GEVREY ESTIMATES FOR ASYMPTOTIC EXPANSIONS OF TORI IN WEAKLY DISSIPATIVE SYSTEMS

ADRIÁN P. BUSTAMANTE AND RAFAEL DE LA LLAVE

Abstract. We consider a singular perturbation for a family of analytic symplectic maps of the annulus possessing a KAM torus. The perturbation introduces dissipation and contains an adjustable parameter. By choosing the adjustable parameter, one can ensure that the torus persists under perturbation. Such models are common in celestial mechanics. In field theory, the adjustable parameter is called *the counterterm* and in celestial mechanics, the *drift*. It is known that there are formal expansions in powers of the perturbation both for the quasi-periodic solution and the counterterm.

We prove that the asymptotic expansions for the quasiperiodic solutions and the counterterm satisfy Gevrey estimates. That is, the n-th term of the expansion is bounded by a power of $n!$. The Gevrey class (the power of $n!$) depends only on the Diophantine condition of the frequency and the order of the friction coefficient in powers of the perturbative parameter.

The method of proof we introduce may be of interest beyond the problem considered here. We consider a modified Newton method in a space of power expansions. As it is custumary in KAM theory, each step of the method is estimated in a smaller domain. In contrast with the KAM results, the domains where we control the Newton method shrink very fast and the Newton method does not prove that the solutions are analytic. On the other hand, by examining carefully the process, we can obtain estimates on the coefficients of the expansions and conclude the series are Gevrey.

1. Introduction

Hamiltonian systems with small dissipation appear as models of many problems of physical interest. Notably, dissipation is a small effect in astrodynamics of planets and satellites [MNF87, Cel13] [1]. In the design of many mechanical devices, eliminating friction is a design goal which is never completely accomplished. Hamiltonian systems with friction also appear as Euler-Lagrange equations of discounted functionals which are natural in finance and in the receding horizon problem in control theory. In such a case the limit of zero discount (equivalent to the limit of zero friction) is of interest. See [Ben88, MHER95, ISM11, DFIZ16] for different studies of the zero dissipation limit in calculus of variations and in control.

Since the friction is small, it is natural to try to study such systems using perturbation theory. Nevertheless, adding a small friction is a very singular perturbation, and periodic/quasi-periodic orbits may disappear for arbitrarily small values or the perturbation. In contrast with Hamiltonian systems that often have sets of quasi-periodic orbits of positive measure (KAM theorem), for dissipative forced systems, there are few periodic or quasi-periodic orbits. These quasi-periodic orbits are known to persist only if one can adjust parameters in the system [Mos67, BHS96, Sev99]. As discussed very clearly in [Mos73], the number of parameters needed is affected by the geometric properties of the systems considered.

1991 *Mathematics Subject Classification.* 35C20, 70K70, 70K43, 37J40, 34K26,30E10.

Key words and phrases. Gevrey estimates, Dissipative systems, quasi-periodic solutions.

Both authors have been supported by NSF grant DMS 1800241.

[1]A problem in astrodynamics which motivate us is the *spin orbit problem* describing approximately the motion of an oblate planet, subject to tidal friction, in a Keplerian orbit [Cel91]

In recent times, for some particular types of dissipative systems – the conformally symplectic systems, see Definition 1 – there is a very systematic KAM theory [CCdlL13] based on geometric arguments. The examples mentioned above (Hamiltonian systems with friction proportional to the momentum and Euler-Lagrange equations of exponetially discounted variational principles) are conformally symplectic. This theory, once we fix a frequency, predicts the changes of parameters and the changes in the solutions needed to obtain a quasi-periodic solution of the prescribed frequency.

The goal of this paper is to study the singular perturbation theories in which the perturbation introduces dissipation.

There are several studies of the singular perturbation theories in dissipation which are particularly relevant for us: The paper [CCdlL17] shows that if one fixes a Diophantine frequency ω (see Definition 11), considers a Hamiltonian system – not necessarily integrable – with a quasi-periodic solution of frequency ω, and introduces a conformally symplectic perturbation (see Definition 1), then there is a (unique under a natural normalization) formal power series expansion for the quasi-periodic solution of frequency ω and for the drift parameter. These series are very similar to the Lindstedt series of classical mechanics. The paper [CCdlL17] also showed that the formal Lindstedt series is the asymptotic expansion of a true solution defined in a complex domain of parameters that does not include any ball around zero (giving an indication that the power series may be divergent). The paper [BC19] studied numericaly these Lindstedt series in a concrete example and the possible domain of analyticity of the function (using Padé as well as non-perturbative methods). The numerical studies in [BC19] lead to the remarkable conjecture that, in the cases examined, the formal power series giving the quasiperiodic solution and the forcing are Gevrey (see Definition 8).

In this paper, for some class of maps (we require that the system is conformaly symplectic and that the non-linearity is a trig. polynomial) we show that the conjecture in [BC19] is true and that the series obtained are indeed Gevrey. The Gevrey class can be bounded depending only on the Diophantine condition of the frequency ω (and the order of the friction in the dissipation). See Theorem 18.

The method of proof we introduce may be of interest beyond the problem considered here and we hope that there are other applications. We consider a Newton method in the space of power expansions. As in KAM theory, each step of the quadratically convergent method is estimated in a domain smaller than the domain of the previous steps. In contrast with KAM theory, the domains where we control the results shrink very fast to a point, so that, at the end we do not obtain any analytic function. On the other hand, by examining carefully the process, we can obtain estimates on the coefficients of the expansions.

Our hypothesis that the non-linearity is a trigonometric polynomial ensures that the coefficients of order N do not change after $\log_2(N)$ steps of the Newton method, so that one can use Cauchy estimates in the domain that is under control after $\log_2(N)$ steps to obtain estimates on the Nth coefficient.

We hope that the hypothesis that the non-linearity is a trigonometric polynomial can be removed at the price of estimating the change of the coefficients in subsequent iterations, but a proof would require a new set of estimates that – if indeed possible – would lengthen the exposition and obscure the main ideas.

The Newton method acting on power series is patterned after the Newton method used in [CCdlL13]. This Newton method takes advantage of remarkable cancellations related to the geometry and introduces the corrections to the torus additively (rather than making changes of variables). The fact that the Newton method in [CCdlL13] does not involve changes of variables makes it possible to lift it to formal power series. We will present full details later.

For simplicity in the treatment, we will deal with maps since the geometric arguments are simpler. The same arguments apply for differential equations, but they are more elaborate. Besides adapting the proof of maps to the case of ODE's, one can deduce rigorously the results for differential equations from the results for maps by taking time-T maps. Note that in this case, the fact that the non-linearity in the time-T map is a trig. polynomial is difficult to express in terms of the original ODE. This is another reason why we would like eventually to get rid of that hypothesis.

1.1. **A preview of the main result.** A model to keep in mind is the so-called dissipative standard map $f_{\varepsilon,\mu_\varepsilon} : \mathbb{T} \times \mathbb{R} \longrightarrow \mathbb{T} \times \mathbb{R}$ given by

$$f_{\varepsilon,\mu_\varepsilon}(x, y) = (x + \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V'(x), \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V'(x)) \tag{1.1}$$

In (1.1), the physical meaning of $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$, $\alpha \in \mathbb{N}$, is dissipation and μ_ε, called the *drift* parameter, has the physical meaning of a forcing. Our assumption on the non-linearity amounts to V being a trigonometric polynomial. The model (1.1) is indeed conformally symplectic in the sense of Definition 1 (see below). The map (1.1) is the model that was used in the numerical experiments in [BC19].

Note that for $\varepsilon = 0$, the map (1.1) is integrable. The integrability of the map at $\varepsilon = 0$ does not play any role in the theoretical results in [CCdlL17], the only assumption needed in [CCdlL17] is that map for $\varepsilon = 0$ is symplectic and has as an invariant torus. For the numerical study in [BC19], the fact that the map for $\varepsilon = 0$ is integrable leads to much more efficient algorithms. In this paper, we will not use explicitly the integrability for $\varepsilon = 0$, but this seems to be the only case where it is possible to verify the assumption on the nonlinearity being a trig polynomial (yet another reason to try to get rid of that hypothesis).

The main result of this paper, Theorem 18, establishes the Gevrey character of the formal power series expansions for the drift parameter μ_ε and for the quasi-periodic orbit of frequency ω of the map (1.1). The rigorous formulation of the main Theorem is given in Section 3, the statements of the main results can be better understood after some preliminary definitions and remarks are given (see Section 2). Here we give an informal statement of our main result: Given a Diophantine frequency ω, the coefficients of the formal power series expansions $\sum K_n\varepsilon^n$ and $\sum \mu_n\varepsilon^n$ for the quasi-periodic orbit and the drift parameter, respectively, satisfy the following Gevrey estimates

$$\|K_n\| \le CR^n n^{(2\tau/\alpha)n} \qquad |\mu_n| \le CR^n n^{(2\tau/\alpha)n}$$

where τ depends on the Diophantine type of ω (see Definition 11) and α is the order of the dissipation $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$.

The model (1.1) can be thought as a numerical time step – using a Verlet-like method – of the spin-orbit problem

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\mu y + \lambda + V'(x) \end{aligned} \tag{1.2}$$

1.2. **Organization of the paper.** The paper is organized as follows. In Section 2 we collect some standard definitions and we also define the function spaces in which the iterative procedure takes place. Also, in the same section we present some geometric identities which allow us to solve the linearized equations of the modified Newton method. In Section 3 we state Theorem 18 and Lemma 22, which are the main results of the paper and establish the Gevrey character of the perturbative expansions of the quasi periodic orbits.

The proof of Theorem 18 is based on a quasi Newton method. In Section 4 we formulate the iterative step of this Newton method, while in Section 5 we provide estimates for the corrections and the new error at one step of the method. Finally, in Section 6, using a KAM like argument, we give estimates for any step of the Newton like procedure and, with them, a proof of Lemma 22 is given establishing the Gevrey character of the perturbative expansions.

2. Preliminaries

In this section we introduce the notations, collect some standard definitions including the Banach spaces and their norms that enter in this paper. This section should be used as a reference.

2.1. **Symplectic properties.**

Let $\mathcal{M} = \mathbb{T}^d \times B$, $B \subseteq \mathbb{R}^d$; endowed with an exact symplectic form Ω. Note that the manifold $\mathcal{M}$ is Euclidean (i.e. the tangent bundle is trivial) and we can compare vectors in different tangent spaces. This is crucial in KAM theory.

We denote by J the matrix associated to the symplectic form Ω, i.e., in coordinates we have $\Omega_x(u, v) = (u, J(x)v)$ where $(\cdot, \cdot)$ denotes the inner product for any $u, v \in T_x\mathcal{M}$. Note that J depends on the choice of the inner product.

Definition 1. *We say that a diffeomorphism defined on an symplectic manifold $(\mathcal{M}, \Omega)$ is conformally symplectic when*

$$f^*\Omega = \lambda\Omega$$

for a number λ, where f^* denotes the standard pull back on forms.

The map (1.1) is conformally symplectic with the conformal factor $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$ and the standard symplectic form $\Omega = dx \wedge dy$ on the cylinder $\mathbb{T} \times \mathbb{R}$.

2.2. **Banach spaces of analytic functions.**

2.2.1. *Analytic functions on the torus.* Given $\rho > 0$ we define the complex extension of the d-dimensional torus as

$$\mathbb{T}^d_\rho = \left\{ z \in \mathbb{C}^d/\mathbb{Z}^d \mid \mathrm{Re}(z_j) \in \mathbb{T},\ |\mathrm{Im}(z_j)| \le \rho \right\}$$

and denote $\mathcal{A}_\rho$ as the vector space of analytic functions defined $\mathrm{int}(\mathbb{T}^d_\rho)$ which can be extended continuously to the boundary of $\mathbb{T}^d_\rho$. $\mathcal{A}_\rho$ is endowed with the norm

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

which makes it into a Banach space.

For vector valued functions, $g = (g_1, g_2, ..., g_d)$, we define the norm

$$\|g\|_\rho = \sqrt{\|g_1\|^2_\rho + \|g_2\|^2_\rho + ... + \|g_d\|^2_\rho}$$

and for $n_1 \times n_2$ matrix valued functions, G, we define

$$\|G\|_\rho = \sup_{v \in \mathbb{R}^{n_2}, |v|=1} \sqrt{\sum_{i=1}^{n_1} \left(\sum_{j=1}^{n_2} \|G_{ij}\|_\rho v_j \right)^2}.$$

We will also need to work with functions of two variables. Denoting $B_\gamma(0) \subseteq \mathbb{C}$ the open ball with center zero and radius γ in the complex plane, define

$$\mathcal{A}_{\rho,\gamma} = \left\{ K : B_\gamma(0) \to \mathcal{A}_\rho \;\middle|\; K \text{ is analytic in } B_\gamma(0) \text{ and can be extended continuously to } \overline{B_\gamma(0)} \right\}$$

endowed with the norm

$$\|K\|_{\rho,\gamma} := \sup_{|\varepsilon| \le \gamma} \|K(\varepsilon)\|_\rho .$$

It is well known that with the norms $\|\cdot\|_{\rho,\gamma}$ and $\|\cdot\|_\rho$ the spaces $\mathcal{A}_{\rho,\gamma}$ and $\mathcal{A}_\rho$ are Banach algebras.

To discuss analyticity properties, we will need to deal with complex values of all the arguments. For phyical applications, we need mainly real variables. Hence, it will be important that the functions we consider have the property that they yield real values for real arguments. The functions that satify this property (real valued for real arguments) is a closed (real) subspace of the above

Banach spaces. All the constructions we use have the property that when applied to real valued functions, they produce real valued functions.

Note that we can think of functions $\mathcal{A}_{\rho,\gamma}$ as analytic functions on $B_\gamma(0)$ taking values on a space of analytic functions of the torus. This point of view is consistent with the interpretation that we are considering families of problems and we are seeking families of solutions.

For typographical reasons from now on we will use the following notation. Given $K \in \mathcal{A}_{\rho,\gamma}$ we denote $K_\varepsilon(\theta) = K(\theta, \varepsilon) := (K(\varepsilon))(\theta)$.

Definition 2. *Let $\mathcal{B}$ a Banach space. Given an analytic function $g : B_\gamma(0) \subseteq \mathbb{C} \longrightarrow \mathcal{B}$, and $n \geq 0$, we say $g(\varepsilon) \sim \mathcal{O}(|\varepsilon|^n)$ if and only if there exists $C > 0$ such that*

$$\|g(\varepsilon)\| \leq C|\varepsilon|^n$$

for ε small enough. Equivalently, $g(\varepsilon) \sim \mathcal{O}(|\varepsilon|^n)$ if and only if $g(\varepsilon) = \sum_{k=n}^{\infty} g_k \varepsilon^k$ for ε small enough and $g_k \in \mathcal{B}$.

2.2.2. *Cauchy estimates.* We recall the classical Cauchy inequalities, see [SZ65].

Lemma 3. *For any $0 < \delta \leq \rho$ and for any function $f \in \mathcal{A}_\rho$ we have*

$$\|D^n f\|_{\rho-\delta} \leq C\delta^{-n} \|f\|_\rho ,$$

where D^n denotes the n-th derivative and

$$|\hat{f}_k| \leq e^{-2\pi|k|\rho} \|f\|_\rho$$

where $|k| = |k_1| + |k_2| + \cdots + |k_n|$ and $\hat{f}$ denotes the Fourier coefficient of f with index k.

As mentioned above we will be working with functions depending upon two variables. The following are Cauchy inequalities in the second variable, ε.

Lemma 4. *For any $0 < r \leq \gamma$ and any function $f \in \mathcal{A}_{\rho,\gamma}$ such that $f_\varepsilon(\theta) = \sum_{n=0}^{\infty} f_n(\theta)\varepsilon^n$ we have*

$$\|f_n\|_\rho \leq \frac{1}{r^n} \|f\|_{\rho,r} .$$

Proof. By Cauchy integral formula

$$f_n(\theta) = \frac{1}{n!}\frac{d^n}{d\varepsilon^n} f(\theta,\varepsilon)\Big|_{\varepsilon=0} = \frac{1}{2\pi i}\int_{|\xi|=r} \frac{f(\theta,\xi)}{\xi^{n+1}} d\xi = \frac{1}{2\pi r^n}\int_0^{2\pi} \frac{f(\theta, re^{i\phi})}{e^{in\phi}} d\phi,$$

thus, $|f_n(\theta)| \leq \dfrac{1}{r^n} \sup\limits_{|\varepsilon|\leq r} |f(\theta,\varepsilon)|$ and $\|f_n\|_\rho \leq \dfrac{1}{r^n} \|f\|_{\rho,r}$. □

Corollary 5. *Assume that $\Delta \in \mathcal{A}_{\rho,\gamma}$ is such that $\Delta_\varepsilon = \sum_{n=N+1}^{\infty} \Delta_n \varepsilon^n$. Let $a, b \in \mathbb{N}$ such that $N \leq a < b \leq \infty$ and denote $\Delta_\varepsilon^{(a,b]} = \sum_{n=a+1}^{b} \Delta_n \varepsilon^n$. Then, for all $0 < r < 1$ we have*

$$\left\|\Delta^{(a,b]}\right\|_{\rho,r\gamma} \leq \frac{r^{a+1}}{1-r} \|\Delta\|_{\rho,\gamma} .$$

Remark 6. *Note that the estimate in Corollary 5 only depends on a, associated with the order of the first term in the expansion of $\Delta^{(a,b]}$.*

2.3. **Formal power series.**

2.3.1. *General definitions.* Formal power series expansions are just expressions of the form

$$\sum_n a_n \varepsilon^n$$

where a_n belong to a Banach space, sometimes a_n are just scalars.

Formal power series are not meant to converge nor to represent a function. They can, however be added, multiplied (using the Cauchy formula for product; note that for a fixed degree, computing the coefficients involves only a finite sum) or substituted one into another.

One can form equations among formal power series. The meaning is, of course, that the coefficients on each side should be the same. This is extremely useful in many areas of mathematics, notably combinatorics. See [Car95], [Cos09] for more details on formal power series.

Many perturbation expansions in Physics or in applied mathematics are based precisely into formulating the solutions of the equations of motion as formal power series and requiring that the equations of motion are satisfied in the sense of power series. Notably, the Lindstedt series were in standard use in astronomy even if they were only shown to converge for some frequencies in [Mos67].

2.3.2. *Asymptotic expansions.* For formal power series, a notion weaker that convergence of the series to a function is that the series is asymptotic to a function.

Definition 7. *We say that a formal power series $\sum a_n \varepsilon^n$ with coefficients a_n in a Banach space X, is an asymptotic expansion to a function $\phi : \mathcal{D} \to X$ when for all $N \in \mathbb{Z}$, there exists C_N such that for all $\rho < \rho_0$*

$$\sup_{\varepsilon \in \mathcal{D}, |\varepsilon| \le \rho} \left\| \sum_{n=0}^{N} a_n \varepsilon^n - \phi(\varepsilon) \right\| \le C_N \rho^{N+1}$$

If the domain $\mathcal{D}$ does not include any ball centered at zero, even if the function ϕ is analytic and bounded on $\mathcal{D}$, this does not imply that the series converges.

Given a function ϕ, the associated expansions may be non unique. The Cauchy example

$$\phi(\varepsilon) = \exp(-\varepsilon^{-2}) \tag{2.1}$$

has an identically zero asymptotic expansion on a domain

$$\mathcal{D}_\delta = \{\varepsilon : |\mathrm{Arg}(\varepsilon)| < \delta\} \tag{2.2}$$

when $\delta < \pi$.

Note that the definition of asymptotic involves the domain $\mathcal{D}$. A series may be asymptotic to a function in a domain but not in a larger domain. For example the zero series is asymptotic to the the Cauchy example 2.1 in the domains $\mathcal{D}_\delta$ as in (2.2) when $\delta < \pi$, but not when $\delta > \pi$.

2.3.3. *Gevrey formal expansions.* Given a formal power series, even if it diverges, it is interesting to study how fast the coefficients grow. The following definition captures some speed of growth that is weaker than convergence, but which nevertheless appears naturally in many applied problems.

Definition 8. *Let $\beta, \rho > 0$. We say that a power series expansion $f = \sum_{n=0}^{\infty} f_n(\theta)\varepsilon^n$, with $f_n \in \mathcal{A}_\rho$, belongs to a Gevrey class (β, ρ) if and only if there exist constants $C \ge 0$, $R \ge 0$, and $n_0 \in \mathbb{N}$ such that*

$$\|f_n\|_\rho \le C R^n n^{\beta n} \quad \textit{for } n \ge n_0, \tag{2.3}$$

and we denote $f \in \mathcal{G}^\beta_\rho$.

Similarly, we say that a power series expansion $\mu = \sum_{n=0}^{\infty} \mu_n \varepsilon^n$, with $\mu_n \in \mathbb{C}^d$, belongs to a Gevrey class β if and only if there exist constants $C \ge 0$, $R \ge 0$, and $n_0 \in \mathbb{N}$ such that

$$|\mu_n| \le C R^n n^{\beta n} \quad \textit{for } n \ge n_0, \tag{2.4}$$

and we denote $\mu \in \mathcal{G}^{\beta}$.

Remark 9. *It is well known that (2.3) in Definition 8 is equivalent to the inequality*

$$\|f_n\|_{\rho} \leq C R^n (n!)^{\beta} \quad \textit{for } n \geq n_0$$

which, in turn, implies the series $\sum_{n=0}^{\infty} \frac{f_n(\theta)}{(n!)^{\beta}} \varepsilon^n$ converges in $\mathcal{A}_{\rho}$ with positive radius of convergence.

This remark makes a connection with the theory of Borel summability. If a series is Gevrey, under some extra conditions, the Borel transform produces a function that is analytic in a sector and the series is asymptotic to this function. See [CGGG07], [Cos09].

Remark 10. *The class of functions that around each point have expansions satisfying Definition 8 has received a lot of interest recently since those functions are related to many deep theorems of Dynamical Systems (KAM, Nekhoroshev). Similar theories (e.g. hypoellipticity) also admit Gevrey classes as natural regularity.*

This paper goes in a different direction. Even if we start with an analytic problem – indeed polynomial! – several objects of interest are only Gevrey. The phemenon that Analytic problems have only Gevrey solutions has appeared in other contexts in dynamics, notably in the study of singular perturbations [CDRSS00]*, the regularity of attractors and fast-slow systems* [FT89, CD91, Bae95]*. Closer to us, in dependence on parameters of solutions of non-linear problems,* [Sau92, Lin92]*, dependence of KAM tori in the frequency* [Pop00]*, or in the theory of parabolic manifolds* [BH08, BFM17]*.*

2.3.4. *A property from number theory.* In KAM theory, some number theoretical properties of frequencies play an important role. We will use the standard:

Definition 11. *For $\nu, \tau > 0$, we say $\omega \in \mathbb{R}^d$ is Diophantine of type (ν, τ) if $\left|e^{2\pi i k\cdot\omega} - 1\right| \geq \nu |k|^{-\tau}$.*

We denote $\omega \in \mathcal{D}(\nu, \tau)$.

2.4. Quasi-periodic orbits.

A quasi-periodic sequence $\{x_n\}_{n\in\mathbb{Z}}$ of frequency $\omega \in \mathbb{R}^d$ in a Euclidean manifold is a sequence which can be expressed in terms of Fourier series.

$$x_n = \sum_{k\in\mathbb{Z}^d} e^{2\pi i k\cdot\omega n} \hat{x}_k = K(n\omega)$$

where $K(\theta) = \sum_{k\in\mathbb{Z}^d} e^{2\pi i k\cdot\theta} \hat{x}_k$.

We can think of the function K as an embedding of the torus $\mathbb{T}^d$ into phase space. If ω does not have any resonances (i.e. $k \cdot \omega \neq 0$ for $k \in \mathbb{Z}^d \setminus \{0\}$, which can always be arranged by reducing d if there is one), then $\{\omega n\}_{n\in\mathbb{Z}}$ is dense on the torus. The map K is often called the *hull function*.

If x_n is an orbit of a map, $x_{n+1} = f(x_n)$ we see that $K(n\omega + \omega) = f(K(n\omega))$. Since $\{\omega n\}_{n\in\mathbb{Z}}$ is dense, this is equivalent to

$$K(\theta + \omega) = f(K(\theta)) \;\forall \theta \in \mathbb{T}^d \tag{2.5}$$

Hence, we see that the set $K(\mathbb{T}^d)$, the image of the standard torus under the embedding K is invariant under f. So, it is customary to describe quasi-periodic solutions as *invariant tori*.

The problem of given a map finding a quasi-periodic solution of frequency ω can be formulated as finding an embedding K solving (2.5). The equation (2.5) will be our fundamental tool to characterize quasi-periodic orbits.

2.5. Set-up of the problem. The invariance equation.

In this section, we describe informally the geometric set up and the geometric meaning of the formulation of our problem. The precise formulation of the main result of this paper (Theorem 22) will be presented in Section 3.

We will be mainly concerned with an analytic family of maps $f_{\varepsilon,\mu} : \mathcal{M} \longrightarrow \mathcal{M}$, such that

$$f^*_{\varepsilon,\mu}\Omega = \lambda(\varepsilon)\Omega$$

where $\varepsilon \in \mathbb{C}$ is a small parameter, $\mu \in \Lambda \subseteq \mathbb{C}^d$ is an internal parameter (the drift parameter), and $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$.

A good example to keep in mind is the dissipative standard map presented in (1.1). Note that, for $\varepsilon = 0$ and for each μ, the maps $f_{0,\mu}$ are symplectic because $\lambda(0) = 1$.

The main assumption in the main Lemma, Lemma 22, is that the map f_{0,μ_0} has an invariant torus in which the motion is a rotation of frequency ω which is Diophantine (see Definition 11). Note that the drift parameter, μ, is chosen to guarantee the persistence of a quasi periodic orbit of a given frequency ω, so we also consider $\mu = \mu_\varepsilon$.

Following the discussion in Section 2.4 and, in particular (2.5), we see that finding a quasi-periodic orbit for $f_{\varepsilon,\mu_\varepsilon}$ is equivalent to finding families of embeddings K_ε and families of parameters μ_ε in such a way that

$$f_{\varepsilon,\mu_\varepsilon} \circ K_\varepsilon(\theta) = K_\varepsilon(\theta + \omega) \tag{2.6}$$

Equation (2.6) should be interpreted as, given the family $f_{\varepsilon,\mu}$ and the frequency ω finding $\mu_\varepsilon, K_\varepsilon$. For this work, the sense in which (2.6) is meant to hold is the meaning of formal power series (the coefficients of ε^n on both sides of (2.6) are identical for all n, as it is customary in the study of Lindstedt series).

Note that the equation (2.6) is highly underdetermined. If $\mu_\varepsilon, K_\varepsilon$ is a solution, changing θ into $\theta + \sigma_\varepsilon$, we obtain that $\mu_\varepsilon, \tilde{K}_\varepsilon$ is also a solution where $\tilde{K}_\varepsilon(\theta) = K_\varepsilon(\theta + \sigma_\varepsilon)$. This change of variables has the physical meaning of choosing a change of origins in the torus.

2.6. **Automatic reducibility.** As it is noted in [CCdlL13], a very useful property of conformally symplectic systems is that solutions to equation (2.6) satisfy the so-called *automatic reducibility*, that is, in a neighborhood of an invariant torus, one can find a system of coordinates in which the linearization of the evolution has constant coefficients.

Lemma 12. *Let $f_\mu : \mathcal{M} \longrightarrow \mathcal{M}$, such that, $f_\mu^*\Omega = \lambda\Omega$, and $K : \mathbb{T}^d \longrightarrow \mathcal{M}$ such that $f_\mu \circ K(\theta) = K(\theta + \omega)$ with ω an irrational vector. If $\mathcal{N} = (DK^\top DK)^{-1}$, then, the $2d \times 2d$ matrix*

$$M(\theta) = \left[DK(\theta) | J^{-1} \circ K(\theta) DK(\theta) \mathcal{N}(\theta)\right] \tag{2.7}$$

satisfies

$$Df_\mu \circ K(\theta) M(\theta) = M(\theta + \omega) \begin{pmatrix} \mathrm{Id} & S(\theta) \\ 0 & \lambda\,\mathrm{Id} \end{pmatrix} \tag{2.8}$$

where $\mathrm{Id} \in \mathbb{R}^{d\times d}$ *and $S(\theta)$ is an explicit algebraic expression involving DK, Df_μ, $J \circ K$, and, $\mathcal{N}$.*

The proof of Lemma 12 is given in [CCdlL13]. The argument is as follows, taking derivative in equation (2.6) one has $Df_\mu \circ K_0(\theta) DK_0(\theta) = DK_0(\theta + \omega)$ which gives the first column in (2.8). The second column comes from the fact that the conformally symplectic property, $f_\mu^*\Omega = \lambda\Omega$, implies that the invariant torus given by equation (2.6) is Lagrangian. Then, using the conformally symplectic geometry the second column can be obtained.

Remark 13. *As it is pointed out in* [CCdlL13] *if K is an approximate solution of* (2.6)*, that is,*

$$f_\mu \circ K(\theta) - K(\theta + \omega) =: E(\theta) \tag{2.9}$$

the relation (2.8) *will hold with an error, R, that can be estimated in terms of the error, $E(\theta)$, of the invariance equation, that is*

$$Df_\mu \circ K(\theta) M(\theta) = M(\theta + \omega) \begin{pmatrix} \mathrm{Id} & S(\theta) \\ 0 & \lambda\,\mathrm{Id} \end{pmatrix} + R(\theta), \tag{2.10}$$

with

$$S(\theta) \equiv P(\theta+\omega)^\top Df\circ K(\theta) J^{-1}\circ K(\theta)P(\theta) - \mathcal{N}(\theta+\omega)^\top \Gamma(\theta+\omega)\mathcal{N}(\theta+\omega)\lambda \tag{2.11}$$
$$P(\theta) \equiv DK(\theta)\mathcal{N}(\theta),$$
$$\Gamma(\theta) \equiv DK(\theta)^\top J^{-1}\circ K(\theta)DK(\theta).$$

Moreover,

$$R(\theta) = \left[DE(\theta)\middle| V(\theta+\omega)(\tilde{B}(\theta)-\lambda\,\mathrm{Id}) + DK(\theta+\omega)(\tilde{S}(\theta)-S(\theta))\right] \tag{2.12}$$

where

$$V(\theta) \equiv J^{-1}\circ K(\theta)DK(\theta)\mathcal{N}(\theta) \tag{2.13}$$
$$\tilde{B}(\theta)-\lambda\,\mathrm{Id} \equiv DK(\theta)^\top J\circ K(\theta)DK(\theta)\tilde{S}(\theta) \tag{2.14}$$
$$\tilde{S}(\theta)-S(\theta) \equiv -\mathcal{N}(\theta+\omega)^\top\Gamma(\theta+\omega)\mathcal{N}(\theta+\omega)(\tilde{B}(\theta)-\lambda\,\mathrm{Id}) \tag{2.15}$$

We note that $\tilde{B}-\mathrm{Id}$ is estimated by the norm of (2.9), thus R in (2.12) can be estimated by the norm of (2.9) as it is shown in Lemma 37. The derivation of the formulas in (2.11), (2.12), and (2.13) can be found in [CCdlL13].

Remark 14. *Observe that when considering K_0, μ_0 satisfying (2.6) and a perturbation K_ε, μ_ε (which could be given in terms of formal power series), equation (2.10) is also satisfied by K_ε, μ_ε but with all the expressions depending on ε (small enough), that is,*

$$Df_{\mu_\varepsilon}\circ K_\varepsilon(\theta)M_\varepsilon(\theta) = M_\varepsilon(\theta+\omega)\begin{pmatrix}\mathrm{Id} & S_\varepsilon(\theta)\\ 0 & \lambda\,\mathrm{Id}\end{pmatrix} + R_\varepsilon(\theta).$$

3. Statement of the main result, Theorem 18

In this section we state the main result, Theorem 18, which gives the Gevrey character of the perturbative expansions of the solutions to equation (2.6). First we introduce a normalization which guarantees the uniqueness of the solutions to equation (2.6).

3.1. **Normalization and local uniqueness.** The centerpiece of this work is the invariance equation

$$f_{\varepsilon,\mu_\varepsilon}\circ K_\varepsilon = K_\varepsilon\circ T_\omega \tag{3.1}$$

where $T_\omega(\theta)=\theta+\omega$. Note that if (K,μ) is a solution of the invariant equation (3.1), then, for any $\sigma\in\mathbb{T}^d$, $(K\circ T_\sigma,\mu)$ is also a solution of (3.1), due to the fact that $K\circ T_\sigma$ parameterizes the same torus as K. So, in order to get uniqueness it is neccesary to impose a normalization condition.

Definition 15. *We say that a torus with embedding K is normalized with respect to K_0 when*

$$\int_{\mathbb{T}^d}\left[M_0^{-1}(\theta)(K(\theta)-K_0(\theta))\right]_d d\theta = 0 \tag{3.2}$$

where the subscript d indicates that we take the first d rows of the $2d\times d$ matrix, and M_0 is constructed from K_0 as in (2.7).

We also recall the following result ([CCdlL13], Proposition 26) which shows that this condition can be imposed without loss of generality for solutions that are close to one another.

Proposition 16. *Let K_0, K be solutions of (3.1) and $||K-K_0||_{C^1}$ be sufficiently small (with respect to quatities depending only on M -computed out of K_0 - and f). Then, there exists $\sigma\in\mathbb{R}^d$, such that $K^{(\sigma)}=K\circ T_\sigma$ satisfies (3.2). Furthermore,*

$$|\sigma|\le C||K-K_0||_{C^1}$$

where the constant C can be chosen to be as close to 1 as desired by assuming that f_μ, K_0, and K_1 are twice differentiable, $DK_0^\top DK_1$ is invertible and $||K - K_0||_{C^0}$ is sufficiently small. The σ thus chosen is locally unique.

Remark 17. *As it is noted in [CCdlL13] the normalization (3.2) works as well when K is only an approximate solution. Then, assuming that K_0 is a solution of equation (3.1), the normalization condition (3.2) for an approximate solution of (3.1) given as power series expansion $\sum_{n=0}^{\infty} K_n(\theta)\varepsilon^n$ is equivalent to the conditions*

$$\int_{\mathbb{T}^d} \left[M_0^{-1}(\theta)K_n(\theta)\right]_d d\theta = 0 \tag{3.3}$$

for all $n \geq 1$.

3.2. **Main Theorem.** Here we present our main theorem, Theorem 18.

Theorem 18 (Main Theorem). *Let $\omega \in \mathcal{D}(\nu, \tau)$. Consider the map $f : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ given by*

$$f_{\varepsilon,\mu_\varepsilon}(x, y) = (x + \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V'(x), \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V'(x)) \tag{3.4}$$

where $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$, $\alpha \in \mathbb{N}$, $V(x)$ is a trigonometric polynomial, $\mu_\varepsilon \in \mathbb{C}$, and $\varepsilon \in \mathbb{C}$. Then, there exists $\rho_0 > 0$ such that the following holds

(A) *There exist formal power series expansions $K_\varepsilon^{[\infty]} = \sum_{j=0}^{\infty} K_j\varepsilon^j$ and $\mu_\varepsilon^{[\infty]} = \sum_{j=0}^{\infty} \mu_j\varepsilon^j$ satisfying $f_{\varepsilon,\mu} \circ K = K(\theta + \omega)$ in the sense of formal power series. More precisely, defining $K_\varepsilon^{[\leq N]} = \sum_{j=0}^{N} K_j\varepsilon^j$ and $\mu_\varepsilon^{[\leq N]} = \sum_{j=0}^{N} \mu_j\varepsilon^j$ for any $N \in \mathbb{N}$ we have*

$$\left\| f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} - K_\varepsilon^{[\leq N]} \circ T_\omega \right\|_{\rho_0} \leq C_N|\varepsilon|^{N+1}. \tag{3.5}$$

where $C_N > 0$. Moreover, if the K_j's satisfy the normalization condition (3.3), then the expansions $K_\varepsilon^{[\infty]}$, $\mu_\varepsilon^{[\infty]}$ are unique.

(B) *The unique formal power series expansions, $K_\varepsilon^{[\infty]}$ and $\mu_\varepsilon^{[\infty]}$, satisfying (3.5) and the normalization (3.3) are such that $K^{[\infty]} \in \mathcal{G}_{\rho_0}^{2\tau/\alpha}$ and $\mu^{[\infty]} \in \mathcal{G}^{2\tau/\alpha}$, i.e., there exists constants L, F, N_0 such that*

$$\|K_n\|_{\rho_0} \leq LF^n n^{(2\tau/\alpha)n} \quad \text{and} \quad |\mu_n| \leq LF^n n^{(2\tau/\alpha)n} \qquad \text{for any } n > N_0. \tag{3.6}$$

The proof of Theorem 18 is an easy consequence of Lemma 22. Proposition 55, given in the Appendix, shows the hypothesis of Lemma 22 are satisfied for maps of the form (3.4). Lemma 22 states the same results as Theorem 18 but in a more general setting.

Remark 19. *It is instructive to compare the results in Theorem 18 with the numerical explorations of [BC19] (see also [BC]). In the case that $\lambda(\varepsilon) = 1 - \varepsilon^3$ and ω is the golden mean, Theorem 18 gives that the expansion satisfies the Gevrey bounds with exponent 2/3. Of course, Theorem 18 gives only an upper bound and lower exponents could also be true. The numerical results in [BC19] and [BC] lead to the conjecture that the expansion $\sum K_n\varepsilon^n$ has some well defined asymptotics*

$$\|K_n\|_\rho^{1/n} \approx Cn^\sigma \tag{3.7}$$

with a slightly smaller Gevery exponent, $\sigma \approx 0.3$. The asymptotics (3.7) is compatible with the results in Theorem 18, but suggests that the results in Theorem 18 are not optimal. We call attention that [BC19] contained an unfortunate typo and the results attributed there to $\|K_n\|^{1/n}$ are actually results for $\|n!K_n\|^{1/n}$, this is corrected in [BC]. The paper [BC] also presents several other patterns in the series (refined versions of (3.7) including oscillations of period 3, studies for other Diophantine numbers, etc.) We hope that the method presented in this paper can lead to studies of these phenomena, hitherto discovered only through numerical implementation.

We think that the argument in Theorem 18 can optimized to lower the Gevrey exponent and get closer to the numerical values, but, since the method of proof is rather novel, we decided to follow the advice "Premature optimization is the root of all evil" [Knu98]*, and present the argument in its simplest form so that it could, perhaps, be applied to other problems.*

For the sake of completeness, before stating the main Lemma we will state a Theorem in [CCdlL17] which assures the existence of formal power series expansions satisfying (3.1) up to any order for conformally symplectic systems.

Theorem 20 ([CCdlL17], Theorem 12)**.** *Let $\mathcal{M} \equiv \mathbb{T}^d \times \mathcal{B}$ with $B \subseteq \mathbb{R}^d$ an open, simply connected domain with smooth boundary; $\mathcal{M}$ is endowed with an analytic symplectic form Ω.*

Let $\omega \in \mathcal{D}(\tau, \nu)$ and consider a family $f_{\varepsilon,\mu}$ of conformally symplectic mappings that satisfy

$$f^*_{\varepsilon,\mu}\Omega = \lambda(\varepsilon)\Omega, \tag{3.8}$$

with $\mu \in \Lambda, \Lambda \subseteq \mathbb{C}^d$, $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$, $\alpha \in \mathbb{N}$ and $\varepsilon \in \mathbb{C}$.

Assume that for $\varepsilon = 0$ the family of maps $f_{0,\mu}$ is symplectic and that for some value μ_0 the map f_{0,μ_0} admits a Lagrangian invariant torus, namely we can find an analytic embedding $K_0 \in \mathcal{A}_\rho(\mathbb{T}^d, \mathcal{M})$, for some $\rho > 0$, such that

$$f_{0,\mu_0} \circ K_0 = K_0 \circ T_\omega. \tag{3.9}$$

Furthermore, assume that the torus K_0 satisfies the following hypothesis:

HND *Let the following non-degeneracy condition be satisfied:*

$$\det \begin{pmatrix} \overline{S_0} & \overline{S_0 (B_{0b})^0} + \overline{\tilde{A}_{01}} \\ 0 & \overline{\tilde{A}_{02}} \end{pmatrix} \neq 0$$

where the $d \times d$ matrix S_0 is defined as

$$\begin{aligned} S_0(\theta) \equiv\; & \mathcal{N}_0(\theta+\omega)^T DK_0(\theta+\omega) Df_{\mu_0,0} \circ K_0(\theta) J^{-1} \circ K_0(\theta) DK_0(\theta) \mathcal{N}_0(\theta) \\ & - \mathcal{N}_0(\theta+\omega)^T DK_0(\theta+\omega)^T J^{-1} \circ K_0(\theta+\omega) DK_0(\theta+\omega) \mathcal{N}_0(\theta+\omega) \end{aligned}$$

with $\mathcal{N} = (DK_0^T DK_0)^{-1}$, the $d \times d$ matrices $\tilde{A}_{01}, \tilde{A}_{02}$ denote the first d and the last d rows of the $2d \times d$ matrix $\tilde{A}_0 = (M_0 \circ T_\omega)^{-1} (D_\mu f_{0,\mu_0} \circ K_0)$, where M_0 is as in (2.7), $(B_{0b})^0$ is the solution (with zero average) of the cohomology equation $(B_{0b})^0 - B_{0b} \circ T_\omega = -(\tilde{A}_{02})^0$, where $(B_{0b})^0 \equiv B_{0b} - \overline{B_{0b}}$ and the overline denotes the average.

Then, we have the following

(A) *There exist a formal power series expansions $K_\varepsilon^{[\infty]} = \sum_{j=0}^\infty K_j \varepsilon^j$ and $\mu_\varepsilon^{[\infty]} = \sum_{j=0}^\infty \mu_j \varepsilon^j$ satisfying (3.9) in the sense of formal power series. More precisely, defining $K_\varepsilon^{[\leq N]} = \sum_{j=0}^N K_j \varepsilon^j$ and $\mu_\varepsilon^{[\leq N]} = \sum_{j=0}^N \mu_j \varepsilon^j$ for any $N \in \mathbb{N}$ and $\rho > 0$, we have*

$$\left\| f_{\varepsilon, \mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} - K_\varepsilon^{[\leq N]} \circ T_\omega \right\|_{\rho'} \leq C_N |\varepsilon|^{N+1}. \tag{3.10}$$

for some $0 < \rho' < \rho$ and $C_N > 0$.

Moreover, if we require the K_j's satisfy the normalization condition (3.3), then the expansions $K_\varepsilon^{[\infty]}$, $\mu_\varepsilon^{[\infty]}$ are unique.

Note that Theorem 20 does not assume that the case $\varepsilon = 0$ is an integrable system, as it is the case for the map (3.4), it suffices that the case $\varepsilon = 0$ is a Hamiltonian system with a KAM torus.

Remark 21. *Denoting*

$$E_\varepsilon^N(\theta) \equiv f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta) - K_\varepsilon^{[\leq N]}(\theta+\omega) \tag{3.11}$$

then (3.10) *can be written as*

$$\left\| E_\varepsilon^N \right\|_{\rho'} \leq C_N |\varepsilon|^{N+1}. \tag{3.12}$$

According to the notation introduced earlier, this means that $E_\varepsilon^N \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ *or* $E_\varepsilon^N = \sum_{j=N+1}^{\infty} E_j \varepsilon^j$ *for* ε *small enough. We denote*

$$E_\varepsilon^{(N,2N]} = \sum_{j=N+1}^{2N} E_j \varepsilon^j$$

the trucated series.

The following lemma, Lemma 22, can be considered as an improvement of Theorem 20 in the sense that it gives Gevrey bounds for the coefficients K_j, μ_j of the unique (under normalization) formal power series expansions $K_\varepsilon^{[\infty]}, \mu_\varepsilon^{[\infty]}$.

Lemma 22 (Main Lemma)**.** *Assume the hypothesis of Theorem 20. Assume also that for any* ε*, small enough, and for any* $N \in \mathbb{N}$ *we have:*

HTP1 $\tilde{E}_{\varepsilon,2}^{(N,2N]}$, $\tilde{A_{\varepsilon,2}^N}$ *are trigonometric polynomials in* θ *of degree at most* aN*,* $a \in \mathbb{N}$*. Where* $\tilde{E}_{\varepsilon,2}^{(N,2N]}$, $\tilde{A_{\varepsilon,2}^N}$ *denote the* $d \times 1$ *and* $d \times d$ *matrices, respectively, given by taking the last* d *rows of the* $2d \times 1$ *matrix* $\tilde{E}_\varepsilon^{(N,2N]} = \left(M_\varepsilon^{[\leq N]} \circ T_\omega\right)^{-1} E_\varepsilon^{(N,2N]}$ *and the* $2d \times d$ *matrix* $\tilde{A}_\varepsilon^N = \left(M_\varepsilon^{[\leq N]} \circ T_\omega\right)^{-1} D_\mu f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}$*, respectively.* $M_\varepsilon^{[\leq N]}$ *is as in* (2.7) *constructed from* $K_\varepsilon^{[\leq N]}$*.*

HTP2 *The* $d \times d$ *matrix*

$$\begin{aligned}\tilde{E}_{\Omega,\varepsilon}^N(\theta) \equiv{}& DK_\varepsilon^{[\leq N]}(\theta+\omega)^\top J \circ K_\varepsilon^{[\leq N]}(\theta+\omega) DK_\varepsilon^{[\leq N]}(\theta+\omega)\\ &- D(f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta))^\top J \circ (f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) D(f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta))\end{aligned} \tag{3.13}$$

is a trigonometric polynomial of degree at most aN*.*

Then, there exist $\rho_0 \leq \rho'$ *such that the unique formal power series expansions,* $K_\varepsilon^{[\infty]}$ *and* $\mu_\varepsilon^{[\infty]}$*, satisfying* (3.10) *and* (3.3) *are such that* $K^{[\infty]} \in \mathcal{G}_{\rho_0}^{2\tau/\alpha}$ *and* $\mu^{[\infty]} \in \mathcal{G}^{2\tau/\alpha}$*, i.e., there exists constants* L*,* F*,* N_0 *such that*

$$\|K_n\|_{\rho_0} \leq LF^n n^{(2\tau/\alpha)n} \quad \text{and} \quad |\mu_n| \leq LF^n n^{(2\tau/\alpha)n} \qquad \text{for any } n > N_0. \tag{3.14}$$

The proof of Lemma 22, given in Section 6.2, is done by means of a Newton like method which acts on finite powers series expansions ($K_\varepsilon^{[\leq N]}$, $\mu_\varepsilon^{[\leq N]}$), this method is described in the next section. We emphasize that this quasi Newton method takes advantage of the conformally symplectic property (see Definitions 1) that maps like (3.4) satisfy.

We also point out that hypothesis **HTP1** and **HTP2** are very natural for the maps considered in Theorem 18. The verification of these hypothesis for the dissipative standard map is described in detail in Proposition 55 of the Appendix. In the general setting in which Lemma 22 is stated, the hypothesis **HTP1** and **HTP2** are needed to be able to get estimates, in balls with center at the origin, for the solutions of the linear equations of the quasi Newton method.

3.3. **Asymptotic estimates for invariance functions.** The formal power series studied in this paper are asymptotic expansions of functions $K_\varepsilon, \mu_\varepsilon$ constructed in [CCdlL17]. The functions $K_\varepsilon, \mu_\varepsilon$ are determined by the condition that they satisfy the invariance equation (3.1) and the normalization (3.3). In this section we argue that the same method we use to prove the Gevrey estimates also shows that the formal series defined here are asymptotic to the functions $K_\varepsilon, \mu_\varepsilon$ with very strong estimates in the remainder, see Theorem 23.

We emphasize that the functions $K_\varepsilon, \mu_\varepsilon$ are not constructed out of the asymptotic expansions by complex analysis methods (Borel summation, resummation of series). They are obtained from the requirement that they satisfy the invariance equation (3.1) and the normalization (3.3). It is an interesting open question whether some resummation of the asymptotic expansions studied here can produce the functions $K_\varepsilon, \mu_\varepsilon$.

The domain of definition of the functions $K_\varepsilon, \mu_\varepsilon$ is rather subtle. In [CCdlL17], it is proved that the domain of definition of $K_\varepsilon, \mu_\varepsilon$ contains a set $\mathcal{G}$ obtained by removing sequence of balls that are dense on curves converging to the origin, in fact, it is rigorously showed that $\mathcal{G}$ is a lower bound on the analyticity domain of the functions $K_\varepsilon, \mu_\varepsilon$. We also point out that the set $\mathcal{G}$ does not contain any ball centered at the origin. Indeed, the set $\mathcal{G}$ does not contain any sector centered at the origin of width bigger than π/α, thus the width of the domain is not enough to apply many methods of complex analysis related to Phragmén-Lindelöf theory. In the other direction, the paper [CCdlL17] contains arguments showing that for generic perturbations one should not expect that the domain of analyticity contains the excluded balls (if the perturbation happens to be identically zero one indeed obtains a larger domain). The paper [BC19] studies numerically the maximal domain of definition of the functions $K_\varepsilon, \mu_\varepsilon$ for the map (3.4) using a variety of methods including Pade summation and continuation methods. Indeed [BC19] conjectured that the series were Gevrey and this was an important motivation for this paper.

The set $\mathcal{G}$ is determined by asking that $\lambda(\varepsilon)$ satisfies a Diophantine condition with respect to ω, more precisely, defining

$$\tilde{\nu} = \tilde{\nu}(\lambda; \omega, \tau) \equiv \sup_{k \in \mathbb{Z}^d \setminus \{0\}} |e^{2\pi i k \cdot \omega} - \lambda|^{-1} |k|^{-\tau} \tag{3.15}$$

one has

$$\mathcal{G} = \mathcal{G}(A; \omega, \tau, N) = \left\{ \varepsilon \in \mathbb{C} \ : \quad \tilde{\nu}(\lambda; \omega, \tau) |\lambda(\varepsilon) - 1|^{N+1} \le A \right\}. \tag{3.16}$$

The basic idea to prove the existence of the functions $K_\varepsilon, \mu_\varepsilon$ is as follows: The formal power expansions produces a sequence of polynomials which satisfy the invariance equation (3.1) rather approximately in a ball. In the intersection of the ball with the set $\mathcal{G}$, we can apply the a-posteriori theorem, Theorem 14 in [CCdlL17], and obtain a true solution of (3.1). Of course, the detailed implementation requires taking into account several other issues such as the absence of monodromy.

In this paper, we will use a very similar technique. As as byproduct of the estimates used in the proof of Lemma 22, we obtain that some truncations of the formal expansion satisfy the invariance equation up to a very small error in appropriate balls. Then, in the intersection of the balls with the set $\mathcal{G}$ we will be able to apply Theorem 20 in [CCdlL13].

More precisely we have:

Theorem 23. *Assuming the hypothesis of Lemma 22 and $n \in (2^h N_0, 2^{h+1} N_0] \cap \mathbb{N}$, then for any $0 < \delta < \rho_0$ the asymptotic expansions in Lemma 22 satisfy*

$$\sup_{\varepsilon \in \mathcal{G}, |\varepsilon| \le \tilde{\gamma}_{h+2}} \left\| \sum_{j=1}^{n} K_j \varepsilon^j - K_\varepsilon \right\|_{\rho_0 - \delta} \le \left(U + V 2^{h(3\tau + 3d)} r^{n+1} r^{2^h N_0} \right) (CD)^h B^{h^2} r^{(2^h - 1) N_0} \left\| E^{N_0} \right\|_{\rho_0} \tag{3.17}$$

where $\hat{C}$ and C are uniform constants and $U = \hat{C}\nu^{-1}\tilde{\nu}^{-1}\delta^{-2(\tau+d)}$, $V = C\nu^{-3}(aN_0)^{2\tau}\rho_0^{-(\tau+3d)}2^{2\tau+6d}$, $D = \nu^{-6}(aN_0)^{4\tau}\rho_0^{-(2\tau+6d)}2^{-(4\tau+12d)}$, $r = 2^{-\tau/\alpha}$, $B = 2^{6\tau+6d}$, *and* $\tilde{\gamma}_h = (2^{-1}\nu)^{1/\alpha}(a2^hN_0)^{-\tau/\alpha}$.

Note that (3.17) can be understood as having super-exponentially small errors in domains decreasing exponentially fast. It is also important to note that almost all constants in (3.17) are given explicitly. The proof of Theorem 23 is given in Section 6.3.

4. Iterative step of the quasi Newton method.

The KAM procedure for the proof of Theorem 22 is based on the application of a quasi Newton method, which is described in Section 4.2. Before describing this procedure we introduce two types of cohomology equations that allow us to solve the linear equations, and obtain estimates, of the modified Newton method. The estimates for each step of the method will be given in Section 5.

4.1. **Estimates for some cohomology equations.**

The iterative step described in Section 4.2 depends on the solution of two cohomology equations. The first equation, (4.1), is very standard in KAM theory. The estimate given in Lemma 24 is well known for the experts in KAM theory, we have decided to include a proof here for the sake of completeness. The second type of cohomology equation we consider, (4.3), it is more complicated to study due to the fact of the appearance of the factor $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$. This factor introduces some restrictions in the set of parameters, ε, for which we are able to obtain estimates.

4.1.1. *Standard cohomology equation.* The first cohology equation we deal with is the following

$$\varphi_\varepsilon(\theta) - \varphi_\varepsilon(\theta + \omega) = \eta_\varepsilon(\theta) \tag{4.1}$$

Lemma 24 below, gives sufficient conditions to solve equation (4.1) and to obtain estimates of its solutions. This estimates are very standard in KAM theory.

Lemma 24. *Let $\omega \in \mathcal{D}(\nu, \tau)$. Assume that $\eta \in \mathcal{A}_{\rho,r}$ is such that $\int_{\mathbb{T}^d} \eta_\varepsilon(\theta)d\theta = 0$. Then, we can find a unique solution of (4.1), φ_ε, that satisfies $\int_{\mathbb{T}^d} \varphi_\varepsilon(\theta)d\theta = 0$. Moreover, if for any $0 < \delta \leq \rho$ we have $\varphi \in \mathcal{A}_{\rho-\delta,r}$, then*

$$\|\varphi\|_{\rho-\delta,r} \leq C\nu^{-1}\delta^{-(\tau+d)} \|\eta\|_{\rho,r}\,.$$

With $C = C(d)$. Furthermore, $\eta_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$ implies $\varphi_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$.

Proof. Expanding in Fourier series the solution to (4.1) is given by $\varphi_\varepsilon(\theta) = \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \frac{\eta_k(\varepsilon)}{1-e^{2\pi i k\cdot\omega}} e^{2\pi i k\cdot\theta}$. Then, using Cauchy estimates one obtains

$$\begin{aligned}
\|\varphi_\varepsilon\|_{\rho-\delta} &\leq \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \frac{|\hat{\eta_k}(\varepsilon)|}{|1-e^{2\pi i k\cdot\omega}|} \left\|e^{2\pi i k\cdot\theta}\right\|_{\rho-\delta} \\
&\leq \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \nu^{-1}|k|^\tau \|\eta_\varepsilon\|_\rho e^{-2\pi|k|\rho} e^{2\pi(\rho-\delta)|k|} \\
&\leq C\nu^{-1} \|\eta_\varepsilon\|_\rho \sum_{j\in\mathbb{N}} j^{\tau+d+1} e^{2\pi\delta j} \\
&\leq C\nu^{-1}\delta^{-(\tau+d)} \|\eta_\varepsilon\|_\rho\,.
\end{aligned} \tag{4.2}$$

The last line gives $\varphi_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$ if $\eta_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$ and taking supremum over ε the result is proved. □

Remark 25. *Equation (4.1) appears very often in KAM theory. When $\varepsilon \in \mathbb{R}$, the paper [Rüs75] contains estimates with a better exponent on δ. That is, in the same situation of Lemma 24, when $\varepsilon \in \mathbb{R}$, one can get $\|\varphi_\varepsilon\|_{\rho-\delta} \leq C\nu\delta^{-\tau}\|\eta_\varepsilon\|_\rho$.*

4.1.2. *Parametric cohomology equation.* The second cohomology equation we are interested in is an equation for $\varphi_\varepsilon : \mathbb{T}^d \to \mathbb{C}$, of the form

$$\lambda(\varepsilon)\varphi_\varepsilon(\theta) - \varphi_\varepsilon(\theta+\omega) = \eta_\varepsilon(\theta) \tag{4.3}$$

where $\eta_\varepsilon : \mathbb{T}^d \to \mathbb{C}$ and $\omega \in \mathbb{R}^d$ are given, ε fixed.

Note that, as it is seen in Lemma 27, solve equation (4.3) presents a small divisors problem. In this case the small divisors depend on the variable ε, that is, equation (4.3) is not expected to have a solution when $\lambda(\varepsilon) = e^{2\pi i k\cdot\omega}$. One approach that has been used to deal with the small divisors in equation (4.3) (see [CCdlL13]) requires to *remove* a set from the complex plane, $\varepsilon \in \mathbb{C}$, where the denominators $\lambda(\varepsilon) - e^{2\pi i k\cdot\omega}$ are small. This gives rise to a set with a complicated structure, $\mathcal{G} \subset \mathbb{C}$, of parameters, ε, in which is possible to find a solution, and estimates, of equation (4.3). One of the properties of the set $\mathcal{G}$ described in [CCdlL13], is that it does not contain any ball with center at the origin. This property is one of the reasons for which we follow a different approach to deal with equation (4.3), to prove the Gevrey estimates in Lemma 22 we rely heavily on being able to obtain estimates of (4.3) for ε in a ball centered at the origin.

The following two Lemmas allow us to obtain estimates in balls centered at $\varepsilon = 0$ for the solution, φ_ε, of equation (4.3) whenever η_ε is a trigonometric polynomial. If the degree of the trig polynomial, η_ε, is aN, Lemma 26 gives a relation between this degree and a domain in which the solution, φ_ε, of (4.3) will be analytic in ε.

Note that the requirement of hypothesis **HTP1** and **HTP2** in Lemma 22 is due to the fact that the quantities given in these hypothesis will be the right hand side of equations of the form (4.3).

Lemma 26. *Let $\omega \in \mathcal{D}(\nu,\tau)$, $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$, $\alpha \geq 1$, and $a, N \in \mathbb{N}$. If $|\varepsilon| \leq \left(\frac{\nu}{2}\right)^{1/\alpha} \frac{1}{(aN)^{\tau/\alpha}}$, then, for $|k| \leq aN$ we have*

$$\left|\lambda(\varepsilon) - e^{2\pi i k\cdot\omega}\right| \geq \frac{\nu}{2}\frac{1}{(aN)^\tau}$$

Proof.

$$|e^{2\pi i k\cdot\omega} - \lambda(\varepsilon)| \geq |e^{2\pi i k\cdot\omega} - 1| - |1-\lambda(\varepsilon)| \geq \frac{\nu}{|k|^\tau} - |\varepsilon|^\alpha \geq \frac{\nu}{(aN)^\tau} - \frac{\nu}{2(aN)^\tau} = \frac{\nu}{2}\frac{1}{(aN)^\tau}$$

□

Lemma 27. *Let $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$, $\alpha \geq 1$, $\omega \in \mathcal{D}(\nu,\tau)$; $a, N \in \mathbb{N}$, and define*

$$\gamma_N = \left(\frac{\nu}{2}\right)^{1/\alpha}\frac{1}{(aN)^{\tau/\alpha}}.$$

Let $\eta \in \mathcal{A}_{\rho,\gamma_N}$ such that $\int_{\mathbb{T}^d}\eta_\varepsilon(\theta)d\theta = 0$ and assume that, for any ε, $\eta_\varepsilon(\theta)$ is a trigonometric polynomial of degree aN in θ. Then, for any $|\varepsilon| \leq \gamma_N$ equation (4.3) has a unique solution, $\varphi_\varepsilon(\theta)$, such that $\int_{\mathbb{T}^d}\varphi_\varepsilon(\theta)d\theta = 0$. Furthermore, if for any $0 < \delta \leq \rho$ we have $\varphi \in \mathcal{A}_{\rho-\delta,\gamma_N}$, then,

$$\|\varphi\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^\tau\delta^{-d}\|\eta\|_{\rho,\gamma_N}.$$

Moreover, if $\eta_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$, then $\varphi_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$.

Proof. Expanding η_ε in Fourier series as $\eta_\varepsilon(\theta) = \sum_{0<|k|\leq aN}\hat{\eta}_k(\varepsilon)e^{2\pi i k\cdot\theta}$ a solution to (4.3) is given by

$$\varphi_\varepsilon(\theta) = \sum_{0<|k|\leq aN}\frac{\hat{\eta}_k(\varepsilon)}{\lambda(\varepsilon) - e^{2\pi i k\cdot\omega}}e^{2\pi i k\cdot\theta}.$$

Using Lemma 26 and Cauchy estimates, one obtains that for any $|\varepsilon| \leq \gamma_N$

$$\begin{aligned}
\|\varphi_\varepsilon\|_{\rho-\delta} &\leq \sum_{0<|k|\leq aN} \frac{|\hat{\eta}_k(\varepsilon)|}{|\lambda(\varepsilon) - e^{2\pi i k\cdot\omega}|} \left\| e^{2\pi i k\cdot\theta} \right\|_{\rho-\delta} \\
&\leq 2(aN)^\tau \nu^{-1} \sum_{0<|k|\leq aN} |\hat{\eta}_k(\varepsilon)| e^{2\pi|k|(\rho-\delta)} \\
&\leq 2(aN)^\tau \nu^{-1} \sum_{0<|k|\leq aN} \|\eta_\varepsilon\|_\rho e^{-2\pi|k|\rho} e^{2\pi|k|(\rho-\delta)} \\
&\leq 2(aN)^\tau \nu^{-1} \|\eta_\varepsilon\|_\rho \sum_{j=1}^{aN} j^{d-1} e^{-2\pi j\delta} \\
&\leq C\nu^{-1}(aN)^\tau \delta^{-d} \|\eta_\varepsilon\|_\rho
\end{aligned} \tag{4.4}$$

Thus, $\|\varphi\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^\tau\delta^{-d}\|\eta\|_{\rho,\gamma_N}$. The last claim comes from (4.4), that is $\varphi_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$ if $\eta_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$. □

4.2. **Formulation of the quasi Newton method.** Every step of the quasi Newton method starts with a solution of equation (3.1) up to order ε^N. That is, assume that

$$K_\varepsilon^{[\leq N]}(\theta) = \sum_{n=0}^N K_n(\theta)\varepsilon^n, \quad \mu_\varepsilon^{[\leq N]} = \sum_{n=0}^N \mu_n \varepsilon^n$$

satisfy the normalization (3.3) and

$$f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta) - K_\varepsilon^{[\leq N]}(\theta+\omega) =: E_\varepsilon^N(\theta)$$

with

$$\left\|E_\varepsilon^N\right\|_\rho \leq C|\varepsilon|^{N+1}.$$

Remark 28. *The first step of the Newton method could start with $K^{[\leq N_0]}$, $\mu^{[\leq N_0]}$, given by Theorem 20, for some N_0.*

Newton's method consists in finding corrections $\Delta_\varepsilon, \mu_\varepsilon$ to $K_\varepsilon^{[\leq N]}$ and $\mu_\varepsilon^{[\leq N]}$ such that the linear approximation of equation (3.1) associated to $K_\varepsilon^{[\leq N]} + \Delta_\varepsilon, \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon$ reduces the error up to quadratic terms. Taking into account that

$$f_{\varepsilon,\mu+\sigma} \circ (K+\Delta) = f_{\varepsilon,\mu} \circ K + [Df_{\varepsilon,\mu} \circ K]\,\Delta + [D_\mu f_{\varepsilon,\mu} \circ K]\,\sigma + O(\|\Delta\|^2) + O(\|\sigma\|^2)$$

the Newton equation is

$$\left[Df_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right]\Delta_\varepsilon - \Delta_\varepsilon \circ T_\omega + \left[D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right]\sigma_\varepsilon = -E_\varepsilon^N. \tag{4.5}$$

Equation (4.5) is not easy to solve due to the fact that $Df_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}$ is not constant. Following an approach similar to that in [CCdlL13], we will not solve (4.5) exactly but we will find approximate solutions that will reduce quadratically the error. The idea is to approximate the solution of (4.5) using the geometric identities introduced in Section 2.6. Considering the change of variables

$$\Delta_\varepsilon = M_\varepsilon^{[\leq N]} W_\varepsilon, \tag{4.6}$$

where $M_\varepsilon^{[\leq N]}$ is as in (2.7) computed from $K_\varepsilon^{[\leq N]}$. Using (2.10) one obtains that (4.5) is equivalent to

$$M_\varepsilon^{[\leq N]} \circ T_\omega \left[\begin{pmatrix} \mathrm{Id} & S_\varepsilon^{[\leq N]} \\ 0 & \lambda(\varepsilon)\,\mathrm{Id} \end{pmatrix} W_\varepsilon - W_\varepsilon \circ T_\omega\right] + \left(D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right)\sigma_\varepsilon = -E_\varepsilon^N - R_\varepsilon^{[\leq N]} W_\varepsilon \tag{4.7}$$

where $R_\varepsilon^{[\le N]}$ is the error (2.12) and $S_\varepsilon^{[\le N]}$ is given in (2.11), both computed from $K_\varepsilon^{[\le N]}$. That is

$$M_\varepsilon^{[\le N]} \equiv \left[DK_\varepsilon^{[\le N]} \mid J^{-1}\circ K_\varepsilon^{[\le N]} DK_\varepsilon^{[\le N]}\mathcal{N}_\varepsilon^{[\le N]}\right] \sim \mathcal{O}(|\varepsilon|^0) \tag{4.8}$$

$$S_\varepsilon^{[\le N]} \equiv {P_\varepsilon^{[\le N]}}^\top Df_{\mu_\varepsilon^{[\le N]},\varepsilon}\circ K_\varepsilon^{[\le N]} J^{-1}\circ K_\varepsilon^{[\le N]} P_\varepsilon^{[\le N]} - \lambda(\varepsilon){\mathcal{N}_\varepsilon^{[\le N]}}^\top \Gamma_\varepsilon^{[\le N]}\mathcal{N}_\varepsilon^{[\le N]} \sim \mathcal{O}(|\varepsilon|^0) \tag{4.9}$$

$$\begin{aligned}
\mathcal{N}_\varepsilon^{[\le N]} &\equiv \left[\left(DK_\varepsilon^{[\le N]}\right)^\top DK_\varepsilon^{[\le N]}\right]^{-1} \sim \mathcal{O}(|\varepsilon|^0), && (4.10)\\
P_\varepsilon^{[\le N]} &\equiv DK_\varepsilon^{[\le N]}\mathcal{N}_\varepsilon^{[\le N]}, &&\\
\Gamma_\varepsilon^{[\le N]} &\equiv {DK_\varepsilon^{[\le N]}}^T J^{-1}\circ K_\varepsilon^{[\le N]} DK_\varepsilon^{[\le N]} && (4.11)
\end{aligned}$$

Since we expect both W_ε and $R_\varepsilon^{[\le N]}$ to be estimated by E_ε^N, see (5.5) and (5.15), the term $W_\varepsilon R_\varepsilon^{[\le N]}$ is quadratic in E_ε^N, thus, we expect that omitting this term in (4.7) will not change the quadratic nature of the method.

In order to be able to get estimates of solutions of cohomology equations of the form (4.3) instead of considering the whole error $E_\varepsilon^N = \sum_{j=N+1}^\infty E_j\varepsilon^j$ we only consider a truncation of this series, that is, we only consider $E_\varepsilon^{(N,2N]} = \sum_{j=N+1}^{2N} E_j\varepsilon^j$.
Taking the above into account our quasi Newton step consist in solving the following equation

$$M_\varepsilon^{[\le N]}\circ T_\omega\left[\begin{pmatrix}\mathrm{Id} & S_\varepsilon^{[\le N]}\\ 0 & \lambda(\varepsilon)\,\mathrm{Id}\end{pmatrix} W_\varepsilon - W_\varepsilon\circ T_\omega\right] + \left(D_\mu f_{\varepsilon,\mu_\varepsilon^{[\le N]}}\circ K_\varepsilon^{[\le N]}\right)\sigma_\varepsilon = -E_\varepsilon^{(N,2N]} \tag{4.12}$$

Remark 29. *The election of the truncation $E_\varepsilon^{(N,2N]}$ in (4.12) has two very important implications for the proof of our result. The first one is that this will yield a new approximate solution which reduces the error quadratically, as a function of ε. Moreover, our model example, the dissipative standard map (1.1), will satisfy hypothesis* **HTP1** *and* **HTP2** *in Lemma 22 due to the fact that the truncation is made. See appendix A.*

In order to construct a solution of equation (4.12), we follow a similar approach as in [CCdlL13]. Defining

$$\tilde{E}_\varepsilon^{\,(N,2N]} := \left(M_\varepsilon^{[\le N]}\circ T_\omega\right)^{-1} E_\varepsilon^{(N,2N]} \sim \mathcal{O}(|\varepsilon|^{N+1}) \tag{4.13}$$

$$\tilde{A}_\varepsilon^N := \left(M_\varepsilon^{[\le N]}\circ T_\omega\right)^{-1} D_\mu f_{\varepsilon,\mu_\varepsilon^{[\le N]}}\circ K_\varepsilon^{[\le N]} \sim \mathcal{O}(|\varepsilon|^0) \tag{4.14}$$

and writing $\tilde{E}_\varepsilon^{(N,2N]} \equiv (\tilde{E}_{\varepsilon,1}^{(N,2N]}, \tilde{E}_{\varepsilon,2}^{(N,2N]})^\top$, where $\tilde{E}_{\varepsilon,1}^{(N,2N]}$ and $\tilde{E}_{\varepsilon,2}^{(N,2N]}$ are the first and last d rows of the $2d\times 1$ matrix $\tilde{E}_\varepsilon^{\,(N,2N]}$. Similarly, write $\tilde{A}_\varepsilon^N = (\tilde{A}_{\varepsilon,1}^N, \tilde{A}_{\varepsilon,2}^N)^\top$ and $W_\varepsilon = (W_{\varepsilon,1}, W_{\varepsilon,2})^\top$. Then (4.12) can be written in components as

$$W_{\varepsilon,1} - W_{\varepsilon,1}\circ T_\omega = -S_\varepsilon^{[\le N]} W_{\varepsilon,2} - \tilde{E}_{\varepsilon,1}^{(N,2N]} - \tilde{A}_{\varepsilon,1}^N\sigma_\varepsilon \tag{4.15}$$

$$\lambda(\varepsilon)W_{\varepsilon,2} - W_{\varepsilon,2}\circ T_\omega = -\tilde{E}_{\varepsilon,2}^{(N,2N]} - \tilde{A}_{\varepsilon,2}^N\sigma_\varepsilon \tag{4.16}$$

Denoting $\overline{W_{\varepsilon,i}}$ as the average of $W_{\varepsilon,i}$, with respect to θ, and $(W_{\varepsilon,i})^0 = W_{\varepsilon,i} - \overline{W_{\varepsilon,i}}$, $i = 1,2$; we can divide the system above into two systems, one for the average and another one for the no-average part, that is

$$\begin{aligned}
0 &= -\overline{S_\varepsilon^{[\le N]}}\,\overline{W_{\varepsilon,2}} - \overline{S_\varepsilon^{[\le N]}(W_{\varepsilon,2})^0} - \overline{\tilde{E}_{\varepsilon,1}^{(N,2N]}} - \overline{\tilde{A}_{\varepsilon,1}^N}\sigma_\varepsilon\\
\varepsilon^3\overline{W_{\varepsilon,2}} &= -\overline{\tilde{E}_{\varepsilon,2}^{(N,2N]}} - \overline{\tilde{A}_{\varepsilon,2}^N}\sigma_\varepsilon
\end{aligned} \tag{4.17}$$

$$
\begin{aligned}
(W_{\varepsilon,1})^0 - (W_{\varepsilon,1})^0 \circ T_\omega &= -(S_\varepsilon^{[\leq N]} W_{\varepsilon,2})^0 - (\tilde{E}_{\varepsilon,1}^{(N,2N]})^0 - (\tilde{A}_{\varepsilon,1}^N)^0 \sigma_\varepsilon \\
\lambda(\varepsilon)(W_{\varepsilon,2})^0 - (W_{\varepsilon,2})^0 \circ T_\omega &= -(\tilde{E}_{\varepsilon,2}^{(N,2N]})^0 - (\tilde{A}_{\varepsilon,2}^N)^0 \sigma_\varepsilon.
\end{aligned}
\tag{4.18}
$$

In order to uncouple systems (4.17) and (4.18) we consider $(W_{\varepsilon,2})^0$ as an affine function of σ_ε, due to (4.18). That is,

$$
(W_{\varepsilon,2})^0 = (B_{a,\varepsilon})^0 + (B_{b,\varepsilon})^0 \sigma_\varepsilon \tag{4.19}
$$

where $(B_{a,\varepsilon})^0$ and $(B_{b,\varepsilon})^0$ are defined as the solutions of

$$
\lambda(\varepsilon)(B_{a,\varepsilon})^0 - (B_{a,\varepsilon})^0 \circ T_\omega = -(\tilde{E}_{\varepsilon,2}^{(N,2N]})^0 \tag{4.20}
$$

$$
\lambda(\varepsilon)(B_{b,\varepsilon})^0 - (B_{b,\varepsilon})^0 \circ T_\omega = -(\tilde{A}_{\varepsilon,2}^N)^0. \tag{4.21}
$$

Due to **HTP1**, and applying Lemma 27, equations (4.20) and (4.21) can be solved and we can get estimates in balls with center at $\varepsilon = 0$. Once that (4.20) and (4.21) are solved, and using (4.19), system (4.17) can be written as

$$
\begin{pmatrix} \overline{S_\varepsilon^{[\leq N]}} & \overline{S_\varepsilon^{[\leq N]} (B_{b,\varepsilon})^0} + \overline{\tilde{A}_{\varepsilon,1}^N} \\ \varepsilon^3 \,\mathrm{Id} & \overline{\tilde{A}_{\varepsilon,2}^N} \end{pmatrix} \begin{pmatrix} \overline{W_{\varepsilon,2}} \\ \sigma_\varepsilon \end{pmatrix} = \begin{pmatrix} -\overline{S_\varepsilon^{[\leq N]} (B_{a,\varepsilon})^0} - \overline{\tilde{E}_{\varepsilon,1}^{(N,2N]}} \\ -\overline{\tilde{E}_{\varepsilon,2}^{(N,2N]}} \end{pmatrix} \tag{4.22}
$$

Remark 30. *Due to* **HND** *in Theorem 20 the matrix in the left hand side of (4.22) is invertible at $\varepsilon = 0$. By the continuity of the determinant, equation (4.22) can be solved for ε small enough and the inverse is analytic in ε.*

Thus, (4.19) and (4.22) yield $\sigma_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ and $W_{\varepsilon,2} = (W_{\varepsilon,2})^0 + \overline{W_{\varepsilon,2}} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$. It remains to find $W_{\varepsilon,1}$, this can be done by solving the equation

$$
(W_{\varepsilon,1})^0 - (W_{\varepsilon,1})^0 \circ T_\omega = -(S_\varepsilon^{[\leq N]} W_{\varepsilon,2})^0 - (\tilde{E}_{\varepsilon,1}^{(N,2N]})^0 - (A_{\varepsilon,1}^N)^0 \sigma_\varepsilon, \tag{4.23}
$$

which can be done due to Lemma 24. To fulfill the normalization condition (3.3) and obtain uniqueness of the coefficients of the perturbative expansions, $\overline{W}_{\varepsilon,1}$ is chosen as

$$
\overline{W}_{\varepsilon,1} = -\left(\int_{\mathbb{T}^d} \left[M_0^{-1}(\theta) DK_\varepsilon^{[\leq N]} \right]_d d\theta \right)^{-1} \int_{\mathbb{T}^d} \left[M_0^{-1}(\theta) \left(DK_\varepsilon^{[\leq N]} (W_{\varepsilon,1})^0 + V_\varepsilon^{[\leq N]} W_{\varepsilon,2} \right) \right]_d d\theta \tag{4.24}
$$

where $V^{[\leq N]} = J^{-1} \circ K_\varepsilon^{[\leq N]} DK_\varepsilon^{[\leq N]} \mathcal{N}_\varepsilon^{[\leq N]}$ is the second *column* of the matrix $M_\varepsilon^{[\leq N]}$, see Remark 17.

Remark 31. *Assuming that $K_\varepsilon^{[\leq N]}$ satisfies the normalization (3.3), then the new approximation $K_\varepsilon^{[\leq N]} + \Delta_\varepsilon$ will satisfy (3.3) if the correction satisfies*

$$
\int_{\mathbb{T}^d} M_0^{-1}(\theta) \Delta_\varepsilon(\theta) d\theta = 0.
$$

Since $\Delta_\varepsilon = M_\varepsilon^{[\leq N]} W_\varepsilon = DK_\varepsilon^{[\leq N]} W_{\varepsilon,1} + V_\varepsilon^{[\leq N]} W_{\varepsilon,2} = DK_\varepsilon^{[\leq N]} \left((W_{\varepsilon,1})^0 + \overline{W_{\varepsilon,1}} \right) + V_\varepsilon^{[\leq N]} W_{\varepsilon,2}$, (4.24) follows from the fact that $\int_{\mathbb{T}^d} \left[M_0^{-1} DK_\varepsilon^{[\leq N]} \overline{W}_{\varepsilon,1} \right]_d d\theta = \int_{\mathbb{T}^d} \left[M_0^{-1} DK_\varepsilon^{[\leq N]} \right]_d d\theta \overline{W_{\varepsilon,1}}$. Note that the $d \times d$ matrix $\int_{\mathbb{T}^d} \left[M_0^{-1}(\theta) DK_\varepsilon^{[\leq N]}(\theta) \right]_d d\theta$ is invertible, for ε small enough, due to the fact that $DK_\varepsilon^{[\leq N]}(\theta)$ is a perturbation of $DK_0(\theta)$ and $\left[M_0^{-1}(\theta) DK_0(\theta) \right]_d = I_{d\times d}$, because $M_0(\theta) = [DK_0(\theta) | V_0(\theta)]$.

This yields, $W_{\varepsilon,1} = (W_{\varepsilon,1})^0 + \overline{W}_{\varepsilon,1} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ and thus

$$\Delta_\varepsilon = M_\varepsilon^{[\leq N]} W_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right) \quad \text{and} \quad \sigma_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right). \tag{4.25}$$

which means that $\Delta_\varepsilon = \sum_{n=N+1}^{\infty} \Delta_n \varepsilon^n$ and $\sigma_\varepsilon = \sum_{n=N+1}^{\infty} \sigma_n \varepsilon^n$. Finally, we take the corrections as

$$\Delta_\varepsilon^{(N,2N]} \equiv \sum_{n=N+1}^{2N} \Delta_n \varepsilon^n \quad \text{and} \quad \sigma_\varepsilon^{(N,2N]} \equiv \sum_{n=N+1}^{2N} \sigma_n \varepsilon^n. \tag{4.26}$$

Therefore, the new approximation is chosen as

$$K_\varepsilon^{[\leq 2N]} := K_\varepsilon^{[\leq N]} + \Delta_\varepsilon^{(N,2N]} \quad \text{and} \quad \mu_\varepsilon^{[\leq 2N]} := \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon^{(N,2N]}. \tag{4.27}$$

Remark 32. *Notice that, due to Lemma 27, the solutions of* (4.20) *and* (4.21) *will satisfy* $(B_{a,\varepsilon})^0 \sim \mathcal{O}(|\varepsilon|^{N+1})$ *and* $(B_{b,\varepsilon})^0 \sim \mathcal{O}(|\varepsilon|^0)$*, because* $(\tilde{E}_{\varepsilon,2}^{(N,2N]})^0 \sim \mathcal{O}(|\varepsilon|^{N+1})$ *and* $(\tilde{A}_{\varepsilon,2}^{N})^0 \sim \mathcal{O}(|\varepsilon|^0)$*. Moreover,* (4.22) *implies that* $\overline{W}_{\varepsilon,2} \sim \mathcal{O}(|\varepsilon|^{N+1})$ *and* $\sigma_\varepsilon \sim \mathcal{O}(|\varepsilon|^{N+1})$*. Thus,* $W_{\varepsilon,2} \sim \mathcal{O}(|\varepsilon|^{N+1})$ *and similarly* $W_{\varepsilon,1} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ *which implies* $\Delta_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$*.*

4.3. **Algorithm for the iterative step.** The procedure described above leads Algorithm 33 for a given Diophantine vector ω and assuming that we are given an analytic family $f_{\varepsilon,\mu_\varepsilon}$. Some steps in the algorithm are denoted as $p \leftarrow q$, meaning that the quantity q is assigned to the variable p.

Algorithm 33. *Given* $K_\varepsilon^{[\leq N]} : \mathbb{T}^n \to \mathcal{M}$*,* $\mu_\varepsilon^{[\leq N]} \in \mathbb{R}^d$*. We perform the following computations:*

$$
\begin{array}{ll}
(1) & E_\varepsilon^N \leftarrow f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} - K_\varepsilon^{[\leq N]} \circ T_\omega \\
(2) & E_\varepsilon^{(N,2N]} \textit{ obtained from } E_\varepsilon^N \textit{ by truncation} \\
(3) & \alpha_\varepsilon \leftarrow DK_\varepsilon^{[\leq N]} \\
(4) & \mathcal{N}_\varepsilon \leftarrow [\alpha_\varepsilon^\top \alpha_\varepsilon]^{-1} \\
(5) & V_\varepsilon \leftarrow J^{-1} \circ K_\varepsilon^{[\leq N]} \alpha_\varepsilon \mathcal{N}_\varepsilon \\
(6) & M_\varepsilon \leftarrow [\alpha_\varepsilon | V_\varepsilon] \\
(7) & \beta_\varepsilon \leftarrow (M_\varepsilon \circ T_\omega)^{-1} \\
(8) & \tilde{E}_\varepsilon^{(N,2N]} \leftarrow \beta_\varepsilon E_\varepsilon^{(N,2N]} \\
(9) & P_\varepsilon \leftarrow \alpha_\varepsilon \mathcal{N}_\varepsilon \\
 & \Gamma_\varepsilon \leftarrow \alpha_\varepsilon^\top J^{-1} \circ K_\varepsilon^{[\leq N]} \alpha_\varepsilon \\
 & S_\varepsilon \leftarrow (P_\varepsilon \circ T_\omega)^\top D f_{\mu_\varepsilon^{[\leq N]},\varepsilon} \circ K_\varepsilon^{[\leq N]} J^{-1} \circ K_\varepsilon^{[\leq N]} P_\varepsilon - \lambda(\varepsilon)(\mathcal{N}_\varepsilon \circ T_\omega)^\top \Gamma_\varepsilon \circ T_\omega (\mathcal{N}_\varepsilon \circ T_\omega) \\
 & \tilde{A}_\varepsilon \leftarrow \beta_\varepsilon D_\mu f_{\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} \\
(10) & (B_{a,\varepsilon})^0 \textit{ solves } \lambda(\varepsilon)(B_{a,\varepsilon})^0 - (B_{a,\varepsilon})^0 \circ T_\omega = -(\tilde{E}_{\varepsilon,2}^{(N,2N]})^0 \\
 & (B_{b,\varepsilon})^0 \textit{ solves } \lambda(\varepsilon)(B_{b,\varepsilon})^0 - (B_{b,\varepsilon})^0 \circ T_\omega = -(\tilde{A}_{\varepsilon,2})^0 \\
(11) & \textit{Find } \overline{W_{\varepsilon,2}}, \sigma_\varepsilon \textit{ by solving} \\
 & \begin{pmatrix} \overline{S_\varepsilon} & \overline{S_\varepsilon (B_{b,\varepsilon})^0} + \overline{\tilde{A}_{\varepsilon,1}} \\ \varepsilon^3 \operatorname{Id} & \overline{\tilde{A}_{\varepsilon,2}} \end{pmatrix} \begin{pmatrix} \overline{W_{\varepsilon,2}} \\ \sigma_\varepsilon \end{pmatrix} = \begin{pmatrix} -\overline{S_\varepsilon (B_{a,\varepsilon})^0} - \overline{\tilde{E}_{\varepsilon,1}^{(N,2N]}} \\ -\overline{\tilde{E}_{\varepsilon,2}^{(N,2N]}} \end{pmatrix} \\
(12) & (W_{\varepsilon,2})^0 = (B_{a,\varepsilon})^0 + (B_{b,\varepsilon})^0 \sigma_\varepsilon \\
(13) & W_{\varepsilon,2} = (W_{\varepsilon,2})^0 + \overline{W_{\varepsilon,2}} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right) \\
(14) & (W_{\varepsilon,1})^0 \textit{ solves } (W_{\varepsilon,1})^0 - (W_{\varepsilon,1})^0 \circ T_\omega = -(S_\varepsilon W_{\varepsilon,2})^0 - (\tilde{E}_{\varepsilon,1}^{(N,2N]})^0 - (\tilde{A}_{\varepsilon,1})^0 \\
(15) & \overline{W_{\varepsilon,1}} = -\left(\int_{\mathbb{T}^d} \left[M_0^{-1}\alpha_\varepsilon\right]_1 d\theta\right)^{-1} \int_{\mathbb{T}^d} \left[M_0^{-1}\left(\alpha_\varepsilon (W_{\varepsilon,1})^0 + V_\varepsilon W_{\varepsilon,2}\right)\right]_1 d\theta \\
(16) & W_{\varepsilon,1} = (W_{\varepsilon,1})^0 + \overline{W_{\varepsilon,1}} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right) \\
(17) & \Delta_\varepsilon \leftarrow M_\varepsilon W_\varepsilon \\
(18) & K_\varepsilon^{[\leq 2N]} \leftarrow K_\varepsilon^{[\leq N]} + \Delta_\varepsilon^{(N,2N]} \\
 & \mu_\varepsilon^{[\leq 2N]} \leftarrow \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon^{(N,2N]}
\end{array}
$$

It is worth to know that all the operations in Algorithm 33 could be implemented in a few lines in a high level computer language.

Remark 34. *Note that Algorithm 33 involves only algebraic operations, compositions, derivatives, truncations, and solving cohomology equations. This implies that if we start with analytic functions then the output will be an analytic function.*

Remark 35. *Note that at each step of the iterative procedure obtained by the quasi Newton method the input will be polynomials of degree N in ε, $K_\varepsilon^{[\leq N]} \equiv \sum_{n=0}^N K_n \varepsilon^n$, and $\mu_\varepsilon^{[\leq N]} = \sum_{n=0}^N \mu_n \varepsilon^n$. The output will be polynomials of degree $2N$ in ε given by*

$$K_\varepsilon^{[\leq 2N]} := K_\varepsilon^{[\leq N]} + \Delta_\varepsilon^{(N,2N]} \quad \textit{and} \quad \mu_\varepsilon^{[\leq 2N]} := \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon^{(N,2N]}.$$

Since, by construction, $\Delta_\varepsilon^{(N,2N]} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ and $\sigma_\varepsilon^{(N,2N]} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$, the first N coefficients $K_1, K_2, ..., K_N$ of the expansion of $K^{[\leq 2N]}$ will be the same coefficients of $K^{[\leq N]}$ and they will not change for any of the next steps. The same also happens for the coefficients of $\mu_\varepsilon^{[\leq 2N]}$. This is a crucial step for proving the main lemma, Lemma 22, since due to the fact that the coefficient up to order N do not change after $\log_2(N)$ steps of the modified Newton method, one can use Cauchy estimates in the domains given by Lemma 27 after $\log_2(N)$ steps to obtain estimates on the N coefficient.

Remark 36. *To iterate the modified Newton method in Algorithm 33 it is needed that the new error* E_ε^{2N} *obtained using the new approximations* $K_\varepsilon^{[\leq 2N]} = K_\varepsilon^{[\leq N]} + \Delta_\varepsilon^{(N,2N]}$ *and* $\mu_\varepsilon^{[\leq 2N]} = \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon^{(N,2N]}$ *satisfies* $E_\varepsilon^{2N} \sim \mathcal{O}\left(|\varepsilon|^{2N+1}\right)$*. This is a consequence of the fact that the new error is quadratic in the original error, as an expansion on* ε*, and this is verified in Lemma 45.*

5. Estimates for the iterative step.

In this section we present the estimates for the corrections given by the Newton step described in Section 4, these estimates are obtained by following the steps in Algorithm 33. Throughout this section we consider maps in the spaces $\mathcal{A}_{\rho,\gamma}$. In the following we will be dealing with equations of the form (4.3) which, accordingly with Lemma 27, can be solved if

$$\varepsilon \leq \gamma_N := \left(\frac{\nu}{2}\right)^{1/\alpha} \frac{1}{(aN)^{\tau/\alpha}}. \tag{5.1}$$

where aN is the degree of the trigonometric polynomial in the right hand side of (4.3).

5.1. **Estimate for the reducibility error.** The following Lemma provides an estimate for the error in the approximate reducibility given by $R_\varepsilon^{[\leq N]}$ as in (2.12) computed from $K_\varepsilon^{[\leq N]}$. The estimates are obtained by studying qualitatively the geometric identities introduced in Section 2.6 and taking into account the uniformity on the variable ε.

Lemma 37. *Let* $N \in \mathbb{N}$, $\omega \in \mathcal{D}(\nu,\tau)$ *and* $f_{\varepsilon,\mu} : \mathcal{M} \to \mathcal{M}$ *be a family of analytic conformally symplectic maps, with* $f_{\varepsilon,\mu}^*\Omega = \lambda(\varepsilon)\Omega$, $\mu \in \Lambda \subseteq \mathbb{C}^d$*. Let* $K^{[\leq N]} \in \mathcal{A}_{\rho,\gamma_N}$ *such that* $K_\varepsilon^{[\leq N]} : \mathbb{T}^d \to \mathcal{M}$ *is an embedding for any* $|\varepsilon| \leq \gamma_N$*. Assume also that, for any* $|\varepsilon| \leq \gamma_N$*,*

i) $K_\varepsilon^{[\leq N]}\left(\mathbb{T}^d_\rho\right) \subset \mathrm{Domain}(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}})$ *and that there exist* $\xi \geq 0$ *such that*

$$\mathrm{dist}\left(K_\varepsilon^{[\leq N]}\left(\mathbb{T}^d_\rho\right), \partial\,\mathrm{Domain}(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}})\right) \geq \xi > 0$$

$$\mathrm{dist}\left(\mu_\varepsilon^{[\leq N]}, \partial\Lambda\right) \geq \xi > 0$$

ii) The approximate invariance equation holds

$$f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} - K_\varepsilon^{[\leq N]} \circ T_\omega = E_\varepsilon^N \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$$

iii)

$$\nu^{-1}(aN)^\tau\delta^{-(d+1)}\left\|E^N\right\|_{\rho,\gamma_N} \ll 1 \tag{5.2}$$

iv) **HTP2** *The* $d \times d$ *matrix*

$$\begin{aligned} E_{\Omega,\varepsilon}^N(\theta) \equiv{}& DK_\varepsilon^{[\leq N]}(\theta+\omega)^\top J \circ K_\varepsilon^{[\leq N]}(\theta+\omega) DK_\varepsilon^{[\leq N]}(\theta+\omega) \\ &- D(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta))^\top J \circ (f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) D(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) \end{aligned} \tag{5.3}$$

is a trigonometric polynomial of degree less than aN*.*

Then

$$R_\varepsilon^{[\leq N]} \sim \mathcal{O}(|\varepsilon|^{N+1}) \tag{5.4}$$

and for any $0 < \delta \leq \rho$ *we have*

$$\left\|R^{[\leq N]}\right\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^\tau\delta^{-(d+1)}\left\|E^N\right\|_{\rho,\gamma_N} \tag{5.5}$$

where $C = C(d, \left\|DK^{[\leq N]}\right\|_{\rho,\gamma_N}, \left\|\mathcal{N}^{[\leq N]}\right\|_{\rho,\gamma_N}, \left\|J \circ K^{[\leq N]}\right\|_{\rho,\gamma_N})$*.*

Proof. Writing $R_\varepsilon^{[\leq N]}$ in terms of $K_\varepsilon^{[\leq N]}$ as in (2.12) yields

$$R_\varepsilon^{[\leq N]}(\theta) = \left[DE_\varepsilon^N(\theta) \mid V_\varepsilon^{[\leq N]}(\theta+\omega)\left(B_\varepsilon(\theta)-\lambda(\varepsilon)\,\mathrm{Id}\right) + DK_\varepsilon^{[\leq N]}(\theta+\omega)\left(\tilde{S}_\varepsilon(\theta) - S_\varepsilon^{[\leq N]}(\theta)\right)\right]$$

with

$$V_\varepsilon^{[\leq N]}(\theta) \equiv J^{-1}\circ K_\varepsilon^{[\leq N]}(\theta) DK_\varepsilon^{[\leq N]}(\theta)\mathcal{N}_\varepsilon^{[\leq N]}(\theta) \tag{5.6}$$

$$B_\varepsilon(\theta)-\lambda(\varepsilon)\,\mathrm{Id} \equiv -E_{L,\varepsilon}^N(\theta+\omega)S_\varepsilon^{[\leq N]}(\theta) \tag{5.7}$$

$$\tilde{S}_\varepsilon(\theta) - S_\varepsilon^{[\leq N]}(\theta) \equiv -\mathcal{N}_\varepsilon^{[\leq N]}(\theta+\omega)^\top \Gamma_\varepsilon^{[\leq N]}(\theta+\omega)\mathcal{N}_\varepsilon^{[\leq N]}(\theta+\omega)\left(B_\varepsilon(\theta)-\lambda(\varepsilon)\,\mathrm{Id}\right) \tag{5.8}$$

where

$$E_{L,\varepsilon}^N(\theta) \equiv DK_\varepsilon^{[\leq N]}(\theta)^\top J\circ K_\varepsilon^{[\leq N]}(\theta)DK_\varepsilon^{[\leq N]}(\theta) \tag{5.9}$$

is the pull back $(K_\varepsilon^{[\leq N]})^*\Omega$ written in coordinates and $\Gamma_\varepsilon^{[\leq N]}$ as in (4.11). We recall that J is the matrix associated to the symplectic form, see Section 2. It is easy to estimate the first column of $R_\varepsilon^{[\leq N]}$ using Cauchy estimates, that is

$$\left\|DE_\varepsilon^N\right\|_{\rho-\delta} \leq C\delta^{-1}\left\|E_\varepsilon^N\right\|_\rho$$

To obtain estimates for the second column of $R_\varepsilon^{[\leq N]}$, due to (5.7) and (5.8), it is enough to get estimates of E_L^N . The estimate for E_L^N is obtained using that $f_{\varepsilon,\mu}^*\Omega = \lambda(\varepsilon)\Omega$. Note that $E_{\Omega,\varepsilon}^N = (K_\varepsilon^{[\leq N]}\circ T_\omega)^*\Omega - (f_{\varepsilon,\mu_\varepsilon^{[\leq N]}}\circ K_\varepsilon^{[\leq N]})^*\Omega$ in coordinates and, since $(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}}\circ K_\varepsilon^{[\leq N]})^*\Omega = \lambda(K_\varepsilon^{[\leq N]})^*\Omega$, we have that E_L^N satisfies the equality

$$E_{L,\varepsilon}^N\circ T_\omega - \lambda(\varepsilon)E_{L,\varepsilon}^N = E_{\Omega,\varepsilon}^N. \tag{5.10}$$

Then, by Lemma 27 and **HTP2** we obtain

$$\left\|E_L^N\right\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^\tau\delta^{-d}\left\|E_\Omega^N\right\|_{\rho-\delta/2,\gamma_N}. \tag{5.11}$$

To get estimates for E_Ω^N, we follow [CCdlL13]. If h and g are smooth maps with range in $\mathcal{M}$, the matrix corresponding to $h^*\Omega - g^*\Omega$ is

$$Dh^\top J\circ hDh - Dg^\top J\circ gDg = (Dh^\top - Dg^\top)J\circ hDh - Dg^\top(J\circ h - J\circ g)Dh + Dg^\top J\circ g(Dh-Dg)$$

Using this formula with $g = f_{\varepsilon,\mu_\varepsilon^{[\leq N]}}\circ K_\varepsilon^{[\leq N]}$, $h = K_\varepsilon^{[\leq N]}\circ T_\omega$ and Cauchy estimates one obtains

$$\left\|E_{\Omega,\varepsilon}^N\right\|_{\rho-\delta/2} \leq C\delta^{-1}\left\|E_\varepsilon^N\right\|_\rho \tag{5.12}$$

which yields $E_{L,\varepsilon}^N, E_{\Omega,\varepsilon}^N \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ and, then, $R_\varepsilon^{[\leq N]} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ and

$$\left\|R^{[\leq N]}\right\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^\tau\delta^{-(d+1)}\left\|E^N\right\|_{\rho,\gamma_N}. \tag{5.13}$$

Note that when the matrix J is constant both **HTP2** and the computations above are significantly simpler than in the general case. □

Remark 38. *We emphasize that, if K_0 satisfies $K_0\circ T_\omega - f_{0,\mu_0}\circ K_0 = 0$ then $DK_0(\theta)^\top J\circ K_0 DK_0(\theta) = 0$ and $K_0(\mathbb{T}^d)$ is a Langrangian manifold, see [CCdlL13]. This implies that the spaces Range$(DK_0(\theta))$ and Range$(J^{-1}\circ K_0(\theta)DK_0(\theta))$ are transversal and this condition makes $M_0(\theta)$ a linear isomorphism. Note that if E_L^N in (5.9) represents the error of the lagrangian character of $K_\varepsilon^{[\leq N]}$, then, if E_L^N is small enough the spaces Range$(DK_\varepsilon^{[\leq N]}(\theta))$ and Range$(J^{-1}\circ K_\varepsilon^{[\leq N]}(\theta)DK_\varepsilon^{[\leq N]}(\theta))$ will be transversal and the matrix $M_\varepsilon^{[\leq N]}$ will define a linear isomorphism. This transversality will be obtained if (5.2) is satisfied and it is given by (5.11) and (5.12).*

5.2. **Estimates for the corrections.** In this sections we obtain estimates for the corrections $\Delta^{(N,2N]}$ and $\sigma^{(N,2N]}$, this estimates are obtained by following the steps in Algorithm 33. First, Lemma 39, we obtain estimates for the corrections Δ_ε, σ_ε and then, using Cauchy estimates, we obtain estimates for the truncations $\Delta^{(N,2N]}$, $\sigma^{(N,2N]}$, Corollary 40.

Consider $\mathcal{C} \subseteq \mathbb{C}^d/\mathbb{Z}^d \times \mathbb{C}^d$ the complexification of $\mathcal{M} = \mathbb{T}^d \times B$.

Lemma 39. *Let $a \in \mathbb{N}$, $0 < \rho < 1$, and δ such that $0 < 2\delta < \rho$. Assume that for any $\varepsilon \in \mathbb{C}$, such that $|\varepsilon| < \gamma_N$, $f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} : \mathcal{C} \to \mathcal{C}$ is an analytic conformally symplectic map with $f^*_{\varepsilon,\mu_\varepsilon^{[\leq N]}}\Omega = \lambda(\varepsilon)\Omega$. Assume also that $K^{[\leq N]} \in \mathcal{A}_{\rho,\gamma_N}$ is such that $K_\varepsilon^{[\leq N]} : \mathbb{T}^d_\rho \to \mathbb{C}^d/\mathbb{Z}^d \times \mathbb{C}^d$ is an embedding. Assume also that for any $|\varepsilon| < \gamma_N$ we have the following:*

i) $K_\varepsilon^{[\leq N]}\left(\mathbb{T}^d_\rho\right) \subset \mathrm{Domain}(f_{\varepsilon,\mu^{[\leq N]}})$ *and that there exist $\xi \geq 0$ such that*

$$\mathrm{dist}\left(K_\varepsilon^{[\leq N]}\left(\mathbb{T}^d_\rho\right), \partial\,\mathrm{Domain}(f_{\varepsilon,\mu^{[\leq N]}})\right) \geq \xi > 0$$

$$\mathrm{dist}\left(\mu_\varepsilon^{[\leq N]}, \partial\Lambda\right) \geq \xi$$

ii) **HND**. *The following non-degeneracy condition holds:*

$$\det\begin{pmatrix} \overline{S_\varepsilon^{[\leq N]}} & \overline{S_\varepsilon^{[\leq N]}(B_{b,\varepsilon})^0} + \overline{\tilde{A}^N_{\varepsilon,1}} \\ \varepsilon^3\,\mathrm{Id} & \overline{\tilde{A}^N_{\varepsilon,2}} \end{pmatrix} \neq 0$$

iii) For any $N \in \mathbb{N}$, the matrices $(\tilde{E}^{(N,2N]}_{\varepsilon,2})^0$ and $(\tilde{A}^N_{\varepsilon,2})^0$ defined in (4.13) and (4.14), are trigonometric polynomials of degree less or equal than aN.

Then, for any $0 < r < 1$ we have

$$W_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right), \quad \sigma_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right) \tag{5.14}$$

$$\|W\|_{\rho-\delta,r\gamma_N} \leq C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N} \tag{5.15}$$

and

$$\sup_{|\varepsilon|\leq r\gamma_N}|\sigma_\varepsilon| \leq C\nu^{-1}(aN)^\tau\delta^{-d}\frac{r^{N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N} \tag{5.16}$$

where $C = C(d, \left\|DK^{[\leq N]}\right\|_{\rho,\gamma_N}, \left\|M^{[\leq N]}\right\|_{\rho,\gamma_N}, \left\|(M^{[\leq N]})^{-1}\right\|_{\rho,\gamma_N}, \left\|\mathcal{N}^{[\leq N]}\right\|_{\rho,\gamma_N}, \mathcal{T}^N)$ and $\mathcal{T}^N$ is defined in (5.20).

Proof. Given that $(\tilde{E}^{(N,2N]}_{\varepsilon,2})^0$ and $(\tilde{A}^N_{\varepsilon,2})^0$ are trigonometric polynomials, by Lemma 27, (4.20), and (4.21); B_a and B_b satisfy the following estimates

$$\begin{aligned}\|B_a\|_{\rho-\delta,r\gamma_N} &\leq C\nu^{-1}(aN)^\tau\delta^{-d}\left\|\tilde{E}_2^{(N,2N]}\right\|_{\rho,r\gamma_N} \\ &\leq C\nu^{-1}(aN)^\tau\delta^{-d}\left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N}\end{aligned} \tag{5.17}$$

and similarly

$$\|B_b\|_{\rho-\delta,r\gamma_N} \leq C\nu^{-1}(aN)^\tau\delta^{-d}\left\|A^N\right\|_{\rho,r\gamma_N}. \tag{5.18}$$

Taking into account that $W_2 = (W_2)^0 + \overline{W_2}$ and $(W_2)^0 = (B_a)^0 + \sigma(B_b)^0$, to have estimates for W_2 we need estimates for $\overline{W_2}$ and σ. Now, according to (4.22) we have

$$\begin{pmatrix}\overline{W_{\varepsilon,2}} \\ \sigma_\varepsilon\end{pmatrix} = \begin{pmatrix} \overline{S_\varepsilon^{[\leq N]}} & \overline{S_\varepsilon^{[\leq N]}(B_{b,\varepsilon})^0} + \overline{\tilde{A}^N_{\varepsilon,1}} \\ \varepsilon^3\,\mathrm{Id} & \overline{\tilde{A}^N_{\varepsilon,2}} \end{pmatrix}^{-1}\begin{pmatrix} -\overline{S_\varepsilon^{[\leq N]}(B_{a,\varepsilon})^0} - \overline{\tilde{E}^{(N,2N]}_{\varepsilon,1}} \\ -\overline{\tilde{E}^{(N,2N]}_{\varepsilon,2}}\end{pmatrix}, \tag{5.19}$$

denoting

$$\mathcal{T}_\varepsilon^N := \left\| \begin{pmatrix} \overline{S_\varepsilon^{[\leq N]}} & \overline{S_\varepsilon^{[\leq N]}(B_{b,\varepsilon})^0} + \overline{\tilde{A}^N_{\varepsilon,1}} \\ \varepsilon^3 \,\mathrm{Id} & \overline{\tilde{A}^N_{\varepsilon,2}} \end{pmatrix}^{-1} \right\| \qquad \text{and} \qquad \mathcal{T}^N = \sup_{|\varepsilon|\leq r\gamma_N} \mathcal{T}_\varepsilon^N \tag{5.20}$$

from (5.19) we have

$$|\sigma_\varepsilon|, |\overline{W_{\varepsilon,2}}| \leq \mathcal{T}_\varepsilon^N \left(\left| \overline{S_\varepsilon^{[\leq N]}(B_{a,\varepsilon})^0} + \overline{\tilde{E}^{(N,2N]}_{\varepsilon,1}} \right| + \left| \overline{\tilde{E}^{(N,2N]}_{\varepsilon,2}} \right| \right) \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right) \tag{5.21}$$

which yields $\sigma_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ and $\overline{W_{\varepsilon,2}} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ because $(B_{a,\varepsilon})^0 \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$ and $\tilde{E}^{(N,2N]}_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$.

Thus

$$\begin{aligned} |\sigma_\varepsilon|, |\overline{W_{\varepsilon,2}}| &\leq \mathcal{T}_\varepsilon^N \left(\left| \overline{S_\varepsilon^{[\leq N]}(B_{a,\varepsilon})^0} \right| + \left| \overline{\tilde{E}^{(N,2N]}_{\varepsilon,1}} \right| + \left| \overline{\tilde{E}^{(N,2N]}_{\varepsilon,2}} \right| \right) \\ &\leq C\mathcal{T}^N \left(\left\| S_\varepsilon^{[\leq N]} \right\|_\rho \left\| (B_{a,\varepsilon})^0 \right\|_{\rho-\delta} + \left\| \tilde{E}^{(N,2N]}_{\varepsilon,1} \right\|_\rho + \left\| \tilde{E}^{(N,2N]}_{\varepsilon,2} \right\|_\rho \right) \end{aligned}$$

for any $0 < \delta < \rho$. Thus, using (4.13) and (5.17) we obtain

$$\sup_{|\varepsilon|\leq r\gamma_N} \left|\overline{W_{\varepsilon,2}}\right| \leq C\nu^{-1}(aN)^\tau \delta^{-d} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N} \tag{5.22}$$

$$\sup_{|\varepsilon|\leq r\gamma_N} |\sigma_\varepsilon| \leq C\nu^{-1}(aN)^\tau \delta^{-d} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N}. \tag{5.23}$$

For $(W_2)^0 = (B_a)^0 + \sigma(B_b)^0$ we have

$$\begin{aligned} \left\|(W_2)^0\right\|_{\rho-\delta,r\gamma_N} &\leq \left\|(B_a)^0\right\|_{\rho-\delta,r\gamma_N} + \sup_{|\varepsilon|\leq r\gamma_N} |\sigma| \left\|(B_b)^0\right\|_{\rho-\delta,r\gamma_N} \\ &\leq C\nu^{-1}(aN)^\tau \delta^{-d} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N} + C\nu^{-2}(aN)^{2\tau}\delta^{-2d} \left\|A^N\right\|_{\rho,r\gamma_N} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N}, \\ &\leq C\nu^{-2}(aN)^{2\tau}\delta^{-2d} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N}. \end{aligned} \tag{5.24}$$

Thus, combining (5.22) and (5.24) we get

$$\|W_2\|_{\rho-\delta,r\gamma_N} \leq C\nu^{-2}(aN)^{2\tau}\delta^{-2d} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N} \tag{5.25}$$

The estimates for $(W_1)^0$ come from (4.23) and Lemma 24, i.e.,

$$\begin{aligned} &\left\|(W_1)^0\right\|_{\rho-2\delta,r\gamma_N} \\ &\leq C\nu^{-1}\delta^{-(\tau+d)} \left[\left\| S^{[\leq N]} \right\|_{\rho-\delta,r\gamma_N} \|W_2\|_{\rho-\delta,r\gamma_N} + \left\| \tilde{E}^{(N,2N]} \right\|_{\rho-\delta,r\gamma_N} + \sup_{|\varepsilon|\leq r\gamma_N} |\sigma_\varepsilon| \left\| \tilde{A}^N \right\|_{\rho-\delta,r\gamma_N} \right] \\ &\leq C\nu^{-1}\delta^{-(\tau+d)} \left[\left\| S^{[\leq N]} \right\|_{\rho,r\gamma_N} \nu^{-2}(aN)^{2\tau}\delta^{-2d} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N} \right. \\ &\qquad \left. + \left\| \left(M^{[\leq N]}\right)^{-1} \right\|_{\rho,r\gamma_N} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N} + \left\|A^N\right\|_{\rho,r\gamma_N} \nu^{-1}(aN)^\tau \rho^{-d} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N} \right] \end{aligned}$$

that is,

$$\left\|(W_1)^0\right\|_{\rho-2\delta,r\gamma_N} \leq C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)} \left\| E^{(N,2N]} \right\|_{\rho,r\gamma_N}. \tag{5.26}$$

Finally, the estimate for $\overline{W_1}$ comes from (4.24), that is

$$\begin{aligned}\sup_{|\varepsilon|\le r\gamma_N} \left|\overline{W_{\varepsilon,1}}\right| &\le C\left(\left\|(W_1)^0\right\|_{\rho-\delta,r\gamma_N} + \|W_2\|_{\rho-\delta,r\gamma_N}\right) \\ &\le C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N}. \end{aligned} \tag{5.27}$$

Putting together (5.25), (5.26), (5.27), and using the Cauchy estimates in Corollary 5 yields the claimed estimate for W. □

Corollary 40. *Assuming the hypothesis of Lemma 37 and Lemma 39, for any $0<\delta<\rho$ and $0<r<1$ we have*

$$\left\|\Delta^{(N,2N]}\right\|_{\rho-\delta,r\gamma_N} \le C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} \tag{5.28}$$

$$\sup_{|\varepsilon|\le r\gamma_N}\left|\sigma_\varepsilon^{(N,2N]}\right| \le C\nu^{-1}(aN)^{\tau}\delta^{-d}\frac{r^{N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} \tag{5.29}$$

Moreover,

$$\left\|\Delta^{(2N,\infty]}\right\|_{\rho-\delta,r\gamma_N} \le C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} \tag{5.30}$$

$$\sup_{|\varepsilon|\le r\gamma_N}\left|\sigma_\varepsilon^{(2N,\infty]}\right| \le C\nu^{-1}(aN)^{\tau}\delta^{-d}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} \tag{5.31}$$

Proof. Using the Cauchy estimates as in Corollary 5 and the estimates in Lemma 39 one obtains

$$\begin{aligned}\left\|\Delta^{(2N,\infty]}\right\|_{\rho-\delta,r^2\gamma_N} &\le \frac{r^{2N+1}}{(1-r)}\|\Delta\|_{\rho-\delta,r\gamma_N} \\ &\le C\frac{r^{2N+1}}{1-r}\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N} \\ &= C\frac{r^{3N+2}}{(1-r)^2}\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\left\|E^N\right\|_{\rho,\gamma_N}\end{aligned}$$

and

$$\begin{aligned}\sup_{|\varepsilon|\le r^2\gamma_N}\left|\sigma_\varepsilon^{(2N,\infty]}\right| &\le \frac{r^{2N+1}}{1-r}\sup_{|\varepsilon|\le r\gamma_N}|\sigma_\varepsilon| \\ &\le \frac{r^{2N+1}}{1-r}C\nu^{-1}(aN)^{\tau}\delta^{-d}\frac{r^{N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N} \\ &= C\nu^{-1}(aN)^{\tau}\delta^{-d}\frac{r^{3N+2}}{(1-r)^2}\left\|E^N\right\|_{\rho,\gamma_N}\end{aligned}$$

The other estimates are obtained similarly. □

5.3. **Non-linear estimates for the quasi-Newton method.** The quasi-Newton procedure in Algorithm 33 can also be described using a convenient operator notation. Defining the error functional

$$\mathcal{E}[K_\varepsilon,\mu_\varepsilon] = f_{\varepsilon,\mu_\varepsilon}\circ K_\varepsilon - K_\varepsilon\circ T_\omega \tag{5.32}$$

and assuming Δ and σ are *small* enough, the Taylor expansion of $\mathcal{E}[K+\Delta,\mu+\sigma]$ is given by

$$\mathcal{E}[K+\Delta,\mu+\sigma] = \mathcal{E}[K,\mu] + D_1\mathcal{E}[K,\mu]\Delta + D_2\mathcal{E}[K,\mu]\sigma + \mathcal{R}[\Delta,\sigma;K,\mu] \tag{5.33}$$

where the Frechet derivatives are given by

$$D_1\mathcal{E}[K_\varepsilon,\mu_\varepsilon]\Delta_\varepsilon = (Df_{\varepsilon,\mu_\varepsilon}\circ K_\varepsilon)\,\Delta_\varepsilon - \Delta_\varepsilon\circ T_\omega \tag{5.34}$$

$$D_2\mathcal{E}[K_\varepsilon,\mu_\varepsilon]\sigma_\varepsilon = (D_\mu f_{\varepsilon,\mu_\varepsilon}\circ K_\varepsilon)\,\sigma_\varepsilon \tag{5.35}$$

and $\mathcal{R}$ is the remainder of the Taylor expansion. Note that $\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}] = E_\varepsilon^N$, with this notation the *classic* Newton method would consist in finding a correction $(\Delta_\varepsilon^{(N,2N]},\mu_\varepsilon^{(N,2N]})$ such that

$$\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}] + D_1\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\Delta_\varepsilon^{(N,2N]} + D_2\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\sigma_\varepsilon^{(N,2N]} = 0. \tag{5.36}$$

As it was explained before, in Section 4.2, the corrections we construct with Algorithm 33 do not satisfy (5.36) but they solve an approximate equation (4.12). The following Lemmas give estimates for the error functional evaluated in the corrected unknowns. First, Lemma 43, we give estimates for the error $\mathcal{E}[K^{[\le N]}+\Delta,\mu^{[\le N]}+\sigma]$ and then, using Cauchy estimates, we obtain the estimates for the error evaluated in the truncated corrections, $\mathcal{E}[K^{[\le N]}+\Delta^{(N,2N]},\mu^{[\le N]}+\sigma^{(N,2N]}]$, Proposition 45.

Remark 41. *We emphasize that to be able to compute $\mathcal{E}[K+\Delta,\mu+\sigma]$ we need both Δ and σ to be* small enough*, so the compositions in* (5.32) *are well defined. In particular Δ and σ need to satisfy $\|\Delta\|,|\sigma|\le\xi$ and we need to choose the domain loss. In Section 6, Lemma 49, we give smallness conditions on the initial error which will guarantee that the compositions will be defined at any step of the iteration. This is very standard in KAM theory.*

Lemma 42. *Assume $0<r<1$ and $0<\delta\le\rho$. Then, under the hypothesis of Lemma 37 and Lemma 39 one has*

$$\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}] + D_1\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\Delta_\varepsilon + D_2\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\sigma_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{2N+1}\right) \tag{5.37}$$

and

$$\begin{aligned}\Big\| \mathcal{E}[K^{[\le N]},\mu^{[\le N]}] + D_1\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]\Delta + D_2\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]\sigma \Big\|_{\rho-\delta,r\gamma_N}& \\ \le \frac{r^{2N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N} + C\nu^{-4}(aN)^{3\tau}\delta^{-(\tau+4d+1)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|^2_{\rho,\gamma_N}&\end{aligned} \tag{5.38}$$

Proof. Note that with the operator notation introduced at the beginning of this section we have $\mathcal{E}(K^{[\le N]},\mu^{[\le N]}) = E^N$. Using (2.10) and taking into account that $\Delta_\varepsilon = M_\varepsilon^{[\le N]}W_\varepsilon$ and that W_ε satisfies (4.12) we have

$$\begin{aligned}
&\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}] + D_1\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\Delta_\varepsilon + D_2\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\sigma_\varepsilon\\
&\quad= E_\varepsilon^N + \left(Df_{\varepsilon,\mu^{[\le N]}}\circ K_\varepsilon^{[\le N]}\right)\Delta_\varepsilon - \Delta_\varepsilon\circ T_\omega + \left(D_\mu f_{\varepsilon,\mu^{[\le N]}}\circ K_\varepsilon^{[\le N]}\right)\sigma_\varepsilon - R_\varepsilon^{[\le N]}\left(M_\varepsilon^{[\le N]}\right)^{-1}\Delta_\varepsilon\\
&\qquad + R_\varepsilon^{[\le N]}\left(M_\varepsilon^{[\le N]}\right)^{-1}\Delta_\varepsilon\\
&\quad= E_\varepsilon^N + M_\varepsilon^{[\le N]}\circ T_\omega\begin{pmatrix}\mathrm{Id} & S_\varepsilon^{[\le N]}\\ 0 & \lambda(\varepsilon)\,\mathrm{Id}\end{pmatrix}\left(M_\varepsilon^{[\le N]}\right)^{-1}\Delta_\varepsilon - \Delta_\varepsilon\circ T_\omega + \left(D_\mu f_{\varepsilon,\mu^{[\le N]}}\circ K_\varepsilon^{[\le N]}\right)\sigma_\varepsilon\\
&\qquad + R_\varepsilon^{[\le N]}\left(M_\varepsilon^{[\le N]}\right)^{-1}\Delta_\varepsilon\\
&\quad= E_\varepsilon^N - E_\varepsilon^{(N,2N]} + R_\varepsilon^{[\le N]}W_\varepsilon \qquad\qquad (5.39)\\
&\quad= E_\varepsilon^{(2N,\infty]} + R_\varepsilon^{[\le N]}W_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{2N+1}\right)
\end{aligned}$$

where $E_\varepsilon^{(2N,\infty]} = \sum_{n=2N+1}^{\infty} E_n \varepsilon^n$. Note that the order of ε in the last line follows from the definition of $E^{(2N,\infty]}$, (5.4), and (5.14).

Then, using the Cauchy estimates of Corollary 5, Lemma 37, and Lemma 39 one obtains

$$\begin{aligned}
&\left\| \mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}] + D_1\mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}]\Delta + D_2\mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}]\sigma \right\|_{\rho-\delta, r\gamma_N} \\
&\leq \left\| E^{(2N,\infty]} \right\|_{\rho-\delta, r\gamma_N} + \left\| R^{[\leq N]} \right\|_{\rho-\delta, r\gamma_N} \|W\|_{\rho-\delta, r\gamma_N} \\
&\leq \frac{r^{2N+1}}{1-r} \left\| E^N \right\|_{\rho,\gamma_N} + C\nu^{-4}(aN)^{3\tau}\delta^{-(\tau+4d+1)} \frac{r^{N+1}}{1-r} \left\| E^N \right\|^2_{\rho,\gamma_N}
\end{aligned}$$

□

Lemma 43. *Assume $0 < r < 1$ and $0 < \delta \leq \rho$. Then, under the hypothesis of Lemma 39 and Lemma 37 we have*

$$\mathcal{E}(K_\varepsilon^{[\leq N]} + \Delta_\varepsilon, \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon) \sim \mathcal{O}\left(|\varepsilon|^{2N+1}\right) \tag{5.40}$$

and

$$\left\| \mathcal{E}[K^{[\leq N]} + \Delta, \mu^{[\leq N]} + \sigma] \right\|_{\rho-\delta, r\gamma_N} \leq \frac{r^{2N+1}}{1-r} \left\| E^N \right\|_{\rho,\gamma_N} + C\nu^{-6}(aN)^{4\tau}\delta^{-(2\tau+6d)} \frac{r^{N+1}}{1-r} \left\| E^N \right\|^2_{\rho,\gamma_N} \tag{5.41}$$

where $C = C\left(\left\| DK^{[\leq N]} \right\|_{\rho,\gamma_N}, \left\| D^2 f_{\mu^{[\leq N]}} \circ K^{[\leq N]} \right\|_{\rho,\gamma_N}, \left\| D^2_\mu f_{\mu^{[\leq N]}} \circ K^{[\leq N]} \right\|_{\rho,\gamma_N} \right)$.

Proof. Note that $\mathcal{R}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}, \Delta_\varepsilon, \sigma_\varepsilon]$ in (5.33) can be estimated using Taylor estimates for the remainder, that is

$$\|\mathcal{R}_\varepsilon\|_\rho \leq C\left(\|\Delta_\varepsilon\|^2_\rho + |\sigma_\varepsilon|^2 \right) \tag{5.42}$$

where C is a constant depending on the norms of the second derivatives of $f_{\varepsilon,\mu}$ evaluated at $K_\varepsilon^{[\leq N]}$ and $\mu_\varepsilon^{[\leq N]}$.

Since $f_{\varepsilon,\mu}$ is assumed to be analytic it is natural to expect the quantities $\left\| D^2 f_{\mu^{[\leq N]}} \circ K^{[\leq N]} \right\|_{\rho,\gamma_N}$, $\left\| D^2_\mu f_{\mu^{[\leq N]}} \circ K^{[\leq N]} \right\|_{\rho,\gamma_N}$ to be close to $\left\| D^2 f_{\mu^{[\leq N_0]}} \circ K^{[\leq N_0]} \right\|_{\rho_0,\gamma_{N_0}}$, $\left\| D^2_\mu f_{\mu^{[\leq N_0]}} \circ K^{[\leq N_0]} \right\|_{\rho_0,\gamma_{N_0}}$, at the first step of the iterations. For now, we assume that C is uniform constant. In Section 6, Lemma 49, we give sufficient conditions on the initial error of the iteration that imply that C can be taken as an uniform constant during all the iterations.

Note that (5.42) yields $\mathcal{R}_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{2N+2}\right)$. This, together with (5.37), gives (5.40). Moreover, taking sup with respect to ε one obtains

$$\begin{aligned}
\|\mathcal{R}\|_{\rho-\delta, r\gamma_N} &\leq C\left(\|\Delta\|^2_{\rho-\delta, r\gamma_N} + \sup_{|\varepsilon|\leq r\gamma_N} |\sigma|^2 \right) \\
&\leq C\left(\left\| M^{[\leq N]} \right\|^2_{\rho,\gamma_N} \|W\|_{\rho-\delta, r\gamma_N} + \sup_{|\varepsilon|\leq r\gamma_N} |\sigma|^2 \right) \\
&\leq C\left(\nu^{-6}(aN)^{4\tau}\delta^{-(2\tau+6d)} \frac{r^{2N+2}}{(1-r)^2} \left\| E^N \right\|_{\rho, r\gamma_N} + \nu^{-2}(aN)^{2\tau}\delta^{-2d} \frac{r^{2N+N}}{(1-r)^2} \left\| E^N \right\|_{\rho, r\gamma_N} \right) \\
&\leq C\nu^{-6}(aN)^{4\tau}\delta^{-(2\tau+6d)} \frac{r^{2N+2}}{(1-r)^2} \left\| E^N \right\|^2_{\rho, r\gamma_N}
\end{aligned}$$

where in the third line we use the inequalities in Lemma 39. Finally, this inequality, Lemma 42, and (5.33) give the result. □

Note that the estimates above are done for the analytic functions Δ and σ. It is only left to get the respective estimates for the truncations $\Delta^{(N,2N]}$ and $\sigma^{(N,2N]}$, which are an easy consequence of the Cauchy inequalities and are given in the following propositions.

Proposition 44. *Assuming the hypothesis of Lemma 37 and Lemma 39, for any $0<\delta<\rho$ and $0<r<1$ we have*

$$\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]+D_1\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\Delta_\varepsilon^{(N,2N]}+D_2\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\sigma_\varepsilon^{(N,2N]}\sim\mathcal{O}\left(|\varepsilon|^{2N+1}\right) \tag{5.43}$$

and

$$\begin{aligned}&\left\|\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]+D_1\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]\Delta^{(N,2N]}+D_2\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]\sigma^{(N,2N]}\right\|_{\rho-\delta,r\gamma_N}\\&\qquad\le C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N}+C\nu^{-4}(aN)^{3\tau}\delta^{-(\tau+4d+1)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|^2_{\rho,\gamma_N}\end{aligned} \tag{5.44}$$

Proof. Recalling the notation $\Delta_\varepsilon^{(a,\infty]}\equiv\sum_{n=a+1}^\infty\Delta_n(\theta)\varepsilon^n$ we have that $\Delta^{(N,2N]}+\Delta^{(2N,\infty]}=\Delta$. Also remember that $E^N=\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]$, then, using the linearity of the Frechet derivatives one obtains

$$\begin{aligned}&\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]+D_1\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\Delta_\varepsilon^{(N,2N]}+D_2\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\sigma_\varepsilon^{(N,2N]}\\&\quad=\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]+D_1\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\Delta_\varepsilon+D_2\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\sigma_\varepsilon\\&\qquad-D_1\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\Delta_\varepsilon^{(2N,\infty]}-D_2\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\sigma_\varepsilon^{(2N,\infty]}\\&\quad=\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]+D_1\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\Delta_\varepsilon+D_2\mathcal{E}[K_\varepsilon^{[\le N]},\mu_\varepsilon^{[\le N]}]\sigma_\varepsilon\\&\qquad-\left(Df_{\varepsilon,\mu^{[\le N]}}\circ K_\varepsilon^{[\le N]}\right)\Delta_\varepsilon^{(2N,\infty]}+\Delta_\varepsilon^{(2N,\infty]}\circ T_\omega-\left(D_\mu f_{\varepsilon,\mu^{[\le N]}}\circ K_\varepsilon^{[\le N]}\right)\sigma_\varepsilon^{(2N,\infty]}\end{aligned}$$

which implies (5.43). Moreover, using the relation above and the estimates in Lemma 42 and Lemma 40 one gets

$$\begin{aligned}&\left\|\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]+D_1\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]\Delta^{(N,2N]}+D_2\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]\sigma^{(N,2N]}\right\|_{\rho-\delta,r\gamma_N}\\&\quad\le\left\|\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]+D_1\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]\Delta+D_2\mathcal{E}[K^{[\le N]},\mu^{[\le N]}]\sigma\right\|_{\rho-\delta,r\gamma_N}\\&\qquad+C(\left\|\Delta^{(2N,\infty]}\right\|_{\rho-\delta,r\gamma_N}+\sup_{|\varepsilon|\le r\gamma_N}\left|\sigma_\varepsilon^{(2N,\infty]}\right|)\\&\quad\le\frac{r^{2N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N}+C\nu^{-4}(aN)^{3\tau}\delta^{-(\tau+4d+1)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|^2_{\rho,\gamma_N}\\&\qquad+C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N}+C\nu^{-1}(aN)^{\tau}\rho^{-d}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N}\\&\quad\le C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N}+C\nu^{-4}(aN)^{3\tau}\delta^{-(\tau+4d+1)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|^2_{\rho,\gamma_N}\end{aligned}$$

∎

Proposition 45. *Assuming the hypothesis of Lemma 37 and Lemma 39, for any $0<\delta<\rho$ and $0<r<1$ we have*

$$\mathcal{E}\left[K_\varepsilon^{[\le N]}+\Delta_\varepsilon^{(N,2N]},\mu_\varepsilon^{[\le N]}+\sigma_\varepsilon^{(N,2N]}\right]\sim\mathcal{O}\left(|\varepsilon|^{2N+1}\right) \tag{5.45}$$

and

$$\left\|\mathcal{E}[K^{[\leq N]}+\Delta^{(N,2N]},\mu^{[\leq N]}+\sigma^{(N,2N]}]\right\|_{\rho-\delta,r\gamma_N} \tag{5.46}$$
$$\leq C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N}+C\nu^{-6}(aN)^{4\tau}\delta^{-(2\tau+6d)}\frac{r^{N+1}}{(1-r^{1/2})^4}\left\|E^N\right\|^2_{\rho,\gamma_N}$$

where $C=C(d,\left\|M^{[\leq N]}\right\|_{\rho,\gamma_N},\left\|\left(M^{[\leq N]}\right)^{-1}\right\|_{\rho,\gamma_N},\left\|\mathcal{N}^{[\leq N]}\right\|_{\rho,\gamma_N},\left\|DK^{[\leq N]}\right\|_{\rho,\gamma_N},\mathcal{T})$*, the constant* C *also depends on the norms of the first and second derivatives of* $f_{\varepsilon,\mu}$ *evaluated at* $K_\varepsilon^{[\leq N]}$ *and* $\mu_\varepsilon^{[\leq N]}$*.*

Proof. The expansion (5.45) follows from using the same argument as in the proof of Lemma 43. We also have

$$\left\|\mathcal{R}\left[K^{[\leq N]},\mu^{[\leq N]},\Delta^{(N,2N]},\sigma^{(N,2N]}\right]\right\|_{\rho-\delta,r\gamma_N}$$
$$\leq C\left(\left\|\Delta^{(N,2N]}\right\|^2_{\rho-\delta,r\gamma_N}+\sup_{|\varepsilon|\leq r\gamma_N}\left|\sigma_\varepsilon^{(N,2N]}\right|^2\right)$$
$$\leq C\left(\nu^{-6}(aN)^{4\tau}\delta^{-(2\tau+6d)}\frac{r^{2N+2}}{(1-r^{1/2})^4}\left\|E^N\right\|^2_{\rho-\delta,r\gamma_N}+\nu^{-2}(aN)^{2\tau}\rho^{-2d}\frac{r^{2N+2}}{(1-r^{1/2})^4}\left\|E^N\right\|^2_{\rho-\delta,r\gamma_N}\right).$$

Combining this estimate with (5.44) in Lemma 44 one gets (5.46). ☐

6. Iteration of the quasi-Newton method.

We start this section giving the choice of parameters which quantify the loss of regularity at any step of the quasi Newton method. Lemma 49 will guarantee that the Newton method is well defined at any step. We note that we have loss of domain in both the variable on the torus, θ, and the variable of the perturbation, ε. In contrast with the regular KAM theory we end up losing much more domain in ε, so that at the end we do not have any ε domain.

6.1. **The iterative procedure.** We denote by $h\in\mathbb{N}$ the number of steps of the quasi Newton method. We consider

$$\delta_h:=\frac{\rho_0}{2^{h+2}}\quad\text{and}\quad\rho_{h+1}:=\rho_h-\delta_h\geq\frac{\rho_0}{2}\quad\text{for } h\geq 1, \tag{6.1}$$

where ρ_h denotes the radius of analyticity in the variable θ at step h, that is, at step h we will be considering functions in the space $\mathcal{A}_{\rho_h}$. Note that $\rho_0=\rho'$ can be the one given in Theorem 20. Since at any step we double the number of coefficients of the Lindstedt expansions, we have,

$$N_h:=2^hN_0 \tag{6.2}$$

and

$$\tilde{\gamma}_h:=\gamma_{N_h}=\left(\frac{\nu}{2}\right)^{1/\alpha}\frac{1}{(aN_h)^{\tau/\alpha}}=\left(\frac{\nu}{2}\right)^{1/\alpha}\frac{1}{(a2^hN_0)^{\tau/\alpha}} \tag{6.3}$$

where $\alpha\in\mathbb{N}$ is the exponent in $\lambda(\varepsilon)=1-\varepsilon^\alpha$, $a\in\mathbb{N}$, and $N_0\in\mathbb{N}$ is a fixed constant to be chosen later. Note that $\tilde{\gamma}_h$ is the radius of the domain of analyticity in the variable ε at step h, that is, at step h we will be considering functions in the space $\mathcal{A}_{\rho_h,\tilde{\gamma}_h}$. Also note that

$$\tilde{\gamma}_{h+1}=2^{-\tau/\alpha}\tilde{\gamma}_h. \tag{6.4}$$

Denoting $K_0:=K^{[\leq N_0]}$ and $\mu_0:=\mu^{[\leq N_0]}$, for $h\geq 1$ we have

$$K_h:=K^{[\leq N_0]}+\Delta^{(N_0,N_1]}+\cdots+\Delta^{(N_{h-1},N_h]}\qquad \mu_h:=\mu^{[\leq N_0]}+\sigma^{(N_0,N_1]}+\cdots+\sigma^{(N_{h-1},N_h]}. \tag{6.5}$$

Furthermore, denoting

$$\Delta_h:=\Delta^{(N_h,N_{h+1}]}\quad\text{and}\quad\sigma_h:=\sigma^{(N_h,N_{h+1}]}\quad\text{for } h\geq 0 \tag{6.6}$$

we have that, for $h \geq 0$

$$K_{h+1} = K_h + \Delta_h \quad \text{and} \quad \mu_{h+1} = \mu_h + \sigma_h. \tag{6.7}$$

Finally, denote also

$$\begin{aligned}
e_h &:= \|\mathcal{E}[K_h, \mu_h]\|_{\rho_h, \tilde{\gamma}_h} = \left\|E^{N_h}\right\|_{\rho_h, \tilde{\gamma}_h} &(6.8)\\
d_h &:= \|\Delta_h\|_{\rho_{h+1}, \tilde{\gamma}_{h+1}} &(6.9)\\
v_h &:= \|D\Delta_h\|_{\rho_{h+1}, \tilde{\gamma}_{h+1}} &(6.10)\\
s_h &:= \sup_{|\varepsilon| \leq \tilde{\gamma}_{h+1}} |\sigma_h(\varepsilon)|. &(6.11)
\end{aligned}$$

Remark 46. *We emphasize the dependence of $\tilde{\gamma}_h$ in N_h, note that $\tilde{\gamma}_h \to 0$ as $N_h \to \infty$ ($h \to \infty$). This implies that this quasi Newton method will not converge in any Banach space $\mathcal{A}_{\rho_h, \tilde{\gamma}_h}$, because the domains in ε shrink to 0, however, at each step we get estimates in balls with positive radius, $\tilde{\gamma}_h$. An analysis of these bounds will provide us with estimates of the coefficients of the expansion. Note also that to start with $e_0 \ll 1$ we require N_0 sufficiently large in the formal power series in Theorem 20.*

Note that with this new notation the estimates in Corollary 40 can be written as

$$d_h \leq \hat{C}_h \nu^{-3} (aN_h)^{2\tau} \delta_h^{-(\tau+3d)} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} e_h \tag{6.12}$$

$$v_h \leq \hat{C}_h \nu^{-3} (aN_h)^{2\tau} \delta_h^{-(\tau+3d+1)} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} e_h \tag{6.13}$$

$$s_h \leq \hat{C}_h \nu^{-1} (aN_h)^{\tau} \delta_h^{-d} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} e_h \tag{6.14}$$

where $\hat{C}_h$ is an explicit constant depending in a polynomial manner on $\|M_h\|_{\rho_h,\tilde{\gamma}_h}$, $\left\|M_h^{-1}\right\|_{\rho_h,\tilde{\gamma}_h}$, $\|\mathcal{N}_h\|_{\rho_h,\tilde{\gamma}_h}$, $\|DK_h\|_{\rho_h,\tilde{\gamma}_h}$, and $\mathcal{T}_h$. Moreover, the non linear estimate (5.46) given in Proposition 45 implies

$$e_{h+1} \leq \tilde{C}_h \nu^{-6} (aN_h)^{4\tau} \delta_h^{-(2\tau+6d)} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} (e_h + e_h^2) \tag{6.15}$$

where $\tilde{C}_h$ is a constant which also depends explicitly on $\|M_h\|_{\rho_h,\tilde{\gamma}_h}$, $\left\|M_h^{-1}\right\|_{\rho_h,\tilde{\gamma}_h}$, $\|\mathcal{N}_h\|_{\rho_h,\tilde{\gamma}_h}$, $\|DK_h\|_{\rho_h,\tilde{\gamma}_h}$, and $\mathcal{T}_h$.

Remark 47. *In the following we will denote C a constant depending on $\nu, \tau, d, \xi, \rho_0, \left|J^{-1}\right|$; and that is a polynomial in $\|M_0\|_{\rho_0,\tilde{\gamma}_0}$, $\left\|M_0^{-1}\right\|_{\rho_0,\tilde{\gamma}_0}$, $\|\mathcal{N}_0\|_{\rho_0,\tilde{\gamma}_0}$, $\|DK_0\|_{\rho_0,\tilde{\gamma}_0}$, and $\mathcal{T}_0$. We will also denote*

$$C_h = \max\left(\hat{C}_h, \tilde{C}_h\right).$$

In Lemma 49 , we give smallness conditions so that $C_h \leq C$ for every $h \geq 0$. Since we are working with expansions near to $(K^{[\leq N_0]}, \mu^{[\leq N_0]})$ it is natural to expect that the quantities $\|M_h\|_{\rho_h,\tilde{\gamma}_h}$, $\left\|M_h^{-1}\right\|_{\rho_h,\tilde{\gamma}_h}$, $\|\mathcal{N}_h\|_{\rho_h,\tilde{\gamma}_h}$, $\|DK_h\|_{\rho_h,\tilde{\gamma}_h}$, and $\mathcal{T}_h$ will be close to $\|M_0\|_{\rho_0,\tilde{\gamma}_0}$, $\left\|M_0^{-1}\right\|_{\rho_0,\tilde{\gamma}_0}$, $\|\mathcal{N}_0\|_{\rho_0,\tilde{\gamma}_0}$, $\|DK_0\|_{\rho_0,\tilde{\gamma}_0}$, and $\mathcal{T}_0$, respectively. For now, we assume that C is large enough, for instance $C > 2C_0$. Here $M_h = M^{[\leq N_h]}$, $\mathcal{N}_h = \mathcal{N}^{[\leq N_h]}$, and $\mathcal{T}_h = \mathcal{T}^{N_h}$ as in (4.8), (4.10), and (5.20).

Considering this uniform constant C on (6.15), and taking N_0 sufficiently large, yields $e_h < 1$ for any $h > 0$, and inequality (6.15) implies

$$e_{h+1} \leq C \nu^{-6} (aN_h)^{4\tau} \delta_h^{-(2\tau+6d)} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} e_h. \tag{6.16}$$

Remark 48. *Due to Remark 47 and the definitions of δ_h, ρ_h, N_h, and $\tilde{\gamma}_h$; the inequality (6.16) can be rewritten as*

$$e_{h+1} \leq C\nu^{-6}(aN_0)^{4\tau}\rho_0^{-(2\tau+6d)}2^{-(4\tau+12d)}\left(2^h\right)^{6\tau+6d}\left(\frac{1}{2^{\tau/\alpha}}\right)^{2^h N_0} e_h$$

or

$$e_{h+1} \leq CDB^h r^{2^h N_0} e_h \tag{6.17}$$

where

$$D = \nu^{-6}(aN_0)^{4\tau}\rho_0^{-(2\tau+6d)}2^{-(4\tau+12d)}, \quad r = 2^{-\tau/\alpha} \quad \text{and} \quad B = 2^{6\tau+6d}.$$

Lemma 49. *Assuming that $2^{3(\tau+3d)+1}CDBr^{N_0} \leq \frac{1}{2}$, $Br^{N_0} < 1$, $N_0^{2\tau}e_0 \ll 1$, and*

$$C\nu^{-3}(aN_0)^{2\tau}\rho_0^{-(\tau+3d+1)}2^{2\tau+6d+2}e_0 \ll 1.$$

Then, for all integers $h \geq 0$ the following properties hold:

$(p1; h)$

$$\|K_h - K_0\|_{\rho_h,\tilde{\gamma}_h} \leq \ell_K N_0^{2\tau} e_0 < \xi$$

$$\sup_{|\varepsilon|\leq\tilde{\gamma}_{h+1}} |\mu_h - \mu_0| \leq \ell_\mu N_0^{\tau} e_0 < \xi$$

with $\ell_K \equiv C\nu^{-3}a^{2\tau}\rho_0^{-(\tau+3d)}2^{2\tau+6d}$ *and* $\ell_\mu \equiv C\nu^{-1}a^{\tau}2^d\rho_0^{-d}$

$(p2; h)$

$$e_h \leq (CD)^h B^{h^2} r^{(2^h-1)N_0} e_0$$

$(p3; h)$

$$C_h \leq C$$

Remark 50. *Note that by (3.12) we have $e_0 \sim \mathcal{O}(N_0^{-(\tau/\alpha)N_0})$, due to the fact that we estimate e_0 in a ball with radius $\tilde{\gamma}_0 \sim \mathcal{O}(N_0^{-\tau/\alpha})$. So the assumptions on the smallness of $N_0 e_0$ are satisfied.*

Proof. Note that $(p1; 0)$, $(p2; 0)$, and $(p3; 0)$ are trivial.

Let us now prove $(p1, H+1)$, $(p2, H+1)$, and $(p3, H+1)$ assuming they are true for $h = 1, 2, ..., H$. Noticing that $2^j \leq 2^{j+1} - 1$, for any $j \geq 0$, and assuming that N_0 is large enough such that

$2^{3(d+\tau)}CDBr^{N_0} \leq \frac{1}{2}$ and $Br^{N_0} < 1$, we have

$$\begin{aligned}
\|K_{H+1} - K_0\|_{\rho_{H+1},\tilde{\gamma}_{H+1}} &= \left\|\Delta^{(N_0,N_1]} + ... + \Delta^{(N_H,N_{H+1}]}\right\|_{\rho_{H+1},\tilde{\gamma}_{H+1}} \\
&\leq \sum_{j=0}^{H} d_j \leq \sum_{j=0}^{H} \hat{C}_j \nu^{-3} (aN_j)^{2\tau} \delta_j^{-(\tau+3d)} r^{N_j} e_j \\
&\leq \sum_{j=0}^{H} C\nu^{-3} (a2^j N_0)^{2\tau} \rho_0^{-(\tau+3d)} 2^{2\tau+6d} 2^{(\tau+3d)j} r^{2^j N_0} e_j \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d)} 2^{2\tau+6d} \sum_{j=0}^{H} 2^{3(d+\tau)j} r^{2^j N_0} \left((CD)^j B^{j^2} r^{(2^j-1)N_0} e_0\right) \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d)} 2^{2\tau+6d} \sum_{j=0}^{H} 2^{3(d+\tau)j} (CD)^j B^{j^2} r^{(2^{j+1}-1)N_0} e_0 \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d)} 2^{2\tau+6d} \sum_{j=0}^{H} 2^{3(d+\tau)j} (CD)^j B^{j^2} r^{2^j N_0} e_0 \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d)} 2^{2\tau+6d} e_0 \sum_{j=0}^{H} \left(2^{3(d+\tau)} CDBr^{N_0}\right)^j \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d)} 2^{2\tau+6d} e_0 \\
&\leq \ell_K N_0^{2\tau} e_0
\end{aligned}$$

Similarly,

$$\begin{aligned}
\sup_{|\varepsilon|\leq\tilde{\gamma}_{H+1}} |\mu_{H+1} - \mu_0| &= \sup_{|\varepsilon|\leq\tilde{\gamma}_{H+1}} \left|\sigma^{(N_0,N_1]} + ... + \sigma^{(N_H,N_{H+1}]}\right| \\
&\leq \sum_{j=0}^{H} s_j \leq \sum_{j=0}^{H} \hat{C}_j \nu^{-1} (aN_j)^{\tau} \delta_j^{-d} r^{N_j} e_j \\
&\leq \sum_{j=0}^{H} C\nu^{-1} (a2^j N_0)^{\tau} \rho_0^{-d} 2^{(j+2)d} r^{2^j N_0} \left((CD)^j B^{j^2} r^{(2^j-1)N_0} e_0\right) \\
&\leq C\nu^{-1} (aN_0)^{\tau} \rho_0^{-d} 2^{2d} \sum_{j=0}^{H} (2^{\tau+d})^j (CD)^j B^{j^2} r^{(2^{j+1}-1)N_0} e_0 \\
&\leq C\nu^{-1} (aN_0)^{\tau} \rho_0^{-d} 2^{2d} \sum_{j=0}^{H} (2^{\tau+d})^j (CD)^j B^{j^2} r^{2^j N_0} e_0 \\
&\leq C\nu^{-1} (aN_0)^{\tau} \rho_0^{-d} 2^{2d} e_0 \sum_{j=0}^{H} \left(2^{\tau+d} CDBr^{N_0}\right)^j \\
&\leq C\nu^{-1} (aN_0)^{\tau} 2^d \rho_0^{-d} e_0 \\
&\leq \ell_\mu N_0^{\tau} e_0.
\end{aligned}$$

Thus, taking N_0 large enough, which makes e_0 small, we get $\ell_K N_0^{2\tau} e_0 < \xi$ and $\ell_\mu N_0^{\tau} e_0 < \xi$.

Since $(p_1; H+1)$ is true, we use the estimate (6.17) given in Remark 48, which is a consequence of the nonlinear estimates given in Lemma 45, that is

$$e_{h+1} = \|\mathcal{E}(K_h + \Delta_h, \mu_h + \sigma_h)\|_{\rho_{h+1},\tilde{\gamma}_{h+1}} \leq CDB^h r^{2^h N_0} e_h \tag{6.18}$$

where D, B, and r are as in Remark 48. This yields,

$$\begin{aligned}
e_{h+1} &\leq CDB^h r^{2^h N_0} e_h \\
&\leq CDB^h r^{2^h N_0} \left((CD)^h B^{h^2} r^{(2^h-1)N_0} e_0 \right) \\
&\leq (CD)^{h+1} B^{h^2+h} r^{(2^{h+1}-1)N_0} e_0 \\
&\leq (CD)^{h+1} B^{(h+1)^2} r^{(2^{h+1}-1)N_0} e_0
\end{aligned}$$

which yields $(p_2, H+1)$. In order to prove $(p_3; H+1)$ note that

$$\|\mathcal{N}_h - \mathcal{N}_0\|_{\rho_h,\tilde{\gamma}_h} \leq \overline{C} \|DK_h - DK_0\|_{\rho_h,\tilde{\gamma}_h} \tag{6.19}$$

$$\|M_h - M_0\|_{\rho_h,\tilde{\gamma}_h} \leq \overline{C} \|DK_h - DK_0\|_{\rho_h,\tilde{\gamma}_h} \tag{6.20}$$

$$\left\|M_h^{-1} - M_0^{-1}\right\|_{\rho_h,\tilde{\gamma}_h} \leq \overline{C} \|DK_h - DK_0\|_{\rho_h,\tilde{\gamma}_h} \tag{6.21}$$

$$|\mathcal{T}_h - \mathcal{T}_0| \leq \overline{C} \|DK_h - DK_0\|_{\rho_h,\tilde{\gamma}_h} \tag{6.22}$$

where $\overline{C}$ is a uniform constant. The above inequalities come from the fact that M_h, $\mathcal{N}_h$, and $\mathcal{T}_h$ are algebraic expressions of DK_h, $Df_{\cdot,\mu_h}$, and $D_\mu f_{\cdot,\mu_h}$; see (4.8), (4.10), (4.9), (5.20). Then,

$$\begin{aligned}
\|DK_{H+1} - DK_0\|_{\rho_{H+1},\tilde{\gamma}_{H+1}} &= \left\| D\Delta^{(N_0,N_1]} + \ldots + D\Delta^{(N_H,N_{H+1}]} \right\|_{\rho_{H+1},\tilde{\gamma}_{H+1}} \\
&\leq \sum_{j=0}^{H} d_j \leq \sum_{j=0}^{H} \hat{C}_j \nu^{-3} (aN_j)^{2\tau} \delta_j^{-(\tau+3d+1)} r^{N_j} e_j \\
&\leq \sum_{j=0}^{H} C\nu^{-3} (a2^j N_0)^{2\tau} \rho_0^{-(\tau+3d+1)} 2^{2\tau+6d+2} 2^{(\tau+3d+1)j} r^{2^j N_0} e_j \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d+1)} 2^{2\tau+6d+2} \sum_{j=0}^{H} 2^{(3d+3\tau+1)j} r^{2^j N_0} \left((CD)^j B^{j^2} r^{(2^j-1)N_0} e_0 \right) \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d+1)} 2^{2\tau+6d+2} \sum_{j=0}^{H} 2^{(3d+3\tau+1)j} (CD)^j B^{j^2} r^{(2^{j+1}-1)N_0} e_0 \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d+1)} 2^{2\tau+6d+2} \sum_{j=0}^{H} 2^{(3d+3\tau+1)j} (CD)^j B^{j^2} r^{2^j N_0} e_0 \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d+1)} 2^{2\tau+6d+2} e_0 \sum_{j=0}^{H} \left(2^{3d+3\tau+1} CDBr^{N_0} \right)^j \\
&\leq C\nu^{-3} (aN_0)^{2\tau} \rho_0^{-(\tau+3d+1)} 2^{2\tau+6d+2} e_0
\end{aligned}$$

where the sum is bounded as in the previous estimates. Taking e_0 small enough, such that $\overline{C}C\nu^{-3}(aN_0)^{2\tau}\rho_0^{-(\tau+3d+1)}2^{2\tau+6d+2}e_0 \ll 1$, we are able to verify $(p3; H+1)$ because C_{H+1} is an algebraic expression of M_H, $\mathcal{N}_H$, and $\mathcal{T}_H$; and taking $C \geq 2C_0$, for example. □

6.2. **Proof of main Lemma 22.** For the proof of the main Lemma we inherit all the notation introduced throughout this section.

Proof. Note that Theorem 20 assures the existence of the Lindstedt series satisfying (6.2). That is, given $K_0 \in \mathcal{A}_\rho$ and $\mu_0 \in \Lambda \subseteq \mathbb{C}$ satisfying $f_{0,\mu_0} \circ K_0 = K_0 \circ T_\omega$ and **HND**, there exists $\rho_0 < \rho$ and power expansions $K_\varepsilon^{[\leq N]}$ and $\mu_\varepsilon^{[\leq N]}$ such that

$$\left\| f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} - K_\varepsilon^{[\leq N]} \circ T_\omega \right\|_{\rho'} \leq C_N |\varepsilon|^{N+1}$$

for any $N \geq 0$. This expansion is unique under the normalization condition (3.3).

If $K_\varepsilon^{[\leq N]}$ and $\mu_\varepsilon^{[\leq N]}$ satisfy hypothesis **HTP1** and **HTP2** then, we can choose N_0 such that $K^{[\leq N_0]}$ and $\mu^{[\leq N_0]}$ satisfy the hypothesis of Lemmas 37 and 39. Also, N_0 needs to be large enough such that $2^{3(\tau+3d)+1}CDBr^{N_0} \leq \frac{1}{2}$, $Br^{N_0} < 1$, $\ell_K N_0^{2\tau} e_0 < \xi$, $\ell_\mu N_0^\tau e_0 < \xi$ and

$$\overline{C}C\nu^{-3}(aN_0)^{2\tau}\rho_0^{-(\tau+3d+1)}2^{2\tau+6d+2}e_0 \ll 1,$$

then Lemma 49 can be applied and this allows us to iterate the quasi Newton method described in Algorithm 33. That is, we can construct the unique formal power series as follows

$$K_\varepsilon^{[\leq N_0]} + \Delta_\varepsilon^{(N_0,2N_0]} + \Delta_\varepsilon^{(2N_0,2^2N_0]} + \cdots + \Delta_\varepsilon^{(2^hN_0,2^{h+1}N_0]} + \cdots$$
$$\mu_\varepsilon^{[\leq N_0]} + \mu_\varepsilon^{(N_0,2N_0]} + \mu_\varepsilon^{(2N_0,2^2N_0]} + \cdots + \mu_\varepsilon^{(2^hN_0,2^{h+1}N_0]} + \cdots$$

Note that by definition of $\tilde{\gamma}_h$ we will have $\tilde{\gamma}_h = r^h\tilde{\gamma}_0$, where $r = 2^{-\tau/\alpha}$ and $\tilde{\gamma}_0 = 2^{-1/\alpha}\nu^{1/\alpha}(aN_0)^{-\tau/\alpha}$, see (6.4). Before giving the detailed computations, note that $\tilde{\gamma}_h \sim (2^hN_0)^{-\tau/\alpha}$ and if $n \in \left(2^hN_0, 2^{h+1}N_0\right] \cap \mathbb{N}$ then

$$(\tilde{\gamma}_h)^{-n} \sim (2^hN_0)^{C(\tau/\alpha)2^hN_0} \sim n^{C(\tau/\alpha)n}.$$

Using this together with Cauchy estimates is expected to yield the Gevrey estimates. More precisely, if $n \in \left(2^hN_0, 2^{h+1}N_0\right] \cap \mathbb{N}$, using Cauchy estimates, (6.12), and $(p2; h)$ we have

$$\begin{aligned}
\|K_n\|_{\frac{\rho_0}{2}} &\leq (\tilde{\gamma}_{h+1})^{-n} \|\Delta_h\|_{\frac{\rho_0}{2},\tilde{\gamma}_{h+1}} \\
&\leq (\tilde{\gamma}_{h+1})^{-n} \|\Delta_h\|_{\rho_{h+1},\tilde{\gamma}_{h+1}} \\
&\leq (r^{h+1}\tilde{\gamma}_0)^{-n} d_h \\
&\leq (r^{h+1}\tilde{\gamma}_0)^{-2^{h+1}N_0} \hat{C}_h \nu^{-3}(aN_h)^{2\tau}\delta_h^{-(\tau+3d)} r^{N_h} e_h \\
&\leq (r^{h+1}\tilde{\gamma}_0)^{-2^{h+1}N_0} C\nu^{-3}\left(a2^hN_0\right)^{2\tau} \rho_0^{-(\tau+3d)}2^{(2\tau+6d)}2^{(\tau+3d)h}r^{2^hN_0}(CD)^hB^{h^2}r^{(2^h-1)N_0}e_0 \\
&\leq C\nu^{-3}\rho_0^{-(\tau+3d)}2^{(2\tau+6d)}(aN_0)^{2\tau}e_0(2^{3\tau+3d}CD)^hB^{h^2}(\tilde{\gamma}_0)^{-2^{h+1}N_0}r^{(-(h+1)2^{h+1}+2^{h+1}-1)N_0} \\
&\leq C\nu^{-3}\rho_0^{-(\tau+3d)}2^{(2\tau+6d)}(aN_0)^{2\tau}e_0(2^{3\tau+3d}CD)^hB^{h^2}(2^{1/\alpha}\nu^{-1/\alpha}(aN_0)^{\tau/\alpha})^{2^{h+1}N_0}r^{-(h2^{h+1}+1)N_0} \\
&\leq \hat{L}\left(2^{3\tau+3d}CDB2^{2/\alpha}\nu^{-2/\alpha}a^{2\tau/\alpha}\right)^{2^hN_0}(N_0^{2\tau/\alpha})^{2^hN_0}(2^{\tau/\alpha})^{(h2^{h+1}+1)N_0} \\
&\leq \hat{L}2^{(\tau/\alpha)N_0}F^{2^hN_0}(N_0^{2\tau/\alpha})^{2^hN_0}(2^{2\tau/\alpha})^{h2^hN_0} \\
&\leq LF^{2^hN_0}(2^hN_0)^{(2\tau/\alpha)2^hN_0} \\
&\leq LF^nn^{(2\tau/\alpha)n}
\end{aligned}$$

where $\hat{L} = C\nu^{-3}\rho_0^{-(\tau+3d)}2^{2\tau+6d}(aN_0)^{2\tau}e_0$, $F = 2^{3\tau+3d+2/\alpha}CDB\nu^{-2/\alpha}a^{2\tau/\alpha}$, and $L = \hat{L}(2^{\tau/\alpha})^{N_0}$. The estimates for μ_n are obtained in a similar way. □

6.3. **Proof of Theorem 23.**

Proof. Inheriting the notation from Lema 49, consider N_0 sufficiently large such that the a-posteriori theorem, Theorem 14 in [CCdlL17], can be applied. That is, N_0 such that

$$\sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_0} \left\|E_\varepsilon^{N_0}\right\|_\rho \le \hat{C}\left(\nu\tilde{\nu}(\lambda;\omega,\tau)\right)^2\delta^{-4(\tau+\delta)}. \tag{6.23}$$

where $\tilde{\nu}(\lambda;\omega,\tau)$ is defined in (3.15). Then, following the discussion in Section (3.3) and applying the a-posteriori theorem, Theorem 14 in [CCdlL17], one obtains

$$\sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_{h+2}} \left\|K_\varepsilon^{[\le 2^hN_0]} - K_\varepsilon\right\|_{\rho_0-\delta} \le \hat{C}\nu^{-1}\tilde{\nu}(\lambda;\omega,\tau)^{-1}\delta^{-2(\tau+d)} \sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_{h+2}} \left\|E_\varepsilon^{2^hN_0}\right\|_{\rho_0}$$

where $\mathcal{G}$ is defined in (3.16).

Now, considering $n \in (2^hN_0, 2^{h+1}N_0] \cap \mathbb{N}$ one has

$$\begin{aligned}
\sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_{h+2}} \left\|K_\varepsilon^{[\le n]} - K_\varepsilon\right\|_{\rho_0-\delta} &\le \sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_{h+2}} \left\|K_\varepsilon^{[\le 2^{h+1}N_0]} - \Delta_\varepsilon^{(n,2^{h+1}N_0]} - K_\varepsilon\right\|_{\rho_0-\delta} \\
&\le \sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_{h+2}} \left\|K_\varepsilon^{[\le 2^{h+1}N_0]} - K_\varepsilon\right\|_{\rho_0-\delta} + \sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_{h+2}} \left\|\Delta^{(n,2^{h+1}N_0]}\right\|_{\rho_0-\delta} \\
&\le \hat{C}\nu^{-1}\tilde{\nu}^{-1}\delta^{-2(\tau+d)} \sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_{h+2}} \left\|E_\varepsilon^{2^hN_0}\right\|_{\rho_0} + \sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde{\gamma}_{h+2}} \left\|\Delta_\varepsilon^{(n,2^{h+1}N_0]}\right\|_{\rho_0-\delta} \\
&\le \hat{C}\nu^{-1}\tilde{\nu}^{-1}\delta^{-2(\tau+d)} \left\|E_\varepsilon^{2^hN_0}\right\|_{\rho,\tilde{\gamma}_{h+2}} + \left\|\Delta_\varepsilon^{(n,2^{h+1}N_0]}\right\|_{\rho_0-\delta,\tilde{\gamma}_{h+2}} \\
&\le \hat{C}\nu^{-1}\tilde{\nu}^{-1}\delta^{-2(\tau+d)} \left\|E_\varepsilon^{2^hN_0}\right\|_{\rho_0,\tilde{\gamma}_h} + \frac{r^{n+1}}{1-r}\left\|\Delta_\varepsilon^{(2^hN_0,2^{h+1}N_0]}\right\|_{\rho_0-\delta,\tilde{\gamma}_{h+1}} \\
&\le \hat{C}\nu^{-1}\tilde{\nu}^{-1}\delta^{-2(\tau+d)}e_h + r^{n+1}d_h \\
&\le \hat{C}\nu^{-1}\tilde{\nu}^{-1}\delta^{-2(\tau+d)}e_h + r^{n+1}C\nu^{-3}(aN_h)^{2\tau}\delta_h^{-(\tau+3d)}r^{N_h}e_h \\
&\le \left(U + C\nu^{-3}(aN_0)^{2\tau}\rho_0^{-(\tau+3d)}2^{2\tau+6d}2^{h(3\tau+3d)}r^{n+1}r^{2^hN_0}\right)(CD)^hB^{h^2}r^{(2^h-1)N_0}e_0 \\
&\le \left(U + V2^{h(3\tau+3d)}r^{n+1}r^{2^hN_0}\right)(CD)^hB^{h^2}r^{(2^h-1)N_0}e_0
\end{aligned}$$

where $U = \hat{C}\nu^{-1}\tilde{\nu}^{-1}\delta^{-2(\tau+d)}$ and $V = C\nu^{-3}(aN_0)^{2\tau}\rho_0^{-(\tau+3d)}2^{2\tau+6d}$ □

Appendix A. The case of the dissipative standard map of Theorem 18

A.1. **Verifying trigonometric polynomial hypothesis for the dissipative standard map.**

Consider the dissipative standard map $f_{\varepsilon,\mu_\varepsilon} : \mathbb{T}\times\mathbb{R} \to \mathbb{T}\times\mathbb{R}$ given by

$$f_{\varepsilon,\mu_\varepsilon}(x,y) = \left(x + \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V(x), \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V(x)\right). \tag{A.1}$$

Where $V(x)$ is a trigonometric polynomial. In this section we verify that maps like (A.1) satisfy **HTP1** and **HTP2** of Lemma 22. For the sake of simplicity in the exposition we do it for the case $\lambda(\varepsilon) = 1 - \varepsilon^3$. The general case for $\alpha \in \mathbb{N}$ is done by very similar computations, fixing the value of $\alpha = 3$ allows an easy analysis of the Lindstedt series.

Note that one has $f_{\varepsilon,\mu}^*\Omega = \lambda(\varepsilon)\Omega$ for the symplectic form $\Omega_{(x,y)} = dx \wedge dy$, so it is conformally symplectic. One can write the map as

$$\begin{aligned} x_{n+1} &= x_n + y_{n+1} \\ y_{n+1} &= \lambda(\varepsilon) y_n + \mu_\varepsilon - \varepsilon V(x_n) \end{aligned}$$

equivalently

$$x_{n+1} - (1+\lambda(\varepsilon))x_n + \lambda(\varepsilon)x_{n-1} - \mu_\varepsilon + \varepsilon V(x_n) = 0. \tag{A.2}$$

Considering a parametric representation of the variable $x_n \in \mathbb{T}$ as $x_n = \theta_n + u_\varepsilon(\theta_n)$, $\theta_n \in \mathbb{T}$; where $u_\varepsilon : \mathbb{T} \to \mathbb{R}$ is a 1-periodic function and assuming that θ_n varies linearly, i.e., $\theta_{n+1} = \theta_n + \omega$, then, (A.2) becomes

$$u_\varepsilon(\theta+\omega) - (1+\lambda(\varepsilon))u_\varepsilon(\theta) + \lambda(\varepsilon)u_\varepsilon(\theta-\omega) + (1-\lambda(\varepsilon))\omega - \mu_\varepsilon + \varepsilon V(\theta + u_\varepsilon(\theta)) = 0 \tag{A.3}$$

If u_ε satisfies (A.3) it is easy to check that $K_\varepsilon : \mathbb{T} \to \mathbb{T} \times \mathbb{R}$, given by

$$K_\varepsilon(\theta) = \begin{pmatrix} \theta + u_\varepsilon(\theta) \\ \omega + u_\varepsilon(\theta) - u_\varepsilon(\theta-\omega) \end{pmatrix},$$

satisfies $f_{\varepsilon,\mu_\varepsilon} \circ K_\varepsilon(\theta) = K(\theta+\omega)$. Therefore, the problem of finding Lindstedt series for quasiperiodic orbits for the map $f_{\varepsilon,\mu_\varepsilon}$ is equivalent to find asymptotic power series to a solution, $(u_\varepsilon, \mu_\varepsilon)$, of (A.3). Using $\lambda(\varepsilon) = 1 - \varepsilon^3$, equation (A.3) becomes

$$u_\varepsilon(\theta+\omega) - (2-\varepsilon^3)u_\varepsilon(\theta) + (1-\varepsilon^3)u_\varepsilon(\theta-\omega) + \varepsilon^3\omega - \mu_\varepsilon + \varepsilon V(\theta + u_\varepsilon(\theta)) = 0. \tag{A.4}$$

Introducing the operator

$$L_\omega u(\theta) = u(\theta+\omega) - 2u(\theta) + u(\theta-\omega),$$

and expanding in power series on ε, i.e., $u_\varepsilon(\theta) = \sum_{n=0}^\infty u_n(\theta)\varepsilon^n$ and $\mu_\varepsilon = \sum_{n=0}^\infty \mu_n \varepsilon^n$ equation (A.4) becomes

$$\begin{aligned} &\sum_{k=0}^{2} \left(L_\omega u_k(\theta) - \mu_k\right)\varepsilon^k - \left(L_\omega u_3(\theta) - \mu_3 + u_0(\theta) - u_0(\theta-\omega) - \omega\right)\varepsilon^3 \\ &\qquad + \sum_{k=4}^{\infty} \left(L_\omega u_k(\theta) - \mu_k + u_{k-3}(\theta) - u_{k-3}(\theta-\omega)\right)\varepsilon^k = -\sum_{k=1}^{\infty} S_{k-1}(\theta)\varepsilon^k \end{aligned} \tag{A.5}$$

Remark 51. *When $V(\theta)$ is a trigonometric polynomial, the coefficients S_n can be computed as follows. Note that $V_k(\theta) = \hat{f}_k e^{2\pi i k\theta}$ satisfies the relation*

$$\frac{d}{d\varepsilon} V_k(\theta + u_\varepsilon(\theta)) = 2\pi i k \frac{d}{d\varepsilon} u_\varepsilon(\theta) V_k(\theta + u_\varepsilon(\theta)). \tag{A.6}$$

Thus, considering

$$V_k(\theta + u_\varepsilon(\theta)) = \sum_{n=0}^{\infty} S_n^k(\theta)\varepsilon^n$$

and (A.6) the coefficients S_n^k satisfy the following relation

$$(n+1)S_{n+1}^k = \sum_{\ell=0}^{n} 2\pi i k(\ell+1)u_{\ell+1}S_{n-\ell}^k, \tag{A.7}$$

and $S_0^k(\theta) = \hat{f}_k e^{2\pi i k\theta}$. Furthermore, if $V(\theta) = \sum_{|k|\le a} \hat{f}_k e^{2\pi k\theta} = \sum_{|k|\le a} V_k(\theta)$ is a trigonometric polynomial of degree a, considering

$$V(\theta + u_\varepsilon(\theta)) = \sum_{n=0}^{\infty} S_n(\theta)\varepsilon^n,$$

the coefficients $S_n(\theta)$ are given by

$$S_n(\theta) = \sum_{|k|\le a} S_n^k(\theta)$$

where S_n^k is given by (A.7).

Remark 52. *Note that if η is a trigonometric polynomial and φ is a solution of the equation $L_\omega\varphi = \eta$ then, φ is a trigonometric polynomial of the same degree as η. This is due to the fact that the Fourier coefficients of φ satisfy $\hat{\varphi}_k = \frac{1}{2(\cos(2\pi k\cdot\omega)-1)}\hat{\eta}_k$. Note that the equation $L_\omega\varphi = \eta$ has a solution if $\int_{\mathbb{T}}\eta(\theta)d\theta = 0$, and this solution is unique if we impose the normalization $\int_{\mathbb{T}}\varphi(\theta)d\theta = 0$.*

Proposition 53. *If $V(\theta)$, in (A.1), is a trigonometric polynomial of degree a, then $u_n(\theta)$ is a trigonometric polynomial of degree an. Furthermore, $S_{n-1}(\theta)$ is a trigonometric polynomial of degree an.*

Proof. Equating the terms of same order in equation (A.5) one gets that for order zero $\mu_0 = 0$ and $u_0(\theta) \equiv 0$. For order 1 we have,

$$L_\omega u_1(\theta) - \mu_1 = -S_0(\theta).$$

So, taking $\mu_1 = 0$, u_1 becomes a trigonometric polynomial of degree a, because $S_0(\theta) = V(\theta)$. Now, for order 2 we have

$$L_\omega u_2(\theta) - \mu_2 = -S_1(\theta),$$

if $\mu_2 = 0$ the right hand side is $S_1(\theta) = \sum_{|k|\le a} S_1^k(\theta) = 2\pi i u_1(\theta)\sum_{|k|\le a} kS_0^k(\theta)$ which is a trigonometric polynomial of degree $2a$, thus u_2 is a trig polynomial of degree $2a$. For order three we have

$$L_\omega u_3(\theta) - \mu_3 + \omega = -S_2(\theta),$$

here we take $\mu_3 = \omega$ and u_3 is a trig polynomial of degree $3a$ because

$$S_2(\theta) = \sum_{|k|\le a} S_2^k(\theta) = \pi i u_1(\theta)\sum_{|k|\le a} kS_1^k(\theta) + 2\pi i u_2(\theta)\sum_{|k|\le a} kS_0^k(\theta)$$

is of degree $3a$; then $u_3(\theta)$ is of degree $3a$. Finally, for $n \ge 4$, assume the claim is valid for any $m < n$ then, the equation of order n is

$$L_\omega u_n(\theta) = \mu_n - u_{n-3}(\theta) + u_{n-3}(\theta-\omega) - S_{n-1}(\theta).$$

So, taking $\mu_n = \int_{\mathbb{T}} S_{n-1}(\theta)d\theta$, u_n can be found and has degree an since, $S_{n-1} = \sum_{|k|\le n} S_{n-1}^k$ and each S_{n-1}^k has degree an due to (A.7). Note u_{n-3} has degree $(n-3)a$. ☐

Corollary 54. *If $V(\theta)$, in (A.1), is a trigonometric polynomial of degree a, then for any fixed ε the sum $\displaystyle\sum_{n=0}^{N} u_n(\theta)\varepsilon^n$ is a trig polynomial of degree aN in θ.*

Note that in this case

$$K_\varepsilon^{[\le N]}(\theta) = \begin{pmatrix} \theta + \sum_{n=0}^N u_n(\theta)\varepsilon^n \\ \omega + \sum_{n=0}^N (u_n(\theta) - u_n(\theta-\omega))\varepsilon^n \end{pmatrix}, \tag{A.8}$$

and using equation (A.5) we have

$$E_\varepsilon^N(\theta) := f_{\varepsilon,\mu^{[\le N]}} \circ K_\varepsilon^{[\le N]}(\theta) - K_\varepsilon^{[\le N]}(\theta+\omega) = \sum_{n=N+1}^{\infty} \begin{pmatrix} S_{n-1}(\theta) \\ S_{n-1}(\theta) \end{pmatrix} \varepsilon^n$$

and therefore, for any fixed ε, $E_\varepsilon^{(N,2N)}(\theta)$ is a trigonometric polynomial of degree $2aN$. Moreover, in this case the matrix $M_\varepsilon^{[\leq N]}(\theta) = \left[DK_\varepsilon^{[\leq N]}(\theta)|J^{-1} \circ K_\varepsilon^{[\leq N]}(\theta)DK_\varepsilon^{[\leq N]}(\theta)\mathcal{N}_\varepsilon^{[\leq N]}(\theta)\right]$ is given by

$$M_\varepsilon^{[\leq N]}(\theta) = \begin{bmatrix} 1 + \sum_{k=0}^N u_k'(\theta)\varepsilon^k & \mathcal{N}_\varepsilon^{[\leq N]}(\theta) \sum_{k=0}^N (u_k'(\theta-\omega) - u_k'(\theta))\varepsilon^k \\ \sum_{k=0}^N (u_k'(\theta) - u_k'(\theta-\omega))\varepsilon^k & \mathcal{N}_\varepsilon^{[\leq N]}(\theta)(1 + \sum_{k=0}^N u_k'(\theta)\varepsilon^k) \end{bmatrix}$$

where $\mathcal{N}_\varepsilon^{[\leq N]}(\theta) = \left((1 + \sum_{k=0}^N u_k'(\theta)\varepsilon^k)^2 + (\sum_{k=0}^N (u_k'(\theta) - u_k'(\theta-\omega))\varepsilon^k)^2\right)^{-1}$. So,

$$\left(M_\varepsilon^{[\leq N]} \circ T_\omega\right)^{-1} = \begin{bmatrix} \left(\mathcal{N}_\varepsilon^{[\leq N]} \circ T_\omega\right)(1 + \sum_{k=0}^N u_k'(\theta+\omega)\varepsilon^k) & \left(\mathcal{N}_\varepsilon^{[\leq N]} \circ T_\omega\right) \sum_{k=0}^N (u_k'(\theta+\omega) - u_k'(\theta))\varepsilon^k \\ \sum_{k=0}^N (u_k'(\theta) - u_k'(\theta+\omega))\varepsilon^k & 1 + \sum_{k=0}^N u_k'(\theta)\varepsilon^k \end{bmatrix}$$

which implies that $\tilde{E}_{\varepsilon,2}^{(N,2N)}$ is a trigonometric polynomial of degree $3aN$. Remember that $\tilde{E}_{\varepsilon,2}^{(N,2N)}$ is the second row of the vector $\tilde{E}_\varepsilon^{(N,2N)} = \left(M_\varepsilon^{[\leq N]} \circ T_\omega\right)^{-1} E_\varepsilon^{(N,2N)}$. Note that $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Furthermore, we have $D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}}(x,y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, then the second row, $\tilde{A}_{\varepsilon,2}^N$, of the vector $\tilde{A}_\varepsilon^N = \left(M_\varepsilon^{[\leq N]} \circ T_\omega\right)^{-1} D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}$ is a trigonometric polynomial of degree aN.

The following proposition summarizes the computations presented above and assures that hypothesis **HTP1** and **HTP2** of the main Lemma 22 are satisfied for the dissipative standard map.

Proposition 55. *For any $N \in \mathbb{N}$, if $V(\theta)$ in (A.1) is a trigonometric polynomial of degree a, then $\tilde{E}_{\varepsilon,2}^{(N,2N)}$ is a trigonometric polynomial of degree $3aN$, $\tilde{A}_{\varepsilon,2}^N$ is a trig polynomial of degree aN, and*

$$\begin{aligned} \tilde{E}_{\Omega,\varepsilon}^N(\theta) \equiv\ & DK_\varepsilon^{[\leq N]}(\theta+\omega)^\top J \circ K_\varepsilon^{[\leq N]}(\theta+\omega) DK_\varepsilon^{[\leq N]}(\theta+\omega) \\ & - D(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta))^\top J \circ (f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) D(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) \end{aligned} \tag{A.9}$$

is a trigonometric polynomial of degree $2aN$.

Proof. It is only left to prove the last claim. Note that $\tilde{E}_{\Omega,\varepsilon}^N(\theta)$ is the expression in coordinates of $(K_\varepsilon^{[\leq N]} \circ T_\omega)^*\Omega - (f_{\varepsilon,\mu^{[\leq N]}} \circ K^{[\leq N]})^*\Omega$. Now, using the fact that $f_{\varepsilon,\mu}$ is conformally symplectic we have $(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]})^*\Omega = K_\varepsilon^{[\leq N]^*} f^*_{\varepsilon,\mu_\varepsilon^{[\leq N]}}\Omega = \lambda(\varepsilon) K_\varepsilon^{[\leq N]^*}\Omega$, which means that, in coordinates

$$\begin{aligned} \tilde{E}_{\Omega,\varepsilon}^N(\theta,\varepsilon) = & DK_\varepsilon^{[\leq N]}(\theta+\omega)^\top J \circ K_\varepsilon^{[\leq N]}(\theta+\omega) DK_\varepsilon^{[\leq N]}(\theta+\omega) \\ & - \lambda(\varepsilon) DK_\varepsilon^{[\leq N]}(\theta)^\top J \circ K_\varepsilon^{[\leq N]}(\theta) DK_\varepsilon^{[\leq N]}(\theta) \end{aligned} \tag{A.10}$$

which is a polynomial of degree $2aN$ due to the fact that J is a constant matrix and

$$DK_\varepsilon^{[\leq N]}(\theta) = \begin{pmatrix} 1 + \sum_{n=0}^N u_n'(\theta)\varepsilon^n \\ \sum_{n=0}^N (u_n'(\theta) - u_n'(\theta-\omega))\varepsilon^n \end{pmatrix}$$

is a trigonometric polynomial of degree aN. ☐

A.2. **Uniqueness.** Note that for $\varepsilon = 0$, $M_0 = I$. Also note that the coefficients of the expansion (A.8) are given by

$$K_n(\theta) = \begin{pmatrix} u_n(\theta) \\ u_n(\theta) - u_n(\theta-\omega) \end{pmatrix} \qquad \text{for } n \geq 1.$$

Therefore, the normalization condition

$$\int_{\mathbb{T}} \left[M_0^{-1} K_n(\theta)\right]_1 d\theta = 0$$

in this case has the form

$$\int_{\mathbb{T}} u_n(\theta) d\theta = 0,$$

which is satisfied by the construction of the $u_n's$. Thus, the expansion given in (A.8) is the only one which satisfies the normalization condition.

Acknowledgements

The authors would like to thank T. M-Seara, I. Baldomá, and V. Naudot for many suggestions and discussions.

References

[Bae95] Claude Baesens. Gevrey series and dynamic bifurcations for analytic slow-fast mappings. *Nonlinearity*, 8(2):179–201, 1995.

[BC] Adrian P Bustamante and Renato C Calleja. Corrigendum and addendum to *Computation of domains of analyticity for the dissipative standard map in the limit of small dissipation*. In progress.

[BC19] Adrian P Bustamante and Renato C Calleja. Computation of domains of analyticity for the dissipative standard map in the limit of small dissipation. *Physica D: Nonlinear Phenomena*, 395:15–23, 2019.

[Ben88] Alain Bensoussan. *Perturbation methods in optimal control*. Wiley/Gauthier-Villars Series in Modern Applied Mathematics. John Wiley & Sons Ltd., Chichester, 1988. Translated from the French by C. Tomson.

[BFM17] Inmaculada Baldomá, Ernest Fontich, and Pau Martín. Gevrey estimates for one dimensional parabolic invariant manifolds of non-hyperbolic fixed points. *Discrete Contin. Dyn. Syst.*, 37(8):4159–4190, 2017.

[BH08] I. Baldomá and A. Haro. One dimensional invariant manifolds of Gevrey type in real-analytic maps. *Discrete Contin. Dyn. Syst. Ser. B*, 10(2-3):295–322, 2008.

[BHS96] H. W. Broer, G. B. Huitema, and M. B. Sevryuk. *Quasi-Periodic Motions in Families of Dynamical Systems*. Order Amidst Chaos. Springer-Verlag, Berlin, 1996.

[Car95] Henri Cartan. *Elementary theory of analytic functions of one or several complex variables*. Dover Publications, Inc., New York, 1995. Translated from the French, Reprint of the 1973 edition.

[CCdlL13] Renato C Calleja, Alessandra Celletti, and Rafael de la Llave. A kam theory for conformally symplectic systems: efficient algorithms and their validation. *Journal of Differential Equations*, 255(5):978–1049, 2013.

[CCdlL17] Renato C Calleja, Alessandra Celletti, and Rafael de la Llave. Domains of analyticity and lindstedt expansions of kam tori in some dissipative perturbations of hamiltonian systems. *Nonlinearity*, 30(8):3151, 2017.

[CD91] Mireille Canalis-Durand. Formal expansion of van der Pol equation canard solutions are Gevrey. In *Dynamic bifurcations (Luminy, 1990)*, volume 1493 of *Lecture Notes in Math.*, pages 29–39. Springer, Berlin, 1991.

[CDRSS00] M. Canalis-Durand, J. P. Ramis, R. Schäfke, and Y. Sibuya. Gevrey solutions of singularly perturbed differential equations. *J. Reine Angew. Math.*, 518:95–129, 2000.

[Cel91] Alessandra Celletti. Stability of satellites in spin-orbit resonances and capture probabilities. In *Predictability, stability, and chaos in N-body dynamical systems (Cortina d'Ampezzo, 1990)*, volume 272 of *NATO Adv. Sci. Inst. Ser. B Phys.*, pages 337–344. Plenum, New York, 1991.

[Cel13] Alessandra Celletti. KAM theory and its applications: from conservative to dissipative systems. In *European Congress of Mathematics*, pages 243–257. Eur. Math. Soc., Zürich, 2013.

[CGGG07] O. Costin, G. Gallavotti, G. Gentile, and A. Giuliani. Borel summability and Lindstedt series. *Comm. Math. Phys.*, 269(1):175–193, 2007.

[Cos09] Ovidiu Costin. *Asymptotics and Borel summability*, volume 141 of *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics*. CRC Press, Boca Raton, FL, 2009.

[DFIZ16] Andrea Davini, Albert Fathi, Renato Iturriaga, and Maxime Zavidovique. Convergence of the solutions of the discounted Hamilton-Jacobi equation: convergence of the discounted solutions. *Invent. Math.*, 206(1):29–55, 2016.

[FT89] C. Foias and R. Temam. Gevrey class regularity for the solutions of the Navier-Stokes equations. *J. Funct. Anal.*, 87(2):359–369, 1989.

[ISM11] Renato Iturriaga and Héctor Sánchez-Morgado. Limit of the infinite horizon discounted Hamilton-Jacobi equation. *Discrete Contin. Dyn. Syst. Ser. B*, 15(3):623–635, 2011.

[Knu98] Donald E. Knuth. *The art of computer programming. Vol. 2.* Addison-Wesley, Reading, MA, 1998. Seminumerical algorithms, Third edition [of MR0286318].

[Lin92] C. H. Lin. Gevrey character of formal solutions of a nonlinear system with a parameter. In *Ordinary and delay differential equations (Edinburg, TX, 1991)*, volume 272 of *Pitman Res. Notes Math. Ser.*, pages 129–133. Longman Sci. Tech., Harlow, 1992.

[MHER95] Edward S. Meadows, Michael A. Henson, John W. Eaton, and James B. Rawlings. Receding horizon control and discontinuous state feedback stabilization. *Internat. J. Control*, 62(5):1217–1229, 1995.

[MNF87] Andrea Milani, Anna Maria Nobili, and Paolo Farinella. *Non-gravitational perturbations and satellite geodesy.* Adam Hilger (IOP Publishing Ltd.), Bristol, 1987.

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

[Mos73] J Moser. Stable and random motions in dynamical systems, volume 77 of annals of mathematics studies, 1973.

[Pop00] G. Popov. Invariant tori, effective stability, and quasimodes with exponentially small error terms. I. Birkhoff normal forms. *Ann. Henri Poincaré*, 1(2):223–248, 2000.

[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*, pages 598–624. Springer, 1975.

[Sau92] David Sauzin. Caractère Gevrey des solutions formelles d'un problème de moyennisation. *C. R. Acad. Sci. Paris Sér. I Math.*, 315(9):991–995, 1992.

[Sev99] M. B. Sevryuk. The lack-of-parameters problem in the KAM theory revisited. In *Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995)*, pages 568–572. Kluwer Acad. Publ., Dordrecht, 1999.

[SZ65] S. Saks and A Zygmund. *Analytic Functions.* (Enlarged. Translated by E J Scott. Monografie Matematyczne, Tom 28), 1965. 2nd edn.

A.P. Bustamante, School of Mathematics, Georgia Institute of Technology
E-mail address,`apb7@math.gatech.edu`

R. de la Llave, School of Mathematics, Georgia Institute of Technology
E-mail address,`rafael.delallave@math.gatech.edu`