DOMAINS OF ANALYTICITY AND GEVREY ESTIMATES OF TORI IN WEAKLY DISSIPATIVE SYSTEMS

A Dissertation
Presented to
The Academic Faculty

By

Adrian Perez Bustamante

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology

December 2021

DOMAINS OF ANALYTICITY AND GEVREY ESTIMATES OF TORI IN WEAKLY DISSIPATIVE SYSTEMS

Thesis committee:

Prof. Rafael de la Llave, Advisor
School of Mathematics
Georgia Institute of Technology

Prof. Chongchun Zeng
School of Mathematics
Georgia Institute of Technology

Prof. Molei Tao
School of Mathematics
Georgia Institute of Technology

Prof. Alessandra Celletti
Department of Mathematics
University of Rome Tor Vergata

Prof. Alex Blumenthal
School of Mathematics
Georgia Institute of Technology

Date approved: August 27, 2021

To my parents

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude to my thesis advisor, Prof. Rafael de la Llave, for all his priceless advice, support, encouragement, and infinite patience during these years. I will always appreciate all his valuable teachings and his endless kindness. I am very fortunate to have him as an advisor.

I want to thank Prof. Alessandra Celletti, Prof. Molei Tao, Prof. Chongchun Zeng, and Prof. Alex Blumenthal for serving as members of my thesis committe. I would also like to thank Prof. Renato Calleja for his friendship and support.

I would also like to thank all my friends at Georgia Tech for all the good moments that we spent together: Bounghun Bock, Renyi Chen, Tim Duff, Surena Hozoori, Bhanu Kumar, Dario Mena, Hyun Ki Min, Haodong Sun, Jiaqi Yang, Yian Yao, Dantong Zhu, and, Haiyu Zou. Special thanks to Jose Acevedo, Hassan Attarchi, Jaemin Park, and Tom Rodewald for being such a good friends.

Last but not least, I thank my parents, Federico and Julia, for all their unconditional support during all these years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In the first Chapter we study a family of dissipative standard maps of the cylinder for which the dissipation is a function of a small complex parameter of perturbation, ε. We compute perturbative expansions formally in ε and use them to estimate the shape of the domains of analyticity of invariant circles as functions of ε. We also give evidence that the functions might belong to a Gevrey class. The numerical computations we perform support conjectures on the shape of the domains of analyticity.

In the second Chapter, we consider a singular perturbation for a family of analytic symplectic maps of the annulus possessing a KAM torus (this family contains the dissipative standard map studied in Chapter 1). The perturbation introduces dissipation and contains an adjustable parameter. We prove that the asymptotic expansions for the quasiperiodic solutions and the adjustable parameter satisfy Gevrey estimates (proving one of the conjectures given in Chapter 1). To prove this result we introduce a novel method that might be of interest beyond the problem considered here.

Chapter 1 is based in joint work with Renato Calleja, see [1, 2]. Chapter 2 is based in joint work with Rafael de la Llave, see [3].

CHAPTER 1

COMPUTATION OF DOMAINS OF ANALYTICITY FOR THE DISSIPATIVE STANDARD MAP IN THE LIMIT OF SMALL DISSIPATION

1.1 Introduction

We study the limit of small dissipation/expansion of a family of conformally symplectic standard maps. Since conformally symplectic systems include Hamiltonian systems when one adds a small friction term that is proportional to the velocity, then this kind of systems are relevant in applications [4]. In particular, we approximate the shape of domains of analyticity of invariant circles of a family of conformally symplectic standard maps of the cylinder, $\mathcal{M} = \mathbb{S}^1 \times \mathbb{R}$, depending on a small parameter, ε, that vanishes as the conformal factor tends to one.

The present work addresses some rigorous results and conjectures in [5] from a numerical point of view. For instance, it is remarkable that invariant circles which are attractors/repellors in the dissipative/expanding case [6], converge in the limit of small dissipation to invariant circles in the symplectic case. It was noted in [5] that the small divisors depend on the complex parameter ε and give rise to regions where the functions parameterizing the circles cannot be analytic with respect to ε but miss by very little. A conjecture in [5] states that the tori are analytic in a domain in the complex ε plane that is obtained by taking from a ball centered at zero, a sequence of small balls with centers along smooth curves passing through the origin. The radii of the excluded balls decreases faster that any power of the distance of the centers of the balls to the origin, see Figure 1.1. In fact, it was rigorously proved in [5] that this domain is a lower bound. The main objective of the present work is to illustrate the results in one example, provide numerical evidence and indications of new results. Our computations indicate that there are singularities which cluster around several points at which one does not expect the functions to be analytic.

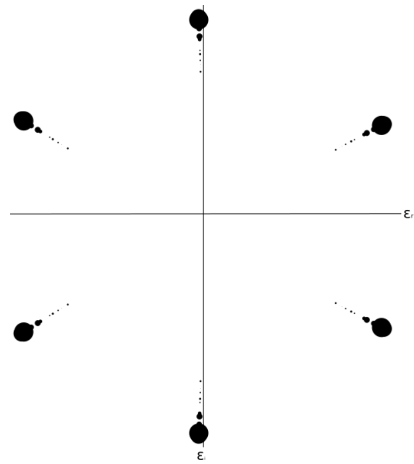

Figure 1.1: Domain of analyticity according to [5].

A common method to compute invariant circles of a map of the cylinder $f_\varepsilon : \mathcal{M} \to \mathcal{M}$, is by computing a parameterization $K_\varepsilon : \mathbb{S}^1 \to \mathcal{M}$ of the invariant circle which satisfies an invariance equation. The invariance equation is

$$f_\varepsilon \circ K_\varepsilon = K_\varepsilon \circ T_\omega$$

with $T_\omega(\theta) = (\theta + \omega)$. The invariance equation states that the dynamics on the invariant circle is conjugated to a rigid rotation of the circle by an irrational number ω. One of the advantages of working with the dissipative standard map is that the parameterization function K_ε can be written in terms of a periodic function, $u_\varepsilon : \mathbb{S}^1 \to \mathbb{R}$, as it is explained in Section 1.2.1. The method we use to find the singularities is to approximate the conjugacy function $u_\varepsilon(\theta)$ by means of a Lindstedt series expansion in ε. The Lindstedt method produces polynomials in ε of high order,

$$u_\varepsilon^{\leq N}(\theta) = \sum_{n=0}^{N} u_n(\theta)\varepsilon^n$$

whose coefficients $u_n : \mathbb{S}^1 \to \mathbb{C}$ are periodic functions. We then use the Lindstedt series of the conjugacies to obtain Padé rational functions whose poles are expected to approximate the poles of the original function u_ε. Padé extrapolation methods of Lindstedt series have been widely used by several authors [7, 8, 9, 10, 11] in the symplectic case. Since the Padé extrapolation method is based on approximating an analytic function with a rational function, the computation of poles is done by approximating the roots of the denominator of the Padé approximant. The denominator is a polynomial that can be of very high degree, and computing its roots depends heavily on numerical precision. Since the computations are very sensitive to precision, we perform them using $\approx 10^3$ digits which allows us to compute singularities for values of ε that are at a distance ≈ 0.3 from $\varepsilon = 0$ in the complex plane. We expect that higher precision together with higher order degree series, would allow us to compute poles that are closer to the origin. However, the method already allows us to have an approximation of the boundary of the domain in regions at which one does not expect the functions to be analytic, as it was predicted by the conjecture in [5], even when the singularities are not very close to $\varepsilon = 0$. We also find conjectures on the rate of growth of the terms of the Lindstedt series.

We note that the shapes of the domains that we present here are remarkably different from what one sees in the symplectic case, see [9, 10, 11, 12]. This is partly due to the fact that in the symplectic case the small divisors do not depend on the conformal factor $b(\varepsilon)$ which in our case is a function of ε.

Some explorations of the shape of the analyticity domains in the dissipative standard map have been performed using the parameterization method in [13], that is very similar to the one described in Section 1.3.3. In [13], it is noticed that the breakdown of invariant tori in the conservative and the dissipative case are similar when the conformal factor b is a constant, [14]. A different behavior in the breakdown of invariant tori involving bundle collapse is observed in the dissipative standard map in [15]. Explorations of the shape of the domains of analyticity in ε in the conservative case with the use of the parameterization method appear in [16, 12].

1.2 Preliminaries

We consider the dissipative standard map defined on the cylinder $\mathcal{M} = \mathbb{S}^1 \times \mathbb{R}$ given by $f_\varepsilon(x_n, y_n) = (x_{n+1}, y_{n+1})$ and

$$\begin{aligned} y_{n+1} &= b_\varepsilon y_n + c_\varepsilon + \varepsilon\, V'(x_n) \\ x_{n+1} &= x_n + y_{n+1}\ , \end{aligned} \tag{1.1}$$

where $y_n \in \mathbb{R}$, $x_n \in \mathbb{S}^1$, $\varepsilon \in \mathbb{R}$, and $V'(x) = \frac{1}{2\pi}\sin(2\pi x)$. Here we consider the case when the dissipative parameter, b_ε, is given by $b_\varepsilon = b(\varepsilon) = 1 - \varepsilon^3$, and the drift parameter $c_\varepsilon = c(\varepsilon)$ is a function that depends on the small parameter ε. Note that adding a dissipation to the system is a very singular perturbation and could lead to the creation of attractors/repellors without quasi periodic motions. For that reason, one has to consider this external parameter, c_ε. The dissipative parameter b_ε coincides with the Jacobian of the function. We note that adding an odd power of epsilon to the b_ε term is the physically relevant choice. For this work, we choose to include a third power since it is the first non–trivial odd case. We note that the Jacobian is the rate of dissipation/expansion of the map (1.1), this rate will be dependent of the parameter ε. In particular, the case $\varepsilon = 0$ coincides with the zero dissipation limit. We must emphasize that we tie the paremeter ε to the non-linearity of the map since in this case $\varepsilon = 0$ also coincides with the integrable case and that will hugely simplify our computation. In particular, by doing this we make sure that the symplectic case, which will be the zero-th order of our series, will be trivial.

In fact, it is discussed in [5] that (1.1) is conformally symplectic. That is, if $\Omega = dy \wedge dx$ is the standard symplectic form of the cylinder, the map f_ε satisfies that

$$f^*\Omega = b_\varepsilon \Omega. \tag{1.2}$$

For certain values of c_ε we know that maps of the form (1.2) have analytic invariant circles corresponding to quasi-periodic orbits with Diophantine rotation numbers, ω. The Lindstedt series analysis in Section 1.3.1 determines that one condition for the the mapping (1.1) to admit an invariant circle is that $c_\varepsilon = \omega\varepsilon^3 + \mathcal{O}(\varepsilon^4)$. In the following, we discuss the properties that the rotation number should satisfy so that one can have quasi-periodic orbits parameterized by a function.

1.2.1 Quasi-periodic orbits

We consider a frequency ω that satisfies the Diophantine condition,

$$|\omega q - p| \geq \nu |q|^{-\tau}, \quad p \in \mathbb{Z}, \quad q \in \mathbb{Z} \setminus \{0\} \tag{1.3}$$

where $\nu \in \mathbb{R}^+$ and $\tau \in \mathbb{R}$ with $\tau \geq 1$.

Quasi periodic orbits of the dissipative standard map (1.1) are found using a parametric representation of the variable $x_n \in \mathbb{S}^1$ as

$$x_n = \theta_n + u_\varepsilon(\theta_n), \quad \theta \in \mathbb{S}^1, \tag{1.4}$$

where $u_\varepsilon : \mathbb{S}^1 \to \mathbb{R}$ is a 1-periodic function. We assume that the variable θ_n varies linearly as $\theta_{n+1} = \theta_n + \omega$ where ω is the rotation frequency.

It follows from equation (1.1) that

$$x_{n+1} - (1 + b_\varepsilon)x_n + b_\varepsilon x_{n-1} - c_\varepsilon + \varepsilon V'(x_n) = 0. \tag{1.5}$$

We look for quasi periodic solutions by finding u_ε, as in (1.4), and $c_\varepsilon = c(\varepsilon)$ such that

$$E_{c_\varepsilon}[u_\varepsilon] = 0 \tag{1.6}$$

where

$$E_{c_\varepsilon}[u_\varepsilon] \equiv u_\varepsilon(\theta+\omega)-(1+b_\varepsilon)u_\varepsilon(\theta)+b_\varepsilon u_\varepsilon(\theta-\omega)+(1-b_\varepsilon)\omega-c_\varepsilon+\varepsilon V'(\theta+u_\varepsilon(\theta)) = 0. \quad (1.7)$$

It is clear that once we find a pair $(u_\varepsilon, c_\varepsilon)$ satisfying (1.6), we can recover the embedding of the quasi-periodic orbit by the parameterization $K_\varepsilon : \mathbb{S}^1 \to \mathcal{M}$,

$$K_\varepsilon(\theta) = \begin{pmatrix} \theta + u_\varepsilon(\theta) \\ \omega + u_\varepsilon(\theta) - u_\varepsilon(\theta-\omega) \end{pmatrix}. \quad (1.8)$$

We remark that the nature of the two unknowns is different since $u_\varepsilon(\theta)$ is a smooth complex 1-periodic function of $\theta \in \mathbb{S}^1$ depending on the complex parameter ε and c_ε is a complex number depending on ε. The conjecture in [5] states that ε is a complex parameter whose range lays in a complex domain that is obtained by taking out from a neighborhood of $\varepsilon = 0$ points inside balls with centers along smooth curves passing through the origin, see Figure 1.1. In [5] there is also a rigorous proof that the domain described in the conjecture is a lower bound that approximates the exact domain o analyticity.

1.3 Methods for computing solutions

We will use two different methods for finding the solution pair $(u_\varepsilon, c_\varepsilon)$ of (1.6). The first method is based on a Lindstedt series approximation of the solutions written as formal power series of the small parameter ε. In our case the small parameter ε will account both for the size of the perturbation and the distance of the conformal factor to the symplectic case. This method produces approximate solutions in the sense that if

$$u_\varepsilon^{\le N}(\theta) = \sum_{k=0}^{N} u_k(\theta)\varepsilon^k \quad \text{and} \quad c^{\le N}(\varepsilon) = \sum_{k=0}^{N} c_k\varepsilon^k \quad (1.9)$$

are polynomials in ε, we say that (1.9) is an approximate solution of order N whenever $\|E_{c^{\leq N}(\varepsilon)}[u_\varepsilon^{\leq N}]\| \sim |\varepsilon|^{N+1}$, where E is the functional defined in (1.6) and $\|\cdot\|$ is the supremum norm over all $\theta \in \mathbb{S}^1$. The Lindstedt series method that we describe in Section 1.3.1 provides a way to construct an approximate solution of any given order $N \in \mathbb{N}$.

In Section 1.3.3, we include an algorithm to find the solution $(u_\varepsilon, c_\varepsilon)$ by means of a Newton method. The method starts form an approximate solution pair (u_a, c_a) so that the norm of $E_{c_a}[u_a]$ is small and provides a correction (v, δ) by imposing that the new solution $(u_a + v, c_a + \delta)$ satisfies the functional equation $E_{c_a+\delta}[u_a + v]$ up to first order in (v, δ). This method can be shown to converge using scales of Banach spaces.

1.3.1 Lindstedt Series

The Lindstedt series method consists of performing a formal series expansion in a small parameter ε. According to (1.6), and the fact that $b(\varepsilon) = 1 - \varepsilon^3$, we look for a solution, $(u_\varepsilon, c_\varepsilon)$, of

$$u_\varepsilon(\theta + \omega) - (2 - \varepsilon^3)u_\varepsilon(\theta) + (1 - \varepsilon^3)u_\varepsilon(\theta - \omega) + \varepsilon^3\omega - c(\varepsilon) = -\varepsilon V'(\theta + u_\varepsilon(\theta)) \quad (1.10)$$

as a power series expansion. That is, we look for solutions

$$u_\varepsilon(\theta) = \sum_{k=0}^{\infty} u_k(\theta)\varepsilon^k \quad \text{and} \quad c(\varepsilon) = \sum_{k=0}^{\infty} c_k\varepsilon^k, \quad (1.11)$$

where each u_n is a function from $\mathbb{S}^1$ to $\mathbb{C}$ and each $c_n \in \mathbb{C}$. This solution can be computed by equating powers of ε in (1.10). Taking the Taylor expansion at $\varepsilon = 0$

$$-\varepsilon V'(\theta + u_\varepsilon(\theta)) = \sum_{k=0}^{\infty} S_k(\theta)\varepsilon^k \quad (1.12)$$

and substituting (1.11) into (1.10), we have that

$$\sum_{k=0}^{\infty} u_k(\theta+\omega)\varepsilon^k-(2-\varepsilon^3)\sum_{k=0}^{\infty} u_k(\theta)\varepsilon^k+(1-\varepsilon^3)\sum_{k=0}^{\infty} u_k(\theta-\omega)\varepsilon^k+\varepsilon^3\omega-\sum_{k=0}^{\infty} c_k\varepsilon^k=\sum_{k=0}^{\infty} S_k(\theta)\varepsilon^k. \quad (1.13)$$

Remark 1. *When $V'(\theta) = \frac{1}{2\pi}\sin(2\pi\theta)$, or a trigonometric polynomial, the $S_k(\theta)$'s can be computed very efficiently in terms of the $u_i(\theta)$'s. Following [17, 18] and denoting $\mathscr{S}(\theta,\varepsilon) = \sin(2\pi(\theta + u_\varepsilon(\theta)))$, $\mathscr{C}(\theta,\varepsilon) = \cos(2\pi(\theta + u_\varepsilon(\theta)))$, the coefficients of the series expansions $\mathscr{S}(\theta,\varepsilon) = \sum_{k=0}^{\infty} \mathscr{S}_k(\theta)\varepsilon^k$ and $\mathscr{C}(\theta,\varepsilon) = \sum_{k=0}^{\infty} \mathscr{C}_k(\theta)\varepsilon^k$ are given by the following recurrence relations,*

$$\begin{aligned}(N+1)\mathscr{S}_{N+1}(\theta) &= 2\pi\sum_{m=0}^{N}(m+1)\mathscr{C}_{N-m}(\theta)u_{m+1}(\theta) \\ (N+1)\mathscr{C}_{N+1}(\theta) &= -2\pi\sum_{m=0}^{N}(m+1)\mathscr{S}_{N-m}(\theta)u_{m+1}(\theta).\end{aligned} \quad (1.14)$$

Thus, $S_k(\theta) = -\mathscr{S}_{k-1}(\theta)$ for $k \geq 1$ and $S_0 = 0$, by (1.12).

Defining the operator

$$L_\omega\varphi(\theta) := \varphi(\theta+\omega) - 2\varphi(\theta) + \varphi(\theta-\omega) \quad (1.15)$$

equation (1.13) can be rewritten as

$$\begin{aligned}\sum_{k=1}^{\infty} S_k(\theta)\varepsilon^k = &\sum_{k=0}^{2}\left(L_\omega u_k(\theta) - c_k\right)\varepsilon^k + \left(L_\omega u_3(\theta) - c_3 + u_0(\theta) - u_0(\theta-\omega) + \omega\right)\varepsilon^3 \\ &+ \sum_{k=4}^{\infty}\left(L_\omega u_k(\theta) - c_k + u_{k-3}(\theta) - u_{k-3}(\theta-\omega)\right)\varepsilon^k,\end{aligned} \quad (1.16)$$

Some properties of the operator L_ω are summarized in the following Lemma. See [19] for details about the proof.

Lemma 2. *Let $\eta : \mathbb{S}^1 \to \mathbb{S}^1$ a continuous function such that $\int_0^1 \eta(\theta)d\theta = 0$. If ω is*

Diophantine as in (1.3)*, then there exists a solution,* $\varphi(\theta)$*, to the equation*

$$L_\omega \varphi(\theta) = \eta(\theta) \tag{1.17}$$

such that $\int_0^1 \varphi(\theta) d\theta = 0$*. In fact, the solution is given by*

$$\varphi(\theta) = \sum_{\ell \in \mathbb{Z} \setminus \{0\}} \frac{\hat{\eta}_\ell}{2(\cos(2\pi\ell\omega) - 1)} e^{2\pi i \ell \theta},$$

where $\hat{\eta}_\ell$ *are the Fourier coefficients of* $\eta(\theta)$*.*

The Lindstedt process is as follows. Matching the coefficients of the same order in (1.16) we obtain the following relations to different orders of ε. The zero-th order term tells us that the coefficients at order zero in ε have to be trivial. The equations are

$$L_\omega u_0(\theta) - c_0 = 0. \tag{1.18}$$

Choosing $c_0 = 0$, then $u_0 \equiv 0$ is the solution given by Lemma 2. This construction is analogous to the zero-th order term in the symplectic case.

Remark 3. *This method has been used in [18, 19, 7, 8, 9, 10, 11] for the symplectic case. That is, making the same process for the standard map,* $(x_{n+1}, y_{n+1}) = (x_n + y_{n+1}, y_n + \varepsilon V'(x_n))$*, gives the following equation to all orders* ε^n*,*

$$L_\omega u_k(\theta) = S_k(\theta) \qquad k \geq 0. \tag{1.19}$$

Moreover, $\int_0^1 S_k(\theta) d\theta = 0$ *for all* $k \geq 0$*. This is a consequence of the symplectic structure and the fact that* $S_k(\theta)$ *depends on the previously computed* $u_0(\theta)$*,* $u_1(\theta)$*, ...,* $u_{k-1}(\theta)$*,* $S_0(\theta)$*,* $S_1(\theta)$*, ...,* $S_{k-1}(\theta)$ *(see Remark 1).*

The first and second orders in ε are also analogous to the symplectic case. For this reason the first two coefficients of c_ε will be trivial.

$$L_\omega u_k(\theta) - c_k = S_k(\theta), \qquad k = 1, 2. \tag{1.20}$$

Choosing $c_1 = 0 = c_2$ the equations are reduced to the non dissipative case and, by Remark 3, the right hand side has zero average. Therefore, we can find solutions $u_1(\theta)$, $u_2(\theta)$.

The third order in ε is the first one that is different from the conservative case.

$$L_\omega u_3(\theta) - c_3 + \omega = S_3(\theta). \tag{1.21}$$

Here we notice that the drift parameter starts playing a rôle in the existence of invariant tori. Taking $c_3 = \omega$, equation (1.21) becomes the same equation as in the symplectic case. Since $S_3(\theta)$ has zero average we find $u_3(\theta)$.

The equations for orders higher than 3 are remarkably different since we have a counter term coming from the previously computed orders. Namely,

$$L_\omega u_k(\theta) = S_k(\theta) - u_{k-3}(\theta) + u_{k-3}(\theta - \omega) + c_k, \qquad k \geq 4. \tag{1.22}$$

Notice that, by construction, $\int_0^1 u_{k-3}(\theta - \omega) d\theta = \int_0^1 u_{k-3}(\theta) d\theta = 0$ (see Lemma 2). Now, taking

$$c_k = -\int_0^1 S_k(\theta) d\theta, \tag{1.23}$$

we can find $u_k(\theta)$ solving (1.22) for all $k \geq 4$.

We have proved the following proposition which is a particular case of part A) of Theorem 12 in [5].

Proposition 4. *For any $N \in \mathbb{N}$, the procedure presented above allows to find an approxi-*

mate solution,

$$u_\varepsilon^{\leq N}(\theta) = \sum_{k=0}^{N} u_k(\theta)\varepsilon^k \quad \textit{and} \quad c^{\leq N}(\varepsilon) = \sum_{k=0}^{N} c_k\varepsilon^k, \tag{1.24}$$

such that

$$\|E_{c^{\leq N}(\varepsilon)}[u_\varepsilon^{\leq N}]\| \leq C|\varepsilon|^{N+1}$$

where E is the functional defined in (1.6).

1.3.2 Padé extrapolation

The domain of analyticity for the solution of (1.6) can be approximated by implementing a Padé method in which we use the approximate solutions obtained by the Lindstedt series constructed in Section 1.3.1.

The Padé method is quite standard and is presented in several places in the literature. Here, we follow the exposition in [20]. A Padé approximant of order $[p/q]$ of a function $g(\varepsilon) = \sum_{i=0}^{\infty} g_i\varepsilon^i$ is a rational function, $P(\varepsilon)/Q(\varepsilon)$, which agrees with g to the highest possible order in ε.

That is,

$$g(\varepsilon) - \frac{P(\varepsilon)}{Q(\varepsilon)} = \mathcal{O}(\varepsilon^{p+q+1}) \tag{1.25}$$

where $P(\varepsilon)$ and $Q(\varepsilon)$ are polynomials of degree p and q respectively, $Q(0) = 1$.

The existence of the polynomials P and Q can be obtained by noticing that (1.25) is equivalent to

$$g(\varepsilon)Q(\varepsilon) = P(\varepsilon) + \mathcal{O}(\varepsilon^{p+q+1})$$

and, then, considering $P(\varepsilon) = \sum_{i=0}^{p} P_i\varepsilon^i$ and $Q(\varepsilon) = \sum_{i=0}^{q} Q_i\varepsilon^i$ the coefficients of the polynomials can be found by solving the following systems of equations

$$\begin{aligned} g_i + \sum_{j=1}^{i} g_{i-j} Q_j &= P_i \qquad 0 \le i \le p \\ g_i + \sum_{j=1}^{q} g_{i-j} Q_j &= 0 \qquad p < i \le p+q. \end{aligned} \tag{1.26}$$

The second equation of (1.26) gives the $Q_j's$, and then we can find the $P_j's$ by substituting in the first equation. Then, the boundary of the domain of analyticity of a function can be approximated by the zeros of Q in the $[p/q]$ Padé approximant.

There are a number of implementations of the Padé methods that are used in a quite standard manner. In the present work we use the implementations included in Version 2.9.0 of GP/PARI, [21].

1.3.3 Newton's method

In this section we summarize an iterative scheme in scales of Banach spaces that can be very well adapted to perform numerical computations. The scheme is based in a Newton iteration starting from approximate solutions to the equation (1.6). We briefly describe the scheme here since details of schemes of these kind and numerical implementations have been described already in the literature [13, 5, 15, 22], and the reader can refer to these works for more details.

We start from an approximate solution (u_a, c_a) of equation (1.6). Namely, we have a solution so that $\|E_{c_a}[u_a]\|$ is small enough. The approximate solution could be obtained by several means. One possibility is starting from the integrable case (for ε close to zero) and perfoming continuation or from a Lindstedt series expansions like the ones obtained in Section 1.3.1. We remark that in the dissipative standard map we are studying, $\varepsilon = 0$ is the point where the map becomes symplectic. Since we use methods for conformally symplectic systems we actually start the continuation from values of ε that are not equal to zero but small.

The Newton algorithm consists of adding a correction (v, δ) to the approximate solution so that supremum norm of (1.6) evaluated in the function plus the corrections, $\|E_{c_a+\delta}[u_a + v]\|$, is of the order of the square of the norm of (1.6) evaluated at the approximate solution,

$$\|E_{c_a+\delta}[u_a + v]\| \leq C\|E_{c_a}[u_a]\|^2.$$

One obtains the correction by solving the linearized equation of $E_{c_a+\delta}[u_a + v]$ for (v, δ) around the approximate solution, (u_a, c_a).

In this case, the equation we have to solve is

$$D_u E_{c_a}[u_a]v - \delta = -E_{c_a}[u_a] \tag{1.27}$$

which involves unbounded operators in Banach spaces (namely $D_u E_{c_a}[u_a]v$) that are actually bounded if one considers that the operators map into Banach spaces of less regularity. It is a standard observation in Nash–Moser theory [23, 24], that to set up a converging iterative Newton scheme it is not necessary to find an exact inverse of the operator $D_u E_{c_a}[u_a]$, but finding an exact inverse of an approximate operator will suffice.

In our case, we will not solve equation (1.27) directly but will solve a modified equation obtained by adding a term that is quadratic in the error. We remark that since a Newton method is a quadratic scheme solving a linear equation then adding a quadratic term also gives a quadratic scheme that solves the same linear equation. Defining $h'(\theta) = 1 + \frac{\partial}{\partial\theta}u_a(\theta)$, one obtains an approximate Newton equation by subtracting the term $vD_u E_{c_a}[u_a]h'$, which is quadratic in the error. The modified Newton equation is,

$$h' D_u E_{c_a}[u_a]v - vD_u E_{c_a}[u_a]h' = -h'(E_{c_a}[u_a] - \delta)\ . \tag{1.28}$$

The l.h.s. of equation (1.28), factorizes into a sequence of operators that are easier to solve numerically, as it is noted in Lemma 5.

This method has been used in several works [25, 26, 19]. Here we only make a reference to the justification in [13], where the reader can refer to for details.

Let the operators $\mathcal{D}_-$, $\mathcal{D}_+^b$ be defined by

$$\begin{aligned} \mathcal{D}_- f(\theta) &= f(\theta-\omega) - f(\theta) \\ \mathcal{D}_+^b f(\theta) &= f(\theta+\omega) - bf(\theta)\ . \end{aligned} \tag{1.29}$$

A small remark is that (1.29) are operators that are diagonal in Fourier space. In the following lemma, we write the modified Newton as a sequence of operators that are either diagonal in Real or Fourier space.

Lemma 5. *The modified Newton equation in* (1.28) *with* $E_{c_a}[u]$ *defined in* (1.6) *is equivalent to*

$$\mathcal{D}_+^b[-h'(\theta)h'(\theta-\omega)\mathcal{D}_-[(h')^{-1}(\theta)v(\theta)]] = -h'(E_{c_a}[u_a](\theta)-\delta). \tag{1.30}$$

Remark 6. *One notices that the operators involved in the l.h.s. of equation* (1.30) *only involve differentiation, multiplication, division, shifting the arguments of functions, and solving the difference equations with constant coefficients in* (1.29). *All this operations can be implemented very efficiently using the computer. For instance if we discretize the periodic functions using* n *uniformly distributed points and we use a Fast Fourier Transform method, the modified Newton step equation can actually be solved in* $O(n\log n)$ *operations.*

The factorization in equation (1.30) suggests an algorithm that is used to solve the modified Newton equation.

Algorithm 7. ***i)*** *Find two functions* φ *and* ν *solving the equations*

$$\mathcal{D}_+^b\varphi(\theta) = -h' E_{c_a}[u] \tag{1.31}$$

and

$$\mathcal{D}_+^b\nu(\theta) = -h'(\theta)\ . \tag{1.32}$$

Notice that if $\varphi(\theta)$ and $\nu(\theta)$ are solutions of (1.31) *and* (1.32)*, respectively, then the equation $\mathcal{D}_+^b(\varphi(\theta) - \delta\nu(\theta)) = -h'(\theta)(E_{c_a}[u_0](\theta) - \delta)$ holds for any $\delta \in \mathbb{C}$. This will allow us to chose a complex number δ so that the average of $\frac{\varphi(\theta)-\delta\nu(\theta)}{h'(\theta)h'(\theta-\omega)}$ vanishes.*

ii) *Choose $\delta \in \mathbb{C}$ such that*

$$\int_{\mathbb{T}} \frac{\varphi(\theta) - \delta\nu(\theta)}{h'(\theta)h'(\theta - \omega)} d\theta = 0 \ .$$

iii) *Obtain w from the solution of the constant coefficient difference equation*

$$\mathcal{D}_- w(\theta) = \frac{\varphi(\theta) - \delta\nu(\theta)}{-h'(\theta)h'(\theta - \omega)} \ . \tag{1.33}$$

Notice that after choosing a δ in step ii) *so that the right hand side has zero average we can always find a periodic function w solving* (1.33) *when the r.h.s. is smooth enough.*

iv) *Construct $v(\theta) = h'(\theta)w(\theta)$ and obtain the improved solution $(\tilde{u}, \tilde{c})$ defined as*

$$\tilde{u}(\theta) = u_a(\theta) + v(\theta) \ , \qquad \tilde{c} = c_a + \delta \ .$$

The observation in remark 6 is that the operators in (1.30) are very efficiently implementable with the use of a computer either in Real or in Fourier space. This efficiency comes from the fact that all the operations involved in the four steps of Algorithm 7 are multiplications, additions and integrals of periodic functions that take only $O(n)$ operations in Real space; and differentiation, shifts and solving cohomology equations with constant coefficients, that take only $O(n)$ operations in Fourier space. Therefore, the most expensive operation in the Algorithm 7 is transforming from Real to Fourier space and back. This can be done in $O(n \log n)$ operations by means of a Fast Fourier Transform.

Remark 8. *We note that the algorithm is guaranteed to converge inside the boundaries of* *the analyticity domain. Indeed, in [12] it was rigorously justified that the algorithm only fails to converge as the continuation reaches the boundary of analyticity. Therefore, the continuation method can also be used to asses the bounds on the domain of ε.*

1.4 Numerical results

In this section we present the results of implementing the methods described in Section 1.3. All the computations in Sections 1.4.1 and 1.4.2 were done using the golden ratio, $\omega = \frac{\sqrt{5}-1}{2}$, which satisfies (1.3), see [19].

1.4.1 Domain of analyticity

The construction of Lindstedt series in Section 1.3.1 was implemented as a numerical algorithm. The statement of Proposition 4 tells us that given any $N \in \mathbb{N}$, the outcome of the method is the pair of polynomials of degree N in (4). The observation of Lemma 2 is that the operator L_ω defined in equation (1.15) is diagonal in Fourier Space and equation 1.17 can be solved for ϕ if we allow to obtain functions with less regularity than the right hand side, η. We find the solution numerically by transforming to Fourier space and solving for the u_k's from expressions (1.20) to (1.22). At every order of the process we obtain the c_k's as a byproduct of imposing the condition that every order should have zero average. The Lindstedt series expansions are used to obtain an approximate solution to the functional equation in (1.6) at some high order.

Approximation of poles of the Lindstedt series

Here we include the poles of the Lindstedt polynomial found with the Padé method. In Figure 1.2, we show the poles of the series approximated by means of the Padé method. It is well known that the Padé method computations are very sensitive to precision, see [20], so we have implemented the computations with extended precision using the software gp/Pari [21]. We show the values of the poles in the ε complex plane, and the complex values of the function $b(\varepsilon) = 1 - \varepsilon^3$. Our computations suggest that the boundary of the analyticity domain has a more complicated structure that what was predicted in [5], compare the left panel of Figure 1.2 with Figure 1.1. Figure 1.3 contains the comparison of the values of the function $b(\varepsilon)$ with the unit circle. We also include zoomed in versions of the values of $b(\varepsilon)$ in Figure 1.3.

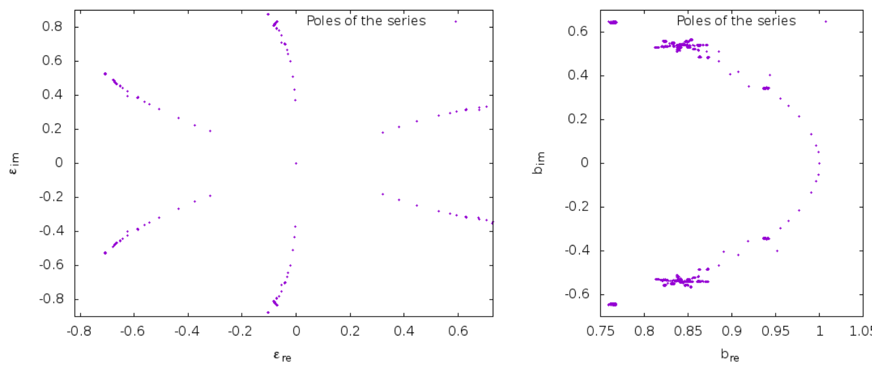

Figure 1.2: Points which are simultaneously poles of Padé approximants of degree [475,475] and [500,500]. The implementation was done with 1000 digits. Left panel: Poles in the complex plane $\varepsilon \in \mathbb{C}$. Right panel: Poles evaluated in the function $b(\varepsilon) = 1 - \varepsilon^3$, with $\varepsilon \in \mathbb{C}$.

Newton method

We used Newton's method and continuation to explore the monodromy of the solutions in the domains. A rigorous result in [5] states that the solutions defined in the domain of analyticity in ε have trivial monodromy. We verified this fact numerically by perfoming continuation of the solutions $(u_\varepsilon, c_\varepsilon)$ around the poles that were previously approximated using the Padé series method described in Section 1.4.1.

We used the approximated poles as centers of circular paths in ε over which we perfomed continuation while solving the invariance equation (1.6) using Algorithm 7. Once the continuation achieves a complete turn around a chosen pole, one verifies that the solution always arrives to the same starting point. This is an effect of the monodromy of the functions being trivial.

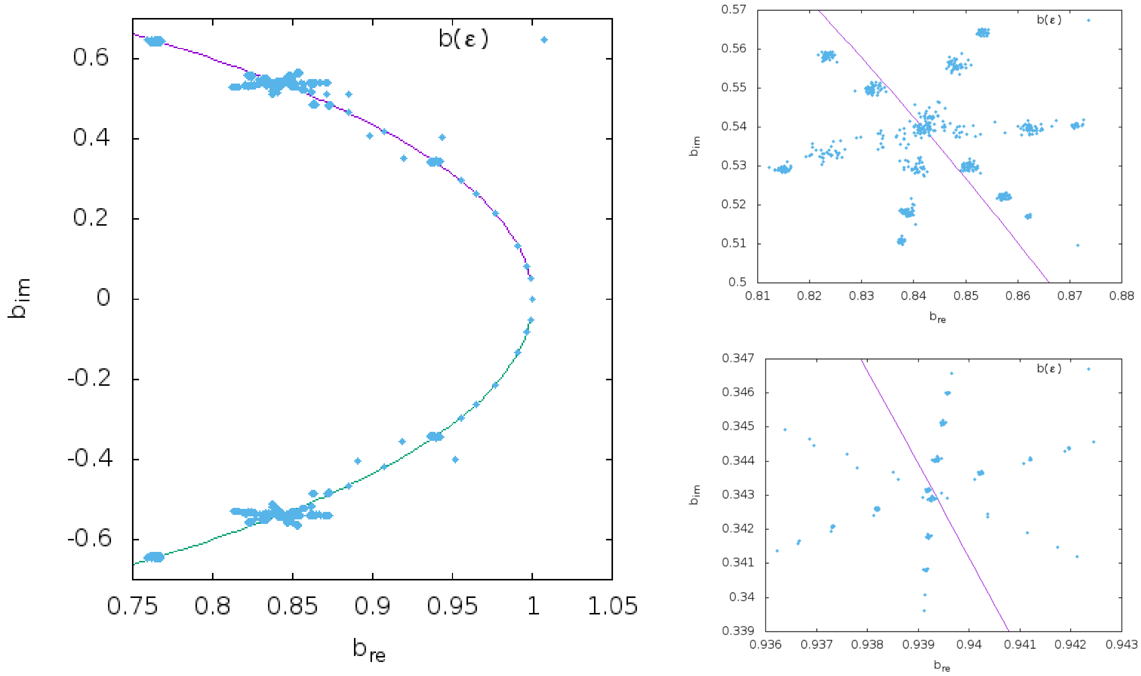

Figure 1.3: The poles compared to the unit circle. Left panel: Evaluation of the poles of the series by the function $b(\varepsilon) = 1 - \varepsilon^3$. Right panels: Two zoomed in versions of the set.

We present several instances of the functions for different parameter values along a circle winding around a pole in Figure 1.5. The path we used to surround the pole is presented in Figure 1.4. The continuation was perfomed using FFTW3, [27], with the libquadmath library, [28]. The radii of the continuation paths were chosen so that the path did not come very close to the poles. Indeed, when the continuation comes close to a pole our implementation of the Newton method becomes degenerate in the sense that one needs to compute quotients of very small quantities. The reason is that when solving equations (1.31) and (1.32), the divisors depending on ε, are below machine precision close to the pole and dividing over those quantities leads to large numerical errors.

1.4.2 Growth of the coefficients of the Lindstedt expansions.

The results in Section 1.4.1 agree with the conjecture in [5] about the domain of analyticity of the parameterization of the quasi periodic orbits. The qualitatively conjectured optimal domain of analyticity of the parameterizations for the map (1.1) does not contain any ball

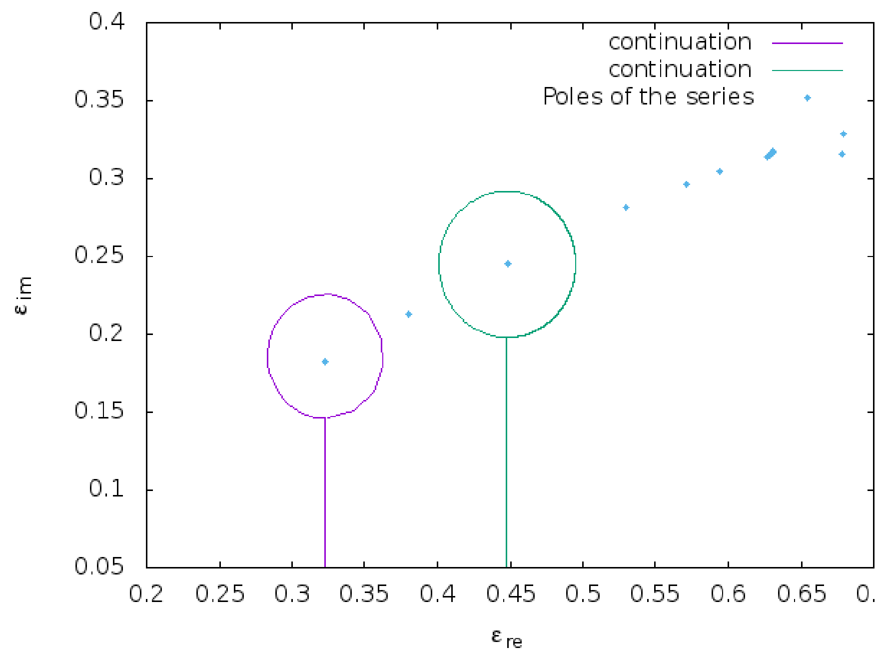

Figure 1.4: Poles of the series and two different continuations done with the Newton algorithm. The continuation is done around the pole in order to illustrate that the monodromy is trivial.

with center at the origin nor angular sectors with width larger than $\pi/3$, so one does not expect the Lindstedt series to converge. The shape of the domain of analyticity suggest the Lindstedt expansions might belong to a Gevrey class.

We recall that a formal power series, $\sum f_n \varepsilon^n$, is σ-Gevrey (at $\varepsilon = 0$), with respect to a norm $\|\cdot\|$, if the coefficients satisfy

$$\|f_n\| \leq C R^n n^{\sigma n}. \tag{1.34}$$

where $\sigma \geq 0$. Equivalently,

$$\frac{1}{n} \log \|f_n\| \sim \sigma \log(n) + \log(R)$$

for n large enough.

Table 1.1: Values of ε and $c(\varepsilon)$ for different instances taken from the small circle in Figure 1.4.

Instance	ε	$c(\varepsilon)$
1	$0.3202966 + i0.1460915$	$0.01994937 - i0.06774761$
2	$0.3008391 + i0.1527000$	$0.009976542 - i0.06136120$
3	$0.2830167 + i0.1871540$	$-0.01146081 - i0.06221038$
4	$0.3122423 + i0.2245263$	$-0.02718298 - i0.08804174$
5	$0.3613448 + i0.1973876$	$0.007928831 - i0.1127768$
6	$0.3242691 + i0.1460201$	$0.02157160 - i0.06953580$

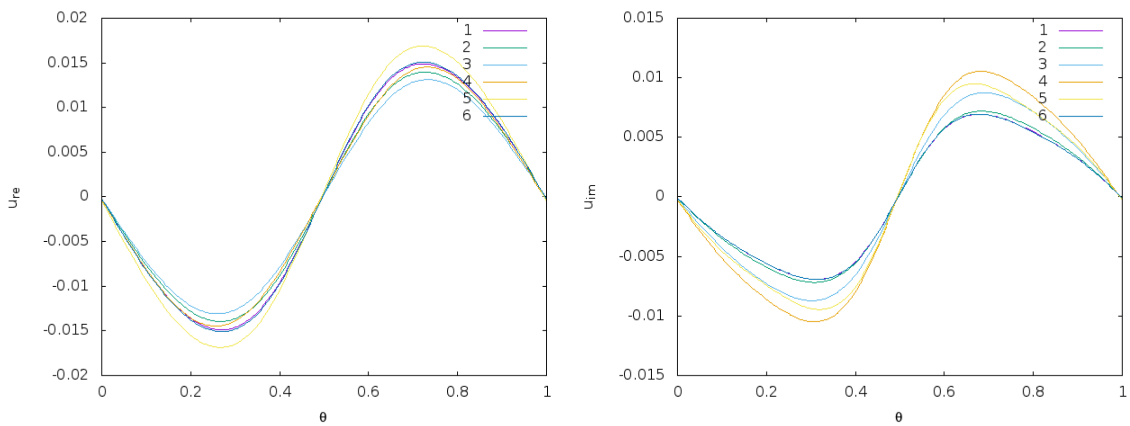

Figure 1.5: Real and imaginary part of different instances of a continuation by the Newton algorithm including the initial and final functions. One observes that there is no monodromy after a full turn around the pole.

Remark 9. *We note that, by Stirling's formula, inequality* (1.34) *is equivalent to*

$$\|f_n\| \leq CR^n (n!)^\sigma. \tag{1.35}$$

If $\sigma = 0$ the power series would define an analytic function of ε.

We approximated several norms of the coefficients $u_k(\theta)$ to have an indication of how far the functions are from being analytic. First, we use the norm on the complex strip of size $\rho > 0$, i.e., $\theta \in \mathbb{S}^1_\rho$ if $|\operatorname{Im}(\theta)| < \rho$. Let $f : \mathbb{S}^1_\rho \to \mathbb{S}^1_\rho$ be a function of $\mathbb{S}^1_\rho$ then the norm we use is

$$\|f\|_\rho = \sum_{\ell \in \mathbb{Z}} |\hat{f}_\ell|^2 e^{2\pi|\ell|\rho}$$

where $\hat{f}_\ell$ are the Fourier coefficients of f.

Since we want to check if the formal series $u_\varepsilon(\theta) = \sum u_k(\theta)\varepsilon^k$ belongs to a Gevrey class with the analytic norms, it is convenient to compute the following expressions as functions of k,

$$A_\rho(k) \equiv \frac{1}{k}\log \|u_k(\theta)\|_\rho, \tag{1.36}$$

and then approximate the constant σ. The values of $A_\rho(k)$ as in (1.36) for different values of ρ are shown in Figure 1.6.

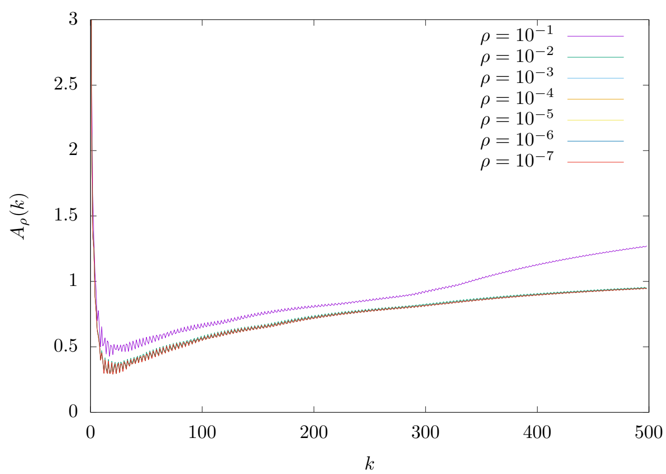

Figure 1.6: Plot of $A_\rho(k)$, $1 \leq k \leq 500$, for the frequency $\omega = \frac{\sqrt{5}-1}{2}$.

We also used Sobolev norms defined for a real number $r > 0$ by the L^2-norm of the r^{th} derivative with respect to θ,

$$\|f\|_{W^r} = \|\partial_\theta^r f\|_{L^2}.$$

The Sobolev norms can also be written in terms of Fourier coefficients as follows,

$$\|f\|_{W^r} = \left(\sum_{k\in\mathbb{Z}} (2\pi k)^{2r}|\hat{f}_k|^2\right)^{1/2}.$$

As in the case for analytic norms we define the following expressions for the Sobolev norms,

$$H^r(k) \equiv \frac{1}{k}\log \|u_k(\theta)\|_{W^r}. \tag{1.37}$$

We include the values of $H^r(k)$ for the coefficients of the approximate solution and several values of r in Figure 1.7.

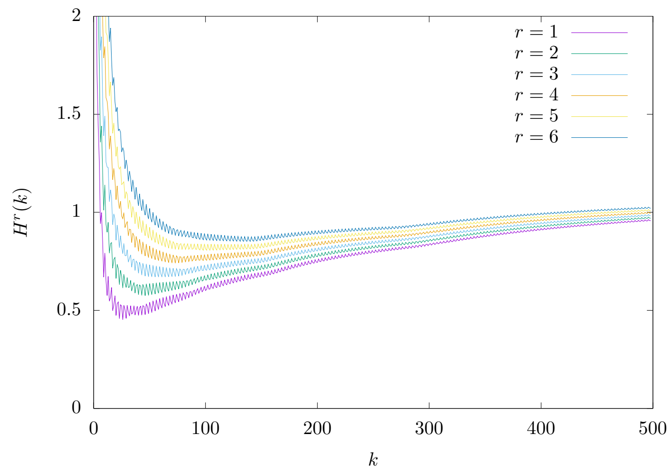

Figure 1.7: Plot of $H^r(k)$, $1 \leq k \leq 500$, for the frequency $\omega = \frac{\sqrt{5}-1}{2}$.

In both cases, the rate of growth of the norms $\|u_k(\theta)\|_B$ with respect to k seem to be large enough to make one think that the formal expansion belong to a Gevrey class. To study the rates of growth more systematically we have performed numerical fits of functions of the form $\log(R) + \sigma \log(k)$ to the data in Figures 1.6 and 1.7. In Tables 1.2 and 1.3 we summarized this numerical fits.

Table 1.2: Numerical fit of a function $\log(R) + \sigma \log(k)$ to the data $A_\rho(k)$ for different values of ρ and frequency $\omega = \frac{\sqrt{5}-1}{2}$. Computations were done using 2^{13} Fourier coefficients and 600 digits of precision. The numerical fit was made in for $100 \leq k \leq 300$.

$e_\rho(k) := A_\rho(k) - (\log(R) + \sigma \log(k))$			
	R	σ	$\|e_\rho\|_\infty$
$\rho = 0.1$	0.672269	0.227899	0.020793
$\rho = 0.01$	0.585740	0.238324	0.019491
$\rho = 0.001$	0.576278	0.240049	0.019325
$\rho = 0.0001$	0.575333	0.240225	0.019280
$\rho = 0.00001$	0.575239	0.240243	0.019282
$\rho = 0.000001$	0.575230	0.240244	0.019279
$\rho = 0.0000001$	0.575229	0.240244	0.019278

We note that the numbers R and σ in Table 1.2 and Table 1.3 are just the raw numbers obtained by fitting numerically functions of the form $\log(R) + \sigma \log(k)$ to the data $A_\rho(k)$ and $H^r(k)$, we are not sure how to assess the reliability of these numbers. Also, we have

Table 1.3: Numerical fit of a function $\log(R) + \sigma \log(k)$ to the data $H^r(k)$ for different values of r and frequency $\omega = \frac{\sqrt{5}-1}{2}$. Computations were done using 2^{13} Fourier coefficients and 600 digits of precision. The numerical fit was made for $100 \leq k \leq 300$.

$e_r(k) := H^r(k) - (\log(R) + \sigma \log(k))$			
	R	σ	$\|e_r\|_\infty$
$r = 1$	0.685071	0.212840	0.020144
$r = 2$	0.816610	0.185284	0.023905
$r = 3$	0.974288	0.157572	0.028145
$r = 4$	1.163403	0.129713	0.032216
$r = 5$	1.390238	0.101731	0.036129
$r = 6$	1.662287	0.073651	0.039905

added a column with a measure of the remainder, $\|e\|_\infty$, between the numerical fit and the data. The measure of these remainders, which looks a little bit worrisome, seems to come from an *oscillatory* behavior in the data, the structure of the remainders is studied in Section 1.4.4.

For the sake of completeness we include a comparison between $A_{10^{-7}}(k)$, $H^6(k)$ and their respective numerical fits, see Figure 1.8. Note that even if the norms considered in $A_\rho(k)$ and $H^r(k)$ are in principle not compatible, the fact that $A_\rho(k)$ and $H^r(k)$ have similar trends seems to indicate that there is a mechanism which is captured for any norm for the functions we study. This suggest that a more detailed study of the structure of this functions could be interesting.

1.4.3 Results for different frequencies

We recall that the computations presented in Section 1.4.2 were done using the frequency $\omega = \frac{\sqrt{5}-1}{2}$. In this section we present the computations also for different frequencies, of the same Diophantine type. We present the results below.

Figure 1.9 contains a plot of $A_\rho(k)$, $\rho = 10^{-7}$, for all the frequencies considered. The plots in Figure 1.9 seems to indicate a logarithm growth for all the frequencies considered. To study more systematically the growth of the coefficients of the Lindstedt series we have also fitted numerically functions of the form $\log(a) + \sigma \log(k)$, the results are summarized

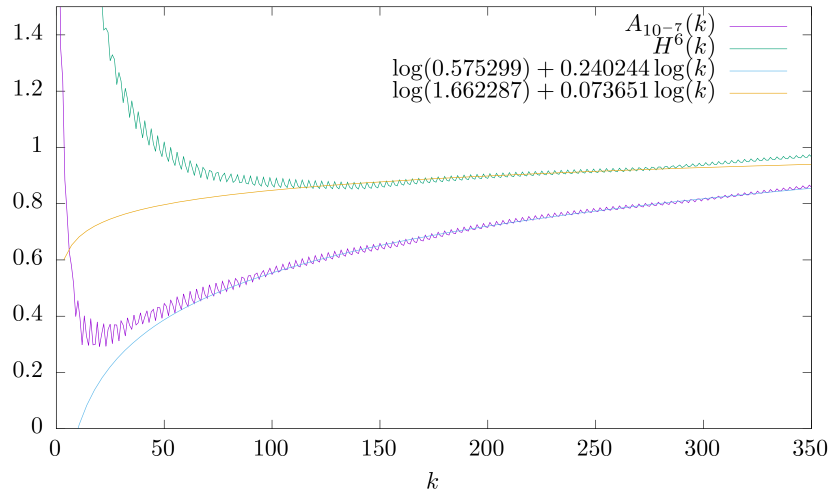

Figure 1.8: Comparison between $H^6(k)$ and $A_{10^{-7}}(k)$ with their respective numerical fits, $\omega = \frac{\sqrt{5}-1}{2}$.

in Table 1.4. We note that all the frequencies considered belong to the same Diophatine class $\mathcal{D}(\nu, 1)$, where $\omega \in \mathcal{D}(\nu, \tau)$ means that $|e^{2\pi i k\omega} - 1| \geq v|k|^{-\tau}$. Figure 1.10 and Figure 1.11 contain comparisons between the quantities A_ρ and their respective numerical fits.

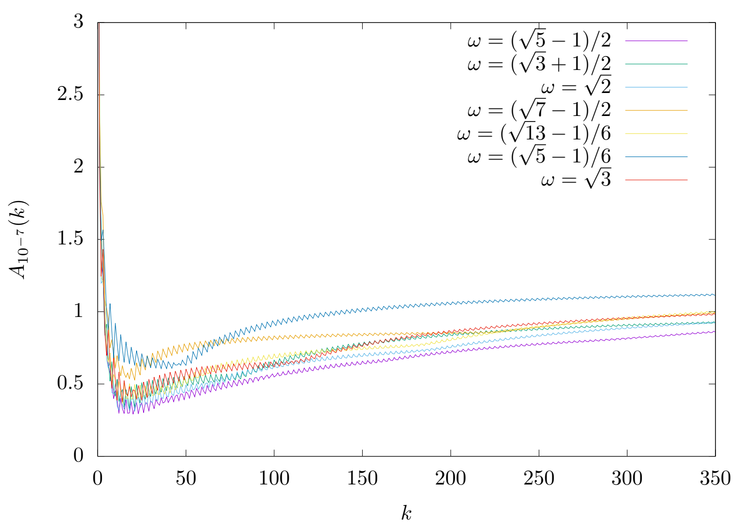

Figure 1.9: Graph of $A_\rho(k)$ for different values of the frequencies ω, $\rho = 10^{-7}$.

The numerical results presented in Sections 1.4.2 and 1.4.3 lead us to think that the Linsdtedt series of the parameterization of quasiperiodic orbits for the dissipative standard map (1.1) belong to a Gevrey class.

Table 1.4: Numerical fit of a function $\log(R) + \sigma \log(k)$ to the data $A_\rho(k)$ for different values of the frequency ω and $\rho = 10^{-7}$. Computations were done using 2^{13} Fourier coefficients and 600 digits of precision. The numerical fit was made in for $100 \leq k \leq 300$.

$e_\omega(k) = A_\rho(k) - \log(R) + \sigma \log(k)$, $\rho = 10^{-7}$			
	R	σ	$\\|e_\omega\\|_\infty$
$\omega = \frac{\sqrt{5}-1}{2} = [0, 1, 1, 1, 1, 1, 1, ...]$	0.575229	0.240244	0.019278
$\omega = \frac{\sqrt{3}-1}{2} = [0, 2, 1, 2, 1, 2, 1, ...]$	0.695887	0.225349	0.047762
$\omega = \sqrt{2} = [1, 2, 2, 2, 2, 2, 2, ...]$	0.583365	0.247799	0.033104
$\omega = \sqrt{3} = [1, 1, 2, 1, 2, 1, 2, 1, ...]$	0.460186	0.307029	0.038801
$\omega = \frac{\sqrt{7}-1}{2} = [0, 1, 4, 1, 1, 4, 1, 1, ...]$	1.300597	0.112924	0.045704
$\omega = \frac{\sqrt{13}-1}{6} = [0, 2, 3, 3, 3, 3, 3, ...]$	0.582937	0.258504	0.047840
$\omega = \frac{\sqrt{5}-1}{6} = [0, 4, 1, 5, 1, 5, 1, 5, ...]$	1.235768	0.158503	0.042327

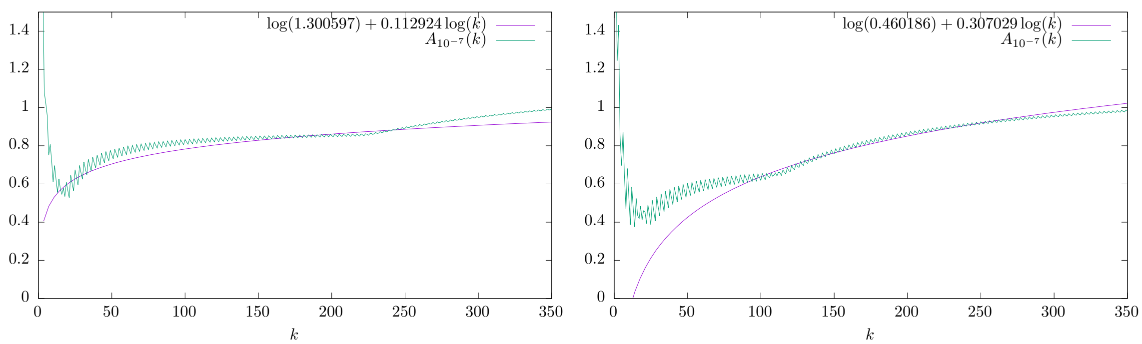

Figure 1.10: Comparison between $A_\rho(k)$, $\rho = 10^{-7}$, and its corresponding numerical fit. Left panel: values for the frequency $\omega = \frac{\sqrt{7}-1}{2}$. Right panel: values for the frequency $\omega = \sqrt{3}$

Conjecture 10. *Given $\omega \in \mathcal{D}(\nu, 1)$, the Lindstedt series, $u_\varepsilon = \sum u_k \varepsilon^k$, of quasi-periodic orbits for the map* (1.1) *belongs to a Gevrey class with Gevrey exponent $\sigma \leq 0.307$. That is, $\|u_n\|_\rho \leq C R^n n^{\sigma n}$ with $\sigma \leq 0.307$ and $\rho \leq 10^{-7}$.*

1.4.4 Some interesting patterns

A careful inspection of Figure 1.6, Figure 1.9, Figure 1.10, and Figure 1.11 shows that the graphs of $A_\rho(k)$ present an *oscillatory behavior* of period three, see Figure 1.12. These *oscillations* are present for all the values of the frequencies considered.

As we mentioned before, the coefficients of the Lindstedt series are determined by solv-

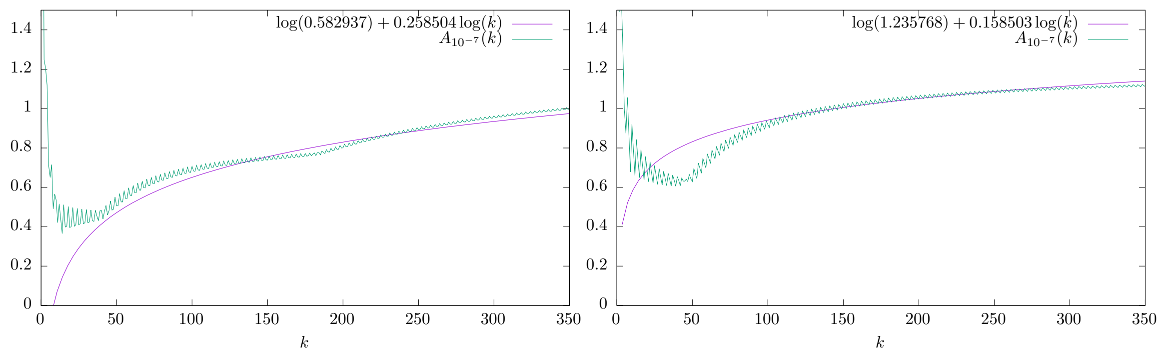

Figure 1.11: Comparison between $A_\rho(k)$, $\rho = 10^{-7}$, and its corresponding numerical fit. Left panel: values for the frequency $\omega = \frac{\sqrt{13}-1}{6}$. Right panel: values for the frequency $\omega = \frac{\sqrt{5}-1}{6}$

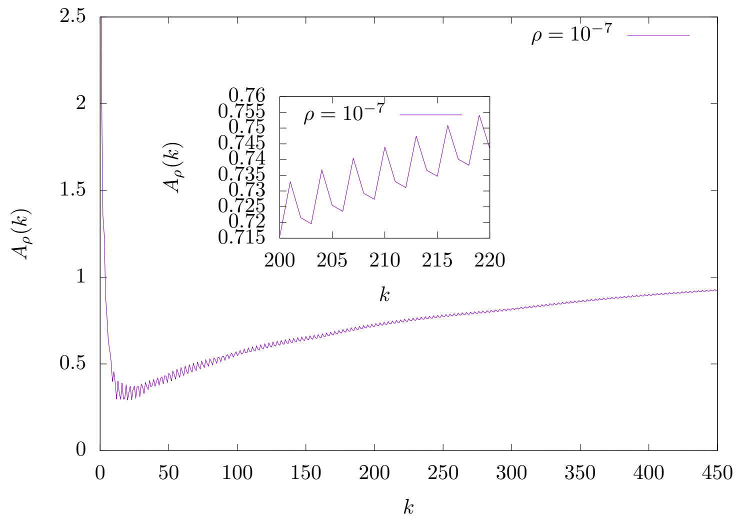

Figure 1.12: Graph of $A_\rho(k)$ for $\rho = 10^{-7}$, $\omega = \frac{\sqrt{5}-1}{2}$. The same oscillatory behavior is also present for $H^r(k)$.

ing equation (1.22) in which the coefficient of order k depends explicitly on the coefficient of order $k-3$. This is due to the power three of ε in the function b_ε. At the same time this phenomenon is independent of the frequency ω we choose. This gives an explanation of the appearance of an oscillating pattern observed in the inset of Figure 1.12 which appears for all the frequencies we considered. However, the computations show that the amplitude of the oscillations decreases as k grows and this oscillating effect fades away.

To study how the amplitude of the oscillations decreases we have *centralized* the oscillations by considering the differences between $A_\rho(k)$ and some moving averages. More precisely, denoting $a_k = A_\rho(k)$, $\rho = 10^{-7}$, we have considered the following *centralizations*

$$x_k = a_k - \frac{1}{5}\sum_{j=k-2}^{k+2} a_j, \qquad z_k = a_k - \frac{1}{3k}\sum_{j=k}^{k+2} j a_j, \tag{1.38}$$

Since the oscillations have period three, the *centralization* x_k is made by subtracting a moving average that captures two periods of the oscillation. The results for x_k are summarized in Figure 1.13 and Figure 1.14. For all the *centralizations* considered it is quite surprising that the amplitude of the oscillations seems to decrease as $k^{-\beta}$, with $\beta \approx 1$. Due to this behavior we consider a second *centralization*, z_k, which assumes that the oscillations decrease as k^{-1}. The results for z_k are summarized in Figure 1.15.

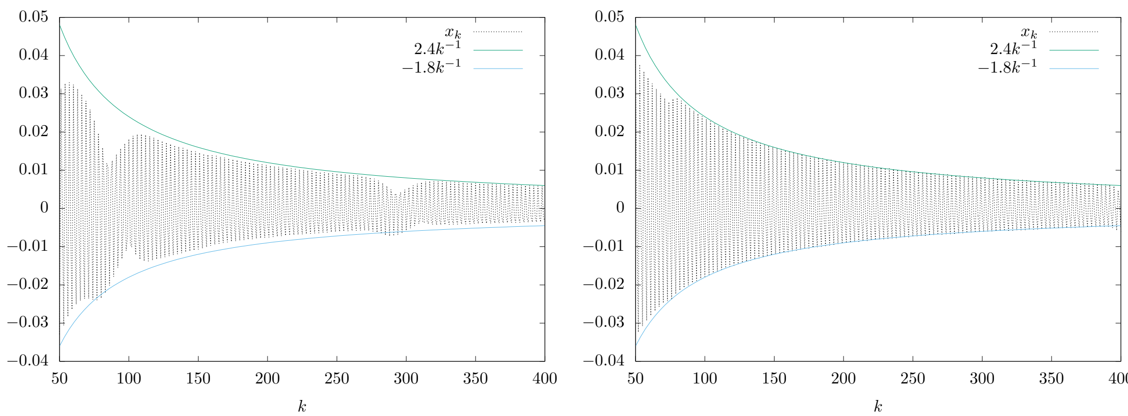

Figure 1.13: Plots of the centralization x_k. Left panel: Plot for the frequency $\frac{\sqrt{5}-1}{2}$. Right panel: Plot for frequency $\frac{\sqrt{3}-1}{2}$.

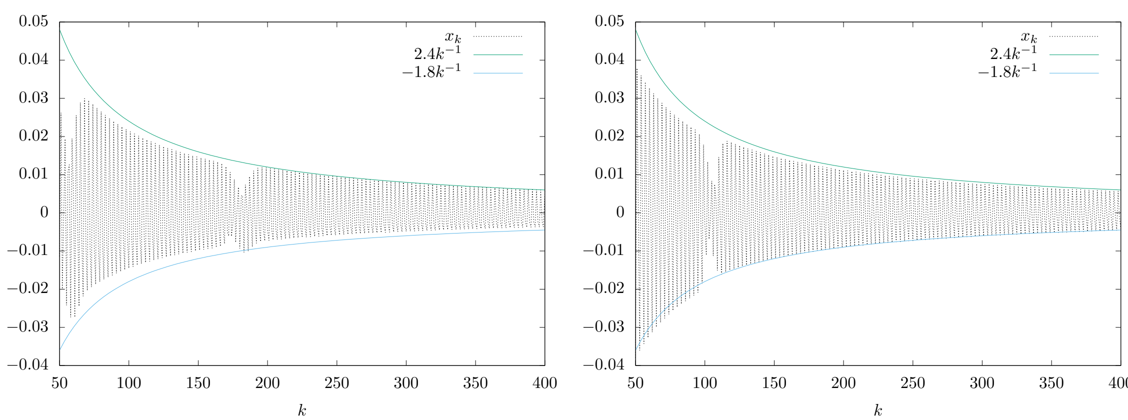

Figure 1.14: Plots oft the centralization x_k. Left panel: Plot for the frequency $\sqrt{2}$. Right panel: Plot for frequency $\sqrt{3}$.

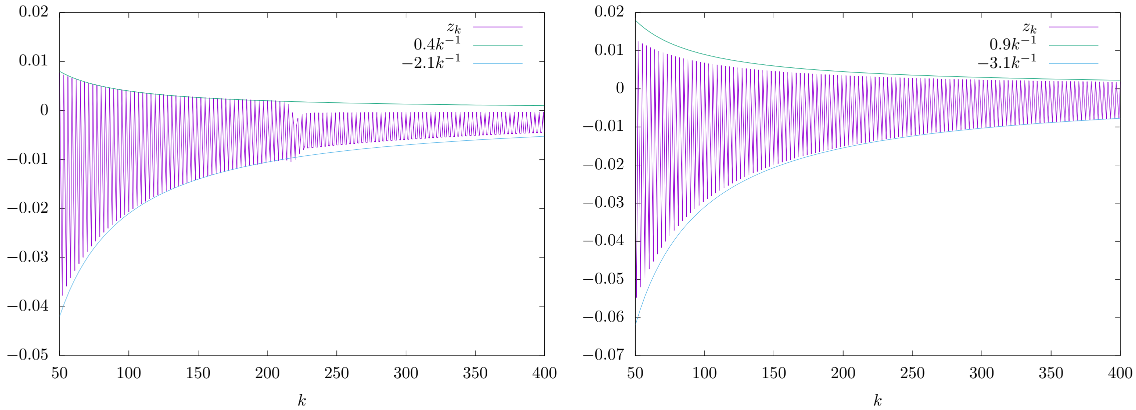

Figure 1.15: Plots of the centralization z_k. Left panel: Plot for the frequency $\frac{\sqrt{7}-1}{2}$. Right panel: Plot for frequency $\frac{\sqrt{5}-1}{6}$.

The results collected in the figures above suggest that the *centralizations* behave like $k^{-\beta} f(k)$ with f a periodic function. This observation motivates the following conjecture.

Conjecture 11. *Let* $A_\rho(k) = \frac{1}{k} \log \|u_k\|_\rho$*, then* $A_\rho(k) \approx \log(R) + \sigma \log(k) + k^{-\beta} f(k)$ *with* $\beta \approx 1$*,* $f(k)$ *a periodic function of period 3, and* $k \gg 1$.

1.5 Validation of the results

To validate the results described above we verified that the cohomology equation (1.22) is satisfied at every order with a suitable error. We also verified, as shown in Proposition 4, that the invariance equation (1.6) satisfies that $\log_{10}(\|E_{c^{\leq N}(\varepsilon)}[u_\varepsilon^{\leq N}]\|_\infty) \sim (N+1)\log_{10}(\varepsilon)$ as long as the error is above machine precision. We recall that $E_{c^{\leq N}(\varepsilon)}[u_\varepsilon^{\leq N}]$ means that we evaluate the operator E, given in (1.6), in the finite expansions $u_\varepsilon^{\leq N} = \sum_{k=1}^{N} u_k \varepsilon^k$ and $c^{\leq N}(\varepsilon) = \sum_{k=0}^{N} c_k \varepsilon^k$.

In Figure 1.16, we show the results of these computations.

For this work, the computations have been performed using 600 digits and 2^ℓ Fourier coefficients, with $10 \leq \ell \leq 13$. Using this precision we have verified that the coefficients u_n of the Lindstedt expansion have a relative error less than 10^{-300} when $n \leq 400$, see Figure 1.17. We have also checked that the functions u_n are trigonometric polynomials of degree n, as predicted in [3], up to an error less than 10^{-200} within the same range of

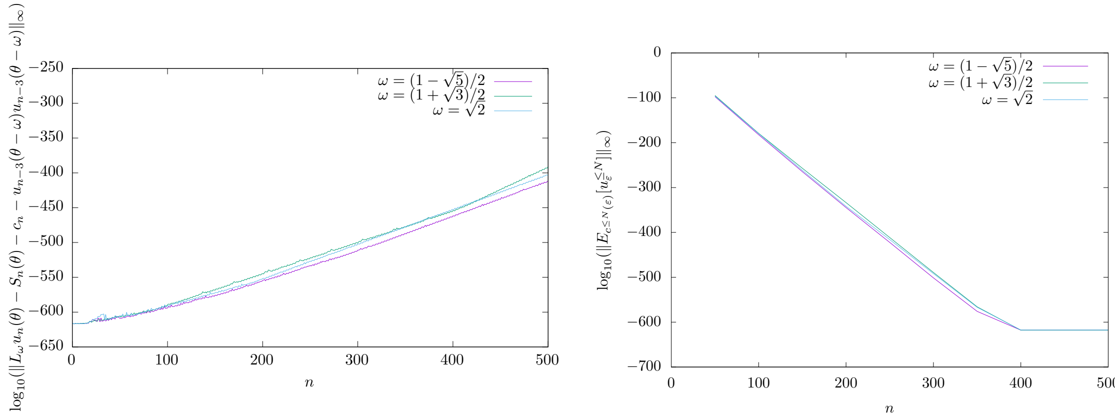

Figure 1.16: Left panel: Plot of $\log_{10}(\|L_\omega u_n(\theta) - S_n(\theta) - c_n - u_{n-3}(\theta-\omega) + u_{n-3}(\theta)\|_\infty)$ for different values of the frequency ω, $1 \leq n \leq 500$. Right panel: Plot of $\log_{10}(\||E_{c^{\leq N}(\varepsilon)}[u_\varepsilon^{\leq N}]\|_\infty)$, with $\epsilon = 10^{-2}$.

parameters. All the computations were done in pari/gp, [29].

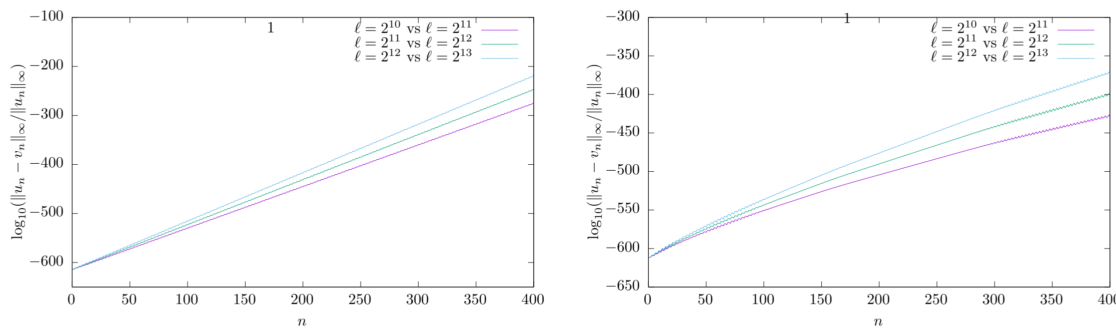

Figure 1.17: Left panel: Graphs of $\log_{10}(\|u_n - v_n\|_\infty)$ where u_n and v_n are the same coefficients of the Lindstedt series but computed using a different number, ℓ, of Fourier coefficients. Right panel: Graphs of the relative errors, $\log_{10}(\|u_n - v_n\|_\infty/\|u_n\|_\infty)$.

CHAPTER 2
GEVREY ESTIMATES FOR ASYMPTOTIC EXPANSIONS OF TORI IN WEAKLY DISSIPATIVE SYSTEMS

2.1 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 [30, 31] [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 [33, 34, 35, 36] for different studies of the zero dissipation limit in calculus of variations and in control.

When the friction is small, it is natural 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 of a fixed frequency are known to persist only if one can adjust parameters in the system [37, 38, 39]. As discussed very clearly in [40], the number of parameters needed is affected by the geometric properties of the systems considered.

In recent times, for some particular types of dissipative systems – the conformally symplectic systems, see Definition 12 – there is a very systematic KAM theory [41] based on geometric arguments. The examples mentioned above (Hamiltonian systems with friction

[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 [32]

proportional to the momentum and Euler-Lagrange equations of exponentially 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 Chapter is to study the singular perturbation theories in which the small parameter also introduces dissipation.

There are several studies of the singular perturbation theories in dissipation which are particularly relevant for us: The paper [5] shows that if one fixes a Diophantine frequency ω (see Definition 22), considers a Hamiltonian system – not necessarily integrable – with a quasi-periodic solution of frequency ω, and introduces a conformally symplectic perturbation (see Definition 12), 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 [5] 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 [1] studied numerically 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 [1] 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 19).

In this thesis, for some class of analytic maps (we require that the system is conformally symplectic and that the non-linearity is a trigonometric polynomial) we show that the conjecture in [1] 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 29.

The Gevrey class of functions 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 Thesis goes in a different direction. Even if we start with an analytic problem – indeed polynomial! – several objects of interest are only Gevrey. The phenomenon that analytic problems have only Gevrey solutions has appeared in other contexts in dynamics, notably in the study of singular perturbations [42], the regularity of attractors and fast-slow systems [43, 44, 45]. Closer to us, in dependence on parameters of solutions of non-linear problems, [46, 47], dependence of KAM tori in the frequency [48], or in the theory of parabolic manifolds [49, 50].

We note that showing that a perturbative expansion is Gevrey allows to obtain good bounds of the error of a finite sum [51]. It also allows the use of resummation methods to extract better results for the series, [52], and it gives insights on the analyticity domains. Indeed, in the Mathematical Physics literature, there has been considerable interest in the Gevrey nature of perturbation theories, often called *factorial bounds, Borel summability*, etc. [53, 54, 55]. We hope that introducing a new method for these questions can have interest in other motivations.

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 [41]. 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 [41] 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.

2.1.1 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

$$f_{\varepsilon,\mu_\varepsilon}(x,y) = (x + \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V'(x), \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V'(x)) \qquad (2.1)$$

In (2.1), 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 (2.1) is indeed conformally symplectic in the sense of Definition 12 (see below). The map (2.1) is the model that was used in the numerical experiments in [1].

Note that for $\varepsilon = 0$, the map (2.1) is integrable. The integrability of the map at $\varepsilon = 0$ does not play any role in the theoretical results in [5], the only assumption needed in [5] is that for $\varepsilon = 0$ the map is symplectic and has as an invariant torus. For the numerical study in [1], the fact that the map for $\varepsilon = 0$ is integrable leads to much more efficient algorithms. In this Thesis, 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 trigonometric polynomial (yet another reason to try to get rid of that hypothesis).

The main result of this Thesis, Theorem 29, 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 (2.1). The rigorous formulation of the main Theorem is given in Section 2.3, the statements of the main results can be better understood after some preliminary definitions and remarks are given (see Section 2.2). 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

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

where τ depends on the Diophantine type of ω (see Definition 22) and α is the order of the dissipation $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$.

The model (2.1) can be thought as a numerical time step – using a Verlet-like method – of the spin-orbit problem

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\mu y + \lambda + V'(x) \end{aligned} \tag{2.2}$$

2.1.2 Organization of the Chapter

The Chapter is organized as follows. In Section 2.2 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 2.3 we state Theorem 29 and Lemma 33, which are the main results of this Thesis and establish the Gevrey character of the perturbative expansions of the quasi periodic orbits.

The proof of Theorem 29 is based on a quasi Newton method. In Section 2.4 we formulate the iterative step of this Newton method, while in Section 2.5 we provide estimates for the corrections and the new error at one step of the method. Finally, in Section 2.6, using a KAM like argument, we give estimates for any step of the Newton like procedure and, with them, a proof of Lemma 33 is given establishing the Gevrey character of the perturbative expansions.

2.2 Preliminaries

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

2.2.1 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 12. *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 (2.1) is conformally symplectic with the conformal factor $\lambda(\varepsilon) = 1 - \varepsilon^\alpha$ and the standard symplectic form $\Omega = dx \wedge dy$ on the cylinder $\mathbb{T} \times \mathbb{R}$.

2.2.2 Banach spaces of analytic functions

Analytic functions on the torus

Given $\rho > 0$ we define the complex extension of the d-dimensional torus as

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

and denote $\mathcal{A}_\rho$ as the vector space of analytic functions defined $\operatorname{int}(\mathbb{T}^d_\rho)$ 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 physical 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 satisfy 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 13. *Let $\mathcal{B}$ be a Banach space. Given an analytic function $g : B_\gamma(0) \subseteq \mathbb{C} \longrightarrow \mathcal{B}$, and $n \ge 0$, we say $g(\varepsilon) \sim \mathcal{O}(|\varepsilon|^n)$ if and only if there exists $C > 0$ such that*

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

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 $g_k \in \mathcal{B}$.

Cauchy estimates.

We recall the classical Cauchy inequalities, see [56].

Lemma 14. *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 15. *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

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

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

Corollary 16. *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 17. *Note that the estimate in Corollary 16 only depends on a, associated with the order of the first term in the expansion of $\Delta^{(a,b]}$.*

2.2.3 Formal power series

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 [57], [58] 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 [37].

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 18. *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* $\rho < \rho_0$

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

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

$$\phi(\varepsilon) = \exp(-\varepsilon^{-2}) \tag{2.3}$$

has an identically zero asymptotic expansion on a domain

$$\mathcal{D}_\delta = \{\varepsilon : |\mathrm{Arg}(\varepsilon)| < \delta\} \tag{2.4}$$

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 2.3 in the domains $\mathcal{D}_\delta$ as in (2.4) when $\delta < \pi$, but not when $\delta > \pi$.

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 19. *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* $C \geq 0$ *,*

$R \geq 0$, and $n_0 \in \mathbb{N}$ such that

$$\|f_n\|_\rho \leq CR^n n^{\beta n} \quad \textit{for } n \geq n_0, \tag{2.5}$$

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 $C \geq 0$, $R \geq 0$, and $n_0 \in \mathbb{N}$ such that

$$|\mu_n| \leq CR^n n^{\beta n} \quad \textit{for } n \geq n_0, \tag{2.6}$$

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

Remark 20. *It is well known that* (2.5) *in Definition 19 is equivalent to the inequality*

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

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 [52], [58].

Remark 21. *The class of functions that around each point have expansions satisfying Definition 19 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 thesis 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 [42], the regularity of attractors and fast-

slow systems [43, 44, 45]. Closer to us, in dependence on parameters of solutions of non-linear problems, [46, 47], dependence of KAM tori in the frequency [48], or in the theory of parabolic manifolds [49, 50].

A property from number theory

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

Definition 22. *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 $\omega \in \mathcal{D}(\nu, \tau)$.

2.2.4 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

$$K(\theta + \omega) = f(K(\theta))\ \forall\theta \in \mathbb{T}^d \tag{2.7}$$

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 (2.7). The equation (2.7) will be our fundamental tool to characterize quasi-periodic orbits.

2.2.5 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 thesis (Theorem 33) will be presented in Section 2.3.

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 (2.1). 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 33, 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 22). 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 2.2.4 and, in particular (2.7), 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

$$f_{\varepsilon,\mu_\varepsilon} \circ K_\varepsilon(\theta) = K_\varepsilon(\theta + \omega) \tag{2.8}$$

Equation (2.8) 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 (2.8) is meant to hold is the meaning of formal power series (the coefficients of ε^n on both sides of (2.8) are identical for all n, as it is customary in the study of Lindstedt series).

Note that the equation (2.8) 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.

2.2.6 Automatic reducibility

As it is noted in [41], a very useful property of conformally symplectic systems is that solutions to equation (2.8) 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 23. *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*

$$M(\theta) = \left[DK(\theta) | J^{-1} \circ K(\theta) DK(\theta) \mathcal{N}(\theta)\right] \tag{2.9}$$

satisfies

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

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 23 is given in [41]. The argument is as follows, taking derivative in equation (2.8) one has $Df_\mu \circ K_0(\theta) DK_0(\theta) = DK_0(\theta+\omega)$ which gives the first column in (2.10). 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 (2.8) is Lagrangian. Then, using the conformally symplectic geometry the second column can be obtained.

Remark 24. *As it is pointed out in [41] if K is an approximate solution of* (2.8)*, that is,*

$$f_\mu \circ K(\theta) - K(\theta+\omega) =: E(\theta) \tag{2.11}$$

the relation (2.10) *will hold with an error, R, that can be estimated in terms of the error, $E(\theta)$, of the invariance equation, that is*

$$Df_\mu \circ K(\theta) M(\theta) = M(\theta+\omega) \begin{pmatrix} \mathrm{Id} & S(\theta) \\ 0 & \lambda\, \mathrm{Id} \end{pmatrix} + R(\theta), \tag{2.12}$$

with

$$\begin{aligned} S(\theta) &\equiv P(\theta+\omega)^\top Df \circ K(\theta) J^{-1} \circ K(\theta) P(\theta) - \mathcal{N}(\theta+\omega)^\top \Gamma(\theta+\omega)\mathcal{N}(\theta+\omega)\lambda \\ & \qquad (2.13) \\ P(\theta) &\equiv DK(\theta)\mathcal{N}(\theta), \\ \Gamma(\theta) &\equiv DK(\theta)^\top J^{-1} \circ K(\theta) DK(\theta). \end{aligned}$$

Moreover,

$$R(\theta) = \Big[DE(\theta) \Big| V(\theta+\omega)(\tilde{B}(\theta) - \lambda\, \mathrm{Id}) + DK(\theta+\omega)(\tilde{S}(\theta) - S(\theta)) \Big] \tag{2.14}$$

where

$$V(\theta) \equiv J^{-1} \circ K(\theta) DK(\theta) \mathcal{N}(\theta) \tag{2.15}$$

$$\tilde{B}(\theta) - \lambda\, \mathrm{Id} \equiv DK(\theta)^\top J \circ K(\theta) DK(\theta) \tilde{S}(\theta) \tag{2.16}$$

$$\tilde{S}(\theta) - S(\theta) \equiv -\mathcal{N}(\theta+\omega)^\top \Gamma(\theta+\omega) \mathcal{N}(\theta+\omega)(\tilde{B}(\theta) - \lambda\, \mathrm{Id}) \tag{2.17}$$

We note that $\tilde{B} - \mathrm{Id}$ *is estimated by the norm of* (2.11)*, thus* R *in* (2.14) *can be estimated by the norm of* (2.11) *as it is shown in Lemma 48. The derivation of the formulas in* (2.13)*,* (2.14)*, and* (2.15) *can be found in [41].*

Remark 25. *Observe that when considering* K_0*,* μ_0 *satisfying* (2.8) *and a perturbation* K_ε*,* μ_ε *(which could be given in terms of formal power series), equation* (2.12) *is also satisfied by* K_ε*,* μ_ε *but with all the expressions depending on* ε *(small enough), that is,*

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

2.3 Statement of the main result, Theorem 29

In this section we state the main result, Theorem 29, which gives the Gevrey character of the perturbative expansions of the solutions to equation (2.8). First we introduce a normalization which guarantees the uniqueness of the solutions to equation (2.8).

2.3.1 Normalization and local uniqueness

The centerpiece of this work is the invariance equation

$$f_{\varepsilon,\mu_\varepsilon} \circ K_\varepsilon = K_\varepsilon \circ T_\omega \tag{2.18}$$

where $T_\omega(\theta) = \theta + \omega$. Note that if (K, μ) is a solution of the invariant equation (2.18), then, for any $\sigma \in \mathbb{T}^d$, $(K \circ T_\sigma, \mu)$ is also a solution of (2.18), 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 26. *We say that a torus with embedding* K *is normalized with respect to* K_0 *when*

$$\int_{\mathbb{T}^d} \left[M_0^{-1}(\theta)(K(\theta) - K_0(\theta))\right]_d d\theta = 0 \tag{2.19}$$

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* (2.9)*.*

We also recall the following result ([41], Proposition 26) which shows that this condition can be imposed without loss of generality for solutions that are close to one another.

Proposition 27. *Let* K_0, K *be solutions of* (2.18) *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* (2.19)*. Furthermore,*

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

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

Remark 28. *As it is noted in [41] the normalization* (2.19) *works as well when* K *is only an approximate solution. Then, assuming that* K_0 *is a solution of equation* (2.18)*, the normalization condition* (2.19) *for an approximate solution of* (2.18) *given as power series expansion* $\sum_{n=0}^{\infty} K_n(\theta)\varepsilon^n$ *is equivalent to the conditions*

$$\int_{\mathbb{T}^d} \left[M_0^{-1}(\theta)K_n(\theta)\right]_d d\theta = 0 \tag{2.20}$$

for all $n \geq 1$*.*

2.3.2 Main Theorem

Here we present our main theorem, Theorem 29.

Theorem 29 (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

$$f_{\varepsilon,\mu_\varepsilon}(x,y) = (x + \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V'(x), \lambda(\varepsilon)y + \mu_\varepsilon - \varepsilon V'(x)) \tag{2.21}$$

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*

$$\left\| f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} - K_\varepsilon^{[\leq N]} \circ T_\omega \right\|_{\rho_0} \leq C_N |\varepsilon|^{N+1}. \tag{2.22}$$

where $C_N > 0$. *Moreover, if the* K_j*'s satisfy the normalization condition* (2.20)*, 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* (2.22) *and the normalization* (2.20) *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*

$$\|K_n\|_{\rho_0} \leq LF^n n^{(2\tau/\alpha)n} \quad \textit{and} \quad |\mu_n| \leq LF^n n^{(2\tau/\alpha)n} \qquad \textit{for any } n > N_0. \tag{2.23}$$

The proof of Theorem 29 is an easy consequence of Lemma 33. Proposition 66, given in the Appendix, shows the hypothesis of Lemma 33 are satisfied for maps of the form (2.21). Lemma 33 states the same results as Theorem 29 but in a more general setting.

Remark 30. *It is instructive to compare the results in Theorem 29 with the numerical explorations of Chapter 1 (see also [1], [2]). In the case that* $\lambda(\varepsilon) = 1 - \varepsilon^3$ *and* ω *is the golden mean, Theorem 29 gives that the expansion satisfies the Gevrey bounds with* *exponent 2/3. Of course, Theorem 29 gives only an upper bound and lower exponents could also be true. The numerical results in Chapter 1 ([1, 2]) lead to the conjecture that the expansion $\sum K_n \varepsilon^n$ has some well defined asymptotics*

$$\|K_n\|_\rho^{1/n} \approx C n^\sigma \tag{2.24}$$

with a slightly smaller Gevery exponent, $\sigma \approx 0.3$. The asymptotics (2.24) *is compatible with the results in Theorem 29, but suggests that the results in Theorem 29 are not optimal. Chapter 1 also presents several other patterns in the series (refined versions of* (2.24) *including oscillations of period* 3*, studies for other Diophantine numbers, etc.) We hope that the method presented in this thesis can lead to studies of these phenomena, hitherto discovered only through numerical implementation.*

We think that the argument in Theorem 29 can be 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" *[17], 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 [5] which assures the existence of formal power series expansions satisfying (2.18) up to any order for conformally symplectic systems.

Theorem 31 ([5], 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

$$f^*_{\varepsilon,\mu}\Omega = \lambda(\varepsilon)\Omega, \tag{2.25}$$

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

$$f_{0,\mu_0} \circ K_0 = K_0 \circ T_\omega. \tag{2.26}$$

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 (2.9)*, $(B_{0b})^0$ 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 (2.26) *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*

$$\left\| f_{\varepsilon, \mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} - K_\varepsilon^{[\leq N]} \circ T_\omega \right\|_{\rho'} \leq C_N |\varepsilon|^{N+1}. \tag{2.27}$$

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

Moreover, if we require the K_j's satisfy the normalization condition (2.20)*, then the expansions $K_\varepsilon^{[\infty]}$, $\mu_\varepsilon^{[\infty]}$ are unique.*

Note that Theorem 31 does not assume that the case $\varepsilon = 0$ is an integrable system, as it is the case for the map (2.21) , it suffices that the case $\varepsilon = 0$ is a Hamiltonian system with a KAM torus.

Remark 32. *Denoting*

$$E_\varepsilon^N(\theta) \equiv f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta) - K_\varepsilon^{[\leq N]}(\theta + \omega) \tag{2.28}$$

then (2.27) *can be written as*

$$\left\| E_\varepsilon^N \right\|_{\rho'} \leq C_N |\varepsilon|^{N+1}. \tag{2.29}$$

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 truncated series.

The following lemma, Lemma 33, can be considered as an improvement of Theorem 31 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 33 (Main Lemma)**.** *Assume the hypothesis of Theorem 31. 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* (2.9) *constructed from $K_\varepsilon^{[\leq N]}$.*

HTP2 *The $d \times d$ matrix*

$$\begin{aligned} \tilde{E}_{\Omega,\varepsilon}^N(\theta) &\equiv DK_\varepsilon^{[\leq N]}(\theta+\omega)^\top J \circ K_\varepsilon^{[\leq N]}(\theta+\omega) DK_\varepsilon^{[\leq N]}(\theta+\omega) \\ &- D(f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta))^\top J \circ (f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) D(f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) \end{aligned} \tag{2.30}$$

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 (2.27) *and* (2.20) *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*

$$\|K_n\|_{\rho_0} \leq LF^n n^{(2\tau/\alpha)n} \quad \text{and} \quad |\mu_n| \leq LF^n n^{(2\tau/\alpha)n} \qquad \text{for any } n > N_0. \tag{2.31}$$

The proof of Lemma 33, given in Section 2.6.2, 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 12) that maps like (2.21) satisfy.

We also point out that hypothesis **HTP1** and **HTP2** are very natural for the maps considered in Theorem 29. The verification of these hypothesis for the dissipative standard map is described in detail in Proposition 66 of the Appendix. In the general setting in which Lemma 33 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.

2.3.3 Asymptotic estimates for invariance functions

The formal power series studied in this thesis are asymptotic expansions of functions $K_\varepsilon, \mu_\varepsilon$ constructed in [5]. The functions $K_\varepsilon, \mu_\varepsilon$ are determined by the condition that they satisfy the invariance equation (2.18) and the normalization (2.20). 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 34.

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 (2.18) and the normalization (2.20). 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 [5], 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 [5] 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 [1] studies numerically the maximal domain of definition of the functions $K_\varepsilon, \mu_\varepsilon$ for the map (2.21) using a variety of methods including Pade summation and continuation methods. Indeed [1] conjectured that the series were Gevrey and this was an important motivation for this thesis.

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

$$\tilde{\nu} = \tilde{\nu}(\lambda; \omega, \tau) \equiv \sup_{k \in \mathbb{Z}^d \setminus \{0\}} |e^{2\pi i k \cdot \omega} - \lambda|^{-1} |k|^{-\tau} \tag{2.32}$$

one has

$$\mathcal{G} = \mathcal{G}(A; \omega, \tau, N) = \left\{ \varepsilon \in \mathbb{C} \ : \quad \tilde{\nu}(\lambda; \omega, \tau) |\lambda(\varepsilon) - 1|^{N+1} \leq A \right\}. \tag{2.33}$$

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 (2.18) 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 [5], and obtain a true solution of (2.18). Of course, the detailed implementation requires taking into account several other issues such as the absence of monodromy.

In this thesis, we will use a very similar technique. As as byproduct of the estimates used in the proof of Lemma 33, 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 [41].

More precisely we have:

Theorem 34. *Assuming the hypothesis of Lemma 33 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 33 satisfy*

$$\sup_{\varepsilon \in \mathcal{G}, |\varepsilon| \leq \tilde{\gamma}_{h+2}} \left\| \sum_{j=1}^{n} K_j \varepsilon^j - K_\varepsilon \right\|_{\rho_0 - \delta} \leq \left(U + V 2^{h(3\tau + 3d)} r^{n+1} r^{2^h N_0} \right) (CD)^h B^{h^2} r^{(2^h - 1) N_0} \left\| E^{N_0} \right\|_{\rho_0} \tag{2.34}$$

where $\hat{C}$ *and* C *are uniform constants and*
$U = \hat{C} \nu^{-1} \tilde{\nu}^{-1} \delta^{-2(\tau + d)}$, $V = C \nu^{-3} (a N_0)^{2\tau} \rho_0^{-(\tau + 3d)} 2^{2\tau + 6d}$,
$D = \nu^{-6} (a N_0)^{4\tau} \rho_0^{-(2\tau + 6d)} 2^{-(4\tau + 12d)}$, $r = 2^{-\tau/\alpha}$, $B = 2^{6\tau + 6d}$*, and*

$\tilde{\gamma}_h = (2^{-1}\nu)^{1/\alpha}(a2^h N_0)^{-\tau/\alpha}$.

Note that (2.34) 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 (2.34) are given explicitly. The proof of Theorem 34 is given in Section 2.6.3.

2.4 Iterative step of the quasi Newton method.

The KAM procedure for the proof of Theorem 33 is based on the application of a quasi Newton method, which is described in Section 2.4.2. 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 2.5.

2.4.1 Estimates for some cohomology equations

The iterative step described in Section 2.4.2 depends on the solution of two cohomology equations. The first equation, (2.35), is very standard in KAM theory. The estimate given in Lemma 35 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, (2.37), 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.

Standard cohomology equation

The first cohology equation we deal with is the following

$$\varphi_\varepsilon(\theta) - \varphi_\varepsilon(\theta + \omega) = \eta_\varepsilon(\theta) \tag{2.35}$$

Lemma 35 below, gives sufficient conditions to solve equation (2.35) and to obtain estimates of its solutions. These estimates are very standard in KAM theory.

Lemma 35. *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* (2.35)*, 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 (2.35) 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

$$\begin{aligned}
\|\varphi_\varepsilon\|_{\rho-\delta} &\leq \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \frac{|\hat{\eta_k}(\varepsilon)|}{|1-e^{2\pi i k\cdot\omega}|} \left\|e^{2\pi i k\cdot\theta}\right\|_{\rho-\delta} \\
&\leq \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \nu^{-1}|k|^\tau \|\eta_\varepsilon\|_\rho e^{-2\pi|k|\rho} e^{2\pi(\rho-\delta)|k|} \\
&\leq C\nu^{-1} \|\eta_\varepsilon\|_\rho \sum_{j\in\mathbb{N}} j^{\tau+d+1} e^{2\pi\delta j} \\
&\leq C\nu^{-1}\delta^{-(\tau+d)} \|\eta_\varepsilon\|_\rho. \qquad (2.36)
\end{aligned}$$

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 36. *Equation* (2.35) *appears very often in KAM theory. When $\varepsilon \in \mathbb{R}$, the paper [59] contains estimates with a better exponent on δ. That is, in the same situation of Lemma 35, when $\varepsilon \in \mathbb{R}$, one can get $\|\varphi_\varepsilon\|_{\rho-\delta} \leq C\nu\delta^{-\tau}\|\eta_\varepsilon\|_\rho$.*

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

$$\lambda(\varepsilon)\varphi_\varepsilon(\theta) - \varphi_\varepsilon(\theta + \omega) = \eta_\varepsilon(\theta) \tag{2.37}$$

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 38, solve equation (2.37) presents a small divisors problem. In this case the small divisors depend on the variable ε, that is, equation (2.37) 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 (2.37) (see [41]) 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 (2.37). One of the properties of the set $\mathcal{G}$ described in [41], 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 (2.37), to prove the Gevrey estimates in Lemma 33 we rely heavily on being able to obtain estimates of (2.37) 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 (2.37) whenever η_ε is a trigonometric polynomial. If the degree of the trigonometric polynomial, η_ε, is aN, Lemma 37 gives a relation between this degree and a domain in which the solution, φ_ε, of (2.37) will be analytic in ε.

Note that the requirement of hypothesis **HTP1** and **HTP2** in Lemma 33 is due to the fact that the quantities given in these hypothesis will be the right hand side of equations of the form (2.37).

Lemma 37. *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 38. *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 (2.37) *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 (2.37) 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 37 and Cauchy estimates, one obtains that for any $|\varepsilon| \leq \gamma_N$

$$\begin{aligned}
\|\varphi_\varepsilon\|_{\rho-\delta} &\leq \sum_{0<|k|\leq aN} \frac{|\hat{\eta}_k(\varepsilon)|}{|\lambda(\varepsilon) - e^{2\pi i k\cdot\omega}|} \left\| e^{2\pi i k\cdot\theta} \right\|_{\rho-\delta} \\
&\leq 2(aN)^\tau \nu^{-1} \sum_{0<|k|\leq aN} |\hat{\eta}_k(\varepsilon)| e^{2\pi|k|(\rho-\delta)} \\
&\leq 2(aN)^\tau \nu^{-1} \sum_{0<|k|\leq aN} \|\eta_\varepsilon\|_\rho e^{-2\pi|k|\rho} e^{2\pi|k|(\rho-\delta)} \\
&\leq 2(aN)^\tau \nu^{-1} \|\eta_\varepsilon\|_\rho \sum_{j=1}^{aN} j^{d-1} e^{-2\pi j\delta} \\
&\leq C\nu^{-1}(aN)^\tau \delta^{-d} \|\eta_\varepsilon\|_\rho
\end{aligned} \tag{2.38}$$

Thus, $\|\varphi\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^\tau\delta^{-d} \|\eta\|_{\rho,\gamma_N}$. The last claim comes from (2.38), that is $\varphi_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$ if $\eta_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^k\right)$. □

2.4.2 Formulation of the quasi Newton method

Every step of the quasi Newton method starts with a solution of equation (2.18) 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 (2.20) 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 39. *The first step of the Newton method could start with $K^{[\leq N_0]}$, $\mu^{[\leq N_0]}$, given by Theorem 31, 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 (2.18) 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

$$\left[Df_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right] \Delta_\varepsilon - \Delta_\varepsilon \circ T_\omega + \left[D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right] \sigma_\varepsilon = -E_\varepsilon^N. \tag{2.39}$$

Equation (2.39) 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 [41], we will not solve (2.39) exactly but we will find approximate solutions that will reduce quadratically the error. The idea is to approximate the solution of (2.39) using the geometric identities introduced in Section 2.2.6. Considering the change of variables

$$\Delta_\varepsilon = M_\varepsilon^{[\leq N]} W_\varepsilon, \tag{2.40}$$

where $M_\varepsilon^{[\leq N]}$ is as in (2.9) computed from $K_\varepsilon^{[\leq N]}$. Using (2.12) one obtains that (2.39) is equivalent to

$$M_\varepsilon^{[\leq N]} \circ T_\omega \left[\begin{pmatrix} \mathrm{Id} & S_\varepsilon^{[\leq N]} \\ 0 & \lambda(\varepsilon)\,\mathrm{Id} \end{pmatrix} W_\varepsilon - W_\varepsilon \circ T_\omega\right] + \left(D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right) \sigma_\varepsilon = -E_\varepsilon^N - R_\varepsilon^{[\leq N]} W_\varepsilon \tag{2.41}$$

where $R_\varepsilon^{[\leq N]}$ is the error (2.14) and $S_\varepsilon^{[\leq N]}$ is given in (2.13), both computed from $K_\varepsilon^{[\leq N]}$. That is

$$M_\varepsilon^{[\leq N]} \equiv \left[DK_\varepsilon^{[\leq N]} \;\middle|\; J^{-1} \circ K_\varepsilon^{[\leq N]} DK_\varepsilon^{[\leq N]} \mathcal{N}_\varepsilon^{[\leq N]}\right] \sim \mathcal{O}(|\varepsilon|^0) \tag{2.42}$$

$$S_\varepsilon^{[\leq N]} \equiv P_\varepsilon^{[\leq N]\top} Df_{\mu_\varepsilon^{[\leq N]},\varepsilon} \circ K_\varepsilon^{[\leq N]} J^{-1} \circ K_\varepsilon^{[\leq N]} P_\varepsilon^{[\leq N]} - \lambda(\varepsilon) \mathcal{N}_\varepsilon^{[\leq N]\top} \Gamma_\varepsilon^{[\leq N]} \mathcal{N}_\varepsilon^{[\leq N]} \sim \mathcal{O}(|\varepsilon|^0) \tag{2.43}$$

$$\mathcal{N}_\varepsilon^{[\leq N]} \equiv \left[\left(DK_\varepsilon^{[\leq N]} \right)^\top DK_\varepsilon^{[\leq N]} \right]^{-1} \sim \mathcal{O}(|\varepsilon|^0), \tag{2.44}$$

$$P_\varepsilon^{[\leq N]} \equiv DK_\varepsilon^{[\leq N]} \mathcal{N}_\varepsilon^{[\leq N]},$$

$$\Gamma_\varepsilon^{[\leq N]} \equiv DK_\varepsilon^{[\leq N]T} J^{-1} \circ K_\varepsilon^{[\leq N]} DK_\varepsilon^{[\leq N]} \tag{2.45}$$

Since we expect both W_ε and $R_\varepsilon^{[\leq N]}$ to be estimated by E_ε^N, see (2.66) and (2.76), the term $W_\varepsilon R_\varepsilon^{[\leq N]}$ is quadratic in E_ε^N, thus, we expect that omitting this term in (2.41) 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 (2.37) 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

$$M_\varepsilon^{[\leq N]} \circ T_\omega \left[\begin{pmatrix} \mathrm{Id} & S_\varepsilon^{[\leq N]} \\ 0 & \lambda(\varepsilon)\,\mathrm{Id} \end{pmatrix} W_\varepsilon - W_\varepsilon \circ T_\omega \right] + \left(D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} \right) \sigma_\varepsilon = -E_\varepsilon^{(N,2N]} \tag{2.46}$$

Remark 40. *The election of the truncation* $E_\varepsilon^{(N,2N]}$ *in* (2.46) *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* (2.1)*, will satisfy hypothesis* **HTP1** *and* **HTP2** *in Lemma 33 due to the fact that the truncation is made. See appendix A.*

In order to construct a solution of equation (2.46), we follow a similar approach as in [41]. Defining

$$\tilde{E}_\varepsilon^{(N,2N]} := \left(M_\varepsilon^{[\leq N]} \circ T_\omega \right)^{-1} E_\varepsilon^{(N,2N]} \sim \mathcal{O}(|\varepsilon|^{N+1}) \tag{2.47}$$

$$\tilde{A}_\varepsilon^N := \left(M_\varepsilon^{[\leq N]} \circ T_\omega\right)^{-1} D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} \sim \mathcal{O}(|\varepsilon|^0) \tag{2.48}$$

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 (2.46) can be written in components as

$$W_{\varepsilon,1} - W_{\varepsilon,1} \circ T_\omega = -S_\varepsilon^{[\leq N]} W_{\varepsilon,2} - \tilde{E}_{\varepsilon,1}^{(N,2N]} - \tilde{A}_{\varepsilon,1}^N \sigma_\varepsilon \tag{2.49}$$

$$\lambda(\varepsilon) W_{\varepsilon,2} - W_{\varepsilon,2} \circ T_\omega = -\tilde{E}_{\varepsilon,2}^{(N,2N]} - \tilde{A}_{\varepsilon,2}^N \sigma_\varepsilon \tag{2.50}$$

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

$$\begin{aligned} 0 &= -\overline{S_\varepsilon^{[\leq N]}}\,\overline{W_{\varepsilon,2}} - \overline{S_\varepsilon^{[\leq N]}(W_{\varepsilon,2})^0} - \overline{\tilde{E}_{\varepsilon,1}^{(N,2N]}} - \overline{\tilde{A}_{\varepsilon,1}^N}\sigma_\varepsilon \\ \varepsilon^3 \overline{W_{\varepsilon,2}} &= -\overline{\tilde{E}_{\varepsilon,2}^{(N,2N]}} - \overline{\tilde{A}_{\varepsilon,2}^N}\sigma_\varepsilon \end{aligned} \tag{2.51}$$

$$\begin{aligned} (W_{\varepsilon,1})^0 - (W_{\varepsilon,1})^0 \circ T_\omega &= -(S_\varepsilon^{[\leq N]} W_{\varepsilon,2})^0 - (\tilde{E}_{\varepsilon,1}^{(N,2N]})^0 - (\tilde{A}_{\varepsilon,1}^N)^0 \sigma_\varepsilon \\ \lambda(\varepsilon)(W_{\varepsilon,2})^0 - (W_{\varepsilon,2})^0 \circ T_\omega &= -(\tilde{E}_{\varepsilon,2}^{(N,2N]})^0 - (\tilde{A}_{\varepsilon,2}^N)^0 \sigma_\varepsilon. \end{aligned} \tag{2.52}$$

In order to uncouple systems (2.51) and (2.52) we consider $(W_{\varepsilon,2})^0$ as an affine function of σ_ε, due to (2.52). That is,

$$(W_{\varepsilon,2})^0 = (B_{a,\varepsilon})^0 + (B_{b,\varepsilon})^0 \sigma_\varepsilon \tag{2.53}$$

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

$$\lambda(\varepsilon)(B_{a,\varepsilon})^0 - (B_{a,\varepsilon})^0 \circ T_\omega = -(\tilde{E}^{(N,2N]}_{\varepsilon,2})^0 \tag{2.54}$$

$$\lambda(\varepsilon)(B_{b,\varepsilon})^0 - (B_{b,\varepsilon})^0 \circ T_\omega = -(\tilde{A}^{N}_{\varepsilon,2})^0. \tag{2.55}$$

Due to **HTP1**, and applying Lemma 38, equations (2.54) and (2.55) can be solved and we can get estimates in balls with center at $\varepsilon = 0$. Once that (2.54) and (2.55) are solved, and using (2.53), system (2.51) can be written as

$$\begin{pmatrix} \overline{S^{[\leq N]}_{\varepsilon}} & \overline{S^{[\leq N]}_{\varepsilon}(B_{b,\varepsilon})^0} + \overline{\tilde{A}^{N}_{\varepsilon,1}} \\ \varepsilon^3 \,\mathrm{Id} & \overline{\tilde{A}^{N}_{\varepsilon,2}} \end{pmatrix} \begin{pmatrix} \overline{W_{\varepsilon,2}} \\ \sigma_\varepsilon \end{pmatrix} = \begin{pmatrix} -\overline{S^{[\leq N]}_{\varepsilon}(B_{a,\varepsilon})^0} - \overline{\tilde{E}^{(N,2N]}_{\varepsilon,1}} \\ -\overline{\tilde{E}^{(N,2N]}_{\varepsilon,2}} \end{pmatrix} \tag{2.56}$$

Remark 41. *Due to* **HND** *in Theorem 31 the matrix in the left hand side of* (2.56) *is invertible at $\varepsilon = 0$. By the continuity of the determinant, equation* (2.56) *can be solved for ε small enough and the inverse is analytic in ε.*

Thus, (2.53) and (2.56) 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

$$(W_{\varepsilon,1})^0 - (W_{\varepsilon,1})^0 \circ T_\omega = -(S^{[\leq N]}_{\varepsilon} W_{\varepsilon,2})^0 - (\tilde{E}^{(N,2N]}_{\varepsilon,1})^0 - (A^{N}_{\varepsilon,1})^0 \sigma_\varepsilon, \tag{2.57}$$

which can be done due to Lemma 35. To fulfill the normalization condition (2.20) and obtain uniqueness of the coefficients of the perturbative expansions, $\overline{W}_{\varepsilon,1}$ is chosen as

$$\overline{W}_{\varepsilon,1} = -\left(\int_{\mathbb{T}^d} \left[M_0^{-1}(\theta) DK^{[\leq N]}_{\varepsilon}\right]_d d\theta\right)^{-1} \int_{\mathbb{T}^d} \left[M_0^{-1}(\theta)\left(DK^{[\leq N]}_{\varepsilon}(W_{\varepsilon,1})^0 + V^{[\leq N]}_{\varepsilon} W_{\varepsilon,2}\right)\right]_d d\theta \tag{2.58}$$

where $V^{[\leq N]} = J^{-1} \circ K^{[\leq N]}_{\varepsilon} DK^{[\leq N]}_{\varepsilon} \mathcal{N}^{[\leq N]}_{\varepsilon}$ is the second *column* of the matrix $M^{[\leq N]}_{\varepsilon}$, see Remark 28.

Remark 42. *Assuming that $K^{[\leq N]}_{\varepsilon}$ satisfies the normalization* (2.20)*, then the new approx-*

imation $K_\varepsilon^{[\leq N]} + \Delta_\varepsilon$ *will satisfy* (2.20) *if the correction satisfies*

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

Since $\Delta_\varepsilon = M_\varepsilon^{[\leq N]}W_\varepsilon = DK_\varepsilon^{[\leq N]}W_{\varepsilon,1} + V_\varepsilon^{[\leq N]}W_{\varepsilon,2} = DK_\varepsilon^{[\leq N]}\left((W_{\varepsilon,1})^0 + \overline{W_{\varepsilon,1}}\right) + V_\varepsilon^{[\leq N]}W_{\varepsilon,2}$, (2.58) *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

$$\Delta_\varepsilon = M_\varepsilon^{[\leq N]}W_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right) \quad \text{and} \quad \sigma_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right). \tag{2.59}$$

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

$$\Delta_\varepsilon^{(N,2N]} \equiv \sum_{n=N+1}^{2N}\Delta_n\varepsilon^n \quad \text{and} \quad \sigma_\varepsilon^{(N,2N]} \equiv \sum_{n=N+1}^{2N}\sigma_n\varepsilon^n. \tag{2.60}$$

Therefore, the new approximation is chosen as

$$K_\varepsilon^{[\leq 2N]} := K_\varepsilon^{[\leq N]} + \Delta_\varepsilon^{(N,2N]} \quad \text{and} \quad \mu_\varepsilon^{[\leq 2N]} := \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon^{(N,2N]}. \tag{2.61}$$

Remark 43. *Notice that, due to Lemma 38, the solutions of* (2.54) *and* (2.55) *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,* (2.56) *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* $\Delta_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$.

2.4.3 Algorithm for the iterative step

The procedure described above leads Algorithm 44 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 44. *Given $K_\varepsilon^{[\leq N]} : \mathbb{T}^n \to \mathcal{M}$, $\mu_\varepsilon^{[\leq N]} \in \mathbb{R}^d$. We perform the following computations:*

(1) $E_\varepsilon^N \leftarrow f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]} - K_\varepsilon^{[\leq N]} \circ T_\omega$

(2) $E_\varepsilon^{(N,2N]}$ *obtained from* E_ε^N *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) $P_\varepsilon \leftarrow \alpha_\varepsilon \mathcal{N}_\varepsilon$

$$\Gamma_\varepsilon \leftarrow \alpha_\varepsilon^\top J^{-1} \circ K_\varepsilon^{[\leq N]} \alpha_\varepsilon$$

$$S_\varepsilon \leftarrow (P_\varepsilon \circ T_\omega)^\top Df_{\mu_\varepsilon^{[\leq N]},\varepsilon} \circ K_\varepsilon^{[\leq N]} J^{-1} \circ K_\varepsilon^{[\leq N]} P_\varepsilon - \lambda(\varepsilon)(\mathcal{N}_\varepsilon \circ T_\omega)^\top \Gamma_\varepsilon \circ T_\omega (\mathcal{N}_\varepsilon \circ T_\omega)$$

$$\tilde{A}_\varepsilon \leftarrow \beta_\varepsilon D_\mu f_{\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}$$

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

$(B_{b,\varepsilon})^0$ *solves* $\lambda(\varepsilon)(B_{b,\varepsilon})^0 - (B_{b,\varepsilon})^0 \circ T_\omega = -(\tilde{A}_{\varepsilon,2})^0$

(11) *Find* $\overline{W_{\varepsilon,2}}, \sigma_\varepsilon$ *by solving*

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

(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) $(W_{\varepsilon,1})^0$ *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) $K_\varepsilon^{[\leq 2N]} \leftarrow K_\varepsilon^{[\leq N]} + \Delta_\varepsilon^{(N,2N]}$

$\mu_\varepsilon^{[\leq 2N]} \leftarrow \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon^{(N,2N]}$

It is worth to note that all the operations in Algorithm 44 could be implemented in a few lines in a high level computer language.

Remark 45. *Note that Algorithm 44 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 46. *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 33, 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 38 after $\log_2(N)$ steps to obtain estimates on the N coefficient.

Remark 47. *To iterate the modified Newton method in Algorithm 44 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 56.*

2.5 Estimates for the iterative step.

In this section we present the estimates for the corrections given by the Newton step described in Section 2.4, these estimates are obtained by following the steps in Algorithm

44. 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 (2.37) which, accordingly with Lemma 38, can be solved if

$$\varepsilon \leq \gamma_N := \left(\frac{\nu}{2}\right)^{1/\alpha} \frac{1}{(aN)^{\tau/\alpha}}. \tag{2.62}$$

where aN is the degree of the trigonometric polynomial in the right hand side of (2.37).

2.5.1 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 (2.14) computed from $K_\varepsilon^{[\leq N]}$. The estimates are obtained by studying qualitatively the geometric identities introduced in Section 2.2.6 and taking into account the uniformity on the variable ε.

Lemma 48. *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) $K_\varepsilon^{[\leq N]}\left(\mathbb{T}_\rho^d\right) \subset \mathrm{Domain}(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}})$ *and that there exist* $\xi \geq 0$ *such that*

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

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

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)

$$\nu^{-1}(aN)^{\tau}\delta^{-(d+1)}\left\|E^N\right\|_{\rho,\gamma_N} \ll 1 \tag{2.63}$$

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

$$\begin{aligned} E^N_{\Omega,\varepsilon}(\theta) &\equiv DK_\varepsilon^{[\leq N]}(\theta+\omega)^\top J \circ K_\varepsilon^{[\leq N]}(\theta+\omega) DK_\varepsilon^{[\leq N]}(\theta+\omega) \\ &- D(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta))^\top J \circ (f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) D(f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}(\theta)) \end{aligned} \tag{2.64}$$

is a trigonometric polynomial of degree less than aN.

Then

$$R_\varepsilon^{[\leq N]} \sim \mathcal{O}(|\varepsilon|^{N+1}) \tag{2.65}$$

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

$$\left\|R^{[\leq N]}\right\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^{\tau}\delta^{-(d+1)}\left\|E^N\right\|_{\rho,\gamma_N} \tag{2.66}$$

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 (2.14) yields

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

with

$$V_\varepsilon^{[\leq N]}(\theta) \equiv J^{-1} \circ K_\varepsilon^{[\leq N]}(\theta) DK_\varepsilon^{[\leq N]}(\theta)\mathcal{N}_\varepsilon^{[\leq N]}(\theta) \tag{2.67}$$

$$B_\varepsilon(\theta) - \lambda(\varepsilon)\,\mathrm{Id} \equiv -E^N_{L,\varepsilon}(\theta+\omega) S_\varepsilon^{[\leq N]}(\theta) \tag{2.68}$$

$$\tilde{S}_\varepsilon(\theta) - S_\varepsilon^{[\leq N]}(\theta) \equiv -\mathcal{N}_\varepsilon^{[\leq N]}(\theta+\omega)^\top \Gamma_\varepsilon^{[\leq N]}(\theta+\omega)\mathcal{N}_\varepsilon^{[\leq N]}(\theta+\omega)\left(B_\varepsilon(\theta) - \lambda(\varepsilon)\,\mathrm{Id}\right) \tag{2.69}$$

where

$$E^N_{L,\varepsilon}(\theta) \equiv DK^{[\leq N]}_\varepsilon(\theta)^\top J \circ K^{[\leq N]}_\varepsilon(\theta) DK^{[\leq N]}_\varepsilon(\theta) \tag{2.70}$$

is the pull back $(K^{[\leq N]}_\varepsilon)^*\Omega$ written in coordinates and $\Gamma^{[\leq N]}_\varepsilon$ as in (2.45). We recall that J is the matrix associated to the symplectic form, see Section 2.2. It is easy to estimate the first column of $R^{[\leq N]}_\varepsilon$ using Cauchy estimates, that is

$$\left\| DE^N_\varepsilon \right\|_{\rho-\delta} \leq C\delta^{-1} \left\| E^N_\varepsilon \right\|_\rho$$

To obtain estimates for the second column of $R^{[\leq N]}_\varepsilon$, due to (2.68) and (2.69), it is enough to get estimates of E^N_L . The estimate for E^N_L is obtained using that $f^*_{\varepsilon,\mu}\Omega = \lambda(\varepsilon)\Omega$. Note that $E^N_{\Omega,\varepsilon} = (K^{[\leq N]}_\varepsilon \circ T_\omega)^*\Omega - (f_{\varepsilon,\mu^{[\leq N]}_\varepsilon} \circ K^{[\leq N]}_\varepsilon)^*\Omega$ in coordinates and, since $(f_{\varepsilon,\mu^{[\leq N]}_\varepsilon} \circ K^{[\leq N]}_\varepsilon)^*\Omega = \lambda(K^{[\leq N]}_\varepsilon)^*\Omega$, we have that E^N_L satisfies the equality

$$E^N_{L,\varepsilon} \circ T_\omega - \lambda(\varepsilon) E^N_{L,\varepsilon} = E^N_{\Omega,\varepsilon}. \tag{2.71}$$

Then, by Lemma 38 and **HTP2** we obtain

$$\left\| E^N_L \right\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^\tau \delta^{-d} \left\| E^N_\Omega \right\|_{\rho-\delta/2,\gamma_N}. \tag{2.72}$$

To get estimates for E^N_Ω, we follow [41]. If h and g are smooth maps with range in $\mathcal{M}$, the matrix corresponding to $h^*\Omega - g^*\Omega$ is

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

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

$$\left\| E^N_{\Omega,\varepsilon} \right\|_{\rho-\delta/2} \leq C\delta^{-1} \left\| E^N_\varepsilon \right\|_\rho \tag{2.73}$$

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

$$\left\|R^{[\leq N]}\right\|_{\rho-\delta,\gamma_N} \leq C\nu^{-1}(aN)^\tau \delta^{-(d+1)} \left\|E^N\right\|_{\rho,\gamma_N}. \tag{2.74}$$

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

Remark 49. *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 Lagrangian manifold, see [41]. 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^N_L in* (2.70) *represents the error of the lagrangian character of $K^{[\leq N]}_\varepsilon$, then, if E^N_L is small enough the spaces Range$(DK^{[\leq N]}_\varepsilon(\theta))$ and Range$(J^{-1} \circ K^{[\leq N]}_\varepsilon(\theta) DK^{[\leq N]}_\varepsilon(\theta))$ will be transversal and the matrix $M^{[\leq N]}_\varepsilon$ will define a linear isomorphism. This transversality will be obtained if* (2.63) *is satisfied and it is given by* (2.72) *and* (2.73).

2.5.2 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 44. First, Lemma 50, 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 51.

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

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

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

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

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

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

iii) **HTP1** *For any* $N \in \mathbb{N}$*, the matrices* $(\tilde{E}^{(N,2N]}_{\varepsilon,2})^0$ *and* $(\tilde{A}^N_{\varepsilon,2})^0$ *defined in* (2.47) *and* (2.48)*, are trigonometric polynomials of degree less or equal than* aN*.*

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

$$W_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right), \quad \sigma_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right) \tag{2.75}$$

$$\|W\|_{\rho-\delta, r\gamma_N} \leq C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N} \tag{2.76}$$

and

$$\sup_{|\varepsilon|\le r\gamma_N}|\sigma_\varepsilon| \leq C\nu^{-1}(aN)^{\tau}\delta^{-d}\frac{r^{N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N} \tag{2.77}$$

where $C = C(d, \left\|DK^{[\le N]}\right\|_{\rho,\gamma_N}, \left\|M^{[\le N]}\right\|_{\rho,\gamma_N}, \left\|(M^{[\le N]})^{-1}\right\|_{\rho,\gamma_N}, \left\|\mathcal{N}^{[\le N]}\right\|_{\rho,\gamma_N}, \mathcal{T}^N)$ *and* $\mathcal{T}^N$ *is defined in* (2.81).

Proof. Given that $(\tilde{E}^{(N,2N]}_{\varepsilon,2})^0$ and $(\tilde{A}^N_{\varepsilon,2})^0$ are trigonometric polynomials, by Lemma 38, (2.54), and (2.55); B_a and B_b satisfy the following estimates

$$\begin{aligned} \|B_a\|_{\rho-\delta,r\gamma_N} &\leq C\nu^{-1}(aN)^\tau\delta^{-d}\left\|\tilde{E}_2^{(N,2N]}\right\|_{\rho,r\gamma_N} \\ &\leq C\nu^{-1}(aN)^\tau\delta^{-d}\left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N} \end{aligned} \tag{2.78}$$

and similarly

$$\|B_b\|_{\rho-\delta,r\gamma_N} \leq C\nu^{-1}(aN)^\tau\delta^{-d}\left\|A^N\right\|_{\rho,r\gamma_N}. \tag{2.79}$$

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 (2.56) we have

$$\begin{pmatrix} \overline{W_{\varepsilon,2}} \\ \sigma_\varepsilon \end{pmatrix} = \begin{pmatrix} \overline{S_\varepsilon^{[\leq N]}} & \overline{S_\varepsilon^{[\leq N]}(B_{b,\varepsilon})^0} + \overline{\tilde{A}_{\varepsilon,1}^N} \\ \varepsilon^3\,\mathrm{Id} & \overline{\tilde{A}_{\varepsilon,2}^N} \end{pmatrix}^{-1} \begin{pmatrix} -\overline{S_\varepsilon^{[\leq N]}(B_{a,\varepsilon})^0} - \overline{\tilde{E}_{\varepsilon,1}^{(N,2N]}} \\ -\overline{\tilde{E}_{\varepsilon,2}^{(N,2N]}} \end{pmatrix}, \tag{2.80}$$

denoting

$$\mathcal{T}_\varepsilon^N := \left\| \begin{pmatrix} \overline{S_\varepsilon^{[\leq N]}} & \overline{S_\varepsilon^{[\leq N]}(B_{b,\varepsilon})^0} + \overline{\tilde{A}_{\varepsilon,1}^N} \\ \varepsilon^3\,\mathrm{Id} & \overline{\tilde{A}_{\varepsilon,2}^N} \end{pmatrix}^{-1} \right\| \qquad \text{and} \qquad \mathcal{T}^N = \sup_{|\varepsilon|\leq r\gamma_N} \mathcal{T}_\varepsilon^N \tag{2.81}$$

from (2.80) we have

$$|\sigma_\varepsilon|, \left|\overline{W_{\varepsilon,2}}\right| \leq \mathcal{T}_\varepsilon^N \left(\left| \overline{S_\varepsilon^{[\leq N]}(B_{a,\varepsilon})^0} + \overline{\tilde{E}_{\varepsilon,1}^{(N,2N]}} \right| + \left| \overline{\tilde{E}_{\varepsilon,2}^{(N,2N]}} \right| \right) \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right) \tag{2.82}$$

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 $\tilde{E}_\varepsilon^{(N,2N]} \sim \mathcal{O}\left(|\varepsilon|^{N+1}\right)$.

Thus

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

for any $0 < \delta < \rho$. Thus, using (2.47) and (2.78) we obtain

$$\sup_{|\varepsilon| \leq r\gamma_N} \left|\overline{W_{\varepsilon,2}}\right| \leq C\nu^{-1}(aN)^{\tau}\delta^{-d} \left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N} \tag{2.83}$$

$$\sup_{|\varepsilon| \leq r\gamma_N} |\sigma_\varepsilon| \leq C\nu^{-1}(aN)^{\tau}\delta^{-d} \left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N}. \tag{2.84}$$

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

$$\begin{aligned} \left\|(W_2)^0\right\|_{\rho-\delta,r\gamma_N} &\leq \left\|(B_a)^0\right\|_{\rho-\delta,r\gamma_N} + \sup_{|\varepsilon| \leq r\gamma_N} |\sigma| \left\|(B_b)^0\right\|_{\rho-\delta,r\gamma_N} \\ &\leq C\nu^{-1}(aN)^{\tau}\delta^{-d}\left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N} + C\nu^{-2}(aN)^{2\tau}\delta^{-2d}\left\|A^N\right\|_{\rho,r\gamma_N}\left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N}, \\ &\leq C\nu^{-2}(aN)^{2\tau}\delta^{-2d}\left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N}. \end{aligned} \tag{2.85}$$

Thus, combining (2.83) and (2.85) we get

$$\|W_2\|_{\rho-\delta,r\gamma_N} \leq C\nu^{-2}(aN)^{2\tau}\delta^{-2d} \left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N} \tag{2.86}$$

The estimates for $(W_1)^0$ come from (2.57) and Lemma 35, i.e.,

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

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

that is,

$$\left\|(W_1)^0\right\|_{\rho-2\delta,r\gamma_N} \leq C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)} \left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N}. \tag{2.87}$$

Finally, the estimate for $\overline{W_1}$ comes from (2.58), that is

$$\begin{aligned}\sup_{|\varepsilon|\le r\gamma_N} \left|\overline{W_{\varepsilon,1}}\right| &\le C\left(\|(W_1)^0\|_{\rho-\delta,r\gamma_N} + \|W_2\|_{\rho-\delta,r\gamma_N}\right) \\ &\le C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\left\|E^{(N,2N]}\right\|_{\rho,r\gamma_N}. \end{aligned} \tag{2.88}$$

Putting together (2.86), (2.87), (2.88), and using the Cauchy estimates in Corollary 16 yields the claimed estimate for W. □

Corollary 51. *Assuming the hypothesis of Lemma 48 and Lemma 50, for any* $0 < \delta < \rho$ *and* $0 < r < 1$ *we have*

$$\left\|\Delta^{(N,2N]}\right\|_{\rho-\delta,r\gamma_N} \le C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} \tag{2.89}$$

$$\sup_{|\varepsilon|\le r\gamma_N}\left|\sigma_\varepsilon^{(N,2N]}\right| \le C\nu^{-1}(aN)^{\tau}\delta^{-d}\frac{r^{N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} \tag{2.90}$$

Moreover,

$$\left\|\Delta^{(2N,\infty]}\right\|_{\rho-\delta,r\gamma_N} \le C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} \tag{2.91}$$

$$\sup_{|\varepsilon|\le r\gamma_N}\left|\sigma_\varepsilon^{(2N,\infty]}\right| \le C\nu^{-1}(aN)^{\tau}\delta^{-d}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} \tag{2.92}$$

Proof. Using the Cauchy estimates as in Corollary 16 and the estimates in Lemma 50 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. □

2.5.3 Non-linear estimates for the quasi-Newton method.

The quasi-Newton procedure in Algorithm 44 can also be described using a convenient operator notation. Defining the error functional

$$\mathcal{E}[K_\varepsilon,\mu_\varepsilon] = f_{\varepsilon,\mu_\varepsilon}\circ K_\varepsilon - K_\varepsilon\circ T_\omega \tag{2.93}$$

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

$$\mathcal{E}[K+\Delta,\mu+\sigma] = \mathcal{E}[K,\mu] + D_1\mathcal{E}[K,\mu]\Delta + D_2\mathcal{E}[K,\mu]\sigma + \mathcal{R}[\Delta,\sigma;K,\mu] \tag{2.94}$$

where the Frechet derivatives are given by

$$D_1\mathcal{E}[K_\varepsilon,\mu_\varepsilon]\Delta_\varepsilon = \left(Df_{\varepsilon,\mu_\varepsilon}\circ K_\varepsilon\right)\Delta_\varepsilon - \Delta_\varepsilon\circ T_\omega \tag{2.95}$$

$$D_2\mathcal{E}[K_\varepsilon,\mu_\varepsilon]\sigma_\varepsilon = \left(D_\mu f_{\varepsilon,\mu_\varepsilon}\circ K_\varepsilon\right)\sigma_\varepsilon \tag{2.96}$$

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

$$\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}] + D_1\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}]\Delta_\varepsilon^{(N,2N]} + D_2\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}]\sigma_\varepsilon^{(N,2N]} = 0. \quad (2.97)$$

As it was explained before, in Section 2.4.2, the corrections we construct with Algorithm 44 do not satisfy (2.97) but they solve an approximate equation (2.46). The following Lemmas give estimates for the error functional evaluated in the corrected unknowns. First, Lemma 54, we give estimates for the error $\mathcal{E}[K^{[\leq N]} + \Delta, \mu^{[\leq N]} + \sigma]$ and then, using Cauchy estimates, we obtain the estimates for the error evaluated in the truncated corrections, $\mathcal{E}[K^{[\leq N]} + \Delta^{(N,2N]}, \mu^{[\leq N]} + \sigma^{(N,2N]}]$, Proposition 56.

Remark 52. *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* (2.93) *are well defined. In particular Δ and σ need to satisfy $\|\Delta\|, |\sigma| \leq \xi$ and we need to choose the domain loss. In Section 2.6, Lemma 60, 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 53. *Assume $0 < r < 1$ and $0 < \delta \leq \rho$. Then, under the hypothesis of Lemma 48 and Lemma 50 one has*

$$\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}] + D_1\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}]\Delta_\varepsilon + D_2\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}]\sigma_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{2N+1}\right) \quad (2.98)$$

and

$$\begin{aligned} &\left\| \mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}] + D_1\mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}]\Delta + D_2\mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}]\sigma \right\|_{\rho-\delta, r\gamma_N} \\ &\qquad \leq \frac{r^{2N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N} + C\nu^{-4}(aN)^{3\tau}\delta^{-(\tau+4d+1)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|^2_{\rho,\gamma_N} \end{aligned} \quad (2.99)$$

Proof. Note that with the operator notation introduced at the beginning of this section we have $\mathcal{E}(K^{[\leq N]}, \mu^{[\leq N]}) = E^N$. Using (2.12) and taking into account that $\Delta_\varepsilon = M_\varepsilon^{[\leq N]} W_\varepsilon$ and that W_ε satisfies (2.46) we have

$$\begin{aligned}
&\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}] + D_1\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}]\Delta_\varepsilon + D_2\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}]\sigma_\varepsilon \\
&\quad = E_\varepsilon^N + \left(Df_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right)\Delta_\varepsilon - \Delta_\varepsilon \circ T_\omega + \left(D_\mu f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right)\sigma_\varepsilon \qquad (2.100) \\
&\qquad - R_\varepsilon^{[\leq N]}\left(M_\varepsilon^{[\leq N]}\right)^{-1}\Delta_\varepsilon + R_\varepsilon^{[\leq N]}\left(M_\varepsilon^{[\leq N]}\right)^{-1}\Delta_\varepsilon \\
&\quad = E_\varepsilon^N + M_\varepsilon^{[\leq N]} \circ T_\omega \begin{pmatrix} \mathrm{Id} & S_\varepsilon^{[\leq N]} \\ 0 & \lambda(\varepsilon)\,\mathrm{Id} \end{pmatrix} \left(M_\varepsilon^{[\leq N]}\right)^{-1}\Delta_\varepsilon - \Delta_\varepsilon \circ T_\omega \qquad (2.101) \\
&\qquad + \left(D_\mu f_{\varepsilon,\mu^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}\right)\sigma_\varepsilon + R_\varepsilon^{[\leq N]}\left(M_\varepsilon^{[\leq N]}\right)^{-1}\Delta_\varepsilon \\
&\quad = E_\varepsilon^N - E_\varepsilon^{(N,2N]} + R_\varepsilon^{[\leq N]} W_\varepsilon \qquad (2.102) \\
&\quad = E_\varepsilon^{(2N,\infty]} + R_\varepsilon^{[\leq N]} W_\varepsilon \sim \mathcal{O}\left(|\varepsilon|^{2N+1}\right)
\end{aligned}$$

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 $E^{(2N,\infty]}$, (2.65), and (2.75).

Then, using the Cauchy estimates of Corollary 16, Lemma 48, and Lemma 50 one obtains

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

□

Lemma 54. *Assume $0 < r < 1$ and $0 < \delta \leq \rho$. Then, under the hypothesis of Lemma 50 and Lemma 48 we have*

$$\mathcal{E}(K_\varepsilon^{[\leq N]} + \Delta_\varepsilon, \mu_\varepsilon^{[\leq N]} + \sigma_\varepsilon) \sim \mathcal{O}\left(|\varepsilon|^{2N+1}\right) \qquad (2.103)$$

and

$$\left\|\mathcal{E}[K^{[\leq N]}+\Delta,\mu^{[\leq N]}+\sigma]\right\|_{\rho-\delta,r\gamma_N}\leq\frac{r^{2N+1}}{1-r}\left\|E^N\right\|_{\rho,\gamma_N}+C\nu^{-6}(aN)^{4\tau}\delta^{-(2\tau+6d)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|^2_{\rho,\gamma_N} \tag{2.104}$$

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

Proof. Note that $\mathcal{R}[K_\varepsilon^{[\leq N]},\mu_\varepsilon^{[\leq N]},\Delta_\varepsilon,\sigma_\varepsilon]$ in (2.94) can be estimated using Taylor estimates for the remainder, that is

$$\left\|\mathcal{R}_\varepsilon\right\|_\rho\leq C\left(\left\|\Delta_\varepsilon\right\|^2_\rho+|\sigma_\varepsilon|^2\right) \tag{2.105}$$

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 $\left\|D^2f_{\mu^{[\leq N]}}\circ K^{[\leq N]}\right\|_{\rho,\gamma_N}$, $\left\|D^2_\mu f_{\mu^{[\leq N]}}\circ K^{[\leq N]}\right\|_{\rho,\gamma_N}$ to be close to $\left\|D^2f_{\mu^{[\leq N_0]}}\circ K^{[\leq N_0]}\right\|_{\rho_0,\gamma_{N_0}}$, $\left\|D^2_\mu f_{\mu^{[\leq N_0]}}\circ K^{[\leq N_0]}\right\|_{\rho_0,\gamma_{N_0}}$, at the first step of the iterations. For now, we assume that C is uniform constant. In Section 2.6, Lemma 60, 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 (2.105) yields $\mathcal{R}_\varepsilon\sim\mathcal{O}\left(|\varepsilon|^{2N+2}\right)$. This, together with (2.98), gives (2.103). Moreover, taking sup with respect to ε one obtains

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

where in the third line we use the inequalities in Lemma 50. Finally, this inequality, Lemma 53, and (2.94) 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 55. *Assuming the hypothesis of Lemma 48 and Lemma 50, for any* $0 < \delta < \rho$ *and* $0 < r < 1$ *we have*

$$\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}] + D_1\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}]\Delta_\varepsilon^{(N,2N]} + D_2\mathcal{E}[K_\varepsilon^{[\leq N]}, \mu_\varepsilon^{[\leq N]}]\sigma_\varepsilon^{(N,2N]} \sim \mathcal{O}\left(|\varepsilon|^{2N+1}\right) \tag{2.106}$$

and

$$\begin{aligned} &\left\|\mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}] + D_1\mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}]\Delta^{(N,2N]} + D_2\mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}]\sigma^{(N,2N]}\right\|_{\rho-\delta, r\gamma_N} \\ &\quad \leq C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\frac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N} + C\nu^{-4}(aN)^{3\tau}\delta^{-(\tau+4d+1)}\frac{r^{N+1}}{1-r}\left\|E^N\right\|^2_{\rho,\gamma_N} \end{aligned} \tag{2.107}$$

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 $E^N = \mathcal{E}[K^{[\leq N]}, \mu^{[\leq N]}]$, then, using the linearity of the Frechet derivatives one obtains

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

which implies (2.106). Moreover, using the relation above and the estimates in Lemma 53 and Lemma 51 one gets

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

□

Proposition 56. *Assuming the hypothesis of Lemma 48 and Lemma 50, for any $0<\delta<\rho$ and $0<r<1$ we have*

$$\mathcal{E}\left[K_\varepsilon^{[\leq N]}+\Delta_\varepsilon^{(N,2N]},\mu_\varepsilon^{[\leq N]}+\sigma_\varepsilon^{(N,2N]}\right]\sim\mathcal{O}\left(|\varepsilon|^{2N+1}\right) \tag{2.108}$$

and

$$
\begin{aligned}
&\left\|\mathcal{E}[K^{[\leq N]}+\Delta^{(N,2N]},\mu^{[\leq N]}+\sigma^{(N,2N]}]\right\|_{\rho-\delta,r\gamma_N} \qquad (2.109)\\
&\quad\leq C\nu^{-3}(aN)^{2\tau}\delta^{-(\tau+3d)}\tfrac{r^{\frac{3}{2}N+1}}{(1-r^{1/2})^2}\left\|E^N\right\|_{\rho,\gamma_N}+C\nu^{-6}(aN)^{4\tau}\delta^{-(2\tau+6d)}\tfrac{r^{N+1}}{(1-r^{1/2})^4}\left\|E^N\right\|^2_{\rho,\gamma_N}
\end{aligned}
$$

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 (2.108) follows from using the same argument as in the proof of

Lemma 54. We also have

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

Combining this estimate with (2.107) in Lemma 55 one gets (2.109). □

2.6 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 60 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.

2.6.1 The iterative procedure.

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

$$\delta_h := \frac{\rho_0}{2^{h+2}} \quad \text{and} \quad \rho_{h+1} := \rho_h - \delta_h \geq \frac{\rho_0}{2} \quad \text{for } h \geq 1, \tag{2.110}$$

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 31. Since at any step we double the number of coefficients of the Lindstedt expansions, we have,

$$N_h := 2^h N_0 \tag{2.111}$$

and

$$\tilde{\gamma}_h := \gamma_{N_h} = \left(\frac{\nu}{2}\right)^{1/\alpha} \frac{1}{(aN_h)^{\tau/\alpha}} = \left(\frac{\nu}{2}\right)^{1/\alpha} \frac{1}{(a2^h N_0)^{\tau/\alpha}} \tag{2.112}$$

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

$$\tilde{\gamma}_{h+1} = 2^{-\tau/\alpha}\tilde{\gamma}_h. \tag{2.113}$$

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

$$K_h := K^{[\leq N_0]} + \Delta^{(N_0,N_1]} + \cdots + \Delta^{(N_{h-1},N_h]} \qquad \mu_h := \mu^{[\leq N_0]} + \sigma^{(N_0,N_1]} + \cdots + \sigma^{(N_{h-1},N_h]}. \tag{2.114}$$

Furthermore, denoting

$$\Delta_h := \Delta^{(N_h,N_{h+1}]} \quad \text{and} \quad \sigma_h := \sigma^{(N_h,N_{h+1}]} \quad \text{for } h \geq 0 \tag{2.115}$$

we have that, for $h \geq 0$

$$K_{h+1} = K_h + \Delta_h \quad \text{and} \quad \mu_{h+1} = \mu_h + \sigma_h. \tag{2.116}$$

Finally, denote also

$$e_h := \|\mathcal{E}[K_h, \mu_h]\|_{\rho_h,\tilde{\gamma}_h} = \left\|E^{N_h}\right\|_{\rho_h,\tilde{\gamma}_h} \tag{2.117}$$

$$d_h := \|\Delta_h\|_{\rho_{h+1},\tilde{\gamma}_{h+1}} \tag{2.118}$$

$$v_h := \|D\Delta_h\|_{\rho_{h+1},\tilde{\gamma}_{h+1}} \tag{2.119}$$

$$s_h := \sup_{|\varepsilon| \leq \tilde{\gamma}_{h+1}} |\sigma_h(\varepsilon)|. \tag{2.120}$$

Remark 57. *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 31.

Note that with this new notation the estimates in Corollary 51 can be written as

$$d_h \le \hat{C}_h \nu^{-3} (aN_h)^{2\tau} \delta_h^{-(\tau+3d)} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} e_h \tag{2.121}$$

$$v_h \le \hat{C}_h \nu^{-3} (aN_h)^{2\tau} \delta_h^{-(\tau+3d+1)} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} e_h \tag{2.122}$$

$$s_h \le \hat{C}_h \nu^{-1} (aN_h)^{\tau} \delta_h^{-d} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} e_h \tag{2.123}$$

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 (2.109) given in Proposition 56 implies

$$e_{h+1} \le \tilde{C}_h \nu^{-6} (aN_h)^{4\tau} \delta_h^{-(2\tau+6d)} \left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} \left(e_h + e_h^2\right) \tag{2.124}$$

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 58. *In the following we will denote C a constant depending on $\nu, \tau, d, \xi, \rho_0, |J^{-1}|$; 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 60 , we give smallness conditions so that $C_h \le C$ for every $h \ge 0$. Since we are working with expansions near to $(K^{[\le N_0]}, \mu^{[\le N_0]})$ 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* (2.42), (2.44), *and* (2.81).

Considering this uniform constant C on (2.124), and taking N_0 sufficiently large, yields $e_h < 1$ for any $h > 0$, and inequality (2.124) implies

$$e_{h+1} \leq C\nu^{-6}(aN_h)^{4\tau}\delta_h^{-(2\tau+6d)}\left(\frac{1}{2^{\tau/\alpha}}\right)^{N_h} e_h. \tag{2.125}$$

Remark 59. *Due to Remark 58 and the definitions of* δ_h, ρ_h, N_h, *and* $\tilde{\gamma}_h$; *the inequality* (2.125) *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

$$e_{h+1} \leq CDB^h r^{2^h N_0} e_h \tag{2.126}$$

where

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

Lemma 60. *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 61. *Note that by* (2.29) *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,

$$\begin{aligned}
\sup_{|\varepsilon|\le\tilde{\gamma}_{H+1}} |\mu_{H+1}-\mu_0| &= \sup_{|\varepsilon|\le\tilde{\gamma}_{H+1}} \left|\sigma^{(N_0,N_1]}+...+\sigma^{(N_H,N_{H+1}]}\right| \\
&\le \sum_{j=0}^{H} s_j \le \sum_{j=0}^{H} \hat{C}_j \nu^{-1}(aN_j)^\tau \delta_j^{-d} r^{N_j} e_j \\
&\le \sum_{j=0}^{H} C\nu^{-1}(a2^jN_0)^\tau \rho_0^{-d} 2^{(j+2)d} r^{2^jN_0} \left((CD)^j B^{j^2} r^{(2^j-1)N_0} e_0\right) \\
&\le C\nu^{-1}(aN_0)^\tau \rho_0^{-d} 2^{2d} \sum_{j=0}^{H} (2^{\tau+d})^j (CD)^j B^{j^2} r^{(2^{j+1}-1)N_0} e_0 \\
&\le C\nu^{-1}(aN_0)^\tau \rho_0^{-d} 2^{2d} \sum_{j=0}^{H} (2^{\tau+d})^j (CD)^j B^{j^2} r^{2^jN_0} e_0 \\
&\le C\nu^{-1}(aN_0)^\tau \rho_0^{-d} 2^{2d} e_0 \sum_{j=0}^{H} \left(2^{\tau+d} CDBr^{N_0}\right)^j \\
&\le C\nu^{-1}(aN_0)^\tau 2^d \rho_0^{-d} e_0 \\
&\le \ell_\mu N_0^\tau e_0.
\end{aligned}$$

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 (2.126) given in Remark 59, which is a consequence of the nonlinear estimates given in Lemma 56, that is

$$e_{h+1} = \|\mathcal{E}(K_h+\Delta_h, \mu_h+\sigma_h)\|_{\rho_{h+1},\tilde{\gamma}_{h+1}} \le CDB^h r^{2^hN_0} e_h \tag{2.127}$$

where D, B, and r are as in Remark 59. 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

$$\|\mathcal{N}_h - \mathcal{N}_0\|_{\rho_h,\tilde{\gamma}_h} \leq \overline{C} \, \|DK_h - DK_0\|_{\rho_h,\tilde{\gamma}_h} \tag{2.128}$$

$$\|M_h - M_0\|_{\rho_h,\tilde{\gamma}_h} \leq \overline{C} \, \|DK_h - DK_0\|_{\rho_h,\tilde{\gamma}_h} \tag{2.129}$$

$$\left\|M_h^{-1} - M_0^{-1}\right\|_{\rho_h,\tilde{\gamma}_h} \leq \overline{C} \, \|DK_h - DK_0\|_{\rho_h,\tilde{\gamma}_h} \tag{2.130}$$

$$|\mathcal{T}_h - \mathcal{T}_0| \leq \overline{C} \, \|DK_h - DK_0\|_{\rho_h,\tilde{\gamma}_h} \tag{2.131}$$

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 (2.42), (2.44), (2.43), (2.81). Then,

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

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. □

2.6.2 Proof of main Lemma

For the proof of the main Lemma, Lemma 33 , we inherit all the notation introduced throughout this section.

Proof. Note that Theorem 31 assures the existence of the Lindstedt series satisfying (2.6.2). 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 (2.20).

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 48 and 50. 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 60 can be applied and this allows us to iterate the quasi Newton method described in Algorithm 44. That is, we can construct the unique formal power series as follows

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

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

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 (2.113). Before giving the detailed computations, note that $\tilde{\gamma}_h \sim (2^h N_0)^{-\tau/\alpha}$ and if $n \in \left(2^h N_0, 2^{h+1} N_0\right] \cap \mathbb{N}$ then

$$(\tilde{\gamma}_h)^{-n} \sim (2^h N_0)^{C(\tau/\alpha) 2^h N_0} \sim n^{C(\tau/\alpha)n}.$$

Using this together with Cauchy estimates is expected to yield the Gevrey estimates. More precisely, if $n \in \left(2^h N_0, 2^{h+1} N_0\right] \cap \mathbb{N}$, using Cauchy estimates, (2.121), and $(p2;h)$ we have

$$
\begin{aligned}
\|K_n\|_{\frac{\rho_0}{2}} &\le (\tilde\gamma_{h+1})^{-n}\,\|\Delta_h\|_{\frac{\rho_0}{2},\tilde\gamma_{h+1}}\\
&\le (\tilde\gamma_{h+1})^{-n}\,\|\Delta_h\|_{\rho_{h+1},\tilde\gamma_{h+1}}\\
&\le (r^{h+1}\tilde\gamma_0)^{-n} d_h\\
&\le (r^{h+1}\tilde\gamma_0)^{-2^{h+1}N_0}\hat C_h \nu^{-3}(aN_h)^{2\tau}\delta_h^{-(\tau+3d)} r^{N_h} e_h\\
&\le (r^{h+1}\tilde\gamma_0)^{-2^{h+1}N_0} C\nu^{-3}\left(a2^hN_0\right)^{2\tau}\rho_0^{-(\tau+3d)}2^{(2\tau+6d)}2^{(\tau+3d)h}r^{2^hN_0}(CD)^hB^{h^2}r^{(2^h-1)N_0}e_0\\
&\le C\nu^{-3}\rho_0^{-(\tau+3d)}2^{(2\tau+6d)}(aN_0)^{2\tau}e_0(2^{3\tau+3d}CD)^hB^{h^2}(\tilde\gamma_0)^{-2^{h+1}N_0}r^{(-(h+1)2^{h+1}+2^{h+1}-1)N_0}\\
&\le C\nu^{-3}\rho_0^{-(\tau+3d)}2^{(2\tau+6d)}(aN_0)^{2\tau}e_0(2^{3\tau+3d}CD)^hB^{h^2}(2^{1/\alpha}\nu^{-1/\alpha}(aN_0)^{\tau/\alpha})^{2^{h+1}N_0}r^{-(h2^{h+1}+1)N_0}\\
&\le \hat L\left(2^{3\tau+3d}CDB2^{2/\alpha}\nu^{-2/\alpha}a^{2\tau/\alpha}\right)^{2^hN_0}(N_0^{2\tau/\alpha})^{2^hN_0}(2^{\tau/\alpha})^{(h2^{h+1}+1)N_0}\\
&\le \hat L 2^{(\tau/\alpha)N_0}F^{2^hN_0}(N_0^{2\tau/\alpha})^{2^hN_0}(2^{2\tau/\alpha})^{h2^hN_0}\\
&\le LF^{2^hN_0}(2^hN_0)^{(2\tau/\alpha)2^hN_0}\\
&\le LF^n n^{(2\tau/\alpha)n}
\end{aligned}
$$

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. □

2.6.3 Proof of Theorem 34

Proof. Inheriting the notation from Lemma 60, consider N_0 sufficiently large such that the a-posteriori theorem, Theorem 14 in [5], can be applied. That is, N_0 such that

$$\sup_{\varepsilon\in\mathcal{G},|\varepsilon|\le\tilde\gamma_0}\left\|E_\varepsilon^{N_0}\right\|_\rho \le \hat C\left(\nu\tilde\nu(\lambda;\omega,\tau)\right)^2\delta^{-4(\tau+\delta)}. \tag{2.132}$$

where $\tilde{\nu}(\lambda;\omega,\tau)$ is defined in (2.32). Then, following the discussion in Section (2.3.3) and applying the a-posteriori theorem, Theorem 14 in [5], one obtains

$$\sup_{\varepsilon\in\mathcal{G},|\varepsilon|\leq\tilde{\gamma}_{h+2}}\left\|K_\varepsilon^{[\leq 2^h N_0]}-K_\varepsilon\right\|_{\rho_0-\delta}\leq\hat{C}\nu^{-1}\tilde{\nu}(\lambda;\omega,\tau)^{-1}\delta^{-2(\tau+d)}\sup_{\varepsilon\in\mathcal{G},|\varepsilon|\leq\tilde{\gamma}_{h+2}}\left\|E_\varepsilon^{2^h N_0}\right\|_{\rho_0}$$

where $\mathcal{G}$ is defined in (2.33).

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

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

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

Appendices

APPENDIX A

EXPLICIT COMPUTATION FOR THE DISSIPATIVE STANDARD MAP

A.1 Verifying trigonometric polynomial hypothesis, HTP1 and HTP2, 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

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

Where $V(x)$ is a trigonometric polynomial. In this section we verify that maps like (A.1) satisfy **HTP1** and **HTP2** of Lemma 33. 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

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

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

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

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

$$\sum_{k=0}^{2} (L_\omega u_k(\theta) - \mu_k)\,\varepsilon^k - (L_\omega u_3(\theta) - \mu_3 + u_0(\theta) - u_0(\theta - \omega) - \omega)\,\varepsilon^3$$
$$+ \sum_{k=4}^{\infty} (L_\omega u_k(\theta) - \mu_k + u_{k-3}(\theta) - u_{k-3}(\theta - \omega))\,\varepsilon^k = -\sum_{k=1}^{\infty} S_{k-1}(\theta)\varepsilon^k \quad \text{(A.5)}$$

Remark 62. *When $V(\theta)$ is a trigonometric polynomial, the coefficients S_n can be com-*

puted 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 $V(\theta) = \sum_{|k|\leq a} \hat{f}_k e^{2\pi k\theta} = \sum_{|k|\leq a} V_k(\theta)$ 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

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

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

Remark 63. *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 64. *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 poly-*

nomial 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 65. *If $V(\theta)$, in* (A.1)*, is a trigonometric polynomial of degree a, then for any* *fixed ε the sum $\sum_{n=0}^{N} u_n(\theta)\varepsilon^n$ is a trig polynomial of degree aN in θ.*

Note that in this case

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

and using equation (A.5) we have

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

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, $\tilde{A}^N_{\varepsilon,2}$, of the vector $\tilde{A}^N_\varepsilon = \left(M_\varepsilon^{[\leq N]} \circ T_\omega\right)^{-1} D_\mu f_{\varepsilon,\mu_\varepsilon^{[\leq N]}} \circ K_\varepsilon^{[\leq N]}$ 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 33 are satisfied for the dissipative standard map.

Proposition 66. *For any $N \in \mathbb{N}$, if $V(\theta)$ in* (A.1) *is a trigonometric polynomial of degree a, then $\tilde{E}^{(N,2N]}_{\varepsilon,2}$ is a trigonometric polynomial of degree $3aN$, $\tilde{A}^N_{\varepsilon,2}$ is a trig polynomial of degree aN, and*

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

is a trigonometric polynomial of degree $2aN$.

Proof. It is only left to prove the last claim. Note that $\tilde{E}^N_{\Omega,\varepsilon}(\theta)$ 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

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

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. $\square$

A.1.1 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.

REFERENCES

[1] A. P. Bustamante and R. C. Calleja, "Computation of domains of analyticity for the dissipative standard map in the limit of small dissipation," *Physica D: Nonlinear Phenomena*, vol. 395, pp. 15–23, 2019.

[2] ——, "Corrigendum and Addendum to Computation of domains of analyticity for the dissipative standard map in the limit of small dissipation [Physica D 395 (2019) 15–23]," *Physica D: Nonlinear Phenomena*, 2021.

[3] A. P. Bustamante and R. de la Llave, "Gevrey estimates for asymptotic expansions of tori in weakly dissipative systems," 2020, preprint, https://arxiv.org/abs/2010.06006.

[4] A. Celletti and L. Chierchia, "Measures of basins of attraction in spin-orbit dynamics," *Celestial Mech. Dynam. Astronom.*, vol. 101, no. 1-2, pp. 159–170, 2008.

[5] R. C. Calleja, A. Celletti, and R. de la Llave, "Domains of analyticity and lindstedt expansions of kam tori in some dissipative perturbations of hamiltonian systems," *Nonlinearity*, vol. 30, no. 8, pp. 3151–3202, 2017.

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

[7] A. Berretti, C. Falcolini, and G. Gentile, "Shape of analyticity domains of Lindstedt series: The standard map," *Phys. Rev. E (3)*, vol. 64, no. 1, part 2, pp. 015202, 4, 2001.

[8] A. Berretti and G. Gentile, "Renormalization group and field theoretic techniques for the analysis of the Lindstedt series," *Regul. Chaotic Dyn.*, vol. 6, no. 4, pp. 389–420, 2001.

[9] A. Celletti and C. Falcolini, "Singularities of periodic orbits near invariant curves," *Phys. D*, vol. 170, no. 2, pp. 87–102, 2002.

[10] A. Berretti and G. Gentile, "Scaling of the critical function for the standard map: Some numerical results," *Nonlinearity*, vol. 17, no. 2, pp. 649–670, 2004.

[11] R. de la Llave and S. Tompaidis, "On the singularity structure of invariant curves of symplectic mappings," *Chaos*, vol. 5, no. 1, pp. 227–237, 1995.

[12] R. Calleja and R. de la Llave, "Computation of the breakdown of analyticity in statistical mechanics models: Numerical results and a renormalization group explanation," *J. Stat. Phys.*, vol. 141, no. 6, pp. 940–951, 2010.

[13] R. Calleja and A. Celletti, "Breakdown of invariant attractors for the dissipative standard map," *Chaos*, vol. 20, no. 1, pp. 013121, 9, 2010.

[14] R. Calleja and R. de la Llave, "A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification," *Nonlinearity*, vol. 23, no. 9, pp. 2029–2058, 2010.

[15] R. Calleja and J.-L. Figueras, "Collision of invariant bundles of quasi-periodic attractors in the dissipative standard map," *Chaos*, vol. 22, no. 3, pp. 033114, 10, 2012.

[16] R. Calleja and R. de la Llave, "Fast numerical computation of quasi-periodic equilibrium states in 1D statistical mechanics, including twist maps," *Nonlinearity*, vol. 22, no. 6, pp. 1311–1336, 2009.

[17] D. E. Knuth, *The art of computer programming. Vol. 2*. Addison-Wesley, Reading, MA, 1998, pp. xiv+762, Seminumerical algorithms, Third edition [of MR0286318], ISBN: 0-201-89684-2.

[18] C. Falcolini and R. de La Llave, "Numerical calculation of domains of analyticity for perturbation theories in the presence of small divisors," *Journal of Statistical Physics*, vol. 67, no. 3, pp. 645–666, 1992.

[19] R. de la Llave, "A tutorial on KAM theory," in *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, ser. Proc. Sympos. Pure Math. Vol. 69, Providence, RI: Amer. Math. Soc., 2001, pp. 175–292.

[20] G. A. Baker and P. R. Graves-Morris, "Padé approximants," vol. 59, 1996.

[21] C. Batut, K. Belabas, D. Benardi, H. Cohen, and M. Olivier, "User's guide to PARI-GP," 1998, see also http://pari.home.ml.org.

[22] R. C. Calleja, A. Celletti, and R. de la Llave, "A KAM theory for conformally symplectic systems: Efficient algorithms and their validation," *J. Differential Equations*, vol. 255, no. 5, pp. 978–1049, 2013.

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

[24] ——, "Moser's implicit function theorem in the framework of analytic smoothing," *Math. Ann.*, vol. 219, no. 2, pp. 105–121, 1976.

[25] A. Celletti and C. L., "Quasi-periodic Attractors in Celestial Mechanics," *Arch. Rational Mech. Anal.*, vol. 191, no. 2, pp. 311–345, 2009.

[26] R. de la Llave, "KAM theory for equilibrium states in 1-D statistical mechanics models," *Ann. Henri Poincaré*, vol. 9, no. 5, pp. 835–880, 2008.

[27] M. Frigo and S. G. Johnson, "The design and implementation of FFTW3," *Proceedings of the IEEE*, vol. 93, no. 2, pp. 216–231, 2005, Special issue on "Program Generation, Optimization, and Platform Adaptation".

[28] Y. Hida, X. S. Li, and D. H. Bailey, *Quad-Double Arithmetic: Algorithms, Implementation, and Application*. 2000.

[29] C. Batut, K. Belabas, D. Bernardi, H. Cohen, and M. Olivier, *User's Guide to PARI-GP*. Université de Bordeaux I, 2000.

[30] A. Milani, A. M. Nobili, and P. Farinella, *Non-gravitational perturbations and satellite geodesy.* Adam Hilger (IOP Publishing Ltd.), Bristol, 1987, ISBN: 0-85274-538-9.

[31] A. Celletti, "KAM theory and its applications: From conservative to dissipative systems," in *European Congress of Mathematics*, Eur. Math. Soc., Zürich, 2013, pp. 243–257.

[32] ——, "Stability of satellites in spin-orbit resonances and capture probabilities," in *Predictability, stability, and chaos in N-body dynamical systems (Cortina d'Ampezzo, 1990)*, ser. NATO Adv. Sci. Inst. Ser. B Phys. Vol. 272, Plenum, New York, 1991, pp. 337–344.

[33] A. Bensoussan, *Perturbation methods in optimal control*, ser. Wiley/Gauthier-Villars Series in Modern Applied Mathematics. Chichester: John Wiley & Sons Ltd., 1988, pp. xiv+573, Translated from the French by C. Tomson, ISBN: 0-471-91994-2.

[34] E. S. Meadows, M. A. Henson, J. W. Eaton, and J. B. Rawlings, "Receding horizon control and discontinuous state feedback stabilization," *Internat. J. Control*, vol. 62, no. 5, pp. 1217–1229, 1995.

[35] R. Iturriaga and H. Sánchez-Morgado, "Limit of the infinite horizon discounted Hamilton-Jacobi equation," *Discrete Contin. Dyn. Syst. Ser. B*, vol. 15, no. 3, pp. 623–635, 2011.

[36] A. Davini, A. Fathi, R. Iturriaga, and M. Zavidovique, "Convergence of the solutions of the discounted Hamilton-Jacobi equation: Convergence of the discounted solutions," *Invent. Math.*, vol. 206, no. 1, pp. 29–55, 2016.

[37] J. Moser, "Convergent series expansions for quasi-periodic motions," *Math. Ann.*, vol. 169, pp. 136–176, 1967.

[38] H. W. Broer, G. B. Huitema, and M. B. Sevryuk, *Quasi-Periodic Motions in Families of Dynamical Systems.* Order Amidst Chaos. Berlin: Springer-Verlag, 1996, pp. xii+196, ISBN: 3-540-62025-7.

[39] 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)*, Dordrecht: Kluwer Acad. Publ., 1999, pp. 568–572.

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

[41] R. C. Calleja, A. Celletti, and R. de la Llave, "A kam theory for conformally symplectic systems: Efficient algorithms and their validation," *Journal of Differential Equations*, vol. 255, no. 5, pp. 978–1049, 2013.

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

[43] C. Foias and R. Temam, "Gevrey class regularity for the solutions of the Navier-Stokes equations," *J. Funct. Anal.*, vol. 87, no. 2, pp. 359–369, 1989.

[44] M. Canalis-Durand, "Formal expansion of van der Pol equation canard solutions are Gevrey," in *Dynamic bifurcations (Luminy, 1990)*, ser. Lecture Notes in Math. Vol. 1493, Springer, Berlin, 1991, pp. 29–39.

[45] C. Baesens, "Gevrey series and dynamic bifurcations for analytic slow-fast mappings," *Nonlinearity*, vol. 8, no. 2, pp. 179–201, 1995.

[46] D. Sauzin, "Caractère Gevrey des solutions formelles d'un problème de moyennisation," *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 315, no. 9, pp. 991–995, 1992.

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

[48] G. Popov, "Invariant tori, effective stability, and quasimodes with exponentially small error terms. I. Birkhoff normal forms," *Ann. Henri Poincaré*, vol. 1, no. 2, pp. 223–248, 2000.

[49] I. Baldomá and A. Haro, "One dimensional invariant manifolds of Gevrey type in real-analytic maps," *Discrete Contin. Dyn. Syst. Ser. B*, vol. 10, no. 2-3, pp. 295–322, 2008.

[50] I. Baldomá, E. Fontich, and P. Martín, "Gevrey estimates for one dimensional parabolic invariant manifolds of non-hyperbolic fixed points," *Discrete Contin. Dyn. Syst.*, vol. 37, no. 8, pp. 4159–4190, 2017.

[51] P. Bonckaert and P. De Maesschalck, "Gevrey and analytic local models for families of vector fields," *Discrete & Continuous Dynamical Systems-B*, vol. 10, no. 2&3, September, p. 377, 2008.

[52] O. Costin, G. Gallavotti, G. Gentile, and A. Giuliani, "Borel summability and Lindstedt series," *Comm. Math. Phys.*, vol. 269, no. 1, pp. 175–193, 2007.

[53] J. Feldman, J. Magnen, V. Rivasseau, and R. Sénéor, "Construction and Borel summability of infrared Φ_4^4 by a phase space expansion," *Comm. Math. Phys.*, vol. 109, no. 3, pp. 437–480, 1987.

[54] S. Graffi and V. Grecchi, "Borel summability of the unequal double well," *Comm. Math. Phys.*, vol. 92, no. 3, pp. 397–403, 1984.

[55] G. Gentile, M. V. Bartuccelli, and J. H. B. Deane, "Summation of divergent series and Borel summability for strongly dissipative differential equations with periodic or quasiperiodic forcing terms," *J. Math. Phys.*, vol. 46, no. 6, pp. 062704, 20, 2005.

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

[57] H. Cartan, *Elementary theory of analytic functions of one or several complex variables*. Dover Publications, Inc., New York, 1995, p. 228, Translated from the French, Reprint of the 1973 edition, ISBN: 0-486-68543-8.

[58] O. Costin, *Asymptotics and Borel summability*, ser. Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics. CRC Press, Boca Raton, FL, 2009, vol. 141, pp. xiv+250, ISBN: 978-1-4200-7031-6.

[59] H. Rüssmann, "On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus," in *Dynamical systems, theory and applications*, Springer, 1975, pp. 598–624.