CONSTRUCTION OF QUASI-PERIODIC SOLUTIONS OF STATE-DEPENDENT DELAY DIFFERENTIAL EQUATIONS BY THE PARAMETERIZATION METHOD II: ANALYTIC CASE

XIAOLONG HE AND RAFAEL DE LA LLAVE

Abstract. We construct analytic quasi-periodic solutions of state-dependent delay differential equations with quasi-periodically forcing. We show that if we consider a family of problems that depends on one dimensional parameters(with some non-degeneracy conditions), there is a positive measure set Π of parameters for which the system admits analytic quasi-periodic solutions.

The main difficulty to be overcome is the appearance of small divisors and this is the reason why we need to exclude parameters. Our main result is proved by a Nash-Moser fast convergent method and is formulated in the a-posteriori format of numerical analysis. That is, given an approximate solution of a functional equation which satisfies some non-degeneracy conditions, we can find a true solution close to it.

This is in sharp contrast with the finite regularity theory developed in [HdlL15]. We conjecture that the exclusion of parameters is a real phenomenon and not a technical difficulty. More precisely,for generic families of perturbations, the quasi-periodic solutions are only finitely differentiable in open sets in the complement of parameters set Π.

1. Introduction

The goal of this paper is to investigate analytic properties of solutions of state-dependent delay differential equations(SD-DDEs).

For functional differential equations with constant delay, [Nus73] proves that for a broad class of analytic equations, many solutions are analytic. Analytic solutions for differential equations with time-varying delay(independent of the state) have been considered before in [LS07, MPN14]. In both papers, an important technique is to conjugate the delay function to a simple

---

Date: November 9, 2015.

2010 *Mathematics Subject Classification.* 70K43, 34K06, 34K27, 34A30, 37J40 .

Key words and phrases. State-dependent delay; parameterization method; analytic solution; foliation-preserving torus map; KAM theory; a-posteriori.

X.He supported by CSC (File No.201406130018) and NSFC(Grant NO. 11271115) .

R. de la Llave supported by NSF grant DMS-1500943.

form(either a linear multiple or a constant rotation). In the reduction to constant delay[MPN14], the dynamical properties of the delay mapping play an important role. In particular, they require to use the theory of smooth conjugacy to rotations[Arn63, Her80]. Hence, small divisors play a role.

As listed in open problems section in [MPN14], the case of state-dependent delay equations is significantly more complicated. Nevertheless, as [MPN14] observes, the time dependent delay can be considered as a linearization version of the theory. Recently, [Kri14] considers a state-dependent delay differential equation in which the delay is defined by the threshold condition. It shows that the globally defined bounded solutions are analytic. We also call attention to [SB03, Remark 3], which presents a KAM theorem for state-dependent delay differential equations on the torus.

In this paper, we look for analytic quasi-periodic solutions of SD-DDEs by the parameterization method[CFdlL03, CFdlL05]. To avoid technical difficulties, we focus on the following scalar quasi-periodic differential equation with state-dependent delay, which contains all essential difficulties

$$\begin{cases} \dot{x}(t) = ax(t) + bx(t - r_\mu(x(t))) + f(\theta) \\ \dot{\theta}(t) = \omega \end{cases} \tag{1.1}$$

where $x \in \mathbb{R}$, $\theta \in \mathbb{T}^d = \mathbb{R}^d/2\pi\mathbb{Z}^d$ and the frequency $\omega \in \mathbb{R}^d$ is rationally independent(i.e. $\omega \cdot k \neq 0$ for $k \in \mathbb{Z}^d - \{0\}$). Later we will need to impose quantitative Diophantine properties on the frequency ω. The *nonhomogeneous* term f is a real analytic periodic function defined on the complex domain

$$\mathcal{D}_\rho = \{\theta = (\theta_1, \cdots, \theta_d) \in \mathbb{C}^d : |\Im\mathrm{m}(\theta_j)| \leq \rho,\ 1 \leq j \leq d\}, \tag{1.2}$$

which means it is analytic in its interior and can be extended continuously to the boundary. We also denote the neighborhood of $\mathbb{T}^d$ with width ρ in the complex space by $\mathbb{T}^d_\rho$, i.e.

$$\mathbb{T}^d_\rho = \{\theta \in \mathbb{C}^d/2\pi\mathbb{Z}^d : |\Im\mathrm{m}(\theta_j)| \leq \rho\}. \tag{1.3}$$

Then we say that f is defined and analytic on $\mathbb{T}^d_\rho$ for some $\rho > 0$. We also introduce the parameter $\mu \in \mathbb{R}$ into our system. Note that (1.1) is a extremely nonlinear equation because the unknown function x appears composed with itself.

Our main result, Theorem 3.1, asserts the existence of analytic quasi-periodic solutions. We require introducing one extra one dimensional parameter and perform *parameter exclusion*. We prove that the set of parameters on which admits analytic quasi-periodic solutions is of positive measure.

It is worth to compare this paper with [HdlL15]. Both papers consider similar problems, but the methods are very different and the results are complementary. Hence the overlap is minimal. The paper [HdlL15] uses a theory of exponential dichotomy in finite regularity setting. Even if it produces finite regularity solutions, it does not need to exclude parameters.

Combining Theorem 3.4 with the results in [HdlL15], we have that given an approximate solution in an one parameter family of problems, we can obtain C^r solutions for all values of the parameter and obtain a positive measure (with open dense complement) set Π of parameters for which the solution is analytic.

We conjecture that this result is essentially sharp in the following sense.

Conjecture 1. *Consider a family of quasi-periodic state-dependent delay differential equations*

$$x'(t) = f(\omega t, x(t), x(t - r_\mu(x(t)))),$$

where f is analytic in its arguments and μ is the parameter.

Assume that for μ_0, we have an analytic quasi-periodic solution and the family of delay functions r_μ satisfy some non-degeneracy conditions(which hold in open sets of their arguments).

Then, there is a sequence of intervals I_j accumulating at μ_0, and a sequence of integers r_j, such that when $\mu \in I_j$, then the K_μ produced in [HdlL15] *is C^{r_j} but not C^{r_j+1}.*

The reason for this conjecture is the observation in [PdlLV03] that a family of delay mappings which satisfies some non-degeneracy conditions will experience phase locking intervals in which there is an exponentially attracting invariant torus of lower dimension. The conditions for these phenomenon are very explicit and can be verified in concrete examples with a finite calculation. Furthermore, these conditions hold in open sets.

These are the generalizations to high dimension of the well-known Arnold tongues in circle mappings. Once one has the existence of an set with nonzero Lyapunov exponents in the delay mapping dynamics, it seems that the argument of [Fen74] will be able to produce that the mapping K is only finitely differentiable.

This phenomenon has been observed in the theory of normally hyperbolic invariant manifolds [HCF+15], where the manifold is analytic if the motion on it is analytically conjugated to an irrational solution and it is only finite differentiable if it contains periodic orbits with a positive Lyapunov exponent. See also [Fen74, Fen77].

1.1. **Motivation for the procedure.**

In the light of parameterization method [CFdlL03, CFdlL05], we look for an embedding $K : \mathbb{T}^d \to \mathbb{R}$ such that

$x(t) = K(\theta + \omega t)$ is a solution of (1.1). It is immediate that $x(t) = K(\theta + \omega t)$ is a solution of (1.1) if and only if K satisfies

$$\partial K(\theta) \cdot \omega = aK(\theta) + bK(\theta - \omega r_\mu(K(\theta))) + f(\theta). \tag{1.4}$$

Instead of treating (1.1) as an evolution equation, we just work on function spaces and find a solution of the functional equation (1.4). An important source of difficulties is the study of the composition operator that to a function K associates $\Psi[K,\mu](\theta) = K(\theta - \omega r_\mu(K(\theta)))$.

If there was any hope to apply a fixed point theorem in a space of analytic functions defined on a domain $\mathcal{U} \subset \mathbb{C}^d$, it would be necessary to have that the range of the composition would be this domain, i.e.

$$\mathcal{U} = \{\theta - \omega r_\mu(K(\theta)) : \theta \in \mathcal{U}\}. \tag{1.5}$$

The condition (1.5) is very hard to satisfy, especially if K is changing.

One case, however in which (1.5) is satisfied is when the diffeomorphism of the torus

$$\varphi[K,\mu] \equiv Id - \omega r_\mu \circ K \tag{1.6}$$

is conjugated to a rotation R_Ω on the torus (see (A.4)). We call the map φ the *delay mapping*.

More precisely, if there exists an analytic function H such that

$$\varphi \circ H = H \circ R_\Omega. \tag{1.7}$$

Then, taking $\mathcal{U} = H(\mathbb{T}^d_\rho)$, we have

$$\varphi(\mathcal{U}) = \varphi(H(\mathbb{T}^d_\rho)) = H(R_\Omega(\mathbb{T}^d_\rho)) = H(\mathbb{T}^d_\rho) = \mathcal{U}.$$

Remark 1.1. *Hence, we can think of* (1.7) *as a way of finding the domain $\mathcal{U}$ where to study φ. It is a result in* [Mos01, Theorem 2.1] *that if a mapping φ sends a domain to itself it is conjugated to a rotation. So the reduction to a rotation to obtain invariance of the domain is a natural procedure.*

Remark 1.2. *From* [SB03, MPN14]*, we also know that the dynamics properties of the delay mapping play an important role in determining the analyticity and non-analyticity of the invariant objects of a differential equation with time varying delay.*

Given K and μ, we observe that the delay mapping φ is a *foliation-preserving torus map*(also called *reparameterization of an irrational flow* [Fay02]), which is a map of form

$$\varphi(\theta) = \theta + \omega\hat{\varphi}(\theta) \tag{1.8}$$

defined on the torus. Here $\hat{\varphi} : \mathbb{T}^d \to \mathbb{R}$ is a scalar function. It is easy to see that the torus maps of the form (1.8) form a group. For more details, see Appendix A.

In the following, we assume the delay function $r_0(z)$ is a constant(i.e. independent of z) so that when $\mu = 0$, equation (1.1) becomes just a constant delay equation, whose existence of quasi-periodic solutions has been well-studied in [LdlL09]. Without loss of generality, we further assume that $r_0 = -1$. Otherwise, we just replace ω by $r_0\omega$, which also satisfies the Diophantine affine condition required in Theorem 3.1.

Geometrically, for sufficient small μ, the delay mapping φ is close to a rotation. From [PdlLV03, Theorem 3.2], we know that there is a near-identity mapping $H : \mathbb{T}^d \to \mathbb{T}^d$ conjugating a foliation-preserving torus map into a rigid rotation with a slight frequency shift $\alpha\omega$. In Appendix A, we present a more quantitative version.

It also should be kept in mind that the transformation H preserves the ω-foliation, i.e.

$$H = Id + \omega h \tag{1.9}$$

where h is a periodic scalar function. With these motivations, especially the particular form of H and the frequency shift in the direction of ω, one will see that the reduction of conjugation equation defined in (1.13) is essentially a one dimensional problem.

1.2. **Formulation of the functional equations.**

We define the operator $\mathscr{F}$ as

$$\mathscr{F}[K,\mu] \equiv \partial K(\theta)\cdot\omega - aK(\theta) - bK\circ\varphi[K,\mu] - f(\theta). \tag{1.10}$$

As indicated above, a zero solution of $\mathscr{F}$ will give a quasi-periodic solution of frequency ω.

Following the motivation in Section 1.1, we supplement the functional equation $\mathscr{F} = 0$ with another auxiliary equation, whose purpose is to make possible to solve the Newton equation for $\mathscr{F}$.

To this end, we simultaneously consider the conjugation problem and define the operator $\mathscr{G}$ as

(1.11)
$$\begin{aligned}\mathscr{G}[K,h,\mu;\alpha] \equiv&\varphi[K,\mu]\circ H - H\circ R_{(\alpha+1)\omega}\\ =&\omega h - \omega h\circ R_{(\alpha+1)\omega} - (\alpha+1)\omega - \omega r_\mu\circ K\circ(Id+\omega h).\end{aligned}$$

Then, we look for solutions of the functional equations

$$\mathscr{F}[K,\mu] = 0, \tag{1.12}$$

$$\mathscr{G}[K,h,\mu;\alpha] = 0. \tag{1.13}$$

We call (1.12) the *invariance* equation and (1.13) the *conjugation* equation.

Even if our main goal is only the invariance equation (1.12), it is remarkable that solving the two euqations is easier that solving just one. Of course, the fact that the equation (1.13) gives us more information. We note that in [HdlL15] there is a procedure to solve (1.12) along but as indicated in Remark 1.1, we do not expect that the solutions in [HdlL15] are always analytic. See Conjecture 1.

Of course, the functional that to a map of the torus associates the conjugacy is rather delicate and one cannot solve it by elementary methods. As a result, one has to resort to KAM methods and has to have external parameters [Arn63, Mos66a, Mos67]. Hence, the procedure we will implement is a Newton iteration for the invariance equation and the conjugation equation. We will consider fixed frequency vector ω and real number α. The unknowns will be thus the parameterization K, the conjugacy to the rotation h and the parameter μ. As it turns out, these two equations(for three unknowns) (1.12) and (1.13) are coupled, but we will be able to solve them.

Note that solving the linearization equation of (1.13) leads to small divisor problem, so that we will have to use a rapidly convergent method to ensure that the iterative process converges. The KAM method has been used in delay differential equations but in a very different way. Typically, the KAM method was used for constant delay equations treated as a dynamical system (for example [FB76, LdlL09, LY12]). A case where the parameterization method was used was [LdlL09] which uses the dynamical interpretation (but only in a first preparatory stage). The KAM method developed in [SB03, Remark 3] also applies to state-dependent delay equations on the torus, which reduces the delay equation to an integral form. We note, however, that [SB03] requires the same dimensional frequency as the torus.

Along the way, we will also have to solve several other difficulties. Notably, the ∂_ω operator does not work well with the composition. Besides, to establish the non-degeneracy condition, we have to take full advantage of the structure of φ as was done in [Van02, PdlLV03].

Our main goal is to seek analytic quasi-periodic solutions of (1.1). Obviously, in the case of $\mu = 0$, they are easy to find. See the formulation of Lindstedt series in Section 4. We want to show that for the majority μ in a neighborhood of zero, there also exist analytic quasi-periodic solutions.

Our results are given in an *a posteriori* format, which plays an important role in numerical analysis. More precisely, given an approximate solution for (1.1) satisfying some non-degeneracy conditions, there is a true solution nearby. For example, we can take as the approximate solutions in the assumptions of Theorem 3.1 the result of a numerical (non-rigorous) calculation. Then, Theorem 3.1 validates the calculations since it shows that there is a true solution nearby. We point out that careful calculations of solutions which could be used as inputs of an a-posteriori theorem already exist in the literature [HDMU12, MKW14, HCHS14].

Unfortunately, the method in this paper is restricted to having just one delay. Indeed, if we transform the variables so that one φ corresponding to a delay becomes a rotation, the terms corresponding to other delays will not in general be simplified.

This paper is organized as follows. In Section 2, we collect some standard definitions and recall some properties, which serve to set the notations. This section could be omitted at a first reading and used as a reference. In Section 3, we present our main results on the existence of analytic quasi-periodic solutions of (1.1) in an a-posteriori format. As an immediate consequence, we also prove that the obtained solutions admit C^1-Whitney regularity, which implies the set of μ validating Theorem 3.1 has positive measure. In Section 4, we formulate a power series of α satisfying the Lindstedt property. In Section 5, we give a detailed analysis on the Newton equations of the functional equations (1.12) and (1.13). With different coordinates, we show a delicate estimate on the loss of analyticity domain to construct a rapidly convergence sequence. In Appendix A, we give a brief introduction to the foliation-preserving torus map, which provides a background and enables us to consider the conjugacy problem.

2. Preliminaries

To formulate the KAM results, we need to introduce the families of Banach spaces on which the operators $\mathscr{F}$ and $\mathscr{G}$ are defined. Hence, we collect several definitions and properties of different spaces of analytic functions in Section 2.1. When studying the conjugation equations, one encounters with the small divisors problem. As standard KAM theory, in Section 2.2, we give the definition of Diophantine conditions and the estimates on the solution of cohomology equations.

Following the notations in KAM theory, we denote the universal constant by C, depending on the dimension of the space, the width of analytic domain, the Sobolev index and the Diophantine exponents. The value of C may be different from line to line.

2.1. **Function spaces.**

Throughout this paper, we will work on analytic function spaces endowed with two different norms-the classical supremum norm and Sobolev norm.

2.1.1. *Supremum norm.*

Recalling the complex domain $\mathcal{D}_\rho$ defined in (1.2), we denote by $\mathcal{A}_\rho$ the Banach space of analytic functions on $\mathcal{D}_\rho$ endowed with the norm

$$\|u\|_\rho = \sup\{|u(\theta)| : \theta \in \mathcal{D}_\rho\}. \tag{2.1}$$

Then by the Cauchy estimate, it is readily seen, if $u \in \mathcal{A}_\rho$, the partial derivative of u with respect to θ_j satisfies

$$\|u_{\theta_j}\|_{\rho-\sigma} \leq \frac{\|u\|_\rho}{\sigma}$$

for all $0 < \sigma < \rho$ and $1 \leq j \leq d$. Furthermore, if u is periodic, it can be expanded into Fourier series

$$u(\theta) = \sum_{k \in \mathbb{Z}^d} \widehat{u}_k e^{2\pi i k\cdot\theta},$$

whose Fourier coefficients $\widehat{u}_k$ satisfy

$$|\widehat{u}_k| \leq \|u\|_\rho e^{-2\pi|k|\rho} \tag{2.2}$$

for all $k \in \mathbb{Z}^d$. Another useful property is the interpolation inequality (Hadamard three circle theorem)

$$\|u\|_{\theta\rho+(1-\theta)\rho'} \leq \|u\|_\rho^\theta \|u\|_{\rho'}^{1-\theta}. \tag{2.3}$$

Given an analytic function ψ mapping $\mathcal{D}_\rho$ to $\mathbb{C}^d$, we denote its range by

$$\mathcal{D}_{\psi,\rho} = \{\psi(\theta) : \theta \in \mathcal{D}_\rho\}. \tag{2.4}$$

2.1.2. *Sobolev norm.* We also find it convenient to introduce norms of analytic functions which can be read off the Fourier coefficients. Since some of our steps are formulated in in Fourier space, these norms will allow us to obtain sharp results and formulate steps as fixed point arguments. For the details, we refer to [CCCdlL15] and the references therein. For simplicity, only the even Sobolev exponents are considered.

Assume $\rho > 0$ and $s \in 2\mathbb{N}$. Let $\mathcal{A}_{\rho,s}(\mathbb{T}^d)$ (or $\mathcal{A}_{\rho,s}$) be the space of real analytic periodic functions defined on the domain $\mathbb{T}^d_\rho$ such that the norm

$$\|u\|^2_{\rho,s} = \sum_{k\in\mathbb{Z}^d} \frac{e^{4\pi|k|\rho}}{B(k,\rho)}((2\pi)^d|k|^2+1)^s|\widehat{u}_k|^2 < \infty, \tag{2.5}$$

where

$$B(k,\rho) = \Pi_{j=1}^d a(k_j,\rho)$$

with

$$a(j,\rho) = \begin{cases} 4\pi|j|, & \text{if } j \neq 0; \\ 1/4\pi\rho, & \text{if } j = 0. \end{cases}$$

When considering $\mathbb{T}^d_\rho$ as a $2d$ dimensional real manifold with boundary, there is a geometric representation of the Sobolev norm norm given by

$$\|u\|^2_{\rho,s} = \int_{\mathbb{T}^d_\rho} |(\Delta_\theta + 1)^{s/2} u(\theta)|^2 d^{2d}\theta = \int_{\mathbb{T}^d_\rho} |(\sum_{n=1}^d \nabla_{\theta_n}\nabla_{\bar\theta_n} + 1)^{s/2} u(\theta)| d^{2d}\theta \tag{2.6}$$

where the bar denotes the complex conjugation. Indeed, (2.6) can be obtained from the inner product

$$\langle u, v\rangle = \int_{\mathbb{T}^d_\rho} \langle u, (\Delta_\theta + 1)^s v\rangle d^{2d}\theta. \tag{2.7}$$

This makes it clear that $\mathcal{A}_{\rho,s}$ is the closed space of complex analytic functions endowed with the norm corresponded to a real $2d$ dimensional Sobolev space. And we can show that $\mathcal{A}_{\rho,s}$ is complete, since the limit in the $\|\cdot\|_{\rho,s}$ of analytic functions is still analytic. Thus $\mathcal{A}_{\rho,s}$ is a Hilbert space under the inner product (2.7). Furthermore, for $s > d$, the Sobolev space $\mathcal{A}_{\rho,s}$ is a Banach algebra under multiplication.

For convenience, we also state some relations between supremum norm (2.1) and the Sobolev norm (2.5).

If $u \in \mathcal{A}_{\rho'}$ with $\rho' > \rho$, then, by (2.2) and Lemma 2.1, we have

$$\begin{aligned} \|u\|^2_{\rho,s} \leq & \sum_{k\in\mathbb{Z}^d} \frac{\|u\|^2_{\rho'}}{B(k,\rho)}((2\pi)^d|k|^2+1)^s e^{-4\pi|k|(\rho'-\rho)} \\ \leq & \|u\|^2_{\rho'} \,(\max\{1,(4\pi\rho)^d\}) \sum_{k\in\mathbb{Z}^d}(2\pi|k|^2+1)^s e^{-4\pi|k|(\rho'-\rho)} \\ \leq & C(\rho'-\rho)^{-(2s+d)}\|u\|^2_{\rho'} \end{aligned} \tag{2.8}$$

where the constant C depends only on ρ, s and d.

On the other hand, if $u \in \mathcal{A}_{\rho,s}$,

$$\begin{aligned} \|u\|_\rho \leq & \sum_{k\in\mathbb{Z}^d} |\widehat{u}_k| e^{2\pi|k|\rho} \leq C\left\{\sum_{k\in\mathbb{Z}^d}((2\pi)^d|k|^2+1)^{-s}B(k,\rho)\right\}^{1/2} \cdot \|u\|_{\rho,s} \\ \leq & C\left\{\sum_{n\geq 0}[(2\pi)^d n^2+1]^{-s} n^d(2n+1)^d\right\}^{1/2} \cdot \|u\|_{\rho,s} \\ \leq & C\|u\|_{\rho,s} \end{aligned} \tag{2.9}$$

when $s > d + \frac{1}{2}$.

2.2. **Diophantine properties and cohomology equation.** There are two types of Diophantine conditions that appear in KAM theory. The first condition appears when we consider KAM theory for flows and the second one for the maps. See [Rüs75, dlL01] for an introduction.

Definition 2.1. *(Diophantine affine) A vector $\omega \in \mathbb{R}^d$ is called Diophantine affine of type (γ, υ) for positive constant γ and υ if*

$$|k\cdot\omega| \geq \gamma|k|^{-\upsilon} \tag{2.10}$$

holds for all $k \in \mathbb{Z}^d - \{0\}$.

Definition 2.2. *(Diophantine) A vector $\Omega \in \mathbb{R}^d$ is called the Diophantine of type (τ, ν) for positive constants τ and ν if*

$$|k \cdot \Omega - l| \geq \tau |k|^{-\nu} \tag{2.11}$$

holds for all $k \in \mathbb{Z}^d - \{0\}$ and $l \in \mathbb{Z}$.

In this paper, we are interested in the number $\alpha \in \mathbb{R}$ such that $\Omega = \alpha\omega$ is Diophantine. Typically, denoting

$$\mathscr{D}(\tau, \nu; \omega) = \{\alpha \in \mathbb{R} : \alpha\omega \text{ is } \textit{Diophantine of type } (\tau, \nu)\}$$

and

$$\mathscr{D}(\nu; \omega) = \bigcup_{\tau > 0} \mathscr{D}(\tau, \nu; \omega),$$

we have the following lemma on the abundance of such numbers.

Lemma 2.1. ([SdlL12]) *If $\omega \in \mathbb{R}^d$ is Diophantine affine of type (γ, υ) and $\nu > d + \upsilon$, then $\mathscr{D}(\nu; \omega)$ is of full Lebesgue measure.*

The proof in [SdlL12] is just an elementary estimate on the regions where the definition of Diophantine fails. The abundance of Diophantine numbers in a constrained linear space has attracted attention in number theory[Spr79]. The papers Kle01,Kle08 present results with sharper exponents.

The following *cohomology equation* (2.12) is standard in KAM theory, which will be used to solve the conjugation equations in Section 5. A detailed proof can be found in [Rüs75].

Lemma 2.2. *Assume that Ω satisfies* (2.11). *Let $Q \in \mathcal{A}_\rho$ with zero average. Then there is a unique solution W of*

$$W(\theta) - W(\theta + \Omega) = Q(\theta) \tag{2.12}$$

such that W has zero average.

Moreover, we have for all $0 < \sigma < \rho$,

$$\|W\|_{\rho-\sigma} \leq C\tau\sigma^{-\nu}\|Q\|_\rho, \tag{2.13}$$

where the constant C depends only on the Diophantine exponent ν and the dimension of the space.

2.3. **Lindstedt series.** We will study (1.1) perturbatively when the delay mapping is close to a rotation. More precisely, we will consider α as a small parameter and will find $K(\alpha)$, $h(\alpha)$ and $\mu(\alpha)$ solving (1.12) and (1.13) in the sense of formal power series in α. To this end, we give the definition for the Lindstedt series.

Definition 2.3. *(Lindstedt series) Let X_ρ and Y_ρ be a scale of Banach spaces and $\mathscr{A} : \Lambda \times X_\rho \to Y_\rho$. Assume that $\mathscr{A}(\lambda_0, x_0) = 0$ for $x_0 \in X_\rho$ and*

$\lambda_0 \in \Lambda$. *We say that the operator* $\mathscr{A}$ *satisfies the Lindstedt property for every solution* (λ_0, x_0) *if there exist formal power series to order M,* $M \in \mathbb{N} \cup \{\infty\}$*,*

$$x(\lambda) = \sum_{j=0}^{M} x_j(\lambda - \lambda_0)^{\otimes j}$$

with $x_j \in Sym_j(\Lambda, X_{\rho-\delta})$*, such that*

$$||\mathscr{A}(\sum_{j=0}^{N} x_j(\lambda - \lambda_0)^{\otimes j}, \lambda)||_{\rho-2\delta} \leq C(N, \delta)|\lambda - \lambda_0|^{N+1}.$$

for $N < M$*.*

The Lindstedt series method is also used to consider the bifurcation problems for delay differential equations. See [CF80].

3. KAM THEOREM

As stated before, our main technical result Theorem 3.1 is in an *a-posteriori* format. Out of it, we will deduce several consequences which are almost automatic. In Theorem 3.1, we will fix α and obtain that given approximate solutions K_0, h_0, μ_0, there are true solutions. We note that if $\alpha, \tilde{\alpha} \in \mathscr{D}(\nu; \omega)$ are close, the solutions corresponding to α is an approximate solution for $\tilde{\alpha}$. Then, applying Theorem 3.1 we can obtain solutions for $\tilde{\alpha}$. Therefore, we can obtain $K(\alpha), h(\alpha), \mu(\alpha)$ defined in a Cantor-like set of parameters which is of positive measure.

Theorem 3.1. *Let* $0 < \rho < \tilde{\rho}$*,* $\sigma = \rho/48$ *and* $d < s \in 2\mathbb{N}$*. Assume the frequency* ω *satisfies the Diophantine affine condition* (2.10)*.*

We further assume the following:

- *(H1) Regularity conditions: The forcing function* f *is an analytic periodic function in* $\mathcal{A}_\rho$*; The delay function* $r_\mu(z)$ *is analytic on the complex domain* $\mathcal{D}^*$ *for any parameter* μ *in some neighborhood* $O \subset \mathbb{R}$ *of zero and* $\mu \mapsto r_\mu$ *is* C^{N+1} *from* O *to the analytic function space on the analyticity domain* $\mathcal{D}^*$ *endowed with the supremum norm;*
- *(H2) Nearly-rotation:* r_μ *is independent of* $z \in \mathcal{D}^*$ *when* $\mu = 0$*. Without loss of generality, we assume* $r_0 = -1$*.*
- *(H3) Initial guess: Let* (K_0, h_0, μ_0) *be an approximate solution of* (1.12) *and* (1.13) *satisfying* $K_0 \circ H_0 \in \mathcal{A}_{\tilde{\rho}}$ *and* $\mu_0 \in O$*. We assume that*

$$||e||_{\tilde{\rho}}\ ,\ ||E \circ H_0||_{\rho,s} < \epsilon \tag{3.1}$$

where

$$\begin{aligned} E &= \mathscr{F}[K_0, \mu_0],\\ e &= \mathscr{G}[K_0, h_0, \mu_0],\\ H_0 &= Id + \omega h_0. \end{aligned} \tag{3.2}$$

(H4) Diophantine condition: $\alpha + 1 \in \mathscr{D}(\tau, \nu; \omega)$. *Namely,* α *satisfies*

$$|k \cdot (\alpha + 1)\omega - l| \geq \tau |k|^{-\nu}, \quad \forall k \in \mathbb{Z}^d - \{0\}, l \in \mathbb{Z};$$

(H5) Non-degeneracy conditions:

$$\|\omega h_0\|_\rho \leq m_1(\sigma, s, d, \rho) < 1, \tag{3.3}$$

$$\langle \mathcal{M}[K_0, h_0, \mu_0] \rangle \neq 0 \tag{3.4}$$

where $\mathcal{M}$ *is given in* (5.33) *and* $\langle \cdot \rangle$ *denotes the average of a periodic function;*

(H6) Composition conditions: Let $\eta = dist(\mathbb{C}^d - \mathcal{D}^*, \mathcal{D}_{K_0 \circ H_0, \rho}) > 0$.

Let $\kappa = 2\nu + 5(s + d/2) + 2$.

Then, if

$$|b| < m_2(K_0, h_0, \|r\|_{C^2}, s, A)$$

and

$$\epsilon < \epsilon^*(K_0, h_0, \|r\|_{C^2}, a, b, s, \xi, d) \cdot \sigma^\kappa,$$

there exist K^*, h^*, μ^* *such that* $h^* \in \mathcal{A}_{\rho/2}$, $K^* \circ H^* \in \mathcal{A}_{\rho/2,s}, \mu^* \in \mathcal{O}$ *and*

$$\mathscr{F}[K^*, \mu^*] = 0, \quad \mathscr{G}[K^*, h^*, \mu^*] = 0,$$

where $H^* = Id + \omega h^*$. *In addition,*

$$\begin{aligned} &\|\omega h^* - \omega h_0\|_{\rho/2} \leq C\sigma^{-\kappa}\epsilon, \\ &\|K^* \circ H^* - K_0 \circ H_0\|_{\rho/2,s} \leq C\sigma^{-\kappa}\epsilon, \\ &|\mu^* - \mu_0| \leq C\sigma^{-\kappa}\epsilon, \end{aligned} \tag{3.5}$$

where C *depends only on* $\|r\|_{C^{N+1}}, h_0, K_0, A, \mu_0, a, b, s, \rho, \omega, \alpha, d, \xi, \tau$ *and* ν.

We also have local uniqueness of the solution of (1.12) *and* (1.13). *More precisely, there exists* $\delta^* > 0$ *such that if* (K_1, h_1, μ_1) *and* (K_2, h_2, μ_2) *satisfy*

$$\mathscr{F}[K_1, \mu_1] = \mathscr{G}[K_1, h_1, \mu_1] = 0, \tag{3.6}$$

$$\mathscr{F}[K_2, \mu_2] = \mathscr{G}[K_2, h_2, \mu_2] = 0, \tag{3.7}$$

and $(H1) - (H2), (H4) - (H5)$.

Then, if

$$\max\left\{\|K_2 \circ H_2 - K_1 \circ H_1\|_{3\hat{\rho}/2}, \|h_2 - h_1\|_{3\hat{\rho}/2}, |\mu_2 - \mu_1|\right\} < \delta^*,$$

we have

$$(K_1, h_1, \mu_1) = (K_2, h_2, \mu_2).$$

To produce the approximate solution, we can use a variety of methods. One case is to apply the Lindstedt series developed in Section 4. We could also take the numerically computed results as the approximate solutions.

Remark 3.1. *If we take the solution of the case* $\mu = 0$ *as the approximate solution, then the non-degenerate condition is simplified by* $\langle \partial_\mu r_\mu \circ K_0|_{\mu=0}\rangle \neq 0$ *since* μ *would be assumed to be small enough. See the formulation of Lindstedt series in Section 4.*

An immediate result of Theorem 3.1 is the Lipschitz continuous property on the parameter α.

Corollary 3.1. *Under the hypothesis of Theorem 3.1, if* $(K(\alpha), h(\alpha), \mu(\alpha))$ *is the true solution of* (1.12) *and* (1.13) *, then* $(K \circ H, h, \mu)$ *is Lipschitz continuous with respect to* $\alpha \in \mathscr{D}(\nu; \omega)$ *in some neighborhood of* ω*.*

Proof: Let $(K(\alpha_1), h(\alpha_1), \mu(\alpha_1))$ and $(K(\alpha_2), h(\alpha_2), \mu(\alpha_2))$ be two different solutions of (1.12) and (1.13). We consider $(K(\alpha_1), h(\alpha_1), \mu(\alpha_1))$ as an approximate solution of (1.12) and (1.13) when $\alpha = \alpha_2$. Indeed, we have

$$\begin{aligned}
&\mathscr{G}[K(\alpha_1), h(\alpha_1), \mu(\alpha_1); \alpha_2]\\
&\quad=\omega h(\alpha_1) - \omega h(\alpha_1) \circ R_{\alpha_2} - \alpha_2 - \omega r_{\mu(\alpha_1)} \circ K(\alpha_1) \circ H(\alpha_1)\\
&\quad=\mathscr{G}[K(\alpha_1), h(\alpha_1), \mu(\alpha_1); \alpha_1] + \omega h(\alpha_1) \circ R_{\alpha_1} - \omega h(\alpha_1) \circ R_{\alpha_2} + \alpha_1 - \alpha_2\\
&\quad=\omega h(\alpha_1) \circ R_{\alpha_1} - \omega h(\alpha_1) \circ R_{\alpha_2} + \alpha_1 - \alpha_2,
\end{aligned}$$

which implies

$$\|\mathscr{G}[K(\alpha_1), h(\alpha_1), \mu(\alpha_1); \alpha_2]\|_{\hat{\rho}} \leq C|\alpha_1 - \alpha_2|.$$

If $|\alpha_1 - \alpha_2|$ is sufficient small, there exists a true solution nearby, which is $(K(\alpha_2), h(\alpha_2), \mu(\alpha_2))$ by the local uniqueness. From (3.5) in Theorem 3.1, we have the Lipschitz continuity of the desired functions. □

We also formulate Lindstedt series for the solutions of (1.12) and (1.13). There are two reasons that we discuss Lindstedt series here. On the one hand, when looking for the power series, we would also encounter the main difficulties appearing in the analysis of Newton equations in Section 5. That is, the small divisor problem in conjugation equation and the solvability of the Newton equation of invariance equation. On the other hand, [CCdlL15] recently develops an easy and efficient way to prove the Whitney regularities of functions with respect to the parameters defined on some Cantor-like sets. The methods in [CCdlL15] are based on an a-posteriori theorem and the existence of Lindstedt series.

Actually, we have

Theorem 3.2. *Let* ω *satisfy the Diophantine condition. Assume* $(H1)$ *and* $(H2)$ *in Theorem 3.1 hold. If* $\langle \partial_\mu r_\mu \circ K^{(0)}|_{\mu=0}\rangle \neq 0$*, then we obtain the formal power series to order* N

$$K = \sum_{j=0}^{N} K^{(j)} \alpha^j, \quad h = \sum_{j=1}^{N} h^{(j)} \alpha^j, \quad \mu = \sum_{j=1}^{N} \mu^{(j)} \alpha^j.$$

Furthermore, the power series satisfy the Lindstedt property. That is, for $0 < \delta < \rho$ *and* $0 \le M < N$*, we have*

$$\|\mathscr{F}^{\le M}[K,\mu]\|_{\rho-\delta} \le C|\alpha|^{M+1}, \quad \|\mathscr{G}^{\le M}[K,h,\mu]\|_{\rho-\delta} \le C|\alpha|^{M+1}$$

where C *depends on* $\rho, \delta, a, b, r_\mu, \omega$ *and* f.

From the proof in Corollary 3.1, it is readily seen that the Lipschitz continuity is a by-product of an a-posteriori result. In [CCdlL15], it is shown that with the Lindstedt series one can obtain higher regularities on the parameters in the sense of Whitney. Actually we have a sharper result than Corollary 3.1.

Theorem 3.3. *Under the assumption of Theorem 3.1, the obtained solutions* $(K(\alpha), h(\alpha), \mu(\alpha))$ *is* C^N*-Whitney smooth with respect to* α *in some neighborhood of zero satisfying* $\alpha + 1 \in \mathscr{D}(\nu;\omega)$.

Proof: Taking $\alpha = 0$, we choose $h = \mu = 0$ and obtain a unique solution $K^{(0)}$. Then by Theorem 3.1, there is a unique solution $(K(\alpha), h(\alpha), \mu(\alpha))$ for $\alpha \in \mathscr{D}(\nu;\omega)$. Let $\alpha^* \in \mathscr{D}(\nu;\omega)$ be small enough. Using the continuity of the non-degeneracy condition and Corollary 3.1, we have $\langle \partial_\mu r_\mu \circ K(\alpha^*)|_{\mu=\mu(\alpha^*)}\rangle \neq 0$. Repeating the procedure in Section 4, there are also power series of order one

$$\begin{aligned}
\widetilde{K}(\alpha) &= K(\alpha^*) + K^{(1)} \cdot (\alpha - \alpha^*),\\
\widetilde{h}(\alpha) &= h(\alpha^*) + h^{(1)} \cdot (\alpha - \alpha^*),\\
\widetilde{\mu}(\alpha) &= \mu^* + \mu^{(1)} \cdot (\alpha - \alpha^*)
\end{aligned}$$

satisfying the Lindstedt property. Namely,

$$\begin{aligned}
&\|\mathscr{F}[\widetilde{K}(\alpha), \widetilde{\mu}(\alpha)]\|_{\rho-\tilde\delta} \le C|\alpha - \alpha^*|^2\\
\text{and}\quad &\|\mathscr{G}[\widetilde{K}(\alpha), \widetilde{h}(\alpha), \widetilde{\mu}(\alpha);\alpha]\|_{\rho-\tilde\delta} \le C|\alpha - \alpha^*|^2.
\end{aligned}$$

Then again by the a-posteriori result in Theorem 3.1, we have

$$\begin{aligned}
&\|K(\alpha) - \widetilde{K}(\alpha)\|_{\rho-2\tilde\delta} \le C|\alpha - \alpha^*|^2,\\
&\|h(\alpha) - \widetilde{h}(\alpha)\|_{\rho-2\tilde\delta} \le C|\alpha - \alpha^*|^2,\\
&|\mu(\alpha) - \widetilde{\mu}(\alpha)| \le C|\alpha - \alpha^*|^2,
\end{aligned}$$

which coincide the definition of C^1-Whitney regularity.

For higher Whitney regularities of solutions K, h, μ with respect to α, we refer to [CCdlL15] for a general discussion. The main technique is that we can find a piecewise C^1 path connecting two points α_1 and α_2 on which all the points satisfy Diophantine condition for complex vectors. □

Finally, we denote Π the set of the parameters μ which guarantees the existence of analytic quasi-periodic solutions. We want to show the abundance of parameters in Π.

Theorem 3.4. *Under the assumptions of Theorem 3.1, there exist positive constants C and r_* such that for all r with $0 < r < r_*$,*

$$|(-r, r) \cap \Pi| \geq Cr \tag{3.8}$$

where $|\cdot|$ denotes the Lebesgue measure.

Proof: By the Whitney extension theorem, we find a C^1 function $\Gamma(\alpha)$ in a neighborhood of zero, whose restriction on $\mathscr{D}(\nu;\omega)$ equal to $\mu(\alpha)$. And we also have

$$\Gamma'(0) = \mu^{(1)} = (\langle \partial_\mu r_\mu \circ K^{(0)})|_{\mu=0}\rangle)^{-1} \neq 0$$

which provides a positive lower Lipschitz constant for Γ. Then the bi-Lipschitz functions allow us to control the measure of their ranges. A standard argument yields inequality (3.8). □

Of course, with higher Whitney regularities of $\mu(\alpha)$, we can get sharper estimates on the measure of Π. See Appendix B in [Van02] on the density of pullbacks.

4. Proof of existence of Lindstedt series(Theorem 3.2)

Following the standard perturbative procedure we write

$$\begin{aligned} K &= \sum_{j\geq 0} K^{(j)}\alpha^j, \\ h &= \sum_{j\geq 0} h^{(j)}\alpha^j, \\ \mu &= \sum_{j\geq 0} \mu^{(j)}\alpha^j. \end{aligned} \tag{4.1}$$

where $K^{(j)}, h^{(j)}, \mu^{(j)}$ are the coefficients of α^j. Substitute (4.1) into invariance equation (1.12) and conjugation equation (1.13) and equate the powers of α.

Equating the coefficient of power α^0 in (1.13), we obtain

$$h^{(0)} - h^{(0)} \circ R_\omega - 1 - r_{\mu^{(0)}} \circ K^{(0)} \circ (Id + \omega h^{(0)}) = 0.$$

Choose $\mu^{(0)} = 0$ and $h^{(0)} = 0$. For the coefficient of power in (1.12), we have

$$\partial K^{(0)} \cdot \omega - aK^{(0)} - bK^{(0)} \circ R_\omega - f = 0. \tag{4.2}$$

Expanding K^0 and f into Fourier series with coefficients $\widehat{K}_k^{(0)}$ and $\widehat{h}_k^{(0)}$ respectively, we obtain

$$\widehat{K}_k^{(0)} = \frac{\widehat{f_k}}{2\pi i k \cdot \omega - a - b e^{2\pi i k\cdot\omega}}. \tag{4.3}$$

It is readily seen that when $|b| < |a|$, the divisor in (4.3) stays away from zero and $|\widehat{K}_k^{(0)}| \le (|a| - |b|)^{-1} |\widehat{f_k}|$. Indeed, the Fourier series $\sum_{k\in\mathbb{Z}^d} \widehat{K}_k^{(0)} e^{2\pi i k(\theta+\omega t)}$ is a real analytic quasi-periodic solution of the constant delay case.

Equating the coefficient of power α^1 in (1.13), we have

$$h^{(1)} - h^{(1)} \circ R_\omega - 1 - \partial_\mu r_\mu(K^{(0)})|_{\mu=0} \cdot \mu^{(1)} = 0.$$

Then if ω satisfies the Diophantine condition and the average of $\partial_\mu r_\mu(K^{(0)})|_{\mu=0}$ is not zero, we obtain from Lemma 2.2 that there is a unique solution $h^{(1)}$ with zero average. For the coefficient of power α^1 in (1.12), we also have

$$\partial K^{(1)} \cdot \omega - aK^{(1)} - bK^{(1)} \circ R_\omega = -bDK^{(0)} \cdot \omega \partial_\mu r_\mu(K^{(0)})|_{\mu=0} \cdot \mu^{(1)}$$

is also solvable since the right hand side is already known.

Of course, one can further continue the calculations to higher order of power series. Indeed, for the coefficients of power α^j with $j > 1$, one has

$$h^{(j)} - h^{(j)} \circ R_\omega = \partial_\mu r_\mu(K^{(0)})|_{\mu=0} \cdot \mu^{(j)} + S^{\le(j-1)} \tag{4.4}$$

and

$$\partial K^{(j)} \cdot \omega - aK^{(j)} - bK^{(j)} \circ R_\omega = T^{\le(j-1)} \tag{4.5}$$

where $S^{\le(j-1)}$ can be explicitly computed. And $T^{\le(n-1)}$ also depends on $\mu^{(j)}$ in (4.4). Therefore, one is able to solve (4.4) and (4.5) by the same arguments under the same assumption that the average of $\partial_\mu r_\mu(K^{(0)})|_{\mu=0}$ does not vanish.

Then by the Taylor estimates, we can easily prove the Lindstedt property for the power series, without involving the Sobolev norm.

5. Iteration procedure for the proof of Theorem 3.1

In this section, we analyze the Newton equations for the invariance and conjugation equations. We show how having an (approximate) solution of the conjugation equation simplifies the Newton step of the invariance equation.

5.1. **Newton equation for the invariance equation** (1.12)**.**

Given the approximate solution K_0, h_0 and μ_0, we formulate and analyze the first Newton iteration step in this subsection. Without causing confusions, we use (K, h, μ) instead of (K_0, h_0, μ_0).

5.1.1. *Formulation of the Newton equation of* (1.12). For the moment, we specify the scheme and ignore several precisions such as domain of operators and regularities. They will be discussed later in Section 5.1.2. In this section, we just want to highlight several remarkable cancellations that make the scheme possible of the invariance equation.

The Newton equation of (1.12) is given by

$$D\mathscr{F}[K,\mu](\Delta K,\Delta\mu) = -E, \tag{5.1}$$

where K and μ are the given approximated solutions and unknowns are ΔK and $\Delta\mu$. If we can find $(\Delta K,\Delta\mu)$ solving (5.1), we expect that $K+\Delta K$ and $\mu+\Delta\mu$ are much better approximate solutions of (1.12).

Under some regularity assumptions which we discuss later, the derivative of $\mathscr{F}$ is

$$\begin{aligned} &D\mathscr{F}[K,\mu](\Delta K,\Delta\mu)\\ &=\partial\Delta K\cdot\omega - a\Delta K - b\Delta K\circ\varphi[K,\mu] - bDK\circ\varphi\cdot\partial_K\varphi[K,\mu]\cdot\Delta K\\ &\quad - bDK\circ\varphi\cdot\partial_\mu\varphi[K,\mu]\cdot\Delta\mu \end{aligned} \tag{5.2}$$

where

$$\begin{cases} \partial_K\varphi[K,\mu]\Delta K = -\omega Dr_\mu\circ K\cdot\Delta K\\ \partial_\mu\varphi[K,\mu]\Delta\mu = -\omega D_\mu r_\mu\circ K\cdot\Delta\mu. \end{cases} \tag{5.3}$$

See [dlLO99] and [Mey75] for the computations of Fréchet derivatives of operators involving composition and their estimates.

Looking at (5.2), we see that the equation (5.1) presents an essential difficulty because it involves the terms $\partial\Delta K\cdot\omega$ and $\Delta K\circ\varphi(K,\mu)$. The problem is that due to the composition with the delay mapping $\varphi[K,\mu]$, it would be hard to find an analyticity domain for both $\partial\Delta K\cdot\omega$ and $\Delta K\circ\varphi(K,\mu)$. As we discussed before in Remark 1.1, the problem of studying analyticity domain invariant under the delay mapping φ is resolved by the conjugation equation.

To this end, denoting

$$V = \Delta K\circ H = \Delta K\circ(Id+\omega h) \tag{5.4}$$

and using (3.2) we have

$$\begin{aligned} &D\mathscr{F}[K,\mu](\Delta K,\Delta\mu)\circ H\\ &\quad=\left[\partial V\cdot\omega - aV - bV\circ R_{(\alpha+1)\omega}\right] + \partial V\cdot((DH)^{-1}-\mathbb{I})\omega + bP\cdot\omega Dr_\mu\circ K\circ H\cdot V\\ &\qquad + bP\cdot\omega\partial_\mu r_\mu\circ K\circ H\cdot\Delta\mu - b\int_0^1 D\Delta K\circ(H\circ R_{(\alpha+1)\omega}+qe)\cdot e\,dq, \end{aligned} \tag{5.5}$$

where

$$P[K,h,e] = DK\circ(e + H\circ R_{(\alpha+1)\omega}). \tag{5.6}$$

From Proposition 2.2, we have

$$\partial V\cdot((DH)^{-1}-\mathbb{I})\omega = Tr\left\{D[V\cdot((DH)^{-1}-\mathbb{I})]\cdot\omega\right\}-V\cdot Tr\left\{D((DH)^{-1}-\mathbb{I})\cdot\omega\right\}$$

in which Tr represents the trace of a matrix and $\mathbb{I}$ is the identity matrix of order d.

We would like to seek an approximate solution of the Newton equation (5.1). More precisely, we will consider the following equation

(5.7)

$$\begin{aligned}-E\circ H =&\mathscr{L}V + Tr\left\{D[V\cdot((DH)^{-1}-\mathbb{I})]\cdot\omega\right\} - V\cdot Tr\left\{D((DH)^{-1}-\mathbb{I})\cdot\omega\right\}\\ &+ bP\cdot\omega Dr_\mu\circ K\circ H\cdot V + bP\cdot\omega\partial_\mu r_\mu\circ K\circ H\cdot\Delta\mu\end{aligned}$$

where the linear operator $\mathscr{L}$ is defined as

$$\mathscr{L}[V] = \partial V\cdot\omega - aV - bV\circ R_{(\alpha+1)\omega} \tag{5.8}$$

and the function spaces will be specified later.

5.1.2. *Analysis of the Newton equation* (5.7). We solve the functional equation (5.7) on the Sobolev space $\mathcal{A}_{\zeta,s}$. Firstly, we prove that the operator $\mathscr{L}$ has a bounded inverse on $\mathcal{A}_{\zeta,s}$ for any positive ζ and $s\in 2\mathbb{N}$. Then we show that the other terms can be treated perturbatively.

To see that $\mathscr{L}$ has a bounded inverse, we have that, for a given $S\in\mathcal{A}_{\zeta,s}$, there exists a V solving $\mathscr{L}V = S$ given by

$$V(\theta) = \sum_{k\in\mathbb{Z}^d}\widehat{V}_k e^{2\pi ik\cdot\theta} = \sum_{k\in\mathbb{Z}^d}(2\pi ik\cdot\omega - a - be^{2\pi ik\cdot(\alpha+1)\omega})^{-1}\widehat{S}_k e^{2\pi ik\cdot\theta}. \tag{5.9}$$

The solution V in (5.9) is obtained by expanding S and V into Fourier series with coefficients $\widehat{S}_k$ and $\widehat{V}_k$ respectively. Note that $\mathscr{L}$ is diagonal in the basis of exponentials

$$\mathscr{L}e^{2\pi ik\cdot\theta} = (2\pi ik\cdot\omega - a - be^{2\pi ik\cdot(\alpha+1)\omega})e^{2\pi ik\cdot\theta}.$$

Obviously, if $|b|<|a|$, then the divisor in (5.9) stays away from zero and satisfies

$$\left|2\pi ik\cdot\omega - a - be^{2\pi ik\cdot(\alpha+1)\omega}\right|^{-1} < (|a|-|b|)^{-1}$$

for all the k, ω and Ω. Then, for the Sobolev norm, one has

$$\begin{aligned}||V||^2_{\zeta,s} =&\sum_{k\in\mathbb{Z}^d}\frac{e^{4\pi|k|\zeta}}{B(k,\zeta)}[(2\pi)^d|k|^2+1]^s\cdot\left|2\pi ik\cdot\omega - a - be^{2\pi ik\cdot(\alpha+1)\omega}\right|^{-2}\cdot|\widehat{S}_k|^2\\ \le&(|a|-|b|)^{-2}||S||^2_{\zeta,s}\end{aligned}$$

or equivalently

$$||\mathscr{L}^{-1}||_{\zeta,s} \le (|a|-|b|)^{-1} \tag{5.10}$$

for any ζ and s.

Next, we consider the expression $\mathscr{L}^{-1}Tr\{D[V\cdot((DH)^{-1}-\mathbb{I})]\cdot\omega\}$. Keeping in mind that $(DH)^{-1} = (\mathbb{I}+\omega Dh)^{-1}$ and that Dh is expected to be small, we can apply the Neumann series to compute $(DH)^{-1}-\mathbb{I}$ and obtain

$$(DH)^{-1} - \mathbb{I} = -\omega Dh \cdot \sum_{j=0}^{\infty}(-\omega Dh)^j.$$

We introduce the notation

$$g = (g_1, \cdots, g_d) = -Dh \cdot \sum_{j=0}^{\infty}(-\omega Dh)^j \tag{5.11}$$

and then

$$(DH)^{-1} - \mathbb{I} = \omega \cdot g. \tag{5.12}$$

By Proposition 2.3 in Appendix B, we have the following formal Fourier series

(5.13)

$$\begin{aligned}&\mathscr{L}^{-1}Tr\left\{D[V\cdot((DH)^{-1}-\mathbb{I})]\cdot\omega\right\}\\ &\quad= \sum_{k\in\mathbb{Z}^d} 2\pi ik\cdot\omega\left(2\pi ik\cdot\omega - a - be^{2\pi ik\cdot(\alpha+1)\omega}\right)^{-1}\sum_{j=1}^{d}\omega_j(\widehat{V\cdot g_j})_k\cdot e^{2\pi ik\cdot\theta}.\end{aligned}$$

Noticing that for all $q\in\mathbb{R}$, when $|t|$ is sufficient large, the estimate

$$\left|\frac{t}{it-a-be^{iq}}\right| = \frac{|t|}{|-a-b\cos q + i(t+b\sin q)|} \le \frac{1}{|1+(b\sin q)/t|} < 2$$

holds and when $|t|$ is bounded,

$$\left|\frac{t}{it-a-be^{iq}}\right| \le \frac{|t|}{|a+b\cos q|}$$

is also bounded. Thus we conclude that

$$\sup_{t,q\in\mathbb{R}}\left|\frac{2\pi it}{2\pi it-a-be^{2\pi iq}}\right| < \infty.$$

Back to the series (5.13), we finally obtain

(5.14)

$$\begin{aligned}&\|\mathscr{L}^{-1}Tr\{D[V\cdot((DH)^{-1}-\mathbb{I})]\cdot\omega\}\|_{\zeta,s}^2\\ &\quad= \sum_{k\in\mathbb{Z}^d}\frac{e^{4\pi|k|\zeta}}{B(k,\zeta)}[(2\pi)^d|k|^2+1]^s\left|\frac{2\pi ik\cdot\omega}{2\pi ik\cdot\omega - a - be^{2\pi ik\cdot\Omega}}\sum_{j=1}^{d}\omega_j(\widehat{V\cdot g_j})_k\right|^2\\ &\quad\le C\sum_{j=1}^{d}\|V\cdot g_j\|_{\zeta,s}^2 \le C\sum_{j=1}^{d}\|g_j\|_{\zeta,s}^2\cdot\|V\|_{\zeta,s}^2\end{aligned}$$

if $\zeta < \rho - \sigma$.

Remark 5.1. *Without the particular form* (5.12)*, the estimate of the Sobolev norm of* $\mathscr{L}^{-1}Tr\{D[V\cdot(D^{-1}H-\mathbb{I})]\cdot\omega\}$ *would diverge. This can be seen from its general expression* (B.2)*, in which the numerator does not contain factor* $\omega\cdot k$*. Hence, for a general form of* $(DH)^{-1}$*, the numerator would grow faster than the denominator for large* k*.*

With the analysis above, we now construct a contraction mapping $\mathscr{B}$ such that the equation (5.7) can be solved using (5.10) and (5.14). To this end, we define the linear operator $\mathscr{B}$ as

(5.15)

$$\begin{aligned}\mathscr{B}[V;K,h,\mu]&\\ &= -\mathscr{L}^{-1}Tr\left\{D[V\cdot((DH)^{-1}-\mathbb{I})]\cdot\omega\right\}+\mathscr{L}^{-1}Tr\left\{V\cdot D((DH)^{-1}-\mathbb{I})\cdot\omega\right\}\\ &\quad - b\mathscr{L}^{-1}\left\{P\cdot\omega Dr_\mu\circ K\circ H\cdot V\right\}\end{aligned}$$

on the Sobolev space $\mathcal{A}_{\zeta,s}$.

5.1.3. *Analysis of analyticity domain of* (5.15). Obviously, we encounter the problem of choosing suitable analyticity domains such that the multiplier $P[K,h,e]$ in (5.15) is well-defined. Thus, we need the following lemma to analyze the loss of domain. Recalling the notation defined in (2.4) we know $\mathcal{D}_{H,\rho}=H(\mathcal{D}_\rho)$ which represents the range of a function H on the analytic strip domain $\mathcal{D}_\rho$ with width $\rho>0$. We will also give several similar lemmas when determining the analytic domain for composition functions.

Lemma 5.1. *If*

$$\sigma^{-1}\|\omega h\|_\rho,\ C\sigma^{-1}\|e\|_\rho<1/2, \tag{5.16}$$

then

$$\begin{aligned}\mathcal{D}_{H\circ R_{(\alpha+1)\omega}+e,\rho-10\sigma}\subseteq&\mathcal{D}_{H,\rho-8\sigma}\subseteq\mathcal{D}_{H\circ R_{(\alpha+1)\omega}+e,\rho-5\sigma}\subseteq\mathcal{D}_{H\circ R_{(\alpha+1)\omega},\rho-2\sigma}\\ =&\mathcal{D}_{H,\rho-2\sigma}\subseteq\mathcal{D}_{H,\rho}.\end{aligned} \tag{5.17}$$

This is really a perturbation argument using that $H=Id+\omega h$ is a Lipschitz perturbation of the identity and one could use the standard Lipschitz implicit function theorem. We give full details to be more quantitative.
Proof: Let $\xi=(1,\cdots,1)_d$. We show one inclusion in (5.17) and divide the argument into several steps.

1) H is injective on $\mathcal{D}_{\rho-2\sigma}$.
For any different z_1,z_2 in $\mathcal{D}_{\rho-2\sigma}$, one has

$$\begin{aligned}|H(z_2)-H(z_1)|&=|z_2-z_1-(\omega h(z_2)-\omega h(z_1))|\\ &\geq |z_2-z_1|-\|\omega Dh\|_{\rho-2\sigma}|z_2-z_1|\\ &>(1-(2\sigma)^{-1}\|\omega h\|_\rho)\cdot|z_2-z_1|\end{aligned}$$

which implies the assertion. As a result, given any $0 \le p \le \rho - 2\sigma$, the mapping M_p

$$M_p^H : \begin{cases} \mathbb{R}^d \to H(\{\theta \in \mathbb{C}^d : \Im\mathfrak{m}(\theta_j) = p, j = 1, \cdots, d\}) \\ x \mapsto H(x + ip\xi) \end{cases}$$

is bijective. And we denote $(\cdot)_j$ the j-th component of a vector.

2) $\Im\mathfrak{m}(M_p^H(x))_j$ is monotone increasing in p for any $j = 1, \cdots, n$.
For any $0 < p_1 < p_2$ and $x \in \mathbb{R}^d$, one has

$$\begin{aligned} &\Im\mathfrak{m}(H(x + ip_2\xi) - H(x + ip_1\xi))_j \\ &\quad = \Im\mathfrak{m}(i(p_2 - p_1)\xi + [\omega h(x + ip_2\xi) - \omega h(x + ip_1\xi)])_j \\ &\quad \ge p_2 - p_1 - \|\omega Dh\|_{\rho-2\sigma}(p_2 - p_1) > 0 \end{aligned}$$

since $\|\omega Dh\|_{\rho-2\sigma} \le (2\sigma)^{-1}\|\omega h\|_\rho < 1$. It should be kept in mind that 2) holds under the assumption of real analyticity of H.

Geometrically, 1) and 2) say that the surface $H(\{\theta : \Im\mathfrak{m}(\theta) = p_1\xi\})$ lies between the real plane and another surface $H(\{\theta : \Im\mathfrak{m}(\theta) = p_2\xi\})$. Similar arguments hold for the case of negative imaginary part.

3) $\mathcal{D}_{H\circ R_{(\alpha+1)\omega}+e,\rho-5\sigma} \subseteq \mathcal{D}_{H\circ R_{(\alpha+1)\omega},\rho-2\sigma}$.
For any $x \in \mathbb{R}^d$, one has

$$\begin{aligned} &\Im\mathfrak{m}(H(x + i(\rho - 2\sigma)\xi + (\alpha + 1)\omega))_j - \Im\mathfrak{m}((H \circ R_{(\alpha+1)\omega} + e)(x + i(\rho - 5\sigma)\xi))_j \\ &\quad =3\sigma + \Im\mathfrak{m}(\omega h(x + i(\rho - 2\sigma)\xi + (\alpha + 1)\omega) - \omega h(x + i(\rho - 5\sigma)\xi + (\alpha + 1)\omega))_j \\ &\qquad + \Im\mathfrak{m}(e(x + i(\rho - 5\sigma)\xi))_j \\ &\quad \ge 3\sigma \cdot (1 - \|\omega Dh\|_{\rho-\sigma} - (3\sigma)^{-1}\|e\|_\rho) > 0. \end{aligned}$$

Observing that the rotation operator $R_{(\alpha+1)\omega}$ does not change the strip domain on the complex space. That is, $R_{(\alpha+1)\omega}(\mathcal{D}_\zeta) = \mathcal{D}_\zeta$ for all $\zeta > 0$.Then, $\mathcal{D}_{H\circ R_{(\alpha+1)\omega},\zeta} = \mathcal{D}_{H,\zeta}$.

The other inclusions in (5.17) are proved in the same way and thus omitted. □

5.1.4. *Solution of the Newton equation* (5.7). Our goal in this subsection is to prove that the operator $\mathscr{B}$ defined in (5.15) is a contraction on the Sobolev space $\mathcal{A}_{\rho-6\sigma,s}$. Actually, we have

Lemma 5.2. *If*

$$\frac{\|\omega h\|_\rho}{\sigma^{2+s+d/2}},\ |b| << 1,$$

then, $\mathscr{B}$ is a contraction on the Sobolev space $\mathcal{A}_{\rho-6\sigma,s}$.

Proof: We first estimate the Sobolev norm of g defined in (5.11) in terms of the supremum norm of h, we get the estimates

$$\|g_j\|_{\rho-2\sigma,s} \le C\sigma^{-(s+d/2)}\|g_j\|_{\rho-\sigma} \le C\sigma^{-(1+s+d/2)}\|\omega h\|_\rho$$

and

$$\|D((DH)^{-1} - \mathbb{I})\|_{\rho-3\sigma,s} \le C\sigma^{-(s+d/2)}\|D((DH)^{-1} - \mathbb{I})\|_{\rho-2\sigma} \le C\sigma^{-(2+s+d/2)}\|\omega h\|_{\rho}.$$

if

$$\frac{C}{\sigma^{2+s+d/2}}\|\omega h\|_{\rho} < 1. \tag{5.18}$$

Furthermore, we have

$$\begin{aligned}\|P[K, h, e]\|_{\rho-6\sigma,s} &\le C\sigma^{-(s+d/2)}\|DK \circ (H \circ R_{(\alpha+1)\omega} + e)\|_{\rho-5\sigma} \le C\sigma^{-(s+d/2)}\|DK \circ H\|_{\rho-2\sigma} \\ &\le C\sigma^{-(1+s+d/2)}\|(DH)^{-1}\|_{\rho-\sigma} \cdot \|K \circ H\|_{\rho}.\end{aligned} \tag{5.19}$$

Then, for any V_1 and V_2 in $\mathcal{A}_{\rho-6\sigma,s}$, one sees from (5.14) and (5.19) that

$$\begin{aligned}&\|\mathcal{B}[V_2 - V_1]\|_{\rho-6\sigma,s} \\ &\le C(\sum_{n=1}^{d}\|g_n\|^2_{\rho-6\sigma,s})^{1/2} \cdot \|V_2 - V_1\|_{\rho-6\sigma,s} + C\|Tr[D((DH)^{-1} - \mathbb{I}) \cdot \omega]\|_{\rho-6\sigma,s} \\ &\quad \times \|V_2 - V_1\|_{\rho-6\sigma,s} + |b| \cdot C\|P\|_{\rho-6\sigma,s} \cdot \|\omega Dr_\mu \circ K \circ H\|_{\rho-6\sigma,s}\|V_2 - V_1\|_{\rho-6\sigma,s} \\ &\le C \cdot \left\{\frac{\|\omega h\|_{\rho}}{\sigma^{2+s+d/2}} + \frac{|b|}{\sigma^{1+s+d/2}}\|(DH)^{-1}\|_{\rho-\sigma} \cdot \|K \circ H\|_{\rho} \cdot \frac{\|Dr_\mu \circ K \circ H\|_{\rho-\sigma}}{\sigma^{1+s+d/2}}\right\} \\ &\quad \times \|V_2 - V_1\|_{\rho-6\sigma,s}.\end{aligned} \tag{5.20}$$

If

$$\frac{\|\omega h\|_{\rho}}{\sigma^{2+s+d/2}},\ |b| << 1$$

such that

$$C \cdot \left\{\frac{\|\omega h\|_{\rho}}{\sigma^{2+s+d/2}} + \frac{|b|}{\sigma^{1+s+d/2}} \cdot \|K \circ H\|_{\rho} \cdot \frac{\|Dr_\mu \circ K \circ H\|_{\rho-\sigma}}{\sigma^{1+s+d/2}}\right\} < \lambda < \frac{1}{2} \tag{5.21}$$

then $\mathcal{B}$ is a contraction on $\mathcal{A}_{\rho-6\sigma,s}$, which satisfies

$$\|\mathcal{B}\|_{\rho-6\sigma,s} < \lambda < \frac{1}{2}. \tag{5.22}$$

□

Hence, from (5.7), one has

$$V = (Id - \mathcal{B})^{-1}\left\{-b\mathcal{L}^{-1}[P \cdot \omega\partial_\mu r_\mu \circ K \circ H \cdot \Delta\mu] - \mathcal{L}^{-1}[E \circ H]\right\}, \tag{5.23}$$

which belongs to $\mathcal{A}_{\rho-6\sigma,s}$. It is necessary to choose a suitable $\Delta\mu$ such that V can be small enough. We will find such $\Delta\mu$ in next subsection under some non-degeneracy conditions.

5.2. **Newton equation for the conjugation equation** (1.13)**.** Likewise, we formulate and analyze the Newton equation for the conjugation equation. Because of the delay mapping preserving the ω−foliation, we actually encounter a one dimensional problem.

5.2.1. *Formulation of the Newton equation of* (1.13). We consider the Newton equation of (1.13), which reads

$$D\mathscr{G}[K, h, \mu](\Delta K, \Delta h, \Delta\mu) = -e. \tag{5.24}$$

The unknowns are ΔK, Δh and $\Delta\mu$. The goal is to solve (5.24) such that $K + \Delta K$, $h + \Delta h$ and $\mu + \Delta\mu$ are better approximate solutions of (1.13).

Under some regularity assumptions stated later, the Fréchet derivative of the operator $\mathscr{G}$ is

(5.25)

$$\begin{aligned} &D\mathscr{G}[K, h, \mu](\Delta K, \Delta h, \Delta\mu) \\ &\quad = \omega\Delta h - \omega\Delta h \circ R_{(\alpha+1)\omega} - \omega Dr_\mu \circ K \circ H \cdot DK \circ H \cdot \omega\Delta h - \omega\partial_\mu r_\mu \circ K \circ H \cdot \Delta\mu \\ &\qquad - \omega Dr_\mu \circ K \circ H \cdot \Delta K \circ H. \end{aligned}$$

Equation (5.24) is very standard and appears in [Mos66a]. We also observe that when the delay mapping φ defined in (1.6) is foliation-preserving(see Appendix A), the equation (5.25) has additional features. That is, the derivative of $\mathscr{G}$ can be written as the product of the ω and a scalar function. Differentiating (3.2) with respect to the variable θ, one also has

$$\begin{aligned} De =& \omega Dh - \omega Dh \circ R_{(\alpha+1)\omega} \\ &- \omega Dr_\mu \circ K \circ (Id + \omega h) \cdot DK \circ H \cdot (\mathbb{I} + \omega Dh). \end{aligned} \tag{5.26}$$

This motivates us introducing a new unknown W in place of Δh. Let

$$W = (\mathbb{I} + \omega Dh)^{-1} \cdot \omega\Delta h = (\mathbb{I} + \omega g) \cdot \omega\Delta h \tag{5.27}$$

and then

$$\omega\Delta h = (\mathbb{I} + \omega Dh)W.$$

Substituting (5.25) and (5.26) into the Newton equation (5.24) yields

$$\begin{aligned} W - W \circ R_{(\alpha+1)\omega} = -(\mathbb{I} + \omega g \circ R_{(\alpha+1)\omega}) \cdot \{&e - \omega\partial_\mu r_\mu \circ K \circ (Id + \omega h) \cdot \Delta\mu \\ &- \omega Dr_\mu \circ K \circ H \cdot \Delta K \circ H + De \cdot W\}, \end{aligned}$$

where now the unknown are W, ΔK and $\Delta\mu$.

Since the term $De{\cdot}W$ is formally quadratic from the equation to be solved, we omit it for the moment and consider the following approximate Newton equation

(5.28)

$$\begin{aligned} W - W \circ R_{(\alpha+1)\omega} = &- (\mathbb{I} + \omega g \circ R_{(\alpha+1)\omega}) \cdot \Big\{e - \omega\partial_\mu r_\mu \circ K \circ (Id + \omega h) \cdot \Delta\mu \\ &-\omega Dr_\mu \circ K \circ H \cdot \Delta K \circ H\Big\}. \end{aligned}$$

Inserting $\Delta K \circ H$ obtained in (5.23) into (5.28), one has

(5.29)
$$\begin{aligned} W - W \circ R_{(\alpha+1)\omega} &= -(\mathbb{I} + \omega g \circ R_{(\alpha+1)\omega}) \cdot \left\{ e + \omega D r_\mu \circ K \circ H \cdot (Id - \mathscr{B})^{-1} \mathscr{L}^{-1}[E \circ H] \right\} \\ &\quad + (\mathbb{I} + \omega g \circ R_{(\alpha+1)\omega}) \cdot \omega \left\{ \partial_\mu r_\mu \circ K \circ H + D r_\mu \circ K \circ H \right. \\ &\quad \left. \times (Id - \mathscr{B})^{-1} \mathscr{L}^{-1}[P \cdot \omega \partial_\mu r_\mu \circ K \circ H] \right\} \cdot \Delta\mu. \end{aligned}$$

We emphasize that the unknowns in (5.29) are W and $\Delta\mu$ and the other terms are already known.

5.2.2. *Analysis of the Newton equation of* (1.13). With the properties of foliation-preserving, we finally reduce (5.29) to a scalare equation problem.

To see this, we find that the initial error for the conjugation equation e can also be written as

$$e = \widetilde{e}\omega, \tag{5.30}$$

where

$$\widetilde{e} = h_0 - h_0 \circ R_\Omega - \alpha - r_{\mu_0} \circ K_0 \circ (Id + \omega h_0).$$

Let $\widetilde{W} = \Delta h + g\omega\Delta h$ and from (5.27) we have

$$W = \widetilde{W}\omega. \tag{5.31}$$

Substituting (5.30) and (5.31) into the cohomology equation (5.29), we obtain the scalar equation

$$\widetilde{W} - \widetilde{W} \circ R_{(\alpha+1)\omega} = -\mathcal{N}[K, h, \mu] + \mathcal{M}[K, h, \mu]\Delta\mu, \tag{5.32}$$

where

(5.33)
$$\begin{aligned} \mathcal{M}[K, h, \mu] =& (1 + g \circ R_{(\alpha+1)\omega} \cdot \omega) \cdot \left\{ \partial_\mu r_\mu \circ K \circ H + D r_\mu \circ K \circ H \right. \\ & \left. \times (Id - \mathscr{B})^{-1} \mathscr{L}^{-1}(DK \circ H \circ R_{(\alpha+1)\omega} \cdot \omega \partial_\mu r_\mu \circ K \circ H) \right\}, \end{aligned}$$

and

(5.34)
$$\mathcal{N}[K, h, \mu] = (1 + g \circ R_{(\alpha+1)\omega} \cdot \omega) \cdot \left\{ \widetilde{e} + D r_\mu \circ K \circ H \cdot (Id - \mathscr{B})^{-1} \mathscr{L}^{-1}[E \circ H] \right\}.$$

Recalling the theory of cohomology equation developed in Section 2.2, to obtain a solution of (5.29), it is necessary and sufficient that the right hand side has zero average. This amounts to

$$-\langle \mathcal{N}[K, h, \mu] \rangle + \langle \mathcal{M}[K, h, \mu] \rangle \Delta\mu = 0, \tag{5.35}$$

where $\langle \cdot \rangle$ denotes the average of a periodic function.

For the first Newton step, by the non-degeneracy condition (3.4)

$$\langle \mathcal{M}[K_0, h_0, \mu_0] \rangle \neq 0,$$

we can solve (5.35) uniquely and obtain a sufficient small $\Delta\mu$. We will give the estimates on the corrections in the next subsection.

Remark 5.2. *Since we will be modifying the functional equation* (1.12) *and* (1.13) *in the iteration process, it is important to realize that the non-degeneracy condition* (3.4) *is an open condition. More precisely, if* (3.4) *holds for some* K_0, h_0, μ_0*, it will also hold for all* K, h, μ *close enough in the usual* C^1 *topology.*

5.3. **Estimates on the corrections.** For Equation (5.35), using the non-degeneracy condition (3.4) one has

$$\Delta\mu = -\langle \mathcal{M}[K_0, h_0, \mu_0]\rangle^{-1} \cdot \langle \mathcal{N}[K_0, h_0, \mu_0]\rangle \tag{5.36}$$

and

$$|\Delta\mu| \leq C\epsilon, \tag{5.37}$$

where the constant C depends on the given approximate solution K_0, h_0, μ_0, A, the delay function r and the norm of $(Id - \mathscr{B})^{-1}$.

Applying Lemma 2.2 to the equation (5.28), one readily obtains a unique W with zero average which belongs to $\mathcal{A}_{\rho-7\sigma}$ and

$$||W||_{\rho-7\sigma} \leq \frac{C}{\sigma^{\nu}}\epsilon\ . \tag{5.38}$$

Combining with (5.27), we get the correction and its estimate as

$$||\omega\Delta h||_{\rho-7\sigma} \leq \frac{C}{\sigma^{\nu}}\epsilon\ . \tag{5.39}$$

Once $\Delta\mu$ and its estimate is obtained, we also have

$$||\Delta K \circ H||_{\rho-6\sigma,s} = ||V||_{\rho-6\sigma,s} \leq C(||P[K, h, e]||_{\rho-6\sigma,s}|\Delta\mu| + ||E \circ H||_{\rho,s}) \leq C\epsilon. \tag{5.40}$$

For the following iteration step, we denote

$$\begin{aligned} K^+ &= K + \Delta K,\\ h^+ &= h + \Delta h,\\ H^+ &= Id + \omega h^+ = H + \omega\Delta h,\\ \mu^+ &= \mu + \Delta\mu. \end{aligned} \tag{5.41}$$

We also give some estimates after introducing the corrections $\Delta K, \omega\Delta h, \Delta\mu$. For convenience, we summarize these standard estimates below in detail and give their proofs in Appendix C.

Proposition 5.1. *If*

$$\sigma^{-(2+s+d/2)}||\omega h||_\rho,\ C\sigma^{-(\nu+1)}\epsilon,\ |b| << 1, \tag{5.42}$$

one has the following estimates:

(i) $\|\omega h^+\|_{\rho-7\sigma} \le \|\omega h\|_\rho + C\sigma^{-\nu}\epsilon$;
(ii) $\|\omega Dh^+\|_{\rho-8\sigma} \le \|\omega h\|_\rho \sigma^{-1} + C\sigma^{-(\nu+1)}\epsilon$;
(iii) $\|DH^+ - \mathbb{I}\|_{\rho-8\sigma} \le \|\omega h\|_\rho \sigma^{-1} + C\sigma^{-(\nu+1)}\epsilon$;
(iv) $\|(DH^+)^{-1} - \mathbb{I}\|_{\rho-8\sigma} \le 2\|\omega h\|_\rho \sigma^{-1} + C\sigma^{-(\nu+1)}\epsilon$;
(v) $\|D((DH^+)^{-1} - \mathbb{I})\|_{\rho-9\sigma} \le C\|\omega h\|_\rho \sigma^{-2} + C\sigma^{-(\nu+2)}\epsilon$;
(vi) $\|D^2 H^+\|_{\rho-9\sigma} \le \|\omega h\|_\rho \sigma^{-2} + C\sigma^{-(\nu+2)}\epsilon$
(vii) $\|K \circ H^+ - K \circ H\|_{\rho-7\sigma} \le C\|DK \circ H\|_{\rho-2\sigma} \cdot \sigma^{-\nu}\epsilon$
(viii) $\|K \circ H^+\|_{\rho-7\sigma} \le [1 + C\sigma^{-(\nu+1)}\epsilon] \cdot \|K \circ H\|_\rho$;
(ix) $\|DK \circ H^+\|_{\rho-8\sigma} \le [\sigma^{-1} + C\sigma^{-(\nu+2)} \cdot \epsilon] \cdot \|K \circ H\|_\rho \cdot (1 + 2\|\omega Dh^+\|_{\rho-8\sigma})$;
(x) $\|\Delta K \circ H^+\|_{\rho-12\sigma,s} \le C\sigma^{-(s+d/2)}\epsilon$;
(xi) $\|DK^+ \circ \varphi[K^+,\mu^+] \circ H^+\|_{\rho-12\sigma,s} \le C\dfrac{\|DK \circ H\|_{\rho-\sigma}}{\sigma^{s+d/2}} + C\|D^{-1}H\|_{\rho-\sigma} \cdot \sigma^{-(1+s+d/2)}\epsilon$;
(xii) $\|\partial_K \varphi[K^+,\mu^+] \circ H^+\|_{\rho-12\sigma,s} \le C\|Dr\|_{\mathcal{D}^*} \cdot \sigma^{-(2s+d)}\epsilon$;
(xiii) $\|\partial_\mu \varphi[K^+,\mu^+] \circ H^+ \cdot \Delta\mu\|_{\rho-12\sigma,s} \le C\|\partial_\mu r\|_{\mathcal{D}^*} \cdot \epsilon$;
(xiv) $\|(\partial_{KK}\varphi[K^+,\mu^+] \cdot (\Delta K)^{\otimes 2}) \circ H^+\|_{\rho-12\sigma,s} \le C\|D^2 r\|_{\mathcal{D}^*} \cdot \sigma^{-(3s+3d/2)}\epsilon^2$;
(xv) $\|(\partial_{K\mu}\varphi[K^+,\mu^+] \cdot \Delta K) \circ H^+ \Delta\mu\|_{\rho-12\sigma,s} \le C\|D\partial_\mu r\|_{\mathcal{D}^*} \cdot \sigma^{-(2s+d)}\epsilon^2$;
(xvi) $\|\partial_{\mu\mu}\varphi[K^+,\mu^+] \circ H^+(\Delta\mu)^{\otimes 2}\|_{\rho-12\sigma,s} \le C\|D_{\mu\mu} r\|_{\mathcal{D}^*} \cdot \sigma^{-(s+d/2)} \cdot \epsilon^2$;
(xvii)

$$\begin{aligned}&\|D^2K^+ \circ \varphi[K^+,\mu^+] \circ H^+\|_{\rho-12\sigma,s}\\ &\qquad \le C\sigma^{-(s+d/2)}\|D^2K \circ H\|_{\rho-2\sigma} + C\|D^2H\|_{\rho-2\sigma} \cdot \|(DH)^{-1}\|^3_{\rho-\sigma} \cdot \sigma^{-(2+s+d/2)}\epsilon.\end{aligned}$$

The estimates $(i) - (vi)$ can be proved using Lemma 5.1 to analyze its analyticity domain. However, when composing with H^+, the estimates $(xi) - (xvii)$ requires a further restriction on the analyticity domain. We will present how these expressions come into our problem in the following subsection. The Lemma 5.3 and Lemma 5.4 will be used to estimate $(vii) - (x)$ and $(xi - xvii)$ in Proposition 5.1 respectively.

5.4. **Estimates on the improved error.** By the assumptions on the smoothness and the Taylor's formula, we have

(5.43)
$$
\begin{aligned}
&\mathscr{G}[K+\Delta K, h+\Delta h, \mu+\Delta\mu] - \mathscr{G}[K,h,\mu] - D\mathscr{G}[K,h,\mu](\Delta K,\Delta h,\Delta\mu)\\
=& -\int_0^1 q \int_0^1 \omega D^2 r_{\mu+pq\Delta\mu}\circ K\circ(H+pq\omega\Delta h)\cdot[DK\circ(H+pq\omega\Delta h)\cdot\omega\Delta h]^{\otimes 2}\\
&+\omega D r_{\mu+pq\Delta\mu}\circ K\circ(H+pq\omega\Delta h)\cdot D^2K\circ(H+pq\omega\Delta h)\cdot(\omega\Delta h)^{\otimes 2}\\
&+2\omega D\partial_\mu r_{\mu+pq\Delta\mu}\circ K\circ(H+pq\omega\Delta h)\cdot[DK\circ(H+pq\omega\Delta h)\cdot\omega\Delta h]\cdot\Delta\mu\\
&+\omega D_\mu^2 r_{\mu+pq\Delta\mu}\circ K\circ(H+pq\omega\Delta h)\cdot\Delta\mu^{\otimes 2}\\
&+\omega D^2 r_{\mu+pq\Delta\mu}\circ(K+pq\Delta K)\circ(H+pq\omega\Delta h)\cdot\Delta K\circ(H+pq\omega\Delta h)\\
&\ \times D(K+pq\Delta K)\circ(H+pq\omega\Delta h)\cdot\omega\Delta h\\
&+\omega D r_{\mu+pq\Delta\mu}\circ(K+pq\Delta K)\circ(H+pq\omega\Delta h)\cdot D\Delta K\circ(H+pq\omega\Delta h)\cdot\omega\Delta h\\
&+\omega D\partial_\mu r_{\mu+pq\Delta\mu}\circ(K+pq\Delta K)\circ(H+pq\omega\Delta h)\cdot\Delta K\circ(H+pq\omega\Delta h)\cdot\Delta\mu\\
&+\omega D^2 r_{\mu+pq\Delta\mu}\circ(K+pq\Delta K)\circ(H+pq\omega\Delta h)\cdot[\Delta K\circ(H+pq\omega\Delta h)]^{\otimes 2}\\
&dp\,dq.
\end{aligned}
$$

Before estimating the difference, we have to show the terms in (5.43) are well-defined. To see this, we show a similar result to Lemma 5.1.

Lemma 5.3. *If the conditions* (5.42) *hold, then*

$$
\mathcal{D}_{H+\omega\Delta h,\rho-10\sigma} \subseteq \mathcal{D}_{H,\rho-8\sigma}. \tag{5.44}
$$

Proof: It is sufficient to show that, for any $x\in\mathbb{R}^d$,

$$
\begin{aligned}
&\Im\mathfrak{m}(H(x+i(\rho-8\sigma)\xi))_j - \Im\mathfrak{m}((H+\omega\Delta h)(x+i(\rho-10\sigma)))_j\\
=&\Im\mathfrak{m}(H(x+i(\rho-8\sigma)))_j - \Im\mathfrak{m}(H(x+i(\rho-10\sigma)))_j - \Im\mathfrak{m}(\omega\Delta h(x+i(\rho-10\sigma)))_j\\
\geq& 2\sigma\cdot(1-\|\omega Dh\|_{\rho-8\sigma} - \frac{\|\omega\Delta h\|_{\rho-7\sigma}}{2\sigma})\\
>& 0
\end{aligned}
$$

□

We now consider the difference (5.43) on the analyticity domain $\mathcal{D}_{\rho-10\sigma}$. Obviously, a trivial estimate yields

(5.45)
$$
\begin{aligned}
&\|\mathscr{G}[K+\Delta K, h+\Delta h, \mu+\Delta\mu] - \mathscr{G}[K,h,\mu] - D\mathscr{G}[K,h,\mu](\Delta K, \Delta h, \Delta\mu)\|_{\rho-10\sigma} \\
&\le |\omega|\cdot\|D^2 r\|_{\mathcal{D}^*}\cdot\|DK\circ H\|^2_{\rho-2\sigma}\cdot C\sigma^{-2\nu}\epsilon^2 + |\omega|\cdot\|Dr\|_{\mathcal{D}^*}\cdot\|D^2K\circ H\|_{\rho-2\sigma}\cdot C\sigma^{-2\nu}\epsilon^2 \\
&\quad + |\omega|\cdot\|D^2 r\|_{\mathcal{D}^*}\cdot\|DK\circ H\|_{\rho-2\sigma}\cdot C\sigma^{-(\nu+1)}\epsilon^2 \\
&\quad + |\omega|\cdot\|D^2 r\|_{\mathcal{D}^*}\cdot C\epsilon^2 + |\omega|\cdot\|D^2 r\|_{\mathcal{D}^*}\cdot\|DK\circ H\|_{\rho-2\sigma}\cdot C\sigma^{-\nu}\epsilon^2 \\
&\quad + |\omega|\cdot\|Dr\|_{\mathcal{D}^*}\cdot C\sigma^{-\nu+1}\epsilon^2 + |\omega|\cdot\|D\partial_\mu r\|_{\mathcal{D}^*}\cdot C\epsilon^2 + |\omega|\cdot\|D^2 r\|_{\mathcal{D}^*}\cdot C\epsilon^2 \\
&\le C(\|DK\circ H\|^2_{\rho-2\sigma} + \|D^2K\circ H\|_{\rho-2\sigma})\cdot\sigma^{-2\nu}\epsilon^2.
\end{aligned}
$$

We recall that $\Delta K, \Delta h$ and $\Delta\mu$ have been chosen in such a way that the Newton equation (5.24) is solved up to an error term $De\cdot W$. Considering the omitted term $De\cdot W$ in the Newton equation (5.2.1), one immediately has

$$
\|De\cdot W\|_{\rho-10\sigma} \le C\|(DH)^{-1}\|_{\rho-\sigma}\cdot\sigma^{-(\nu+1)}\epsilon^2. \tag{5.46}
$$

Then, from (5.28), (5.45) and (5.46), we have

(5.47)
$$
\begin{aligned}
&\|\mathscr{G}[K+\Delta K, h+\Delta h, \mu+\Delta\mu]\|_{\rho-10\sigma} \\
&\le \|\mathscr{G}[K,h,\mu] + D\mathscr{G}[K,h,\mu](\Delta K, \Delta h, \Delta\mu)\|_{\rho-10\sigma} \\
&\quad + \|\mathscr{G}[K+\Delta K, h+\Delta h, \mu+\Delta\mu] - \mathscr{G}[K,h,\mu] - D\mathscr{G}[K,h,\mu](\Delta K, \Delta h, \Delta\mu)\|_{\rho-10\sigma} \\
&\le C(\|(DH)^{-1}\|_{\rho-\sigma} + \|DK\circ H\|^2_{\rho-2\sigma} + \|D^2K\circ H\|_{\rho-2\sigma})\cdot\sigma^{-2\nu}\epsilon^2\ .
\end{aligned}
$$

We denote the new error of the conjugation equation (1.11) by

$$
e^+ = \mathscr{G}[K+\Delta K, h+\Delta h, \mu+\Delta\mu] \tag{5.48}
$$

and thus

$$
\|e^+\|_{\rho-10\sigma} \le C\sigma^{-2\nu}\epsilon^2. \tag{5.49}
$$

Now we are ready to give the estimates on the new error in the invariance equation after introducing the new correction ΔK. Again one has the formal expression

(5.50)

$$\begin{aligned}
&\mathscr{F}[K+\Delta K,\mu+\Delta\mu]-\mathscr{F}[K,\mu]-D\mathscr{F}[K,\mu](\Delta K,\Delta\mu)\\
=&-b\int_0^1 q\int_0^1 2D\Delta K\circ\varphi_{pq}\cdot[\partial_K\varphi_{pq}\cdot\Delta K+\partial_\mu\varphi_{pq}\cdot\Delta\mu]\\
&+D^2(K+pq\Delta K)\circ\varphi_{pq}\cdot[(\partial_K\varphi_{pq}\cdot\Delta K)^{\otimes 2}+2\partial_K\varphi_{pq}\cdot\Delta K\cdot\partial_\mu\varphi_{pq}\cdot\Delta\mu\\
&\quad+(\partial_\mu\varphi_{pq}\cdot\Delta\mu)^{\otimes 2}]\\
&+D(K+pq\Delta K)\circ\varphi_{pq}\cdot[\partial_{KK}\varphi_{pq}\cdot(\Delta K)^{\otimes 2}+2\partial_{K\mu}\varphi_{pq}\cdot\Delta K\Delta\mu\\
&\quad+\partial_{\mu\mu}\varphi_{pq}\cdot(\Delta\mu)^{\otimes 2}]\,dpdq
\end{aligned}$$

where

$$\varphi_{pq}=\varphi[K+pq\Delta K,\mu+pq\Delta\mu].$$

Furthermore, when composing (5.50) with H^+, we show some terms explicitly as follows:

(E1)

$$\begin{aligned}
\varphi_{pq}\circ H^+ =&H^+-\omega r_{\mu+pq\Delta\mu}\circ(K+pq\Delta K)\circ H^+\\
=&-\varphi[K,\mu^+]\circ H^+-(\varphi[K,\mu+pq\Delta\mu]-\varphi[K,\mu^+])\circ H^+\\
&-\omega pq\int_0^1 Dr_{\mu+pq\Delta\mu}\circ(K+pqt\Delta K)\circ H^+\cdot\Delta K\circ H^+dt\\
=&-H^+\circ R_{(\alpha+1)\omega}+e^+\\
&-\omega(1-pq)\int_0^1\partial_\mu r_{\mu+[pq+(1-pq)t]\Delta\mu}\circ K\circ H^+\cdot\Delta\mu\,dt\\
&-\omega pq\int_0^1 Dr_{\mu+pq\Delta\mu}\circ(K+pqt\Delta K)\circ H^+\cdot\Delta K\circ H^+dt
\end{aligned}\tag{5.51}$$

(E2)

$$\begin{aligned}
[\partial_K\varphi_{pq}\cdot\Delta K+\partial_\mu\varphi_{pq}\cdot\Delta\mu]\circ H^+=&-\omega Dr_{\mu+pq\Delta\mu}\circ(K+pq\Delta K)\circ H^+\cdot\Delta K\circ H^+\\
&-\partial_\mu r_{\mu+pq\Delta\mu}\circ(K+pq\Delta K)\circ H^+\cdot\Delta\mu
\end{aligned}$$

(E3)

$$[\partial_{KK}\varphi_{pq}\cdot(\Delta K)^{\otimes 2}]\circ H^+=-\omega D^2r_{\mu+pq\Delta\mu}\circ(K+pq\Delta K)\circ H^+\cdot(\Delta K\circ H^+)^{\otimes 2}$$

which either determine the analyticity domain or are useful for the error estimates.

The remanning terms are similar to (3) and thus omit. From lemma 5.3, one has

$$\mathcal{D}_{H^+,\rho-10\sigma}\subseteq\mathcal{D}_{H,\rho-8\sigma}.$$

Furthermore, since

(5.52)
$$\begin{aligned}\|Dr_{\mu+\Delta\mu}\circ(K+\Delta K)\circ H^{+}\cdot\Delta K\circ H^{+}\|_{\rho-10\sigma} &\le C\|\partial_\mu r\|_{\mathcal{D}^*}\cdot\|\Delta K\circ H^{+}\|_{\rho-10\sigma}\\ &\le C\|\partial_\mu r\|_{\mathcal{D}^*}\cdot\|\Delta K\circ H\|_{\rho-6\sigma}\le C\|\partial_\mu r\|_{\mathcal{D}^*}\cdot\|\Delta K\circ H\|_{\rho-6\sigma,s}\\ &\le C\|\partial_\mu r\|^2_{\mathcal{D}^*}\cdot\|P\|_{\rho-6\sigma,s}\cdot\epsilon\end{aligned}$$

and

$$\|\partial_\mu r_{\mu+\Delta\mu}\circ K\circ H^{+}\cdot\Delta\mu\|_{\rho-10\sigma}\le C\|\partial_\mu r\|_{\mathcal{D}^*}\cdot\epsilon, \tag{5.53}$$

we have the following lemma from the expression (5.51) of $\varphi[K^+,h^+]\circ H^+$.

Lemma 5.4. *If the conditions* (5.42) *hold, then*

$$\mathcal{D}_{\varphi[K^+,\mu^+]\circ H^+,\rho-10\sigma}\subseteq\mathcal{D}_{H,\rho-8\sigma}. \tag{5.54}$$

The smallness of ϵ also ensure that $\mu+\Delta\mu$ and $\mathcal{D}_{K^+\circ H^+,\rho-10\sigma}$ does not run out of the analyticity domain of $\partial_\mu r$ and Dr. Using Lemma 5.4, we are able to estimate the terms in (5.50).

By Proposition 5.1, we have

(5.55)
$$\begin{aligned}&\|\{\mathscr{F}[K+\Delta K,\mu+\Delta\mu]-\mathscr{F}[K,\mu]-D\mathscr{F}[K,\mu](\Delta K,\Delta\mu)\}\circ H^{+}\|_{\rho-12\sigma,s}\\ &\le C\|D\Delta K\circ\varphi_{pq}\circ H^{+}\|_{\rho-12\sigma,s}\cdot\Big\{\|(\partial_K\varphi[K^+,\mu^+]\cdot\Delta K)\circ H^{+}\|_{\rho-12\sigma,s}\\ &\quad+\|\partial_\mu\varphi[K^+,\mu^+]\circ H^{+}\|_{\rho-12\sigma,s}\times|\Delta\mu|\Big\}+\|D^2K^{+}\circ\varphi[K^+,\mu^+]\circ H^{+}\|_{\rho-12\sigma,s}\\ &\quad\times\{\|(\partial_K\varphi[K^+,\mu^+]\cdot\Delta K)\circ H^{+}\|_{\rho-12\sigma,s}+\|\partial_\mu\varphi[K^+,\mu^+]\circ H^{+}\|_{\rho-12\sigma,s}\\ &\qquad\times|\Delta\mu|\}^2+\|DK^{+}\circ\varphi[K^+,\mu^+]\circ H^{+}\|_{\rho-12\sigma,s}\cdot\\ &\quad\times\{\|(\partial_{KK}\varphi[K^+,\mu^+]\cdot(\Delta K)^{\otimes 2})\circ H^{+}\|_{\rho-12\sigma,s}\\ &\qquad+2\|(\partial_{K\mu}\varphi[K^+,\mu^+]\cdot\Delta K)\circ H^{+}\|_{\rho-12\sigma,s}\cdot|\Delta\mu|\\ &\qquad+\|\partial_{\mu\mu}\varphi[K^+,\mu^+]\circ H^{+}\|_{\rho-12\sigma,s}\cdot|\Delta\mu|^2\}\\ &\le C\frac{\epsilon^2}{\sigma^{5(s+d/2)+2}}.\end{aligned}$$

For the omitted term in the Newton equation of $\mathscr{F}$, it is readily seen

$$\mathcal{D}_{H^{-1}\circ H^+,\rho-10\sigma}=H^{-1}(\mathcal{D}_{H^+,\rho-10\sigma})\subseteq H^{-1}(\mathcal{D}_{H,\rho-8\sigma})=\mathcal{D}_{\rho-8\sigma}$$

and then

(5.56)
$$\begin{aligned}&\left\|b\int_0^1[D\Delta K\circ(H\circ R_{(\alpha+1)\omega}+se)\cdot e]ds\circ H^{-1}\circ H^{+}\right\|_{\rho-12\sigma,s}\\ &\quad\le C\sigma^{-2(s+d/2)}\|D\Delta K\circ(H\circ R_{(\alpha+1)\omega}+e)\|_{\rho-8\sigma}\cdot\epsilon\le C\sigma^{-2(s+d/2)}\|D\Delta K\circ H\|_{\rho-7\sigma}\cdot\epsilon\\ &\quad\le C\sigma^{-(2s+d+1)}\epsilon^2.\end{aligned}$$

Then, similarly to (5.47), we have the new error of $\mathscr{F}$ under the corrections

$$\|\mathscr{F}[K+\Delta K,\mu+\Delta\mu]\circ H^{+}\|_{\rho-12\sigma,s}\le C\frac{\epsilon^{2}}{\sigma^{5(s+d/2)+2}}. \tag{5.57}$$

Let

$$\begin{aligned} E^{+} =&\mathscr{F}[K+\Delta K,\mu+\Delta\mu],\\ \rho^{+} =&\rho-12\sigma,\\ \epsilon^{+} =&C\sigma^{-\kappa}\epsilon^{2}, \end{aligned} \tag{5.58}$$

where $\kappa = 2\nu+5(s+d/2)+2$. Combining the analysis above, we have proven

Lemma 5.5. *Assume* $\|\omega h\|_{\rho}$ *and* $|b|$ *are small enough such that* (5.21) *and* (5.42) *hold. Furthermore, the approximate solutions* K, h, μ *satisfy the non-degeneracy condition in Theorem 3.1. Then if*

$$\frac{C}{\sigma^{\nu+1}}\epsilon<1,$$

we have

$$\|E^{+}\circ H^{+}\|_{\rho^{+},s}\le\epsilon^{+}=C\sigma^{-\kappa}\epsilon^{2},$$

and

$$\|e^{+}\|_{\rho^{+}}\le\epsilon^{+}=C\sigma^{-\kappa}\epsilon^{2}.$$

6. Proof of theorem 3.1

In this section, we complete the proof of our main result Theorem 3.1. The openness of the non-degeneracy condition enables us to iterate the Newton steps. By the analysis in Section 5, we prove the convergence of the iteration sequences. This is very standard in KAM theory.

6.1. **Proof of the convergence.** From the standard techniques in KAM theory, we use the subscript n to denote the n-th step for the Newton iterations. More precisely, we choose the loss of the analyticity domain σ_n as $\sigma_n = 2^{-(n-1)}\sigma$ and $\sigma = \rho/48$. Let $\rho_{n+1} = \rho_n - 12\sigma_{n+1}$ and $\rho_0 = \rho$. Inductively, we assume the errors $E_n = \mathscr{F}[K_n, \mu_n]$ and $e_n = \mathscr{G}[K_n, h_n, \mu_n]$, satisfy $\|E_n \circ H_n\|_{\rho_n,s} \le \epsilon_n$ and $\|e\|_{\rho_n} \le \epsilon_n$, where $K_0 = K, h_0 = h, \mu_0 = \mu$ and $\epsilon_0 = \epsilon$. Noted that K_n and h_n are inductively defined by $K_n = K_{n-1} + \Delta K_{n-1}$ and $h_n = h_{n-1} + \Delta h_{n-1}$. Furthermore, we also assume that

$$\epsilon_n = C\sigma_n^{-\kappa}\epsilon_{n-1}^2. \tag{6.1}$$

Generally, if (6.1) holds for all n, it is easy to show that ϵ_n approaches zero when ϵ is small enough. Indeed, denoting $\tilde{\epsilon}_n = C\sigma^{-\kappa}2^{\kappa(n+1)}\epsilon_n$, from (6.1) one has $\tilde{\epsilon}_{n+1} = \tilde{\epsilon}_n^2$, which implies

$$\tilde{\epsilon}_n = [C(2/\sigma)^{\kappa}\epsilon]^{2^n}. \tag{6.2}$$

Then if

$$C(2/\sigma)^{\kappa}\epsilon < 1, \tag{6.3}$$

ϵ_n obviously approaches zero and satisfies

$$\sum_{n=1}^{\infty} \tilde{\epsilon}_n \le \sum_{n=1}^{\infty} [C(2/\sigma)^{\kappa}\epsilon]^n \le C(2/\sigma)^{\kappa}\epsilon$$

To prove the $(n+1)$-th step, it suffices to verify the conditions in Proposition 5.5. All together, we are led to show the difference $K_n \circ H_n - K \circ H$ and $\omega h_n - \omega h$ are small enough so that the non-degeneracy and contraction conditions hold. Furthermore, we also need to show $D^2 H_n$ and $(DH_n)^{-1}$ are uniformly bounded along all the iterations. Obviously, by (vii) and (x) in Proposition 5.1, one has

$$\begin{aligned}
&\|K_n \circ H_n - K \circ H\|_{\rho_n, s} \\
\le& \sum_{j=1}^{n} \|K_{j-1} \circ H_j - K_{j-1} \circ H_{j-1}\|_{\rho_j, s} + \|\Delta K_{j-1} \circ H_j\|_{\rho_j, s} \\
\le& \sum_{j=1}^{n} C\left(\frac{\epsilon_{j-1}}{\sigma_j^{\nu+s+d/2}} + \frac{\epsilon_{j-1}}{\sigma_j^{s+d/2}}\right) \\
\le& \sum_{j=1}^{\infty} \tilde{\epsilon}_j
\end{aligned} \tag{6.4}$$

and

$$\|\omega h_n - \omega h\|_{\rho_n} \le \sum_{j=1}^{n} \|\omega \Delta h_{j-1}\|_{\rho_j} \le \sum_{j=1}^{\infty} C \frac{\epsilon_{j-1}}{\sigma_j^{\nu}} \le \sum_{j=1}^{\infty} \tilde{\epsilon}_j. \tag{6.5}$$

Likewise, the uniformly boundedness of $D^2 H_n$, $(DH_n)^{-1}$, $P[K_n, h_n, e_n]$ can be proved by applying our detailed analysis in Proposition 5.1. Since ρ_n decreasing to $\rho/2$, for the convergence of H_n and $K_n \circ H_n$, it is sufficient to apply the same estimates in (6.4) and (6.5) to show that both sequences are Cauchy on the uniform analyticity domain $\mathcal{D}_{\rho/2}$, which is an immediate result of the convergence of $\sum_{j=1}^{\infty} \tilde{\epsilon}_j$.

6.2. **Proof of local uniqueness.** For the local uniqueness, we assume that there are two solutions (K_1, h_1, μ_1) and (K_2, h_2, μ_2) of the invariance equation (1.12) and the conjugation equation (1.13), which also satisfy the non-degeneracy conditions.

We can write

$$\mathscr{F}[K_2, \mu_2] = \mathscr{F}[K_1, \mu_1] + D\mathscr{F}[K_1, \mu_1](K_2 - K_1, \mu_2 - \mu_1) + R^{\mathscr{F}} \tag{6.6}$$

and
(6.7)
$$\mathscr{G}[K_2,h_2,\mu_2] = \mathscr{G}[K_1,h_1,\mu_1]+D\mathscr{G}[K_1,h_1,\mu_1](K_2-K_1,h_2-h_1,\mu_2-\mu_1)+R^{\mathscr{G}}$$
where $R^{\mathscr{F}}$ and $R^{\mathscr{G}}$ are the Taylor remainder for $\mathscr{F}$ and $\mathscr{G}$. Let

$$\Delta K = K_2 - K_1, \quad \Delta H = H_2 - H_1 = \omega(h_2 - h_1), \quad \Delta\mu = \mu_2 - \mu_1.$$

We just repeat the computations in (5.50)-(5.55) to estimate $R^{\mathscr{F}}$. It is readily seen that $R^{\mathscr{F}}$ has the same expression (5.50). When composing $R^{\mathscr{F}}$ with H_1, we need to specify the analyticity domain of $(\partial_K\varphi_{pq}\cdot\Delta K)\circ H_1$, $(\partial_\mu\varphi_{pq}\cdot\Delta\mu)\circ H_1$ and $\varphi_{pq}\circ H_1$, where

$$\begin{aligned}(\partial_K\varphi_{pq}\cdot\Delta K)\circ H_1 &= -\omega Dr_{\mu+pq\Delta\mu}\circ(K_1+pq\Delta K)\circ H_1\cdot\Delta K\circ H_1,\\ (\partial_\mu\varphi_{pq}\cdot\Delta\mu)\circ H_1 &= -\omega\partial_\mu r_{\mu+pq\Delta\mu}\circ(K_1+pq\Delta K)\circ H_1\cdot\Delta\mu,\end{aligned}$$

and

$$\begin{aligned}\varphi_{pq}\circ H_1 =&\varphi[K_1,\mu_1]\circ H_1+\varphi_{pq}\circ H_1-\varphi[K_1,\mu_1]\circ H_1\\ =&H_1\circ R_{(\alpha+1)\omega}+\int_0^1 \partial_K\varphi_{pqt}\circ H_1\cdot pq\Delta K\circ H_1+\partial_\mu\varphi_{pqt}\circ H_1\cdot pq\Delta\mu\, dt.\end{aligned}$$

Noticing that, for $0\le t\le 1$, the following estimate

$$\|H_1+t\Delta H-H_2\|_{\hat\rho} = \|(t-1)\Delta H\|_{\hat\rho}\le C\|h_2-h_1\|_{\hat\rho}.$$

holds, which yields

$$\mathcal{D}_{H_1+t\Delta H,\hat\rho-6\hat\sigma}\subset\mathcal{D}_{H_1,\hat\rho-4\hat\sigma}\subset\mathcal{D}_{H_1+t\Delta H,\hat\rho-2\hat\sigma}\subset\mathcal{D}_{H_2,\hat\rho-\hat\sigma} \tag{6.8}$$

if $\|h_2-h_1\|_{\hat\rho}$ is small enough. Immediately, one has

$$\begin{aligned}\|K_2\circ H_1-K_2\circ H_2\|_{\hat\rho-3\hat\sigma,s} &= \|\int_0^1 DK_2\circ(H_1+t\Delta H)\cdot\Delta H\, dt\|_{\hat\rho-3\hat\sigma,s}\\ &\le \sup_{0\le t\le 1}\|DK_2\circ(H_1+t\Delta H)\|_{\hat\rho-3\hat\sigma,s}\cdot|\omega|\cdot\|h_2-h_1\|_{\hat\rho}\\ &\le C\hat\sigma^{-(s+d/2)}\|DK_2\circ H_2\|_{\hat\rho-\hat\sigma}\cdot\|h_2-h_1\|_{\hat\rho}\\ &\le C\hat\sigma^{-(1+s+d/2)}\|K_2\circ H_2\|_{\hat\rho,s}\cdot\|h_2-h_1\|_{\hat\rho}.\end{aligned}$$

which implies
(6.9)
$$\begin{aligned}\|\Delta K\circ H_1\|_{\hat\rho-3\hat\sigma,s} &\le \|K_2\circ H_2-K_1\circ H_1\|_{\hat\rho-3\hat\sigma,s}+\|K_2\circ H_1-K_2\circ H_2\|_{\hat\rho-3\hat\sigma,s}\\ &\le \|K_2\circ H_2-K_1\circ H_1\|_{\hat\rho-3\hat\sigma,s}+C\hat\sigma^{-(1+s+d/2)}\|K_2\circ H_2\|_{\hat\rho,s}\cdot\|h_2-h_1\|_{\hat\rho}\end{aligned}$$

Then the smallness of $\Delta K\circ H_1$ and $\Delta\mu$ in the integrand of $\varphi_{pq}\circ H_1$ yields the analyticity domain inclusion

$$\mathcal{D}_{\varphi_t\circ H_1,\hat\rho-4\hat\sigma}\subset\mathcal{D}_{H_1,\hat\rho-3\hat\sigma} \tag{6.10}$$

for any $0\le t\le 1$.

Let

$$\delta_\rho = \max\left\{ \|K_2 \circ H_2 - K_1 \circ H_1\|_\rho, \|h_2 - h_1\|_\rho, |\mu_2 - \mu_1| \right\}.$$

Combining the equations (6.8)-(6.10), one has

$$\|R^{\mathscr{F}}\|_{\hat\rho - 4\hat\sigma, s} \le C\hat\sigma^{-(2s+d+2)}\delta_{\hat\rho}^2.$$

Similarly, for the Taylor remainder of $\mathscr{G}$, $R^{\mathscr{G}}$ has the same formula with (5.43) and satisfies

$$\|R^{\mathscr{G}}\|_{\hat\rho - 6\hat\sigma} \le C\hat\sigma^{-(2s+d+3)}\delta_{\hat\rho}^2.$$

Repeating the procedure in Section 5, we also obtain the equation (5.29) but replacing E and e by $R^{\mathscr{F}}$ and $R^{\mathscr{G}}$ respectively. Then we have

$$\begin{aligned} &|\mu_2 - \mu_1| \le C\hat\rho^{-(2s+d+3)}\delta_{\hat\rho}^2, \\ &\|\omega \Delta h\|_{2\hat\rho/3} \le C\hat\rho^{-(\nu+2s+d+3)}\delta_{\hat\rho}^2, \\ &\|(K_2 - K_1)\circ H_1\|_{\hat\rho/2, s} \le C\hat\rho^{-(2s+d+3)}\delta_{\hat\rho}^2. \end{aligned} \tag{6.11}$$

Noticing that

$$\begin{aligned} \|K_2 \circ H_2 - K_1 \circ H_1\|_{\hat\rho/2} &\le C\|K_2 \circ H_2 - K_2 \circ H_1\|_{\hat\rho/2,s} + C\|\Delta K \circ H_1\|_{\hat\rho/2,s} \\ &\le C\hat\rho^{-(\nu+3s+\frac{3d}{2}+3)}\delta_{\hat\rho}^2 + C\hat\rho^{-(2s+d+3)}\delta_{\hat\rho}^2, \end{aligned}$$

it is readily seen that

$$\delta_{\hat\rho/2} \le C\hat\rho^{-(\nu+3s+\frac{3d}{2}+3)}\delta_{\hat\rho}^2.$$

Combining the interpolation inequality (2.3) to $\delta_{\hat\rho}$, i.e.,

$$\delta_{\hat\rho}^2 \le C\delta_{\hat\rho/2}\delta_{3\hat\rho/2}$$

we have

$$\delta_{\hat\rho/2} \le \frac{C\delta_{3\hat\rho/2}}{\hat\rho^{\nu+3s+\frac{3d}{2}+3}} \cdot \delta_{\hat\rho/2}.$$

Then if $\delta_{3\hat\rho/2}$ is small enough such that $C\hat\rho^{-(\nu+3s+\frac{3d}{2}+3)}\delta_{3\hat\rho/2} < 1$, we have $\delta_{\hat\rho/2} = 0$, which implies the local uniqueness.

Acknowledgments

We thank Prof. Renato C. Calleja for discussions and Profs. Humphries and Krauskopf for sharing their numerical work in progress. X. He would also thank School of Mathematics at Georgia Inst. of Technology for excellent working conditions as a visiting graduate student. X. He is supported by the China Scholarship Council(File No.201406130018) and NSFC(Grant No. 11271115). R. de la Llave has been supported by NSF Grant DMS-1500943.

This work was initiated while both authors were enjoying the hospitality of the JLU-GT Joint Institute for Theoretical Science. We are very grateful for this opportunity.

Appendix A. Foliation-preserving torus map

There are quite extensive literature on the dynamical properties of torus maps both in mathematics and physics. They appear in ergodic theory , Schrödinger's equation with a quasi-periodic potential, bifurcation of quasi-periodic tori and etc. The original treatment appears in [Arn63, Mos66b, Mos66a]. For a more modern presentation, see [KH95] and the references therein.

As we all know, the universal cover of $\mathbb{T}^d$ is $\mathbb{R}^d$ with the covering map

$$\pi : \mathbb{R}^d \to \mathbb{T}^d, \ \ \pi(x) = x \mod 1.$$

Therefore, for any continuous torus map

$$T : \mathbb{T}^d \to \mathbb{T}^d, \tag{A.1}$$

we can lift T to $\widetilde{T} : \mathbb{R}^d \to \mathbb{R}^d$ such that the diagram commutes as

$$T \circ \pi = \pi \circ \widetilde{T}. \tag{A.2}$$

Moreover, $\widetilde{T}$ has the form as

$$\widetilde{T}(x) = Ax + F(x)$$

where A is a $d \times d$ integer valued matrix and F is a periodic function. It is noted that any torus map has infinitely many lifts, differing by an integer vector.

If ω is a eigenvector of A, any torus map of the form $T_f = A + \omega f$ with periodic scalar function f has the property that it preserves the foliations $\{x + \omega t : x \in \mathbb{R}^d, t \in \mathbb{R}\}$. We refer to such maps as *ω-foliation preserving torus map*. These maps are also called *reparameterization of linear flow*[Fay02]. More particularly, when $T_f = Id + \omega f$, the torus map T_f preserves each equivalence class of $\mathbb{T}^d/\omega\mathbb{R}$, where the equivalence relation is defined by

$$x \sim y \Leftrightarrow x - y \in \omega\mathbb{R}.$$

Indeed, for any $z = x + \omega t$, one has

$$T_f(z) = z + \omega f(z) = x + \omega(t + f(x + \omega t)). \tag{A.3}$$

Furthermore, the set of ω-foliation preserving torus maps has group structure under the composition operator. We denote Diff($\mathbb{T}^d$) the diffeomorphism on the torus $\mathbb{T}^d$.

Remark 1.1. *Note that when ω is irrational($\omega \cdot k \neq 0, \forall k \in \mathbb{Z}^d - 0$), each of the leaves of the foliation is dense. Hence, there is no quotient manifold. The maps that preserve irrational foliations have infinitesimal Lie symmetries, but they cannot been "reduced" to a lower dimensional system.*

Lemma 1.1. *Let Ξ be the subset of* Diff $(\mathbb{T}^d)$*, in which the element T_f has the form of $T_f = Id + \omega f$. Then Ξ is a subgroup of* Diff $(\mathbb{T}^d)$ *under the composition $T_f \circ T_g = T_{g+f\circ T_g}$ and the inverse of T_f is given by $T_{-f\circ T_f^{-1}}$.*

The preservation property reduces the dynamics of F to be essentially one dimensional problem.

The simplest example of ω-foliation preserving torus map is the rigid rotations $R_{\alpha\omega} : \mathbb{T}^d \to \mathbb{T}^d$ defined by

$$R_{\alpha\omega}(x) = (x + \alpha\omega) \mod 1 \tag{A.4}$$

where $\alpha \in \mathbb{R}$. These maps are clearly invertible and analytic and their dynamics are easy to understand. If ω is non-resonant, then each leaves of foliation wind densely on $\mathbb{T}^d$. Due to the simplicity of rotations, one may be interested in which classes of torus map can be conjugated to a rotation torus map like (A.4).

It is also noteworthy that if a ω-foliation preserving torus map T_f can be conjugated to a rotation R_Ω, then there exists a α such that $\Omega = \alpha\omega$. We show a simple argument as a justification. If there exists H and Ω such that $H^{-1} \circ T_f \circ H = R_\Omega$, then

$$\frac{1}{n}(T_f^n - Id) = \frac{1}{n}\sum_{j=0}^{n-1} f \circ T_f^{n-1} \cdot \omega$$

which implies the limit of $(T_f^n - Id)/n$, if exists, has the form of $\alpha\omega$. On the other hand, T_f^n also satisfies

$$T_f^n = H \circ R_{n\Omega} \circ H^{-1} = Id + n\Omega + h(R_{n\Omega} \circ H^{-1}) + H^{-1} - Id$$

which implies $(T_f^n - Id)/n$ approaches Ω since the left terms remain bounded. Thus we have $\Omega = \alpha\omega$. Furthermore, to maintain the foliation preserving structure, we also have $H \in \Xi$, i.e. $H = Id + \omega h$. More precisely, in the reduction procedure of a ω-foliation preserving map, all the transformations should be chosen from the subgroup Ξ.

The conjugation problems of maps of the torus is studied in great detail in [Mos66a]. The analogy with Lie algebras is also discussed. Since the the set of foliation preserving maps is a subgroup of the group of diffeomorphism, there are some differences. The most notable is the fact that to study the conjugacies of foliation preserving maps, we only need to adjust one dimensional parameter, while the general case requires as parameters of the same dimension as the torus. See also [Van02, PdlLV03].

Appendix B. Some technical lemmas

In this appendix, we give several elementary facts used in the formulation of our main results. Firstly, we cite a well-known result on the infinite sum containing the exponential decay terms.

Lemma 2.1. [BMS76] Assume $0 < \delta < 1$ and $\nu > 1$, then

$$\sum_{k\in\mathbb{Z}^d} |k|^\nu e^{-2|k|\delta} < \left(\frac{\nu}{e}\right)^\nu \frac{1}{\delta^{\nu+d}}(1+e)^d. \tag{B.1}$$

Next we show a differentiation formula on the product rule of matrix value functions.

Proposition 2.2. *Given $\omega \in \mathbb{R}^d$ and a matrix $A = (a_{ij})_{d\times d}$. If $V, a_{ij} \in C^1(\mathbb{R}^d, \mathbb{R})$, then*

$$DV \cdot A \cdot \omega = Tr[D(V \cdot A)\cdot\omega) - V \cdot D(A\omega)].$$

Proof: Denoting $C = (c_{ij})_{d\times d} = D(V \cdot A)\cdot\omega$ and $M = (m_{ij})_{d\times d} = V \cdot D(A\omega)$. Then one has

$$c_{ij} = \frac{\partial V}{\partial x_j}\sum_{l=1}^d a_{il}\omega_l + V \cdot \sum_{l=1}^d \frac{\partial a_{il}}{\partial x_j}\omega_l$$

and

$$m_{ij} = V \cdot \sum_{l=1}^d \frac{\partial a_{il}}{\partial x_j}\omega_l.$$

The result is readily obtained by observing that

$$DV \cdot A \cdot \omega = \sum_{j=1}^d \frac{\partial V}{\partial x_j}\sum_{l=1}^d a_{jl}\omega_l.$$

For the formal Fourier series (5.13), we prove the general result.

Proposition 2.3. *Assume $A = (a_{ij})_{d\times d}$ with formal Fourier series as*

$$a_{ij}(\theta) \sim \sum_{k\in\mathbb{Z}^d} \widehat{a}_{ij;k} e^{2\pi i k\cdot\theta}.$$

Then,

$$\mathscr{L}^{-1}Tr[D(A\omega)](\theta) \sim \sum_{k\in\mathbb{Z}^d}\sum_{1\le j\le d}\sum_{1\le l\le d} \frac{2\pi i k_j \omega_l \cdot \widehat{a}_{jl;k}}{2\pi i k\cdot\omega - a - be^{2\pi i k\cdot\Omega}} e^{2\pi i k\cdot\theta} \tag{B.2}$$

where the operator $\mathscr{L}^{-1}$ is given in Section 5.1.1.

The verification is straightforward and thus omitted. When

$$A = V \cdot \omega \cdot (g_1, \cdots, g_d),$$

we have $a_{jl;k} = \omega_j(\widehat{V \cdot g_l})_k$ and inserting the particular coefficients into the above sum yields the expression (5.13).

Appendix C. Proof of Proposition 5.1

The estimates on the corrections are standard in KAM theory. Not surprisingly, the techniques used here are the Cauchy inequality (2.1.1) and the relation of two different norms (2.8) and (2.9). In order to prove the convergence of the Newton iterations, we need to give delicate estimates on the corrections.

Proof of Proposition 5.1: We show some direct computation as follows.

(i) $\|\omega h^+\|_{\rho-7\sigma} \le \|\omega h\|_\rho + \|\omega \Delta h\|_{\rho-7\sigma} \le \|\omega h\|_\rho + C\sigma^{-\nu}\epsilon$.

(ii) $\|\omega D h^+\|_{\rho-8\sigma} \le \|\omega D h\|_{\rho-\sigma} + \|\omega D\Delta h\|_{\rho-8\sigma} \le \sigma^{-1}\|\omega h\|_\rho + C\sigma^{-(\nu+1)}\epsilon$.

(iii)
$$\begin{aligned}\|DH^+ - \mathbb{I}\|_{\rho-8\sigma} \le&\|DH^+ - DH\|_{\rho-8\sigma} + \|DH - \mathbb{I}\|_{\rho-\sigma}\\ \le&\sigma^{-1}\|\omega h\|_\rho + C\sigma^{-(\nu+1)}\epsilon.\end{aligned}$$

(iv) Since $\sigma^{-1}\|\omega h\|_\rho$ is small enough, the inverse of DH does exist on $\mathcal{D}_{\rho-\sigma}$ and satisfies
$$\begin{aligned}\|D^{-1}H - \mathbb{I}\|_{\rho-\sigma} = \|\sum_{n=1}^{\infty}(-\omega Dh)^n\|_{\rho-\sigma} \le \|\omega Dh\|_{\rho-\sigma}\cdot(1-\|\omega h\|_{\rho-\sigma})^{-1}\\ < 2\|\omega Dh\|_{\rho-\sigma} < 2\sigma^{-1}\|\omega h\|_\rho.\end{aligned}$$
Likewise, we also have
$$\|D^{-1}(H+\omega\Delta h) - \mathbb{I}\|_{\rho-8\sigma} \le 2\|\omega Dh^+\|_{\rho-8\sigma} \le 2\sigma^{-1}\|\omega h\|_\rho + C\sigma^{-(\nu+1)}\epsilon$$
if $\sigma^{-\nu}\epsilon$ is sufficient small.

(v) Cauchy inequality from (iv).

(vi)
$$\begin{aligned}\|D^2H^+\|_{\rho-9\sigma} \le&\|D^2H\|_{\rho-3\sigma} + \|D^2H^+ - D^2H\|_{\rho-9\sigma}\\ \le&\sigma^{-2}\|\omega h\|_\rho + \sigma^{-2}\|\omega\Delta h\|_{\rho-7\sigma}.\end{aligned}$$

(vii)
$$\begin{aligned}\|K\circ H^+ - K\circ H\|_{\rho-7\sigma} =& \|\int_0^1 DK\circ(H+q\omega\Delta h)\cdot\omega\Delta h\,dq\|_{\rho-7\sigma}\\ \le& \int_0^1 \|DK\circ H\|_{\rho-2\sigma}\cdot\|\omega\Delta h\|_{\rho-7\sigma}\,dq\\ \le& C\|DK\circ H\|_{\rho-2\sigma}\cdot\sigma^{-\nu}\epsilon.\end{aligned}$$

(viii)
$$\begin{aligned}\|K\circ H^+\|_{\rho-7\sigma} \le&\|K\circ H^+ - K\circ H\|_{\rho-7\sigma} + \|K\circ H\|_\rho\\ \le&[1+C\sigma^{-(\nu+1)}\epsilon]\cdot\|K\circ H\|_\rho.\end{aligned}$$

(ix) Since

$$\begin{aligned}\|D(K\circ H^+)-D(K\circ H)\|_{\rho-8\sigma}\leq&\sigma^{-1}\|K\circ H^+-K\circ H\|_{\rho-7\sigma}\\ \leq&C\sigma^{-(\nu+2)}\|K\circ H\|_\rho\cdot\epsilon\end{aligned}$$

and

$$\begin{aligned}\|D(K\circ H^+)\|_{\rho-8\sigma}\leq&\|D(K\circ H^+)-D(K\circ H)\|_{\rho-8\sigma}+\|D(K\circ H)\|_{\rho-\sigma}\\ \leq&\sigma^{-1}\|K\circ H\|_\rho+C\sigma^{-(\nu+2)}\|K\circ H\|_\rho\cdot\epsilon,\end{aligned}$$

then one has

$$\begin{aligned}\|DK\circ H^+\|_{\rho-8\sigma}\leq&\|D(K\circ H^+)\cdot D^{-1}H^+\|_{\rho-8\sigma}\\ \leq&[\sigma^{-1}+C\sigma^{-(\nu+2)}\cdot\epsilon]\cdot\|K\circ H\|_\rho\cdot(1+2\|\omega Dh^+\|_{\rho-8\sigma}).\end{aligned}$$

(x)

$$\|\Delta K\circ H^+\|_{\rho-12\sigma,s}\leq C\sigma^{-(s+d/2)}\|\Delta K\circ H^+\|_{\rho-10\sigma}\leq C\sigma^{-(s+d/2)}\|\Delta K\circ H\|_{\rho-6\sigma,s}.$$

The following estimates would also need the domain analysis in Lemma 5.3.

(xi) Since

$$\begin{aligned}\|D\Delta K\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}\leq& C\sigma^{-(s+d/2)}\|D\Delta K\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-10\sigma}\\ \leq& C\sigma^{-(s+d/2)}\|D\Delta K\circ H\|_{\rho-8\sigma}\leq C\sigma^{-(s+d/2)}\|D(\Delta K\circ H)\|_{\rho-8\sigma}\cdot\|D^{-1}H\|_{\rho-8\sigma}\\ \leq& C\sigma^{-(1+s+d/2)}\|D^{-1}H\|_{\rho-\sigma}\cdot\|\Delta K\circ H\|_{\rho-6\sigma,s}\end{aligned}$$

and

$$\begin{aligned}\|DK\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}\leq& C\sigma^{-(s+d/2)}\|DK\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-10\sigma}\\ \leq&C\sigma^{-(s+d/2)}\|DK\circ H\|_{\rho-8\sigma}\leq C\sigma^{-(s+d/2)}\|DK\circ H\|_{\rho-\sigma},\end{aligned}$$

one has

$$\begin{aligned}&\|DK^+\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}\\ &\quad\leq\|DK\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}+\|D\Delta K\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}\\ &\quad\leq C\sigma^{-(s+d/2)}\|DK\circ H\|_{\rho-\sigma}+C\sigma^{-(1+s+d/2)}\|D^{-1}H\|_{\rho-\sigma}\cdot\|\Delta K\circ H\|_{\rho-6\sigma,s}.\end{aligned}$$

(xii)

$$\begin{aligned}&\|Dr_{\mu+\Delta\mu}\circ(K+\Delta K)\circ H^+\|_{\rho-12\sigma,s}\cdot\|\Delta K\circ H^+\|_{\rho-12\sigma,s}\\ &\quad\leq C\sigma^{-(2s+d)}\|Dr_{\mu^+}\circ K^+\circ H^+\|_{\rho-10\sigma}\cdot\|\Delta K\circ H\|_{\rho-6\sigma}\\ &\quad\leq C\|Dr\|_{\mathcal{D}^*}\cdot\sigma^{-(2s+d)}\|\Delta K\circ H\|_{\rho-6\sigma,s}.\end{aligned}$$

(xiii) $\|\partial_\mu r_{\mu+\Delta\mu}\circ(K+\Delta K)\circ H^+\cdot\Delta\mu\|_{\rho-12\sigma,s}\leq C\|\partial_\mu r\|_{\mathcal{D}^*}\cdot\epsilon.$

(xiv)

$$\begin{aligned}
&\|(\partial_{KK}\varphi[K^+,\mu^+]\cdot(\Delta K)^{\otimes 2})\circ H^+\|_{\rho-12\sigma,s}\\
&\quad\le \|D^2 r_{\mu+\Delta\mu}\circ(K+\Delta K)\circ H^+\cdot(\Delta K\circ H^+)^{\otimes 2}\|_{\rho-12\sigma,s}\\
&\quad\le C\sigma^{-(s+d/2)}\|D^2 r\|_{\mathcal{D}^*}\cdot\|\Delta K\circ H^+\|^2_{\rho-12\sigma,s}\\
&\quad\le C\|D^2 r\|_{\mathcal{D}^*}\cdot\sigma^{-(3s+3d/2)}\|\Delta K\circ H\|^2_{\rho-6\sigma,s},
\end{aligned}$$

(xv)

$$\begin{aligned}
&\|(\partial_{K\mu}\varphi[K^+,\mu^+]\cdot\Delta K)\circ H^+\Delta\mu\|_{\rho-12\sigma,s}\\
&\quad\le \|D\partial_\mu r_{\mu+\Delta\mu}\circ(K+\Delta K)\circ H^+\cdot\Delta K\circ H^+\|_{\rho-12\sigma,s}\cdot|\Delta\mu|\\
&\quad\le C\sigma^{-(s+d/2)}\|D\partial_\mu r\|_{\mathcal{D}^*}\cdot\|\Delta K\circ H^+\|_{\rho-12\sigma,s}\cdot|\Delta\mu|\\
&\quad\le C\|D\partial_\mu r\|_{\mathcal{D}^*}\cdot\sigma^{-(2s+d)}\|\Delta K\circ H\|_{\rho-6\sigma,s}\cdot\epsilon,
\end{aligned}$$

(xvi)

$$\begin{aligned}
\|\partial_{\mu\mu}\varphi[K^+,\mu^+]\circ H^+(\Delta\mu)^{\otimes 2}\|_{\rho-12\sigma,s}&\le \|D_{\mu\mu}r_{\mu+\Delta\mu}\circ(K+\Delta K)\circ H^+\|_{\rho-12\sigma,s}\cdot|\Delta\mu|^2\\
&\le C\|D_{\mu\mu}r\|_{\mathcal{D}^*}\cdot\sigma^{-(s+d/2)}\cdot\epsilon^2,
\end{aligned}$$

(xvii) Since

$$\begin{aligned}
&\|D^2\Delta K\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}\le C\sigma^{-(s+d/2)}\|D^2\Delta K\circ H\|_{\rho-8\sigma}\\
&\quad\le C\sigma^{-(s+d/2)}\|[D^2(\Delta K\circ H)-D\Delta K\circ H\cdot D^2 H]\cdot(D^{-1}H)\|_{\rho-8\sigma}\\
&\quad\le C\sigma^{-(s+d/2)}\{\|D^2(\Delta K\circ H)\|_{\rho-8\sigma}+\|D\Delta K\circ H\|_{\rho-8\sigma}\cdot\|D^2 H\|_{\rho-2\sigma}\}\cdot\|D^{-1}H\|^2_{\rho-\sigma}\\
&\quad\le C\|D^2 H\|_{\rho-2\sigma}\cdot\|D^{-1}H\|^3_{\rho-\sigma}\cdot\sigma^{-(2+s+d/2)}\|\Delta K\circ H\|_{\rho-6\sigma,s}
\end{aligned}$$

and

$$\begin{aligned}
\|D^2 K\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}&\le C\sigma^{-(s+d/2)}\|D^2 K\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-10\sigma}\\
&\le C\sigma^{-(s+d/2)}\|D^2 K\circ H\|_{\rho-8\sigma}\\
&\le C\sigma^{-(s+d/2)}\|D^2 K\circ H\|_{\rho-2\sigma},
\end{aligned}$$

then

$$\begin{aligned}
&\|D^2 K^+\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}\\
&\quad\le \|D^2 K\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}+\|D^2\Delta K\circ\varphi[K^+,\mu^+]\circ H^+\|_{\rho-12\sigma,s}\\
&\quad\le C\sigma^{-(s+d/2)}\|D^2 K\circ H\|_{\rho-2\sigma}+C\|D^2 H\|_{\rho-2\sigma}\cdot\|D^{-1}H\|^3_{\rho-\sigma}\cdot\sigma^{-(2+s+d/2)}\\
&\qquad\times\|\Delta K\circ H\|_{\rho-6\sigma,s}
\end{aligned}$$

which is bounded if ϵ is sufficient small.

□

References

[Arn63] V. I. Arnol′d. Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspehi Mat. Nauk*, 18(5 (113)):13–40, 1963.

[BMS76] N. N. Bogoljubov, Ju. A. Mitropoliskii, and A. M. Samoĭlenko. *Methods of accelerated convergence in nonlinear mechanics*. Hindustan Publishing Corp., Delhi; Springer-Verlag, Berlin-New York, 1976. Translated from the Russian by V. Kumar and edited by I. N. Sneddon.

[CCCdlL15] Renato Calleja, Alessandra Celletti, Livia Corsi, and Rafael de la Llave. Response solutions for quasi-periodically forced, dissipative wave equations. 2015.

[CCdlL15] Renato Calleja, Alessandra Celletti, and Rafael de la LLave. Whitney regularity, monogenic dependence of solutions in kam theory: An a-posteriori approach. 2015.

[CF80] A. Casal and M. Freedman. A Poincaré-Lindstedt approach to bifurcation problems for differential-delay equations. *IEEE Trans. Automat. Control*, 25(5):967–973, 1980.

[CFdlL03] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces. *Indiana Univ. Math. J.*, 52(2):283–328, 2003.

[CFdlL05] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds. III. Overview and applications. *J. Differential Equations*, 218(2):444–515, 2005.

[dlL01] Rafael de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001.

[dlLO99] R. de la Llave and R. Obaya. Regularity of the composition operator in spaces of Hölder functions. *Discrete Contin. Dynam. Systems*, 5(1):157–184, 1999.

[Fay02] Bassam R. Fayad. Weak mixing for reparameterized linear flows on the torus. *Ergodic Theory Dynam. Systems*, 22(1):187–201, 2002.

[FB76] V. Ī. Fodčuk and M. S. Borteĭ. Quasiperiodic solutions of functional-differential equations. *Ukrain. Mat. Ž.*, 28(3):352–365, 429, 1976.

[Fen77] Neil Fenichel. Asymptotic stability with rate conditions. II. *Indiana Univ. Math. J.*, 26(1):81–93, 1977.

[Fen74] Neil Fenichel. Asymptotic stability with rate conditions. *Indiana Univ. Math. J.*, 23:1109–1137, 1973/74.

[HCF+15] Alex Haro, Marta Canadell, Jordi-Lluıs Figueras, Alejandro Luque, and Josep Maria Mondelo. *The parameterization method for invariant manifolds: from theory to effective computations*. 2015.

[HCHS14] D.A. Humphries, A.R.and Bernucci, Renato Calleja, N. Homayounfar, and M. Snarski. Periodic solutions of a singularly perturbed delay differential equation with two state-dependent delays. 2014.

[HdlL15] Xiaolong He and Rafael de la Llave. Construction of quasi-periodic solutions of state-dependent delay differential equations by the parameterization method i: finite differentiable, hyperbolic case. 2015.

[HDMU12] A. R. Humphries, O. A. DeMasi, F. M. G. Magpantay, and F. Upham. Dynamics of a delay differential equation with multiple state-dependent delays. *Discrete Contin. Dyn. Syst.*, 32(8):2701–2727, 2012.

[Her80] Michael-Robert Herman. Résultats récents sur la conjugaison différentiable. In *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)*, pages 811–820. Acad. Sci. Fennica, Helsinki, 1980.

[KH95] Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.

[Kri14] Tibor Krisztin. Analyticity of solutions of differential equations with a threshold delay. In *Recent advances in delay differential and difference equations*, volume 94 of *Springer Proc. Math. Stat.*, pages 173–180. Springer, Cham, 2014.

[LdlL09] Xuemei Li and Rafael de la Llave. Construction of quasi-periodic solutions of delay differential equations via KAM techniques. *J. Differential Equations*, 247(3):822–865, 2009.

[LS07] Jian Liu and Jianguo Si. Analytic solutions for a class of differential equation with delays depending on state. *Applied Mathematics and Computation*, 186(1):261 – 270, 2007.

[LY12] Xuemei Li and Xiaoping Yuan. Quasi-periodic solutions for perturbed autonomous delay differential equations. *J. Differential Equations*, 252(6):3752–3796, 2012.

[Mey75] K. R. Meyer. The implicit function theorem and analytic differential equations. In *Dynamical systems—Warwick 1974 (Proc. Sympos. Appl. Topology and Dynamical Systems, Univ. Warwick, Coventry, 1973/1974; presented to E. C. Zeeman on his fiftieth birthday)*, pages 191–208. Lecture Notes in Math., Vol. 468. Springer, Berlin, 1975.

[MKW14] F. M. G. Magpantay, N. Kosovalić, and J. Wu. An age-structured population model with state-dependent delay: derivation and numerical integration. *SIAM J. Numer. Anal.*, 52(2):735–756, 2014.

[Mos66a] Jürgen Moser. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:499–535, 1966.

[Mos66b] Jürgen Moser. A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:265–315, 1966.

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

[Mos01] Jürgen Moser. *Stable and random motions in dynamical systems*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2001. With special emphasis on celestial mechanics, Reprint of the 1973 original, With a foreword by Philip J. Holmes.

[MPN14] John Mallet-Paret and Roger D. Nussbaum. Analyticity and nonanalyticity of solutions of delay-differential equations. *SIAM J. Math. Anal.*, 46(4):2468–2500, 2014.

[Nus73] Roger D. Nussbaum. Periodic solutions of analytic functional differential equations are analytic. *Michigan Math. J.*, 20:249–255, 1973.

[PdlLV03] Nikola P. Petrov, Rafael de la Llave, and John A. Vano. Torus maps and the problem of a one-dimensional optical resonator with a quasiperiodically moving wall. *Phys. D*, 180(3-4):140–184, 2003.

[Rüs75] Helmut 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 (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, pages 598–624.Lecture Notes in Phys., Vol. 38. Springer, Berlin, 1975.

[SB03] A. M. Samoilenko and E. P. Belan. Quasiperiodic solutions of differential-difference equations on a torus. *J. Dynam. Differential Equations*, 15(2-3):305–325, 2003. Special issue dedicated to Victor A. Pliss on the occasion of his 70th birthday.

[SdlL12] Xifeng Su and Rafael de la Llave. KAM theory for quasi-periodic equilibria in one-dimensional quasi-periodic media. *SIAM J. Math. Anal.*, 44(6):3901–3927, 2012.

[Spr79] Vladimir G. Sprindžuk. *Metric theory of Diophantine approximations*. V. H. Winston & Sons, Washington, D.C.; A Halsted Press Book, John Wiley & Sons, New York-Toronto, Ont.-London, 1979. Translated from the Russian and edited by Richard A. Silverman, With a foreword by Donald J. Newman, Scripta Series in Mathematics.

[Van02] J. Vano. *A Whitney-Zehnder implicit function theorem*. PhD thesis, University of Texas at Austin, 2002. http://www.ma.utexas.edu/mp_arc 02–276.

College of Mathematics and Econometrics, Hunan University, Changsha Hunan 410082, People's Republic of China

School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta GA 30332, USA
E-mail address: hexiaolong@hnu.edu.cn, xiaolong.he@math.gatech.edu

School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta GA 30332, USA
E-mail address: rafael.delallave@math.gatech.edu