# Weakly hyperbolic invariant tori for two dimensional quasiperiodically forced maps in a degenerate case *

Tingting Zhang[1,3], Àngel Jorba[2] and Jianguo Si[3] †

[1]College of Mathematics, Qingdao University,
Qingdao, Shandong 266071, People's Republic of China
E-mail: ting_little@163.com

[2]Departament de Matemàtiques i Informàtica,
Universitat de Barcelona, Barcelona, Spain
E-mail: angel@maia.ub.es

[3]School of Mathematics, Shandong University,
Jinan, Shandong 250100, People's Republic of China
E-mails: ting_little@163.com, sijgmath@sdu.edu.cn

**Abstract**. In this work we consider a class of degenerate analytic maps of the form

$$\begin{cases} \bar{x} = x + y^m + \epsilon f_1(x, y, \theta, \epsilon) + h_1(x, y, \theta, \epsilon), \\ \bar{y} = y + x^n + \epsilon f_2(x, y, \theta, \epsilon) + h_2(x, y, \theta, \epsilon), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

where $mn > 1, n \geq m$, $h_1$ and $h_2$ are of order $n+1$ in $z$, and $\omega = (\omega_1, \omega_2, \ldots, \omega_d) \in \mathbb{R}^d$ is a vector of rationally independent frequencies. It is shown that, under a generic non-degeneracy condition on $f$, if $\omega$ is Diophantine and $\epsilon > 0$ is small enough, the map has at least one weakly hyperbolic invariant torus.

*Keywords:* Quasiperiodic systems, degenerate fixed points, weakly hyperbolic invariant torus, KAM iteration.

## 1 Introduction and main results

The existence and persistence of invariant manifolds are fundamental topics in nonlinear dynamical systems. Geometrically, invariant tori describe the quasiperiodic motions for

*T.Z. and J.S. have been supported by the National Natural Science Foundation of China (Grant Nos. 11171185, 10871117 ) and SDNSF (Grant No. ZR2010AM013). A.J. has been supported by the Spanish grant MTM2015-67724-P and the Catalan grant 2014 SGR 1145. The authors want to thank the anonymous referee for his detailed comments.

†Corresponding author. E-mail address: sijgmath@sdu.edu.cn or sijgmath@126.com

dynamical systems. Indeed, quasiperiodic forcing is not only a natural extension of periodic forcing, it also occurs in many physically relevant situations as there are many systems subject to external forcing depending on several frequencies. For instance, Harper map with quasiperiodic potential and the quasiperiodically forced Arnold circle map serve as models of quasiperiodic crystals respectively in [1] and [5].

In this paper, we consider a map of the form

$$\begin{aligned} (F,\omega): \qquad & \mathbb{R}^2\times\mathbb{T}^d \to \mathbb{R}^2\times\mathbb{T}^d,\\ & (z,\theta)\longmapsto (F(z,\theta,\epsilon),\theta+\omega), \end{aligned}$$

where $\mathbb{T}^d=\mathbb{R}^d/\mathbb{Z}^d$, $F$ is analytic with respect to $z,\theta$ and $\epsilon$, and $\omega$ is a vector of rationally independent frequencies. In this case, $\mathbb{R}^2$ is called the fiber space, and $F$ the fiber map. This kind of skew product has been studied in many works [1, 3, 5, 6, 9, 11] (see also references therein). Many of these works focus on breakdown of invariant tori into strange nonchaotic attractors. In this paper, we study the persistence of invariant tori under perturbation in a degenerate situation: we assume that $F(0,\theta,0)=0$ ($\forall\theta\in\mathbb{T}^d$), that $D_zF(0,\theta,0)$ does not depend on $\theta$, and that $\mathrm{Spec}\,(D_zF(0,\theta,0))=\{1\}$. Degenerate systems appear commonly in Celestial Mechanics [12, 13, 4, 2], and degenerate volume preserving maps are also studied in [14, 19]. The setting considered in this paper is the following:

$$\begin{cases} \bar{x}=x+\Omega y^m+\epsilon f_1(x,y,\theta,\epsilon)+h_1(x,y,\theta,\epsilon),\\ \bar{y}=y+x^n+\epsilon f_2(x,y,\theta,\epsilon)+h_2(x,y,\theta,\epsilon),\\ \bar{\theta}=\theta+\omega, \end{cases}$$

where $\Omega>0$, $mn>1, n\geq m$ and $\omega=(\omega_1,\omega_2,\dots,\omega_d)\in\mathbb{R}^d$. Moreover, $f_i$ and $h_i$ are of the form

$$f_1(x,y,\theta,\epsilon)=\sum_{0\leq i+j\leq n} f_{1ij}(\theta,\epsilon)x^iy^j,\qquad h_1(x,y,\theta,\epsilon)=\sum_{i+j\geq n+1} h_{1ij}(\theta,\epsilon)x^iy^j, \tag{1.1}$$

$$f_2(x,y,\theta,\epsilon)=\sum_{0\leq i+j\leq n} f_{2ij}(\theta,\epsilon)x^iy^j,\qquad h_2(x,y,\theta,\epsilon)=\sum_{i+j\geq n+1} h_{2ij}(\theta,\epsilon)x^iy^j, \tag{1.2}$$

with $f(0,0,\theta,\epsilon)\neq 0$ and $h(0,0,\theta,\epsilon)=0$, where $f=(f_1,f_2)^T$ and $h=(h_1,h_2)^T$. Therefore, if $\epsilon=0$, $u(\theta)=0$ is a parabolic invariant torus. We say that $f$ are lower order terms, and $h$ higher order terms. Here the minimum order of $h_1$ is $n+1$ rather than $m+1$, since $x$ in the solution has larger size than $y$ has if $n>m$ and we expect an uniform size of the high order terms to make the results correct. Throughout this paper, and without loss of generality, we assume that $n\geq m$. If $n<m$, an analogous result can be obtained.

The value of $\Omega$ can be set equal to 1 by using suitable scaling factors in $x$ and $y$. Hence, for convenience, let's consider the following map:

$$\begin{cases} \bar{x}=x+y^m+\epsilon f_1(x,y,\theta,\epsilon)+h_1(x,y,\theta,\epsilon),\\ \bar{y}=y+x^n+\epsilon f_2(x,y,\theta,\epsilon)+h_2(x,y,\theta,\epsilon),\\ \bar{\theta}=\theta+\omega. \end{cases} \tag{1.3}$$

We focus on invariant tori with a prescribed vector of fixed frequencies $\omega$. Hence they can be seen as the response of the autonomous system (when $\epsilon=0$) to the effect of the quasiperiodic forcing (when $\epsilon > 0$). Analogous situations are discussed in [7, 8, 10, 15, 16, 17, 18] for the elliptic or weakly hyperbolic cases.

Moreover, the results of this paper can be applied to some specific class of degenerate differential equations by means of a suitable Poincaré section. A construction of a skew product transformation derived from a quasiperiodic vector field will be found in Appendix A. In particular, we consider the differential equation

$$\begin{cases} \dot{x} = y^m + \epsilon l_1(x, y, t) + q_1(x, y, t), \\ \dot{y} = x^n + \epsilon l_2(x, y, t) + q_2(x, y, t), \end{cases} \tag{1.4}$$

as an example (see details in Appendix A), where the time dependence is quasiperiodic. As a result, when $m = 1, n > 1$, the Poincaré map of (1.4) defined by a complete revolution of one of the angles of the quasiperiodic time-dependence has the following form

$$\begin{cases} \bar{x} = x + y + \Omega x^n + \epsilon f_1(x, y, \theta, \epsilon) + h_1(x, y, \theta, \epsilon), \\ \bar{y} = y + x^n + \epsilon f_2(x, y, \theta, \epsilon) + h_2(x, y, \theta, \epsilon), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{1.5}$$

and when $n \geq m > 1$, it has the form

$$\begin{cases} \bar{x} = x + y^m + \epsilon f_1(x, y, \theta, \epsilon) + h_1(x, y, \theta, \epsilon), \\ \bar{y} = y + x^n + \epsilon f_2(x, y, \theta, \epsilon) + h_2(x, y, \theta, \epsilon), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

where $f$ contains lower order terms and $h$ contains higher order terms. These degenerate maps are the topic of this paper.

The existence of a weakly hyperbolic quasiperiodic solution for the case $m \neq 1$ and $m$, $n$ both odd is mentioned as an open problem in Remark 4 of [16]. In this paper, we provide an answer for this problem. What we do is even more general because $m$ and $n$ can also be even. Hence, it is not just a generalization of results in [15, 16, 17]. Moreover, we stress that the ideas of the proof are totally different from the proofs in [15, 16, 17].

When $m = 1$ and $n$ is odd, a quasiperiodic solution of the differential equations (1.4) is equivalent to an invariant torus of the map (1.5). Taking $\Omega = 1$, we obtain the map

$$\begin{cases} \bar{x} = x + y + x^n + \epsilon f_1(x, y, \theta, \epsilon) + h_1(x, y, \theta, \epsilon), \\ \bar{y} = y + x^n + \epsilon f_2(x, y, \theta, \epsilon) + h_2(x, y, \theta, \epsilon), \\ \bar{\theta} = \theta + \omega. \end{cases} \tag{1.6}$$

Under suitable hypothesis, the results in this paper show that this map has also at least one hyperbolic invariant torus. Therefore, the cases studied in [15] and [16] by Xu are included here. Moreover, $m$ and $n$ are not necessarily odd as they are in [15] and [16].

As it is usual, we need a Diophantine condition to control the small divisors appearing during the KAM iterations.

**Definition 1.1.** *We say $\omega \in \mathbb{R}^d$ is a Diophantine vector of type $(c, \gamma)$ if only if*

$$|(k, \omega) - l| \geq \frac{c}{|k|^\gamma}, \quad \forall k \in \mathbb{Z}^d \backslash \{0\}, \quad l \in \mathbb{Z},$$

*where $|k| = |k_1| + \cdots + |k_d|$, $c > 0$ and $\gamma \geq d$. Let $DC(c, \gamma)$ be the set of the Diophantine vectors of type $(c, \gamma)$.*

The following theorems are the main results of this paper. In order to state the theorems simply and clearly, we denote by $[f]$ the average of $f(\theta)$ with respect to $\theta$.

**Theorem 1.2.** *Let $c > 0$, $\gamma \geq d$, $r > 0$, $\rho > 0$. Suppose that $\omega \in DC(c, \gamma)$, and that $h_i$, $f_i$ are real analytic functions in $x, y, \epsilon$ on an open set of the origin, analytic in $\theta$ on an open complex strip of the real line, and that they have the form (1.1) and (1.2) respectively, where $i = 1, 2$. Moreover, we assume*

$$[f_{100}] \begin{cases} < 0 & \text{if } m \text{ is even,} \\ \neq 0 & \text{if } m \text{ is odd,} \end{cases}$$

*and*

$$[f_{200}] \begin{cases} < 0 & \text{if } n \text{ is even,} \\ \neq 0 & \text{if } n \text{ is odd.} \end{cases}$$

*Then there exists a sufficiently small $\epsilon_0 > 0$, such that if $\epsilon < \epsilon_0$ then the map (1.3) has at least one weakly hyperbolic and analytic invariant torus.*

**Theorem 1.3.** *Let $c > 0$, $\gamma \geq d$, $r > 0$, $\rho > 0$. Suppose that $\omega \in DC(c, \gamma)$, and that $h_i$, $f_i$ are real analytic functions in $x, y, \epsilon$ on an open set of the origin, analytic in $\theta$ on an open complex strip of the real line, and that they have the form (1.1) and (1.2) respectively, where $i = 1, 2$. Moreover, we assume*

$$[f_{200}] \begin{cases} < 0 & \text{if } n \text{ is even,} \\ \neq 0 & \text{if } n \text{ is odd.} \end{cases}$$

*Then there exists a sufficiently small $\epsilon_0 > 0$, such that if $\epsilon < \epsilon_0$ then the map (1.6) has at least one weakly hyperbolic and analytic invariant torus.*

**Remark 1.4.** For concreteness, if $m$ and $n$ are both even, we get two weakly hyperbolic and analytic invariant tori, otherwise, we just get one.

**Remark 1.5.** In the nondegenerate case $m = n = 1$, $h_i$ and $f_i$ do not need to be analytic. It is enough if they are of class $C^k$ with $k \geq 1$. Then, the corresponding invariant torus is also of class $C^k$. The proof can be found in Appendix B.

**Remark 1.6.** The proof of Theorem 1.3 is analogous to the proof of Theorem 1.2. Hence, in this work we mainly discuss the proof of Theorem 1.2.

## 2 Sketch of the proof

To simplify the notation, we will not write the dependence of $f$ and $g$ on $\epsilon$. As we will see, this dependence does not have any impact on the proofs. Hence, we consider the map

$$\begin{cases} \bar{x} = x + y^m + \epsilon f_1(x, y, \theta) + h_1(x, y, \theta), \\ \bar{y} = y + x^n + \epsilon f_2(x, y, \theta) + h_2(x, y, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{2.1}$$

where $nm > 1$ and $n \geq m$. Moreover, $f_i$ and $h_i$ are as in (1.1) and (1.2) skipping $\epsilon$, with $f(0, 0, \theta) \neq 0$ and $h(0, 0, \theta) = 0$. If $\epsilon = 0$, the fiber map of (2.1) has a fixed point at the origin.

It is natural to consider the average map of the fiber map of (2.1)

$$\begin{cases} \bar{x} = x + y^m + \epsilon[f_1](x, y) + [h_1](x, y), \\ \bar{y} = y + x^n + \epsilon[f_2](x, y) + [h_2](x, y), \end{cases} \tag{2.2}$$

where $[f]$ denotes the average of $f$ with respect to $\theta$, which is

$$[f](x, y) = \int_{\mathbb{T}^d} f(x, y, \theta)\, d\theta.$$

If $\epsilon = 0$, the map (2.2) has a fixed point at the origin. However, if $\epsilon > 0$, the fixed point may split into several fixed points and we want to show that at least one of them is real. These fixed points are roots of the combined equations

$$\begin{cases} y^m + \epsilon[f_1](x, y) + [h_1](x, y) = 0, \\ x^n + \epsilon[f_2](x, y) + [h_2](x, y) = 0. \end{cases} \tag{2.3}$$

By rescaling the variables as follows

$$x = \epsilon^{\frac{1}{n}} \tilde{x}, \quad y = \epsilon^{\frac{1}{m}} \tilde{y}, \tag{2.4}$$

equation (2.3) becomes

$$\begin{cases} \tilde{y}^m + [f_{100}] + \tau \hat{f}_1(\tilde{x}, \tilde{y}) + \tau \hat{h}_1(\tilde{x}, \tilde{y}) = 0, \\ \tilde{x}^n + [f_{200}] + \tau \hat{f}_2(\tilde{x}, \tilde{y}) + \tau \hat{h}_2(\tilde{x}, \tilde{y}) = 0, \end{cases} \tag{2.5}$$

where $\tau = \epsilon^{\frac{1}{n}}$. Assume that $\tilde{F}(\tilde{z}, \tau) = 0$ denotes the combined equations (2.5) where $\tilde{z} = (\tilde{x}, \tilde{y})^T$. It is easy to see that $\tilde{F}$ is real analytic on $\tilde{z}$ and $C^1$ on $\tau$. In order to use the Implicit Function Theorem to show the existence of solutions for $\tau \neq 0$, we assume

$$[f_{100}] \begin{cases} < 0 & \text{if } m \text{ is even}, \\ \neq 0 & \text{if } m \text{ is odd}, \end{cases} \tag{2.6}$$

and

$$[f_{200}] \begin{cases} < 0 & \text{if } n \text{ is even}, \\ \neq 0 & \text{if } n \text{ is odd}. \end{cases} \tag{2.7}$$

Let us consider $\tilde{F}(\tilde{z}, 0) = 0$, which is

$$\begin{cases} \tilde{y}^m + [f_{100}] = 0, \\ \tilde{x}^n + [f_{200}] = 0. \end{cases} \tag{2.8}$$

Equations (2.8) have at least a nonzero real root by conditions (2.6) and (2.7). Here if $n$ is even we denote by $(-[f_{200}])^{\frac{1}{n}}$ the positive root of the second equation, and if $m$ is even, we also denote by $(-[f_{100}])^{\frac{1}{m}}$ the positive root of the first equation. In summary, if $m$ and $n$ are both odd we obtain the real root $((-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}})$, if $m$ is odd and $n$ is even we have the real solutions $(\pm(-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}})$, if $m$ is even and $n$ odd we have the real roots $((-[f_{200}])^{\frac{1}{n}}, \pm(-[f_{100}])^{\frac{1}{m}})$ and, finally, if $m$ and $n$ are both even we obtain four real roots, $(\pm(-[f_{200}])^{\frac{1}{n}}, \pm(-[f_{100}])^{\frac{1}{m}})$.

Denote a real root by $\tilde{z}_{00} = (\tilde{x}_{00}, \tilde{y}_{00})$. As the matrix

$$D_{\tilde{z}}\tilde{F}(\tilde{z}_{00}, 0) = \begin{pmatrix} 0 & m(\tilde{y}_{00})^{m-1} \\ n(\tilde{x}_{00})^{n-1} & 0 \end{pmatrix}$$

is regular, the Implicit Function Theorem ensures that, for each $\tau$ close enough to 0, there exist a value $\tilde{z}_0(\tau)$ such that $\tilde{F}(\tilde{z}_0(\tau), \tau) = 0$ and, moreover,

$$\tilde{z}_0(\tau) = (\tilde{x}_0(\tau), \tilde{y}_0(\tau)) = (\tilde{x}_{00}, \tilde{y}_{00}) + \mathcal{O}(\tau).$$

In this work we just consider the roots $\tilde{z}_{00} = (\tilde{x}_{00}, \tilde{y}_{00})$ which make the eigenvalues of the matrix $D_{\tilde{z}}\tilde{F}(\tilde{z}_{00}, 0)$ real. If $m$ and $n$ are both even, we get two appropriate roots $((-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}})$ and $(-(-[f_{200}])^{\frac{1}{n}}, -(-[f_{100}])^{\frac{1}{m}})$, otherwise, one root $((-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}})$. Hence, the conditions (2.6) and (2.7) guarantee that there is at least one root such that the eigenvalues of the matrix $D_{\tilde{z}}\tilde{F}(\tilde{z}_{00}, 0)$ are real. Without loss of generality, we take the root $((-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}})$ as an example. In this case, the corresponding solution of $\tilde{F}(\tilde{z}, \tau) = 0$ is

$$(\tilde{x}_0(\tau), \tilde{y}_0(\tau)) = ((-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}}) + \mathcal{O}(\tau).$$

Substituting this solution into the rescaling transformation (2.4), we have a corresponding real root of the combined equations (2.3)

$$z_0(\tau) = (x_0(\tau), y_0(\tau)) = (\tau\tilde{x}_0(\tau), \tau^{\frac{n}{m}}\tilde{y}_0(\tau)),$$

which is also a real fixed point of the averaged map (2.2).

By a translation

$$\Psi : \begin{cases} x \mapsto x + x_0(\tau), \\ y \mapsto y + y_0(\tau), \end{cases}$$

the map (2.1) becomes

$$\begin{cases} \bar{x} = x + (y + y_0)^m + \tau^n f_1(x + x_0, y + y_0, \theta) + h_1(x + x_0, y + y_0, \theta), \\ \bar{y} = y + (x + x_0)^n + \tau^n f_2(x + x_0, y + y_0, \theta) + h_2(x + x_0, y + y_0, \theta), \\ \bar{\theta} = \theta + \omega. \end{cases} \tag{2.9}$$

As $\tau$ is a small parameter, if $m < n$ the first part of the map (2.9) is more important than the second one. Hence, we rescale the variables as follows

$$x \mapsto \tau^{\beta}x, \quad y \mapsto \tau^{n\beta}y,$$

where $\beta = \frac{n-m}{2m(n-1)}$. Then,

$$\begin{cases} \bar{z} = (I + \delta^{n-1}(A_0 + C(\delta)) + \delta^n\breve{Q}(\theta, \delta))z + \delta^n\breve{g}(\theta, \delta) + \breve{h}(z, \theta, \delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{2.10}$$

being $\delta = \tau^{\frac{2mn-n-m}{2m(n-1)}}$ and

$$A_0 = \begin{pmatrix} 0 & m(-[f_{100}])^{\frac{m-1}{m}} \\ n(-[f_{200}])^{\frac{n-1}{n}} & 0 \end{pmatrix}.$$

Let $B_0$ be a regular matrix such that $B_0^{-1}A_0B_0 = D_0 = \mathrm{diag}(\lambda_1^0, \lambda_2^0)$. Making a change of variables $z \mapsto B_0 z$, the map (2.10) becomes

$$\begin{cases} \bar{z} = (I + \delta^{n-1}D(\delta) + \delta^n Q(\theta,\delta))z + \delta^n g(\theta,\delta) + h(z,\theta,\delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{2.11}$$

where $D(\delta) = D_0 + B_0^{-1}C(\delta)B_0$ and $h$ is of second order in $z$, and of order 0 in $\delta$.

We are going to apply a sequence of transformations such that the final fiber map has a fixed point at the origin. The main idea is based on a KAM iteration. Before starting a KAM iteration, we will simplify the map (2.11) so that this iteration can be carried out in an easier way. The main idea is to make the size of the independent term $g$ smaller. What we exactly do is to translate the variables by an approximation of the invariant torus. This approximation $u(\theta,\delta)$ is obtained from the linearization

$$\begin{cases} \bar{z} = (I + \delta^{n-1}D(\delta))z + \delta^n g(\theta,\delta), \\ \bar{\theta} = \theta + \omega. \end{cases}$$

If we look for the quasiperiodic solution $u$ of this system, we face the small divisors $|e^{2\pi\sqrt{-1}(k,\omega)} - 1 - \delta^{n-1}\lambda_i|$ which are larger than $\frac{c}{|k|^\gamma}$ if $\lambda_i$ is real and $\omega \in DC(c,\gamma)$. As we will see in Lemma 4.3, $u$ is of order $\delta^n$ when $[g] = 0$. After applying the transformation

$$\begin{cases} z = z^* + u, \\ \theta = \theta, \end{cases} \tag{2.12}$$

we obtain

$$\begin{cases} \bar{z}^* = (I + \delta^{n-1}D + \delta^n Q^1(\theta,\delta))z^* + \delta^{2n} g^*(\theta,\delta) + h^*(z^*,\theta,\delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{2.13}$$

where $Q^1 = Q + \frac{D_z h(u,\theta,\delta)}{\delta^n}$, $g^* = \frac{1}{\delta^n}Qu + \frac{1}{\delta^{2n}}h(u,\theta,\delta)$ and $h^* = h(z^*+u,\theta,\delta) - h(u,\theta,\delta) - D_z h(u,\theta,\delta)z^*$. To have a quadratically convergent scheme, we need to find a new change of variables making $\delta^n Q^1$ smaller. To this end, we rewrite the map (2.13) as

$$\begin{cases} \bar{z}^* = (I + \delta^{n-1}A^* + \delta^n Q^*(\theta,\delta))z^* + \delta^{2n} g^*(\theta,\delta) + h^*(z^*,\theta,\delta), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

where $A^* = D + \delta[Q^1]$, $Q^* = Q^1 - [Q^1]$. Making the change of variables

$$\begin{cases} z^* = (I + \delta^n P)z_+, \\ \theta = \theta, \end{cases} \tag{2.14}$$

where $P$ is a quasiperiodic matrix determined by the condition

$$\bar{P} + \delta^{n-1}\bar{P}A^* - P - \delta^{n-1}A^*P = Q^*,$$

the map (2.13) becomes

$$\begin{cases} \bar{z}_+ = (I + \delta^{n-1}A^* + \delta^{2n}\tilde{Q}(\theta,\delta))z_+ + \delta^{2n}\tilde{g}(\theta,\delta) + \tilde{h}(z_+,\theta,\delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{2.15}$$

where $\tilde{Q}(\theta,\delta) = (I+\delta^n\bar{P})^{-1}[\bar{P}^2(I+\delta^{n-1}A^*) + Q^*P - \bar{P}(Q^*+P+\delta^{n-1}A^*P)]$, $\tilde{g}(\theta,\delta) = (I+\delta^n\bar{P})^{-1}g^*$ and $\tilde{h} = (I+\delta^n\bar{P})^{-1}h^*((I+\delta^nP)z_+,\theta,\delta)$. Here we have the small divisors $|e^{2\pi(k,\omega)\sqrt{-1}} - 1 - \delta^{n-1}(\lambda_i - \lambda_j)|$ which are larger than $\frac{c}{|k|^\gamma}$ for $\omega \in DC(c,\gamma)$.

Now, for the linear part of (2.15),

$$\begin{cases} \bar{z}_+ = (I+\delta^{n-1}A^*)z_+ + \delta^{2n}\tilde{g}(\theta,\delta), \\ \bar{\theta} = \theta+\omega, \end{cases}$$

we have that, as $[\tilde{g}] \neq 0$, the invariant torus is not of order $\delta^{2n}$ (for details see Lemma 4.3). Then after two transformations (the first for making the independent term smaller, and the second for making the new term $\delta^{2n}\tilde{Q}$ smaller), we obtain

$$\begin{cases} \bar{z} = (I+\delta^{n-1}A_1^*(\delta) + \delta^{3n+1}Q_1^*(\theta,\delta))z + \delta^{3n+1}g_1^*(\theta,\delta) + h_1^*(z,\theta,\delta), \\ \bar{\theta} = \theta+\omega. \end{cases} \tag{2.16}$$

Here we use Lemma 4.2 to show that all the corresponding eigenvalues are real numbers and this allows to control the divisors $|e^{2\pi\sqrt{-1}(k,\omega)} - 1 - \delta^{n-1}\lambda_i|$ when $\lambda_i$ changes.

Replacing $\delta^{n-1}$ by $\eta$, the map (2.16) can be written as

$$\begin{cases} \bar{z} = (I+\eta A_1^*(\delta) + \eta^3\delta^4Q_1^*(\theta,\delta))z + \eta^3\delta^4g_1^*(\theta,\delta) + h_1^*(z,\theta,\delta), \\ \bar{\theta} = \theta+\omega. \end{cases} \tag{2.17}$$

Now, we split $\eta^3\delta^4g_1^*(\theta,\delta)$ into the two factors $\eta^2$ and $\eta\delta^4g_1^*(\theta,\delta)$. Here we use the part $\eta^2$ to deal with the $\eta$ appearing in $\eta A_1^*(\delta)$, and let the part $\eta\delta^4g_1^*(\theta,\delta)$ be a KAM iteration term which means that in each KAM step we make the size of this term smaller than it is in the above step. We also rewrite $\eta^3\delta^4Q_1^*(\theta,\delta)$ in the same way. Then the map (2.17) becomes

$$\begin{cases} \bar{z} = (I+\eta A_1(\eta) + \eta^2Q_1(\theta,\eta))z + \eta^2g_1(\theta,\eta) + h_1(z,\theta,\eta), \\ \bar{\theta} = \theta+\omega, \end{cases} \tag{2.18}$$

where $A_1(\eta) = A_1^*(\delta)$, $Q_1(\theta,\eta) = \eta\delta^4Q_1^*(\theta,\delta)$, $g_1(\theta,\eta) = \eta\delta^4g_1^*(\theta,\delta)$, $h_1(z,\theta,\eta) = h_1^*(z,\theta,\delta)$. Moreover, it is easy to see that $Q_1$ and $g_1$ are of order $\eta$. We are going to apply a sequence of transformations such that the fiber map of the final map has a fixed point at the origin. We take the map (2.18) as the initial map in KAM iteration.

In the $j$-th KAM step, we have a map of the form

$$\begin{cases} \bar{z} = (I+\eta A_j(\eta) + \eta^2Q_j(\theta,\eta))z + \eta^2g_j(\theta,\eta) + h_j(z,\theta,\eta), \\ \bar{\theta} = \theta+\omega, \end{cases} \tag{2.19}$$

where $\|Q_j\| \le \nu_j$, $\|g_j\| \le \nu_j$ ($\nu_1 = \eta$). By transformations of the type (2.12) and (2.14), the map (2.19) becomes

$$\begin{cases} \bar{z} = (I+\eta A_{j+1} + \eta^2Q_{j+1})z + \eta^2g_{j+1} + h_{j+1}, \\ \bar{\theta} = \theta+\omega, \end{cases}$$

where $\|Q_{j+1}\| \le \nu_{j+1}$, $\|g_{j+1}\| \le \nu_{j+1}$. Roughly speaking, we will have $\nu_{j+1} \simeq \nu_j^2$ (the exact iterative formula of $\nu_j$ can be found in section 6). In view of $\nu_1 = \eta$, $Q_{j+1}$ and $g_{j+1}$ are of order $\eta^{2^j}$. Then, the scheme will be convergent to a map

$$\begin{cases} \bar{z} = (I+\eta A_\infty)z + h_\infty(z,\theta), \\ \bar{\theta} = \theta+\omega, \end{cases}$$

which has a hyperbolic fixed point at the origin. Therefore, the original map has a hyperbolic invariant torus near the origin, more precisely, near the point $((-\epsilon[f_{200}])^{\frac{1}{n}}, (-\epsilon[f_{100}])^{\frac{1}{m}})$ which is the main part of a root of the averaged map (2.2). Moreover, the eigenvalues of $I + \eta A_\infty$ are of the form of $1 \pm \mathcal{O}(\eta)$, so we say that the invariant torus is weakly hyperbolic. Furthermore, when $m$ and $n$ are both even, we can get one more weakly hyperbolic invariant torus near the point $(-(-\epsilon[f_{200}])^{\frac{1}{n}}, -(-\epsilon[f_{100}])^{\frac{1}{m}})$ if we take the point $(-(-[f_{200}])^{\frac{1}{n}}, -(-[f_{100}])^{\frac{1}{m}})$ which is also an appropriate root of $\tilde{F}(\tilde{z}, 0) = 0$.

# 3 Notations

For $z \in \mathbb{R}^2$, $\|z\|$ denotes the sup norm of $z$ and, if $A$ is a matrix, $\|A\|$ is the corresponding sup norm. We denote the complex torus by $\mathbb{T}^d_{\mathbb{C}} = \mathbb{C}^d/\mathbb{T}^d$. For $f(z,\theta) : \mathbb{R}^2 \times \mathbb{T}^d \to \mathbb{R}$, $[f](z)$ denotes the average of $f$ with respect to $\theta$.

Let us define

$$B_\alpha(A_0) = \{A \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2) : \quad \|A - A_0\| < \alpha\},$$

$$D(r,\rho) = \{(z,\theta) \in \mathbb{R}^2 \times \mathbb{T}^d_{\mathbb{C}} : \quad \|z\| < r, \ |\Im(\theta_i)| < \rho \ \ i = 1, 2, \cdots, d\},$$

$$\mathbb{T}^d_\rho = \{\theta \in \mathbb{T}^d_{\mathbb{C}} : \ |\Im(\theta_i)| < \rho, \quad i = 1, 2, \cdots, d\}, \quad \Delta_{\delta_0} = (0, \delta_0],$$

being $\Im(\theta)$ the imaginary part of $\theta$.

Suppose $f(z,\theta) : \mathbb{R}^2 \times \mathbb{T}^d_{\mathbb{C}} \to \mathbb{C}$ is an analytic function defined on $D(r,\rho)$, with

$$\|f(z,\theta)\|_{r,\rho} = \sup_{(z,\theta)\in D(r,\rho)} |f(z,\theta)|.$$

We denote by $C^\omega_{r,\rho}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R})$ the set of analytic functions on $D(r,\rho)$ such that $f(\mathbb{R}^2 \times \mathbb{T}^d) \subset \mathbb{R}$ and by $C^k_{r,\rho}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R})$ the set of $C^k$ functions. If $f(z,\theta) : \mathbb{R}^2 \times \mathbb{T}^d_{\mathbb{C}} \to \mathbb{C}^2$ is analytic on $D(r,\rho)$, then

$$\|f(z,\theta)\|_{r,\rho} = \max\{\|f_1(z,\theta)\|_{r,\rho}, \|f_2(z,\theta)\|_{r,\rho}\},$$

and we define $C^\omega_{r,\rho}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ similarly to $C^\omega_{r,\rho}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R})$. If $f(z,\theta,\delta)$ is analytic with respect to $z$ and $\theta$ on $D(r,\rho)$ and continuous with respect to $\delta$ on $\Delta_{\delta_0}$, we define

$$\|f(z,\theta,\delta)\|_{r,\rho,\delta_0} = \sup_{\delta\in\Delta_{\delta_0}} \sup_{(z,\theta)\in D(r,\rho)} \|f(z,\theta,\delta)\|.$$

Given $k \in \mathbb{N}$, we denote by $C^k(\mathbb{T}^d_\rho, \mathbb{R}^2)$ the set of $C^k$ functions $u : \mathbb{T}^d_\rho \to \mathbb{C}^2$ such that $u(\mathbb{T}^d) \subset \mathbb{R}^2$, endowed with the norm

$$\|u\|_{C^k} = \sup_{|i|\le k} \sup_{\theta\in\mathbb{T}^d_\rho} \|D^i u(\theta)\|,$$

where $\|\cdot\|$ denotes sup norm of $D^i u(\theta)$. $C^\omega(\mathbb{T}^d_\rho, \mathbb{R}^2)$ is defined in the same way, with the sup norm

$$\|u\|_\rho = \sup_{\theta\in\mathbb{T}^d_\rho} \|u(\theta)\|.$$

## 4 Some technical lemmas

In order to prove the main theorem, we will first give some lemmas.

**Lemma 4.1.** *(Lemma 2.8 in [8]) Let $D_0$ be a diagonal matrix with 2 different nonzero eigenvalues $\lambda_1^0$, $\lambda_2^0$, and $\mu$, $\alpha$ be two positive values such that $|\lambda_1^0| > 2\mu$, $|\lambda_2^0| > 2\mu$, $|\lambda_1^0 - \lambda_2^0| > 2\mu$ and $\alpha < 2\mu/5$. Then if $\|A - D_0\| < \alpha$, the following conditions hold:*

1. *Spec$(A) = \{\lambda_1, \lambda_2\}$, and $|\lambda_1| > \mu$, $|\lambda_2| > \mu$, $|\lambda_1 - \lambda_2| > \mu$.*
2. *There exists a nonsingular matrix $B$ such that $B^{-1}AB = D = diag(\lambda_1, \lambda_2)$, where $\|B\| \le 2$ and $\|B^{-1}\| \le 2$.*

The proof of this Lemma can be found in [8].

**Lemma 4.2.** *Let $D$ be a diagonal matrix with 2 different nonzero real eigenvalues $\lambda_1$, $\lambda_2$. Then if $\|A\| < \frac{|\lambda_1 - \lambda_2|}{2}$, the eigenvalues of $D + A$ are real numbers.*

*Proof.* Assume $\lambda_2 > \lambda_1$ and

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Let $\hat{\lambda}_1$, $\hat{\lambda}_2$ be the eigenvalues of $D + A$, and let $U_1$, $U_2$ be the following disks

$$U_1 = \{u \in \mathbb{C} : |u - (\lambda_1 + a_{11})| < |a_{12}|\},$$

$$U_2 = \{u \in \mathbb{C} : |u - (\lambda_2 + a_{22})| < |a_{21}|\}.$$

By Gerschgorin theorem, if $U_1 \bigcap U_2 = \emptyset$, $\hat{\lambda}_1$ lies in one of the disks, and $\hat{\lambda}_2$ lies in the other disk. This implies that $\hat{\lambda}_1$, and $\hat{\lambda}_2$ must be real numbers.

In view of $\|A\| < \frac{|\lambda_1 - \lambda_2|}{2}$, we obtain

$$|a_{11}| + |a_{12}| + |a_{21}| + |a_{22}| < \lambda_2 - \lambda_1,$$

which implies

$$a_{11} + |a_{12}| + |a_{21}| - a_{22} < \lambda_2 - \lambda_1.$$

Then we have

$$\lambda_1 + a_{11} + |a_{12}| < \lambda_2 + a_{22} - |a_{21}|,$$

which implies $U_1 \bigcap U_2 = \emptyset$. Hence, the eigenvalues of $D + A$ are real numbers.

$\square$

**Lemma 4.3.** *Consider the following map*

$$\begin{cases} \bar{z} = (I + \delta^q A + \delta^l Q(\theta))z + \delta^l g(\theta) + h(z, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{4.1}$$

*where $l > q \ge 1$, $A \in B_\alpha(D_0), Q(\theta)$, $g(\theta) \in C^\omega(\mathbb{T}^d_\rho, \mathbb{R}^2)$, $h(z, \theta) \in C^\omega_{r,\rho}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ and $\omega \in DC(c, \gamma)$. Assume that the eigenvalues of $A$ are real and*

$$[Q] = 0, \quad h(0, \theta) = 0, \quad D_z h(0, \theta) = 0, \quad \|D_{zz} h(z, \theta)\|_{r,\rho} \le K.$$

*Let $0 < \rho^* < \rho$ such that $\sigma = \rho - \rho^* \le 1$. Then there exists a function $u(\theta) \in C^\omega(\mathbb{T}^d_{\rho^*}, \mathbb{R}^2)$ such that the transformation $(z, \theta) = H_1(z^*, \theta) = (z^* + u(\theta), \theta)$, $H_1 : D(r^*, \rho^*) \to D(r, \rho)$, conjugates the map (4.1) to*

$$\left\{ \begin{array}{l} \bar{z}^* = (I + \delta^q A^* + \delta^l Q^*(\theta))z^* + \delta^l g^*(\theta) + h^*(z^*, \theta), \\ \bar{\theta} = \theta + \omega, \end{array} \right. \tag{4.2}$$

*where $0 < r^* \le r - \|u\|_{\rho^*}$, $Q^*(\theta)$, $g^*(\theta) \in C^\omega(\mathbb{T}^d_{\rho^*}, \mathbb{R}^2)$, $h^* \in C^\omega_{r^*,\rho^*}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$, $Q^*$ has zero average and the following bounds hold:*

*1.*

$$\|u\|_{\rho^*} \le \left\{ \begin{array}{ll} \delta^{l-q} L_1 \|g\|_\rho \frac{1}{c\sigma^{d+\gamma}}, & \text{if } [g] \ne 0, \\ \delta^l L_1 \|g\|_\rho \frac{1}{c\sigma^{d+\gamma}}, & \text{if } [g] = 0, \end{array} \right.$$

*2.* $\|A^*\| \le \|A\| + \delta^l \|Q\|_\rho + K\|u\|_{\rho^*}$,

*3.* $\|Q^*\|_{\rho^*} \le 2\|Q\|_\rho + \frac{2}{\delta^l} K \|u\|_{\rho^*}$,

*4.* $\|g^*\|_{\rho^*} \le \|Q\|_\rho \|u\|_{\rho^*} + \frac{K}{2\delta^l} \|u\|^2_{\rho^*}$,

*5.* $h^*(0, \theta) = 0, \quad D_{z^*} h^*(0, \theta) = 0, \quad \|D_{z^* z^*} h^*\|_{r^*, \rho^*} \le K$.

*Proof.* Let $u(\theta)$ be the solution of the linear map

$$\left\{ \begin{array}{l} \bar{u} = (I + \delta^q A)u + \delta^l g(\theta), \\ \bar{\theta} = \theta + \omega. \end{array} \right. \tag{4.3}$$

First of all, let us prove that the following bound holds:

$$\|u\|_{\rho^*} \le \left\{ \begin{array}{ll} \delta^{l-q} L_1 \|g\|_\rho \frac{1}{c\sigma^{d+\gamma}}, & \text{if } [g] \ne 0, \\ \delta^l L_1 \|g\|_\rho \frac{1}{c\sigma^{d+\gamma}}, & \text{if } [g] = 0. \end{array} \right. \tag{4.4}$$

In view of $A \in B_\alpha(D_0)$, let $B$ be the matrix found in Lemma 4.1. Then the change of variables $(u, \theta) = (Bv, \theta)$ conjugates the map (4.3) to

$$\left\{ \begin{array}{l} \bar{v} = (I + \delta^q D)v + \delta^l f(\theta), \\ \bar{\theta} = \theta + \omega. \end{array} \right.$$

where $D$ is a diagonal matrix and $f(\theta) = B^{-1} g$. Expanding $f_i$ and $v_i$ as

$$f_i(\theta) = \sum_{k \in \mathbb{Z}^d} f_i^k e^{2\pi\sqrt{-1}(k,\theta)}, \quad v_i(\theta) = \sum_{k \in \mathbb{Z}^d} v_i^k e^{2\pi\sqrt{-1}(k,\theta)},$$

where $i = 1, 2$, we obtain that $v_i^k = \frac{\delta^l f_i^k}{e^{2\pi\sqrt{-1}(k,\omega)} - 1 - \delta^q \lambda_i}$. As $\lambda_i$ is a (real) eigenvalue of $A$ and $\omega \in DC(c, \gamma)$, if $k \ne 0$ we have

$$|e^{2\pi\sqrt{-1}(k,\omega)} - 1 - \delta^q \lambda_i| \ge |(k, \omega) - p| \ge \frac{c}{|k|^\gamma},$$

where $p$ is a suitable integer. Then we have the following bounds:

$$|v_i^k| \le \left\{ \begin{array}{ll} \frac{\delta^{l-q} \|f\|_\rho}{\mu}, & \text{if } k = 0, \\ \delta^l \|f\|_\rho \frac{|k|^\gamma}{c} e^{-\rho|k|}, & \text{if } k \ne 0, \end{array} \right.$$

when $[g] \neq 0$, $\mu$ is defined in Lemma 4.1, and

$$|v_i^k| \leq \begin{cases} 0, & \text{if } k = 0, \\ \delta^l \|f\|_\rho \frac{|k|^\gamma}{c} e^{-\rho|k|}, & \text{if } k \neq 0, \end{cases}$$

when $[g] = 0$.

When $[g] \neq 0$, we use Lemma 2.6 in [8] to have

$$\|v\|_{\rho^*} \leq \delta^{l-q} \|f\|_\rho \left[ \frac{1}{\mu} + \frac{\delta^q L}{c\sigma^{d+\gamma}} \right],$$

and when $[g] = 0$,

$$\|v\|_{\rho^*} \leq \delta^l \|f\|_\rho \frac{L}{c\sigma^{d+\gamma}},$$

where $L = \frac{20d\chi(d+\gamma)}{3}$ is a constant depending on $d$, $\gamma$, with $\chi(s) = \left(\frac{s-1}{e}\right)^{s-1}\sqrt{s-1}$. For $c \leq 1$ and $\sigma \leq 1$, we obtain

$$\|v\|_{\rho^*} \leq \begin{cases} \delta^{l-q} \left[ \frac{1}{\mu} + L \right] \|f\|_\rho \frac{1}{c\sigma^{d+\gamma}}, & \text{if } [g] \neq 0, \\ \delta^l L \|f\|_\rho \frac{1}{c\sigma^{d+\gamma}}, & \text{if } [g] = 0. \end{cases}$$

According to $\|f\|_\rho \leq \|B^{-1}\| \|g\|_\rho$, $\|u\|_{\rho-\sigma} \leq \|B\| \|v\|_{\rho-\sigma}$, the inequality (4.4) holds and $L_1 = 4\left[\frac{1}{\mu} + L\right]$.

Applying the translation $z = z^* + u(\theta)$, the fiber map of (4.1) becomes

$$\bar{z}^* = (I + \delta^q A + \delta^l Q_1(\theta))z^* + \delta^l g^*(\theta) + h^*(z^*, \theta),$$

where $Q_1 = Q + \frac{D_z h(u,\theta)}{\delta^l}$, $g^* = Qu + \frac{h(u,\theta)}{\delta^l}$ and $h^* = h(z^* + u, \theta) - h(u, \theta) - D_z h(u, \theta)z^*$. As $h$ is analytic on $D(r, \rho)$, $h(0, \theta) = 0$, $D_z h(0, \theta) = 0$, and $\|D_{zz} h(z, \theta)\|_{r,\rho} \leq K$, it is easy to see that $\|h(z, \theta)\|_{r,\rho} \leq \frac{K}{2}\|z\|^2$ and $\|D_z h(z, \theta)\|_{r,\rho} \leq K\|z\|$. Therefore,

$$\|Q_1\|_{\rho^*} \leq \|Q\|_\rho + \frac{1}{\delta^l} K \|u\|_{\rho^*},$$

and, similarly,

$$\|g^*\|_{\rho^*} \leq \|Q\|_\rho \|u\|_{\rho^*} + \frac{K}{2\delta^l} \|u\|_{\rho^*}^2,$$

$$\|D_{z^*z^*} h^*\|_{r^*,\rho^*} \leq \|D_{zz} h(z^* + u, \theta)\|_{r,\rho} \leq K.$$

As usual, we denote the average of $Q_1$ by $[Q_1]$. If $Q^* = Q_1 - [Q_1]$ and defining $A^* = A + \delta^{l-q}[Q_1]$, then we obtain the map (4.2). The bounds on $A^*$ and $Q^*$ follow from $\|[Q_1]\|_{\rho^*} \leq \|Q_1\|_{\rho-\sigma}$ and $\|Q^*\|_{\rho^*} \leq 2\|Q_1\|_{\rho^*}$. □

**Lemma 4.4.** *Consider the following map*

$$\begin{cases} \bar{z^*} = (I + \delta^q A^* + \delta^l Q^*(\theta))z^* + \delta^l g^*(\theta) + h^*(z^*, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{4.5}$$

*where $l > q \geq 1$, $A^* \in B_\alpha(D_0)$, $Q^*(\theta)$, $g^*(\theta) \in C^\omega(\mathbb{T}^d_{\rho^*}, \mathbb{R}^2)$, $h^* \in C^\omega_{r^*,\rho^*}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ and $\omega \in DC(c, \gamma)$. Assume that the eigenvalues of $A^*$ are real and*

$$[Q^*] = 0, \quad h^*(0, \theta) = 0, \quad D_{z^*} h^*(0, \theta) = 0, \quad \|D_{z^*z^*} h^*(z^*, \theta)\|_{r^*,\rho^*} \leq K.$$

*Let $0 < \rho_+ < \rho^*$ such that $\sigma = \rho^* - \rho_+ \le 1$. Then there exists a change of variables $H_2 : D(r_+, \rho_+) \to D(r^*, \rho^*)$, where $(z^*, \theta) = H_2(z_+, \theta) = ((I + \delta^l P(\theta))z_+, \theta)$, $\|P\|_{\rho_+} \le L_2 \|Q^*\|_{\rho^*} \frac{1}{c\sigma^{d+\gamma}}$, and $L_2 = \frac{320 d\chi(d+\gamma)}{3}$, such that it conjugates the map (4.5) to*

$$\begin{cases} \bar{z}_+ = (I + \delta^q A_+ + \delta^l Q_+(\theta))z_+ + \delta^l g_+(\theta) + h_+(z_+, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{4.6}$$

*where $P(\theta)$ is analytic on $\mathbb{T}^d_{\rho_+}$ and $Q_+$ has zero average. Moreover, the following bounds hold:*

1. $\|A_+\| \le \|A^*\| + \frac{2\delta^{2l-q}}{1-\delta^l \|P\|_{\rho_+}}[(1 + \delta^q \|A^*\|)\|P\|^2_{\rho_+} + \|Q^*\|_{\rho^*}\|P\|_{\rho_+}]$,
2. $\|Q_+\|_{\rho_+} \le \frac{4\delta^l}{1-\delta^l \|P\|_{\rho_+}}[(1 + \delta^q \|A^*\|)\|P\|^2_{\rho_+} + \|Q^*\|_{\rho^*}\|P\|_{\rho_+}]$,
3. $\|g_+\|_{\rho_+} \le \frac{1}{1-\delta^l \|P\|_{\rho_+}} \|g^*\|_{\rho^*}$,
4. $\|D_{z_+z_+} h_+\|_{r_+,\rho_+} \le K \frac{(I+\delta^l \|P\|_{\rho_+})^2}{2(1-\delta^l \|P\|_{\rho_+})}$,

*where $0 < (I + \delta^l P)r_+ \le r^*$ and $\delta$ is small enough such that $\delta^l L_2 \|Q^*\|_{\rho-\sigma} \frac{1}{c^2 \sigma^{d+2\gamma}} < 1$.*

*Proof.* Under the transformation $H_2$, the fiber map of (4.5) becomes

$$\begin{aligned} \bar{z}_+ \;\; = \;\; & (I + \delta^q A^*)z_+ + \delta^l (P + \delta^q A^* P + Q^* - \bar{P} - \delta^q \bar{P} A^*)z_+ \\ & + \delta^{2l}(I + \delta^l \bar{P})^{-1}[\bar{P}^2(I + \delta^q A^*) + Q^* P - \bar{P}(Q^* + P + \delta^q A^* P)]z_+ \\ & + \delta^l (I + \delta^l \bar{P})^{-1} g^* + (I + \delta^l \bar{P})^{-1} h^* \circ H_2(z_+, \theta), \end{aligned}$$

where we denote $\bar{P} = P(\theta + \omega)$, and it satisfies the homological equation

$$\bar{P} + \delta^q \bar{P} A^* - P - \delta^q A^* P = Q^*.$$

Therefore,

$$\begin{cases} \bar{z}_+ = (I + \delta^q A^* + \delta^l Q_2(\theta))z_+ + \delta^l g_2(\theta) + h_2(z_+, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{4.7}$$

where $Q_2(\theta) = \delta^l (I + \delta^l \bar{P})^{-1}[\bar{P}^2(I + \delta^q A^*) + Q^* P - \bar{P}(Q^* + P + \delta^q A^* P)]$, $g_2(\theta) = (I + \delta^l \bar{P})^{-1} g^*$ and $h_2 = (I + \delta^l \bar{P})^{-1} h^* \circ H_2(z_+, \theta)$. To estimate the bounds on these terms, let us solve the homological equation.

Let $B$ be the matrix found in Lemma 4.1, which satisfies $B^{-1} A^* B = D$. If $P = BSB^{-1}$, then the homological equation becomes $\bar{S} + \delta^q \bar{S} D - S - \delta^q DS = B^{-1} Q^* B =: R$, where $R$ has zero average. As $D$ is a $2 \times 2$ diagonal matrix with eigenvalues $\lambda_1$ and $\lambda_2$, which are real, the upper equation can be solved by dealing with 4 equations. Assume $S = (s_{ij})$, then the entries $s_{ij}$ satisfy

$$s^k_{ij} = \frac{r^k_{ij}}{e^{2\pi(k,\omega)\sqrt{-1}} - 1 - \delta^q(\lambda_i - \lambda_j)}, \quad k \in \mathbb{Z}^d \setminus \{0\},$$

where

$$|e^{2\pi(k,\omega)\sqrt{-1}} - 1 - \delta^q(\lambda_i - \lambda_j)| \ge \frac{c}{|k|^\gamma},$$

when $k \neq 0$. When $k = 0$, $s_{ij}^k = 0$ (recall that $Q^*$ has zero average). Then we have

$$\|S\|_{\rho_+} \leq \frac{20d\chi(d+\gamma)}{3} \frac{1}{c^2\sigma^{d+\gamma}} \|R\|_{\rho^*}.$$

Noticing that $\|P\|_{\rho_+} \leq \|B\| \|S\|_{\rho_+} \|B^{-1}\|$, and $\|R\|_{\rho^*} \leq \|B^{-1}\| \|Q^*\|_{\rho^*} \|B\|$, we obtain $\|P\|_{\rho_+} \leq L_2 \|Q^*\|_{\rho^*} \frac{1}{c\sigma^{d+\gamma}}$, with $L_2 = \frac{320d\chi(d+\gamma)}{3}$.

Now, we are going to estimate the functions in map (4.7). According to the bound of $\|P\|_{\rho_+}$ and if $\delta$ is small enough it is easy to see that $\delta^l \|P\|_{\rho_+} < 1$. Hence we have

$$\|(I + \delta^l P)^{-1}\|_{\rho_+} \leq \sum_{i=0}^{\infty} \delta^{il} \|P\|_{\rho_+}^i \leq \frac{1}{1 - \delta^l \|P\|_{\rho_+}}.$$

Therefore,

$$\|Q_2\|_{\rho_+} \leq \frac{2\delta^l}{1 - \delta^l \|P\|_{\rho_+}} [(1 + \delta^q \|A^*\|) \|P\|_{\rho_+}^2 + \|Q^*\|_{\rho^*} \|P\|_{\rho_+}],$$

$$\|g_2\|_{\rho_+} \leq \frac{1}{1 - \delta^l \|P\|_{\rho_+}} \|g^*\|_{\rho^*},$$

and

$$\|D_{z_+z_+} h_2\|_{r_+,\rho_+} \leq K \frac{(I + \delta^l \|P\|_{\rho_+})^2}{1 - \delta^l \|P\|_{\rho_+}},$$

if $(1 + \delta^l \|P\|_{\rho_+}) r_+ \leq r^*$.

Let $Q_+ = Q_2 - [Q_2]$. Then the initial map becomes (4.6), where $A_+ = A^* + \delta^{l-q}[Q_2]$, $g_+ = g_2$, $h_+ = h_2$ and, moreover,

$$\|A_+\| \leq \|A^*\| + \delta^{l-q} \|Q_2\|_{\rho_+} \leq \|A^*\| + \frac{2\delta^{2l-q}}{1 - \delta^l \|P\|_{\rho_+}} [(1 + \delta^q \|A^*\|) \|P\|_{\rho_+}^2 + \|Q^*\|_{\rho^*} \|P\|_{\rho_+}],$$

and

$$\|Q_+\|_{\rho_+} \leq \frac{4\delta^l}{1 - \delta^l \|P\|_{\rho_+}} [(1 + \delta^q \|A^*\|) \|P\|_{\rho_+}^2 + \|Q^*\|_{\rho^*} \|P\|_{\rho_+}].$$

□

# 5 KAM theorems

To prove Theorem 1.2, we reduce the original map (2.1) to a map of the form (see the details in Section 6)

$$\begin{cases} \bar{z} = (I + \delta^{n-1} A(\delta) + \delta^n Q(\theta, \delta)) z + \delta^n g(\theta, \delta) + h(z, \theta, \delta), \\ \bar{\theta} = \theta + \omega. \end{cases} \tag{5.1}$$

We are not going to start the KAM iteration from map (5.1). In the following theorem, we simplify (5.1) such that these iterations are easier to perform.

**Theorem 5.1.** *Assume that $0 < c < 1$, $\gamma \geq d$, $0 < r < 1$, $0 < \rho < 1$, $0 < \delta_0 \leq 1$ and $\omega \in DC(c, \lambda)$. Suppose that $A(\delta) = D_0 + \mathcal{O}(\delta)$ is $C^1$ with respect to $\delta$, and that $Q(\theta, \delta)$, $g(\theta, \delta)$ and $h(z, \theta, \delta)$ are continuous in $\delta$. We also assume that for a fixed $\delta \leq \delta_0$, $Q(\theta, \delta)$, $g(\theta, \delta) \in C^\omega(T^d_\rho, \mathbb{R}^2)$, $h(z, \theta, \delta) \in C^\omega_{r,\rho}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ and that the following conditions hold:*

$$[Q] = 0, \quad [g] = 0, \quad h(0, \theta, \delta) = 0, \quad D_z h(0, \theta, \delta) = 0, \quad \|D_{zz}h(z, \theta)\|_{r,\rho,\delta} \leq K.$$

*Then there exists a transformation $H \in C^\omega_{r_1,\rho_1}(\mathbb{T}^d \times \mathbb{R}^2, \mathbb{R}^2)$, continuous on $\delta < \delta_1$ where $\delta_1 \leq \delta_0$, such that this transformation conjugates the map*

$$\begin{cases} \bar{z} = (I + \delta^{n-1}A(\delta) + \delta^n Q(\theta, \delta))z + \delta^n g(\theta, \delta) + h(z, \theta, \delta), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

*to*

$$\begin{cases} \bar{z} = (I + \delta^{n-1}A_1^*(\delta) + \delta^{3n+1}Q_1^*(\theta, \delta))z + \delta^{3n+1}g_1^*(\theta, \delta) + h_1^*(z, \theta, \delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{5.2}$$

*where $Q_1^*$ has zero average, $Q_1^*(\theta, \delta)$, $g_1^*(\theta, \delta) \in C^\omega(T^d_{\rho_1}, \mathbb{R}^2)$, $h^*(z, \theta, \delta) \in C^\omega_{r_1,\rho_1}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ continuous on $\delta < \delta_1$, $r_1 = r/2$, $\rho_1 = \rho/2$ and $\|A_1^*\| \leq \|A\| + \mathcal{O}(\delta)$.*

*Proof.* We recall that $D_0 = \operatorname{diag}(\lambda_1^0, \lambda_2^0)$, where $\lambda_1^0 = -\lambda_2^0$. There exists a positive number $\mu$ such that $|\lambda_1^0| > 2\mu$, which implies $|\lambda_2^0| > 2\mu$ and $|\lambda_1^0 - \lambda_2^0| > 4\mu$. Let $\alpha$ be a value such that $0 < \alpha < \frac{2\mu}{5}$. As $A(\delta) = D_0 + \mathcal{O}(\delta)$, there exists a positive constant $0 < \delta^* \leq \delta_0$ such that $A \in B_\alpha(D_0)$. In view of $\alpha < \frac{2\mu}{5} < \frac{|\lambda_1^0 - \lambda_2^0|}{2}$, it is easy to see that, if $A \in B_\alpha(D_0)$, Lemma 4.2 implies that $A$ has real eigenvalues. Using Lemma 4.3 with $[g] = 0$, $r^* = r - \hat{r}$, $\rho^* = \rho - \sigma$, $\hat{r} = \frac{r}{8}$ and $\sigma = \frac{\rho}{8}$, we have a change of variables $(z, \theta) = H_1(z^*, \theta, \delta) = (z^* + u(\theta, \delta), \theta)$, where $H_1 \in C^\omega_{r-\hat{r},\rho-\sigma}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$, that conjugates the map (5.1) to

$$\begin{cases} \bar{z}^* = (I + \delta^{n-1}A^*(\delta) + \delta^n Q^*(\theta, \delta))z^* + \delta^n g^*(\theta, \delta) + h^*(z^*, \theta, \delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{5.3}$$

where $A^*$, $Q^*$, $g^*$ and $h^*$ have the properties obtained in Lemma 4.3. Here, $\delta < \delta^{**}$ where $\delta^{**}$ is a positive constant such that $\delta^{**} \leq \delta^*$ and $(\delta^{**})^l L_1 \|g\|_\rho \frac{1}{c\sigma^{d+\gamma}} \leq \hat{r}$.

From the estimate on $\|A^*\|$, it follows that there exists $\delta_+ \leq \delta^{**}$ such that, if $\delta < \delta_+$, $A^* \in B_\alpha(D_0)$ and $\delta^l L_2 \|Q^*\|_{\rho-\sigma} \frac{1}{c^2\sigma^{d+2\gamma}} \leq \hat{r}$. Then we are able to use Lemma 4.4 to obtain a new change of variables $(z^*, \theta) = H_2(z_+, \theta, \delta) = ((I + \delta^n P)z_+, \theta)$ that conjugates the map (5.3) to

$$\begin{cases} \bar{z}_+ = (I + \delta^{n-1}A_+(\delta) + \delta^n Q_+(\theta, \delta))z^* + \delta^n g_+(\theta, \delta) + h_+(z_+, \theta, \delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{5.4}$$

where $A_+$, $Q_+$, $g_+$ and $h_+$ have the properties obtained in Lemma 4.4 with $r_+ = r - 2\hat{r}, \rho_+ = \rho - 2\sigma$. Rewriting (5.4) by using $z$ in place of $z_+$, $Q = \frac{Q_+}{\delta^n}$, $g = \frac{g_+}{\delta^n}$ and $h = h_+$, we have

$$\begin{cases} \bar{z} = (I + \delta^{n-1}A(\delta) + \delta^{2n}Q(\theta, \delta))z^* + \delta^{2n}g(\theta, \delta) + h(z, \theta, \delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{5.5}$$

where we have the bound $\|A\| \leq \|D_0\| + \mathcal{O}(\delta)$. Moreover, $Q(\theta, \delta), g(\theta, \delta) \in C^\omega(\mathbb{T}^d_{\rho-2\sigma}, \mathbb{R}^2)$, $h(z, \theta, \delta) \in C^\omega_{r-2\hat{r},\rho-2\sigma}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ for $\delta < \delta_+$. Let $H_{12} = H_1 \circ H_2$. Then $H_{12} \in C^\omega_{r-2\hat{r},\rho-2\sigma}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ conjugates the map (5.1) to (5.5).

Similarly, if $\delta$ is small enough, $A \in B_\alpha(D_0)$. We apply Lemmas 4.3 and 4.4 again but with $[g] \neq 0$ which affects the size of the translation in Lemma 4.3. Then we obtain a change of variables $H_{34} \in C^\omega_{r_1,\rho_1}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ which conjugates the map (5.5) to (5.2) where $Q_1^*(\theta, \delta)$, $g_1^*(\theta,\delta) \in C^\omega(\mathbb{T}^d_{\rho_1}, \mathbb{R}^2)$, $h_1^*(z,\theta,\delta) \in C^\omega_{r_1,\rho_1}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ with $r_1 = r/2, \rho_1 = \rho/2, \delta < \delta_1$ and $\|A_1^*\| \leq \|D_0\| + \mathcal{O}(\delta)$.

Hence, there is a transformation $H = H_{12} \circ H_{34} \in C^\omega_{r_1,\rho_1}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$ that conjugates the map (5.1) to (5.2). $\square$

The KAM iteration will be applied to the following map,

$$\begin{cases} \bar{z} = (I + \eta A + \eta^2 Q(\theta))z + \eta^2 g(\theta) + h(z,\theta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{5.6}$$

where

$$Q(\theta), g(\theta) \in C^\omega(\mathbb{T}^d_\rho, \mathbb{R}^2), \quad h(z,\theta) \in C^\omega_{r,\rho}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2), \tag{5.7}$$

$$[Q] = 0, \quad \|Q\|_\rho \leq \nu, \quad \|g\|_\rho \leq \nu, \tag{5.8}$$

$$h(0,\theta) = 0, \quad D_z h(0,\theta) = 0, \quad \|D_{zz} h(z,\theta)\|_{r,\rho} \leq K. \tag{5.9}$$

We denote by

$$r_+ = r - 2\hat{r}, \quad \rho_+ = \rho - 2\sigma,$$

$$E = \frac{K\nu}{c^2 \sigma^{2d+2\lambda}}, \quad \nu_+ = L_0 E^2,$$

$$K_+ = (1 + 2\hat{r})^3 K.$$

Then, the iterative theorem is as follows.

**Theorem 5.2.** *Assume that $0 < r,\ \rho,\ c \leq 1$, $\gamma \geq d$, $0 < 2\hat{r} < r$, $0 < 2\sigma < \rho$, $K \geq 1$ and $\nu$ is small enough. Consider the map (5.6) satisfying (5.7)-(5.9), with $\|A - D_0\| < \alpha - \eta\nu^2 - \eta L_1 E$, $\omega \in DC(c,\gamma)$, and $\eta L_3 E \leq \hat{r}$. Then, there exists a transformation $\Phi \in C^\omega_{r_+,\rho_+}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$, real analytic in $D(r_+, \rho_+)$, such that it conjugates the map (5.6) to $(5.6)_+$ which satisfies the inductive properties $(5.7)_+$-$(5.9)_+$ with parameters $\nu_+$, $r_+$, $\rho_+$, $K_+$ and $\|A_+\| \leq \|A\| + \eta^2\nu + \eta L_1 E + \eta^2 \nu_+$, where*

$$\Phi(\cdot,\cdot) : (z_+, \theta) \in D(r_+, \rho_+) \longmapsto ((I + \eta^2 P(\theta))z_+ + u(\theta), \theta) \in D(r,\rho),$$

*being $\eta^2 \|P\|_{\rho_+} \leq \eta L_3 E$ and $\|u\|_{\rho_+} \leq \eta L_3 E$.*

*Proof.* As $\|A - D_0\| < \alpha - \eta\nu^2 - \eta L_1 E$, Lemmas 4.3 and 4.4 imply that the change of variables $\Phi = H_1 \circ H_2(z_+, \theta)$ conjugates (5.6) to $(5.6)_+$. Now let us compute the corresponding bounds.

As $\Phi(z_+,\theta) = ((I + \eta^2 P)z_+ + u(\theta), \theta)$, with

$$\|u\|_{\rho-\sigma} \leq \eta L_1 \frac{\nu}{c\sigma^{d+\gamma}}$$

and

$$\eta^2 \|P\|_{\rho_+} \leq \eta^2 L_2 \|Q^*\|_{\rho-\sigma} \frac{1}{c\sigma^{d+\gamma}} \leq 2\eta^2 L_2 \frac{\nu}{c\sigma^{d+\gamma}} + 2\eta L_1 L_2 \frac{K\nu}{c^2\sigma^{2d+2\gamma}} \leq \eta L_3 E,$$

with $L_3 = 2(L_2 + L_1 L_2)$, we obtain

$$\Phi: \ D(r_+, \rho_+) \to D(r - 2\hat{r}^2, \rho_+) \subset D(r, \rho),$$

if $\eta L_3 E \leq \hat{r}$. It is easy to see that $\eta^2 \|P\|_{\rho_+} \leq \frac{1}{2}$. Then,

$$\|D_{z_+ z_+} h_2\|_{r_+, \rho_+} \leq K \frac{(I + \eta^2 \|P\|_{\rho_+})^2}{1 - \eta^2 \|P\|_{\rho_+}} \leq K(I + 2\eta^2 \|P\|_{\rho_+})^3 \leq (1 + 2\hat{r})^3 K = K_+.$$

If $\eta$ is small enough such that $\eta\|A\| \leq \frac{1}{2}$, then

$$\begin{aligned} \|A_+\| &\leq \|A^*\| + \frac{2\eta^3}{1 - \eta^2 \|P\|_{\rho_+}} [(1 + \eta\|A^*\|)\|P\|_{\rho_+}^2 + \|Q^*\|_{\rho-\sigma} \|P\|_{\rho_+}] \\ &\leq \|A\| + \eta^2 \nu + \eta L_1 E + \eta^2 L_4 E^2, \end{aligned}$$

$$\|Q_+\|_{\rho-\sigma} \leq \frac{4\eta^2}{1 - \eta^2 \|P\|_{\rho_+}} [(1 + \eta\|A^*\|)\|P\|_{\rho_+}^2 + \|Q^*\|_{\rho-\sigma} \|P\|_{\rho_+}] \leq 2\eta L_4 E^2,$$

$$\|g_+\|_{\rho-\sigma} \leq \frac{1}{1 - \eta^2 \|P\|_{\rho_+}} \|g^*\|_{\rho-\sigma} \leq (L_1^2 + \eta L_1) E^2 \leq L_4 E^2,$$

where $L_4 = 24 L_3^2$. If $L_0$ is large enough, such that $2L_4 E^2 \leq L_0 E^2 = \nu_+$, then we obtain that

$$\|A_+\| \leq \|A\| + \eta^2 \nu + \eta L_1 E + \eta^2 \nu_+,$$

$$\|Q_+\|_{\rho_+} \leq \nu_+,$$

$$\|g_+\|_{\rho_+} \leq \nu_+.$$

$\square$

# 6 Proof of the Main Theorem (Theorem 1.2)

Considering the map (1.3) which has form of

$$\begin{cases} \bar{x} = x + y^m + \epsilon f_1(x, y, \theta) + h_1(x, y, \theta), \\ \bar{y} = y + x^n + \epsilon f_2(x, y, \theta) + h_2(x, y, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

when $1 \leq m \leq n$, $n > 1$. Let us start by considering the following equations

$$\begin{cases} y^m + \epsilon [f_1](x, y) + [h_1](x, y) = 0, \\ x^n + \epsilon [f_2](x, y) + [h_2](x, y) = 0, \end{cases} \tag{6.1}$$

where $[\cdot]$ denotes the average with respect to $\theta$. By rescaling the variables as follows,

$$x = \epsilon^{\frac{1}{n}} \tilde{x}, \quad y = \epsilon^{\frac{1}{m}} \tilde{y},$$

equation (6.1) becomes

$$\begin{cases} \tilde{y}^m + [f_{100}] + \tau \hat{f}_1(\tilde{x}, \tilde{y}) + \tau \hat{h}_1(\tilde{x}, \tilde{y}) = 0, \\ \tilde{x}^n + [f_{200}] + \tau \hat{f}_2(\tilde{x}, \tilde{y}) + \tau \hat{h}_2(\tilde{x}, \tilde{y}) = 0, \end{cases} \tag{6.2}$$

where $\tau = \epsilon^{\frac{1}{n}}$,

$$\hat{f}_1 = \sum_{1 \le i+j \le n} [f_{1ij}] \tau^{i+\frac{n}{m}j-1} \tilde{x}^i \tilde{y}^j, \quad \hat{h}_1 = \sum_{i+j \ge n+1} [h_{1ij}] \tau^{i+\frac{n}{m}j-n-1} \tilde{x}^i \tilde{y}^j,$$

and

$$\hat{f}_2 = \sum_{1 \le i+j \le n} [f_{2ij}] \tau^{i+\frac{n}{m}j-1} \tilde{x}^i \tilde{y}^j, \quad \hat{h}_2 = \sum_{i+j \ge n+1} [h_{2ij}] \tau^{i+\frac{n}{m}j-n-1} \tilde{x}^i \tilde{y}^j.$$

Assume that $\tilde{F}(\tilde{z}, \tau) = 0$ denotes the combined equations (6.2) where $\tilde{z} = (\tilde{x}, \tilde{y})^T$. It is clear that $\tilde{F}$ is real analytic on $\tilde{z}$ and $C^1$ on $\tau$. If

$$[f_{100}] \begin{cases} < 0 & \text{if } m \text{ is even}, \\ \neq 0 & \text{if } m \text{ is odd}, \end{cases}$$

and

$$[f_{200}] \begin{cases} < 0 & \text{if } n \text{ is even}, \\ \neq 0 & \text{if } n \text{ is odd}, \end{cases}$$

then, for the equation $\tilde{F}(\tilde{z}, 0) = 0$, we obtain one real root if $m$ and $n$ are both odd, 2 real roots if one of $m, n$ is even, or 4 real roots otherwise. From now on, we consider the root $((-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}})$ as an example. Here if $n$ is even, we denote by $(-[f_{200}])^{\frac{1}{n}}$ the positive root, and if $m$ is even, $(-[f_{100}])^{\frac{1}{m}}$ denotes the positive root. As the Jacobian matrix $D_{\tilde{z}}\tilde{F}((-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}}, 0)$ is invertible, the Implicit Function Theorem implies that there exists a constant $\tau_0$, and a function $\tilde{z}_0$,

$$\tilde{z}_0(\tau) = ((-[f_{200}])^{\frac{1}{n}}, (-[f_{100}])^{\frac{1}{m}}) + \mathcal{O}(\tau) =: (\tilde{x}_0(\tau), \tilde{y}_0(\tau)),$$

such that if $|\tau| < \tau_0$, it satisfies $F(\tilde{z}_0(\tau), \tau) = 0$. Moreover, $\tilde{z}_0$ is a $C^1$ function of $\tau < \tau_0$. Let $x_0(\tau) = \tau \tilde{x}_0(\tau)$, $y_0(\tau) = \tau^{\frac{n}{m}} \tilde{y}_0(\tau)$, then $(x_0(\tau), y_0(\tau))$ is the solution of equation (6.1) when $|\tau| < \tau_0$.

Making the change of the variables

$$\Psi : \begin{cases} x \mapsto x + x_0(\tau), \\ y \mapsto y + y_0(\tau), \\ \theta \mapsto \theta, \end{cases}$$

on map (1.3) we obtain

$$\begin{cases} \bar{x} = x + (y + y_0)^m + \tau^n f_1(x + x_0, y + y_0, \theta) + h_1(x + x_0, y + y_0, \theta), \\ \bar{y} = y + (x + x_0)^n + \tau^n f_2(x + x_0, y + y_0, \theta) + h_2(x + x_0, y + y_0, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

which can be written as

$$\begin{cases} \bar{x} = x + P_1(x + x_0, y + y_0, \theta), \\ \bar{y} = y + P_2(x + x_0, y + y_0, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

where $P_1(x, y, \theta) = y^m + \tau^n f_1(x, y, \theta) + h_1(x, y, \theta)$ and $P_2(x, y, \theta) = x^n + \tau^n f_2(x, y, \theta) + h_2(x, y, \theta)$. Rescaling the variables as follows

$$x \mapsto \tau^\beta x, \quad y \mapsto \tau^{n\beta} y,$$

where $\beta = \frac{n-m}{2m(n-1)}$, we obtain

$$\begin{cases} \bar{x} = x + \tau^{-\beta}(\tau^{n\beta}y + y_0)^m + \tau^{n-\beta}f_1(\tau^\beta x + x_0, \tau^{n\beta}y + y_0, \theta) + \tau^{-\beta}h_1(\tau^\beta x + x_0, \tau^{n\beta}y + y_0, \theta), \\ \bar{y} = y + \tau^{-n\beta}(\tau^\beta x + x_0)^n + \tau^{n-n\beta}f_2(\tau^\beta x + x_0, \tau^{n\beta}y + y_0, \theta) + \tau^{-n\beta}h_2(\tau^\beta x + x_0, \tau^{n\beta}y + y_0, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

which is equivalent to

$$\begin{cases} \bar{z} = (I + \tau^{\frac{2mn-m-n}{2m}}A_0 + \tau^{\frac{n(2mn-m-n)}{2m(n-1)}}\hat{Q}(\theta,\tau))z + \tau^{\frac{n(2mn-m-n)}{2m(n-1)}}\hat{g}(\theta,\tau) + \hat{h}(z,\theta,\tau), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

where

$$A_0 = \begin{pmatrix} 0 & m(-[f_{100}])^{\frac{m-1}{m}} \\ n(-[f_{200}])^{\frac{n-1}{n}} & 0 \end{pmatrix},$$

$$\hat{Q} = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix},$$

with

$$\begin{aligned} q_{11} &= \tau^{n\beta}D_1f_1(x_0,y_0,\theta) + \tau^{n\beta-n}D_1h_1(x_0,y_0,\theta), \\ q_{12} &= \tau^{(2n-1)\beta}D_2f_1(x_0,y_0,\theta) + \tau^{(2n-1)\beta-n}D_2h_1(x_0,y_0,\theta), \\ q_{21} &= \tau^{\beta}D_1f_2(x_0,y_0,\theta) + \tau^{\beta-n}D_1h_2(x_0,y_0,\theta), \\ q_{22} &= \tau^{n\beta}D_2f_2(x_0,y_0,\theta) + \tau^{n\beta-n}D_2h_2(x_0,y_0,\theta), \end{aligned}$$

and

$$\hat{g} = \begin{pmatrix} \tau^{\frac{n-m}{2m}}f_1(x_0,y_0,\theta) + \tau^{\frac{n-m}{2m}-n}h_1(x_0,y_0,\theta) \\ f_2(x_0,y_0,\theta) + \tau^{-n}h_2(x_0,y_0,\theta) \end{pmatrix},$$

$$\hat{h}(z,\theta,\tau) \quad = \quad \begin{pmatrix} \tau^{-\beta_1}[P_1(\tau^\beta x + x_0, \tau^{n\beta}y + y_0, \theta) - P_1(x_0,y_0,\theta) - D_zP_1(x_0,y_0,\theta)z] \\ \tau^{-n\beta}[P_2(\tau^\beta x + x_0, \tau^{n\beta}y + y_0, \theta) - P_2(x_0,y_0,\theta) - D_zP_2(x_0,y_0,\theta)z] \end{pmatrix}.$$

Let us define $\delta = \tau^{\frac{2mn-n-m}{2m(n-1)}}$. Then,

$$\begin{cases} \bar{z} = (I + \delta^{n-1}(A_0 + C(\delta)) + \delta^n\breve{Q}(\theta,\delta))z + \delta^n\breve{g}(\theta,\delta) + \breve{h}(z,\theta,\delta), \\ \bar{\theta} = \theta + \omega, \end{cases} \tag{6.3}$$

where

$$C(\delta) = \delta[\hat{Q}](\tau), \quad \breve{Q}(\theta,\delta) = \hat{Q}(\theta,\tau) - [\hat{Q}](\tau), \quad \breve{g}(\theta,\delta) = \hat{g}(\theta,\tau), \quad \breve{h}(z,\theta,\delta) = \hat{h}(z,\theta,\tau).$$

It is clear that $C(\delta)$ is $C^1$ with respect to $\delta$, and that $\breve{Q}$, $\breve{g}$ and $\breve{h}$ are continuous in $\delta$. Moreover,

$$[\breve{Q}] = 0, \quad [\breve{g}] = 0, \quad \breve{h}(0,\theta,\delta) = 0, \quad D_z\breve{h}(0,\theta,\delta) = 0, \quad \|D_{zz}\breve{h}(z,\theta,\delta)\|_{r,\rho} \le K.$$

Let $B_0$ be a regular matrix such that $B_0^{-1}A_0B_0 = D_0 = \mathrm{diag}(\lambda_1^0, \lambda_2^0)$. Making a change of variables $z \mapsto B_0 z$, the map (6.3) becomes

$$\left\{ \begin{array}{l} \bar{z} = (I + \delta^{n-1}D(\delta) + \delta^n Q(\theta,\delta))z + \delta^n g(\theta,\delta) + h(z,\theta,\delta), \\ \bar{\theta} = \theta + \omega, \end{array} \right. \tag{6.4}$$

where $D(\delta) = D_0 + B_0^{-1}C(\delta)B_0 = D_0 + \mathcal{O}(\delta)$, $Q(\theta,\delta) = B_0^{-1}\breve{Q}(\theta,\delta)B_0$, $g(\theta,\delta) = B_0^{-1}\breve{g}(\theta,\delta)B_0$ and $h(\theta,\delta) = B_0^{-1}\breve{h}(B_0 z,\theta,\delta)$. In fact, after these transformations, the domain of $z$ changes. For the moment being, and to simplify the notation, we still use $r$ as the bound of $z$ (later we will account for the changes in this domain). If $\delta$ is small enough, $\delta \le \delta_0$, by Theorem 5.1 there exists a transformation that conjugates the map (6.4) to the map

$$\left\{ \begin{array}{l} \bar{z} = (I + \delta^{n-1}A_1^*(\delta) + \delta^{3n+1}Q_1^*(\theta,\delta))z + \delta^{3n+1}g_1^*(\theta,\delta) + h_1^*(z,\theta,\delta), \\ \bar{\theta} = \theta + \omega, \end{array} \right.$$

where $Q_1^*$ has zero average, $Q_1^*(\theta,\delta), g_1^*(\theta,\delta) \in C^\omega(T^d_{\rho_1}, \mathbb{R}^2)$, $h_1^*(z,\theta,\delta) \in C^\omega_{r_1,\rho_1}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$, with continuous dependence on $\delta < \delta_1$, and $\|A_1^*\| \le \|D\| + \mathcal{O}(\delta)$, where $r_1 = r/2, \rho_1 = \rho/2, \delta_1 \le \delta_0$. Let $\delta_2$ be a value such that $\delta_2 \le \delta_1$ and that, for any $\delta < \delta_2$, $A_1^* \in B_{\frac{\alpha}{2}}(D_0)$.

Let us denote $\delta^{n-1}$ as $\eta$, and let us take $\nu_1 = \eta$. Then,

$$\left\{ \begin{array}{l} \bar{z} = (I + \eta A_1 + \eta^2 Q_1(\theta))z + \eta^2 g_1(\theta) + h_1(z,\theta), \\ \bar{\theta} = \theta + \omega, \end{array} \right. \tag{6.5}$$

which satisfies $(5.7)_1$-$(5.9)_1$ with $A_1 \in B_{\frac{\alpha}{2}}(D_0)$.

Now we choose the inductive parameters to be able to iterate the KAM steps indefinitely. The choice is:

$$r_1 = r/2, \quad \hat{r}_1 = \frac{r_1}{8}, \quad \hat{r}_{j+1} = \frac{\hat{r}_j}{2}, \quad r_{j+1} = r_j - 2\hat{r}_j,$$

$$\rho_1 = \rho/2, \quad \sigma_1 = \frac{\rho_1}{8}, \quad \sigma_{j+1} = \frac{\sigma_j}{2}, \quad \rho_{j+1} = \rho_j - 2\sigma_j,$$

$$\nu_1 = \eta, \quad E_j = \frac{K_j \nu_j}{c^3 \sigma_j^{2d+3\gamma}}, \quad \nu_{j+1} = L_0 E_j^2,$$

$$a_1 = 0, \quad a_{j+1} = a_j + \eta^2 \nu_j + \eta L_1 E_j + \eta^2 \nu_{j+1},$$

$$K_1 = K, \quad K_{j+1} = (1 + 2\hat{r}_j)^3 K_j,$$

$$D_j = D(r_j, \rho_j).$$

It is easy to see that $K_j$ converges,

$$\lim_{j\to\infty} K_j = M > 0.$$

Moreover, we have

$$\frac{E_{j+1}}{E_j} = 2^{2d+3\gamma}(1 + 2\hat{r}_j)^3 L_0 \left(\frac{E_j}{E_{j-1}}\right)^2 \le L^* \left(\frac{E_j}{E_{j-1}}\right)^2 \le \left(L^* \frac{E_2}{E_1}\right)^{2^{j-1}},$$

implying that

$$E_{j+1} \leq \left(L^* \frac{E_2}{E_1}\right)^{2^j-1} E_1.$$

If $\nu_1$ is small enough such that $\frac{2^{2d+3\gamma} L^* L_1 K_1^2 (1+2\hat{r}_1)^3}{c^6 \sigma_1^{4d+6\gamma}} \nu_1 < \frac{1}{2}$, then

$$\sum_{j=1}^{\infty} E_j \leq 2E_1, \quad \sum_{j=1}^{\infty} \nu_j \leq 2E_1.$$

Now, we are going to show that, if $\nu_1$ is small enough such that

$$\frac{2(L_1+2)K_1}{c^3 \sigma_1^{2d+3\gamma}} \nu_1 < \min(\frac{\alpha}{2}, \frac{r_1}{8}),$$

then $\|A_j - D_0\| \leq \alpha - \eta\nu_j^2 - \eta L_1 E_j$ and $\eta L_3 E_j \leq \hat{r}_j$ for all $j \geq 1$. From the definition of $E_i$ and $\hat{r}_i$, it follows that $\eta L_3 E_i \leq \hat{r}_i$ for all $i \geq 1$ if $\frac{2(L_1+2)K_1}{c^3 \sigma_1^{2d+3\gamma}} \nu_1 < \frac{r_1}{8}$.

For $j = 1$, it is obvious as $A_1 \in B_{\frac{\alpha}{2}}(D_0)$. For $j > 1$, let us proceed by induction: assume that $\|A_i - D_0\| \leq \alpha - \eta\nu_i^2 - \eta L_1 E_i$ holds for any $1 \leq i < j$, and let us see that it holds for $j$.

According to the inductive assumption, using Theorem 5.2 $(j-1)$ times, we obtain that $\|A_j - D_0\| \leq \frac{\alpha}{2} + a_j$. As the sequence $\{K_i\}_i$ is convergent, if $\frac{2(L_1+2)K_1}{c^3 \sigma_1^{2d+3\gamma}} \nu_1 < \frac{\alpha}{2}$ we have that $\lim_{i\to\infty} a_i < \alpha/2$. As the sequence $\{a_i\}_i$ is increasing, we obtain

$$\|A_j - A_0\| \leq \frac{\alpha}{2} + a_j < \alpha - (\frac{\alpha}{2} - a_j) \leq \alpha - (a_{j+1} - a_j) < \alpha - \eta\nu_j^2 - \eta L_1 E_j.$$

Hence, for all $j \geq 1$, we have that $\|A_j - D_0\| \leq \alpha - \eta\nu_j^2 - \eta L_1 E_j$ and $\eta L_3 E_j \leq \hat{r}_j$ hold which allows to use Theorem 5.2 to find a sequence of changes of variables $\Phi_j$, which is real analytic, such that

$$\Phi_j : D_{j+1} \to D(r_j - 2\hat{r}_j^2, \rho_j) \subset D_j.$$

Moreover, $\Phi_j$ is of the form $z_j = (I + \eta^2 P_j(\theta)) z_{j+1} + u_j(\theta)$, $\theta = \theta$, where $\eta^2 \|P\|_{\rho_{j+1}} \leq \eta L_3 E_j$, $\|u\|_{\rho_{j+1}} \leq \eta L_3 E_j$, which means

$$\|\Pi_1 \circ (\Phi_j - id)\|_{r_{j+1},\rho_{j+1}} \leq 2\eta L_3 E_j, \quad \|\frac{\partial}{\partial z_{j+1}} \Pi_1 \circ (\Phi_j - id)\|_{r_{j+1},\rho_{j+1}} \leq \eta L_3 E_j.$$

Let $\Phi^j = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{j-1}$, where $\Phi_0 = id$. $\Phi^j$ is analytic on $D_j$ and transforms the initial map into

$$\begin{cases} \bar{z}_j = (I + \eta A_j + \eta^2 Q_j(\theta)) z_j + \eta^2 g_j(\theta) + h_j(z_j, \theta), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

where

$$Q_j(\theta), g_j(\theta) \in C^w(\mathbb{T}^d_{\rho_j}, \mathbb{R}^2), \quad h(z_j, \theta) \in C^w_{r_j,\rho_j}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2),$$
$$[Q_j] = 0, \quad \|Q_j\|_{\rho_j} \leq \nu_j, \quad \|g_j\|_{\rho_j} \leq \nu_j,$$

$$h_j(0,\theta)=0,\quad D_{z_j}h_j(0,\theta)=0,\quad \|D_{z_jz_j}h_j(z_j,\theta)\|_{r_j,\rho_j}\le K_j.$$

For $\lim_{j\to\infty} r_j=\frac{r_1}{2}$ and $\lim_{j\to\infty}\rho_j=\frac{\rho_1}{2}$, the domain $D(r_j,\rho_j)$ converges to the domain $D(\frac{r_1}{2},\frac{\rho_1}{2})$. In order to prove the convergence of the KAM iteration, we need to verify that all the related sequences are convergent with the norm $\|\cdot\|_{\frac{r_1}{2},\frac{\rho_1}{2}}$.

As

$$\|\Pi_1\circ(\Phi_j-id)\|_{r_{j+1},\rho_{j+1}}\le 2\eta L_3E_j,\quad \|\frac{\partial}{\partial z_{j+1}}\Pi_1\circ(\Phi_j-id)\|_{r_{j+1},\rho_{j+1}}\le \eta L_3E_j,$$

the derivative of $\Phi^j$ is bounded from above and below by two numbers, $L_5=\prod_{i=0}^{\infty}(1-\eta L_3E_i)$ and $L_6=\prod_{i=0}^{\infty}(1+\eta L_3E_i)$. This implies that the sequence $\{\Phi^j\}_j$ (where $\Phi^j$ is a diffeomorphism on $D_j$) converges uniformly on $D_\infty=D(\frac{r_1}{2},\frac{\rho_1}{2})$ to a diffeomorphism $\Phi^\infty$.

As $\lim_{j\to\infty}a_j<\alpha/2$, the sequence $A_j$ converges to $A_\infty$ and $\|A_\infty\|<\|D_0\|+\alpha$.

It is easy to see that $\|\eta^2Q_j\|_{\frac{\rho_1}{2}}\to 0$ and that $\|\eta^2g_j\|_{\frac{\rho_1}{2}}\to 0$. As the sequence $K_j$ converges, we have that $\|D_{z_jz_j}h_j\|_{\frac{r_1}{2},\frac{\rho_1}{2}}\le M$, which means that $\|h_i\|_{\frac{r_1}{2},\frac{\rho_1}{2}}\le\frac{M}{8}r_1^2$. Therefore, the sequence $\{h_j\}_j$ converges. Consequently, the transformation $\Phi^\infty$ conjugates the map (6.5) to

$$\begin{cases}\bar z_\infty=(I+\eta A_\infty)z_\infty+h_\infty(z_\infty,\theta),\\ \bar\theta=\theta+\omega,\end{cases}\tag{6.6}$$

where $h_\infty(0,\theta)=0$, $D_{z_\infty}h_\infty(0,\theta)=0$. As the map (6.6) has $z_\infty=0$ as a weakly hyperbolic fixed point, the initial map has a weakly hyperbolic invariant torus near the origin.

# Appendices

## A The Poincaré map of a degenerate differential equation

Here, we show how to obtain a skew product transformation from a quasiperiodic time-dependent vector field. The considered differential equation is

$$\begin{cases}\dot x=y^m+\epsilon l_1(x,y,t)+q_1(x,y,t)=\hat G_1(z,t,\epsilon),\\ \dot y=x^n+\epsilon l_2(x,y,t)+q_2(x,y,t)=\hat G_2(z,t,\epsilon),\end{cases}\tag{A.1}$$

where $1\le m\le n\ne 1$, $l$ are the lower order terms, $q$ the higher order terms, $z=(x,y)^T$ and $\hat G_i$ is an analytic quasiperiodic function with $d+1$ frequencies. Let us define $(\hat G_1(z,t,\epsilon),\hat G_2(z,t,\epsilon))^T=\hat G(z,t,\epsilon)$. Then we have $\hat G(z,t,\epsilon)=G(z,\theta_1,\theta_2,\ldots,\theta_{d+1},\epsilon)$ where $G$ is 1-periodic in each $\theta_i$. Moreover, $\dot\theta_i=\bar\omega_i$, and $\bar\omega=(\bar\omega_1,\bar\omega_2,\ldots,\bar\omega_{d+1})$ is a vector of rationally independent frequencies. It is natural to consider the lift of the differential equation to $\mathbb{R}^2\times\mathbb{T}^{d+1}$,

$$\begin{cases}\dot z=G(z,\hat\theta,\epsilon),\\ \dot{\hat\theta}=\bar\omega.\end{cases}$$

Let $\theta_{d+1} = 0$ be the Poincaré section and let us denote by $z(t; z_0, \theta_1^0, \dots, \theta_d^0, 0, \epsilon)$ the solution of the differential equation, where $z(0; z_0, \theta_1^0, \dots, \theta_d^0, 0, 0) = z_0$. Then the Poincaré map is defined on $\mathbb{R}^2 \times \mathbb{T}^d$,

$$\begin{cases} \bar{z}_0 = z(1/\bar{\omega}_{d+1}; z_0, \theta_1^0, \dots, \theta_d^0, 0, \epsilon), \\ \bar{\theta}_i^0 = \theta_i^0 + \omega_i, \quad i = 1, 2, \dots, d, \end{cases}$$

where $\omega_i = \bar{\omega}_i / \bar{\omega}_{d+1}$. Let us denote the Poincaré map as

$$\begin{cases} \bar{z} = P(z, \theta, \epsilon), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

with $z = (x, y)^T$ and $\omega = (\omega_1, \dots, \omega_d)$. As $P(z, \epsilon)$ is analytic with respect to $z$ and $\epsilon$, we can write $P(z, \epsilon) = P(z, 0) + \mathcal{O}(\epsilon)$ and $P = (P_1, P_2)^T$ where

$$P_i(z, 0) = \sum_{|\alpha| \geq 0} \frac{\partial^\alpha P_i(0, 0)}{\alpha!} z^\alpha,$$

where $i = 1, 2$ and $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^2$ is a two dimensional multi-index,

$$|\alpha| = |\alpha_1| + |\alpha_2|, \quad \alpha! = \alpha_1! \alpha_2!,$$

and we use the standard notation

$$z^\alpha = x^{\alpha_1} y^{\alpha_2}, \quad \partial^\alpha = \partial_x^{\alpha_1} \partial_y^{\alpha_2}.$$

Let us denote by $\varphi(t, z, \epsilon) = (\varphi_1(t, z, \epsilon), \varphi_2(t, z, \epsilon))^T$ the solution of (A.1) with initial condition $\varphi(0, z, \epsilon) = z$. As the Poincaré map is $P(z, \epsilon) = \varphi(1/\bar{\omega}_{d+1}, z, \epsilon)$, we have that $P(0, 0) = 0$. Moreover,

$$\begin{cases} \dot{\varphi}_1(t, z, \epsilon) = \hat{G}_1(\varphi(t, z, \epsilon), t, \epsilon), \\ \dot{\varphi}_2(t, z, \epsilon) = \hat{G}_2(\varphi(t, z, \epsilon), t, \epsilon). \end{cases} \tag{A.2}$$

Differentiating (A.2) with respect to $z$ and exchanging the derivations on $t$ and $z$, we obtain the first order variational equations,

$$\frac{d}{dt} \frac{\partial \varphi(t, z, 0)}{\partial z} = D_1 \hat{G}(\varphi(t, z, 0), t, 0) \frac{\partial \varphi(t, z, 0)}{\partial z}, \quad \frac{\partial \varphi(0, z, 0)}{\partial z} = I,$$

where $D_1 \hat{G}(z, t, \epsilon) = \frac{\partial \hat{G}(z, t, \epsilon)}{\partial z}$. Similarly, we can obtain higher order variational equations. Moreover, we have $\partial^\alpha \varphi(0, z, 0) = 0$ for $|\alpha| > 1$.

When $n > m = 1$, the solution of the first order variational equations is

$$\frac{\partial \varphi(t, z, 0)}{\partial z}|_{z=0} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

implying that

$$\begin{pmatrix} \partial^{(1,0)} P_1(0, 0) & \partial^{(0,1)} P_1(0, 0) \\ \partial^{(1,0)} P_2(0, 0) & \partial^{(0,1)} P_2(0, 0) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{\bar{\omega}_{d+1}} \\ 0 & 1 \end{pmatrix}.$$

For $1 < |\alpha| \le n$, $\alpha_1 \ne n$, as $\partial^\alpha \hat{G}_i(z,t,0)|_{z=0} = 0$, we have

$$\frac{d}{dt}\partial^\alpha \varphi(t,z,0)|_{z=0} = D_1\hat{G}(z,t,0)|_{z=0}\partial^\alpha\varphi(t,z,0)|_{z=0}, \quad \partial^\alpha\varphi(0,z,0) = 0,$$

which means that $\partial^\alpha\varphi(t,z,0)|_{z=0} = 0$ and this implies that $\partial^\alpha P(0,0) = 0$. For $\alpha = (n,0)$, $\partial^\alpha \hat{G}(z,t,0)|_{z=0} = (0, n!)^T$, we have

$$\frac{d}{dt}\partial^\alpha \varphi(t,z,0)|_{z=0} = D_1\hat{G}(z,t,0)|_{z=0}\partial^\alpha\varphi(t,z,0)|_{z=0} + (0, n!(\frac{\partial\varphi_1}{\partial x}|_{z=0})^n)^T, \quad \partial^\alpha\varphi(0,z,0) = 0.$$

For $\frac{\partial\varphi_1}{\partial x}|_{z=0} = 1$, we have $\partial^\alpha\varphi(t,z,0)|_{z=0} = (n!t^2/2, n!t)^T$, which means

$$\frac{\partial^\alpha P(0,0)}{\alpha!} = \left(\frac{1}{(\bar{\omega}_{d+1})^2}, \frac{1}{\bar{\omega}_{d+1}}\right)^T.$$

As a result, when $m = 1$, $n > 1$, the Poincaré map of the differential equation (A.1) has the following form

$$\begin{cases} \bar{x} = x + y + \frac{1}{(\bar{\omega}_{d+1})^2}x^n + \epsilon\breve{l}_1(x,y,\theta,\epsilon) + \breve{q}_1(x,y,\theta), \\ \bar{y} = y + \frac{1}{\bar{\omega}_{d+1}}x^n + \epsilon\breve{l}_2(x,y,\theta,\epsilon) + \breve{q}_2(x,y,\theta), \\ \bar{\theta} = \theta + \omega. \end{cases}$$

Using suitable scaling factors in $x$ and $y$ and dividing $\epsilon\breve{l}$ into lower order terms and high order terms, we obtain the map

$$\begin{cases} \bar{x} = x + y + \Omega x^n + \epsilon f_1(x,y,\theta,\epsilon) + h_1(x,y,\theta,\epsilon), \\ \bar{y} = y + x^n + \epsilon f_2(x,y,\theta,\epsilon) + h_2(x,y,\theta,\epsilon), \\ \bar{\theta} = \theta + \omega, \end{cases}$$

where $f$ are lower order terms and $h$ higher order terms.

When $n \ge m > 1$, which implies $D_1\hat{G}(z,t,0)|_{z=0} = 0$, the solution of the first order variational equations is $\frac{\partial\varphi(t,0,0)}{\partial z} = I$, and then,

$$\begin{pmatrix} \partial^{(1,0)}P_1 & \partial^{(0,1)}P_1 \\ \partial^{(1,0)}P_2 & \partial^{(0,1)}P_2 \end{pmatrix} = I.$$

For $1 < |\alpha| \le m, \alpha_1 \ne m, \alpha_2 \ne m$, as $\partial^\alpha\hat{G}_i(z,t,0)|_{z=0} = 0$ we have $\partial^\alpha P(0,0) = 0$. For $\alpha = (0,m)$, $\partial^\alpha\hat{G}(z,t,0)|_{z=0} = (m!, 0)^T$, we have

$$\frac{d}{dt}\partial^\alpha \varphi(t,z,0)|_{z=0} = D_1\hat{G}(z,t,0)|_{z=0}\partial^\alpha\varphi(t,z,0)|_{z=0} + (m!(\frac{\partial\varphi_2}{\partial y}|_{z=0})^n, 0)^T, \quad \partial^\alpha\varphi(0,z,0) = 0,$$

where $\frac{\partial\varphi_2}{\partial y}|_{z=0} = 1$, implying that $\partial^\alpha\varphi(t,z,0)|_{z=0} = (m!t, 0)^T$, which means

$$\frac{\partial^\alpha P(0,0)}{\alpha!} = \left(\frac{1}{\bar{\omega}_{d+1}}, 0\right)^T.$$

If $n = m$, for $\alpha = (m,0)$, we obtain

$$\frac{\partial^\alpha P(0,0)}{\alpha!} = \left(0, \frac{1}{\bar{\omega}_{d+1}}\right)^T.$$

If $n \neq m$, for $m < |\alpha| \leq n, \alpha_1 \neq n$, we have $\partial^\alpha P(0,0) = 0$, for $\alpha = (n, 0)$,

$$\frac{\partial^\alpha P(0,0)}{\alpha!} = \left(0, \frac{1}{\bar{\omega}_{d+1}}\right)^T.$$

As a result, we obtain the corresponding Poincaré map of (A.1) as

$$\left\{\begin{array}{l} \bar{x} = x + y^m + \epsilon f_1(x, y, \theta, \epsilon) + h_1(x, y, \theta, \epsilon), \\ \bar{y} = y + x^n + \epsilon f_2(x, y, \theta, \epsilon) + h_2(x, y, \theta, \epsilon), \\ \bar{\theta} = \theta + \omega. \end{array}\right.$$

under some scaling and rearrangement when $n \geq m > 1$.

# B Proof of Remark 1.4 (Nondegenerate Case)

When $m = n = 1$, we rewrite the map (1.3) using $z = (x, y)^T$, $f = (f_1, f_2)^T$, and $h = (h_1, h_2)^T$. We obtain

$$\left\{\begin{array}{l} \bar{z} = Az + \epsilon f(z, \theta) + h(z, \theta), \\ \bar{\theta} = \theta + \omega, \end{array}\right.$$

where $f, h \in C^k_{r,\rho}(\mathbb{R}^2 \times \mathbb{T}^d, \mathbb{R}^2)$, and

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

which can be diagonalized. Let $C$ be a regular matrix such that $C^{-1}AC = \text{diag}(0, 2) = D$.

Now, we can prove that there exists an invariant torus by using the Implicit Function Theorem. Let $X = C^k(\mathbb{T}^d_\rho, \mathbb{R}^2)$, $E = \{u \in X : \|u\|_{C^k} < r\}$. Obviously, $X$ is a Banach space with norm $\|\cdot\|_{C^k}$ and $E$ is an open subset of $X$. Let

$$F(u, \epsilon) = u(\theta + \omega) - Au(\theta) - \epsilon f(u, \theta) - h(u, \theta).$$

Then we get $F(0,0) = 0$, $F : E \times \mathbb{R} \to X$, and

$$\begin{aligned} D_u F(0,0) : \quad & X \to X \\ & \phi(\theta) \longmapsto \phi(\theta + \omega) - A\phi(\theta). \end{aligned}$$

Then,

1. $F \in C^k(E \times \mathbb{R}, X)$.

2. The linear map $D_u F(0,0)$ is injective: If $\phi(\theta + \omega) - A\phi(\theta) = 0$, let $\phi(\theta) = Cv(\theta)$. We have $v(\theta + \omega) - Dv(\theta) = 0$ implying $v(\theta) = 0$ and, finally, $\phi(\theta) = Cv(\theta) = 0$.

3. The linear map $D_u F(0,0)$ is exhaustive: If $\varphi(\theta) \in X$, we have to show that there exists an unique $\phi \in X$ such that $\phi(\theta + \omega) - A\phi(\theta) = \varphi(\theta)$. Let $\phi(\theta) = Cv(\theta)$ and $C^{-1}\varphi(\theta) = \hat{\varphi}(\theta)$. Let us consider the equation $v(\theta+\omega) - Dv(\theta) = \hat{\varphi}(\theta)$, which implies $v_1(\theta) = \hat{\varphi}_1(\theta - \omega)$ and
$$v_2(\theta + \omega) - 2v_2(\theta) = \hat{\varphi}_2(\theta). \tag{B.1}$$

Let us define

$$Tu = \frac{1}{2}u(\theta+\omega) - \frac{1}{2}\hat{\varphi}_2(\theta).$$

Obviously, $T : C^k(\mathbb{T}^d_\rho, \mathbb{R}) \to C^k(\mathbb{T}^d_\rho, \mathbb{R})$ is a contractive map: for any $u_1$, $u_2 \in C^k(\mathbb{T}^d_\rho, \mathbb{R})$, we have

$$\|Tu_1 - Tu_2\|_{C^k} \leq \frac{1}{2}\|u_1 - u_2\|_{C^k}.$$

As $C^k(\mathbb{T}^d_\rho, \mathbb{R})$ is a Banach space, we can use the Fixed Point Theorem to solve equation (B.1), and the solution $v_2$ is unique. As $\phi(\theta) = Cv(\theta)$, we obtain an unique $\phi \in X$ such that $\phi(\theta+\omega) - (I+A)\phi(\theta) = \varphi(\theta)$ for any $\varphi(\theta) \in X$.

Now we can use the Implicit Function Theorem to show that there exist neighbourhoods $I_0$ of $\epsilon = 0$, $V_0$ of $u = 0$ and an operator $U(\theta, \cdot) : I_0 \to V_0$, such that $F(U(\theta,\epsilon),\epsilon) = 0$, and $F(u,\epsilon) = 0$ if and only if $u = U(\theta,\epsilon)$, for all $(u,\epsilon) \in V_0 \times I_0$. Moreover, $U$ is necessarily $C^k$. Therefore, there exists a constant $\epsilon_0$ such that if $\epsilon < \epsilon_0$, the map (1.3) has a $C^k$ invariant torus which is hyperbolic with normal eigenvalues close to 0 and 2.

## References


[1] M. Ding, C. Grebogi, and E. Ott. Evolution of attractors in quasiperiodically forced systems: From quasiperiodic to strange non-chaotic to chaotic. *Phys. Rev. A*, 39(5):2277, 1989.

[2] I. Baldomà, E. Fontich and P. Martín. Stable manifolds for parabolic points through the parameterization method. *Preprint.*

[3] P. Glendinning. The non-smooth pitchork bifurcation. *Discrete Contin. Dyn. Syst. Ser. B*, 6(4):1–7, 2002.

[4] M. Guardia, P. Martín and T.M. Seara. Oscillatory motions for the restricted planar circular three body problem. *Invent. math.*, 1–76 2015.

[5] À. Haro and J. Puig. Strange nonchaotic attractors in Harper maps. *Chaos*, 16(3):033127, 12, 2006.

[6] À. Jorba, P. Rabassa, and J.C. Tatjer. Superstable periodic orbits of 1d maps under quasi-periodic forcing and reducibility loss. *Discrete Contin. Dyn. Syst. Ser. A*, 34(2):589–597, 2014.

[7] À. Jorba and C. Simó. On the reducibility of linear differential equations with quasiperiodic coefficients. *J. Differential Equations*, 98:111–124, 1992.

[8] À. Jorba and C. Simó. On quasiperiodic perturbations of elliptic equilibrium points. *SIAM J. Math. Anal.*, 27(6):1704–1737, 1996.

[9] À. Jorba and J.C. Tatjer. A mechanism for the fractalization of invariant curves in quasi-periodically forced 1-D maps. *Discrete Contin. Dyn. Syst. Ser. B*, 10(2-3):537–567, 2008.

[10] À. Jorba and J. Villanueva. On the persistence of lower dimensional invariant tori under quasi-periodic perturbations. *J. Nonlinear Sci.*, 7:427–473, 1997.

[11] L.M. Lerman. On remarks of skew products over irrational rotation. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 15(11):3675, 2005.

[12] R. Martínez and C. Pinyol. Parabolic orbits in the elliptic restricted three body problem. *J. Differential Equations*, 111(2):299–339, 1994.

[13] R. Martínez and C. Simó. Invariant manifolds at infinity of the RTBP and the boundaries of bounded motion. *Regul. Chaotic Dyn.*, 19(6):745–765, 2014.

[14] U. Vaidya and I. Mezić. Existence of invariant tori in three dimensional maps with degeneracy. *Phys. D*, 241(13):1136–1145, 2012.

[15] J. Xu. On small perturbation of two-dimensional quasi-periodic systems with hyperbolic-type degenerate equilibrium point. *J. Differential Equations*, 250(1):551–571, 2011.

[16] J. Xu. On quasi-periodic perturbations of hyperbolic-type degenerate equilibrium point of a class of planar system. *Discrete Contin. Dyn. Syst. Ser. A*, 33(6):2593–2619, 2013.

[17] J. Xu and S. Jiang. Reducibility for a class of nonlinear quasi-periodic differential equations with degenerate equilibrium point under small perturbation. *Ergodic Theory Dynam. Systems*, 31(2):599–611, 2011.

[18] J. Xu and Q. Zheng. On the reducibility of linear differential equations with quasiperiodic coefficients which are degenerate. *Proc. Amer. Math. Soc.*, 126(5):1445–1451, 1998.

[19] J. You. A KAM theorem for hyperbolic-type degenerate lower dimensional tori in hamiltonian systems. *Comm. Math. Phys.*, 192:145–168, 1998.