A non-perturbative theorem on conjugation of torus diffeomorphisms to rigid rotations

A. González-Enríquez *

Abstract

The problem of conjugation of torus diffeomorphisms to rigid rotations is considered here. Rather than assuming that the diffeomorphisms are close to rotations, we assume that the conjugacy equation has an approximate solution. First, it is proved that if the rotation vector is Diophantine and the invariance error function of the approximate solution has sufficiently small norm, then there exists a true solution nearby. The previous result is used to prove that if an element of a family of diffeomorphisms $\{f_\mu\}_\mu$ is conjugated to a rigid rotation with Diophantine rotation vector, then there exists a Cantor set $\mathcal{C}$ of parameters such that for each $\mu \in \mathcal{C}$ the diffeomorphism f_μ is conjugated to a Diophantine rigid rotation with rotation vector that depends on $\mu \in \mathcal{C}$ in a Whitney-smooth way.

Keywords: *conjugation of torus maps, KAM theory, non-perturbative conjugation, Whitney differentiability.*

1 Introduction

In this paper we consider the problem of conjugating diffeomorphisms of the d-dimensional torus $\mathbb{T}^d \stackrel{\text{def}}{=} (\mathbb{R}/\mathbb{Z})^d$ to rigid rotations. The case of diffeomorphisms that are perturbations of rotations have been widely, studied [Arn65, Her83, Mos66b, Zeh75]. However, the diffeomorphisms considered here are not assumed to be close to a rotation. Instead we assume that the conjugacy equation has an approximate solution. More precisely, it is well known [BHS96b, Chapter 2], [Mos66b, Zeh75] that to conjugate a given torus diffeomorphism to a rotation with fixed rotation vector it is necessary to add parameters to 'correct' the diffeomorphism. Hence the conjugacy equation one deals with is the following:

$$f \circ h = h \circ T_\omega + \lambda\,, \tag{1}$$

where f is the diffeomorphism of $\mathbb{T}^d$ to be conjugated to the rotation T_ω on $\mathbb{T}^d$, with ω Diophantine, h is the unknown conjugation and $\lambda \in \mathbb{T}^d$ is the unknown 'correction'. In our

*Dept. Mathematics, University of Texas at Austin 'Current address: Via Tiepolo 1, 62032 Camerino (MC), Italy'. email: alego@math.utexas.edu

setting, by an approximate solution of (1) we mean a couple (h_0, λ_0) with h_0 in a suitable function space and $\lambda_0 \in \mathbb{T}^d$, such that the error function

$$e_0 \stackrel{\text{def}}{=} f \circ h_0 - h_0 \circ T_\omega - \lambda_0 \tag{2}$$

is sufficiently small in an appropriate norm.

In Section 2 of [Mos66c] Moser describes an iterative method to construct solutions of non-linear functional equations satisfying certain conditions (see (2.1) in [Mos66c]), which we call *group structure*. Under the assumption of existence of an initial solution of the functional equation, the necessary condition to formally define the iterative method is the solvability of the linearised equation at the initial solution (see equation (2.6) in [Mos66c]). Moser's method is multiplicative in the sense that at each step the approximate solution is computed by using the composition operator (see equation (2.3) in [Mos66c]).

A variation of Moser's method is considered by Zehnder in Section 5 of [Zeh75], where the author explains how to construct an iterative method to solve conjugation problems that have a group structure (see (5.1) in [Zeh75]). The main assumption Zehnder requires is the invertibility of the linearised operator at solutions of the functional equation (see page 134 in [Zeh75]). Zehnder's method is additive in the sense that at each step of the iterative procedure, the new approximate solution is computed as the sum of the previous one and an increment, which is obtained by solving –approximately– the linearised equations.

The main idea of Moser's and Zehnder's methods is the following. If a solution of the functional equation is known and the linearised operator at the initial solution has a *right inverse*, then, using the group structure one obtains that the linearised operator at an approximate solution has an *approximate* right inverse (for a definition see [Ham82, Zeh75]). Thereby, the linearised equations are approximately solvable and a modified Newton method can be constructed to solve the non-linear problem.

Both Moser's and Zehnder's method were used to study conjugation problems in a perturbative setting. In particular, they assume that the identity map is a solution of the unperturbed problem. The main observation we make here is that in the case of conjugation of torus diffeomorphisms to rotations, it is not necessary to assume either that the diffeomorphism is a perturbation of a rigid rotation or the initial – approximate – solution is the identity. What one actually needs is the existence of an approximate solution of the conjugacy equation satisfying an appropriate non-degeneracy condition. We show that if a non-degenerate approximate solution of the conjugacy equation exists then an additive iterative procedure can be constructed to solve the non-linear problem.

It turns out that the linearised equation of the conjugacy equation at an (approximate) solution is *(approximately) reducible*. More precisely, the derivatives of an approximate solution provide a change of variables which takes the linearised equation into a difference equation with constant coefficients.

We study the conjugation equation (1) in the case that f is analytic. Under the assumption of the existence of an approximate solution (h_0, λ_0) of (1) – with h_0 analytic and satisfying a non-degenerate condition – we prove that if the error function e_0 in (2) is sufficiently small, where the bound is explicitly given in terms of the initial data $(f, \omega, h_0, \lambda_0)$, then there is a true solution of (1) which is close to the initial one (Theorem 1). Although we do not assume that the diffeomorphism f in (1) is close to a rotation, we need the linear part of the error e_0 to be equal to zero. This enables us to improve the nonlinear part of h_0 by solving a difference equation.

By using Moser's smoothing technique [Mos66c, Sal86, Zeh75] a finitely differentiable version of Theorem 1 is obtained in [GEV05], where both f and h_0 are assumed to be only finitely differentiable.

Remark 1. *Given a torus diffeomorphism f and a Diophantine rotation vector ω, notice that if (h, λ) satisfy equation (1) then for any $\theta_0 \in \mathbb{T}^d$ the couple $(h \circ T_{\theta_0}, \lambda)$ is also a solution of (1). We adopt the criterion that two solutions (h, λ) and $(\hat{h}, \lambda)$ are equivalent whenever $h(\theta) = \hat{h}(\theta + \theta_0)$ for some $\theta_0 \in \mathbb{T}^n$, since they only differ in the arbitrary choice of the origin of the phases. Here we prove that solutions of (1) are locally unique modulo the above equivalence criterion.*

Equation (1) is also studied in the case of a C^2-parametric family[1] $\{ f_\mu \}_\mu$ of analytic torus diffeomorphisms. We prove in Theorem 4 that if there is an element f_{μ_0} of the family which is conjugated to the rigid rotation T_{ω_0} with Diophantine rotation ω_0, then there is a Cantor set $\mathcal{C}$ of parameters in a sufficiently small neighbourhood of μ_0, such that for each $\mu \in \mathcal{C}$ the corresponding element of the family f_μ is also conjugated to a rigid rotation $T_{\omega(\mu)}$. Moreover we prove that the mapping $\omega(\mu)$ is Whitney-C^2 and also give an approximation of order 2, with respect to $(\mu - \mu_0)$, of $\omega(\mu)$ in a neighbourhood of μ_0.

In order to prove Theorem 4 we use the technique of *'borrowing parameters'*. The idea is similar to that of Moser [Mos66a, Mos67] also developed in [BHTB90, BHS96a], see also [BHS96b, Section 2.3]. It consists on adding parameters and let the frequencies vary on Diophantine vectors in a small neighbourhood of ω_0. Then, after solving the problem with the borrowed parameters, 'pay for' them. That is, find conditions for which the frequency can be obtained from the added parameters.

More precisely, consider the conjugacy equation (1) but replace f with f_μ, and let ω run over Diophantine vectors in a sufficiently small neighbourhood of ω_0. So that the considered conjugation problem is the following

$$f_\mu \circ h - h \circ T_\omega - \lambda = 0\,, \tag{3}$$

where (μ, ω) are known and close to (μ_0, ω_0), and (h, λ) are the unknowns.

Assuming that there is an approximate solution $(\mu_0, \omega_0, h_0, \lambda_0)$ of (3), with h_0 satisfying a non-degenerate condition, we prove – using Theorem 1 and local uniqueness – that there are functions $H(\mu, \omega)$ and $\Lambda(\mu, \omega)$, defined on a sufficiently small neighbourhood of (μ_0, ω_0), with μ in the set of parameters and ω Diophantine, satisfying the following properties:

1. $(\mu, \omega, h, \lambda) = (\mu, \omega, H(\mu, \omega), \Lambda(\mu, \omega))$ satisfies (3).

2. Functions $H(\mu, \omega)$ and $\Lambda(\mu, \omega)$ are Whitney-C^2.

[1] Definition 4 in Section 2.1.

Then Theorem 4 follows applying a version of the Whitney Extension Theorem[2] to Λ and the classical Implicit Function Theorem to the extended function $\hat{\Lambda}$ to obtain the frequency mapping $\omega(\mu)$ such that $\Lambda(\mu, \omega(\mu)) = 0$.

An outline of the paper follows. In Section 2 we state the main results. In Section 3 we describe, briefly and without technical details, the procedure used to prove the results of this work. Once proved that the linearised equations are approximately solvable, the proof of existence and local uniqueness of solutions of (1) is quite standard in KAM theory. So we do not include it here completely and only sketch the main steps in Sections 4.1 and 4.2. Section 4.3 contains a proof of the existence of the Cantor set $\mathcal{C}$ with the properties described above. In Section 5 we prove of the existence and Whitney differentiability of the functions $H(\mu, \omega)$ and $\Lambda(\mu, \omega)$.

2 Setup and statement of the results

2.1 Notation and definitions

Before stating the results of this work we introduce some notation and definitions. $C^0(\mathbb{T}^d)$ denotes the set of functions $u : \mathbb{R}^d \to \mathbb{R}^d$, which are $\mathbb{Z}^d$-periodic and such that

$$\| u\|_{C^0(\mathbb{T}^d)} \stackrel{\text{def}}{=} \sup_{x\in\mathbb{R}^d} | u(x)| < \infty \,,$$

where $|\cdot|$ represents the maximum norm on the spaces $\mathbb{R}^m$ and $\mathbb{C}^m$, *i.e.*

$$|x| \stackrel{\text{def}}{=} \max_{j=1,\ldots,m} |x_j| \qquad \text{for} \quad x = (x_1, \ldots, x_m) \in \mathbb{C}^m.$$

Similar notation for the norm is also used for real or complex matrices of arbitrary dimension.

Given $\rho \geq 0$ consider the complex strip

$$\mathbb{T}^d_\rho \stackrel{\text{def}}{=} \{z = x + iy \in \mathbb{C}^d : |y\,| \leq \rho\,\} \,.$$

Let $\mathcal{A}\left(\mathbb{T}^d_\rho,\, C^0\right)$ be the Banach space of functions $u : \mathbb{T}^d_\rho \to \mathbb{C}^d$ which are $\mathbb{Z}^d$-periodic, real analytic on the interior of $\mathbb{T}^d_\rho$, continuous on the boundary of $\mathbb{T}^d_\rho$, and such that

$$\|u\|_\rho \stackrel{\text{def}}{=} \sup_{\theta\in\mathbb{T}^d_\rho} |u(\theta)| < \infty \,.$$

For $r > 0$, $\mathcal{A}\left(\mathbb{T}^d_\rho,\, C^r\right)$ denotes the subset of $\mathcal{A}\left(\mathbb{T}^d_\rho,\, C^0\right)$ for which the following holds:

$$\|u\|_{\rho,\, C^r} \stackrel{\text{def}}{=} \sup_{|k|\leq r} \left\{ \left\|D^k u\right\|_\rho \right\} < \infty, \qquad \text{if} \quad r \in \mathbb{N},$$

[2]See for example Theorem 4 in Chapter VI in [Ste70] or Theorem 6.15 in [BHS96b].

and letting $r = p + \alpha$ with $p \in \mathbb{N}$ and $\alpha \in (0,1)$,

$$\|u\|_{\rho, C^r} \stackrel{\text{def}}{=} \sup_{|k|\le p} \left\{ \left\|D^k u\right\|_\rho \right\} + \sup_{\substack{x,y\in\mathbb{T}^d_\rho, x\neq y\\ |k|=p}} \left\{ \frac{|D^k f(x) - D^k f(y)|}{|x-y|^\alpha} \right\} < \infty\,,$$

where $\mathbb{N}$ denotes the set of natural numbers.

Given a continuous torus map $f : \mathbb{T}^d \to \mathbb{T}^d$, a *lift* of f to $\mathbb{R}^d$ (the universal cover of $\mathbb{T}^d$) is a continuous map $\hat{f} : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\pi \circ \hat{f} = f \circ \pi\,, \tag{4}$$

where π is the covering map

$$\pi : \mathbb{R}^d \to \mathbb{T}^d, \qquad \pi(x) = x \mod \mathbb{Z}^d. \tag{5}$$

Proposition 1. *Given a continuous torus map $f : \mathbb{T}^d \to \mathbb{T}^d$ any lift $\hat{f} : \mathbb{R}^d \to \mathbb{R}^d$ has the form*

$$\hat{f}(x) = A\,x + u(x) \tag{6}$$

where $A \in \mathcal{M}_{d\times d}(\mathbb{Z})$ and $u \in C^0\left(\mathbb{T}^d\right)$ (here $\mathcal{M}_{d\times d}(\mathbb{Z})$ is the set $d \times d$ integer valued matrices). Moreover, if f has additional regularity the corresponding periodic function u has the same regularity.

Proof. Let $\hat{f}$ satisfy (4), then for each $x \in \mathbb{R}^d$, and $k \in \mathbb{Z}^d$ there exists a vector $s \in \mathbb{Z}^d$ depending on k (for continuity s is independent of x) such that

$$\hat{f}(x+k) = \hat{f}(x) + s\,.$$

Then for each $i = 1, \ldots, d$, there exists a vector $A_i \in \mathbb{Z}^d$ such that

$$\hat{f}(x+e_i) = \hat{f}(x) + A_i$$

where e_i is the vector in $\mathbb{R}^d$ with zero in all its coordinates but the i-th which is equal to 1. Then for any $x \in \mathbb{R}^d$ and $k \in \mathbb{Z}^d$

$$\hat{f}(x+k) = \hat{f}(x) + A\,k\,,$$

where A is the matrix with columns A_i. Hence the function $u : \mathbb{R}^d \to \mathbb{R}^d$ defined by

$$u(x) \stackrel{\text{def}}{=} \hat{f}(x) - A\,x$$

is $\mathbb{Z}^d$-periodic and satisfies (6). □

Remark 2. *Let $f : \mathbb{T}^d \to \mathbb{T}^d$ be continuous, and let $\hat{f}$ and A be as in Proposition 1, then*

$$\hat{f}(x+k) = \hat{f}(x) + A\,k\,.$$

Moreover, if $\Gamma \in \mathbb{T}^d$ is a closed curve in $\mathbb{T}^d$ with winding numbers $k = (k_1, \ldots, k_d)$, then the transformed curve $f(\Gamma)$ has winding numbers given by Ak.

Remark 3. *Let $D^0(\mathbb{T}^d)$ denote the set of C^0-diffeomorphism on $\mathbb{R}^d$ that can be written in the form (6). Notice that any diffeomorphism $\hat{f} \in D^0(\mathbb{T}^d)$ defines a torus diffeomorphism f on $\mathbb{T}^d$ such that $f \circ \pi = \pi \circ \hat{f}$, with π as in (5). Even though lifts of continuous torus diffeomorphism are not unique, they differ from a constant vector in $\mathbb{Z}^d$. This enables us to work with lifts of torus maps. For notational simplicity we use the same letter to denote the torus map and a lift of it.*

Given a map $u \in C^0(\mathbb{T}^d)$ the *average* of it is defined by

$$\operatorname{avg}\{u\}_\theta \stackrel{\text{def}}{=} \int_{\mathbb{T}^d} u(x)dx\,.$$

The above notation is extended to matrix or vector valued-functions G with components $G_{i,j} \in C^0(\mathbb{T}^d)$ by integrating component-wise.

Definition 1. *Let $v \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$, and $B \in \mathcal{M}_{d\times d}(\mathbb{Z})$. We say that $h = B + v$ is non-degenerate if it satisfies the following conditions:*

1. *The matrix $Dh(\theta)$ is invertible for any $\theta \in \mathbb{T}^d_\rho$,*
2. *The matrix $\Phi \stackrel{\text{def}}{=} \operatorname{avg}\{\, Dh(\theta)^{-1}\,\}_\theta$ is invertible.*
3. *There exist two positive numbers $\eta = \eta(h)$ and $\tilde{\eta} = \tilde{\eta}(h)$ such that*

$$\left|\,\Phi^{-1}\,\right| \le \eta\,, \qquad \left\|\,[\,Dh(\theta)]^{-1}\,\right\|_\rho \le \tilde{\eta}\,.$$

Definition 2. *Given $\gamma > 0$ and $\sigma \ge d$, we define $D(\gamma, \sigma)$ as the set of frequency vectors $\omega \in \mathbb{R}^d$ satisfying the Diophantine condition:*

$$|\,k\cdot\omega - m| \ge \gamma\,|k|_1^{-\sigma} \qquad \forall k \in \mathbb{Z}^d \setminus \{0\},\, m \in \mathbb{Z},$$

where $|k|_1 = |k_1| + \cdots + |k_d|$.

Definition 3. *Let $\Omega \subset \mathbb{R}^n$ and $\ell > 0$ be such that $m < \ell \le m+1$, for $m \in \mathbb{Z}_+$ (the set of non-negative integer numbers). The function g defined on Ω is Whitney-C^ℓ if there exist functions $\{g^{(j)} : 0 \le |j|_1 \le m\}$ defined on Ω, and a constant M so that for all $|j|_1 \le m$*

$$\left|\,g^{(j)}(x)\right| \le M \qquad \forall x \in \Omega, \tag{7}$$

and if

$$R_j(x,y) \stackrel{\text{def}}{=} g^{(j)}(x) - \sum_{|i+j|_1 \le m} \frac{1}{i\,!}\, g^{(i+j)}(y)\,(x-y)^i\,,$$

where $x^j \stackrel{\text{def}}{=} x_1^{j_1} \ldots x_n^{j_n}$, then

$$|R_j(x,y)| \le M\;|x-y|^{\ell - |j|_1} \qquad \forall x, y \in \Omega\,. \tag{8}$$

Remark 4. *As it is well known, the Whitney derivatives are not unique.*[3] *For this reason if $\Omega \neq \mathbb{R}^n$ when talking about a Whitney-C^ℓ function g, we refer to a family of functions $\{g_j\}_{0\leq|j|_1\leq m}$ with $g_0 = g$, as in Definition 3.*

Definition 4. *A C^ℓ-parametric family of analytic functions on $\mathbb{T}^d$ with parameter $\mu \in \Xi \subset \mathbb{R}^s$, denoted by $\{u_\mu : \mu \in \Xi \subset \mathbb{R}^s\}$ (or $\{u_\mu\}_\mu$ for shortness) is a function u defined on $\Xi \times \mathbb{T}^d_\rho$, for some $\rho > 0$, such that for each $\mu \in \Xi$ the map $u_\mu \stackrel{\text{def}}{=} u(\mu, \cdot) \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$, and such that for each $\theta \in \mathbb{T}^d_\rho$, $u(\cdot, \theta)$ is C^ℓ on Ξ.*

We denote by $\partial_1 u(\mu, \theta)$ and $\partial_2 u(\mu, \theta)$, respectively, the partial derivatives of u with respect to μ and with respect to θ.

Given two Banach spaces B_1 and B_2 we denote by $L(B_1, B_2)$ the space of linear transformations defined on B_1 and with range in B_2.

2.2 Statement of the results

In this section we state the results of this work. Briefly, Theorem 1 ensures that if there is an approximate solution (h_0, λ_0) of the conjugation problem (1) with h_0 non-degenerate, and if the norm of the error function is sufficiently small, then there is a true solution which is close to the initial one. Theorem 2 deals with the uniqueness of the solutions of (1).

In Theorems 3 and 4 we consider the conjugation problem (3) for a C^2-parametric family of diffeomorphisms $\{f_\mu\}_{\mu\in\Xi}$. Theorem 3 says that if there is a solution of (3) $(\mu_0, \omega_0; h_0, \lambda_0)$ with $\omega_0 \in D(\gamma, \sigma)$ and h_0 non-degenerate then there is a set Ω in a small neighbourhood of (μ_0, ω_0) in $\Xi \times D(\gamma, \sigma)$, such that for each $(\mu, \omega) \in \Omega$ there is a solution $(\mu, \omega; h(\mu, \omega), \lambda(\mu, \omega))$ of (3). Moreover, the functions $h(\mu, \omega), \lambda(\mu, \omega)$ are Whitney-C^2.

Theorem 4 states that if $\lambda_0 = 0$ in Theorem 3, then there is a Cantor set $\mathcal{C}$ in a small neighbourhood of μ_0 in Ξ, and a function ω defined on $\mathcal{C}$ such that for each $\mu \in \mathcal{C}$ there is a solution $(\mu, \omega(\mu); h(\mu, \omega(\mu)), 0)$ of (3).

Theorem 1. *Let $\omega \in D(\gamma, \sigma)$, for some $\gamma > 0$ and $\sigma \geq d$. Assume that $\varrho, \rho > 0$, $\lambda_0 \in \mathbb{R}^d$, and f, $h_0 \in D^0(\mathbb{T}^d)$ are lifts of the form (6), that is, $f = A + u$ and $h_0 = B + v_0$. Define the error function*

$$e_0 \stackrel{\text{def}}{=} f \circ h_0 - h_0 \circ T_\omega - \lambda_0 \,.$$

Assume that the following hypotheses hold

1. $u \in \mathcal{A}\left(\mathbb{T}^d_{2\varrho\rho}, C^2\right)$, $v_0 \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$, *and* $\|\operatorname{Im} h_0\|_\rho < \varrho\rho$.
2. $e_0 \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$.
3. *h_0 is non-degenerate in the sense of Definition 1 with*

$$\left|\Phi^{-1}\right| \leq \eta_0 \,, \qquad \left\|[Dh_0(\theta)]^{-1}\right\|_\rho \leq \tilde{\eta}_0 \qquad\qquad \textit{for some} \quad \tilde{\eta}_0, \eta_0 > 0.$$

[3] See Section 2 of Chapter VI in [Ste70].

Then there exists a constant $\kappa > 0$, depending on d, σ, γ^{-2}, ϱ, η, $\tilde{\eta}$, $|B| + |Dv_0|_\rho$, and $|u|_{2\,\varrho\rho, C^2}$, such that if $q \geq 2(\sigma+1)$, $0 < \delta \leq \rho/8$, and

$$\kappa\, \delta^{-q}\, |\, e_0|_\rho < \min(1, \varrho)\,,$$

then there exists a constant vector $\lambda^* \in \mathbb{R}^d$ and a diffeomorphism $h^* \in D^0(\mathbb{T}^d)$ with $(h^* - h_0) \in \mathcal{A}\left(\mathbb{T}^d_{\rho-4\delta}, C^0\right)$, and such that

$$f \circ h^* = h^* \circ T_\omega + \lambda^*\,.$$

Moreover, the following inequalities hold:

$$\begin{aligned} |h^* - h_0|_{\rho-4\delta} &\leq \kappa\, \delta^{-\sigma}\ |\, e_0\,|_\rho\ , \\ |Dh^* - Dh_0|_{\rho-4\delta} &\leq \kappa\, \delta^{-(\sigma+1)}\ |\, e_0\,|_\rho\ , \\ |\lambda^* - \lambda_0| &\leq \kappa\ |\, e_0\,|_\rho\,. \end{aligned}$$

Remark 5. *It follows from the proof of Theorem 1 that the constant κ in Theorem 1 is an increasing function of the initial data ϱ, η_0, $\tilde{\eta}_0$, $|Dv_0|_\rho$, and $|u|_{2\,\varrho\rho, C^2}$. Moreover, the dependence on γ^{-2} of κ is of the form $\kappa = \gamma^{-2}\,\tilde{\kappa}$ where $\tilde{\kappa}$ does not depend on γ^{-2}.*

As we already mentioned in Remark 1, two solutions of (1) of the form (h_1, λ) and $(h_1 \circ T_{\theta_0}, \lambda)$, for some $\theta_0 \in \mathbb{R}^d$, are considered to be equivalent. Within this equivalence class the local uniqueness is the following:

Theorem 2. *Let ω, ϱ, ρ and f be as Theorem 1. Assume that $\lambda_1, \lambda_2 \in \mathbb{R}^d$, and $v_1, v_2 \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$ are such that for $i = 1, 2$ $h_i \stackrel{\text{def}}{=} B + v_i \in D^0(\mathbb{T}^d)$ satisfies hypotheses of Theorem 1, and that (λ_i, h_i) is a solution of (1). Then there exists a constant $c > 0$ depending on d, σ, γ^{-1}, $|u|_{\rho, C^2}$, $\left\|Dh_1^{-1}\right\|_\rho$, $\left|\left(\operatorname{avg}\left\{Dh_1^{-1}\right\}_\theta\right)^{-1}\right|_\rho$, and $|B| + \|Dv_1\|_\rho$, such that if*

$$c\, \rho^{-\sigma}\ \|h_1 - h_2\|_\rho < 1\,, \qquad \textit{and} \qquad c\ \|\lambda_1 - \lambda_2\|_\rho < 1\,,$$

then $\lambda_1 = \lambda_2$, and there exists an initial phase $\theta_0 \in \mathbb{R}^n$, such that $h_1 \circ T_{\theta_0} = h_2$ on $\mathbb{T}^d_{\rho/2}$.

Theorem 3. *Let $\{u_\mu\}_{\mu\in\Xi}$ be a C^2-parametric family of analytic functions on $\mathbb{T}^d$ with $\Xi \subset \mathbb{R}^s$ such that for each $\mu \in \Xi$, $u_\mu \stackrel{\text{def}}{=} u(\mu, \cdot) \in \mathcal{A}\left(\mathbb{T}^d_{2\,\varrho\,\rho}, C^2\right)$ and*

$$\left\|\, D^2 u(\mu, \theta)\right\|_{2\,\varrho\,\rho} < \Upsilon.$$

Let $\mu_0 \in \Xi$, and $\omega_0 \in D(\gamma, \sigma)$ for some $\sigma \geq d$ and $\gamma > 0$. Assume that for $\mu \in \Xi$, $f_\mu \stackrel{\text{def}}{=} A + u_\mu \in D^0(\mathbb{T}^d)$, and that there exist $h_0 = B + v_0 \in D^0(\mathbb{T}^d)$ and $\lambda_0 \in \mathbb{R}^d$, such that

1. *$v_0 \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$.*
2. *$\|\operatorname{Im}\,(h_0)\|_\rho < \rho\,\varrho$.*
3. *$f_{\mu_0} \circ h_0 = h_0 \circ T_{\omega_0} + \lambda_0$.*

4. h_0 *is non-degenerate (Definition 1) with* $\eta_0 = \eta(h_0)$ *and* $\tilde\eta_0 = \tilde\eta(h_0)$

Then there exists a positive constant κ*, depending on* d*,* σ*.* γ^{-1}*,* ϱ*,* Υ*,* η_0*,* $\tilde\eta_0$*, and* $\| Dh_0(\theta) \|_\rho$*, such that if* $q \geq 4\sigma + 2$ *and* $\varepsilon > 0$ *satisfy*

$$\kappa \, \rho^{-q} \, \varepsilon^2 < \min(1, \varrho) \, ,$$

then for each $(\mu,\omega) \in \Omega(\mu_0,\omega_0;\varepsilon) \stackrel{\text{def}}{=} \{(\mu,\omega) \in \Xi \times D(\gamma,\sigma) : |\, (\mu-\mu_0, \omega-\omega_0) \,| < \varepsilon\}$ *, there exist* $\lambda(\mu,\omega) \in \mathbb{R}^d$ *and* $v(\mu,\omega) \in \mathcal{A}\left(\mathbb{T}^d_{\rho/2}, C^0\right)$ *such that* $h(\mu,\omega) \stackrel{\text{def}}{=} B + v(\mu,\omega)$ *is non-degenerate in the sense of Definition 1 and*

$$f_\mu \circ h(\mu,\omega) = h(\mu,\omega) \circ T_\omega + \lambda(\mu,\omega) \, .$$

Define the functions

$$\begin{aligned} &H : \Omega(\mu_0,\omega_0;\varepsilon) \to \mathcal{A}\left(\mathbb{T}^d_{\rho/2}, C^0\right), \qquad && H(\mu,\omega) = h(\mu,\omega) \, , \\ &\Lambda : \Omega(\mu_0,\omega_0;\varepsilon) \to \mathbb{R}^d, && \Lambda(\mu,\omega) = \lambda(\mu,\omega) \, . \end{aligned} \tag{9}$$

There are functions a*,* b*,* α*, and* β*, defined on* $\Omega(\mu_0,\omega_0;\varepsilon)$ *such that for each* $(\mu,\omega) \in \Omega(\mu_0,\omega_0;\varepsilon)$*,*

$$\begin{aligned} a(\mu,\omega) &\in L\left(\mathbb{R}^s, \mathcal{A}\left(\mathbb{T}^d_{\rho/2}, C^0\right)\right), \qquad & \alpha(\mu,\omega) &\in L\left(\mathbb{R}^s, \mathbb{R}^d\right) \\ b(\mu,\omega) &\in L\left(\mathbb{R}^d, \mathcal{A}\left(\mathbb{T}^d_{\rho/2}, C^0\right)\right), & \beta(\mu,\omega) &\in L\left(\mathbb{R}^d, \mathbb{R}^d\right) \, , \end{aligned}$$

and such that the functions $\{\, H, a, b \,\}$ *and* $\{\, \Lambda, \alpha, \beta \,\}$ *are Whitney-*C^2 *on* $\Omega(\mu_0,\omega_0;\varepsilon)$ *(see Definition 3).*

Furthermore, for each $(\mu,\omega) \in \Omega(\mu_0,\omega_0;\varepsilon)$*,* $\beta(\mu,\omega)$ *is invertible and*

$$\beta(\mu,\omega)^{-1} = - \ \text{avg}\left\{ (\partial_\theta \, h(\mu,\omega)(\theta))^{-1} \right\}_\theta \, .$$

Remark 6. *Informally, we can think of* a *and* α *as the partial derivatives with respect to* μ*, and of* b *and* β *as the partial derivatives with respect to* ω *of, respectively,* H *and* Λ*.*

The proof of the Whitney differentiability we present here follows [dlLGE] where the authors prove Whitney differentiability of arbitrary order of a family of invariant tori for exact symplectic maps. In fact, if in Theorem 3, one assumes that u_μ *is a* C^ℓ*-parametric family of analytic maps on* $\mathbb{T}^d$*, then it is possible to prove – following [dlLGE] – that the mappings* H *and* Λ *are Whitney-*C^ℓ*. The idea is the following. First one constructs Lindstedt series in a neighbourhood of* (μ_0,ω_0)*. and then prove that the Lindstedt coefficients satisfy Definition 3. Informally, the Lindstedt coefficient of order* k *give us the Whitney derivative of order* k*. In this paper we restrict ourselves to the case* $\ell = 2$*, we hope to come back to general case in a future work.*

Theorem 4. *Assume that the hypotheses of Theorem 3 hold and that moreover* $\lambda_0 = 0$. *Then there exists a Cantor set* $\mathcal{C} \subset \Xi$ *and a function* $\omega : \mathcal{C} \to D(\gamma, \sigma)$ *such that, for* $\mu \in \mathcal{C}$, *one has* $(\mu, \omega(\mu)) \in \Omega(\mu_0, \omega_0; \varepsilon)$ *and*

$$f_\mu \circ h(\mu, \omega(\mu)) = h(\mu, \omega(\mu)) \circ T_{\omega(\mu)} \,.$$

Furthermore if α *and* β *are as in Theorem 3 and* ω_1 *is defined by*

$$\omega_1(\mu) \stackrel{\text{def}}{=} -\beta(\mu, \omega(\mu))^{-1} \alpha(\mu, \omega(\mu)) \,, \qquad \mu \in \mathcal{C},$$

then $\{\, \omega, \omega_1 \,\}$ *is Whitney-*C^2 *on* $\mathcal{C}$. *In particular, for each each* $\mu \in \mathcal{C}$

$$\left| \,\omega(\mu) - \left[\omega_0 \, - \, \beta(\mu_0, \omega_0)^{-1} \alpha(\mu_0, \omega_0) \, (\mu - \mu_0) \right] \right| \, \leq \kappa \, |\, \mu - \mu_0 \,|^2 \,,$$

where κ *is a constant depending on* d, σ. γ^{-1}, ϱ, Υ, η_0, $\tilde{\eta}_0$, *and* $\|\, Dh_0(\theta) \,\|_\rho$.

Theorems 1, 2 and 4 are proved in Section 4 and Theorem 3 is proved in Section 5.

3 Sketch of the procedure

The conjugation problem we are dealing with can be written in a functional setting as follows. Let u_μ be a C^2-parametric family of maps with parameter $\mu \in \Xi \subset \mathbb{R}^s$, such that for each $\mu \in \Xi$ the map $f_\mu \stackrel{\text{def}}{=} A + u_\mu$ belongs to $D^0(\mathbb{T}^d)$. For $(\mu, \omega; h, \lambda) \in \left(\Xi \times \mathbb{R}^d\right) \times D^0(\mathbb{T}^d) \times \mathbb{R}^d$ we define the functional

$$\mathcal{F}(\mu, \omega; h, \lambda) \stackrel{\text{def}}{=} f_\mu \circ h - h \circ T_\omega - \lambda \tag{10}$$

In the case that both μ and $\omega \in D(\gamma, \sigma)$ are fixed, we set $f = f_\mu$ and denote

$$F(h, \lambda) \stackrel{\text{def}}{=} \mathcal{F}(f, \omega; h, \lambda) \,. \tag{11}$$

We denote by $D_1\mathcal{F}(\mu, \omega; h, \lambda)$ and by $D_2\mathcal{F}(\mu, \omega; h, \lambda)$ the Fréchet derivative of $\mathcal{F}$ with respect to (μ, ω) and with respect to (h, λ), respectively. Similarly, $DF(h, \lambda)$ denotes the Fréchet derivative of F with respect to (h, λ).

Roughly, Theorem 1 states that if there is a sufficiently 'good' approximate solution (h_0, λ_0) of the functional equation

$$F(h, \lambda) = 0 \,, \tag{12}$$

and if h_0 is non-degenerate (see Definition 1), then it is possible to find a true solution of (12), which is close to the initial one. Theorem 2 states that the solutions of (12) are locally unique (see Remark 1).

In Section 3.1 we show that if an approximate solution of (12) is non-degenerate, then the corresponding linearised operator has an approximate right inverse [Ham82, Zeh75].

Moreover for a true non-degenerate solution of (12) the corresponding linearised operator is invertible – of course with some loss of domain of analyticity.

As it is proved in [Ham82, Mos66c, Mos66b, Zeh75], the existence of an approximate right inverse of the linearised equations is enough to prove the existence of a true solution of (12) by constructing a modified Newton method. The local uniqueness stated in Theorem 2 follows from the fact that the approximate solutions of the linear equations corresponding to (12) are unique when the initial phase is fixed.

An implicit function theorem for the functional equation

$$\mathcal{F}(\mu,\omega;h,\lambda)=0\,. \tag{13}$$

is given in Theorem 3, which states moreover that the implicit functions $h(\mu,\omega)$ and $\lambda(\mu,\omega)$ are Whitney-C^2. In Section 5.1 we show that the existence of such implicit functions is guaranteed by Theorem 1 and the invertibility properties of the linear operator

$$D_2\mathcal{F}(\mu,\omega;h,\lambda)(\Delta,\Delta\lambda)=\partial_2 f(\mu,h)\,\Delta-\Delta\circ T_\omega-\Delta\lambda\,,$$

which is essentially the same as $DF(h,\lambda)(\Delta,\Delta\lambda)$ for (μ,ω) fixed.

We proceed as follows. For $\varepsilon>0$ define $\Omega(\,\mu_0,\omega_0;\varepsilon)$ as in Theorem 3 and assume that $(\mu_0,\omega_0;h_0,\lambda_0)$ is a solution of (13), with h_0 non-degenerate. For ε sufficiently small and $(\mu,\omega)\in\Omega(\,\mu_0\omega_0;\varepsilon)$ we take $(\mu,\omega;h_0,\lambda_0)$ as an approximate solution of (13). Define the polynomials

$$H^{\leq 1}\,(\mu_0,\omega_0;x,y)=h(\mu_0,\omega_0)+a(\mu_0,\omega_0)\,(x-\mu_0)+b(\mu_0,\omega_0)\,(y-\omega_0)\,,$$

and

$$\Lambda^{\leq 1}\,(\mu_0,\omega_0;x,y)=h(\mu_0,\omega_0)+\alpha(\mu_0,\omega_0)\,(x-\mu_0)+\beta(\mu_0,\omega_0)\,(y-\omega_0)\,,$$

where the functions $a(\mu_0,\omega_0)$, $b(\mu_0,\omega_0)$, $\alpha(\mu_0,\omega_0)$, and $\beta(\mu_0,\omega_0)$, are solutions of the linear equations

$$\begin{aligned}D_2\mathcal{F}\,(\,\mu_0,\omega_0;h_0,\lambda_0\,)\,[\,a(\mu_0,\omega_0),\alpha(\mu_0,\omega_0)\,]&=-\partial_1 u(\mu_0,h_0(\theta))\,,\\ D_2\mathcal{F}\,(\,\mu_0,\omega_0;h_0,\lambda_0\,)\,[\,b(\mu_0,\omega_0),\beta(\mu_0,\omega_0)\,]&=-\partial_1 u(\mu_0,h_0(\theta))\,.\end{aligned} \tag{14}$$

It turns out that if ε is sufficiently small and $(\mu,\omega)\in\Omega(\,\mu_0\omega_0;\varepsilon)$, then

$$\big(\mu,\omega\,;\;H^{\leq 1}\,(\mu_0,\omega_0;\mu,\omega)\,,\Lambda^{\leq 1}\,(\mu_0,\omega_0;\mu,\omega)\,\big) \tag{15}$$

is an approximate solution of (13) with error of order 2 with respect to ε. Taking ε sufficiently small and applying Theorem 1 to the approximate solution (15) we obtain the existence of a true solution $(\mu,\omega;h(\mu,\omega),\lambda(\mu,\omega))$ of (13) for each $(\mu,\omega)\in\Omega(\mu_0,\omega_0;\varepsilon)$. In Section 5.2 we prove that if ε is sufficiently small then for each $(\mu,\omega)\in\Omega(\mu_0,\omega_0;\varepsilon)$. the equations obtained by replacing $(\mu_0,\omega_0;h_0,\lambda_0)$ with $(\mu,\omega;h(\mu,\omega),\lambda(\mu,\omega))$ in (14) have solutions $a(\mu,\omega)$, $b(\mu,\omega)$, $\alpha(\mu,\omega)$, and $\beta(\mu,\omega)$.

The fact that the non-degeneracy condition given in Definition 1 is an open property enables us to obtain uniform estimates of the involved quantities in such a way that the size of ε will depend only on the initial data ω_0, μ_0, and h_0. The Whitney differentiability of the functions H and Λ in (9) follows from the uniform estimates and from the local uniqueness of the solutions of (12).

Theorem 4 states that if the element f_{μ_0} of the family $\{f_\mu\}_{\mu\in\Xi}$ is conjugated to the rigid rotation T_{ω_0}, with $\omega_0 \in D(\gamma,\sigma)$, and the conjugation map h_0 is non-degenerate, then in a sufficiently small neighbourhood of μ_0 there is a Cantor set of parameters $\mathcal{C}$ such that for each $\mu \in \mathcal{C}$, the diffeomorphism f_μ is conjugated to the rigid rotation $T_{\omega(\mu)}$, with $\omega(\mu_0) = \omega_0$. Theorem 4 is proved in Section 4.3. The proof follows from Theorem 3 by applying first the a version of the Whitney Extension Theorem [Whi34] to the equation $\Lambda(\mu,\omega) = 0$ and then the classical Implicit Function Theorem to the extended equation $\hat{\Lambda}(\mu,\omega) = 0$.

3.1 The linearised equation

Let (h,λ) be an approximate solution of (12) with function error defined by

$$e(\theta) \stackrel{\text{def}}{=} F(h,\lambda)(\theta)\,. \tag{16}$$

Then, by taking derivatives with respect to θ we have the following fundamental equality

$$Df(h(\theta))Dh(\theta) = Dh(\theta+\omega) + De(\theta)\,. \tag{17}$$

Equality (17) enables us to solve approximately the linearised equation

$$DF(h,\lambda)(\Delta,\Delta\lambda) = Df(h(\theta))\,\Delta - \Delta\circ T_\omega - \Delta\lambda = g \tag{18}$$

in the case $g \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$, for some $\rho > 0$.

Lemma 2. *Let $\omega \in D(\gamma,\sigma)$ and $f = A + u \in D^0(\mathbb{T}^d)$, with $u \in \mathcal{A}\left(\mathbb{T}^d_{\varrho\rho}, C^0\right)$, for some $\varrho\,,\rho > 0$. Let $h = B + v \in D^0(\mathbb{T}^d)$, with $v \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$, be non-degenerate (Definition 1) and such that $\left|\operatorname{Im} h\right|_\rho < \varrho\rho$. Assume (h,λ) is an approximate solution of (12) with function error e defined by (16) such that $e \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$. Then for any $g \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$ the linear equation (18) has a unique approximate solution $(\Delta,\Delta\lambda) \in \mathcal{A}\left(\mathbb{T}^d_{\rho-\delta}, C^0\right) \times \mathbb{R}^d$ satisfying*

$$\operatorname{avg}\left\{\,(Dh(\theta))^{-1}\,\Delta\right\}_\theta = 0\,,$$

$$\Delta\lambda = \operatorname{avg}\left\{Dh(\theta)^{-1}\right\}_\theta^{-1}\,\operatorname{avg}\left\{Dh(\theta+\omega)^{-1}g\right\}_\theta\,, \tag{19}$$

and

$$\begin{aligned}
\|\Delta\|_{\rho-\delta} &\le c\,\gamma^{-1}\,\delta^{-\sigma}\,\|g\|_\rho\,,\\
\|\,D\Delta\|_{\rho-2\delta} &\le c\,\gamma^{-1}\,\delta^{-(\sigma+1)}\,\|g\|_\rho\,,\\
|\Delta\lambda| &\le c\,\|g\|_\rho\,.
\end{aligned}$$

where c is a constant depending on d, σ, $\left|(\mathrm{avg}\,\{Dh(\theta)^{-1}\}_\theta)^{-1}\right|$, $\|Dh(\theta)\|_\rho$, and $\|Dh(\theta)^{-1}\|_\rho$. Moreover, the following holds

$$\|DF(h,\lambda)(\Delta,\Delta\lambda)-g\|_{\rho-\delta} = \left\|De(\theta)\, Dh(\theta)^{-1}\,\Delta\right\|_{\rho-\delta} \leq \hat{c}\,\gamma^{-1}\,\delta^{-\sigma}\ \|e\|_\rho\ \|g\|_\rho\ .$$

Proof. Since $Dh(\theta)$ is invertible we can define the change of variables $\Delta(\theta) = Dh(\theta)\xi(\theta)$. Then using (17) we have that in the new variables (18) becomes

$$Dh(\theta+\omega)\,[\,\xi-\xi\circ T_\omega\,] + De(\theta)\,\xi\, - \Delta\lambda = g. \tag{20}$$

From (19) we have

$$\mathrm{avg}\left\{Dh(\theta+\omega)^{-1}\,(g+\Delta\lambda)\right\}_\theta = 0\,.$$

Hence [dlL01, Rüs75] the linear equation

$$\xi-\xi\circ T_\omega = Dh(\theta+\omega)^{-1}\,(g+\Delta\lambda)\ ,$$

has a unique solution with $\mathrm{avg}\,\{\xi\}_\theta = 0$ satisfying

$$\begin{aligned} \|\xi\|_{\rho-\delta} &\leq \tilde{c}\,\delta^{-\sigma}\ \Big(\|g\|_\rho + |\Delta\lambda|\Big)\ ,\\ \|D\xi(\theta)\|_{\rho-2\delta} &\leq \tilde{c}\,\delta^{-(\sigma+1)}\ \Big(\|g\|_\rho + |\Delta\lambda|\Big)\ , \end{aligned} \tag{21}$$

where $\tilde{c}$ is a constant depending on d, σ, $\|Dh(\theta)^{-1}\|_\rho$.

Lemma 2 follows from (20), (21), and (19). □

Remark 7. *Notice that if the error function e in (16) is identically zero, then Lemma 2 provides an exact solution of the linear equation (18).*

3.2 The non-degeneracy condition

From Lemma 2 we know that the linearised equation (18) can be solved approximately in the case that h is non-degenerate (Definition 1). In order to use Lemma 2 to construct a modified Newton method to solve the nonlinear equation (12) we need that at each step the corrections to the approximate solution are again non-degenerate. The non-degeneracy condition we are considering here amounts to invertibility of matrices. So the openness of the non-degeneracy condition follows from the fact that invertibility of matrices is an open property.

Lemma 3 states the openness property of the non-degeneracy condition and gives estimates which enable us obtain uniform bounds of the constants in the modified Newton method described in Section 4.1 and in proof of Theorem 3 (Section 5).

Lemma 3. *Assume that h is non-degenerate and let η and $\tilde{\eta}$ be as in Definition 1. Let $\Delta \in \mathcal{A}\left(\mathbb{T}^d_\rho, C^0\right)$ be such that*

$$\tilde{\eta}\ \|\,D\Delta\|_\rho \leq 1/2\,, \qquad \textit{and} \qquad 2\,\eta\,\tilde{\eta}^2\ \|\,D\Delta\|_\rho \leq 1/2. \tag{22}$$

Then $h_1 \stackrel{\rm def}{=} h + \Delta$ *is non-degenerate and*

$$\left\|Dh_1(\theta)^{-1}\right\|_\rho \leq \tilde{\eta} + 2\,\tilde{\eta}^2\ \|D\Delta(\theta)\|_\rho\,,$$
$$\left|\operatorname{avg}\left\{Dh_1(\theta)^{-1}\right\}_\theta^{-1}\right|_\rho \leq \eta + 2\,\eta^2\ \|D\Delta(\theta)\|_\rho\,.$$

Proof. Let $\theta \in \mathbb{T}^d_\rho$, the first inequality in (22) implies $[\,I_d - Dh(\theta)^{-1}D\Delta(\theta)\,]$ is invertible, where I_d represents the $d \times d$ identity matrix. Performing some simple computations one has

$$Dh_1(\theta)^{-1} = Dh(\theta)^{-1} - \left[\,I_d - Dh(\theta)^{-1}D\Delta(\theta)\,\right]^{-1} Dh(\theta)^{-1}D\Delta(\theta)\,Dh(\theta)^{-1}\,,$$

from which Lemma 3 follows. □

4 Proof of theorems 1, 2 and 4

4.1 Proof of Theorem 1

Let ω, ρ, ϱ, $f = A + u$, h_0, λ_0, and e_0 be as in Theorem 1. In this section we describe a modified Newton method to solve the non-linear equation (12). The method converges to a true solution if $\|e_0\|_\rho$ is sufficiently small. The proof of convergence is quite standard in KAM theory, so we only give an outline of the main steps.

Starting with the approximate solution (h_0, λ_0) of (12) we construct a sequence of approximate solutions

$$h_{n+1} = h_n + \varphi_n\,, \qquad \lambda_{n+1} = \lambda_n + \Delta\lambda_n\,.$$

The key point on constructing the modified Newton method, as it is explained in [Mos66c] and [Zeh75], is to solve approximately the corresponding linearised equation (18). Lemma 2 provides such approximate solution. Moreover, if $\|e_0\|$ is sufficiently small, Lemma 3 enables us to iterate the method. Lemma 4 provides one step of the modified newton method and it is a immediate consequence of lemmas 2, and 3 and Taylor's Theorem.

Lemma 4. *Assume that* $\varrho_n, \rho_n > 0$, $\lambda_n \in \mathbb{R}^d$, *and* $h_n \in D^0(\mathbb{T}^d)$ *are given and define the error function*

$$e_n \stackrel{\rm def}{=} F(h_n, \lambda_n)\,. \tag{23}$$

Assume that the following hypotheses hold:

S1(n). $\|Dh_n\|_{\rho_n} \leq \tau_n$.

S2(n). $\|\operatorname{Im} h_n\|_{\rho_n} < \zeta_n$, *with* $\rho/2 \leq \rho_n \leq \rho$ *and* $\zeta_n = \varrho\,\rho \sum_{k=0}^{n} 2^{-k}$.

S3(n). $e_n \in \mathcal{A}\left(\mathbb{T}^d_{\rho_n}, C^0\right)$.

S4(n). *h_n is non-degenerate (Definition 1): The matrix $Dh_n(\theta)$ is invertible for any $\theta \in \mathbb{T}^d_{\rho_n}$. The matrix $\Phi_n \stackrel{\text{def}}{=} \text{avg}\{ Dh_n(\theta)^{-1} \}_\theta$ is invertible. Moreover, there exist two positive numbers η_n and $\tilde{\eta}_n$ such that*

$$\left| (Dh_n(\theta))^{-1} \right|_{\rho_n} \leq \tilde{\eta}_n \,, \qquad \left| \Phi_n^{-1} \right| \leq \eta_n \,.$$

Then there exists a vector $\Delta\lambda_n \in \mathbb{R}^d$ and a function $h_{n+1} \in D^0(\mathbb{T}^d)$ such that $h_{n+1} - h_n \in \mathcal{A}\left(\mathbb{T}^d_{\rho_n - 2\delta_n}, C^0\right)$ with $0 < \delta_n < \rho_n/2$, and such that if $\rho_{n+1} = \rho_n - 2\,\delta_n$, then the following estimates hold

$$\begin{aligned} \| h_{n+1} - h_n\|_{\rho_{n+1}} &\leq M_n\, \delta_n^{-\sigma} \, \|e_n\|_{\rho_n} \,, \\ \|Dh_{n+1} - Dh_n\|_{\rho_{n+1}} &\leq M_n\, \delta_n^{-(\sigma+1)} \, \|e_n\|_{\rho_n} \,, \\ | \lambda_{n+1} - \lambda_n| &\leq \eta_n\, \tilde{\eta}_n \, \|e_n\|_{\rho_n} \,, \end{aligned}$$

where M_n is a constant which depends on d, σ, γ^{-2}, η_n, $\tilde{\eta}_n$, and $\| Dh_n\|_{\rho_n}$.

Moreover the following estimate holds

$$|e_{n+1}|_{\rho_{n+1}} \leq \hat{M}_n\, \delta_n^{-2\sigma} \, |e_n|^2_{\rho_n} \,,$$

where e_{n+1} is defined by (23) by replacing n with $n+1$ and $\hat{M}_n$ is a constant which depends on M_n, and $\| u \|_{2\,\varrho\,\rho, C^2}$.

Furthermore, there exists a constant $\tilde{M}_n$, depending on M_n, $\tilde{\eta}_n$, and η_n such that if

$$\tilde{M}_n\, 2^{(n+1)}\, \delta_n^{-(\sigma+1)} \, |e_n|_{\rho_n} < \min(1,\, \varrho)$$

then properties $S(n+1)$ hold with

$$\begin{aligned} \tau_n &\stackrel{\text{def}}{=} \tau_n + \varrho\, 2^{-(n+1)} \,, \\ \zeta_{n+1} &\stackrel{\text{def}}{=} \zeta_n + \rho\,\varrho\, 2^{-(n+1)} \,, \\ \eta_{n+1} &\stackrel{\text{def}}{=} \eta_n \left(1 + 2^{-(n+1)}\right) \,, \\ \tilde{\eta}_{n+1} &\stackrel{\text{def}}{=} \tilde{\eta}_n \left(1 + 2^{-(n+1)}\right) \,. \end{aligned}$$

Remark 8. *Let M_n $\tilde{M}_n$ and $\hat{M}_n$ be as in Lemma 4. Define*

$$\kappa_n \stackrel{\text{def}}{=} \max\left(M_n, \tilde{M}_n, \hat{M}_n\right) \,.$$

Performing the proof of Lemma 4 one realizes that κ_n is increasing with respect to η_n, $\tilde{\eta}_n$, and $|Dh_n|_{\rho_n}$. Let us write explicitly dependence on these variables as follows:

$$\kappa_n = \vartheta\left(\eta_n, \tilde{\eta}_n, \|Dh_n\|_{\rho_n}\right),$$

Now the proof of Theorem 1 follows the same lines of the proof of Theorem 1.1 in [Zeh75] using Lemma 4 and defining

$$\kappa_\infty = \vartheta\left(3\,\eta, 3\,\tilde{\eta}\,, \|Dh\|_\rho + 2\,\varrho\right).$$

4.2 Proof of Theorem 2

The local uniqueness stated in Theorem 2 is proved by using the fact that, in the case that (h, λ) is a solution of (12), the linearised equation (18) has a solution which is unique if the average of $Dh(\theta)^{-1}\Delta$ is fixed (see Lemma 5 bellow). We only include here the statement of the uniqueness of the linearised equation (18). The complete proof of Theorem 2 is not reported here because it is essentially the same as in Section 6 of [dlLGJV05], where the authors prove local uniqueness of invariant tori for exact symplectic maps.

Lemma 5. *Assume that the hypotheses of Lemma 2 hold and that $e = F(h, \lambda) \equiv 0$. Let $(\Delta, \Delta\lambda)$ be as in Lemma 2. Then for any solution $(\hat{\Delta}, \Delta\lambda)$ of the linear equation (18), the following holds:*

$$\Delta = \hat{\Delta} - Dh(\theta)\ \mathrm{avg}\left\{ Dh(\theta)^{-1}\hat{\Delta} \right\}_\theta .$$

4.3 Proof of Theorem 4

Assume that the hypotheses of Theorem 3 holds, and let $\Omega\,(\mu_0, \omega_0; \varepsilon)$, H and Λ be as in Theorem 3. Theorem 4 in Chapter VI in [Ste70] (a version of the Whitney Extension Theorem) ensures the existence of an extension of Λ, say $\hat{\Lambda}$, which is Whitney-C^2 on $\mathbb{R}^s \times \mathbb{R}^d$ and such that for any $(\mu, \omega) \in \Omega\,(\mu_0, \omega_0; \varepsilon)$ the following holds:

1. $\hat{\Lambda}(\mu, \omega) = \Lambda(\mu, \omega)$,
2. $\hat{\Lambda}(\mu_0, \omega_0) = 0$,
3. $\partial_1 \left[\hat{\Lambda}(x, y) \right]_{(x,y)=(\mu,\omega)} = \alpha(\mu, \omega)$,
4. $\partial_2 \left[\hat{\Lambda}(x, y) \right]_{(x,y)=(\mu,\omega)} = \beta(\mu, \omega) = -\ \left[\mathrm{avg}\left\{ [\, \partial_\theta\, h(\mu, \omega)(\theta)\,]^{-1} \right\}_\theta \right]^{-1}$ is invertible.

Then, the classical Implicit Function Theorem implies the existence of a neighbourhood $V \subset \mathbb{R}^s$ of μ_0 and a unique Whitney-C^2 map $\omega : V \to \mathbb{R}^d$ such that for all $x \in V$

$$\hat{\Lambda}(x, \omega\,(x)) = 0\,.$$

Define

$$\mathcal{C} \stackrel{\text{def}}{=} \omega^{-1}\left(\{\, y \in D(\gamma, \sigma) : |\, y - \omega_0| <\ \varepsilon \,\} \right) \bigcap \{\, \mu \in \Xi : |\, \mu - \mu_0| <\ \varepsilon \,\}\,.$$

Then for any $\mu \in \mathcal{C}$ we have

$$f \circ h(\mu, \omega(\mu)) = h(\mu, \omega(\mu)) \circ T_{\omega(\mu)}\,.$$

Moreover, $\omega(\mu_0) = \omega_0$ and for any $\mu \in \mathcal{C}$

$$\begin{aligned} D\,[\,\omega(x)\,]_{x=\mu} &= -\,\partial_2 \left[\hat{\Lambda}(x, y) \right]^{-1}_{(x,y)=(\mu,\omega(\mu))} \partial_1 \left[\hat{\Lambda}(x, y) \right]_{(\mu,\omega(\mu))=(\mu,\omega(\mu))} \\ &= -\beta(\mu, \omega)^{-1}\, \alpha(\mu, \omega(\mu))\,. \end{aligned}$$

This finishes the proof of Theorem 4.

5 Proof of Theorem 3

Assume that the hypotheses of Theorem 3 hold. Throughout this section $\delta \stackrel{\text{def}}{=} \rho/32$ and for $\varepsilon > 0$, $B(\mu_0, \omega_0; \varepsilon)$ denotes the open ball of radius ε and centre (μ_0, ω_0).

The proof of Theorem 3 is divided into two parts. First, in Section 5.1 we prove that there exist two functions H and Λ defined in an ε-neighbourhood of (μ_0, ω_0) in $\Xi \times D(\gamma, \sigma)$ such that for each (μ, ω) in such neighbourhood, $(\mu, \omega; H(\mu, \omega), \Lambda(\mu, \omega))$ is a solution of (13). Since the non-degeneracy condition in Definition 1 is an open property (see Lemma 3) we obtain uniform bounds, which enables us to choose ε depending on μ_0, ω_0, and h_0, but independent on μ, ω. Second, in Section 5.2 we prove the Whitney regularity of H and Λ.

5.1 Existence of the implicit functions

We prove the existence of the implicit functions $H(\mu, \omega)$ and $\Lambda(\mu, \omega)$ in Theorem 3 in two parts. The first part consists in finding ε sufficiently small such that for any $(x, y) \in B(\mu_0, \omega_0; \varepsilon)$ with $x \in \Xi$, the equation

$$\mathcal{F}(x, y; h, \lambda) = 0 \tag{24}$$

has an approximate solution for which the error function has norm of order 2 with respect to $|(x - \mu_0, y - \omega_0)|$,. This is done in Lemma 6 bellow.

In the second part we choose ε sufficiently small such that if $(x, y) \in B(\mu_0, \omega_0; \varepsilon)$ with $x \in \Xi$, and $y \in D(\gamma, \sigma)$, then Theorem 1 implies existence of a solution of (24) (this is done in Lemma 7). Moreover, in Lemma 8 we give estimates of the norm of the difference between the functions $H(\mu, \omega)$, $\Lambda(\mu, \omega)$, and h_0, λ_0.

Lemma 6. *Let f_μ, μ_0, ω_0, h_0 and λ_0 be as in Theorem 3. Then there exist two polynomials*

$$H^{\leq 1}(\mu_0, \omega_0; x, y)(\theta) = h_0(\theta) + a(\mu_0, \omega_0)(\theta)\,(x - \mu_0) + b(\mu_0, \omega_0)(\theta)(y - \omega_0)\,, \tag{25}$$

and

$$\Lambda^{\leq 1}(\mu_0, \omega_0; x, y) = \lambda_0 + \alpha(\mu_0, \omega_0)\,(x - \mu_0) + \beta(\mu_0, \omega_0)(y - \omega_0)\,, \tag{26}$$

with

$$\begin{aligned} \|\, a(\mu_0, \omega_0)\|_{\rho-\delta} &\leq \kappa_1\, \gamma^{-1}\, \delta^{-\sigma}\, \|\partial_1\, f\,(\mu_0, h_0(\theta))\|_\rho \,, \\ \|\, b(\mu_0, \omega_0)\|_{\rho-\delta} &\leq \kappa_1\, \gamma^{-1}\, \delta^{-\sigma}\, \|\partial_\theta\, h_0(\theta)\|_\rho \,, \end{aligned} \tag{27}$$

$$\begin{aligned} |\, \alpha(\mu_0, \omega_0)| &\leq \kappa_1\, \|\partial_1 f\,(\mu_0, h_0(\theta))\|_\rho \,, \\ |\, \beta(\mu_0, \omega_0)| &\leq \kappa_1\, \|\partial_\theta\, h_0(\theta)\|_\rho \,, \end{aligned} \tag{28}$$

where κ_1 depends on d, σ, η_0, $\tilde{\eta}_0$, $\|\partial_\theta h_0(\theta)\|_\rho$.

Assume that $\varepsilon > 0$ is such that

$$\kappa_2\, \delta^{-(\sigma+1)}\, \varepsilon \leq \min(1, \varrho)\,, \tag{29}$$

where

$$\kappa_2 \stackrel{\text{def}}{=} 4\,\kappa_1\,\max\left(1,\tilde\eta_0,\eta_0\,\tilde\eta_0^2,\eta_0\right)\,\gamma^{-1}\,\left(\|\partial_1\,u\,(\mu_0,h_0(\theta))\|_\rho+\|\partial_\theta\,h_0(\theta)\|_\rho\right).$$

Then for each $(x,y)\in B(\mu_0,\omega_0;\varepsilon)$ *with* $x\in\Xi$

$$\left\|\,\mathcal{F}\left(x,y;H^{\leq 1}(\mu_0,\omega_0;x,y),\Lambda^{\leq 1}(\mu_0,\omega_0;x,y)\right)\right\|_{\rho-2\delta}\leq\hat\kappa_2\;\delta^{-2\sigma}\;|\,(x-\mu_0,y-\omega_0)\,|^2\;, \tag{30}$$

where $\hat\kappa_2$ *depends on* d, κ_1^2, γ^{-2}, $\|\,\partial_\theta h_0(\theta)\,\|_\rho$, *and* $\|\,\partial_2^2 u_{\mu_0}(\theta)\|_{2\,\varrho\,\rho}$.

Moreover, for each $(x,y)\in B(\mu_0,\omega_0;\varepsilon)$ *the map* $H^{\leq 1}(\mu_0,\omega_0;x,y)$ *is non-degenerate and the following estimates hold*

$$\left\|\,\partial_\theta\,H^{\leq 1}(\mu_0,\omega_0;x,y)(\theta)\,\right\|_{\rho-2\delta}\leq\|\partial_\theta h_0(\theta)\|_\rho+\varrho/2\,, \tag{31}$$

$$\begin{aligned}\left\|\,\left[\,\partial_\theta\,H^{\leq 1}(\mu_0,\omega_0;x,y)(\theta)\,\right]^{-1}\right\|_{\rho-2\delta}&\leq\tilde\eta_0\left(1+2^{-1}\right)\\ \left|\operatorname{avg}\left\{\,\left[\,\partial_\theta\,H^{\leq 1}(\mu_0,\omega_0;x,y)(\theta)\,\right]^{-1}\right\}_\theta^{-1}\right|&\leq\eta_0\left(1+2^{-1}\right)\,.\end{aligned} \tag{32}$$

Proof. It is clear that in order to satisfy (30) the coefficients of the polynomials (25) and (26) have to be such that

$$\frac{\partial}{\partial(x,y)}\left[\,\mathcal{F}\left(x,y;H^{\leq 1}(\mu_0,\omega_0;x,y),\Lambda^{\leq 1}(\mu_0,\omega_0;x,y)\right)\right]_{(x,y)=(\mu_0,\omega_0)}=0\,. \tag{33}$$

Performing some simple computations one sees that formally the coefficients a, b, α, and β have to satisfy the following linear equations

$$\begin{aligned}D_2\mathcal{F}(\mu_0,\omega_0;h_0,\lambda_0)\,[\,a(\mu_0,\omega_0),\alpha(\mu_0,\omega_0)\,]&=-\partial_1 u\,(\mu_0,h_0(\theta))\;,\\ D_2\mathcal{F}(\mu_0,\omega_0;h_0,\lambda_0)[\,b(\mu_0,\omega_0),\beta(\mu_0,\omega_0)\,]&=\partial_\theta h_0(\theta+\omega_0)\,.\end{aligned}$$

Then estimates (27) and (28) follow from Lemma 2 by choosing

$$\operatorname{avg}\left\{\,[\,Dh_0(\theta)\,]^{-1}\,a(\mu_0,\omega_0)\right\}_\theta=\operatorname{avg}\left\{\,[\,Dh_0(\theta)\,]^{-1}\,b(\mu_0,\omega_0)\right\}_\theta=\,0\,.$$

If (29) holds and $(x,y)\in B(\mu_0,\omega_0;\varepsilon)$), from Cauchy's inequalities and (27) we have

$$\begin{aligned}\left\|\,\operatorname{Im}\,\left(H^{\leq 1}(\mu_0,\omega_0;x,y)\right)\right\|_{\rho-2\delta}&\leq\|\operatorname{Im}\,\partial_\theta h_0(\theta)\|_\rho+\rho\,\left\|\,\partial_\theta\left[\,H^{\leq 1}(\mu_0,\omega_0;x,y)-h_0(\theta)\,\right]\right\|_{\rho-2\delta}\\ &\leq\varrho+\kappa_2\,\varepsilon\,\rho\\ &<2\,\varrho\,\rho\,.\end{aligned}$$

Therefore the composition $f(x,H^{\leq 1}(\mu_0,\omega_0;x,y))$ is well defined for each $(x,y)\in B(\mu_0,\omega_0;\varepsilon)$, with $x\in\Xi$. Hence (30) follows from Taylor's Theorem, estimates (33), (27), (28), and from the following inequality

$$\|h_0\circ T_y-[\,h_0\circ T_{\omega_0}-\partial_\theta h_0(\theta+\omega_0)\,(y-\omega_0)\,]\,\|_{\rho-2\,\delta}\leq\delta^{-1}\;\|\partial_\theta h_0(\theta)\|_\rho\;|y-\omega_0|^2\;,$$

that is obtained using Cauchy's inequalities.

Estimate (31) is obtained by taking derivatives with respect to θ in (25) and using (27) and (29). The non-degeneracy of $H^{\leq 1}(\mu_0,\omega_0;x,y)$ and estimates (32) follow from (27), (29), and Lemma 3. □

We now apply Theorem 1 to the approximate solutions of (13) given by Lemma 6:

$$\left(\mu, \omega; H^{\leq 1}(\mu_0, \omega_0; \mu, \omega), \Lambda^{\leq 1}(\mu_0, \omega_0; \mu, \omega)\right) .$$

Lemma 7. *Let*

$$\Omega(\mu_0, \omega_0; \varepsilon) \stackrel{\text{def}}{=} \{ (\mu, \omega) \in \Xi \times D(\gamma, \sigma) : |(\mu - \mu_0, \omega - \omega_0)| < \varepsilon \}$$

and let $\hat{\kappa}_2$ be as in Lemma 6. There exists a constant κ_3, depending on d, σ, γ^{-2}, ϱ, $\tilde{\eta}_0$, η_0, and $\hat{\kappa}_2$, such that if $q \geq 2(\sigma + 1)$ and

$$\kappa_3 \, \delta^{-(q+2\sigma)} \, \varepsilon^2 < \min(1, \varrho) , \tag{34}$$

then for each $(\mu, \omega) \in \Omega(\mu_0, \omega_0; \varepsilon)$ there exists a solution $(\mu, \omega; h(\mu, \omega), \lambda(\mu, \omega))$ of the equation (13) such that $h(\mu, \omega) = B + v(\mu, \omega)$, with $v(\mu, \omega) \in \mathcal{A}\left(\mathbb{T}^d_{\rho - 6\delta}, C^0\right)$ and

$$\begin{aligned}
\left\| h(\mu, \omega) - H^{\leq 1}(\mu_0, \omega_0; \mu, \omega) \right\|_{\rho - 6\delta} &\leq \kappa_3 \, \delta^{-3\sigma} \; |(\mu - \mu_0, \omega - \omega_0)|^2 \\
\left\| \partial_\theta h(\mu, \omega)(\theta) - \partial_\theta H^{\leq 1}(\mu_0, \omega_0; \mu, \omega)(\theta) \right\|_{\rho - 6\delta} &\leq \kappa_3 \, \delta^{-(3\sigma+1)} \; |(\mu - \mu_0, \omega - \omega_0)|^2 \\
\left| \lambda(\mu, \omega) - \Lambda^{\leq 1}(\mu_0, \omega_0; \mu, \omega) \right| &\leq \kappa_3 \, \delta^{-2\sigma} \; |(\mu - \mu_0, \omega - \omega_0)|^2 .
\end{aligned}$$

Moreover, if ε also satisfies

$$\kappa_4 \, \delta^{-(3\sigma+1)} \, \varepsilon^2 \; < \; \min(1, \varrho) , \tag{35}$$

where

$$\kappa_4 \stackrel{\text{def}}{=} 2^3 \, \kappa_3 \max\left(1, \tilde{\eta}_0, 2\eta_0 \, \tilde{\eta}_0^2, \eta_0\right) ,$$

then for each $(\mu, \omega) \in B(\mu_0, \omega_0; \varepsilon)$ $h(\mu, \omega)$ is non-degenerate and the following estimates hold

$$\begin{aligned}
\|\partial_\theta h(\mu, \omega)(\theta)\|_{\rho - 6\delta} &\leq \|\partial_\theta h_0(\theta)\|_\rho + \varrho\left(2^{-1} + 2^{-2}\right) , \\
\left\| [\, \partial_\theta h(\mu, \omega)(\theta) \,]^{-1} \right\|_{\rho - 6\delta} &\leq \tilde{\eta}_0 \left(1 + 2^{-1} + 2^{-2}\right) \\
\left\| \operatorname{avg} \left\{ [\, \partial_\theta h(\mu, \omega)(\theta) \,]^{-1} \right\}_\theta^{-1} \right\|_{\rho - 6\delta} &\leq \eta_0 \left(1 + 2^{-1} + 2^{-2}\right) .
\end{aligned} \tag{36}$$

Proof. This is an immediate consequence of Theorem 1, Lemma 6 and Lemma 3. □

Lemmas 6 and 7 yield

Lemma 8. *Let $\varepsilon > 0$ be such that lemmas 6 and 7 hold and such that*

$$\delta^{-2\sigma} \, \varepsilon \; < 1 . \tag{37}$$

Then for each $(\mu, \omega) \in \Omega\,(\mu_0, \omega_0; \varepsilon)$, the functions $h(\mu, \omega)$ and $\lambda(\mu, \omega)$ defined in Lemma 7 satisfy the following estimates

$$\begin{aligned}
\| \, h(\mu, \omega) - h_0 \, \|_{\rho - 6\delta} &\leq \; \kappa_5 \, \delta^{-\sigma} \; |(\mu - \mu_0, \omega - \omega_0)| , \\
| \, \lambda(\mu, \omega) - \lambda_0 \, | &\leq \; \kappa_5 \; |(\mu - \mu_0, \omega - \omega_0)| ,
\end{aligned}$$

where

$$\kappa_5 \stackrel{\text{def}}{=} \kappa_3 + \kappa_1 \, \gamma^{-1} \; \left(\|\partial_1 u(\mu_0, h_0(\theta))\|_\rho + \|\partial_\theta h_0(\theta)\|_\rho \right) .$$

5.2 Whitney differentiability

For ε as in Lemma 7 the implicit mappings $H : \Omega(\mu_0, \omega_0; \varepsilon) \to D^0(\mathbb{T}^d)$ and $\Lambda : \Omega(\mu_0, \omega_0; \varepsilon) \to \mathbb{R}^d$ in (9) are well defined. We are now ready to prove that – taking ε sufficiently small– the polynomials $H^{\leq 1}(\mu, \omega; x, y)$ and $\Lambda^{\leq 1}(\mu, \omega; x, y)$ corresponding to the solution $(\mu, \omega; h(\mu, \omega), \lambda(\mu, \omega))$ are well defined for all $(\mu, \omega) \in \Omega(\mu_0, \omega_0; \varepsilon)$. Informally, this means that the partial derivatives of the mappings H and Λ are well defined for $(\mu, \omega) \in \Omega(\mu_0, \omega_0; \varepsilon)$.

We emphasise that in order to prove the Whitney regularity of the functions $H(\mu, \omega)$ and $\Lambda(\mu, \omega)$ we have to allow changes on the initial phases (see Lemma 11 bellow), However, this is not a problem because we are working within the equivalence class described in Remark 1.

Throughout this section we assume that Lemma 7 holds and that $\varepsilon_1 > 0$ satisfies (29), (34), (35) and (37).

Lemma 9. *For each $(\mu, \omega) \in \Omega(\mu_0, \omega_0; \varepsilon_1)$ let $h(\mu, \omega)$ and $\lambda(\mu, \omega)$ be as in Lemma 7. There are two polynomials*

$$H^{\leq 1}(\mu, \omega; x, y)(\theta) = h(\mu, \omega) + a(\mu, \omega)(\theta)\,(x - \mu) + b(\mu, \omega)(\theta)(y - \omega)\,,$$

and

$$\Lambda^{\leq 1}(\mu, \omega; x, y) = \lambda(\mu, \omega) + \alpha(\mu, \omega)\,(x - \mu) + \beta(\mu, \omega)(y - \omega)\,,$$

where the coefficients $a(\mu, \omega), b(\mu, \omega)$, $\alpha(\mu, \omega)$, and $\beta(\mu, \omega)$ are the solutions of the linear equations

$$\begin{aligned} D_2\mathcal{F}(\mu, \omega; h(\mu, \omega), \lambda(\mu, \omega))\,[\,a(\mu, \omega), \alpha(\mu, \omega)\,] &= -\,\partial_1 u\,(\mu, h(\mu, \omega)(\theta))\,, \\ D_2\mathcal{F}(\mu, \omega; h(\mu, \omega), \lambda(\mu, \omega))[\,b(\mu, \omega), \beta(\mu, \omega)\,] &= \partial_\theta h(\mu, \omega)(\theta + \omega)\,, \end{aligned} \tag{38}$$

satisfying

$$\begin{aligned} \operatorname{avg}\left\{\,[\,Dh(\mu, \omega)(\theta)\,]^{-1}\,a(\mu, \omega)\right\}_\theta &= 0, \\ \operatorname{avg}\left\{\,[\,Dh(\mu, \omega)(\theta)\,]^{-1}\,b(\mu, \omega)\right\}_\theta &= 0. \end{aligned} \tag{39}$$

Moreover, if κ_5 is as in Lemma 8, then there exists a constant κ_6, depending on d, σ, ϱ, η_0, $\tilde{\eta}_0$, κ_5, $\|\partial_\theta h_0(\theta)\|_\rho$, $\|\partial_1 u(\mu_0, h_0(\theta))\|_\rho$, $\|\partial_2\partial_1\, u\,(\mu_0, h_0)\|_\rho$, and $\|\partial_1^2 u\,(\mu_0, h_0)\|_\rho$, such that for any $(\mu, \omega) \in \Omega(\mu_0, \omega_0; \varepsilon)$, with $0 < \varepsilon < \varepsilon_1$,

$$\|h(\mu, \omega)\|_{\rho - 8\delta} \leq \|h_0\|_\rho + \kappa_6\,\delta^{-\sigma}\,\varepsilon\,, \tag{40}$$

$$\begin{aligned} \|\,a(\mu, \omega)\|_{\rho - 8\,\delta} &\leq \kappa_6\,\gamma^{-1}\,\delta^{-\sigma}\,, \quad |\,\alpha(\mu, \omega)| \leq \kappa_6\,, \\ \|\,b(\mu, \omega)\|_{\rho - 8\delta} &\leq \kappa_6\,\gamma^{-1}\,\delta^{-\sigma}\,, \quad |\,\beta(\mu, \omega)| \leq \kappa_6\,. \end{aligned} \tag{41}$$

Furthermore, for any $(x, y) \in B(\mu_0, \omega_0; \varepsilon)$ the following estimates hold:

$$\left\|H^{\leq 1}(\mu, \omega; x, y) - h(\mu, \omega)\right\|_{\rho - 8\delta} \leq \kappa_6\,\gamma^{-1}\,\delta^{-\sigma}\,\,|\,(x - \mu, y - \omega)\,|\,, \tag{42}$$

$$\left|\Lambda^{\leq 1}(\mu, \omega; x, y) - \lambda(\mu, \omega)\right| \leq \kappa_6\,\,|\,(x - \mu, y - \omega)\,|\,,$$

Proof. The existence of the solutions of (38) is guaranteed by Lemma 2. Moreover, from estimates in (36) we have that there is a constant κ_7, depending on d, σ, η_0, $\tilde{\eta}_0$ and ϱ, such that

$$\begin{aligned} \| a(\mu,\omega)\|_{\rho-8\delta} &\leq \kappa_7\,\gamma^{-1}\,\delta^{-\sigma}\ \|\,\partial_1 u(\mu,h(\mu,\omega)(\theta))\,\|_{\rho-6\delta}\ , \\ \| b(\mu,\omega)\|_{\rho-8\delta} &\leq \kappa_7\,\gamma^{-1}\,\delta^{-\sigma}\ \|\,\partial_\theta h(\mu,\omega)(\theta)\,\|_{\rho-7\delta}\ , \end{aligned} \tag{43}$$

and

$$\begin{aligned} |\,\alpha(\mu,\omega)| &\leq \kappa_7\ \|\,\partial_1 u(\mu,h(\mu,\omega)(\theta))\,\|_{\rho-6\delta}\ , \\ |\,\beta(\mu,\omega)| &\leq \kappa_7\ \|\,\partial_\theta h(\mu,\omega)(\theta)\,\|_{\rho-6\delta}\ . \end{aligned} \tag{44}$$

Using Lemma 8 one has

$$\|\partial_1\, u\,(\mu,h(\mu,\omega)) - \partial_1\, u\,(\mu_0,h_0)\|_{\rho-6\delta} \leq \hat{\kappa}_7\,\delta^{-\sigma}\,\varepsilon\,,$$

where

$$\hat{\kappa}_7 \stackrel{\text{def}}{=} \kappa_7\ \|\partial_2\partial_1\, u\,(\mu_0,h_0)\|_\rho + \left\|\partial_1^2 u\,(\mu_0,h_0)\right\|_\rho\ .$$

Therefore

$$\begin{aligned} \|\partial_1\, u\,(\mu,h(\mu,\omega))\|_{\rho-6\delta} &\leq \|\partial_1\, u\,(\mu_0,h_0)\|_\rho + \hat{\kappa}_7\,\delta^{-\sigma}\,\varepsilon\ \leq \|\partial_1\, u\,(\mu_0,h_0)\|_\rho + \hat{\kappa}_7\,, \\ \|\partial_\theta h(\mu,\omega)(\theta)\,\|_{\rho-7\delta}\ &\leq \|\partial_\theta h_0(\theta)\,\|_\rho\ +\ \kappa_5\,\delta^{-(\sigma+1)}\,\varepsilon\ \leq \|\partial_\theta h_0(\theta)\,\|_\rho\ +\ \kappa_5\,. \end{aligned} \tag{45}$$

Lemma 9 follows from (43), (44), and (45). □

Lemma 10. *Let κ_6 be as in Lemma 9 and assume that (37) holds. Then there exists a constant κ_8, depending on κ_6, η_0, and $\tilde{\eta}_0$ such that if $\varepsilon \leq \varepsilon_1$ and*

$$\kappa_8\,\delta^{-(\sigma+1)}\,\varepsilon\ \leq\ \min\,(1,\ \varrho)\ , \tag{46}$$

then for each $(x,y) \in B(\mu_0,\omega_0;\varepsilon)$ the map $H^{\leq 1}(\mu,\omega;x,y)$, defined in Lemma 9, is non-degenerate and the following estimates hold

$$\left\|\,\partial_\theta\, H^{\leq 1}(\mu,\omega;x,y)(\theta)\,\right\|_{\rho-8\delta}\ \leq \|\partial_\theta h_0(\theta)\|_\rho + \varrho\left(2^{-1}+2^{-2}+2^{-3}\right)\ , \tag{47}$$

$$\begin{aligned} \left\|\,\left[\,\partial_\theta\, H^{\leq 1}(\mu,\omega;x,y)(\theta)\,\right]^{-1}\,\right\|_{\rho-8\delta} &\leq \tilde{\eta}_0\left(1+2^{-1}+2^{-2}+2^{-3}\right) \\ \left|\,\mathrm{avg}\left\{\,\left[\,\partial_\theta\, H^{\leq 1}(\mu,\omega;x,y)(\theta)\,\right]^{-1}\,\right\}_\theta^{-1}\right| &\leq \eta_0\left(1+2^{-1}+2^{-2}+2^{-3}\right)\ . \end{aligned} \tag{48}$$

Moreover, for each $(x,y) \in B(\mu_0,\omega_0;\varepsilon)$ with $x \in \Xi$,

$$\left\|\,\mathcal{F}\left(x,y;H^{\leq 1}(\mu,\omega;x,y),\Lambda^{\leq 1}(\mu,\omega;x,y)\right)\,\right\|_{\rho-8\delta}\ \leq \hat{\kappa}_8\ \delta^{-2\,\sigma}\ |\,(x-\mu,x-\omega)\,|^{\,2}\ , \tag{49}$$

where $\hat{\kappa}_8$ depends on d, κ_6^2, γ^{-2}, ϱ, $\|\,\partial_\theta h_0(\theta)\,\|_\rho$, and $\sup\limits_{\mu\in\Xi}\|\,\partial_\theta^2 u_\mu(\theta)\|_{2\,\varrho\,\rho}$.

Proof. Estimates (47) and (48) follow from (42), (36) and Lemma 3 by choosing

$$\kappa_8 \overset{\text{def}}{=} 2^6\,\kappa_6\,\gamma^{-1} \max\left(1, 2\,\tilde{\eta}_0, \eta_0\,\tilde{\eta}_0, 2\,\eta_0\right)\,.$$

Estimate (49) follows from Taylor's Theorem, inequality (42), and Cauchy's inequalities The independence of $\hat{\kappa}_8$ on $\partial_\theta h(\mu,\omega)$ follows from the first inequality in (36). □

Remark 9. *For ε as in Lemma 10. Notice that, since the coefficients $\alpha(\mu,\omega)$ and $\beta(\mu,\omega)$ of $\Lambda^{\leq 1}(\mu,\omega;x,y)$ in Lemma 9 satisfy (38), from Lemma 2 we have that for any $(\mu,\omega)\in\Omega\,(\mu_0,\omega_0;\varepsilon)$*

$$\begin{aligned}
\alpha(\mu,\omega) &= \operatorname{avg}\left\{(\partial_\theta\, h(\mu,\omega)(\theta))^{-1}\right\}_\theta^{-1} \operatorname{avg}\left\{(\partial_\theta\, h(\mu,\omega)(\theta+\omega))^{-1}\,\partial_1 f(\mu, h(\mu,\omega)(\theta))\right\}_\theta,\\
\beta(\mu,\omega) &= -\operatorname{avg}\left\{(\partial_\theta\, h(\mu,\omega)(\theta))^{-1}\right\}_\theta^{-1}\,,
\end{aligned}$$

where we have defined $h(\mu_0,\omega_0)\overset{\text{def}}{=} h_0$. Hence $\beta(\mu,\omega)$ is invertible and

$$\beta(\mu,\omega)^{-1} = -\operatorname{avg}\left\{(\partial_\theta\, h(\mu,\omega)(\theta))^{-1}\right\}_\theta\,.$$

Let $\varepsilon>0$ such that Lemma 10 holds. For $(\mu,\omega)\in\Omega(\mu_0,\omega_0;\varepsilon)$ let $h(\mu,\omega)$ and $\lambda(\mu,\omega)$ be as in Lemma 7 and $a(\mu,\omega)$, $b(\mu,\omega)$, $\alpha(\mu,\omega)$, and $\beta(\mu,\omega)$ be as in Lemma 9. In order to show that $\{H,a,b\}$ and $\{\Lambda,\alpha,\beta\}$ are Whitney-C^2 we have to prove that (7) and (8) in Definition 3 hold. Notice that, if ε satisfies (46), then (7) follows from (40), and (41).

It is not difficult to see that in order to prove (8) for any multi-index $j\in\mathbb{Z}_+^s\times\mathbb{Z}_+^d$ with $|j|_1=1,2$, it is enough to prove that the functions a, b, α, β are Lipschitz and that there exists a constant M such that the following estimates hold, for some $\rho'>0$ and any $(\mu,\omega),(\bar{\mu},\bar{\omega})\in\Omega(\mu_0,\omega_0;\varepsilon)$:

$$\begin{aligned}
\left\|h(\mu,\omega) - H^{\leq 1}(\bar{\mu},\bar{\omega};\mu,\omega)\right\|_{\rho'} &\leq M\,\left|(\mu-\bar{\mu},\omega-\bar{\omega})\right|^2\,,\\
\left|\lambda(\mu,\omega) - \Lambda^{\leq 1}(\bar{\mu},\bar{\omega};\mu,\omega)\right| &\leq M\,\left|(\mu-\bar{\mu},\omega-\bar{\omega})\right|^2\,.
\end{aligned}\tag{50}$$

In order to prove (50) we have to allow changes on the initial phases (see Lemmas 11 and 12 bellow, and also see Remark 1).

Lemma 11. *Let $\varepsilon>0$ be such that Lemma 10 holds, let $h(\mu,\omega)$ and $\lambda(\mu,\omega)$ be as in Lemma 7 and $a(\mu,\omega)$, $b(\mu,\omega)$, $\alpha(\mu,\omega)$, and $\beta(\mu,\omega)$ be as in Lemma 9.*

Then there exists a constant κ_9, depending on d, σ, γ^{-2}, ϱ, η_0, $\tilde{\eta}_0$, $\|\,\partial_\theta h_0(\theta)\,\|_\rho$, and $\sup\limits_{\mu\in\Xi}\|\partial_2^2\,u_\rho\|_{2\,\varrho\,\rho}$ such that if $q\geq 2(\sigma+1)$ and

$$\kappa_9\,\delta^{-(q+2\sigma)}\,\varepsilon^2 \leq \min(\,1,\ \varrho\,)\,,\tag{51}$$

then for each $(\mu,\omega),(\bar{\mu},\bar{\omega})\in\Omega(\mu_0,\omega_0;\varepsilon)$ the following estimate holds

$$\left|\,\lambda(\mu,\omega) - \Lambda^{\leq 1}(\bar{\mu},\bar{\omega};\omega,\mu\,)\,\right| \leq \kappa_9\,\delta^{-2\sigma}\,\left|\,(\bar{\mu}-\mu,\bar{\omega}-\omega)\,\right|^2.\tag{52}$$

Moreover, there exists a vector $\Theta(\omega,\mu;\bar{\mu},\bar{\omega})$ such that if

$$\hat{H}^{\leq 1}(\bar{\mu},\bar{\omega};\mu,\omega)\overset{\text{def}}{=} H(\bar{\mu},\bar{\omega};\mu,\omega)\circ T_{\Theta(\omega,\mu;\bar{\mu},\bar{\omega})}\,,$$

then

$$\left\|\,h(\mu,\omega) - \hat{H}^{\leq 1}(\bar{\mu},\bar{\omega};\mu,\omega)\,\right\|_{\rho-14\delta} \leq \kappa_9\,\delta^{-3\sigma}\,\left|\,(\bar{\mu}-\mu,\bar{\omega}-\omega)\,\right|^2.\tag{53}$$

Proof. First of all notice that Lemma 10 implies that for any $(\mu,\omega),(\bar\mu,\bar\omega)\in\Omega(\mu_0,\omega_0;\varepsilon)$, $\big(\mu,\omega;\ H^{\le 1}(\bar\mu,\bar\omega;\mu,\omega),\Lambda^{\le 1}(\bar\mu,\bar\omega;\mu,\omega)\big)$ is an approximate solution of (13) such that its error is bounded as in (49). Moreover, $H^{\le 1}(\bar\mu,\bar\omega;\mu,\omega)$ is non-degenerate and the hypotheses of Theorem 1 hold. The uniform estimates (47) and (48) and Theorem 1 yield the existence of a constant κ_{10}, depending on d, σ, γ^{-2}, ϱ, η_0, $\tilde\eta_0$, $\|\,\partial_\theta h_0(\theta)\,\|_\rho$, and $\sup_{\mu\in\Xi}\|\partial_2^2\, u_\mu\|_{2\,\varrho\,\rho}$, such that if (51) holds then there exist a function $v^*(\mu,\omega)\in\mathcal{A}\left(\mathbb{T}^d_{\rho-12\delta},\,C^0\right)$ and a vector $\lambda^*(\mu,\omega)\in\mathbb{R}^d$ such that if $h^*(\mu,\omega)\stackrel{\text{def}}{=}B+v^*(\mu,\omega)$, then

$$\mathcal{F}\left(\mu,\omega;h^*(\mu,\omega),\lambda^*(\mu,\omega)\right)=0\,,$$

$$\big\|\,h^*(\mu,\omega)-H^{\le 1}(\bar\mu,\bar\omega;\mu,\omega)\big\|_{\rho-12\delta}\le\ \kappa_{10}\,\delta^{-3\sigma}\ |\,(\bar\mu-\mu,\bar\omega-\omega)\,|^2\,, \tag{54}$$

and

$$\big|\,\lambda^*(\mu,\omega)-\Lambda^{\le 1}(\bar\mu,\bar\omega;\mu,\omega)\big|\le\ \kappa_{10}\,\delta^{-2\sigma}\ |\,(\bar\mu-\mu,\bar\omega-\omega)\,|^2\,. \tag{55}$$

Hence $(\mu,\omega;h(\mu,\omega),\lambda(\mu,\omega))$ and $(\mu,\omega;h^*(\mu,\omega),\lambda^*(\mu,\omega))$ are solutions of (13) obtained from two different initial approximate solutions:

$$\begin{array}{ccc}\big(\mu,\omega;\ H^{\le 1}(\mu_0,\omega_0;\mu,\omega),\Lambda^{\le 1}(\mu_0,\omega_0;\mu,\omega)\big) & & \big(\mu,\omega;\ H^{\le 1}(\bar\mu,\bar\omega;\mu,\omega),\Lambda^{\le 1}(\bar\mu,\bar\omega;\mu,\omega)\big)\\ \Big\downarrow{\scriptstyle\text{Theorem 1}} & & \Big\downarrow{\scriptstyle\text{Theorem 1}}\\ \mathcal{F}\left(\mu,\omega;\ h(\mu,\omega),\lambda(\mu,\omega)\right)=0 & & \mathcal{F}\left(\mu,\omega;\ h^*(\mu,\omega),\lambda^*(\mu,\omega)\right)=0\,.\end{array}$$

We are going to prove now that if ε is sufficiently small, then these solutions are equal (modulo the equivalence class defined in Remark 1). Let us first estimate the difference of the initial approximate solutions: From Lemma 7 and estimates (42) and (8) for any $(\mu,\omega),(\bar\mu,\bar\omega)\in\Omega(\mu_0,\omega_0;\varepsilon)$ we have

$$\big\|\,H^{\le 1}(\mu_0,\omega_0;\mu,\omega)-H^{\le 1}(\bar\mu,\bar\omega;\mu,\omega)\big\|_{\rho-8\delta}\le\ \gamma^{-1}\kappa_{11}\,\delta^{-\sigma}\,\varepsilon\,,$$

and

$$\big|\,\Lambda^{\le 1}(\mu_0,\omega_0;\mu,\omega)-\Lambda^{\le 1}(\bar\mu,\bar\omega;\mu,\omega)\big|\le\ \kappa_{11}\,\varepsilon\,,$$

where

$$\kappa_{11}\stackrel{\text{def}}{=}\kappa_2\,+\kappa_5+2\,\kappa_6$$

with κ_2, κ_5, κ_6 as in Lemma 7, Lemma 8, and Lemma 9 respectively.

Hence if $\kappa_9\ \ge\ \max\left(\kappa_{11}\,\gamma^{-1},\kappa_{10}\right)$ and condition (51) holds, then Theorem 2 implies $\lambda(\bar\mu,\bar\omega)=\lambda^*(\bar\mu,\bar\omega)$ and the existence of an initial phase $\Theta(\mu,\omega;\bar\mu,\bar\omega)$ such that

$$\big\|\,h(\mu,\omega)\ -\ h^*(\mu,\omega)\circ T_{\Theta(\mu,\omega;\bar\mu,\bar\omega)}\,\big\|_{\rho-14\sigma}=0\,.$$

Therefore (53) and (52) follow from (54) and (55). ∎

Now we prove that the functions a, b, α, and β defined in Lemma 9 are Lipschitz on $\Omega\,(\mu_0,\omega_0;\varepsilon)$, with $\varepsilon>0$ sufficiently small.

Lemma 12. *Assume that $\varepsilon>0$ is such that Lemma 11 holds. Let Υ be as in Theorem 3, and let κ_9, $\hat{H}^{\leq 1}(\bar{\mu},\bar{\omega},x,y)$ and $\Theta(\mu,\omega;\bar{\mu},\bar{\omega})$ be as in Lemma 11. Define*

$$\hat{a}\,(\bar{\mu},\bar{\omega}) \stackrel{\text{def}}{=} a(\bar{\mu},\bar{\omega})\circ T_{\Theta(\mu,\omega;\bar{\mu},\bar{\omega})}\,,$$
$$\hat{b}\,(\bar{\mu},\bar{\omega}) \stackrel{\text{def}}{=} a(\bar{\mu},\bar{\omega})\circ T_{\Theta(\mu,\omega;\bar{\mu},\bar{\omega})}\,.$$

Then there exists a constant κ_{12}, depending on the same variables as κ_9 and on Υ, such that for any $(\mu,\omega),(\bar{\mu},\bar{\omega})\in\Omega\,(\mu_0,\omega_0;\varepsilon)$ the following estimates hold:

$$\begin{aligned}\|\,a(\mu,\omega)-\hat{a}(\bar{\mu},\bar{\omega})\,\|_{\rho-16\delta} &\leq \kappa_{12}\,\delta^{-(2\sigma+1)}\,\left|\,(\mu-\bar{\mu},\omega-\bar{\omega})\,\right|\,,\\ \left\|\,b(\mu,\omega)-\hat{b}(\bar{\mu},\bar{\omega})\,\right\|_{\rho-16\delta} &\leq \kappa_{12}\,\delta^{-(2\sigma+1)}\,\left|\,(\mu-\bar{\mu},\omega-\bar{\omega})\,\right|\,,\end{aligned}\tag{56}$$

$$\begin{aligned}|\,\alpha(\mu,\omega)-\alpha(\bar{\mu},\bar{\omega})| &\leq\ \kappa_{12}\,\left|\,(\mu-\bar{\mu},\omega-\bar{\omega})\,\right|\,,\\ |\,\beta(\mu,\omega)-\beta(\bar{\mu},\bar{\omega})\,| &\leq\ \kappa_{12}\,\left|\,(\mu-\bar{\mu},\omega-\bar{\omega})\,\right|\,.\end{aligned}\tag{57}$$

Proof. Performing some computations we obtain

$$D_2\mathcal{F}(\mu,\omega;h(\mu,\omega),\lambda(\mu,\omega))\,[\,a(\mu,\omega)-\hat{a}(\bar{\mu},\bar{\omega}),\alpha(\mu,\omega)-\alpha(\bar{\mu},\bar{\omega})\,]\;=\mathcal{R}_1\,(\mu,\omega;\bar{\mu},\bar{\omega})\;,\tag{58}$$

and

$$D_2\mathcal{F}(\mu,\omega;h(\mu,\omega),\lambda(\mu,\omega))\left[\,b(\mu,\omega)-\hat{b}(\bar{\mu},\bar{\omega}),\beta(\mu,\omega)-\beta(\bar{\mu},\bar{\omega})\,\right]\;=\mathcal{R}_2\,(\mu,\omega;\bar{\mu},\bar{\omega})\;,\tag{59}$$

where

$$\begin{aligned}\mathcal{R}_1\,(\mu,\omega;\bar{\mu},\bar{\omega}) \stackrel{\text{def}}{=} &\left[\,\partial_2 u\left(\mu,\hat{H}^{\leq 1}(\bar{\mu},\bar{\omega};\mu,\omega)\right)\,-\partial_2 u(\mu,h(\mu,\omega))\right]\,\hat{a}(\bar{\mu},\bar{\omega})+\\ &+\,\left[\,\partial_1 u\left(\mu,\hat{H}^{\leq 1}(\bar{\mu},\bar{\omega};\mu,\omega)\right)\,-\partial_1 u(\mu,h(\mu,\omega))\right]\,+\\ &+\,\partial_x\,\left[\,\mathcal{G}(x,y)\,\right]_{(x,y)=(\mu,\omega)}\;,\end{aligned}$$

and

$$\begin{aligned}\mathcal{R}_2\,(\mu,\omega;\bar{\mu},\bar{\omega}) \stackrel{\text{def}}{=} &\left[\,\partial_2 u\left(\mu,\hat{H}^{\leq 1}(\bar{\mu},\bar{\omega};\mu,\omega)\right)\,-\partial_2 u(\mu,h(\mu,\omega))\right]\,\hat{b}(\bar{\mu},\bar{\omega})+\\ &+\,\left[\,\partial_\theta h(\mu,\omega)\circ T_\omega-\partial_\theta\hat{H}^{\leq 1}(\bar{\mu},\bar{\omega};\mu,\omega)\circ T_\omega\right]\,+\\ &+\,\partial_y\,\left[\,\mathcal{G}(x,y)\,\right]_{(x,y)=(\mu,\omega)}\;,\end{aligned}$$

where for any $(x,y)\in B(\mu,\omega_0;\varepsilon)$ with $x\in\Xi$

$$\mathcal{G}(x,y)\stackrel{\text{def}}{=}\mathcal{F}(x,y;H^{\leq 1}(\mu,\omega;x,y),\Lambda^{\leq 1}(\mu,\omega;x,y))-\mathcal{F}(x,y;\hat{H}^{\leq 1}(\bar{\mu},\bar{\omega};x,y),\Lambda^{\leq 1}(\bar{\mu},\bar{\omega};x,y))\,.$$

Now we estimate the right hand side of (58) and (59). Performing some computations, using that $\hat{a}\,(\bar{\mu},\bar{\omega})$, $\hat{b}\,(\bar{\mu},\bar{\omega})$, $\alpha\,(\bar{\mu},\bar{\omega})$ and $\beta\,(\bar{\mu},\bar{\omega})$ satisfy (38) and using estimates (40), (41), and (42) one obtains that there exists a constant κ_{13}, depending on Υ, γ^{-1}, κ_6 (with κ_6 is as in Lemma 9) such that

$$\left\| \partial_x \left[\mathcal{F}\left(x, y; \hat{H}^{\leq 1}(\bar{\mu},\bar{\omega}; x, y), \Lambda^{\leq 1}(\bar{\mu},\bar{\omega}; x, y) \right) \right]_{(x,y)=(\mu,\omega)} \right\|_{\rho-14\delta} \leq \kappa_{13}\delta^{-2\sigma} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right| ,$$

and

$$\left\| \partial_y \left[\mathcal{F}\left(x, y; \hat{H}^{\leq 1}(\bar{\mu},\bar{\omega}; x, y), \Lambda^{\leq 1}(\bar{\mu},\bar{\omega}; x, y) \right) \right]_{(x,y)=(\mu,\omega)} \right\|_{\rho-14\delta} \leq \kappa_{13}\,\delta^{-2\sigma} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right| .$$

Then Lemma 10 implies

$$\left\| \partial_x \left[\mathcal{G}(x,y) \right]_{(x,y)=(\mu,\omega)} \right\|_{\rho-14\delta} \leq \kappa_{13}\,\delta^{-2\sigma} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right| ,$$

$$\left\| \partial_y \left[\mathcal{G}(x,y) \right]_{(x,y)=(\mu,\omega)} \right\|_{\rho-14\delta} \leq \kappa_{13}\,\delta^{-2\sigma} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right| ,$$

Moreover, Lemma 11 implies for $i = 1, 2$

$$\left\| \partial_i u \left(\mu, \hat{H}^{\leq 1}\,(\bar{\mu},\bar{\omega}; \mu,\omega) \right) - \partial_i u\,(\mu, h(\mu,\omega)) \right\|_{\rho-14\delta} \leq \Upsilon\,\kappa_9\,\delta^{-3\sigma} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right|^2 ,$$

and

$$\left\| \partial_\theta h(\mu,\omega) - \partial_\theta \hat{H}^{\leq 1}\,(\bar{\mu},\bar{\omega}; \mu,\omega) \right\|_{\rho-15\delta} \leq \kappa_9\,\delta^{-(3\sigma+1)} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right|^2 .$$

Hence if $\varepsilon > 0$ also satisfies condition (37) we have

$$\| \mathcal{R}_1\,(\mu,\omega; \bar{\mu},\bar{\omega}) \|_{\rho-15\delta} \leq \kappa_{14}\,\delta^{-2\sigma} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right| \tag{60}$$

and

$$\| \mathcal{R}_2\,(\mu,\omega; \bar{\mu},\bar{\omega}) \|_{\rho-15\delta} \leq \kappa_{14}\;\delta^{-2\sigma} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right| , \tag{61}$$

where κ_{14} is a constant depending on Υ, κ_6, κ_9, and κ_{13}.

Lemma 2 and estimates (60), and (61) imply

$$\begin{aligned} \| a(\mu,\omega) - \hat{a}(\bar{\mu},\bar{\omega}) \|_{\rho-16\delta} &\leq \tilde{\kappa}_{14}\,\delta^{-(2\sigma+1)} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right| + \left| \varphi_1\,(\mu,\omega; \bar{\mu},\bar{\omega}) \right| \\ \left\| b(\mu,\omega) - \hat{b}(\bar{\mu},\bar{\omega}) \right\|_{\rho-16\delta} &\leq \tilde{\kappa}_{14}\,\delta^{-(2\sigma+1)} \left| (\mu-\bar{\mu}, \omega-\bar{\omega}) \right| + \left| \varphi_2\,(\mu,\omega; \bar{\mu},\bar{\omega}) \right| , \end{aligned} \tag{62}$$

where

$$\begin{aligned} \varphi_1\,(\mu,\omega; \bar{\mu},\bar{\omega}) &\stackrel{\text{def}}{=} \text{avg} \left\{ \left[\partial_\theta h(\mu,\omega)(\theta) \right]^{-1} \left[a(\mu,\omega) - \hat{a}(\bar{\mu},\bar{\omega}) \right] \right\}_\theta \\ \varphi_2\,(\mu,\omega; \bar{\mu},\bar{\omega}) &\stackrel{\text{def}}{=} \text{avg} \left\{ \left[\partial_\theta h(\mu,\omega)(\theta) \right]^{-1} \left[b(\mu,\omega) - \hat{b}(\bar{\mu},\bar{\omega}) \right] \right\}_\theta . \end{aligned}$$

Notice that from condition (39) and estimates (41) in Lemma 9, for $i = 1, 2$ we have

$$|\,\varphi_i\,| \le \left\| \,[\,\partial_\theta h(\mu,\omega)(\theta)\,]^{-1} - \left[\,\partial_\theta \hat{h}(\bar{\mu},\bar{\omega})(\theta)\,\right]^{-1} \right\|_0 \kappa_6\,\gamma^{-1}\,\delta^{-\sigma}\,,$$

where $\hat{h}(\bar{\mu},\bar{\omega}) \stackrel{\text{def}}{=} h(\bar{\mu},\bar{\omega}) \circ T_{\Theta(\mu,\omega;\bar{\mu},\bar{\omega})}$. Using condition (37) and estimates (42) and (53) we have

$$\left\| \partial_\theta h(\mu,\omega)(\theta) \, - \partial_\theta \hat{h}(\bar{\mu},\bar{\omega})(\theta) \,\right\|_0 \le (\kappa_9 + \kappa_2)\;\delta^{-(\sigma+1)}\;|\,(\mu-\bar{\mu},\omega-\bar{\omega})\,|\;,$$

from which we have

$$\left\| \,[\,\partial_\theta h(\mu,\omega)(\theta)\,]^{-1} - \left[\,\partial_\theta \hat{h}(\bar{\mu},\bar{\omega})(\theta)\,\right]^{-1} \right\|_0 \le\; \kappa_{15}\,\delta^{-(\sigma+1)}\;|\,(\mu-\bar{\mu},\omega-\bar{\omega})\,|\;,$$

where thanks to second inequality in (36), the constant κ_{15} depends on $\left\|[\partial_\theta h_0(\theta)]^{-1}\right\|$, $\tilde{\eta}_0$, κ_9, γ^{-1}, and κ_6, but it does not depend either on (μ,ω) or $(\bar{\mu},\bar{\omega})$.

Hence for $i = 1, 2$ we have

$$|\,\varphi_i\,| \le \kappa_{16}\,\delta^{-(2\sigma+1)}\;|\,(\mu-\bar{\mu},\omega-\bar{\omega})\,|\;. \tag{63}$$

Estimates (56) and (57) follow from (58), (59), (60), (61), (62), and (63), and Lemma 2. The constant κ_{15} depends on d, σ, γ^{-1}, $\|\partial_\theta h_0(\theta)\|_\rho$, ϱ, κ_{14}, and κ_{16} but independent of (μ,ω) and $(\bar{\mu},\bar{\omega})$. □

The Whitney regularity stated in Theorem 3 follows from (40), (41), (52), (53), (56), and (57).

Acknowledgement

I am deeply grateful to Prof. R. de la Llave who taught me the techniques used in this paper. I also thank R. Vitolo for many discussions on the subject.

References

[Arn65] V. I. Arnol'd. Small denominators. I. mappings of the circumference onto itself. *Amer. Math. Soc. Transl. Ser. 2*, 46:213–284, 1965. English translation: *Amer. Math. Sos. Transl. (2)*, 46:213–284, 1965.

[BHS96a] H. W. Broer, G. B. Huitema, and M. B. Sevryuk. Families of quasi-periodic motions in dynamical systems depending on parameters. In *Nonlinear dynamical systems and chaos (Groningen, 1995)*, pages 171–211. Birkhäuser, Basel, 1996.

[BHS96b] H. W. Broer, G. B. Huitema, and M. B. Sevryuk. *Quasi-Periodic Motions in Families of Dynamical Systems.* Order Amidst Chaos. Springer-Verlag, Berlin, 1996.

[BHTB90] H. W. Broer, G. B. Huitema, F. Takens, and B. L. J. Braaksma. Unfoldings and bifurcations of quasi-periodic tori. *Mem. Amer. Math. Soc.*, 83(421):viii+175, 1990.

[dlL01] Rafael de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001.

[dlLGE] R. de la Llave and A. González-Enríquez. Whitney differentiability and condensation of KAM tori for systems not written in action-angle variables. *Manuscript*.

[dlLGJV05] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.

[GEV05] A. González-Enríquez and J. Vano. An estimate of smoothing and composition with applications to conjugation problems. *Preprint,* `http://www.ma.-utexas.edu/mp_arc 05-418` , 2005.

[Ham82] R. S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1):65–222, 1982.

[Her83] M.-R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1*, volume 103 of *Astérisque*. Société Mathématique de France, Paris, 1983.

[Mos66a] J. Moser. On the theory of quasiperiodic motions. *SIAM Rev.*, 8:145–172, 1966.

[Mos66b] J. Moser. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:499–535, 1966.

[Mos66c] J. Moser. A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:265–315, 1966.

[Mos67] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169:136–176, 1967.

[Rüs75] H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical Systems, Theory and Applications (Battelle Rencontres, Seattle, Wash., 1974)*, pages 598–624. Lecture Notes in Phys., Vol. 38, Berlin, 1975. Springer.

[Sal86] D. Salamon. The Kolmogorov-Arnold-Moser theorem. *Zürich preprint*, 1986.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions.* Princeton University Press, Princeton, N.J., 1970. Princeton Mathematical Series, No. 30.

[Whi34] H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.*, 36(1):63–89, 1934.

[Zeh75] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.