JOURNAL OF DIFFERENTIAL EQUATIONS **152**, 1–29 (1999)
Article ID jdeq.1998.3515, available online at http://www.idealibrary.com on IDEAL®

Perturbations of Lower Dimensional Tori for Hamiltonian Systems*

Jiangong You

Department of Mathematics, Nanjing University, Nanjing 210093, People's Republic of China
E-mail: jyou@nju.edu.cn

Received June 20, 1997; revised April 16, 1998

This paper provides a generalized Kolmogorov–Arnold–Moser theorem for lower dimensional tori in Hamiltonian systems, which applies to multiple normal frequency case. The proof is based on Newton's iteration method and a generalized version of small divisor conditions. © 1999 Academic Press

1. INTRODUCTION AND RESULT

We start with a real analytic perturbation

$$H(x, y, u) = \mathscr{H}(y, u) + \mathscr{P}(x, y, u) \tag{1.1}$$

of a real analytic Hamiltonian $\mathscr{H}(y, u)$ in a complex neighborhood in $C^{2n} \times C^{2m}$, with the symplectic structure $\sum_{i=1}^{n} dx_i \wedge dy_i + \sum_{i=1}^{m} du_i \wedge du_{-i}$, of a $2n$ real domain $u = 0$, $y \in D \subset R^n$ where $u = (u_1, u_{-1}, ..., u_m, u_{-m}) \in R^{2m}$, $(x, y) \in T^n \times R^n$ and D is an open set of R^n.

Assume that

$$\mathscr{H}_u(y, 0) = 0, \qquad \det \mathscr{H}_{uu}(y, 0) \neq 0,$$

and $\mathscr{P}$ is small. The unperturbed Hamiltonian system defined by $\mathscr{H}$ possesses an invariant subspace $u = 0$ foliated by a family of invariant tori $y = y_0$, $u = 0$ and the flow on each torus is given by $x(t) = x_0 + \mathscr{H}_y(y_0, 0)\, t$.

Expanding the Hamiltonian (1.1) in the neighborhood of $u = 0$. The Hamiltonian (1.1) is reduced to

$$H(x, y, u) = \mathscr{H}(y, 0) + \tfrac{1}{2} \langle \mathscr{H}_{uu}(y, 0)\, u, u \rangle + P(x, y, u) + O(u^3).$$

* This work was partially Supported by Alexander von Humboldt Foundation in Germany and a grant of the National Natural Science Foundation of China.

0022-0396/99 $30.00
Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

For the unperturbed system, the local normal behavior of the invariant torus $y=y_0, u=0$ is determined by the matrix $\mathscr{H}_{uu}(y_0,0)$. The reader is referred to [6] for a detailed motivation.

Further we suppose that $\mathscr{H}(y,0)$ is non-degenerate, i.e., $\mathscr{H}_{yy}(y,0)$ is nonsingular. Linearizing the Hamiltonian at the neighborhood of torus $T^n \times \{y=y_0\} \times \{u=0\}$, we arrive at a family of perturbed integrable Hamiltonians,

$$H=N+P=\langle \omega, y\rangle + \tfrac{1}{2}\langle A(\omega)\, u, u\rangle + P, \tag{1.2}$$

where $(x, y, u) \in T^n \times R^n \times R^{2m}$, $A(\omega)=\mathscr{H}_{uu}(\mathscr{H}_y^{-1}(\omega), 0)$ is a $2m\times 2m$ symmetric matrix and $P=\mathscr{P}+O_2(y)+O(yu)+O_3(u)$. We will treat $\omega = H_y(y_0, 0)$ as an independent parameter. This setting has been frequently used by many authors. In this paper, we state our results for (1.2) where ω is an independent parameters varying over a positive measure set $\mathcal{O}$.

If all eigenvalues of JA are not on the imaginary axis (J being the standard symplectic matrix in R^{2m}), the torus is called hyperbolic. In this case, for any given $\omega=(\omega_1, ..., \omega_n) \in \mathcal{O}$ satisfying the Diophantine condition

$$|\langle k, \omega\rangle| > \gamma\, |k|^{-\tau}, \qquad \text{for} \quad \tau > n-1, \quad 0 \neq k \in Z^n, \tag{1.3}$$

Moser [12], Graff [6], and Zehnder [24] proved that there is a ω^* close to ω such that (1.2) at ω^* possesses an invariant torus with prescribed frequencies ω if perturbation is sufficiently small whether the eigenvalues of JA are different or not. From this, it follows that (1.1) has a Cantor family of invariant tori if the perturbation is small.

If all the eigenvalues of $JA(\omega)$ belong to $\mathrm{i}R^1\backslash 0$, the torus is called elliptic. So far $\mathscr{H}_{uu}(y,0)$ has different eigenvalues has been considered extensively. More precisely, for a system with the following Hamiltonian

$$H=N+P=\sum_{j=1}^{n} \omega_j\, y_j + \tfrac{1}{2} \sum_{j=1}^{m} \Omega_j(\omega)(u_j^2+u_{-j}^2)+P, \tag{1.4}$$

Melnikov [11] in 1967 announced that for a positive Lebesgue measure subset $\mathcal{O}_\gamma \subset \mathcal{O}$, $(1.4)_{\omega\in\mathcal{O}_\gamma}$ possesses a n dimensional invariant torus with frequencies satisfying the non-resonant conditions

$$|\langle k, \tilde{\omega}\rangle + \langle l, \tilde{\Omega}\rangle| > \tfrac{1}{2}\gamma\, |k|^{-\tau}, \qquad |l| \leqslant 2 \tag{1.5}$$

for $k\in Z^n$, $l \in Z^m$, $|k|+|l| \neq 0$, $\tau > n-1$ where $(\tilde{\omega}, \tilde{\Omega})=(\tilde{\omega}_1, ..., \tilde{\omega}_n, \tilde{\Omega}_1, ..., \tilde{\Omega}_m)$ is close to (ω, Ω), provided the perturbation is sufficiently small ($\tilde{\omega}$ are usually referred as 'inner frequencies' while $\tilde{\Omega}$ as 'normal frequencies'). The complete proof was carried out fifteen years later by Eliasson, Kuksin, and Pöschel [7, 8, 14]. In this case, only the measure estimate is available. One cannot say if (1.4) has a torus with prescribed frequencies.

The proof of the previous mentioned works for (1.4) heavily depends on the normal form $\sum_{j=1}^{n} \omega_j y_j + \frac{1}{2}\sum_{j=1}^{m} \Omega_j(\omega)(u_j^2 + u_{-j}^2)$. However, $JA(\omega) = J\mathcal{H}_{uu}(y_0, 0)$ might have multi-eigenvalues and the above normal form may not be available.

More recently, developing Craig and Wayne's method [5], Bourgain [2] proved the existence of quasi-periodic solutions for Hamiltonian systems with Hamiltonian (1.4) provided that (1.5) holds for $|l| \leqslant 1$. Such approach applies to some PDEs with periodic boundary condition [1]. His proof is based on the Liapounov–Schmidt reduction introduced by Craig and Wayne [5] and some sophisticated estimates needed to control the inverse of matrices with singular sites.

In this paper, we start from the real analytical Hamiltonian systems (1.2) with a general non-singular symmetric matrix A. The goal of this paper is to prove, under a mild non-degenerate condition which applies to multiple normal frequency case, that there is a nonempty subset $\mathcal{O}_\gamma \subset \mathcal{O}$, such that (1.2) at $\omega \in \mathcal{O}_\gamma$ is equivalent to

$$N_\infty = e_\infty + \langle \omega_\infty(\omega), y \rangle + \langle A_\infty(\omega) u, u \rangle + O(y^2 + |u|^3 + |yu|), \quad (1.6)$$

and thus (1.2) at $\omega \in \mathcal{O}_\gamma$ possesses a n dimensional invariant torus provided that the perturbation P is analytic and small enough. Although Bourgain [2] also started from the Hamiltonian (1.4), but the normal form in (1.6) seems not essential for his approach. Thus the persistence result for our case might have been known essentially (see [2]). The advantage of our approach by KAM theory is that it provides not only the persistence but also a local normal form in the neighborhood of the obtained torus, which might be helpful for a better understanding of the dynamics. For example, in case A in the unperturbed system is definite positive, the obtained torus is linearly stable, since A_∞, a small perturbation of A, is also definite positive.

The main idea of this paper is the following: we do each step of KAM iteration with unperturbed Hamiltonians $N_\nu = \langle \omega, y \rangle + \langle A_\nu u, u \rangle$ where matrix $A_\nu = (a_{ij}^\nu(\omega))$ is not diagonal. Certainly JA might be normalized into a Jordan normal form by an algebraic argument (see, e.g., [3]), but the normalized operation depends singularly on parameters which will cause troubles for measure estimates. In order to make the KAM machinery work for this situation, we have to modify the classical small denominator condition (1.5) by a more general condition (spelled in (2.3) below). In the next section, we will explain the ideas in more details.

Remark. For the sake of simplicity, in the following $|\cdot|$ denotes the absolute value for complex numbers, Euclidean norm for vectors, the determinant for matrices and Lebesgue measure for sets. We also denote by $|\cdot|_a$ the absolute value of the determinant for matrices.

Let $\pm \mathrm{i}\Omega_i$, $i=1, ..., m$ be eigenvalues of JA. We say that the unperturbed Hamiltonian is *non-degenerate* if the following conditions are satisfied

$$|\{\omega: |\langle k, \omega\rangle + \langle l, \Omega(\omega)\rangle| = 0\}| = 0 \tag{1.7}$$

for any $|k| \neq 0$, $|l| \leqslant 2$ where $l \in Z^m$, $\Omega = (\Omega_1, ..., \Omega_m)$.

This kind of non-degeneracy conditions has been proposed by Pöschel [15] for the special case (1.4). We stress that here the normal form in (1.4) is not required, also (1.7) for $k=0$ does not need to be satisfied as it is instead required in previous version of KAM theorems (see for example [15]). The latter observation might be important since it allows to apply the results to the multiple normal frequency case.

It is known that the eigenvalues might depend non-smoothly on the parameter ω even JA is analytic in ω. For the convenience of the measure estimates, we will use the following conditions which is *equivalent* to (1.7)

$$\left| \begin{array}{l} |\langle k, \omega\rangle|, |\mathrm{i}\langle k, \omega\rangle I_{2m} - JA(\omega)| = 0, \text{ or} \\ \{\omega: ||\mathrm{i}\langle k, \omega\rangle I_{4m^2} - I_{2m} \otimes (JA(\omega)) - (JA(\omega)) \otimes I_{2m}| = 0\} \end{array} \right| = 0, \quad \text{for } k \neq 0, \tag{1.8}$$

where $J = \mathrm{Diag}(J_2, ..., J_2)$, with $J_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is the standard symplectic structure matrix in R^{2m}, I_N is $N \times N$ unit matrix, $\otimes$ denotes the tensor (or direct) product of matrices (see Appendix for explicit definition). The equivalence of (1.7) and (1.8) will be proven in the Appendix.

Remark. (1.8) is satisfied if some derivatives (w.r.t. ω) of $|\mathrm{i}\langle k, \omega\rangle I - JA|$ and $|\mathrm{i}\langle k, \omega\rangle - I_{2m} \otimes (JA(\omega) - (JA(\omega)) \otimes I_{2m}|$ are away from 0 (see [21] or [23]). For example if A is a constant matrix, the $2m$ order derivatives of $|\mathrm{i}\langle k, \omega\rangle I - JA|$ and the $4m^2$ order derivatives of $|\mathrm{i}\langle k, \omega\rangle I_{4m^2} - I_{2m} \otimes (JA(\omega)) - (JA(\omega)) \otimes I_{2m}|$ are bigger than 1 for $k \neq 0$, and thus (1.8) is satisfied. In case the Hamiltonian does not depend analytically on ω, (1.8) has to be replaced by

$$|D_\omega^{N_1} |\mathrm{i}\langle k, \omega\rangle I_{2m} - JA(\omega)|| \geqslant C > 0$$

$$|D_\omega^{N_2} |\mathrm{i}\langle k, \omega\rangle I_{4m^2} - I_{2m} \otimes (JA(\omega)) - (JA(\omega)) \otimes I_{2m}|| \geqslant C > 0,$$

for some integers N_1, N_2.

From now on, we consider the complex extension of the real Hamiltonian (1.2) on the complex neighborhood $D(r, s) \times \Pi_h$ endowed with the standard symplectic structure $\sum_{i=1}^n dx_i \wedge dy_i + \sum_{i=1}^m du_i \wedge du_{-i}$, where

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

Imx is the image part of x and $|\cdot|$ is the *sup*-norm in C^n, and Π_h is a complex neighborhood of $\mathcal{O}$:

$$\Pi_h = \{\omega \in C^n : |\omega - \mathcal{O}| < h\}. \tag{1.9}$$

The norm of P on $D(r, s) \times \Pi_h$ is defined to be

$$\|P\|_{D \times \Pi_h} = \sup_{D \times \Pi_h} |P|.$$

To state our result, we assume that

$$\max_{|l| \leqslant 8m^2} \left| \frac{\partial^l a_{ij}}{\partial \omega^l} \right| \leqslant L$$

on $\Pi_h(K)$.

The following theorem is the main result of this paper:

THEOREM 1. *Suppose that the real analytic Hamiltonian* (1.2) *satisfies the non-degenerate condition* (1.7). *Then for a given γ there is a small constant ε_0 depending on n, τ, r, L, γ, h, such that if the complex extension of P in $D(r, s)$ satisfies*

$$\frac{1}{s^2} \|P\|_{D(r, s) \times \Pi_h} = \varepsilon \leqslant \varepsilon_0,$$

we have the following conclusions: there exists a non-empty Cantor set $\mathcal{O}_\gamma \subset \mathcal{O}$, a Whitney-smooth map

$$\Phi : T^n \times \mathcal{O}_\gamma \to T^n \times R^n \times R^{2m},$$

and a diffeormophism $\tilde{\omega}: \mathcal{O}_\gamma \to R^n$, such that the map Φ restricted to $T^n \times \{\omega\}$ is a real analytic invariant torus with frequencies $\tilde{\omega}$ for the Hamiltonian H at ω. Moreover

$$|\mathcal{O} - \mathcal{O}_\gamma| \to 0, \qquad \textit{as} \quad \gamma \to 0.$$

For convenience, Theorem 1 will be split into two theorems (Theorem 2 and Theorem 3) in Section 4. The parameter γ in Theorem 1 will play the role of the Diophantine constant in the KAM iteration. If one knows more information for the unperturbed system, the estimate $|\mathcal{O} - \mathcal{O}_\gamma| \to 0$, as $\gamma \to 0$ can be made quantitative as $|\mathcal{O} - \mathcal{O}_\gamma| = O(\gamma^{1/4m^2})$. The reader is referred to Theorem 3 for details.

Remark. Analyticity in space variables is not necessary but it simplifies considerably the proof. Analyticity in parameters is also not necessary if one use Whitney's extension theorem [19]. However we do need H

depending C^{8m^2}-smoothly on parameter ω in Whitney's sense (compare with [14] where only Lipschitz continuous is required). The reason is that the eigenvalues of JA might be not Lipschitz continuous. We have to find an alternative way to estimate the measure of 'bad' frequencies at each KAM step. In order to get the desired measure estimate, we require our Hamiltonian is sufficiently smooth in parameters. We do not know whether the restriction is essential or not.

Remark. A bit more general result can be proven, without any essential difficulties, in the case $\omega(\xi)$ depends on a parameter ξ provided the non-degeneracy conditions (1.8) hold. The case $\omega(\xi)$ is Lipomorphism was considered in [15]. Non-Lipomorphism case was considered recently by many authors [4, 13, 17, 20–22]. The key point is to find a geometric restriction for the initial frequencies $\omega(\xi)$ so that the intersection of $\mathcal{O}_\gamma$ and the image set of $\omega(\xi)$ has positive measure with respect to the induced measure in the image set of $\omega(\xi)$.

Remark. This version of KAM theorem will be generalized to infinite dimensional case in a forthcoming paper [4a] for constructing *linearly stable* quasi-periodic solutions of partial differential equations. The existence result has been obtained by Craig and Wayne [5] (periodic solutions), Bourgain [3] (also quasi-periodic solutions). By KAM theory, we can also provide a normal form for the obtained solutions.

2. KAM STEP

In this section, we will first outline the main ideas beyond the proof of Theorem 1 and then we will give details about one step of the KAM iteration.

2.1. *Main Ideas and an Outline of KAM Step*

The proofs are based on KAM theory which involves an infinite sequence of change of variables. The philosophy of KAM theory is to construct a series of coordinate transformation which makes the perturbation smaller and smaller at the cost of excluding a small set of parameters. Finally convergence is proved and the existence of invariant tori is established. In previous KAM theorems, the perturbed system is reduced to the very nice form

$$N = \langle \omega, y \rangle + \sum_{i=1}^{m} \Omega_i (u_i^2 + u_{-i}^2), \tag{2.1}$$

with different Ω_i, which surely has an invariant torus but also surely requires more restrictions on the unperturbed Hamiltonian.

Observe that to get persistence, requiring the above normal form is not necessary. For example, it would be enough to reduce the perturbed Hamiltonian into the form

$$N = \langle \omega, y \rangle + \langle Au, u \rangle.$$

As a return, in principle, the unperturbed systems can be more general. The cost is the KAM machinery will be a little bit more complicated than the classical one since the non-resonant relations are now more general as we do not require (1.7) to be satisfied for $k = 0$. We also remark that A could even be x-dependent, but this leads to some difficult techniques due to Bourgain to control the inverse of a linear operator.

Based on the above understanding and motivated by problems in Hamiltonian dynamical systems and in partial differential equations recently considered by Bourgain, we provide a KAM theorem for the case where A is a general nonsingular constant matrix. In some sense, our theorem is in the middle of the previous KAM theorems and Bourgain's approach. We require more informations for the eigenvalues of A than Bourgain, but less than those used in the previous KAM theorems. Obviously, the information we can provide for the persistent tori is more detailed than that obtained by Bourgain but less than that obtained in previous KAM theorems. An outcome of our approach is that we do provide a normal form for the obtained torus while this is not so with the Bourgain's approach. Note that our approach also covers the multiple eigenvalue case, which makes it possible to generalize it to infinite dimensional space and provide a KAM theorem for partial differential equations with periodic boundary condition considered by Bourgain for proving the existence of linear stable quasi-periodic solutions. This has been done in a forthcoming paper [4a].

In order to proceed the proofs with an general matrix A, we will set up one step of KAM iteration as follows: at each KAM step, we only eliminate the x-dependent term in a suitable truncation R of P (defined by (2.11)), which allows us, in particular, not impose any small denominator condition for Fourier modes with $k = 0$ (This is a very important point since it allows to treat multiple eigenvalue cases). The x-independent part is of the form

$$\langle P_{0l0}, y \rangle, \quad \langle P_{002} u, u \rangle, \quad \langle P_{001}, u \rangle,$$

where $P_{0l0}, P_{002}, P_{001}$ are the coefficients of the expansion of P. Note that the first two terms can be put into the new normal form part, while the third term can be eliminated by a linear change of coordinates since A is non-singular.

Now the problem is how to eliminate the x-dependent terms in R assuming A to be a general non-singular matrix. To do this, we have to solve a matrix equation of the form (see Lemma 2.4 for the exact form of the equations)

$$AX + XB = C,$$

which involves small denominators. By a classical result in the matrix theory such an equation is solvable if and only if $I \otimes A^T + B \otimes I$ is non-singular, where A^T denotes the transpose of A (See the Appendix or [10], p. 256). To control the norm of the solution, we need the determinant of $I \otimes A^T + + B \otimes I$ not to be too small: this leads to our small denominator condition (2.3).

In the following we describe one step of KAM iteration in more details. As we will see, at each step of the KAM scheme, a Hamiltonian

$$H_\nu = N_\nu + P_\nu$$

defined in $D(r_\nu, s_\nu) \times \Pi_\nu$ is considered near a n dimensional torus $\{y = 0, u = 0\}$, where

$$\frac{1}{s_\nu^2} \|P_\nu\|_{D(r_\nu, s_\nu) \times \Pi_\nu} \leqslant \varepsilon_\nu,$$

$$N_\nu = e_\nu + \langle \omega_\nu, y \rangle + \frac{1}{2} \langle A_\nu u, u \rangle = e_\nu + \sum_{i=1}^{m} \omega_i^\nu y_i + \frac{1}{2} \sum_{i, j = \pm 1}^{\pm m} a_{ij}^\nu u_i u_j.$$

In what follows, the Hamiltonian without subscripts denotes the Hamiltonian in νth step, while those with subscripts $+$ denotes the Hamiltonian of $\nu + 1$th step. Thus we consider the Hamiltonian

$$\begin{aligned} H = N + P &= e + \langle \omega, y \rangle + \tfrac{1}{2} \langle Au, u \rangle + P \\ &= e + \sum_{i=1}^{n} \omega_i y_i + \tfrac{1}{2} \sum_{i, j = \pm 1}^{\pm m} a_{ij} u_i u_j + P \end{aligned} \tag{2.2}$$

defined in $D(\rho, s) \times \Pi_h(K)$ where $\Pi_h(K)$ is the complex h-neighborhood of $\mathcal{O}_\gamma(K)$. $\mathcal{O}_\gamma(K)$ is the set of $\omega \in \mathcal{O}$ satisfying the following small denominator condition

$$\begin{gathered} |\langle k, \omega \rangle| > \frac{\gamma}{|k|^\tau}, \qquad |\Delta I_{2m} - JA|_a > \frac{\gamma}{|k|^\tau} \\ |\Delta I_{4m^2} - I_{2m} \otimes (JA(\omega)) - (JA(\omega)) \otimes I_{2m}|_a > \frac{\gamma}{|k|^\tau} \end{gathered} \tag{2.3}$$

for $\tau > 4m^2(n-1)$, $0 \neq |k| \leqslant K$ where $\Delta = \Delta(k, \omega) = \mathrm{i}\langle k, \omega\rangle$. We recall that $|\cdot|_a$ denotes the absolute value of the determinant for matrices. We assume that a_{ij} has the following bound

$$\max_{|l| \leqslant 8m^2} \left|\frac{\partial^l a_{ij}}{\partial \omega^l}\right| \leqslant L, \tag{2.4}$$

on $\Pi_h(K)$, and that

$$\frac{1}{s^2} \|P\|_{D \times \Pi_h(K)} \leqslant \varepsilon. \tag{2.5}$$

Henceforth, we denote by

$$\begin{aligned} &\alpha = \varepsilon^{1/3}, \qquad h = \varepsilon^{1/(8m^2+1)}, \qquad s_+ = \alpha s, \qquad \varepsilon_+ = \gamma^{-2}\Gamma(\tfrac{1}{2}(\rho - \rho_+))\varepsilon^{4/3} \\ &K = \min\left\{K \in Z \colon \int_K^\infty x^n e^{-x\rho}\, dx \leqslant \varepsilon\right\}, \end{aligned} \tag{2.6}$$

where $\rho > \rho_+$ and for $t > 0$,

$$\Gamma(t) = \sum_{k \neq 0} |k|^{4m^2(2m-1)^2 + \tau} e^{-|k|t}.$$

The purpose of this section is to find a change of variables defined in a smaller domain $D_+ \times \Pi_+$, such that the transformed Hamiltonian $H_+ = N_+ + P_+$ has the same form as H and satisfies all the above iterative assumptions with $s_+, \varepsilon_+, \rho_+, \gamma_+, L_+$.

For this purpose, we first rewrite H as

$$H = N + R + (P - R),$$

where R is a suitable truncation of P (see (2.11) for the explicit definition). $P - R$ can be made smaller by shrinking the domain.

Then we will find a special F, defined in a smaller domain $D(\rho_+, \alpha s)$, such that the time 1 map ϕ_F^1 of the Hamiltonian vector field X_F transforms H into a new normal form with a smaller perturbation P_+.

More precisely, by Taylor series expansion, we have

$$\begin{aligned} H \circ \phi_F^1 &= (N + R) \circ \phi_F^1 + (P - R) \circ \phi_F^1 \\ &= N + \{N, F\} + R \\ &\quad + \int_0^1 (1 - t)\{\{N + R, F\}, F\} \circ \phi_F^t\, dt + \{R, F\} + (P - R) \circ \phi_F^1 \\ &= N_+ + \{N, F\} + R - \frac{1}{2\pi}\int_0^{2\pi} R\, dx + \sum_{i=\pm 1}^{\pm m} R_{00i} u_i + P_+, \end{aligned} \tag{2.7}$$

where

$$N_{+}=e+\langle\omega, y\rangle+\tfrac{1}{2}\langle Au, u\rangle+\int_0^{2\pi} R\,dx-\sum_{i=\pm 1}^{\pm m} R_{00i}u_i,$$

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

and R_{00i} are defined after (2.11). $\{\cdot,\cdot\}$ is Poisson bracket of smooth functions

$$\begin{aligned}\{G_1, G_2\}&=\sum\left(\frac{\partial G_1}{\partial x_i}\frac{\partial G_2}{\partial y_i}-\frac{\partial G_1}{\partial y_i}\frac{\partial G_2}{\partial x_i}\right)+\sum_{i=\pm 1}^{\pm m}\operatorname{Sgn}(i)\frac{\partial G_1}{\partial u_i}\frac{\partial G_2}{\partial u_{-i}}\\&=\sum\left(\frac{\partial G_1}{\partial x_i}\frac{\partial G_2}{\partial y_i}-\frac{\partial G_1}{\partial y_i}\frac{\partial G_2}{\partial x_i}\right)+\langle\nabla_u G_1, J\nabla_u G_2\rangle. \qquad (2.8)\end{aligned}$$

What we will do is to find a F, which solves all x-dependent terms in R as well as the u-linear terms in $(1/2\pi)\int_0^{2\pi} R\,dx$. More precisely, we will solve

$$\{N, F\}+R-\frac{1}{2\pi}\int_0^{2\pi} R\,dx+\sum_{i=\pm 1}^{\pm m} R_{00i}u_i=0, \qquad (2.9)$$

in case that N is defined with a general non-singular matrix A in the following Subsection 2.4. Certainly, we have to prove the new perturbation P_+ is much smaller than P.

2.2. *Truncating Perturbations*

Expanding P into the Fourier–Taylor series

$$P=\sum_{k, l, p} P_{klp}e^{\mathrm{i}(k, x)}y^l u^p,$$

where the sum is taken over

$$k=(k_1, ..., k_n)\in Z^n, \qquad l=(l_1, ..., l_n)\in Z_+^n, \qquad p\in Z_+^{2m}.$$

Let R be the truncation of P with $2\,|l|+|p|\leqslant 2$, $|k|\leqslant K$, i.e.,

$$\begin{aligned}R(x, y, u)=P_0+P_1+P_2=&\sum_{|k|<K,\,|l|\leqslant 1} P_{kl0}e^{\mathrm{i}(k, x)}y^l\\&+\sum_{|k|<K,\,|p|=1} P_{k0p}e^{\mathrm{i}(k, x)}u^p+\sum_{|k|<K,\,|p|=2} P_{k0p}e^{\mathrm{i}(k, x)}u^p. \qquad (2.10)\end{aligned}$$

Rewrite R as

$$R(x, y, u) = P_0 + P_1 + P_2 = \sum_{|l| \leqslant 1,\ |k| \leqslant K} P_{kl0} e^{i(k,x)} y^l + \sum_{|k| \leqslant K} \langle R_{ku}, u \rangle e^{i(k,x)} + \sum_{|k| \leqslant K} \langle R_{kuu} u, u \rangle e^{i(k,x)}, \tag{2.11}$$

where R_{ku} denotes the $2m$ vector (R_{k0i}) with $R_{k0i} = P_{k0p}$ for $|p| = 1$ and $u^p = u_i$, R_{kuu} denotes the $2m \times 2m$ symmetric matrix (R_{k0ij}) with $R_{k0ij} = \frac{1}{2}(1 + \delta_i^j)\, P_{k0p}$ for $|p| = 2$ and $u^p = u_i u_j$.

By Cauchy estimate,

$$|R_{k0i}| \leqslant \varepsilon s e^{-|k|\rho}, \qquad |R_{k0ij}| \leqslant \varepsilon e^{-|k|\rho}, \tag{2.12}$$

for $i, j = 1, ..., 2m$.

Remark. For simplicity, in the following, '$\lessdot$' stands for '$< c$' with a constant c independent of iteration steps, and '$a \ll b$' means 'there is a constant c independent of iteration steps such that $a < cb$'.

The following estimates come from Cauchy estimates.

LEMMA 2.1.

$$|R|_{D(\rho_+, (1/2)s) \times \Pi_h} \lessdot |P|_{D \times \Pi_h} = \varepsilon s^2. \tag{2.13}$$

Moreover, in a smaller domain $D(\rho_+, \alpha s)$, *we have*

$$|P - R| \lessdot \varepsilon^2 s^2 \tag{2.14}$$

if K *is the minimum integer satisfying* $\int_K^\infty x^n e^{-(\rho - \rho_+)x}\, dx \leqslant \varepsilon$.

Proof. (2.16) directly follows from Cauchy inequality. Now we prove (2.14). Note that

$$P - R = \sum_{|k| > K} P_{klp} e^{i(k,x)} y^l u^p + \sum_{|k| \leqslant K,\ 2|l| + |p| \geqslant 3} P_{klp} e^{i(k,x)} y^l u^p.$$

(2.14) follows from the following two estimates:

$$\left| \sum_{|k| \geqslant K} P_{klp} e^{i(k,x)} y^l u^p \right| \leqslant \sum_{|k| \geqslant K} |P|\, e^{-|k|\rho} e^{|k|\rho_+} \leqslant \varepsilon s^2 \sum_{\kappa \geqslant K} \kappa^n e^{-\kappa(\rho - \rho_+)}$$

$$\leqslant \varepsilon s^2 \int_K^\infty x^n e^{-(\rho - \rho_+)x}\, dx \leqslant \varepsilon^2 s^2 \tag{2.15}$$

since $\int_K^\infty x^n e^{-(\rho-\rho_+)x}\,dx \leqslant \varepsilon$, and

$$\begin{aligned}\left|\sum_{|k|\leqslant K,\, 2|l|+|p|\geqslant 3} P_{klp} e^{i(k,x)} y^l u^p\right| &= \left|\int \sum_{2|l|+|p|=3} \frac{\partial^{|l|+|p|}}{\partial y^l\, \partial u^p} P\right| \lessdot \frac{\alpha^3 s^3}{s^3}\,|P| \\ &\lessdot \alpha^3 \varepsilon s^2 = \varepsilon^2 s^2.\end{aligned} \tag{2.16}$$

in $D(\rho_+, \alpha s)$, where $\int$ stands for $\underbrace{\int_0^y \cdots \int_0^y}_{l} \underbrace{\int_0^u \cdots \int_0^u}_{p}$ with $2|l|+|p|=3$. ∎

2.3. *Inverse of Matrices*

For a $N\times N$ matrix $M=(m_{ij})$, we denote by $|M|$ its determinant. Consider M as a linear operator on $(R^N, |\cdot|)$ where $|x| = \max |x_i|$. Let $\|M\|$ be its operator norm. It is known $\|M\|$ is equivalent to norm $N \max |m_{ij}|$. Since a constant depends only on the space dimension and two fixed norms is irrelevant, we will simply denote $\|M\| = N \max |m_{ij}|$.

Now we consider matrices $\Delta I_{2m} - JA$ and $\Delta I_{4m^2} - I_{2m} \otimes (JA(\omega)) - (JA(\omega)) \otimes I_{2m}$ introduced in the first section.

LEMMA 2.2. *For $\omega \in \mathcal{O}_\gamma(K)$ defined in* (2.3), *we have*

$$\begin{aligned}\|(\Delta I_{2m} - JA(\omega))^{-1}\| &\lessdot \frac{|k|^{4m^2+\tau}}{\gamma} \\ \|(\Delta I_{4m^2} - I_{2m} \otimes (JA(\omega)) - (JA(\omega)) \otimes I_{2m})^{-1}\| &\lessdot \frac{|k|^{4m^2+\tau}}{\gamma}\end{aligned} \tag{2.17}$$

where $\mathcal{O}_\gamma$ is the set of ω satisfying (2.3).

Proof. The inequalities follows from the assumption (2.3), the definition of $\|M\|$ and the formula

$$M^{-1} = \frac{1}{|M|} \operatorname{adj} M$$

for any nonsingular matrix M where adj M is the adjoint of M. ∎

LEMMA 2.3. *Let Π_h be a h-neighborhood of $\mathcal{O}_\gamma(K)$ in C^n. Then for any $\omega \in \Pi_h(K)$, we have*

$$\begin{aligned}|\Delta^{-1}| \lessdot \frac{|k|^{4m^2+\tau}}{\gamma}, \qquad \|(\Delta I_{2m} - JA)^{-1}\| &\lessdot \frac{|k|^{4m^2+\tau}}{\gamma} \\ \|(\Delta I_{4m^2} - I_{2m} \otimes (JA(\omega)) - (JA(\omega)) \otimes I_{2m})^{-1}\| &\lessdot \frac{|k|^{4m^2+\tau}}{\gamma}\end{aligned} \tag{2.18}$$

for $0 \neq |k| \leqslant K$, *if* $h\,|k|^{4m^2+\tau} \max\{K, L+1\} = \varepsilon^{1/(4m^2+1)}\,|k|^{4m^2+\tau} \max\{K, L+1\} \ll \gamma$.

Proof. For any $\omega \in \Pi_h(K)$, there is a $\omega_0 \in \mathcal{O}_\gamma(K)$ such that $|\omega - \omega_0| \leqslant h$. Let $M = \Delta(k, \omega) - JA(\omega)$, $M_0 = \Delta(k, \omega_0) - JA(\omega_0)$, $M_1 = M - M_0$. In view of (2.4), $\|M_1\| \lessdot h \max\{K, L\}$. It follows that

$$\begin{aligned}\|M^{-1}\| &= \|(M_0 + M_1)^{-1}\| \lessdot \|M_0^{-1}\| \cdot \|(I + M_0^{-1} M_1)^{-1}\| \\ &\lessdot \frac{\|M_0^{-1}\|}{1 - \|M_0^{-1}\| \cdot \|M_1\|} \lessdot \|M_0^{-1}\| \lessdot \frac{|k|^{4m^2+\tau}}{\gamma} \end{aligned} \tag{2.19}$$

since $1/(1 - \|M_0^{-1}\| \cdot \|M_1\|) < 2$ if $h\,|k|^{4m^2+\tau} \max\{K, L\} = \varepsilon^{1/(4m^2+1)} K^{4m^2+\tau} \max\{K, L\} \ll \gamma$.

Similarly, we can prove other inequalities. ∎

2.4. *Solving the Homogeneous Equation*

Let $[R] = (1/2\pi) \int_0^{2\pi} R\,dx$, $\tilde{R} = R - [R]$, $N_+ = N + \int_0^{2\pi} R\,dx - \sum_{i=\pm 1}^{\pm m} R_{00i} u_i$, i.e.,

$$N_+ = N + P_{000} + \sum_{i,\, j = \pm 1}^{\pm m} R_{00ij} u_i u_j + \langle P_{010}, y\rangle, \tag{2.20}$$

where R_{00i}, R_{00ij} are defined after (2.11).

Now we look for a Hamiltonian Function F of the form

$$\begin{aligned} F(x, y, u) &\\ &= F_0 + F_1 + F_2 \\ &= \sum_{|l| \leqslant 1,\, 0 \neq |k| \leqslant K} f_{kl0} e^{\mathrm{i}(k, x)} y^l + \sum_{k \neq 0} \langle F_{ku}, u\rangle\, e^{\mathrm{i}(k, x)} + \sum_{k \neq 0} \langle F_{kuu} u, u\rangle\, e^{\mathrm{i}(k, x)} \\ &= \sum_{|l| \leqslant 1,\, k \neq 0} f_{kl0} e^{\mathrm{i}(k, x)} y^l + \sum_{k \neq 0} f_{k0i} e^{\mathrm{i}(k, x)} u_i + \sum_{k \neq 0} f_{k0ij} e^{\mathrm{i}(k, x)} u_i u_j, \end{aligned} \tag{2.21}$$

which solves the homogeneous equation (2.9).

LEMMA 2.4. *The homogeneous equation* (2.9) *is equivalent to*

$$\begin{aligned} F_{kl0} &= \Delta^{-1}(k)\, P_{kl0}, \qquad k \neq 0, \\ (\Delta(k) I + AJ)\, F_{ku} + R_{ku} &= 0, \\ (\Delta(k) I + AJ)\, F_{kuu} - F_{kuu} JA + R_{kuu} &= 0, \qquad k \neq 0. \end{aligned} \tag{2.22}$$

Proof. Insert F into (2.9), we arrive at the following equations

$$\begin{aligned} \{N, F_0\} + P_0 - \langle P_{01}, y\rangle &= 0, \qquad k \neq 0 \\ \{N, F_1\} + P_1 &= 0, \\ \{N, F_2\} + P_2 - \langle R_{0uu} u, u\rangle &= 0, \qquad k \neq 0. \end{aligned} \tag{2.23}$$

The first equation in (2.22), by comparing the coefficients, is obviously equivalent to the first equation in (2.23). To solve $\{N, F_1\} + P_1 = 0$, we note that

$$\begin{aligned} \{N, F_1\} &= \Delta(k)\, F_1 + \langle \nabla_u N, J\nabla_u F_1\rangle \\ &= \sum_{|k| \leqslant K} (\langle \Delta(k)\, F_{ku}, u\rangle + \langle Au, JF_{ku}\rangle)\, e^{\mathrm{i}(k, x)} \\ &= \sum_{|k| \leqslant K} (\langle \Delta(k)\, F_{ku}, u\rangle + \langle AJF_{ku}, u\rangle)\, e^{\mathrm{i}(k, x)} \\ &= \sum_{|k| \leqslant K} \langle (\Delta(k)\, I + AJ)\, F_{ku}, u\rangle\, e^{\mathrm{i}(k, x)}. \end{aligned} \tag{2.24}$$

It follows that F_{ku} are determined by the following $2m$ linear algebraic system

$$(\Delta(k) I + AJ)\, F_{ku} + R_{ku} = 0.$$

Similarly, from

$$\begin{aligned} \{N, F_2\} &= \Delta(k)\, F_2 + \langle \nabla_u N, J\nabla_u F_2\rangle \\ &= \sum_{0 \neq |k| \leqslant K} (\langle \Delta(k)\, F_{kuu} u, u\rangle + \langle Au, 2JF_{kuu} u\rangle)\, e^{\mathrm{i}(k, x)} \\ &= \sum_{0 \neq |k| \leqslant K} (\langle \Delta(k)\, F_{kuu} u, u\rangle + \langle (AJF_{kuu} - F_{kuu} JA) u, u\rangle)\, e^{\mathrm{i}(k, x)} \\ &= \sum_{0 \neq |k| \leqslant K} \langle (\Delta(k)\, F_{kuu} + AJF_{kuu} - F_{kuu} JA) u, u\rangle\, e^{\mathrm{i}(k, x)} \end{aligned} \tag{2.25}$$

it follows that, F_{kuu} is determined by the following matrix equation

$$(\Delta(k) I + AJ)\, F_{kuu} - F_{kuu} JA + R_{kuu} = 0, \qquad k \neq 0, \tag{2.26}$$

where F_{kuu}, R_{kuu} are symmetric $2m \times 2m$ matrices with elements F_{k0ij}, R_{k0ij} respectively. ∎

In order to solve (2.26), we need the following result in matrix theory.

LEMMA 2.5. *Let A, B, C are $n\times n, m\times m, n\times m$ matrices respectively, and X is an $n\times m$ unknown matrix. The matrix equation*

$$AX+XB=C$$

is solvable if and only if $I_m\otimes A^T+B\otimes I_n$ is nonsingular. Moreover,

$$\|X\|<\|(I_m\otimes A^T+B\otimes I_n)^{-1}\|\cdot\|C\|.$$

In fact, the matrix equation is equivalent to a bigger vector equation $(I\otimes A+B^T\otimes I)\,X'=C'$ by listing the elements of X, C as vectors. We refer to the Appendix of this paper or [10], p. 256 for a detailed proof.

COROLLARY 1. (2.29) *has a symmetric matrix solution if* (2.18) *is satisfied.*

Proof. (2.26) has a solution provided that (2.18) is satisfied is a consequence of Lemma 2.5. Note that F_{kuu}^T obeys the same equation as F_{kuu}. It follows that $F_{kuu}=F_{kuu}^T$. ∎

Moreover, we have the following estimates for F which play a key role in proving the smallness of the new perturbation. Let $D_i=D(\rho_+ +(i-1/4)(\rho-\rho_+), (i/4)s)$, $0<i\leqslant 4$ with $\rho_+<\rho$.

LEMMA 2.6.

$$\frac{1}{s^2}\|F\|_{D_3\times\Pi_h}\lessdot\gamma^{-1}\Gamma\varepsilon.$$

Here and after Γ stands for $\Gamma(\frac{1}{2}(\rho-\rho_+))$.

Proof. By (2.12), Lemma 2.3 and Lemma 2.4, we have

$$\begin{aligned}
|f_{kl0}| &\leqslant |\Delta|^{-1}\,|P_{kl}|\lessdot\gamma^{-1}\,|k|^{\tau}\,e^{-|k|\rho}\,\varepsilon s^{2-2\,|l|}, \qquad k\neq 0,\\
|f_{k0i}| &= |(\Delta(k,\omega)\,I+AJ)^{-1}\,R_{ku}|\leqslant\|(\Delta(k,\omega)I-JA)^{-1}\|\cdot|R_{ku}|\\
&\lessdot\gamma^{-1}\,|k|^{(4m^2+1)\tau}e^{-|k|\,\rho}\,\varepsilon s,\\
|f_{k0ij}| &\leqslant\|F_{kuu}\|\leqslant\|(\Delta(k,\omega)\,I-I\otimes(JA)-(JA)\otimes I)^{-1}\|\cdot\|R_{kuu}\|\\
&\lessdot\gamma^{-1}\,|k|^{4m^2(2m-1)^2+\tau}e^{-|k|\,\rho}\varepsilon s^2, \qquad k\neq 0.
\end{aligned}\tag{2.28}$$

It follows that

$$\frac{1}{s^2}\|F\|_{D_3\times\Pi_h}\leqslant\sum\frac{1}{s^2}(|f_{kl0}|\cdot|y^l|\cdot|e^{i(k,x)}|+|f_{k0i}u_i|\cdot|e^{i(k,x)}|$$
$$+|f_{k0ij}u_iu_j|\cdot|e^{i(k,x)}|)$$
$$\lessdot\sum_{k\neq 0}\gamma^{-1}|k|^{4m^2(2m-1)^2+\tau}e^{-(1/2)|k|(\rho-\rho_+)}\varepsilon\lessdot\gamma^{-1}\Gamma\varepsilon. \quad ∎ \tag{2.29}$$

By Cauchy inequality, we have

$$\frac{1}{\rho-\rho_+}\|F_x\|,\qquad s^2\|F_y\|,\qquad s\|F_u\|\lessdot\gamma^{-1}\Gamma\varepsilon s^2, \tag{2.30}$$

uniformly on $D_2\times\Pi_h$.

Let

$$\|D^mF\|_{D\times\Pi_h}=\max\left\{\left|\frac{\partial^{|i|+|l|+p}}{\partial x^i\,\partial y^l\,\partial u^p}F\right|_{D\times\Pi_h},\ |i|+|l|+|p|=m\right\}.$$

Note that F is polynomial in y of order 1, in u of order 2. From the Cauchy inequality, it also follows that

$$\|D^mF\|_{D_1\times\Pi_h}\lessdot\gamma^{-1}\Gamma\varepsilon,$$

for any $m\geqslant 2$.

2.5. *The New Normal Form and New Frequencies*

ϕ_F^1 defined above transforms H into $H_+=N_++P_+$ (see (2.7) and (2.9)) with

$$N_+=e_++\langle\omega_+,y\rangle+\tfrac{1}{2}\langle A_+u,u\rangle$$
$$=e_++\sum_{i=1}^{n}\omega_i+y_i+\tfrac{1}{2}\sum_{i,j=\pm1}^{\pm m}a_{ij}^+u_iu_j, \tag{2.32}$$

where

$$e_+=e+P_{000},\qquad\omega_+=\omega+P_{0l0}(|l|=1),\qquad A_+=A+R_{0uu}.$$

Since $\int_K^\infty x^n e^{-x(\rho-\rho_+)}\,dx \leqslant \varepsilon$, $|\omega_{+i}-\omega_i|$, $|a_{ij}^+ - a_{ij}| \leqslant 2\varepsilon$, we have

$$\begin{aligned}|\det(A_+)| &> |\det A - 4^{m^2}(2m)!\,(|k|+4m^2L)^{4m^2-1}\,\varepsilon \\ &> \tfrac{1}{2}|\det(A)|, \\ |\Delta(k,\omega+P_{0l00})\,I - JA_+| &> \gamma K^{-\tau} - 4^{m^2}(2m)!\,(|k|+4m^2L)^{4m^2-1}\,\varepsilon \\ &> \gamma_+ K^{-\tau},\end{aligned} \tag{2.34}$$

for $0 \neq k \leqslant K$, if $K^{4m^2-1}\varepsilon \ll \min\{\gamma-\gamma_+, |\det(A)|\}$.

Similarly, we have

$$\begin{aligned}&|\Delta(k,\omega+P_{0l00})I - I\otimes(JA_+) - (JA_+)\otimes I| \\ &\quad > \gamma K^{-\tau} - 4^{m^2}(2m)!\,(|k|+4m^2L)^{4m^2-1}\,\varepsilon \\ &\quad > \gamma_+ K^{-\tau}\end{aligned} \tag{2.35}$$

for $0 \neq k \leqslant K$, if $K^{4m^2-1}\varepsilon \ll \min\{\gamma-\gamma_+, |\det(A)|\}$.

The following estimate due to Cauchy estimates is for the proof of Theorem 3:

$$\left(\frac{h}{2}\right)^{|l|}\left|\frac{\partial^l(\omega_+-\omega)}{\partial\omega^l}\right|_{\Pi_{h/4}}, \qquad \left(\frac{h}{2}\right)^{|l|}\left|\frac{\partial^l(a_{ij}^+-a_{ij})}{\partial\omega^l}\right|_{\Pi_{h/4}} \lessdot \varepsilon. \tag{2.36}$$

2.6. *Estimates for the New Perturbation*

To finish one cycle of iteration, the only thing remained is to estimate the new error term. In the next lemma, we give some estimates for ϕ_F^t. The following (2.37) will be used to prove our coordinate transformations is well defined. (2.38) is for proving the convergence of the iteration.

Let $D_{(i/2)\alpha} = D(\rho_+ + (i-1/2)(\rho-\rho_+), (i/2)s_+)$. We have

LEMMA 2.7.

$$\phi_F^t\colon D_{(1/2)\alpha} \to D_\alpha, \qquad 0 \leqslant t \leqslant 1, \tag{2.37}$$

if $\varepsilon \ll (\frac{1}{2}\gamma\Gamma^{-1})^{3/2}$. *Moreover,*

$$\|D\phi_F^1 - Id\|_{D_{(1/2)\alpha}} < \varepsilon, \qquad \|D^2\phi_F^1\|_{D_{(1/2)\alpha}} < c. \tag{2.38}$$

Proof. To get the estimates for ϕ_F^t, we start from the integral equation,

$$\phi_F^t = id + \int_0^t X_F \circ \phi_F^s\,ds$$

$\phi_F^t\colon D_{(1/2)\alpha} \to D_\alpha$, $0 \leqslant t \leqslant 1$, follows directly from (2.30). Since

$$D\phi_F^1 = Id + \int_0^1 (DX_F)\, D\phi_F^s\, ds = Id + \int_0^1 J(D^2F)\, D\phi_F^s\, ds,$$

it follows that

$$\|D\phi_F^1 - Id\| \leqslant 2\, \|D^2F\| \lessdot \gamma^{-1}\Gamma\varepsilon. \tag{2.39}$$

The estimates of second order derivative $D^2\phi_F^1$ follows from (2.31). ∎

By the definition of ϕ_F^1 and Lemma 2.7, we know

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

is well defined in $D_{(1/2)\alpha}$. Moreover, we have the following estimates

$$\begin{aligned}
\|P_+\|_{D_{(1/2)\alpha}} &= \left\| \int_0^1 \{R_t, F\} \circ \phi_F^t\, dt + \{R, F\} + (P - R) \circ \phi_F^1 \right\|_{D_{(1/2)\,\alpha}} \\
&= \left\| \int_0^1 \{R_t, F\} \circ \phi_F^t\, dt \right\|_{D_{(1/2)\,\alpha}} + \|\{R, F\}\|_{D_{(1/2)\alpha}} + \|(P-R) \circ \phi_F^1\|_{D_{(1/2)\alpha}} \\
&\leqslant \|\{R_t, F\}\|_{D_\alpha} + \|\{R, F\}\|_{D_\alpha} + \|(P-R)\|_{D_\alpha} \\
&\lessdot \gamma^{-2}\Gamma\varepsilon^2 s^2,
\end{aligned} \tag{2.40}$$

where $R_t = (1-t)\{N + R, F\}$.

That is, there is a big constant c, independent of iteration steps, such that

$$\|P_+\| \leqslant c\gamma^{-2}\Gamma\varepsilon^2 s^2, \tag{2.41}$$

which implies

$$\frac{1}{s_+^2}\, \|P_+\| \leqslant c\gamma^{-2}\Gamma\varepsilon^{4/3} \leqslant c\varepsilon_+. \tag{2.42}$$

One circle of KAM step is finished.

3. ITERATION AND CONVERGENCE

For any given $s_0, \varepsilon_0, \rho_0$, we define some sequences inductively depending on $s_0, \varepsilon_0, \rho_0$

$$\rho_\nu = r\left(1 - \sum_{i=2}^{\nu+1} 2^{-i}\right),$$

$$\varepsilon_\nu = c\gamma_{\nu-1}^{-2}\Gamma\left(\frac{1}{2}(\rho_{\nu-1} - \rho_\nu)\right)\varepsilon_{\nu-1}^{4/3}.$$

$$\alpha_\nu = \varepsilon_\nu^{1/3}, \qquad h_\nu = \varepsilon_\nu^{1/(8m^2+1)},$$

$$\delta_\nu = \sum_{i=1}^{\nu} 2\frac{\varepsilon_i}{h_i^{8m^2}} = \sum_{i=0}^{\nu-1} 2\varepsilon_i^{(1/8m^2+1)}, \qquad L_\nu = L + \delta_\nu,$$

$$s_\nu = \frac{1}{2}\alpha_{\nu-1}s_{\nu-1} = 2^{-\nu}\left(\prod_{i=0}^{\nu-1}\varepsilon_i\right)^{1/3} s_0,$$

$$K_\nu = \min\left\{K\colon \int_K^\infty K^{n\tau}e^{-K(\rho_{\nu-1}-\rho_\nu)} \leqslant \varepsilon_\nu\right\}$$

$$\gamma_\nu = \gamma\left(1 - \sum_{i=2}^{\nu+1} 2^{-i}\right),$$

$$D_\nu = D(r_\nu, s_\nu), \tag{3.1}$$

where c is the constant in (2.42).

As a matter of fact,

$$\Psi(r) = \prod_{i=1}^{\infty} \Gamma(\tfrac{1}{2}(\rho_i - \rho_{i-1}))^{(3/4)^i}$$

is a well-defined finite function of r (see, e.g., [14]).

Summarizing conclusions of last section, we have the following iteration lemma.

LEMMA 3.1. *Suppose that ε_0 is small enough so that*

$$\varepsilon_0 \leqslant \min\left\{\frac{\gamma^2}{4c\Psi(r)}, \left(\frac{\gamma}{c\Gamma(\frac{1}{4}r)}\right)^2\right\},$$

$$\varepsilon_0^{1/(4m^2+1)}K_0^{4m^2+\tau}\max\{K_0, L+1\} \leqslant c^{-1}\gamma \tag{3.2}$$

$$K_0^{8m^2-1}\varepsilon_0 \leqslant c^{-1}\min\left\{\frac{1}{4}\gamma, |\det(A)|\right\}$$

for a big constant c depending on m, n, τ, L, γ, r. Then the following holds for all ν: Let

$$N_\nu = e_\nu + \langle \omega_\nu(\omega), y\rangle + \langle A_\nu(\omega)u, u\rangle = e_\nu + \langle \omega_\nu, y\rangle + \sum_{i,j=\pm 1}^{\pm m} a_{ij}^\nu u_i u_j$$

be a normal form with parameters w satisfying

$$|\Delta(k,\omega_\nu)| > \frac{\gamma_\nu}{|k|^\tau}, \qquad |\Delta(k,\omega_\nu)\, I_{2m} - JA_\nu|_a > \frac{\gamma_\nu}{|k|^\tau}$$
$$|\Delta(k,\omega_\nu)\, I_{4m^2} - I_{2m}\otimes(JA_\nu) - (JA_\nu)\otimes I_{2m}|_a > \frac{\gamma_\nu}{|k|^\tau} \tag{3.3}$$

on a closed set $\mathcal{O}_\nu$ *of* R^n *for* $|k| \leqslant K_\nu$. *Moreover, suppose that* $\omega_\nu(\omega)$, $a^\nu_{ij}(\omega)$ *are real analytic and satisfy*

$$\left|\frac{\partial^{8m^2}(\omega_\nu - \omega)}{\partial\omega^{8m^2}}\right| \leqslant \delta_\nu, \qquad \left|\frac{\partial^{8m^2} a^\nu_{ij}}{\partial\omega^{8m^2}}\right| \leqslant L_\nu$$

on the complex h_ν *neighborhood* Π_ν *of* $\mathcal{O}_\nu$.

Finally,

$$\frac{1}{s_\nu^2}\, \|P_\nu\|_{D_\nu\times\Pi_\nu} \leqslant \varepsilon_\nu.$$

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

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

where

$$\mathcal{R}^{\nu+1}_k(\gamma_{\nu+1})$$
$$= \left\{\omega\in\mathcal{O}_\nu : \left| \begin{array}{c} |\langle k,\omega_{\nu+1}\rangle|,\ |\mathrm{i}\langle k,\omega_{\nu+1}\rangle\, I_{2m} - JA_{\nu+1}|_a \leqslant \gamma_{\nu+1}/|k|^\tau, \quad or \\ |\mathrm{i}\langle k,\omega_{\nu+1}\rangle\, I_{4m^2} - I_{2m}\otimes(JA_{\nu+1}) - (JA_{\nu+1})\otimes I_{2m}|_a \leqslant \gamma_{\nu+1}/|k|^\tau \end{array}\right.\right\},$$

with $\omega_{\nu+1} = \omega_\nu + P^\nu_{0l0}$, *and a symplectic change of variables*

$$\Phi_\nu : D_{\nu+1}\times\Pi_{\nu+1} \to D_\nu, \tag{3.4}$$

such that $H_{\nu+1} = H_\nu \circ \Phi_\nu$, *defined on* $D_{\nu+1}\times\Pi_{\nu+1}$, *has the form*

$$H_{\nu+1} = e_{\nu+1} + \langle\omega_{\nu+1}, y\rangle + \sum_{i,j=\pm1}^{\pm m} a^{\nu+1}_{ij} u_i u_j + P_{\nu+1}, \tag{3.5}$$

satisfying

$$\max_{|l|\leqslant 8m^2}\left|\frac{\partial^l(\omega_{\nu+1}(\omega)-\omega)}{\partial\omega^l}\right|_{\Pi_{\nu+1}}\leqslant\delta_{\nu+1},$$
$$\max_{|l|\leqslant 8m^2}\left|\frac{\partial^l a_{ij}^{\nu+1}(\omega)}{\partial\omega^l}\right|_{\Pi_{\nu+1}}\leqslant L_{\nu+1}, \tag{3.6}$$

$$\frac{1}{s_{\nu+1}^2}\|P_{\nu+1}\|_{D_{\nu+1}\times\Pi_{\nu+1}}\leqslant\varepsilon_{\nu+1}. \tag{3.7}$$

on the complex $h_{\nu+1}$ neighborhood $\Pi_{\nu+1}$ of $\mathcal{O}_{\nu+1}$.

4. PROOF OF THEOREM 1

Since Theorem 1 is the combination of the following Theorem 2 and Theorem 3, we only need to give the proofs for the latter two theorems. Theorem 2 is actually the analytic part of the KAM theory, which is independent of Theorem 3, the geometric part of the KAM theory. In fact, we will run KAM machinery first for the perturbed Hamiltonian systems no matter the unperturbed system is non-degenerate or not. Then we use the non-degeneracy condition (1.8) to prove that a positive measure set of the parameter will survives the KAM iteration.

THEOREM 2. *Suppose that the Hamiltonian* (1.2) *with a non-singular matrix A is real analytic. Then for any given γ, there is a small constant ε_0 depending on n, τ, r, L, γ, h, such that if the complex extension of P in $D(r, s)$ satisfies*

$$\frac{1}{s^2}\|P\|_{D(r,s)\times\Pi_h}=\varepsilon\leqslant\varepsilon_0,$$

we have the following conclusions: there exists a Cantor set $\mathcal{O}_\gamma\subset\mathcal{O}$, a smooth family of torus embedding

$$\Phi\colon T^n\times\mathcal{O}_\gamma\to T^n\times R^n\times R^{2m},$$

and a diffeomorphism $\tilde{\omega}\colon\mathcal{O}_\gamma\to R^n$, such that the map Φ restricted to $T^n\times\{\omega\}$ is a real analytic embedding of a rotational torus with frequencies $\tilde{\omega}$ for the Hamiltonian H at ω.

Moreover, there are smooth matrices A_ν on $\mathcal{O}_\nu$ for $\nu \geqslant 0$ with $A_0 = A$, $\mathcal{O}_0 = \mathcal{O}$, such that $\mathcal{O} - \mathcal{O}_\gamma \subset \bigcup \mathcal{R}_k^\nu(\gamma)$, where

$$\mathcal{R}_k^\nu(\gamma) = \left\{ \omega : \left| \begin{array}{c} |\langle k, \omega \rangle|, \ |\mathrm{i}\langle k, \omega \rangle I_{2m} - JA_\nu|_a \leqslant \gamma/|k|^\tau, \qquad or \\ |\mathrm{i}\langle k, \omega\rangle I_{4m^2} - I_{2m} \otimes (JA_\nu) - (JA_\nu) \otimes I_{2m}|_a \leqslant \gamma/|k|^\tau \end{array} \right. \right\}$$

for $K_{\nu-1} \leqslant |k| < K_\nu$.

From Theorem 2, one can not get any information about the size of $\mathcal{O}_\gamma$ since (1.7) is not assumed. The next theorem shows that $\mathcal{O}_\gamma$ in Theorem 2 is quite large if (1.7) holds.

THEOREM 3. *Under the assumptions of Theorem* 1, *if the Hamiltonian H is non-degenerate in the sense of* (1.7), *then beside the conclusions in Theorem* 2, *we also have*

$$|\mathcal{O} - \mathcal{O}_\gamma| \to 0, \qquad as \quad \gamma \to 0.$$

Moreover, if $|\bigcup \mathcal{R}_k^0| \leqslant c\gamma^{1/4m^2}|k|^{-\tau}$, then

$$|\mathcal{O} - \mathcal{O}_\gamma| = O(\gamma^{1/4m^2}).$$

Proof of Theorem 2. Suppose that the assumptions of Theorem 1 are satisfied. To apply the iteration lemma with $\nu = 0$, we set

$$s_0 = s, \qquad N_0 = N, \qquad P_0 = P, \qquad \gamma_0 = \gamma, \qquad L_0 = L,$$

$$\mathcal{O}_0 = \left\{ \omega \in \mathcal{O} : \left| \begin{array}{c} |\langle k, \omega\rangle|, \ |\mathrm{i}\langle k, \omega\rangle I_{2m} - JA|_a > \gamma/|k|^\tau, \qquad \text{or} \\ |\mathrm{i}\langle k, \omega\rangle I_{4m^2} - I_{2m} \otimes (JA) - (JA) \otimes I_{2m}|_a > \gamma/|k|^\tau, \quad 0 \neq |k| \leqslant K_0 \end{array} \right. \right\}.$$

Taking ε_0 satisfying (3.2), then the iteration lemma applies. Inductively, we obtain the following sequences

$$D_\nu \times \mathcal{O}_\nu \subset D_{\nu-1} \times \mathcal{O}_{\nu-1},$$

$$\Psi^\nu = \Phi_1 \circ \cdots \circ \Phi_\nu : D_\nu \times \mathcal{O}_{\nu-1} \to D_0, \qquad \nu \geqslant 1,$$

$$H \circ \Psi^\nu = H_\nu = N_\nu + P_\nu.$$

Let $\mathcal{O}_\gamma = \bigcap_{\nu=0}^\infty \mathcal{O}_\nu$. Similar to the argument in [14], in view of Lemma 2.7, it concludes that N_ν, Ψ^ν, $D\Psi^\nu$ converges uniformly on $D_\infty \times \mathcal{O}_\gamma = D(\frac{1}{2}r, 0, 0) \times \mathcal{O}_\gamma$ with

$$N_\infty = e_\infty + \langle \omega_\infty, y\rangle + \langle A_\infty u, u\rangle = e_\infty + \langle \omega_\infty, y\rangle + \sum_{i,j=\pm 1}^{\pm m} a_{ij}^\infty u_i u_j.$$

Since

$$\varepsilon_\nu = c\gamma_{\nu-1}^{-2}\Gamma(\tfrac{1}{2}(\rho_{\nu-1}-\rho_\nu))\,\varepsilon_{\nu-1} \leqslant (c\gamma^{-2}\Psi(r)\,\varepsilon_0)^{(4/3)^\nu}.$$

By the choice of ε_0 in (3.2), $\varepsilon_\nu \to 0$.

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

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

The convergence of Ψ^ν, $D\Psi^\nu$, X_H, X_{H_ν} implies that one can take limit in (4.1) and arrive at

$$\phi_H^t \circ \Psi^\infty = \Psi^\infty \circ \phi_{H_\infty}^t \tag{4.2}$$

on $D(\frac{1}{2}r, 0, 0) \times \mathcal{O}_\gamma$, where

$$\Psi^\infty \colon \mathcal{O}_\gamma \times T^n \to R^n \times T^n \times R^{2m}.$$

It follows from (4.2),

$$\phi_H^t(\Psi^\infty(\{\omega\} \times T^n)) = \Psi^\infty \phi_{N_\infty}^t(\{\omega\} \times T^n) = \Psi^\infty(\{\omega\} \times T^n),$$

for $\omega \in \mathcal{O}_\gamma$. That means $\Psi^\infty(\{\omega\} \times T^n)$ is an embedding invariant torus of the original perturbed Hamiltonian system at $\omega \in \mathcal{O}_\gamma$. We remark here the frequencies of $\Psi^\infty(\{\omega\} \times T^n)$ is slightly different from ω. The normal behavior of the invariant torus is governed by the matrix $A^\infty = (a_{ij}^\infty(\omega))$. ∎

Proof of Theorem 3. By the iteration lemma,

$$\mathcal{O} - \mathcal{O}_\gamma \subset \bigcup_{\nu=0}^{\infty} \mathcal{R}_k^\nu(\gamma_\nu),$$

where

$$\mathcal{R}_k^{\nu+1}(\gamma_{\nu+1}) = \left\{ \omega \in \mathcal{O}_\nu \colon \begin{array}{c} |\langle k, \omega_{\nu+1}\rangle|,\ |\mathrm{i}\langle k, \omega_{\nu+1}\rangle\, I_{2m} - JA_{\nu+1}|_a \leqslant \gamma_{\nu+1}/|k|^\tau, \quad \text{or} \\ |\mathrm{i}\langle k, \omega_{\nu+1}\rangle\, I_{4m^2} - I_{2m} \otimes (JA_{\nu+1}) - (JA_{\nu+1}) \otimes I_{2m}|_a \leqslant \gamma_{\nu+1}/|k|^\tau \end{array} \right\}$$

$K_\nu \leqslant |k| \leqslant K_{\nu+1}$.

In order to estimate the measure of $\mathcal{R}_k^\nu$, we need the following lemma, which has been proven in [21, 23]. Similar estimate is also used by Bourgain [3].

LEMMA 4.1. *Suppose that $g(u)$ is a C^m function on the closure $\bar{I}$, where $I \subset R^1$ is an interval. Let $I_h = \{u \mid |g(u)| < h\}$, $h > 0$. If for some constant*

$d>0$, $|g^{(m)}(u)|\geqslant d$ for $\forall u\in I$, then $|I_h|\leqslant ch^{1/m}$, where $|I_h|$ denotes the Lebesgue measure I_h and $c=2(2+3+\cdots+m+d^{-1})$.

In the following, we estimate the Lebesgue measure of $\mathscr{R}_k^{\nu+1}$ for fixed k.

LEMMA 4.2. *There is a K^* depending on n, L, m, such that if $|k|\geqslant K^*$, then*

$$|\mathscr{R}_k^{\nu+1}(\gamma)|\lessdot c_2\gamma^{1/4m^2}\,|k|^{-\tau/4m^2}.$$

Proof. Note that

$$\mathscr{R}_k^{\nu+1}=\mathscr{R}_{k1}^{\nu+1}+\mathscr{R}_{k2}^{\nu+1}+\mathscr{R}_{k3}^{\nu+1},$$

where

$$\mathscr{R}_{k1}^{\nu+1}(\gamma_{\nu+1})=\left\{\omega\in\mathcal{O}_\nu\colon |\langle k,\omega_{\nu+1}(\omega)\rangle|\leqslant\frac{\gamma_{\nu+1}}{|k|^\tau}\right\},$$

$$\mathscr{R}_{k2}^{\nu+1}(\gamma_{\nu+1})=\left\{\omega\in\mathcal{O}_\nu\colon |\,|\mathrm{i}\langle k,\omega_{\nu+1}(\omega)\rangle\, I_{2m}-JA_{\nu+1}|_a\leqslant\frac{\gamma_{\nu+1}}{|k|^\tau}\right\},$$

$$\mathscr{R}_{k3}^{\nu+1}(\gamma_{\nu+1})=\left\{\omega\in\mathcal{O}_\nu\colon ||\mathrm{i}\langle k,\omega_{\nu+1}(\omega)\rangle\, I_{4m^2}-I_{2m}\otimes(JA_{\nu+1})\right.$$
$$\left.-(JA_{\nu+1})\otimes I_{2m}|_a\leqslant\frac{\gamma_{\nu+1}}{|k|^\tau}\right\}.$$

We estimate the measure for the most complicated set $\mathscr{R}_{k3}^{\nu+1}(\gamma_{\nu+1})$, the other two are similar. Let

$$M_k(\omega)=|\mathrm{i}\langle k,\omega_{\nu+1}(\omega)\rangle\, I_{4m^2}-I_{2m}\otimes(JA_{\nu+1})-(JA_{\nu+1})\otimes I_{2m}|_a^2$$

defined in $\mathcal{O}_\nu$. It follows that

$$\mathscr{R}_{k3}^{\nu+1}(\gamma_{\nu+1})\subset\left\{\omega\in\mathcal{O}_\nu\colon M_k(\omega)\leqslant\frac{\gamma_{\nu+1}^2}{|k|^{2\tau}}\right\}.$$

As a matter of fact, $M_k(\omega)$ is a polynomial of $\langle k,\omega\rangle$ of the form

$$M_k=\langle k,\omega\rangle^{8m^2}+\sum_{l\leqslant 8m^2-1}p_l(\omega)\langle k,\omega\rangle^l,$$

where the coefficients depend on m and $A_{\nu+1}$.

By (3.6), we have

$$\left|\frac{\partial^l(\omega_{\nu+1}(\omega)-\omega)}{\partial\omega_+^l}\right|\leqslant\frac{1}{2},$$
$$\left|\frac{\partial^l a_{ij}^{\nu+1}(\omega)}{\partial\omega^l}\right|\leqslant L_{\nu+1}\leqslant L+1. \tag{4.3}$$

Without loss of generality, we assume $k_1\geqslant K_\nu/n$ if $|k|\geqslant K_\nu$. It follows that

$$\left|\frac{d^{8m^2}}{d\omega_1^{8m^2}}M_k(\omega)\right|\geqslant(8m^2)!\,|k_1|^{8m^2}(1-O(|k_1^{-1}|)). \tag{4.4}$$

It concludes that there is an integer $K^*(n, m, L)$ such that if $|k|\geqslant K^*$,

$$\frac{d^{8m^2}}{d\omega_{+1}^{8m^2}}|M_k(\omega)|_a>\frac{1}{2}.$$

From Lemma 4.1, it follows that

$$|\mathscr{R}_{k3}^{\nu+1}(\gamma_{\nu+1})|\leqslant c\left(\frac{\gamma_{\nu+1}}{|k|^\tau}\right)^{1/4m^2}\leqslant c\gamma^{1/4m^2}\,|k|^{-\tau/4m^2},$$

where $c=(\frac{1}{2}m(m+1)+\frac{1}{2})\,D^{n-1}$ with D the diameter of $\mathcal{O}$. ∎

We arrive at

$$\left|\bigcup_{|k|\geqslant K^*}\mathscr{R}_k^{\nu+1}(\gamma_{\nu+1})\right|\leqslant\sum_{|k|\geqslant K^*}|R_k^{\nu+1}(\gamma_{\nu+1})|$$
$$\leqslant 3\sum_{|k|\geqslant K^*}c\gamma^{1/4m^2}\,|k|^{-\tau/4m^2}=O(\gamma^{1/4m^2})$$

if $\tau>4m^2(n-1)$.

To prove Theorem 2, we let ε_0 is small enough so that $K_0\geqslant K^*$ (the relation between K_0 and ε_0 is defined by (3.1)). It follows that

$$\mathscr{R}_k^0(\gamma)$$
$$=\left\{\omega\in\mathcal{O}:\left|\begin{array}{c}|\langle k,\omega\rangle|,\ |\mathrm{i}\langle k,\omega\rangle\,I_{2m}-JA|_a\leqslant\gamma/|k|^\tau,\quad\text{or}\\ |\mathrm{i}\langle k,\omega\rangle\,I_{4m^2}-I_{2m}\otimes(JA(\omega))-(JA(\omega))\otimes I_{2m}|_a\leqslant\gamma/|k|^\tau\end{array}\right.\right\}$$

for $0<|k|\leqslant K_0$. By the non-degeneracy condition (1.8),

$$\left|\bigcup_{|k|\leqslant K_0}\mathscr{R}_k^0(\gamma)\right|\to 0,\qquad\text{as}\quad\gamma\to 0.$$

Since $\mathcal{O}-\mathcal{O}_\gamma \subset \mathcal{R}_k^0(\gamma) \bigcup_{|k|\geqslant K^*} \mathcal{R}_k^{\nu+1}(\gamma_{\nu+1})|$, we know

$$|\mathcal{O}-\mathcal{O}_\gamma| \to 0.$$

It is clear that $|\bigcup_{|k|\leqslant K_0} \mathcal{R}_k^0(\gamma)| = O(\gamma^{1/4m^2})$ implies $|\mathcal{O}-\mathcal{O}_\gamma| = O(\gamma^{1/4m^2})$.

APPENDIX

In this section, we list some results in matrix theory which is used in this paper. Although the proofs are quite elementary and can be found in many textbooks in matrix theory (see, e.g., [10]), we sketch them for the convenience of the reader.

DEFINITION 5.1. The tensor product (or direct product) of two matrices A_{mn}, B_{kl} is a $(mk)\times(nl)$ defined by

$$A\otimes B=(a_{ij}B)=\begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \cdots & \cdots & \cdots \\ a_{n1}B & \cdots & a_{mn}B \end{pmatrix}.$$

In the following, we always assume that $A=(a_{ij})$, $B=(b_{ij})$, $C=(c_{ij})$ are $n\times n$, $m\times m$, $n\times m$ matrices respectively. $X=(x_{ij})$ is a $n\times m$ unknown matrix. The eigenvalues of A, B are denoted by $\lambda_1, ..., \lambda_n$ and $\mu_1, ..., \mu_m$ respectively.

For convenience, we also represent A, C, X as $A=(A_1, ..., A_n)$, $C=(C_1, ..., C_m)$, $X=(X_1, ..., X_m)$ respectively. Denote by $C'=(C_1^T, ..., C_m^T)^T$, $X'=(X_1^T, ..., X_m^T)^T$ the corresponding nm-vectors.

The object under consideration is the following matrix equation

$$AX+XB=C, \tag{5.1}$$

and vector equation

$$(I_m\otimes A+B^T\otimes I_n)\, X'=C', \tag{5.2}$$

LEMMA 5.1. *The matrix equation* (5.1) *is solvable if and only if the vector equation* (5.2) *is solvable.*

Proof. Rewrite (5.1) as

$$A(X_1, ..., X_m)+XB=(C_1, ..., C_m),$$

which is

$$AX_j + \sum_{i=1}^{m} X_i b_{ij} = C_j, \qquad j = 1, ..., m.$$

Written in another compact form, it is just the vector equation (5.2). ∎

COROLLARY 2. *(5.1) is solvable if and only if $I_m \otimes A + B^T \otimes I_n$ is nonsingular, or equivalently $I_m \otimes A^T + B \otimes I_n$ is nonsingular. Moreover,*

$$\|X\| \leqslant \|(I_m \otimes A^T + B \otimes I_n)^{-1}\| \cdot \|C\|.$$

LEMMA 5.2. *(5.1) is solvable for any given C if and only if*

$$\lambda_i + \mu_j \neq 0, \qquad \text{for} \quad i = 1, ..., n, \quad j = 1, ..., m.$$

Proof. Let T_1, T_2 are nonsingular matrices such that

$$T_1^{-1} A T_1 = \tilde{A} = \mathrm{Diag}(J_1, ..., J_s); \qquad T_2^{-1} B T_2 = \tilde{B} = \mathrm{Diag}(J'_1, ..., J'_{s'}),$$

where J_i, J'_j are Jordan blocks of the form

$$\begin{pmatrix} * & 0 & \cdots & 0 \\ 1 & * & 0 & 0 \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & * & 0 \\ 0 & 0 & 1 & * \end{pmatrix}.$$

(5.1) is reduced to

$$\tilde{A}(T_1^{-1} X T_2) + (T_1^{-1} X T_2)\, \tilde{B} = T_1^{-1} C T_2.$$

Without loss of generality, we assume that A and B are in Jordan normal form. By comparing the elements in the position $(1, m)$ in equation (5.1), we have

$$x_{1m} = \frac{1}{\lambda_1 + \mu_m} c_{1m}.$$

$x_{1,m-1}, ..., x_{11}; x_{2m}, ..., x_{21}; ...; x_{nm}, ..., x_{n1}$ can be solved inductively due to the Jordan form if $\lambda_i + \mu_j \neq 0$ for all $i = 1, ..., n$, $j = 1, ... m$. ∎

LEMMA 5.3. *The eigenvalues of $I_m \otimes A^T + B \otimes I_n$ are*

$$\lambda_{ij} = \lambda_i + \lambda_j, \qquad i = 1, ..., n, \quad j = 1, ... m.$$

Proof. By Lemma 5.2, (5.1) is solvable if and only if $\lambda_i + \mu_j \neq 0$. Together with Lemma 5.1 and Corollary 2, it follows that $I_m \otimes A^T + B \otimes I_n$ is nonsingular if and only if $\lambda_i + \mu_j \neq 0$, for $i = 1, ..., n$, $j = 1, ...m$. That means

$$\lambda_{ij} = \lambda_i + \mu_j, \qquad i = 1, ..., n, \quad j = 1, ...m,$$

are eigenvalues of $I_m \otimes A + B^T \otimes I_n$. ∎

LEMMA 5.4. *The eigenvalues of* $\mathrm{i}\langle k, \omega\rangle I_{4m^2} - I_{2m} \otimes (JA(\omega)) - (JA(\omega)) \otimes I_{2m}$ *are* $\mathrm{i}(\langle k, \omega\rangle \pm \Omega_i \pm \Omega_j)$, $i, j = 1, ..., m$.

Proof. Since the eigenvalues of JA are in pairs $\pm \mathrm{i}\Omega_i$, $i = 1, ..., m$. It follows that the eigenvalues of $\mathrm{i}\langle k, \omega\rangle I_{2m} - JA$ are $\mathrm{i}(\langle k, \omega\rangle \pm \Omega_i)$, $i = 1, ..., m$. Rewrite $\mathrm{i}\langle k, \omega\rangle I_{4m^2} - I_{2m} \otimes (JA) - (JA) \otimes I_{2m}$ as

$$(\mathrm{i}\langle k, \omega\rangle I_{2m} - JA) \otimes I_{2m} - I_{2m} \otimes (JA).$$

By Lemma 5.3, its eigenvalues are $\mathrm{i}(\langle k, \omega\rangle \pm \Omega_i \pm \Omega_j)$, $i, j = 1, ..., m$. ∎

COROLLARY 3. *The non-degeneracy conditions* (1.7) *are equivalent to* (1.8)

ACKNOWLEDGMENTS

This work was done during the author's visit at the Mathematisches Institut der Technischen Universität München and Mathematisches Institut A der Universität Stuttgart, and he is grateful for the hospitality of both institutes. The author thanks J. Pöschel and J. Scheurle for various valuable discussions. The revised version was written during the author's visit at Universita di Roma Tre under a grant of the Italian CNR. He thanks L. Chierchia for discussions on the presentation of the paper. Finally, he is deeply indebted to the referee for his carefully reading of the previous version of the manuscript and for many suggestions.

REFERENCES

1. J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, *Internat. Math. Res. Notices* (1994), 475–497.
2. J. Bourgain, On Melnikov's persistency problem, *Math. Res. Lett.* **4** (1997), 445–458.
3. A. D. Bruno, Analytic form of differential equations, *Trudy Moskov. Mat. Obshch.* **25** (1971) [in Russian]; *Trans. Moscow Math. Soc.* **25** (1971), 131–288.
4. C.-Q. Cheng and Y.-S. Sun, Existence of KAM tori in degenerate Hamiltonian systems, *J. Differential Equations* **114** (1994), 288–335.

4a. L. Chierchia and J. You, KAM tori for 1D nonlinear wave equations with periodic boundary conditions, preprint, 1998.

5. W. Craig and C. E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, *Commun. Pure Appl. Math.* **46** (1993), 1409–1498.
6. S. M. Graff, On the continuation of stable invariant tori for Hamiltonian systems, *J. Differential Equations* **15** (1974), 1–69.
7. L. H. Eliasson, Perturbations of stable invariant tori for Hamiltonian systems, *Ann. Scuola Norm. Sup. Pisa* **15** (1988), 115–147.
8. S. B. Kuksin, "Nearly Integrable Infinite Dimensional Hamiltonian Systems," Lecture Notes in Math., Vol. 1556, Springer-Verlag, Berlin, 1993.
9. S. B. Kuksin and J. Pöschel, Invariant cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation, *Ann. of Math.* **142** (1995), 149–179.
10. P. Lancaster, "Theory of Matrices," Academic Press, New York and London, 1969.
11. V. K. Melnikov, On some cases of the conservation of conditionally periodic motions under a small change of the Hamiltonian function, *Soviet Math. Dokl.* **6** (1965), 1592–1596.
12. J. Moser, Convergent series expansions for quasiperiodic motions, *Math. Ann.* **169** (1) (1967), 136–176.
13. I. O. Parasyuk, Coisotropic invariant tori of Hamiltonian systems in the quasiclassical theory of motion of a conduction electon, *Ukrain. Math. J.* **42** (3) (1990), 308–312.
14. J. Pöschel, On elliptic lower dimensional tori in Hamiltonian systems, *Math. Z.* **202** (1989), 559–608.
15. J. Pöschel, A KAM-Theorem for some Nonlinear Partial Differential Equations, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **23** (1996), 119–148.
16. J. Pöschel, Integrability of Hamiltonian systems on Cantor Sets, *Comm. Pure Appl. Math.* **25** (1982), 653–695.
17. H. Rüssmann, On twist Hamiltonians, talk on the Colloque International: Mécanique céleste et systemes hamiltonians, Marseille, 1990.
18. C. E. Wayne, Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, *Comm. Math. Phys.* **127** (1990), 479–528.
19. H. Whitney, Analytical extension of differentiable functions defined on closed set, *Trans. Amer. Math. Soc.* **36** (1934), 63–89.
20. Z. Xia, The existence of the invariant tori in the volume preserving diffeomorphism, *Ergodic Theory Dynam. Systems* **12** (1992), 621–631.
21. J. Xu, J. You, and Q. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* **226** (1997).
22. J.-C. Yoccoz, Travaux de Herman sur les tores invariants, *Sem. Bourbaki* **754** (1991–1992), 311–344; *Asterique* **206** (1992).
23. J. You, A KAM theorem for hyperbolic-type degenerate lower dimensional tori in Hamiltonian systems, *Comm. Math. Phys.* **192** (1998), 145–168.
24. E. Zehnder, Generalized implicit function theorems with applications to some small divisor problem, I and II, *Comm. Pure Appl. Math.* **28** (1975), 91–140; (1976), 49–111.