_j \|_{j^*} j^* \). If \( \frac{\varepsilon}{\alpha^{N+2} \alpha^r+1} \) is sufficiently small, by Lemmas 2.1 and 2.2, it follows that for \( |t| \leq 1 \)

\[
\frac{1}{\sigma} \| X_{\bar{F}}^r - \text{id} \|_{r, \delta(\bar{p} - 2 \varepsilon, \frac{\varepsilon}{\alpha^r})}, \| D^2 X_{\bar{F}}^r - \text{id} \|_{r, \alpha^{(s-3) \varepsilon}} \leq c \varepsilon ME,
\]

with \( E = \frac{\varepsilon}{\alpha^{N+2} \alpha^r+1} \). By Cauchy's inequality

\[
\sigma \| D^2 X_{\bar{F}}^r \|_{r, \alpha^{(s-4) \varepsilon}} \leq c \varepsilon ME,
\]

where the norm

\[
\| D^2 X_{\bar{F}}^r \|_{r, \alpha^{(s-4) \varepsilon}} = \sup_{W^{*} \neq 0} \frac{|| D^2 X_{\bar{F}}^r W ||_{r, \alpha^{(s-4) \varepsilon}}}{|| W ||_{r, \alpha^{(s-4) \varepsilon}}}.
\]

Under the change \( \Phi = X_{\bar{F}}^r, (\bar{H} + R) \circ \Phi = \bar{H} + R_+ \), where \( R_+ = \int_0^1 \{(1-t) \bar{H} + tR, F\} \circ X_{\bar{F}}^r dt \).

Thus, \( H \circ \Phi = \bar{H} + R_+ + (P - R) \circ \Phi = \bar{H} + P_+ \), where the new perturbation

\[
P_+ = R_+ + (P - R) \circ \Phi = (P - R) \circ \Phi + \int_0^1 \{ \bar{R}(t), F\} \circ X_{\bar{F}}^r dt,
\]

with \( \bar{R}(t) = (1-t) \bar{H} + tR \). Hence, the Hamiltonian vector field of the new perturbation is

\[
X_{P_+} = (X_{\bar{F}}^r)^* (X_{\bar{F}}^r) + \int_0^1 (X_{\bar{F}}^r)^*[X_{\bar{F}}^r, X_{\bar{F}}^r] dt,
\]

where \([ \cdot, \cdot \] denotes the commutator of vector fields, \((X_{\bar{F}}^r)^*\) is the cotangent mapping of \(X_{\bar{F}}^r\).

By the constructing of \(X_{\bar{F}}^r\), it follows that for \( \forall \xi \in \Pi_+ \),

\[
X_{\bar{F}}^r(\cdot, \xi): D(s-5 \alpha, \eta \bar{r}) \to D(s-4 \alpha, 2 \eta \bar{r}), \text{ for } |t| \leq 1.
\]

For the estimates of \(X_{P_+}\) we need the following lemma.

**Lemma 2.3.** If the Hamiltonian vector field \( W(\cdot, \xi) \) on \( V = D(s-4 \alpha, 2 \eta \bar{r}) \) depends on the parameter \( \xi \in \Pi_+ \) with \( \| W \|_{\cdot, \nu} < + \infty \), and \( \Phi = X_{\bar{F}}^r: U = D(s-5 \alpha, \eta \bar{r}) \to V \), then \( \Phi^* W = (D \Phi)^{-1} W \circ \Phi \) and if \( E \) is small, we have \( \| \Phi^* W \|_{\cdot, \nu} \leq c \| W \|_{\cdot, \nu} \).

**Proof.** Since the change \( \Phi \) is symplectic, it follows that \( \Phi^* W = (D \Phi)^{-1} W \circ \Phi \) and

\[
\| \Phi^* W \|_{\cdot, \nu} \leq \| (D \Phi)^{-1} \|_{\cdot, \nu} \| W \circ \Phi \|_{\cdot, \nu}.
\]
By (2.5) and $\eta^3 = E$, if $E$ is sufficiently small such that
\[
\frac{1}{\sigma} \|\mathbf{D} - id\|_{\psi, \gamma}^* \|\mathbb{D}\Phi - Id\|_{\psi, \gamma}^* \leq cM\eta^2 \leq \frac{1}{2},
\]
then by Lemma A6 in Appendix of the second paper it follows that $\|W\mathbf{D}\Phi\|_{\psi, \gamma}^* \leq c\|W\|_{\psi, \gamma}^*$.
Again we have
\[
\|\mathbb{D}\Phi\|_{\psi, \gamma}^* \leq 1 + \|\mathbb{D}\Phi - Id\|_{\psi, \gamma}^* + (\|\mathbb{D}\Phi - Id\|_{\psi, \gamma}^*)^2 + \cdots \leq 2.
\]
Hence, $\|\mathbf{D}^* W\|_{\psi, \gamma}^* \leq c\|W\|_{\psi, \gamma}^*$.

Now we estimate $X_{p, \gamma}$. By Lemma 2.3, if $E$ is sufficiently small
\[
\|X_{p, \gamma}^*\|_{\psi, \gamma}^* \leq c\|X_{p, \gamma}^*\|_{\psi, \gamma}^* + \int_0^1 \|\mathbb{D}\Phi\|_{\psi, \gamma}^* \|X_{p, \gamma}^*\|_{\psi, \gamma}^* \, dt.
\]
Since $p > p$, by Cauchy's inequality and Lemma 2.1, we have
\[
\|\mathbb{D}\Phi\|_{\psi, \gamma}^* \|X_{p, \gamma}^*\|_{\psi, \gamma}^* \leq \frac{1}{\eta^3} \|\mathbb{D}\Phi\|_{\psi, \gamma}^* \|X_{p, \gamma}^*\|_{\psi, \gamma}^* + \frac{cM\eta^2}{\eta^3} \leq \frac{cM\eta^2}{\eta^3}.
\]
Combining (2.3) we have
\[
\|X_{p, \gamma}^*\|_{\psi, \gamma}^* \leq c\eta + \frac{cM\eta^2}{\eta^3} \leq cM\eta,
\]
where we choose $\eta^3 = E = \frac{\varepsilon}{\alpha^{N+2}\sigma^{N+1}}$. Let $r_\gamma = \rho, \quad \sigma_\gamma = s - s\rho, \quad \varepsilon = cM\eta^2$. By (2.5) it follows
\[
\frac{1}{\sigma} \|X_{p, \gamma}^* - id\|_{\psi, \gamma}^* \|\mathbb{D}\Phi\|_{\psi, \gamma}^* \|X_{p, \gamma}^*\|_{\psi, \gamma}^* \leq cM\varepsilon.
\]
Moreover,
\[
\|X_{p, \gamma}^*\|_{\psi, \gamma}^* \leq \varepsilon.
\]

So far we have already given a complete process of one circle of KAM iteration. Below we choose a sequence of the pertinent parameters so that we can iterate the KAM step infinitely.

Let $\varepsilon_1 = \varepsilon, \quad r_1 = r, \quad s_1 = s, \quad \sigma_1 = \frac{s_1}{20}, \quad \alpha_1 = \alpha, \quad M(1) = (M_1 + M_2)\sigma^{N+1}$. Thus, $E_1 = \frac{\varepsilon_1}{\alpha_1^{N+2}\sigma_1^{N+1}}$ and
\[
\eta_1 = E_1^{1/3}.
\]
Let $\varepsilon_2 = \frac{cM}{\alpha_1^{N+2}\sigma_1^{N+1}} \varepsilon_1^{4/3}, \quad r_{m+1} = \eta_1 r_m, \quad s_{m+1} = s_m - 5\sigma_m, \quad \sigma_{m+1} = \frac{1}{2} \sigma_m, \quad \alpha_m = \frac{\alpha}{m^{1/3}}, \quad E_m = \frac{\varepsilon_m}{\alpha_m^{N+2}\sigma_m^{N+1}}, \quad \eta_m = E_m^{1/3}, \quad D_m = (D_m, \quad r_m\) and $M(m) = (M_1 + M_2 + 4(\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_m - 1))^{N+1}$. Let $\Pi_m = \left\{ \mathbf{c} \right\}$
\[
\mathbf{c} = \mathbf{A}_m \mathbf{e}_1, \quad \forall k \in \mathbb{Z}^*, \quad \forall l \in \mathbb{Z}^*, \quad \mathbf{e}_1 \neq 0, \quad |k| \neq l \neq 2, \quad |l| \leq 2, \quad \text{where} \quad \omega_1 = \omega.
\]
\( \Omega_{1} = \Omega \), \( \omega_{m} \) and \( \Omega_{m} \) are frequency vector and normal frequency vector at \( m \)th KAM step, which are given later.

By (2.7) we have the map \( \Phi_{m} : D_{m+1} \rightarrow D_{m} \) satisfying

\[
\frac{1}{\sigma_{m}} \| \Phi_{m} - Id \|^*_{r_{m}, \sigma_{m}, D_{m+1}} \| \partial \Phi_{m} - Id \|^*_{r_{m}, \sigma_{m}, D_{m+1}} \leq cM(m)E_{m},
\]

where \( \| \cdot \|^* \) denotes the norm \( \| \cdot \|^*_{\vec{r}, \sigma_{m}, \Pi_{\varepsilon}} \) with \( \Pi_{\varepsilon} = \bigcap_{m \geq 1} \Pi_{\varepsilon} \).

Let \( \Phi^{m} = \Phi_{1} \circ \Phi_{2} \cdots \circ \Phi_{m} (m \geq 1) \) and \( \Phi^{0} = Id \). Thus, \( H_{m} = H \circ \Phi^{m-1} = \tilde{H}_{m} + P_{m} \), where \( \tilde{H}_{m} = \langle \omega_{m}, y \rangle + \langle \Omega_{m} z, \bar{z} \rangle \), \( P_{m} \) is the perturbation of \( m \)th KAM step. By (2.8), \( \| X_{r_{m}} \|^*_{r_{m}, \sigma_{m}, D_{m}} \leq \varepsilon_{m} \).
By (2.4), \( \omega_{m} \) and \( \Omega_{m} \) satisfy

\[
\| \omega_{m+1} - \omega \|^*_{r_{m}, \sigma_{m}, D_{m}} \| \Omega_{m+1} - \Omega \|^*_{r_{m}, \sigma_{m}, D_{m}} \leq 2 \varepsilon_{m},
\]

and so

\[
\| \omega_{m+1} - \omega \|^*_{r_{m}, \sigma_{m}, D_{m}} \| \Omega_{m+1} - \Omega \|^*_{r_{m}, \sigma_{m}, D_{m}} \leq 2(\varepsilon_{1} + \varepsilon_{2} + \cdots + \varepsilon_{m}).
\]

Let \( h_{m} = \frac{\varepsilon_{m}}{(\sigma_{m}^{2} + \sigma_{m+1}^{3})^{3}} \). Thus, \( h_{m} \leq cM(m)h_{m}^{4} \). We may require \( M(m) \leq M = (M_{1} + M_{2} + 4)^{N+1} \). Thus \( c_{m} h_{m} = \frac{\varepsilon_{m}}{(\sigma_{m}^{2} + \sigma_{m+1}^{3})^{3}} \) and \( c_{m} h_{m} \leq (c_{m})^{4/3} \) with \( c = (cM)^{3} \). Let \( h_{1} \leq \frac{1}{2c} \). Since \( \varepsilon_{m} = (\alpha_{m}^{2} + \sigma_{m+1}^{3})^{3} \), here without loss of generality, we suppose \( \alpha = 1 \), \( \sigma_{m} = 1 \), by (2.11) it follows that \( M(m) \leq M \). So \( M(m) \) may be absorbed into constants.

Now we prove that \( \{ \Phi^{m} \} \) is convergent on \( D_{*} \times \Pi_{\varepsilon} \), where \( D_{*} = D(\frac{1}{2} s, \{ 0, 0, 0 \}), 
\]

\( D \left( \frac{1}{2} s \right) = \left\{ x: \| \text{Im} x \| < \frac{1}{2} s \right\} \).

By the constructing of \( \Phi_{m} \), \( \Phi_{m} \) maps \( D_{m+1} \) into \( D(s_{m} - 4\sigma_{m}, 2\eta_{m} r_{m}) \subseteq D(s_{m} - 2\sigma_{m}, \frac{1}{2} r_{m}) \).

Since in the weight norm \( \| \cdot \|_{r_{m}} \) the distance from \( D(s_{m} - 4\sigma_{m}, 2\eta_{m} r_{m}) \) to the boundary of \( D \left( s_{m} - 2\sigma_{m}, \frac{1}{2} r_{m} \right) \) is more than \( \sigma_{m} \), by Lemma A6 of the second paper, if \( h_{1} \) is sufficiently small (this holds if \( \varepsilon \) sufficiently small) such that \( cME \) is small enough for \( \forall m \geq 1 \), then we have \( \| \Phi_{m-1} \circ \Phi_{m} - Id \|^*_{r_{m-1}, \sigma_{m}, D_{m+1}} \leq \| \Phi_{m-1} - Id \|^*_{r_{m-1}, \sigma_{m}, D_{m}} \).
Inductively it easily follows that for any \( j \geq 0 \) and \( m \geq 1 \),

\[
\| \Phi_{m-1} \circ \Phi_{m} \circ \Phi_{m+1} \cdots \circ \Phi_{m+j} - Id \|^*_{r_{m-1}, \sigma_{m}, D_{m+j+1}} \leq \| \Phi_{m} - Id \|^*_{r_{m-1}, \sigma_{m}, D_{m}}.
\]

Since \( \Phi_{m+1} = \Phi_{m} \circ \Phi_{m+1} \), by the mean theorem of differentiation we have

\[
\| \Phi_{m+1} - \Phi_{m} \|^*_{r_{m}, \sigma_{m}, D_{m+2}} \leq \| \partial \Phi_{m} \|^*_{r_{m}, \sigma_{m}, D_{m+1}} \| \Phi_{m+1} - Id \|^*_{r_{m}, \sigma_{m}, D_{m+2}}.
\]
By definition we easily know that for \( r \geq s \), \( ||AB||_{r,s} \leq ||A||_{r} \cdot ||B||_{s} \). So by (2.12) and (2.9) we have

\[
||D\Phi^{*}\|_{r, \frac{1}{2}, p_{m}, \frac{1}{2}, \frac{1}{2}, \ldots} \leq \prod_{j=1}^{m} (1 + cE_{j}) \leq \prod_{j=1}^{m} (1 + c\delta_{j}) < + \infty.
\]

Since \( D = \bigcap_{m \geq 1} D_{m} \), by (2.9) it follows that \( \{\Phi^{*}\} \) is convergent on \( D_{*} \times \Pi_{*} \).

Completely similar to ref. [3] we can prove that \( \{D\Phi^{*}\} \) and \( \{D^{*}\Phi^{*}\} \) are also convergent on \( D_{*} \times \Pi_{*} \). By noticing that \( \Phi^{*} \) consists of at most one-order terms of \( y \) and at most two-order terms of \( z, \bar{z} \), \( \{\Phi^{*}\} \) is actually convergent on \( D \left( \frac{1}{2}, \frac{1}{2}, s \right) \times \Pi_{*} \).

Let \( \lim_{m \to +\infty} \Phi^{*} = \Phi_{*} \) and \( H_{*} = \bar{H}_{*} + P_{*} \), where \( \bar{H}_{*} = \langle \omega_{0}, y \rangle + \langle \Omega_{z}, \bar{z} \rangle \). Since \( \omega_{*} = \lim_{m \to +\infty} \omega_{m} \), by (2.11) it follows that \( |\omega - \omega_{*}| \leq c \varepsilon \). Since \( ||X_{P_{*}}||_{m, \frac{1}{2}, \frac{1}{2}} \leq \varepsilon_{m} \) and \( \lim_{m \to +\infty} ||X_{P_{*}}||_{m, \frac{1}{2}, \frac{1}{2}} = 0 \), it follows that \( X_{P_{*}} = 0 \) on \( D_{*} \times \Pi_{*} \). In fact, \( \frac{\partial^{m+1}}{\partial y^{m} \partial z^{2} \partial z^{2}} |_{P_{*}} = 0 \) for \( 2|m| + |q + q| \leq 2 \). By (2.12) it follows that \( ||\Phi_{*} - id|| \leq c \varepsilon \).

Now we estimate the measure of \( \Pi_{*} \) to prove that it is nonempty. By Theorem B,

\[
\text{mes}(\Pi_{*} - \Pi_{m}) \leq c (\text{diam}\Pi)^{m-1} \frac{\alpha^{m}}{m^{2}}.
\]

So \( \text{mes}(\Pi_{*} - \Pi_{m}) \leq \sum_{m=1}^{\infty} \text{mes}(\Pi_{*} - \Pi_{m}) \leq c (\text{diam}\Pi)^{m-1} \alpha^{m} \). This implies that if \( \alpha \) is sufficiently small, \( \Pi_{*} \) is nonempty.

References