Arch. Rational Mech. Anal. 231 (2019) 971–1044
Digital Object Identifier (DOI) https://doi.org/10.1007/s00205-018-1293-6

# *An A Posteriori KAM Theorem for Whiskered Tori in Hamiltonian Partial Differential Equations with Applications to some Ill-Posed Equations*

Rafael de la Llave & Yannick Sire

*Communicated by* V. Šverák

**Abstract**

The goal of this paper is to develop a KAM theory for tori with hyperbolic directions, which applies to Hamiltonian partial differential equations, even to some ill-posed ones. The main result has an *a-posteriori* format, that is, we show that if there is an approximate solution of an invariance equation which also satisfies some non-degeneracy conditions, then there is a true solution nearby. This allows, besides dealing with the quasi-integrable case, for the validation of numerical computations or formal perturbative expansions as well as for obtaining quasi-periodic solutions in degenerate situations. The a-posteriori format also has other automatic consequences (smooth dependence on parameters, bootstrap of regularity, etc.). We emphasize that the non-degeneracy conditions required are just quantities evaluated on the approximate solution (no global assumptions on the system such as twist). Hence, they are readily verifiable in perturbation expansions. We will pay attention to the quantitative relations between the sizes of the approximation and the non-degeneracy conditions. This will allow us to prove what experts call *small twist theorems* (the non-degeneracy conditions vanishes as the perturbation goes to zero but much slower than the error of the approximation). The method of proof is based on an iterative method for solving a functional equation for the parameterization of the torus expressing that the range of the parameterization admits an evolution and is invariant. We also solve functional equations for bundles which imply that are invariant under the linearization. The iterative method does not use transformation theory nor action-angle variables. The main result does not assume that the system is close to integrable. More surprisingly, we do not need that the equations we study define an evolution for all initial conditions and are well posed. Even if the systems we study do not admit solutions for all initial conditions, we show that there is a

R.L. Supported in part by National Science Foundation DMS-1800241.

systematic way to choose initial conditions on which one can define an evolution which is quasi-periodic. We first develop an abstract theorem. Then, we show how this abstract result applies to some concrete examples. The examples considered in this paper are the scalar Boussinesq equation and the Boussinesq system (both are PDE models that aim to describe water waves in the long wave limit). For these equations we construct *small amplitude* time quasi-periodic solutions which are even in the spatial variable. The strategy for the abstract theorem is inspired by that in Fontich et al. (Electron Res Announc Math Sci 16:9–22, 2009; J Differ Equ 246(8):3136–3213, 2009). The main part of the paper is to study infinite dimensional analogues of dichotomies which applies even to ill-posed equations and which is stable under addition of unbounded perturbations. This requires that we assume smoothing properties. We also present very detailed bounds on the change of the splittings under perturbations.

## Contents

## 1. Introduction

The goal of this paper is to develop a KAM theory for tori with hyperbolic directions, which applies to Hamiltonian partial differential equations, even to some ill-posed ones. The main result, Theorem 3.5, is stated in an a-posteriori format, that is, we formulate invariance equations and show that approximate solutions that satisfy some explicit non-degeneracy conditions lead to a true solution. This a-posteriori format leads automatically to several consequences (see Section 3.6.2) and can be used to justify numerical solutions and asymptotic expansions. We note that the results do not assume that the equations we consider define evolutions and indeed we present examples of quasi-periodic solutions in some well known ill-posed equations. See Sections 10, 11.

Adapting dynamical systems techniques to evolutionary PDE's has to overcome several technical difficulties. For starters, since the PDE's involve unbounded operators, the standard theories of existence, uniqueness developed for ordinary differential equations does not apply. As and is well known by now, there are systematic ways of defining the evolution for some systems using, for example, semigroup theory [14,36,62,72]. If semigroup theory applies, many dynamical systems techniques can be adapted in the generality of semigroups (see the pioneering work of [40] and more modern treatises [11,14,18,22,38,41,55,66,74,75]).

Besides the analytic difficulties, adapting ODE techniques to PDE's has to face several geometric difficulties. Some arguments widely used in dynamical systems fail to hold. For instance, symplectic structures on infinite-dimensional spaces (see for instance [3,17]) could lack several important properties. Relatedly, methods based on transformation theory—very common in dynamical systems—have to overcome severe technical properties [44,46–49]. Some recent methods on PDE that have avoided transformation theory are [4,7,20,23,24]. A final difficulty is that, when working near an equilibrium point, the action-angle variables are singular (even in finite dimensions) [35,44].

In the approach of this paper, we sidestep many of the above difficulties. We do not apply transformation theory, action-angle variables and the only use of geometric properties is in a very weak sense.

The main novelty of this paper is that the method applies to some ill-posed equations. The ill-posed equations we study cannot be interpreted as evolution equations.

Roughly speaking, we assume that one can define evolution forward in some (infinite codimension) space and one can define the evolution backward in another (also infinite codimension) space. We assume that when the evolutions can be defined, they are smoothing (that is, the evolved functions are smoother than the initial data).

From the dynamical point of view, the fact that the dynamics can be defined in the future (and smoothing) can be interpreted as the existence of a contractive space in the future, and similarly for the past. An important part of the assumption is that the spaces where the future or past dynamics can be defined span the whole phase space.

The existence of partial evolutions in different spaces is heuristically similar to the notion of hyperbolicity in dynamical systems (not to be confused with hyperbolicity in the sense of PDE). The fact that we can define smoothing evolutions roughly means that the Fourier modes corresponding to this space contract (either on the future or on the past), which is the dynamical notion of hyperbolicity.

The tori we consider in this paper are *whiskered*, that is the linearized dynamics (suitably interpreted) have many hyperbolic directions, indeed, as many directions as is possible to be compatible with the preservation of the symplectic structure. There is a rich KAM theory for finite dimensional whiskered tori [37,81] or for lower dimensional tori with elliptic directions [30,52,71,78]. A treatment of normally elliptic tori by methods similar to those here is in [51]. In this paper, however, we have to develop very different strategies to cope with the infinite dimensional nature of the problem and with the ill-posedness of the equations.

The rough idea—following the program outlined in [32]—is that we try to use methods inspired by the hyperbolicity theory in the hyperbolic directions and then use the sophisticated methods involving small divisors and geometry only on the center directions.

The adaptation of dynamical systems methods to the current setup present severe challenges. Some methods (for example graph transform, index theory methods, etc.), which are very useful in ODEs, require taking arbitrary initial conditions, so that, even at the level of heuristics, we need to make severe adjustments. Notably, the invariance equations for graphs etc. under some partial evolution have to be supplemented by equations ensuring that the initial conditions allow us to define the evolution. We also have to take care of formalizing the smoothing properties of the evolutions.

The final result will be given in an *a-posteriori* format. That is, given some function that solves the invariance equations approximately, if we can verify some non-degeneracy conditions, we are sure that there is a true solution near by. In particular, if we can verify the non-degeneracy conditions, we can justify numerical solutions, or formal expansions. (We will indeed justify some formal expansions in Sections 10 and 11.)

We were motivated by several concrete problems. An especially important one is the *long wave* approximations to water waves problems (the water waves problem, of course, is a free boundary problem and the surface evolves through equations that involve pseudo-differential operators, which were treatable only in the XX century, hundreds of years after the problem was posed so that many PDE approximations were systematically derived).

In Sections 10, 11, we present our result for some long-wave models that happen to be ill-posed. We produce solutions taking advantage of the a posteriori format of the main theorem. We construct formal expansions (Lindstedt series for low amplitude solutions) which satisfy the invariance equation very approximately. Even if the non-degeneracy conditions also deteriorate when the amplitude goes to zero, they do so at a much smaller rate. It should be remarked that the solutions we construct are in the range where the approximations involved are valid so that the solutions of the equations we produce will be approximate solutions of the water wave problem.

There are other problems in applied mathematics that can be reduced to the main abstract theorem; most notably, elliptic problems in cylindrical domains [45,54,65]. More tentatively, it seems that mean field games with noise are very close to the formalism developed here [1]. One can also hope to study other ill posed equations such as state dependent delay equations.

It should also be remarked that the solutions produced here can be shown to control a large part of the phase space. Even if the solutions we produce are finite dimensional, they contain infinite dimensional stable/unstable manifolds [10].

### *1.1. An Informal Description of the Main Result*

We now state our main result in an informal way and refer the reader to Section 3 for a rigorous statement. We hope that having a short overview of the structure will help to get a global road map that may be obscured by the details.

We consider an evolutionary PDE, which we write, symbolically, as

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{X} \circ u, \tag{1}$$

where $\mathcal{X}$ will be a differential and possibly non-linear operator.

We search for a solution under parametrization of the form $u(t) = K(\omega t)$ where $\omega$ is a frequency vector and $K$ a map from a complex strip $D_\rho$ of width $\rho > 0$ with values in a Banach space $X$.

We observe that $u(t) = K(\omega t)$ is a solution of (1) if and only if

$$\partial_\omega K(\theta) - \mathcal{X} \circ K(\theta) = 0, \tag{2}$$

where $\partial_\omega = \omega \cdot \nabla_\theta$. Note that that (2) implies that the range of $K$ consists of initial conditions that lead to solutions—not trivial in the case of ill posed equations—and that the solutions thus obtained never leave the set and that the motion in the set is equivalent to a rotation. Note that if the space $X$ consists of sufficiently differentiable functions, the solutions obtained will satisfy the equations in the strong sense.

We need the following assumptions:

- On the structure and regularity properties of $\mathcal{X}$.
- On the existence of an approximate solution.
- On Diophantine properties of the frequency vector.
- On non-degeneracy conditions on the approximate solution.

The structural assumptions we make are very conveniently stated following the two space approach of [40]. We consider two spaces of functions: $X$ consisting of smoother functions and $Y$ consisting of less smooth functions. The nonlinear part of the operator $\mathcal{X}$ will map $X$ to $Y$ (think that the operator $\mathcal{N}$ is a nonlinear differential operator of order $s$ and that $Y$ consists of functions with $s$ derivatives less than the functions in $X$). We will assume that $\mathcal{N}$ will be analytic in the sense of functions between Banach spaces.

We will also assume that the partial evolutions, forward and backward, map $Y$ to $X$ with quantitative bounds. As is typical in KAM theory we will also need that the frequencies satisfy number theoretic properties and that there is a twist condition.

The following is a more explicit formulation of the result, with some explicit precisions omitted:

**Theorem 1.1.** *Suppose that equation* (1) *satisfies the following structural assumptions* **H** *(described precisely in Section 3. ): writing $\mathcal{X} = \mathcal{A} + \mathcal{N}$ with $\mathcal{A}$ linear, we assume that*

- *$\mathcal{N}$ is analytic from $X$ to $Y$ and that $\mathcal{N}(0) = 0$, $D\mathcal{N}(0) = 0$;*
- *We can find two closed spaces inside of $X$, $X = X^{cs} + X^{cu}$ where $\mathcal{A}|_{X^{cs}}$ generates a forward evolution semigroup and $\mathcal{A}|_{X^{cu}}$ generates a backward evolution semigroup;*
- *The semigroups above extend to closed subspaces of $Y$ and map $Y$ to $X$ with some quantitative bounds. (This is a precise formulation of the intuitive idea that the semigroups giving the partial evolution are smoothing);*
- *Let $\omega$ be a Diophantine vector in $\mathbb{R}^\ell$ as in Definition 3.1;*
- *Assume that we are given an embedding $K_0 : \mathbb{T}^\ell \to X$ such that*
  - *It satisfies some non degeneracy conditions*
  - *Defining*

$$E_0 = \partial_\omega K_0 - \mathcal{X} \circ K_0,$$

    *we have that $E_0$ is small enough.*

*Then, there is a true solution $K$ of* (2) *so that $u(t) = K(\omega t)$ is a solution of the PDE.*

*Furthermore we have that $K - K_0$ are close (in some appropriate analytic norm) if $E_0$ is small (in some other norm). In particular, if $K_0$ is an embedding, we can ensure that the $K$ is also an embedding.*

We have, of course, not specified the norms that make precise the *"small enough"* statements above. We anticipate that, as typical in KAM theory, the spaces entering the smallness assumptions are more regular than the spaces in the conclusions.

The smallness conditions of the error with respect to the non-degeneracy conditions will be very explicit formulas.

We also note that (besides Diophantine conditions on $\omega$), the non-degeneracy conditions needed are all very explicit. They are obtained from the approximate solution itself taking derivatives, algebraic operations and averages. There are no global assumptions on the PDE (such as the customary twist conditions in dynamical systems). The main input of the theorem is an approximate solution and we conclude that there is a true solution close to it. Such theorems are called *a posteriori* in numerical analysis. Of course, full details will be given later but we thought that it would be good to give a feeling on the hypothesis.

The theorem also provides a form of local uniqueness of the solutions of the invariance equation.

To obtain the results for the applications to concrete equations, we will just take as the approximate solutions the result of some formal asymptotic expansions.

### *1.2. Organization of the Paper*

This paper is organized as follows: in Section 2 we present an overview of the method, describing the steps we will take, but ignoring some important precisions (for example domains of the operators), and proofs. In Section 3 we start developing the precise formulation of the results. We first present an abstract framework in the generality of equations defined in Banach spaces, including the abstract hypothesis. The general abstract results are stated in Section 3.6.1 and in Section 3.6.3 we discuss how to apply the results to some concrete examples. Some possible extensions are discussed in Section 3.6.2. The rest of the paper is devoted to the proof of the results following the strategy mentioned in the previous sections. One of the main technical results, which could have other applications, is the persistence of hyperbolic evolutions with smoothing properties; for this see Section 6.

## 2. Overview of the Method

In this section, we present a quick overview describing informally the steps of the method. We present the equations that need to be solved and the manipulations that need to be done ignoring issues such as domain of operators, estimates. These precisions will be taken up in Section 3. This section can serve as motivation for Section 3 since we use the formal manipulations to identify the issues that need to be resolved by a precise formulation.

We will discuss first abstract results, but in Sections 10 and 11, we will show that the abstract result applies to concrete examples.

One example to keep in mind and which has served as an important motivation for us is the Boussinesq equation

$$u_{tt} = \mu u_{xxxx} + u_{xx} + (u^2)_{xx} \quad x \in \mathbb{T},\ \ t \in \mathbb{R}, \mu > 0. \tag{3}$$

In Section 11, we will also consider the Boussinesq system. Other models in the literature which fit our scheme are the Complex Ginzburg–Landau equation and the derivative Complex Ginzburg–Landau equation for values of the parameters in suitable ranges.

**Remark 2.1.** There are several equations called the Boussinesq equation in the literature (in Section 11 we also present the Boussinesq system), notably the Boussinesq equation for fluids under thermal buoyancy. The paper [53] uses the name Boussinesq equation for $u_{tt} = -u_{xxxx} + (u^2)_{xx}$ and shows it is integrable in some sense made precise in that paper. Note that this equation is very different from (3) because of the sign of the fourth space derivative and (less importantly), the absence of the term with the second derivative. The sign of the fourth derivative term causes that the wave propagation properties of (3) and the equation in [53] are completely different.

Sometimes people refer to (3) with $\mu > 0$ as the *"bad"* Boussinesq equation, and call the equation with $\mu < 0$, the *"good"* Boussinesq equations. We note that the case $\mu > 0$ considered here is the case that appears in water waves (see [5, Equation (26)] ).

**Remark 2.2.** We note that the fourth derivative in (3) is just the next term in the long wave expansion of the water wave problem (which is not a PDE, but rather a free boundary problem). It would be natural to look also to higher order expansions in the space frequency which would lead to higher order evolution equations. The equation to order 6 seems to be well-posed but for the equation of order 8 is again ill-posed. Equations similar to (3) appear in many long wave approximations for waves. See [13,21] for modern discussions.

The special solutions of (3) which are in the range of validity of the long wave approximation are good approximate solutions of the water wave problem, but they are analyzable by PDE methods rather than the free boundary methods required by the original problem. [19,50]. Note that the solutions produced here lie in the regime (low amplitude, long wave) where the equation (3) was derived, so that they provide approximate solutions to the water wave problem.

## *2.1. The Evolution Equation*

We consider an evolutionary PDE, which we write symbolically as

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{X} \circ u, \tag{4}$$

where $\mathcal{X}$ will be a differential and possibly non-linear operator. This will, of course, require assumptions on domains etc., which we will take up in Section 3. For the moment, we will just say that $\mathcal{X}$ is defined in a domain inside a Banach space $X$. We will write

$$\mathcal{X}(u) = \mathcal{A}u + \mathcal{N}(u), \tag{5}$$

where $\mathcal{A}$ is linear and $\mathcal{N}$ is a nonlinear and possibly unbounded operator.

The differential equations $\dot{u} = \mathcal{A}u$ will not be assumed to generate dynamical evolution for all initial conditions (we just assume that it generates forward and backward evolutions when restricted to appropriate subspaces). Of course, we will not assume that (4) defines an evolution either. Lack of solutions for all the initial conditions will not be a severe problem for us since we will only try to produce some specific solutions.

The meaning in which (4) is to hold will be in the classical sense since we can produce very smooth solutions. As we will see we will take the space $X$ to consist of very differentiable functions so that the derivatives can be taken in the elementary classical sense. As intermediate steps, we will also find useful some solutions in the *mild* sense, satisfying some integral equations formally equivalent to (4). The mild solutions require less regularity in $X$. Again, we emphasize that we will only try to produce some specific solutions so that we will not need to discuss existence for general initial data.

We will assume that the nonlinear operator $\mathcal{N}$ is *"sub-dominant"* with respect to the linear part. This will be formulated later in Section 3, but we anticipate that this means roughly that $\mathcal{A}$ is of higher order than $\mathcal{N}$ and that the evolution generated by $\mathcal{A}$ when restricted to appropriate sub-spaces gains more derivatives than the order of $\mathcal{N}$. We will formulate all this precisely later.

We will follow [40] and formulate these effects by saying that the operator $\mathcal{N}$ is an analytic function from a domain $\mathcal{U} \subset X$ ($X$ is a Hilbert space of smooth functions) to $Y$ ($Y$ is a Hilbert space consisting of functions less smooth than the functions in $X$). Moreover, the forward/backward partial evolution operators map their domains in $Y$ back to $X$ with some quantitative bounds.

In the applications that we present in Sections 10 and 11, the equations we consider are polynomial but the method can deal with more general nonlinearities.[1]

### *2.2. The Linearized Evolution Equations*

Note that, in this set up we can define a linearized evolution equation around a curve $u(t)$ in $X$, that is

$$\frac{\mathrm{d}\xi}{\mathrm{d}t} = D\mathcal{X} \circ u(t)\xi \equiv \mathcal{A}\xi + D\mathcal{N}(u(t))\xi. \tag{6}$$

The equations (6) are to be considered as evolution equations for $\xi$ while $u(t)$ is given and fixed. The meaning of the term $D\mathcal{N}$ could be understood if $\mathcal{N}$ is a differentiable operator from $X$ to $Y$.

Of course, when $u(t)$ is solution of the evolution equation (4), equations (6) are the variational equations for the evolution. In our case, the evolution is not assumed to exist and, much less, the variational equations are assumed to provide a description of the effect of the initial conditions on the variation. We use these equations (6) even when $u(t)$ is not a solution of the evolution equation (4) and we will show that they are indeed a tool to modify an approximate solution $u(t)$ into a true solution.

[1] The equations we consider are taken from the literature of approximations of water waves. In these derivations, it is customary to expand the non-linearity and keep only the lower order terms.

Notice that (6) is non-autonomous, linear non-homogeneous, but that the existence of solutions is not guaranteed for all the initial conditions (even if the time dependent term is omitted).

In the finite dimensional case, equations of the form (6) even when $u(t)$ is not a solution are studied when performing a Newton method to construct a solution; for example in multiple shooting. Here, we will use (6) in a similar way. We will see that (6) can be studied using that $\mathcal{A}$ is dominant and has a splitting (and that $u(t)$ is not too wild).

### *2.3. The Invariance Equation*

Given a fixed $\omega \in \mathbb{R}^\ell$ that satisfies some good number theoretic properties (formulated precisely in Section 3.3), we will be seeking an embedding $K : \mathbb{T}^\ell \to X$ in such a way that

$$\mathcal{X} \circ K = DK \cdot \omega. \tag{7}$$

Note that if (7) holds, then, for any $\theta_0 \in \mathbb{T}^\ell$, $u(t) = K(\omega t + \theta_0)$ will be solution of (4). Hence, when we succeed in producing a solution of (7), we will have a $\ell$-parameter family of quasi-periodic solutions. The meaning of these parameters is the origin of the phase as is very standard in the theory of quasi-periodic functions.

Note that we will be looking for solutions obtained from embeddings of the torus. The geometric objects we seek for will be indexed by $\theta \in \mathbb{T}^\ell$. Nevertheless, when we consider evolutions we will need to consider arguments $\omega t$. Similar remarks will happen for other geometric objects such as invariant bundles and the evolution of linearized equations. We should think that these geometric objects are based in the $\mathbb{T}^\ell$ and that their evolution is given by a straight motion on $\mathbb{T}^\ell$.

### *2.4. Outline of the Main Result*

The main ingredient of the main result, Theorem 3.5, is that we will assume given an approximate solution $K_0$ of (7). That is, we are given an embedding $K_0$ in such a way that

$$\mathcal{X} \circ K_0 - DK_0 \cdot \omega \equiv e \tag{8}$$

is small enough. We will also assume that the linearized evolution satisfies some non-degeneracy assumptions. The conclusion is that there is a true solution close to the original approximate solutions. Theorems of these form in which we start from an approximate solution and conclude the existence of a true one are often called *"a posteriori"* theorems.

In the concrete equations that we consider in the applications, the approximate solutions will be constructed using Lindstedt series.

The sense in which the error $e$ is small requires defining appropriate norms, which will be taken up in Section 3. The precise form of the non-degeneracy conditions will be motivated by the following discussion which specifies the steps we will perform for the Newton method for the linearized equation

$$\frac{\mathrm{d}u}{\mathrm{d}t} = D\mathcal{X} \circ K_0(\theta + \omega t)u. \tag{9}$$

The non-degeneracy conditions have two parts. We first assume that for each $\theta \in \mathbb{T}^\ell$, the linearized equation satisfies some spectral properties. These spectral properties mean roughly that there are solutions of (9) that decrease exponentially in the future (stable solutions), others that decrease exponentially in the past (unstable solutions), and some center directions that can grow or decrease with a smaller exponential rate. The span of these three class of solutions is the whole space. We will also assume that the evolutions, when they can be defined, gain regularity.

In the ODE case, this means that the linearized equation admits an exponential trichotomy in the sense of [73].

In the PDE case, there are some subtleties not present in the ODE case. For instance, the vector field is not differentiable and is only defined on a dense subset. In our case, the difference with ODE theory is more drastic.

We will not assume that (9) defines an evolution for all time and all the initial conditions. We will however assume that (9) admits a solution forward in time for initial conditions in a space (the center stable space) and backwards in time for the another space (the center unstable space). We will furthermore assume that the center stable and center unstable spaces span the whole space, and they have a finite dimensional intersection (we will also assume that they have a finite angle, which we will formulate as saying that the projections corresponding to the splitting into the subspaces are bounded). We emphasize that we will not assume that the evolution forward of (9) can be defined outside of the center stable space nor that the backward evolution can be defined outside of the center unstable space.

Furthermore, we will assume that the evolutions defined in these spaces are smoothing with quantitative bounds. Of course, these subtleties are only present when we consider evolutions generated by unbounded operators and are not present in the ODE case.

A crucial result for us is Lemma 6.1, which shows that this structure (the trichotomy with smoothing) is stable under the addition of unbounded terms of lower order. We also present very quantitative estimates on the change of the structure under perturbations. Note that the result is also presented in an a-posteriori format so that we can use just the existence of an approximate invariant splitting.

The smoothing properties along the stable directions overcome the loss of regularity of the perturbation. Hence, we can obtain a persistence of the spaces under unbounded perturbations of lower order. A further argument shows the persistence of the smoothing properties. The result in Lemma 6.1 can be considered as a generalization of the finite dimensional result on stability of exponential dichotomies for allowing unbounded perturbations. An important consequence of this is that, when $\mathcal{N}(u)$ is small enough (in an appropriate sense), we can transfer the hyperbolicity from $\mathcal{A}$ to the approximate solution, which is the way that we construct the approximately hyperbolic solutions in the applications.

We will need to assume that in the center directions, there is some geometric structure that leads to some cancellations (sometimes called *automatic reducibility*). These cancellations happen because of the symplectic structure. We note that, in our case, we only need a very weak form of symplectic structure, namely that it can be made sense of in a finite dimensional space consisting of rather smooth functions. Note that for a rotational torus the infinitesimal perturbations do not grow in the tangential directions. The preservation of the geometric structure also implies that some of the perpendicular directions evolve not faster than linearly. Hence, the tori we consider are never normally hyperbolic and that for $\ell$-dimensional tori, the space of directions with subexponential growth is at least $2\ell$ dimensional. We will assume that the tori are as hyperbolic as possible while preserving of the symplectic structure, that is, the set of directions with subexponential growth is precisely $2\ell$ dimensional. These tori are called *whiskered* in the finite dimensional case.

**Remark 2.3.** The geometric properties we assume is the preservation of a symplectic structure, but this preservation is assumed to happen only on a very weak sense. The forms are only assumed to make sense on restrictions to finite dimensional spaces and also that the evolution preserves it in a finite dimensional invariant space.

This notion of symplectic form (which is general enough to encompassing several applications) is very weak and does not allow us to recover some of the standard results in symplectic geometry such as the Darboux theorems, etc. .

Fortunately, the method used in this paper does not require many symplectic properties. We do not rely on transformation theory. We only use the geometry to construct a good system of coordinates in a finite dimensional space and to show that a finite dimensional averages vanishes.

**Remark 2.4.** We note that (9) is formally the variation equation giving the derivative of the flow of the evolution equation. This interpretation is very problematic since the equations we will be interested in do not define necessarily a flow.

An important part of the effort in Section 3 consists in defining these structures in the restricted framework considered in this paper when many of the geometric operations used in the finite dimensional case are not available.

We also need to make assumptions that are analogues of the twist conditions in finite dimensions; see Definition 3.4. The twist condition we will require is just that a finite dimensional matrix is invertible. The matrix is computed explicitly on the approximate solution and does not require any global considerations on the differential equation.

### *2.5. Overview of the Proof*

The method of proof will be to show that, under the hypotheses we are making, a quasi-Newton method for equation (7) started in the initial guess, converges to a true solution. We emphasize that the unknown in equation (7) is the embedding $K$ of $\mathbb{T}^\ell$ into a Banach space $X$. The main method of proof will be to describe an iterative method of Nash–Moser type which will be quadratically convergent, but will involve weakening of the norm in each step.

Hence, we will need to introduce families of Banach spaces of embeddings (the proof of the convergence will be patterned after the corresponding proofs [57,80]).

For simplicity, we will only consider spaces of analytic embeddings. Note that the regularity of the embedding $K$ as a function of their argument $\theta \in \mathbb{T}^\ell$ is different from the regularity of the functions $K(\theta) \in X$. The $K(\theta)$ will be functions of the

$x$ variable. The space $X$ encodes the regularity with respect to another variable (denoted by $x$ ) and $K : \mathbb{T}^\ell \to X$ may have a different regularity than that of the functions in $X$. Indeed, we will consider also other Banach spaces $Y$ consisting of functions of smaller regularity in $x$.

The Newton method consists in solving the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta(\theta + \omega t) - D\mathcal{X} \circ K_0(\theta + \omega t)\Delta(\theta + \omega t) = -e, \tag{10}$$

and then taking $K_0 + \Delta$ as an improved solution.

Clearly, (10) is a non-homogeneous version of (9). Hence, the spectral properties of (9) will play an important role in the solution of (10) by the variations of constants formula. Following [31,32], we will show that using the trichotomy, we can decompose (10) into three equations, each one of them corresponding to one of the invariant subspaces.

The equations along the stable and unstable directions can be readily solved using the variation of parameters formula also known as Duhamel formula (which holds in the generality of semigroups) since the exponential contraction and the smoothing allow us to represent the solution as a convergent integral.

The equations along the center direction, as usual, are much more delicate. We will be able to show the geometric properties to establish the *automatic reducibility*. That is, we will show that there is an explicit change of variables that reduces the equation along the center direction to the standard cohomology equations over rotations (up to an error which is quadratic–in the Nash–Moser sense—in the original error in the invariance equation). It is standard that we can solve these cohomology equations under Diophantine assumptions on the rotation and that we can obtain *tame* estimates in the standard meaning of KAM theory [56,57,80]. One geometrically delicate point is that the cohomology equations admit solutions provided that certain averages vanish. The vanishing of these averages over perturbations is related to the exactness properties of the flow. Even if this is, in principle, much more delicate in the infinite dimensional case, it will turn out to be very similar to the finite dimensional case, because we will work on the restriction to the center directions which are finite dimensional. The procedure is very similar to that in [32].

We will not solve the linearized equations in center direction exactly; we will solve them up to an error which is quadratic in the original error. The resulting modified Newton method will still lead to a quadratically small error in the sense of Nash–Moser theory and can be used as the basis of a quadratically convergent method.

Once we have the Newton-like step under control we need to show that the step can be iterated infinitely often and it converges to the solution of the problem. This part of the argument is very standard.

A necessary step in the strategy is to show the stability of the non-degeneracy assumptions. The stability of the twist conditions is not difficult since it amounts to the invertibility of a finite dimensional matrix, depending on the solution. The stability of spectral theory is reminiscent of the standard stability theory for trichotomies [43,73] but it requires significantly more work since we need to use the smoothing properties of the evolution semigroups to control the fact that the perturbations are unbounded. Then, we need to recover the smoothing properties to be able to solve the cohomology equations. For this functional analysis set up, we have found very inspiring the *"two spaces approach"* of [40] and some of the geometric constructions of [15,16,40,64]. Since the present method is part of an iterative procedure, we will need very detailed estimates of the change.

We note that rather than presenting the main result as a persistence result, we prove an a-posteriori result showing that an approximate invariant structure implies the existence of a truly invariant one and we bound the distance between the original approximation and the truly invariant one. This, of course, implies immediately the persistence results since, given a system which has an exact solution of the invariance equations, we can consider this solution as an approximate solution for all systems close to the original one and apply the a-posteriori result. More importantly for us, we can use the a-posteriori format to validate the outcome of some formal expansions such as Lindstedt series. (See Section 10, 11).

## 3. The Precise Framework for the Results

In this section we formalize the framework for our abstract results. As indicated above, we will present carefully the technical assumptions on domains, etc. of the operators under consideration, and the symplectic forms. We will formulate spectral non-degeneracy conditions and the twist non-degeneracy assumption.

In Section 3.6 we will state our main abstract result, Theorem 3.5. The proof will be obtained in the subsequent Sections. Then, in Sections 10 and 11 we will show how the abstract theorem applies to several examples. The abstract framework has been chosen so that the examples fit into it, so that the reader is encouraged to refer to these sections for motivation. Of course, the abstract framework has been formulated with the goal that it applies to other problems in a more or less direct manner. We leave these to the reader.

We note that the formalism we use is inspired by the *two-space formalism* of [40]. We consider two Hilbert spaces $X$ and $Y$. The differential operators, which are unbounded from a space to itself will be very regular operators considered as operators from $X$ to $Y$. Some evolutions will have smoothing properties and map $Y$ to $X$ with good bounds.

### *3.1. The Evolution Equation*

We will consider an evolution equation as in (4) and (5).

We assume

**H1** There are two complex Hilbert spaces

$$X \hookrightarrow Y,$$

with continuous embedding. The space $X$ (resp. $Y$) is endowed with the norm $\|.\|_X$ (resp. $\|.\|_Y$)

We denote by $\mathcal{L}(X_1, X_2)$ the space of bounded linear operators from $X_1$ to $X_2$.

We will assume furthermore that $X$ is dense in $Y$. We will assume in applications that $\mathcal{A}$ and $\mathcal{N}$ are such that they map real functions into real functions; it will be part of the conclusions that the solutions of the invariance equations we obtain also give real values for real arguments.

**H2** The non-linear part $\mathcal{N}$ of (5) is an analytic function from $X$ to $Y$.

We recall that the definition of an analytic function is that it is locally defined by a norm convergent sum of multilinear operators. Since we will be considering an implicit function theorem, it suffices to consider just one small neighborhood and a single expansion in multi-linear operators. The examples in Sections 10 and 11 have nonlinearities which are just polynomials (finite sums of multilinear operators).

**Remark 3.1.** In our case, it seems that some weaker assumptions would work. It would suffice that $\mathcal{X} \circ K(\theta)$ is analytic for any analytic embedding $K$. In many situations this is equivalent to the stronger definition [42, Chapter III]. In the main examples that we will consider and in other applications, the vector field $\mathcal{X}$ is a polynomial.

**Remark 3.2.** It also seems possible that one could deal with finite differentiable problems. For the experts, we note that there are two types of KAM smoothing techniques: either smoothing only the solutions in the iterative process (single smoothing) [9,70] or smoothing also the problems (double smoothing) [57,80]. In general, double smoothing techniques produce better differentiability in the results. On the other hand, in this case, the approximation of the problems seems fraught with difficulties (how to define smoothings in infinite dimensional spaces, for unbounded operators is difficult). Nevertheless, single smoothing methods do not seem to have any problem. Of course, if the non-linearities have some special structure (for example they are obtained by composing with a non-linear function) it seems that a double smoothing could also be applied.

**Remark 3.3.** Note that the structure of $\mathcal{X}$ assumed in (5) allows us to estimate always the errors in $Y$, even if the unknown $K$ takes in $X$.

This is somewhat surprising since the loss of derivatives from $X$ to $Y$ is that of the subdominant term $\mathcal{N}$. We expect that the results of applying $\mathcal{A}$ to elements in $X$ does not lie in $Y$.

Nevertheless, using the structure in (5) and the smoothing properties we will be able to show by induction that if the error is in $Y$ at one step of the iteration, we can estimate the $Y$ norm of the error in subsequent steps of the iteration. Note that the new error is the error in the Taylor approximation of $\mathcal{X} \circ (K + \Delta)$, which is the error in the Taylor approximation of $\mathcal{N} \circ (K + \Delta)$.

Of course, we also need to ensure that the initial approximation satisfies this hypothesis. In the practical applications considered in this paper, we will just take a trigonometric polynomial as the initial approximate solution.

### *3.2. Symplectic Properties*

We will need that there is some exact symplectic structure. In our method, this does not play a very important role, we just use the preservation of the symplectic structure to derive certain identities in the (finite dimensional) center directions. These are called *automatic reducibility* and use the exactness to show that some (finite dimensional) averages vanish (*vanishing lemma*) so that we can prove the result without adjusting parameters.

We will assume that there is a (exact) symplectic form in the space $X$ and that the evolution equation (4) can be written in Hamiltonian form in a suitable weak sense, which we will formulate now.

Motivated by the examples in Sections 10 and 11 and others in the literature, we will assume that the symplectic form is just a constant operator over the whole space $X$ (notice that we can identify all the tangent spaces). We will not consider the possibility that the symplectic form depends on the position. Note that heuristically, the fact that the symplectic form is constant ensures $d\Omega = 0$ and, because we are considering a Banach space, the Poincaré lemma would give $\Omega = d\alpha$. We will need only weak forms of these facts. General symplectic forms in infinite dimensions may present surprising phenomena not present in finite dimensions [3,17,44]. Fortunately, we only need very few properties in finite dimensional subspaces in a very weak sense.

**H3** There is an anti-symmetric bounded operator $\Omega : X \times X \to \mathbb{C}$ taking real values on real vectors.

The operator $\Omega$ is assumed to be non-degenerate in the sense that $\Omega(u, v) = 0\ \forall v \in X$, implies $u = 0$.

$\Omega$ will be referred to as *the symplectic form*.

As we mentioned above, we are assuming that the symplectic form is constant.

In some of the applications, $\Omega$ could be a differential operator or the inverse of a differential operator. When $\Omega$ is a differential operator, the fact that $\Omega$ is bounded only means that we are considering a space $X$ consisting of functions with high enough regularity. The form $\Omega$ could be unbounded in $L^2$ or in spaces consisting of functions with lower regularity than the functions in $X$.

Notice that given a $C^1$ embedding $K$ of $\mathbb{T}^\ell$ to $X$ we can define the pull-back of $\Omega$ by the customary formula

$$K^*\Omega_\theta(a, b) = \Omega(DK(\theta)a, DK(\theta)b). \tag{11}$$

The form $K^*\Omega$ is a form on $\mathbb{T}^\ell$. If $K$ is $C^r$ as a mapping form $\mathbb{T}^\ell$ to $X$ (in our applications it will be analytic), the form $K^*\Omega$ will be $C^{r-1}$.

**H3.1** We will assume that $\Omega$ is exact in the sense that, for all $C^2$ embeddings $K : \mathbb{T}^\ell \to X$ we have

$$K^*\Omega = d\alpha_K \tag{12}$$

with $\alpha_K$ a one-form on the torus.

In the applications we will have that $\alpha_K = K^*\alpha$ for some 1-form in $X$. Note that if $\Omega$ is not constant, we will need that $\alpha$ depends on the position.

**H4** There is an analytic function $H : X \to \mathbb{C}$ such that for any $C^1$ path $\gamma : [0, 1] \to X$, we have

$$H(\gamma(1)) - H(\gamma(0)) = \int_0^1 \Omega(\mathcal{X}(\gamma(s)), \gamma'(s))\, \mathrm{d}s. \tag{13}$$

Note that **H4** is a weak form of the standard Hamilton equations $i_{\mathcal{X}}\Omega = dH$. We take the Hamiltonian equations and integrate them along a path to obtain (13).

A consequence of **H3** and **H4** is we have that for any closed loop $\Gamma$ with image in $\mathbb{T}^\ell$

$$\int_\Gamma i_{\mathcal{X}\circ K} K^*\Omega = 0. \tag{14}$$

**Remark 3.4.** The formulation of (13) is a very weak version of the Hamilton equation. In particular, it is somewhat weaker than the formulation in [49], but on the other hand, we will assume more hyperbolicity properties than in [49].

**3.2.1. Some Remarks on the Notation for the Symplectic Form** The symplectic form can be written as

$$\Omega(u, v) = \langle u, Jv\rangle_Z,$$

where $Z$ is a Hilbert space and $\langle\cdot,\cdot\rangle_Z$ denotes the inner product in $Z$ and $J$ is a (possibly unbounded) operator in $Z$—but bounded from $X$ to $Z$.

Once we have defined the operator $J$, we can talk about the operator $J^{-1}$ if it is defined in some domain.

The evolution equations can be written formally

$$\frac{\mathrm{d}u}{\mathrm{d}t} = J^{-1}\nabla H(u), \tag{15}$$

where $\nabla H$ is the gradient understood in the sense of the metric in $Z$. In the concrete applications here, we will take $Z = L^2$, $X = H^m$, $Y = H^{m-a}$ for large enough $m$. Of course, in well posed systems we can take sometimes $X = Y$, but even in parabolic equations (see [40]) it finds useful to distinguish the spaces.

We recall that the definition of a gradient (which is a vector field) requires a metric to identify differentials with vector fields. This is true even in finite dimensions. In infinite dimensions, there are several more subtleties such as the way that the derivative is to be understood. Hence, we will not use much the gradient notation and the operator $J$ except in Section 7, which is finite dimensional.

**Remark 3.5.** In the Physical literature (and in the traditional calculus of variations) it is very common to take $Z$ to be always $L^2$, even if the functions in the space $X$ or $Y$ are significantly more differentiable. In some ways the space $Z = L^2$ is considered as fixed and the spaces $X, Y$ are mathematical choices, so that the association of the symplectic form to a symplectic operator is always done with a different inner product $Z$. The book [61] contains a systematic treatment of the use of gradients associated to Sobolev inner products.

### *3.3. Diophantine Properties*

We will consider frequencies that satisfy the standard Diophantine properties.

**Definition 3.1.** Given $\kappa > 0$ and $\nu \geqq \ell - 1$, we define $\mathcal{D}(\kappa, \nu)$ as the set of frequency vectors $\omega \in \mathbb{R}^\ell$ satisfying the Diophantine condition

$$|\omega \cdot k|^{-1} \leqq \kappa |k|^\nu, \quad \text{for all } k \in \mathbb{Z}^\ell - \{0\}, \tag{16}$$

where $|k| = |k_1| + \cdots + |k_\ell|$. We denote

$$\mathcal{D}(\nu) = \cup_{\kappa > 0} \mathcal{D}(\kappa, \nu).$$

It is well known that when $\nu > \ell$, the set $\mathcal{D}(\nu)$ has full Lebesgue measure.

### *3.4. Spaces of Analytic Mappings from the Torus*

We will denote by $D_\rho$ the complex strip of width $\rho$, that is

$$D_\rho = \left\{ z \in \mathbb{C}^\ell / \mathbb{Z}^\ell : \ |\mathrm{Im}\, z_i| < \rho \ i = 1, \ldots, \ell \right\}.$$

We introduce the following $C^m$-norm for $g$ with values in a Banach space $W$:

$$|g|_{C^m(\mathcal{B}), W} = \sup_{0 \leqq |k| \leqq m} \sup_{z \in \mathcal{B}} ||D^k g(z)||_W.$$

Let $\mathcal{H}$ be a Banach space and consider $\mathcal{A}_{\rho,\mathcal{H}}$ the set of continuous functions on $\overline{D_\rho}$, analytic in $D_\rho$ with values in $\mathcal{H}$. We endow this space with the norm

$$\|u\|_{\rho,\mathcal{H}} = \sup_{z \in D_\rho} \|u(z)\|_{\mathcal{H}}.$$

$(\mathcal{A}_{\rho,\mathcal{H}}, \|\cdot\|_{\rho,\mathcal{H}})$ is well known to be a Banach space. Some particular cases which will be important for us are when the space $\mathcal{H}$ is a space of linear mappings (for example projections).

We will also need some norms for linear operators. Fix $\theta \in D_\rho$ and consider $A(\theta)$ a continuous linear operator from $\mathcal{H}_1$ into $\mathcal{H}_2$, two Banach spaces. Then we define $\|A\|_{\rho,\mathcal{H}_1,\mathcal{H}_2}$ as

$$\|A\|_{\rho,\mathcal{H}_1,\mathcal{H}_2} = \sup_{z \in D_\rho} \|A(z)\|_{\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)},$$

where $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ denotes the Banach space of linear continuous maps from $\mathcal{H}_1$ into $\mathcal{H}_2$ endowed with the supremum norm.

**Definition 3.2.** Let $\mathbb{T}^\ell = \mathbb{R}^\ell / \mathbb{Z}^\ell$ and $f \in L^1(\mathbb{T}^\ell, \mathcal{H})$ where $\mathcal{H}$ is some Banach space. We denote avg $(f)$ its average on the $\ell$-dimensional torus, that is

$$\mathrm{avg}\,(f) = \int_{\mathbb{T}^\ell} f(\theta)\, \mathrm{d}\theta.$$

**Remark 3.6.** Of course, in the previous definition, since $\mathcal{H}$ might be an infinite-dimensional space, the above integral, in principle, has to be understood as a Dunford integral. Nevertheless, since we will consider rather smooth functions, it will agree with simple approaches such as Riemann integrals.

### *3.5. Non-degeneracy Assumptions*

This section is devoted to the non-degeneracy assumptions associated to approximate solutions $K$ of (8). We first deal with the spectral non degeneracy conditions. The crucial object is the linearization equation around a map $K$ given by

$$\frac{\mathrm{d}\Delta}{\mathrm{d}t} = A(\theta + \omega t)\Delta, \tag{17}$$

where $A(\theta) = D(\mathcal{X} \circ K)(\theta)$ is an operator mapping $X$ into $Y$.

Roughly, we want to assume that there is a splitting of the space into directions on which the evolution corresponding to the linearized equation can be defined either forwards or backwards in time and that the evolutions thus defined are smoothing in their domain (that is, when the evolutions can be defined, they produce functions which are smoother than the initial data). We anticipate that in Section 6, we will present other conditions that imply Definition 3.3. We will just need to assume approximate versions of the invariance.

**Definition 3.3. Spectral non degeneracy** We will say that an embedding $K : D_\rho \to X$ is spectrally non degenerate if for every $\theta$ in $D_\rho$, we can find a splitting

$$X = X^s_\theta \oplus X^c_\theta \oplus X^u_\theta \tag{18}$$

with associated bounded projection $\Pi^{s,c,u}_\theta \in \mathcal{L}(X, X)$ and where $X^{s,c,u}_\theta$ are in such a way that:

- **SD1** The mappings $\theta \to \Pi^{s,u,c}_\theta$ are in $\mathcal{A}_{\rho,\mathcal{L}(X,X)}$ (in particular, analytic).
- **SD2** The space $X^c_\theta$ is finite dimensional with dimension $2\ell$. Furthermore the restriction of the operator $J$ to $X^c_\theta$ denoted $J_c$ induces a symplectic form on $X^c_\theta$ which is preserved by the evolution on $X^c_\theta$ (see below).
- **SD3** We can find families of operators

$$\begin{aligned} U^s_\theta(t) &: Y^s_\theta \to X^s_{\theta+\omega t} \quad t > 0\\ U^u_\theta(t) &: Y^u_\theta \to X^u_{\theta+\omega t} \quad t < 0\\ U^c_\theta(t) &: Y^c_\theta \to X^c_{\theta+\omega t} \quad t \in \mathbb{R} \end{aligned} \tag{19}$$

  such that:
  - **SD3.1** The families $U^{s,c,u}_\theta$ are cocycles over the rotation of angle $\omega$ (cocycles are the natural generalization of semigroups for non-autonomous systems)

$$U^{s,c,u}_{\theta+\omega t}(t')U^{s,c,u}_\theta(t) = U^{s,c,u}_\theta(t + t'). \tag{20}$$

  - **SD3.2** The operators $U^{s,c,u}_\theta$ are smoothing in the time direction where they can be defined and they satisfy assumptions in the quantitative rates. There exist $\alpha_1, \alpha_2 \in [0, 1)$, $\beta_1, \beta_2, \beta_3^+, \beta_3^- > 0$ and $C_h > 0$ independent of $\theta$ such that the evolution operators are characterized by the following rate conditions:

$$\|U^s_\theta(t)\|_{\rho,Y,X} \leqq C_h e^{-\beta_1 t} t^{-\alpha_1}, \qquad t > 0, \tag{21}$$

$$\|U_\theta^u(t)\|_{\rho,Y,X} \leqq C_h e^{\beta_2 t}|t|^{-\alpha_2}, \qquad t<0, \tag{22}$$

$$\begin{aligned} \|U_\theta^c(t)\|_{\rho,X,X} &\leqq C_h e^{\beta_3^+ t}, \qquad t>0 \\ \|U_\theta^c(t)\|_{\rho,X,X} &\leqq C_h e^{\beta_3^- |t|}, \qquad t<0, \end{aligned} \tag{23}$$

with $\beta_1 > \beta_3^+$ and $\beta_2 > \beta_3^-$.

– **SD3.3** The operators $U_\theta^{s,u,c}$ are fundamental solutions of the variational equations in the sense that

$$\begin{aligned} U_\theta^s(t) &= Id + \int_0^t A(\theta+\omega\sigma)U_\theta^s(\sigma)\,\mathrm{d}\sigma \quad t>0 \\ U_\theta^u(t) &= Id + \int_0^t A(\theta+\omega\sigma)U_\theta^u(\sigma)\,\mathrm{d}\sigma \quad t<0 \\ U_\theta^c(t) &= Id + \int_0^t A(\theta+\omega\sigma)U_\theta^c(\sigma)\,\mathrm{d}\sigma \quad t\in\mathbb{R}. \end{aligned} \tag{24}$$

Notice that we will assume that for $\theta$ real, the functions are real.

**Remark 3.7.** Note that as consequence of the integral equations and the rate conditions (21), (22), (23) we have, using just the triangle inequality

$$||U_\theta^s(t)||_{\rho,Y,Y} \le 1 + \int_0^t A s^{-\alpha_1} e^{-\beta_1 s}\,\mathrm{d}s.$$

Proceeding similarly for the others, we obtain

$$\begin{aligned} \|U_\theta^s(t)||_{\rho,Y,Y} &\leqq \tilde{C}_h e^{-\beta_1 t} \qquad t>0, \\ \|U_\theta^u(t)\|_{\rho,Y,Y} &\leqq \tilde{C}_h e^{\beta_2 t}, \qquad t<0, \\ \|U_\theta^c(t)\|_{\rho,Y,Y} &\leqq \tilde{C}_h e^{\beta_3^+ t}, \qquad t>0 \\ \|U_\theta^c(t)\|_{\rho,Y,Y} &\leqq \tilde{C}_h e^{\beta_3^- |t|}, \qquad t<0. \end{aligned} \tag{25}$$

**Remark 3.8.** We are not aware of any general argument that would show that

$$\begin{aligned} \|U_\theta^s(t)||_{\rho,X,X} &\leqq \tilde{C}_h e^{-\beta_1 t} \qquad t>0, \\ \|U_\theta^u(t)\|_{\rho,X,X} &\leqq \tilde{C}_h e^{\beta_2 t}, \qquad t<0, \\ \|U_\theta^c(t)\|_{\rho,X,X} &\leqq \tilde{C}_h e^{\beta_3^+ t}, \qquad t>0 \\ \|U_\theta^c(t)\|_{\rho,X,X} &\leqq \tilde{C}_h e^{\beta_3^- |t|}, \qquad t<0 \end{aligned} \tag{26}$$

follow from the other assumptions. We would be happy to hear about such argument. One can, however, clearly have that since $||U_\theta^s(t)||_{X,X} \le ||U_\theta^s(t)||_{Y,X}$, so that the semigroups are exponentially decreasing for large $t$.

A notable case, which happens in practice, when one can deduce (26) is when the spaces X and Y are Hilbert spaces. In such a case, taking Hilbert space adjoints in (3.7) we obtain

$$U^s_\theta(t)^* = Id + \int_0^t U^s_\theta(\sigma)^* A(\theta + \omega\sigma)^* \,\mathrm{d}\sigma \quad t > 0,$$

and using the fact that the adjoints preserve the norm, we can easily obtain bounds for $U^s_\theta(t)^*$ in the same way as (25).

**Remark 3.9.** We remark that when the equation preserves a symplectic structure, we can have, naturally,

$$\beta_3^+ = \beta_3^-, \qquad \beta_1 = \beta_2. \tag{27}$$

Conversely, if (27) is satisfied, the center direction automatically preserves a symplectic structure (see Lemma 7.3).

We anticipate that the results in Section 6 on persistence of trichotomies (a fortiori dichotomies) with smoothing are developed without assuming that the equation is Hamiltonian and, hence apply also to dissipative equations. Similarly, the solutions of linearized equations in the hyperbolic directions developed in Section 5 are obtained without using the Hamiltonian structure. The Hamiltonian structure is used only to deal with the linearized equations in the center direction in Section 7.

Let us comment on the previous spectral non-degeneracy conditions.

The first observation is that, if we assume that the spaces $X, Y$ are Sobolev spaces of high enough index (so that the functions in them are $C^r$ for $r$ high enough) then we have that (24) holds in a classical sense if it holds in the sense of mild solutions (the sense of integral equations). In the applications we have in mind, the above remark applies. We will to take as spaces $X$, $Y$ spaces consisting of functions with enough derivatives so that the solutionw we produce satisfy the equations in the classical sense.

Then, (24) is just a weak form of

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} U^s_\theta(t) &= A(\theta + \omega t) U^s_\theta(t) \quad t > 0\\
\frac{\mathrm{d}}{\mathrm{d}t} U^u_\theta(t) &= A(\theta + \omega t) U^u_\theta(t) \quad t < 0\\
\frac{\mathrm{d}}{\mathrm{d}t} U^c_\theta(t) &= A(\theta + \omega t) U^c_\theta(t) \quad t \in \mathbb{R}.
\end{aligned} \tag{28}$$

Often (24) is described as saying that the derivatives in (28) are understood in the mild sense.

Making sense of the integrals in (24) is immediate after some reflection. Our conditions just require the existence of an evolution for positive and negative times on certain subspaces. The important conditions on these evolutions are the characterization of the splitting by rates (21)–(22), expressing the fact that the operators are bounded and smoothing from $Y$ into $X$ (recall that $X \hookrightarrow Y$). If the system were autonomous, such properties would hold under some spectral assumptions on the operator $A(\theta)$ (bisectoriality or generation of strongly continuous semi-groups, see [62]).

Since the spaces $X^c_\theta$ and $Y^c_\theta$ are finite dimensional and of the same dimension, the evolution $U^c_\theta(t)$ can be considered as an operator from $Y^c_\theta$ to $Y^c_{\theta+t\omega}$.

In the finite dimensional case (or in the cases where there is a well defined evolution), property **SD.1** follows from the contraction rates assumption **SD.3** by a fixed point argument in spaces of analytic functions (see [39]). In our case, we have not been able to adapt the finite dimensional argument, that is why we have included it as an independent assumption (even if may end up be redundant). We note that **SD.1**, **SD.3** are clearly true when $\mathcal{N} \equiv 0$ and in this paper we will show that both properties are stable under perturbations, hence **SD.3** will hold for all small enough $u$. This suffices for our purposes, so we will not pursue the question of whether **SD.1** can be obtained from **SD.3** in general.

The fact that $\Omega|_{X^c_\theta}$ is non-degenerate (which is a part of **SD.2**) follows from the rate conditions **SD.3**, as we show in Lemma 7.3.

One situation when all the above abstract properties are satisfied is when the evolution is given just by the linear part $\mathcal{A}$, that is $\mathcal{N} \equiv 0$. The assumptions of our set up are verified if the spectrum of $\mathcal{A}$ is just eigenvalues of finite multiplicity and the spectrum is the union of a sector around the positive axis, another sector around the negative axis and a finite set of eigenvalues of finite multiplicity around the imaginary axis. Then, the stable space is the spectral projection over the sector in the negative real axis, the unstable space will be the spectral projection over the sector along the positive axis and the center directions will be the spectral space associated to the eigenvalues in the finite set. There are many examples of linear operators appearing in applications that satisfy these properties.

It will be important that the main result of Section 6 in these structures persists when we add a lower order perturbation which is small enough. Indeed, we will show that if we find splittings that satisfy them approximately enough, there is true splitting nearby. This would allow us to validate numerical computations, formal expansions, etc..

**3.5.1. The Twist Condition** As it is standard in KAM theory, one has to impose another non-degeneracy assumption, namely the twist condition. This is the object of the next definition. Notice that it amounts to a finite dimensional matrix being invertible. It is identical to the conditions that were used in the finite dimensional cases [27,31].

**Definition 3.4.** Denote $N(\theta)$ the $\ell \times \ell$ matrix such that $N(\theta)^{-1} = DK(\theta)^\perp DK(\theta)$.

Denote $P(\theta) = DK(\theta)N(\theta)$.

Let $J_c$ stand for restriction of symplectic operator $J$ to $X^c_\theta$. We will show in Lemma 7.3 that the form $\Omega_c \equiv \Omega|_{X^c_\theta}$ is non-degenerate so that the operator $J_c$ is invertible.

We now define the twist matrix $S(\theta)$ (the motivation will become apparent in Section 7, but it is identical to the definition in the finite dimensional case in [27,31]). The average of the matrix

$$S(\theta) = N(\theta)DK(\theta)^\perp[J_c^{-1}\partial_\omega(DK\,N) - AJ_c^{-1}(DK\,N)](\theta) \tag{29}$$

is non-singular.

We note that the matrix $S$ in (29) is a very explicit expression that can be computed out of the approximate solution of the invariant equation and the invariant bundles just taking derivatives, projections and performing algebraic operations. Thus it is easy to verify in applications when we are given an approximate solution.

We will say that an embedding is non-degenerate (and we denote it $K \in ND(\rho)$) if it is non-degenerate in the sense of Definitions 3.3 and 3.4.

**Remark 3.10.** As it will become apparent in the proof, the twist condition has a very clear geometric meaning, namely that the frequency of the quasiperiodic motions changes when we change the initial conditions in a direction (conjugate to the tangent to the torus).

Note that, given an invariant torus, we can consider it as an approximate solution for similar frequencies and that the twist condition also holds.

Using the a-posteriori theorem shows that under the conditions, we have many tori with similar frequencies near to the torus.

**3.5.2. Description of the Iterative Step** Once the two non-degeneracy conditions are met for the initial guess of the modified Newton method, the iterative step goes as follows:

(1) We project the cohomological equations with respect to the invariant splitting.
(2) We then solve the equations for the stable and unstable subspaces.
(3) We then solve the equation on the center subspace. This involves small divisor equations. We note that solving the equation in the center requires to use the exactness so that we can show that the equations are solvable.
(4) To be able to iterate we will need to show that the corrections also satisfy the non-degeneracy conditions (with only some slightly worse quantitative assumptions). This amounts to showing the stability of the spectral non-degeneracy conditions, and developing explicit estimates of the changes in the properties given the changes on the embedding.

## *3.6. Statement of the Results*

**3.6.1. General Abstract Results** The following theorem (3.5) is the main result of this paper. It provides the existence of an embedding $K$ for equation (7) under some non-degeneracy conditions for the initial guess. We stress here that Theorem 3.5 is in an *a posteriori* format (an approximate solution satisfying non-degeneracy conditions implies the existence of a true solution close to it). As already pointed out in the papers [31–33], this format allows one to validate many methods that construct approximate solutions, including asymptotic expansions or numerical solutions. We also note that it has several automatic consequences presented in Section 3.6.2.

**Theorem 3.5.** *Suppose assumptions* **H1**, **H2**, **H3** *are met; let* $\omega \in \mathcal{D}(\kappa, \nu)$ *for some* $\kappa > 0$ *and* $\nu \geq \ell - 1$. *Assume that*

- $K_0$ *satisfies the non-degeneracy Conditions 3.3 and 3.4 for some* $\rho_0 > 0$.

- *The range of $K_0$ acting on a complex extension of the torus is well inside of $\mathcal{U}$ the domain of analyticity of $\mathcal{N}$ introduced in* **H2**. *More precisely,*

$$\operatorname{dist}_X(K_0(D_\rho), X \setminus \mathcal{U}) \geq r > 0.$$

*That is, if $x = K_0(\theta)$, $\theta \in D_{\rho_0}$ and $||x - y||_X \leq \rho_0$, then $y \in \mathcal{U}$.*

*Define the initial error*

$$E_0 = \partial_\omega K_0 - \mathcal{X} \circ K_0.$$

*Then there exists a constant $C > 0$ depending on $l$, $\nu$, $\rho_0$, $|\mathcal{X}|_{C^1(B_r)}$, $\|DK_0\|_{\rho_0,X}$, $\|N_0\|_{\rho_0}$, $\|S_0\|_{\rho_0}$, (where $S_0$ and $N_0$ are as in Definition 3.4 replacing $K$ by $K_0$) and the norms of the projections $\|\Pi^{c,s,u}_{K_0(\theta)}\|_{\rho_0,Y,Y}$ such that, if $E_0$ satisfies the estimates*

$$C|\text{avg}\,(S_0)^{-1}|^2\kappa^4\delta^{-4\nu}\|E_0\|_{\rho_0,Y} < 1$$

*and*

$$C|\text{avg}\,(S_0)^{-1}|^2\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0,Y} < r,$$

*where $0 < \delta \leq \min(1, \rho_0/12)$ is fixed, then there exists an embedding $K_\infty \in ND(\rho_\infty := \rho_0 - 6\delta)$ such that*

$$\partial_\omega K_\infty(\theta) = \mathcal{X} \circ K_\infty(\theta)). \tag{30}$$

*Furthermore, we have the estimate*

$$\|K_\infty - K_0\|_{\rho_\infty,X} \leqq C|\text{avg}\,(S_0)^{-1}|^2\kappa^2\delta^{-2\nu}\|E_0\|_{\rho_0,Y}. \tag{31}$$

*The torus $K_\infty$ is also spectrally non degenerate in the sense of Definition 3.3 with $\rho$ in Definition 3.3 replaced by $\rho_\infty$ and with other constants differing from those of $K_0$ modifying by an amount bounded by $C\|E_0\|_{\rho_0}$.*

*Furthermore, if we have two solutions $K_1$, $K_2$ satisfying (7) and spectrally non-degenerate in the sense of Definition 3.3 and that satisfy*

$$\|K_1 - K_2\|_{\rho_\infty,X} \leqq C|\text{avg}\,(S_0)^{-1}|^2\kappa^2\delta^{-2\nu}. \tag{32}$$

*Then, there exists $\sigma \in \mathbb{R}^\ell$ such that*

$$K_1(\theta) = K_2(\theta + \sigma). \tag{33}$$

The statement that $K_\infty$ satisfies the Definition 3.3 is a consequence of the estimates in Section 6.

The uniqueness statement will be proved in Section 8. It is exactly the same as the one in the finite dimensional case in [31].

**3.6.2. Some Consequences of the A-Posteriori Format** The a-posteriori format leads immediately to several consequences. When we have systems that depend on parameters, observing that the solution for a value of the parameter is an approximate solution for similar values of the parameters, one obtains *Lipschitz dependence on parameters, including the frequency*.

If one can obtain Lindstedt expansions in the parameters, one can obtain Taylor expansions. If the parameter ranges over $\mathbb{R}^n$, this is the hypothesis of the converse Taylor theorem [2,60] so that one obtains *smooth dependence on parameters*. In the case that the parameters range on a closed set, we obtain one of the conditions of the Whitney extension theorem. Some general treatments are [8,76].

In many perturbative solutions, one gets that the twist condition is small but that the error is much smaller. Note that in the main result, we presented explicitly that the smallness conditions on the error are proportional to the square of the twist condition. Hence, we obtain the *small twist condition*. Note also that the twist condition required is not a global condition on the map, but rather a condition that is computed on the approximate solution. Indeed, we will take advantage of this feature in the sections on applications.

The abstract theorem can be applied to several spaces. Some spaces of low regularity (for example $H^m$) and others with high regularity (for example analytic). The existence results are more powerful in the high regularity spaces and the local uniqueness is more powerful in the low regularity spaces.

Given a sufficiently regular solution, one can obtain an analytic approximate solution by truncating the Fourier series, which leads to an analytic solution, which has to be the original one. Hence, one can *bootstrap the regularity*. See [9] for an abstract version.

**3.6.3. Results for Concrete Equations** Consider the following one-dimensional Boussinesq equation subject to periodic boundary conditions, that is

$$u_{tt} = \mu u_{xxxx} + u_{xx} + (u^2)_{xx} \quad x \in \mathbb{T}, \; t \in \mathbb{R}. \tag{34}$$

Looking for solutions of the linearization of the form $u(x,t) = e^{2\pi i(kx+\omega(k)t)}$ we obtain the following relation between frequencies and wave vectors (also called dispersion relation):

$$\omega^2(k) = -\mu |k|^4 (2\pi)^2 + |k|^2. \tag{35}$$

We see that for large $|k|$, $\omega(k) \approx \pm 2i\pi\mu^{1/2}|k|^2$. Hence, the Fourier modes may grow at an exponential rate and the rate is quadratic in the index of the mode, so that even analytic functions evolving under the linearized equation leave instaneously even spaces of distributions. The non-linear term does not restore the well posedness (see Remark 10.1.) The previous equation (34) is Hamiltonian on $L^2(\mathbb{T})$. Indeed, we introduce first the skew-symmetric operator

$$J^{-1} = \begin{pmatrix} 0 & \partial_x \\ \partial_x & 0 \end{pmatrix}$$

and define

$$H_\mu(u, v) = \int_0^1 \frac{1}{2}\left\{u^2 + v^2 - \mu(\partial_x u)^2\right\} + \frac{1}{3}u^3.$$

Therefore, equation (34) can be written

$$\dot{z} = J^{-1}\nabla H_\mu(z), \ \ z = (u, v), \tag{36}$$

where $\nabla$ has to be understood with respect to the inner product in $L^2(\mathbb{T})$. Note, however that when $\mu$ is small enough, there are several values of $k$, which for which $\omega(k)$ is real. We denote by $\omega^0$ the vector whose components are all the real frequencies that appear:

$$\begin{aligned} &\omega^0 = (\omega(k_1), \omega(k_2), \ldots, \omega(k_\ell)); \\ &\{k_1, \ldots k_\ell\} = \{k \in \mathbb{Z} \,|\, k > 0; -\mu|k|^4(2\pi)^2 + |k|^2 \geq 0\}. \end{aligned} \tag{37}$$

We can think of $\omega^0$ as the frequency vector of the motions for very small amplitude.

Note that the equation (34) conserves the quantity $\int_0^1 \partial_t u(t, x)\mathrm{d}x$ (called the momentum). Hence $\int_0^1 u(t, x)\mathrm{d}x$ (the center of mass) evolves linearly in time.

We can always change to a system of coordinates in which $\int_0^1 \partial_t u(t, x)\mathrm{d}x = 0$. Hence, in this system $\int_0^1 u(t, x)\mathrm{d}x = cte$. By adding the constant we can assume without loss of generality that $\int_0^1 u(t, x)\,\mathrm{d}x = 0$.

Hence we will assume (without loss of generality) that

$$\begin{aligned} &\int_0^1 \partial_t u(t, x)\mathrm{d}x = 0 \\ &\int_0^1 u(t, x)\mathrm{d}x = 0. \end{aligned} \tag{38}$$

**Remark 3.11.** We emphasize that the two parts of (38) are not two independent equations. The first one is just a derivative with respect to time of the second. Even if the relation is formal, it makes sense when we are dealing with polynomial approximate solutions.

We also note that the equation (34) leaves invariant the space of functions which are symmetric around $x$ (it does not leave invariant the space of functions antisymmetric around $x$). Hence, we can consider the equation as defined on the space of general functions or in the space of symmetric functions:

$$u(t, x) = u(t, -x). \tag{39}$$

The main difference between the symmetric and the general case is that center space is of different dimension.

We introduce Sobolev-type spaces $H^{\rho,m}(\mathbb{T})$ for $\rho > 0$ and $m \in \mathbb{N}$ being the space of analytic functions $f$ in $D_\rho$ such that the quantity

$$\|f\|^2_{\rho,m} = \sum_{k\in\mathbb{Z}} |f_k|^2 e^{4\pi\rho|k|}(|k|^{2m} + 1)$$

is finite, and where $\{f_k\}_{k\in\mathbb{Z}}$ are the Fourier coefficients of $f$. Let

$$X = H^{\rho,m}(\mathbb{T}) \times H^{\rho,m-2}(\mathbb{T}) \tag{40}$$

for $m \geqq 2$.

We state the following result:

**Theorem 3.6.** *Choose $\ell \in \mathbb{N}$, (the number of degrees of freedom we wil consider)*

*Consider the interval $I_\ell \subset \mathbb{R}^+$ such that if $\mu \in I_\ell$, the center space for $\mu > 0$ in* (34) *has dimension $2\ell \geqq 2$.*

*Fix a Diophantine exponent $\nu > \ell$ , a regularity exponent $m > 5/2$ and a positive analyticity radius $\rho_0$.*

*Then, for all $\mu \in I_\ell \setminus S$, where $S$ is a finite set, there exist a Cantor set $\mathcal{C}$ of frequencies in $\mathbb{R}^\ell$ and quasi-periodic solutions of* (34)*. These solutions correspond to whiskered tori.*

*More precisely, for each $\omega \in \mathcal{C}$, there exists an embedding $K_\omega$, and an analytic function from $D_{\rho_0} \to H^{\rho,m}(\mathbb{T}) \times H^{\rho,m-2}(\mathbb{T})$ solving* (7) *with frequency $\omega$.*

*The frequencies are asymptotically very abundant for small amplitude in the sense that we can write $\omega$ and $K_\omega$ as a Lipschitz function of the the amplitudes $A \in \mathbb{R}^\ell$.*

*Denoting $B_\varepsilon(0)$ the ball of radius $\varepsilon$ around zero in $\mathbb{R}^\ell$ one defines $\mathcal{A}_\varepsilon = \{A \mid \omega(A) \in \mathcal{C}, A \in B_\varepsilon(0)\}$.*

*Furthermore, we have as $\varepsilon \to 0$*

$$\frac{|\mathcal{A}_\varepsilon|}{|B_\varepsilon(0)|} \to 1.$$

As we have mentioned before, it is a consequence of the a-posteriori format of Theorem 3.5 that the mapping $\omega \to K_\omega$ is Lipschitz when $K$ are given the topology of analytic embeddings from $D_{\rho'}$ to $X$ when $\rho' < \rho_0$.

The reason why we call the $A$'s amplitudes is that the embeddings $K_\omega$ have the form $K_\omega(\theta) = \sum_{j=1}^{\ell} A_j \cos(jx)\cos(\theta_j) + O(A^2)$. Indeed, we will develop a systematic procedure to compute expansions of $A$.

In Section 10.5 we present a complete proof of Theorem 3.6.

Informally, following the standard Lindstedt procedure, for $\varepsilon$ small we find families of approximate solutions up to an error which is smaller than an arbitrarily large power of $\varepsilon$.

We can also verify that the non-degeneracy assumptions hold with a condition number which is a fixed power of $\varepsilon$. If $\varepsilon$ is very small one can allow frequencies with a large Diophantine constant, and obtain that the functions are analytic in a very large domain.

**Remark 3.12.** We expect that Theorem 3.6 can be greatly expanded (a wider range of parameters, removing the symmetry conditions) by just performing longer calculations using the Lindstedt method.

We also note that the dependence which of the frequencies on the amplitude, which we claim only to be Lipschitz, is actually $C^\infty$ in the sense of Whitney.

We hope to come back to this problem in future work.

**Remark 3.13.** Note that the case $\ell = 1$ amounts to periodic orbits so that there are no small denominators. In this case, one can use simpler fixed point theorems. There are already numerical computer assisted proofs in this case [12]. A general framework for this and for related results is in [34]. In this case, the results could be stronger (see Theorem 3.7 since we will not need to exclude parameters). We present a proof of the case $\ell = 1$ in Section 10.6.4.

A simple proof of the existence of periodic orbits for any $\ell$ can be obtained from a different argument. Using [26], we conclude that there is a finite dimensional manifold and that the system restricted to it is Hamiltonian. Then, one can apply the result of the existence of Lyapunov orbits [28,59,77]. Of course, using this method, one can only show existence for a small amplitude. The numerical methods of [12] can continue to large values and obtain information.

Similar results will be proved for other equations such as the Boussinesq system of water waves (see Section 11). The system under consideration is

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & -\partial_x - \mu\partial_{xxx} \\ -\partial_x & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \partial_x(uv) \\ 0 \end{pmatrix}, \tag{41}$$

where $t > 0$ and $x \in \mathbb{T}$. System (41) has a Hamiltonian structure given by

$$J^{-1} = \begin{pmatrix} 0 & \partial_x \\ \partial_x & 0 \end{pmatrix}$$

and

$$H_\mu(u, v) = \int_0^1 \frac{1}{2}\Big\{u^2 + v^2 - \mu(\partial_x v)^2\Big\} + \int_0^1 uv^2.$$

In this case, one has to take

$$X = H^{\rho,m}(\mathbb{T}) \times H^{\rho,m+1}(\mathbb{T})$$

and

$$Y = H^{\rho,m-1}(\mathbb{T}) \times H^{\rho,m}(\mathbb{T}).$$

The elementary linear analysis around the $(0, 0)$ equilibrium has been performed in [26]. The dispersion relation is given by

$$\omega(k) = \pm|k|2\pi i\sqrt{1 - 4\pi^2\mu k^2} \quad k \in \mathbb{Z}. \tag{42}$$

We take the principal determination of the square root. We denote by $\omega^0$ the vector whose components are all the real frequencies that appear

$$\begin{aligned} &\omega^0 = (\omega(k_1), \omega(k_2), \ldots, \omega(k_\ell)); \\ &\{k_1, \ldots k_\ell\} = \{k \in \mathbb{Z} \mid k > 0;\ 1 - 4\pi^2\mu k^2 \geq 0\}. \end{aligned} \tag{43}$$

For the Boussinesq systems, we have the same result as Theorem 3.6 for $\ell = 1$ (periodic orbits). In such a case, we do not need to exclude values of $\mu$.

**Theorem 3.7.** *Fix a positive analyticity radius $\rho_0$. Choose $\mu > 0$ so that the center direction is of half-dimension $\ell = 1$. Then, for all $\omega$ sufficiently close to $\omega^0$, we can find a periodic orbit of frequency $\omega$.*

The Proof of Theorem 3.7 will be done in Section 11. Of course, we believe that the result is true for any $\ell$, but it will take longer to prove.

Again, the main task will be to develop a Lindstedt expansion which produces errors that decrease to an arbitrarily high order, while the non-degeneracy conditions are satisfied at a fixed finite order.

## 4. The Linearized Invariance Equation

The crucial ingredient of the Newton method is to solve the linearized operator around an embedding $K$. This is motivated because one can hope to improve the solution of (4). Notice that the appearance of the linearized evolution does not have a dynamical motivation. The linearized equation does not appear as measuring the change of the evolution with respect to the initial conditions; it appears as the linearization of (4).

Let us denote

$$\mathcal{F}_\omega(K) = \partial_\omega K - \mathcal{X} \circ K. \tag{44}$$

Clearly, the invariance equation (7) can be written concisely as $\mathcal{F}_\omega(K) = 0$:

We prove the following result:

**Lemma 4.1.** *Consider the linearized equation*

$$D\mathcal{F}_\omega(K)\Delta = -E. \tag{45}$$

*There exists a constant $C$ that depends on $\nu$, $l$, $\|DK\|_{\rho,X}$, $\|N\|_\rho$, $\|\Pi_\theta^{s,c,u}\|_{\rho,Y,Y}$, $|(\mathrm{avg}\,(S))^{-1}|$ and the hyperbolicity constants such that assuming that $\delta \in (0, \rho/2)$ satisfies*

$$C\kappa\delta^{-(\nu+1)}\|E\|_{\rho,Y} < 1, \tag{46}$$

*we have that*

*A) There exists an approximate solution $\Delta$ of (45) and $\tilde{E}(\theta)$ such that $\Delta$ exactly solves*

$$D_K\mathcal{F}_\omega(K)\Delta = -E + \tilde{E} \tag{47}$$

*with, for all $\delta \in (0, \rho/2)$, the following estimates:*

$$\begin{aligned} \|\tilde{E}\|_{\rho-\delta,Y} &\leqq C\kappa^2\delta^{-(2\nu+1)}\|E\|_\rho\|\mathcal{F}_\omega(K)\|_{\rho,Y}, \\ \|\Delta\|_{\rho-2\delta,X} &\leqq C\kappa^2\delta^{-2\nu}\|E\|_{\rho,Y}, \\ \|D\Delta\|_{\rho-2\delta,X} &\leqq C\kappa^2\delta^{-2\nu-1}\|E\|_{\rho,Y}. \end{aligned} \tag{48}$$

*B) If $\Delta_1$ and $\Delta_2$ are approximate solutions of the linearized equation (45) in the sense of (47), then there exists $\alpha \in \mathbb{R}^\ell$ such that for all $\delta \in (0, \rho)$,*

$$\|\Delta_1 - \Delta_2 - DK(\theta)\alpha\|_{\rho-\delta,X} \leqq C\kappa^2\delta^{-(2\nu+1)}\|E\|_{\rho,Y}\|\mathcal{F}_\omega(K)\|_{\rho,Y}. \tag{49}$$

The previous Lemma is the cornerstone of the KAM iteration and the goal of the following sections is to prove this result. We will also need to prove that the non-degeneracy conditions are preserved under the iteration and that the constants measuring the non-degeneracy deteriorate only slightly. This will follow from the quantitative estimates developed in Section 6.

Note that (47), (48) is the main ingredient of several abstract implicit function theorems which lead to the existence of a solution. See, for example, [80], or, particularly, [9, Appendix A], for implicit function theorems based on the existence of approximate inverses with tame bounds.

Note also that in part (2) of Lemma 4.1 we have established some uniqueness for the solutions of the linearized equation. In Section 8 we will show how this can be used to prove rather directly the uniqueness result in Theorem 4.1.

The proof of Lemma 4.1 is based on decomposing the equation into equations along the invariant bundles assumed to exist in the hypothesis that the approximate solution satisfies Definition 3.3. In the hyperbolic directions we will roughly use the variations of parameters formula, but we will have to deal with the fact that the perturbations are unbounded. In the center directions, we will have to use the number theory and the geometry. Fortunately, the center space is finite dimensional.

The theory of solutions of the linearized equation is developed in Sections 5 and 7 and Lemma 4.1 is obtained just putting together Lemma 5.1 and the results in Section 7.4.

We also note that the estimates on the solutions of the linearized equation in the hyperbolic directions will be important in the perturbation theory of the bundles, which is needed to show that the linearized equation can be applied repeatedly.

For coherence of the presentation, we have written together all the results requiring hyperbolic technology (the solution in the hyperbolic directions and the perturbation theory of bundles). Of course, we hope that the sections can be read independently in the order preferred by the reader.

## 5. Solutions of Linearized Equations on the Stable and Unstable Directions

In this Section we develop the study of linearized equations of a system with splitting. Lemma 5.1 will be one of the ingredients in Lemma 4.1.

**Lemma 5.1.** *We assume that $\mathcal{A} : D_\rho \to \mathcal{L}(X, Y)$ is an analytic function admitting an invariant splitting in the sense that the space $X$ has an analytic family of splittings*

$$X = X_\theta^s \oplus X_\theta^c \oplus X_\theta^u$$

*(we say that a splitting is analytic when the associated projections depend on $\theta$ in an analytic way) invariant in the following sense: we can find families of operators $\{U_\theta^s(t)\}_{t>0}$, $\{U_\theta^c(t)\}_{t\in\mathbb{R}}$ , $\{U_\theta^u(t)\}_{t<0}$ with domains $X_\theta^s$, $X_\theta^c$, $X_\theta^u$ respectively. These families are analytic in $\theta, t$ (in the usual sense of analyticity of operators into $X$, that is, that they admit convergent Taylor expansions centered in any $\theta, t$). They satisfy*

$$U_\theta^{s,c,u}(t)X_\theta^{s,c,u} = X_{\theta+\omega t}^{s,c,u}. \tag{50}$$

*Let $\Pi_\theta^{s,c,u}$ the projections associated to this splitting. Assume furthermore that there exist $\beta_1, \beta_2, \beta_3^\pm > 0$, $\alpha_1, \alpha_2 \in (0, 1)$ and $C_h > 0$ independent of $\theta \in \overline{D_\rho}$ satisfying that $\beta_3^+ < \beta_1$ and $\beta_3^- < \beta_2$, and such that the splitting is characterized by the following rate conditions:*

$$\begin{aligned} \|U_\theta^s(t)\|_{\rho,Y,X} &\leqq C_h \frac{e^{-\beta_1 t}}{t^{\alpha_1}}, && t > 0,\\ \|U_\theta^u(t)\|_{\rho,Y,X} &\leqq C_h \frac{e^{\beta_2 t}}{|t|^{\alpha_2}}, && t < 0,\\ \|U_\theta^c(t)\|_{\rho,X,X} &\leqq C_h e^{\beta_3^+|t|}, && t > 0\\ \|U_\theta^c(t)\|_{\rho,X,X} &\leqq C_h e^{\beta_3^-|t|}, && t < 0. \end{aligned} \tag{51}$$

*Let* $F^{s,u} \in \mathcal{A}_{\rho,Y}$ *taking values in* $Y^s$ *(resp.* $Y^u$*). Consider the equations*

$$\partial_\omega \Delta^{u,s}(\theta) - \mathcal{A}(\theta)\Delta^{u,s}(\theta) = F^{u,s}(\theta). \tag{52}$$

*Then there are unique bounded solutions for* (52) *which are given by the formulas*

$$\Delta^s(\theta) = \int_0^\infty U^{s,u}_{\theta-\omega\tau}(\tau) F^s(\theta - \omega\tau)\,\mathrm{d}\tau \tag{53}$$

*and*

$$\Delta^u(\theta) = \int_{-\infty}^0 U^{s,u}_{\theta-\omega\tau}(\tau) F^u(\theta - \omega\tau)\,\mathrm{d}\tau. \tag{54}$$

*Furthermore, the following estimate holds:*

$$\|\Delta^{s,u}\|_{\rho,X^{s,u}_\theta} \leqq C\|\Pi^{s,u}_\theta\|_{\rho,Y,Y}\|F\|_{\rho,Y^s_\theta}.$$

**Remark 5.1.** The assumptions of the previous Lemma are very similar to the standard setup of the theory of dichotomies, but we have to take care of the fact that the evolution operators are smoothing and the perturbations unbounded.

**Proof.** The proof is based on the integration of the equation along the characteristics by using $\theta + \omega t$. We give the proof for the stable case, the unstable case being symmetric (for negative times). Furthermore, the proof is similar to the one in [31], up to some modifications of the functional spaces. Denote $\tilde{\Delta}^s(t) = \Delta^s(\theta + \omega t)$. By the variation of parameters formula (Duhamel formula), which is valid for mild solutions (see [62]), one has

$$\tilde{\Delta}^s(t) = U^s_\theta(t)\tilde{\Delta}^s(0) + \int_0^t U^s_{\theta+\omega z}(t - z) F^s(\theta + \omega z)\,\mathrm{d}z. \tag{55}$$

**Remark 5.2.** A full proof of (55) for mild solutions can be found in [62]. Nevertheless, it is illuminating to understand the heuristic reasons.

Heuristically, formula (55) is obtained by superposing the effect of $F(\theta + \omega z)$ measured at the final time. The effect of $F(\theta + \omega z)$ is applied when the solution is at $\theta + \omega z$, so, it propagates with with $U_{\theta+\omega z}$.

More carefully, but still postponing some details, one can observe that since, according to (28), $\frac{\mathrm{d}}{\mathrm{d}t}U^s_{\theta+\omega z}(t-z) = \mathcal{A}(\theta+\omega z)U^s_{\theta+\omega z}(t-z)$, one can deduce that (55) indeed solves the equation by taking derivatives inside the integral.

Now that we have a formula for the solution $\tilde{\Delta}(t)$, we want to find a formula for $\Delta : \mathbb{T}^\ell \to X$ so that $\tilde{\Delta}(t) = \Delta(\omega t)$. In our future applications $\Delta$ will be the correction of the embedding we seek.

Since the formula (55) is valid for all $\theta \in D_\rho \supset \mathbb{T}^\ell$, we can use it by substituting $\theta$ with $\theta - \omega t$ and then we have

$$\Delta^s(\theta) = U^s_{\theta-\omega t}(t)\Delta^s(\theta - \omega t) + \int_0^t U^s_{\theta-\omega t}(t) F^s(\theta - \omega t))\,\mathrm{d}z.$$

By the previous bounds on the semi-group, we have, if there is a bounded solution for $\Delta^s$, that $U_{\theta-\omega t}(t)\Delta^s(\theta - \omega t)$ goes to 0 exponentially fast when $t$ goes to $\infty$.

Also we obtain that the integrand in the previous formula goes to zero sufficiently fast so that the integral can be taken for $t \to \infty$.

Hence we obtain the following representation:

$$\Delta^s(\theta) = \int_0^\infty U_{\theta-\omega\tau}(\tau) F^s(\theta - \omega\tau)\, \mathrm{d}\tau. \tag{56}$$

We now estimate the integral in (56) to show that it converges, to establish bounds, and to show that it defines the analytic in $\theta$ and that, indeed, it provides a solution.

Notice that the operator $U_\theta^s(t)$ maps $Y_\theta^s$ into $X_\theta^s$ continuously and that the following estimate holds for every $\theta \in D_\rho$ and every $t > 0$:

$$\|U_\theta^s(t) F^s(\theta)\|_{X_\theta^s} \leqq \frac{C}{t^{\alpha_1}} e^{-\beta_1 t} \|F^s(\theta)\|_{Y_\theta^s}.$$

The exponential bound in **SD3.2** ensures the convergence at infinity of the integral and the fact that $\alpha_1 \in (0, 1)$ ensures the convergence at 0 gives the desired bound.

The unstable case can be obtained by reversing the direction of time or given by a direct proof which is identical to the present one. □

## 6. Perturbation Theory of Hyperbolic Bundles in an Infinite-Dimensional Framework

In this section we develop a perturbation theory for hyperbolic bundles and their smoothing properties. We consider a slightly more general framework than the one introduced in the previous sections since we hope that the results in this section could be useful for other problems (e.g in dissipative PDE's). In particular, we note that we only assume that the spaces $X$ and $Y$ are Banach spaces. Also, we do not need to assume that the (unbounded) vector field $\mathcal{X}$ giving the equation is Hamiltonian. In agreement with previous results, we note that we do not assume that the equations define an evolution for all initial conditions. We only assume that we can define evolutions in the future (or on the past) of the linearization in some spaces. This is obvious for the linear operator and in this section we will show that this is persistent under small perturbations.

The theory of perturbations of bundles for evolutions in infinite dimensional spaces has a long history; see for example [40, 64]. A treatment of partial differential equations has already been considered in the literature, for example in [15, 16].

Our treatment has several important differences with the above mentioned works; among them: 1) We study the stability of smoothing properties; 2) We take advantage of the fact that the dynamics in the base is a rotation, so that we obtain results in the analytic category, which are false when the dynamics in the base is more complicated; 3) We present our main results in an a-posteriori format, which, of course, implies the standard persistence results but has other applications such as validating numerical or asymptotic results; 4) We present very quantitative estimates of the changes of the splitting and its merit figures under perturbations;

this is needed for our applications since we use it as an ingredient of an iterative process and we need to show that it converges.

The main result in this Section is Lemma 6.1, which shows the invariant splittings and their smoothing properties when we change the linearized equation. Of course, in the applications in the iterative Nash–Moser method, the change of the equation will be induced by a change in the approximate solution.

**Lemma 6.1.** *Assume that $A(\theta)$ is an analytic family of linear maps as before. Let $\tilde{A}(\theta)$ be another family such that $\|\tilde{A} - A\|_{\rho,X,Y}$ is small enough. Then there exists a family of analytic splittings*

$$X = \tilde{X}_\theta^s \oplus \tilde{X}_\theta^c \oplus \tilde{X}_\theta^u,$$

*which is invariant under the linearized equations*

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta = \tilde{A}(\theta + \omega t)\Delta$$

*in the sense that the following holds:*

$$\tilde{U}_\theta^{s,c,u}(t)\tilde{X}_\theta^{s,c,u} = \tilde{X}_{\theta+\omega t}^{s,c,u}.$$

*We denote by $\tilde{\Pi}_\theta^{s,c,u}$ the projections associated to this splitting. Then there exist $\tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3^+, \tilde{\beta}_3^- > 0$, $\tilde{\alpha}_1, \tilde{\alpha}_2 \in (0, 1)$ and $\tilde{C}_h > 0$ independent of $\theta$ satisfying $\tilde{\beta}_3 < \tilde{\beta}_1$, $\tilde{\beta}_3 < \tilde{\beta}_2$ and such that the splitting is characterized by the following rate conditions:*

$$\begin{aligned}
\|\tilde{U}_\theta^s(t)\|_{\rho,Y,X} &\leqq \tilde{C}_h \frac{e^{-\tilde{\beta}_1 t}}{t^{\tilde{\alpha}_1}}, \qquad t > 0,\\
\|\tilde{U}_\theta^u(t)\|_{\rho,Y,X} &\leqq \tilde{C}_h \frac{e^{\tilde{\beta}_2 t}}{|t|^{\tilde{\alpha}_2}}, \qquad t < 0,\\
\|\tilde{U}_\theta^c(t)\|_{\rho,X,X} &\leqq \tilde{C}_h e^{\tilde{\beta}_3^+ t}, \qquad t > 0\\
\|\tilde{U}_\theta^c(t)\|_{\rho,X,X} &\leqq \tilde{C}_h e^{\tilde{\beta}_3^- |t|}, \qquad t < 0.
\end{aligned} \tag{57}$$

*Furthermore the following estimates hold:*

$$\|\tilde{\Pi}_\theta^{s,c,u} - \Pi_\theta^{s,c,u}\|_{\rho,Y,Y} \leqq C\|\tilde{A} - A\|_{\rho,X,Y}, \tag{58}$$

$$|\tilde{\beta}_i - \beta_i| \leqq C\|\tilde{A} - A\|_{\rho,X,Y}, \qquad i = 1, 2, 3^\pm, \tag{59}$$

$$\tilde{\alpha}_i = \alpha_i, \qquad i = 1, 2 \tag{60}$$

$$\tilde{C}_h = C_h. \tag{61}$$

**Proof.** We want to find invariant subspaces for the linearized evolution equation. We concentrate on the stable subspace, the theory for the other bundles being similar. We do this by finding a family of linear maps indexed by $\theta$, denoted $\mathcal{M}_\theta$ : $X_\theta^s \to X_\theta^{cu} \equiv X_\theta^c \oplus X_\theta^u$ in such a way that the graph of $\mathcal{M}_\theta$ is invariant under the equation. Note that since we do not assume that the equation defines a flow, the fact that we can evolve the elements in the graph in the future is an important part of the conclusions. We will also show that the family of maps depends analytically in $\theta$.

**Step 1: Construction of the invariant splitting.**

We will consider first the case of the stable bundle. The others are done identically. We will first try to characterize the initial conditions of the linearized evolution equation that lead to a forward evolution which is a contraction. We will see that these lie in a space. We will formulate the new space as the graph of a linear function $\mathcal{M}_\theta$ from $X_\theta^s$ to $X_\theta^{cu}$. We will show that if such a characterization was possible, $\mathcal{M}_\theta$ would have to satisfy some equations. To do this, we will formulate the problem of existence of forward solutions and the invariance of the bundle as two (coupled) fixed point problems (see (67) and (68).) One fixed point problem will formulate the invariance of the space, and the other fixed point problem the existence of forward solutions. We will show that, in some appropriate spaces, these two fixed point problems can be studied using the contraction mapping principle. The definition of the spaces will be somewhat elaborate since they will also encode the analytic dependence on the initial conditions, which is natural if we want to show the analytic dependence on $\theta$ of the invariant spaces.

Note that this step is significantly different from the problems in dynamical systems. In the finite dimensional problems, we do not need to worry about choosing the initial conditions so the the evolution can be defined. In our case, however, we need to select carefully the initial conditions so that the evolution can be defined.

**Remark 6.1.** Since the main tool will be a contraction argument, it follows that the main result is an a-posteriori result. Given approximate solutions of the invariance equations (obtained e.g numerically or through formal expansions, etc.) one can find a true solution close to the approximate one. We leave to the reader the recasting of Lemma 6.1 in this style.

Now, we implement in detail the above strategy. We first derive the functional equations, then specify the spaces.

We start by considering the linearized equation with an initial phase $\theta$. For subsequent analysis, it will be important to study the dependence on $\theta$ of the solutions. Eventually, we will show that the new invariant spaces depend analytically on $\theta$. This will translate in the geometric properties of the bundles. Consider

$$\frac{\mathrm{d}}{\mathrm{d}t} W_\theta(t) = \tilde{A}(\theta + \omega t) W_\theta(t). \tag{62}$$

Note that we use the index $\theta$ to indicate that we are considering the equation with initial phase $\theta$.

We write (62) as

$$\frac{\mathrm{d}}{\mathrm{d}t} W_\theta(t) = A(\theta + \omega t) W_\theta(t) + B(\theta + \omega t) W_\theta(t), \tag{63}$$

with $B = \tilde{A} - A$. Denote $\gamma = \|\tilde{A} - A\|_{\rho,X,Y} \equiv \|B\|_{\rho,X,Y}$, which we will assume to be small.

We recall that this is an equation for $W_\theta$ and that we are not assuming solutions to exist. Indeed, one of our goals is to work out conditions that ensure that forward solutions exist. Hence, we will manipulate the equation (62) to obtain some conditions.

We compute the evolution of the projections of $W_\theta(t)$ along the invariant bundles by the linearized equation when $B \equiv 0$. For $\sigma = s, c, u$ we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\Pi^\sigma_{\theta+\omega t} W_\theta(t)\right) &= \left(\omega \cdot \partial_\theta \Pi^\sigma_{\theta+\omega t}\right) W_\theta(t) + \Pi^\sigma_{\theta+\omega t}\left(\frac{\mathrm{d}}{\mathrm{d}t} W_\theta(t)\right) \\
&= \left(\omega \cdot \partial_\theta \Pi^\sigma_{\theta+\omega t}\right) W_\theta(t) + \Pi^\sigma_{\theta+\omega t} A(\theta+\omega t) + \Pi^\sigma_{\theta+\omega t} B(\theta+\omega t) W_\theta(t) \\
&= A^\sigma(\theta+\omega t)\Pi^\sigma_{\theta+\omega t} W_\theta(t) + \Pi^\sigma_{\theta+\omega t} B(\theta+\omega t) W_\theta(t).
\end{aligned} \tag{64}$$

In the last line of (64), we have used that the calculation in the first two lines of (64) is also valid when $B = 0$ and that, in that case, the invariance of the bundles under the $A$ evolution implies that all the terms appearing can be subsumed into $A^\sigma$ which only depends on the projection on the bundle.

Of course the same calculation is valid for the projections over the center-unstable (and center-stable, etc.) bundles. We denote by $\Pi^{cu}_\theta = \Pi^c_\theta + \Pi^u_\theta$ the projection over the center-unstable bundle. Note that $\Pi^s_\theta + \Pi^{cu}_\theta = \mathrm{Id}$.

Our goal now is to try to find a subspace in which the solutions of (62) (equivalently (64)) can be defined forward in time.

We will assume that this space where solutions can be defined is given as the graph of a linear function $\mathcal{M}_\theta$ from $X^s_\theta$ to $X^c_\theta \oplus X^u_\theta$.

That is, we introduce the notation $W^{cu}_\theta(t) = \Pi^{cu}_\theta W_\theta(t)$, $W^s_\theta(t) = \Pi^s_\theta W_\theta(t)$ and we will assume that if we consider one initial condition of the form

$$W_\theta(0) = (W^s_\theta(0), \mathcal{M}_\theta W_\theta(0)),$$

then there is a solutions of (62) with the initial condition above. Moreover, the solutions have the form

$$\begin{aligned}
W_\theta(t) &= (W^s_\theta(t), W^{cu}_\theta(t)) \\
&= (W^s_\theta(t), \mathcal{M}_{\theta+\omega t} W^s_\theta(t)) = (Id + \mathcal{M}_{\theta+\omega t}) W^s_\theta(t).
\end{aligned}$$

Of course, we will have to determine the evolution of $W^s_\theta(t)$ and of $M_\theta$.

We will have to show that the linear subspace of $X$ where the evolution can be defined and which is invariant can indeed be found and show that it depends analytically on $\theta$. To do so, we start by formulating the equations satisfied by the unknowns $W^s_\theta(t)$, $\mathcal{M}_\theta$ that express the existence of the evolution and the invariance of the graph. These two equations, will, of course, be coupled.

For any $T > 0$, if there were solutions of the equation satisfied by $W^{cu}_\theta$ we would have Duhamel's formula. Then, imposing that it is in the graph, we get

$$\begin{aligned}
\mathcal{M}_\theta W^s_\theta(0) &= W^{cu}_\theta(0) \\
&= U^u_{\theta+\omega T}(-T)\mathcal{M}_{\theta+\omega T} W^s_\theta(T) \\
&+ \int_0^T U^u_{\theta+\omega t}(t-T)\Pi^{cu}_{\theta+\omega t} B(\theta+\omega t)(Id + \mathcal{M}_{\theta+\omega t}) W^s_\theta(t)\,\mathrm{d}t.
\end{aligned} \tag{65}$$

Similarly, one has

$$W^s_\theta(t) = U^s_{\theta+\omega t} W^s_\theta(0) + \int_0^t U^s_{\theta+\omega(t-\tau)}(t-\tau)\Pi^{cu}_{\theta+\omega\tau} B(\theta+\omega\tau)(Id + \mathcal{M}_{\theta+\omega\tau}) W^s_\theta(\tau)\,\mathrm{d}t. \tag{66}$$

Notice that the fact that (66) is linear implies that if its solutions are unique, then $W^s_\theta(t)$ depends linearly on $W^s_\theta(0)$ (it depends very nonlinearly on $\mathcal{M}_\theta$). We will write $W^s_\theta(t) = \mathcal{N}_\theta(t) W^s_\theta(0)$ where $\mathcal{N}_\theta(t)$ is a linear operator.

We have then

$$\mathcal{M}_\theta = U^u_{\theta+\omega T}(-T)\mathcal{M}_{\theta+\omega T}\mathcal{N}_\theta(T) + \int_0^T U^u_{\theta+\omega t}(t-T)\Pi^{cu}_{\theta+\omega t} B(\theta+\omega t)(Id + \mathcal{M}_{\theta+\omega t})\mathcal{N}_\theta(t)\,\mathrm{d}t. \tag{67}$$

Similarly, we have that (66) is implied by

$$\mathcal{N}^s_\theta(t) = U^s_{\theta+\omega t}(0) + \int_0^t U^s_{\theta+\omega(t-\tau)}(t-\tau)\Pi^{cu}_{\theta+\omega\tau} B(\theta+\omega\tau)(Id + \mathcal{M}_{\theta+\omega\tau})\mathcal{N}^s_\theta(\tau). \tag{68}$$

We can think of (67) and (68) as equations for the two unknowns $\mathcal{M}$ and $\mathcal{N}_\theta$ where $\mathcal{M}$ will be a function of $\theta$ and $\mathcal{N}$ a function of $\theta, t$.

Note that (67) and (68) are already written as fixed point equations for the operators defined by the right hand side of the equations. It seems intuitively clear that the right hand side of the equations will be contractions since the linear terms involve a factor $B$ which we are assuming is small. Of course, to make this intuition precise, we have to specify appropriate Banach spaces and carry out some estimates. After the spaces are defined, the estimates are somewhat standard and straightforward. We point out that operators similar to (67) appear in the perturbation theory of hyperbolic bundles and operators similar to (68) appear in the theory of perturbations of semigroups. The integral equations are also very common in the study of neutral delay equations.

### *6.1. Definition of Spaces*

Let $\rho > 0$. For $\theta \in \overline{D_\rho}$ we denote by $\mathcal{L}(X^s_\theta, X^{cu}_\theta)$ the space of bounded linear maps from $X^s_\theta$ into $X^{cu}_\theta$. We considered it endowed with the standard supremum norm of linear operators.

Denote also by $\mathcal{L}_\rho(X^s, X^{cu})$ the space of analytic mappings from $D_\rho$ into the space of linear operators in $X$ that to each $\theta \in D_\rho$, assign a linear operator in $\mathcal{L}(X^s(\theta), X^{cu}(\theta))$. We also require from the maps in $\mathcal{L}_\rho(X^s, X^{cu})$ that they extend continuously to the boundary of $D_\rho$. We endow $\mathcal{L}_\rho(X^s, X^{cu})$ with the topology of the supremum norm, which makes it into a Banach space.

We also introduce the standard $C^0([0,T], \mathcal{L}_\rho(X^s, X^{cu}))$, endowed with the supremum norm. For each $\theta \in D_\rho$ we denote $C^0_\theta([0,T], \mathcal{L}(X^s, X^c))$ the space of continuous functions which for every $t \in [0, T]$, assign a linear operator in $\mathcal{L}(X^s_{\theta+\omega t}, X^{cu}_{\theta+\omega t})$. Of course, the space is endowed with the supremum norm. For typographical reasons, we will abbreviate the above spaces to $C^0$ and $C^0_\rho$. It is a standard result that the above spaces are Banach spaces when endowed with the above norms.

### 6.2. Some Elementary Estimates

We denote by $\mathcal{T}_1$, $\mathcal{T}_2$ the operators given by the R. H. S. of the equations (67) and (68), respectively. For typographical reasons, we just denote $\|B\| = \sup_{\theta\in D_\rho} \|B(\theta)\|_{X,Y}$.

Using just the triangle inequality and bounds on the semi-group $U^s_\theta$, we have

$$\begin{aligned}
&\|\mathcal{T}_2(\mathcal{M},\mathcal{N}) - \mathcal{T}_2(\tilde{\mathcal{M}},\tilde{\mathcal{N}})\|_{C^0} \le C\Big((1+\|\mathcal{M}\|_{\mathcal{L}_\rho})\|B\|)\|\mathcal{N}-\tilde{\mathcal{N}}\|_{C^0} \\
&\quad + \max(\|\mathcal{N}\|_{C^0}, \|\tilde{\mathcal{N}}\|_{C^0})\Big)\|\mathcal{M}-\tilde{\mathcal{M}}\|_{C^0} \\
&\quad \|\mathcal{T}_1(\mathcal{M},\mathcal{N}) - \mathcal{T}_1(\tilde{\mathcal{M}},\tilde{N})\|_{C^0} \\
&\quad \le \Big(C_h T^{-\alpha_1} e^{-\beta_1 T}\|\mathcal{M}\|_{\mathcal{L}_\rho} + C(1+\|\mathcal{M}\|_{\mathcal{L}_\rho}\|B\|\Big)\|\mathcal{N}-\tilde{\mathcal{N}}\|_{C^0_\rho} \\
&\quad + C_h T^{-\alpha_1} e^{-\beta_1 T}\|\tilde{M}-\mathcal{M}\|_{\mathcal{L}_\rho} + C\|B\|\max(\|\mathcal{N}\|_{C^0}, \|\tilde{\mathcal{N}}\|_{C^0})\|\mathcal{M}-\tilde{\mathcal{M}}\|_{C^0}.
\end{aligned}$$

Since $\|U^s_\theta\| \le A$, we choose $\mathcal{S} = \{(\mathcal{N},\mathcal{M}) \le 2A\}$. We first fix $T$ large enough so that $CT^{-\alpha_1}Te^{-\beta_1 T} \le 10^{-2}$. Then, we see that if $\|B\|$ is small enough, $(\mathcal{T}_1, \mathcal{T}_2)(\mathcal{S}) \subset \mathcal{S}$.

Furthermore, under another smallness condition in $\|B\|$, using the previous bounds, we see that $(\mathcal{T}_1, \mathcal{T}_2)$ is a contraction in $S$.

Therefore, with the above choices we can get solutions of (67) and (68) which are sufficient conditions to obtain a forward evolution and that the graph is invariant under this evolution.

### 6.3. Some Small Arguments to Finish the Construction of the Invariant Subspaces

Since we have the function $W$ defined in all $D_\rho$, it follows that the function $\mathcal{W}(t) = W(\theta+\omega t)$ is defined for all time as desired. The argument also shows that for a fixed $\theta$, the function solves the linearized equation for a short time. Of course, the argument can be done in the same way for other dichotomies running the time backwards. Hence we obtain the stability of the splittings $X^{sc}$ and $X^u$. The space $X^c$ can be reconstructed as $X^c = X^{cu} \cap X^{sc}$.

**Step 2. Estimates on the projections.** To get the bounds for the projections we use the same argument as in [31]. We only give the argument for the stable subspace. Let $\mathcal{M}^{cu}_\theta$ be the linear map whose graph gives $\tilde{X}^{cu}_\theta$.

We write

$$\Pi^s_\theta \xi = (\xi^s, 0), \qquad \tilde{\Pi}^s_\theta \xi = (\tilde{\xi}^s, \mathcal{M}^s_\theta \tilde{\xi}^s),$$
$$\Pi^{cu}_\theta \xi = (0, \xi^{cu}), \qquad \tilde{\Pi}^{cu}_\theta \xi = (\mathcal{M}^{cu}_\theta \tilde{\xi}^{cu}, \tilde{\xi}^{cu}),$$

and then

$$\xi^s = \tilde{\xi}^s + \mathcal{M}_\theta^{cu} \tilde{\xi}^{cu},$$
$$\xi^{cu} = \mathcal{M}_\theta^s \tilde{\xi}^s + \tilde{\xi}^{cu}.$$

Since $\mathcal{M}_\theta^s$ and $\mathcal{M}_\theta^{cu}$ are $O(\gamma)$ in $\mathcal{L}(X, X)$ we can write

$$\begin{pmatrix} \tilde{\xi}^s \\ \tilde{\xi}^{cu} \end{pmatrix} = \begin{pmatrix} \mathrm{Id} & \mathcal{M}_\theta^{cu} \\ \mathcal{M}_\theta^s & \mathrm{Id} \end{pmatrix}^{-1} \begin{pmatrix} \xi^s \\ \xi^{cu} \end{pmatrix}$$

and then deduce that

$$\|(\tilde{\Pi}_\theta^s - \Pi_\theta^s)\xi\|_Y \le \|(\tilde{\xi}^s - \xi^s, \mathcal{M}_\theta^s \tilde{\xi}^s)\|_Y \le C\gamma.$$

We recall that the projections depend on the whole splitting, not just on the space considered, so that the changes on $\Pi^s$ are affected by the changes in all the spaces.

**Step 3. Stability of the smoothing properties.**

In this step, we will show that the smoothing properties of the cocycles are preserved under the lower order perturbations considered before. That is, we will show that if we define the evolutions in the invariant spaces constructed in Step 1 above, they satisfy bounds of the form in **SD.3** but with slightly worse parameters. To be able to apply this repeatedly, it will be important for us to develop estimates on the change of the constants as a function of the correction.

We will first study the stable case. The unstable case is studied in the same way, just reversing the direction of time. The maps $U_\theta^s$ and $\tilde{U}_\theta^s$ satisfy the variational equations

$$\frac{\mathrm{d}U_\theta^s}{\mathrm{d}t} = A(\theta + \omega t) U_\theta^s(t)$$

and

$$\frac{\mathrm{d}\tilde{U}_\theta^s}{\mathrm{d}t} = \tilde{A}(\theta + \omega t) \tilde{U}_\theta^s(t).$$

Since $(U_\theta^s - \tilde{U}_\theta^s)(0) = 0$, one has by the variation of parameters formula

$$\tilde{U}_\theta^s(t) = U_\theta^s(t) + \int_0^t U_\theta^s(t-\tau)(\tilde{A} - A)(\theta + \omega\tau)\tilde{U}_\theta^s(\tau)\,\mathrm{d}\tau \tag{69}$$

for $t \ge 0$.

Let $\mathcal{C}_{\alpha,\beta,\rho}(X)$ be the space of continuous functions from $(0, \infty)$ into the space $\mathcal{A}_{\rho,\mathcal{L}(X,X)}$ endowed with the norm

$$|||U|||_{\alpha,\beta,\rho} = \sup_{\substack{\theta \in D_\rho \\ t > 0}} ||U(\theta(t))||_{Y,X} e^{\beta} t^{\alpha}.$$

We fix $\tilde{A}$, $A$ and $U^s_\theta$ and consider the left hand-side of (69) as an operator on $\tilde{U}^s_\theta$, that is denote

$$\mathcal{T}\bar{U}^s_\theta(t) = U^s_\theta(t) + \int_0^t U^s_\theta(t-\tau)(\tilde{A} - A)(\theta + \omega\tau)\tilde{U}^s_\theta(\tau)\,\mathrm{d}\tau.$$

Hence (69) is just a fixed point equation. We note that the operator $\mathcal{T}$ is affine in its argument. We write it as $\mathcal{T}(U^s_\theta) = \mathcal{O} + \mathcal{L}(U^s_\theta)$ where $\mathcal{O}$ is a constant vector and $\mathcal{L}$ is a linear operator. To show that $\mathcal{T}$ is a contraction, it suffices to estimate the norm of $\mathcal{L}$. We have

$$|||\mathcal{L}U_1 - \mathcal{L}U_2|||_{\alpha,\beta,\rho} \leqq C\gamma\left(t^\alpha e^{\beta t}\int_0^t \frac{e^{-\beta_1(t-\tau)}}{(t-\tau)^{\alpha_1}} e^{-\beta\tau}\tau^{-\alpha}\mathrm{d}\tau\right)|||U_1 - U_2|||_{\alpha,\beta,\rho}.$$

We now estimate

$$C(t) = t^\alpha e^{\beta t}\int_0^t \frac{e^{-\beta_1(t-\tau)}}{(t-\tau)^{\alpha_1}} e^{-\beta\tau}\tau^{-\alpha}\mathrm{d}\tau.$$

We have

$$C(t) = t^\alpha \int_0^t \frac{e^{(\beta-\beta_1)(t-\tau)}}{(t-\tau)^{\alpha_1}}\tau^{-\alpha}\mathrm{d}\tau.$$

Changing variables, one gets

$$C(t) = t^\alpha \int_0^t \frac{e^{(\beta-\beta_1)z}}{(t-z)^{\alpha}} z^{-\alpha_1}\mathrm{d}z.$$

We now choose $\beta$ such that $\beta < \beta_1$ denoting $\beta = \beta_1 - \varepsilon$. Making the change of variables $z = tu$ in the integral, one gets

$$C(t) = t^{1-\alpha_1}\int_0^1 \frac{e^{-\varepsilon tu}}{(1-u)^\alpha} u^{-\alpha_1}\mathrm{d}u.$$

This is clearly bounded for $t \leqq 1$ since $\alpha \in (0,1)$ and $1 - \alpha_1 > 0$. We now consider the case $t > 1$. There exists a constant $C >$ universal such that the following estimate holds:

$$e^{-t\varepsilon u} \leqq \frac{C}{(1 + t\varepsilon u)^{1-\alpha_1}}$$

for any $t, u \geqq 0$. Therefore we estimate for $t > 1$

$$C(t) \leqq Ct^{1-\alpha_1}\int_0^1 \frac{\mathrm{d}u}{(1-u)^\alpha u^{\alpha_1}(1+\varepsilon tu)^{1-\alpha_1}},$$

which is uniformly bounded as $t$ goes to $\infty$. Recalling that $|||\mathcal{L}U_1 - \mathcal{L}U_2|||_{\rho,\alpha_1,\beta_1} \leqq C\gamma$ where $C$ is the constant we just computed, we obtain that $\mathcal{L}$ is a contraction in the space $\mathcal{C}_{\alpha_1,\beta,\rho}(X)$ for any $\beta < \beta_1$ and any $\alpha_1 \in (0,1)$ when $\gamma$ is sufficiently small. □

The first consequence of Proposition 6.1 is that in the iterative step the small change of $K$ produces a small change in the invariant splitting and in the hyperbolicity constants.

**Corollary 6.2.** *Assume that $K$ satisfies the hyperbolic non-degeneracy Condition 3.3 and that $\|K - \tilde{K}\|_{\rho,X}$ is small enough. If we denote $\tilde{A}(\theta) = D\mathcal{X}(K)$, there exists an analytic family of splitting for $\tilde{K}$, that is*

$$X = X^s_{\tilde{K}(\theta)} \oplus X^c_{\tilde{K}(\theta)} \oplus X^u_{\tilde{K}(\theta)},$$

*which is invariant under the linearized equation* (17) *(replacing $K$ by $\tilde{K}$) in the sense that*

$$\tilde{U}^\sigma_\theta(t) X^\sigma_{\tilde{K}(\theta)} = X^\sigma_{\tilde{K}(\theta+\omega t)}. \qquad \sigma = s, c, u.$$

*We denote $\Pi^s_{\tilde{K}(\theta)}$, $\Pi^c_{\tilde{K}(\theta)}$ and $\Pi^u_{\tilde{K}(\theta)}$ the projections associated to this splitting. There exist $\tilde{\beta}_1$, $\tilde{\beta}_2$, $\tilde{\beta}_3^+, \tilde{\beta}_3^- > 0$, $\tilde{\alpha}_1, \tilde{\alpha}_2 \in (0,1)$ and $\tilde{C}_h > 0$ independent of $\theta$ satisfying $\tilde{\beta}_3^+ < \tilde{\beta}_1$, $\tilde{\beta}_3^- < \tilde{\beta}_2$ and such that the splitting is characterized by the following rate conditions:*

$$\begin{aligned}
\|\tilde{U}^s_\theta(t)\|_{\rho,Y,X} &\leqq \tilde{C}_h \frac{e^{-\tilde{\beta}_1 t}}{t^{\tilde{\alpha}_1}}, \qquad t > 0,\\
\|\tilde{U}^u_\theta(t)\|_{\rho,Y,X} &\leqq \tilde{C}_h \frac{e^{\tilde{\beta}_2 t}}{|t|^{\tilde{\alpha}_2}}, \qquad t < 0,\\
\|\tilde{U}^c_\theta(t)\|_{\rho,X,X} &\leqq \tilde{C}_h e^{\tilde{\beta}_3^+ t}, \qquad t > 0\\
\|\tilde{U}^c_\theta(t)\|_{\rho,X,X} &\leqq \tilde{C}_h e^{\tilde{\beta}_3^- |t|}, \qquad t < 0.
\end{aligned}$$

*Furthermore, the following estimates hold:*

$$\|\Pi^{s,c,u}_{\tilde{K}(\theta)} - \Pi^{s,c,u}_{K(\theta)}\|_{\rho,Y,Y} \leqq C\|\tilde{K} - K\|_{\rho,X}, \tag{70}$$

$$|\tilde{\beta}_i - \beta_i| \leqq C\|\tilde{K} - K\|_{\rho,X}, \qquad i = 1,2,3, \tag{71}$$

$$|\tilde{\alpha}_i - \alpha_i| \leqq C\|\tilde{K} - K\|_{\rho,X}, \qquad i = 1,2 \tag{72}$$

$$\tilde{C}_h = C_h. \tag{73}$$

**Proof.** We just take $A(\theta) = D\mathcal{X}(K(\theta))$, $\tilde{A}(\theta) = D\mathcal{X}(\tilde{K}(\theta))$, $X^{s,c,u}_{K(\theta)} = X^{s,c,u}_\theta$, $X^{s,c,u}_{\tilde{K}(\theta)} = \tilde{X}^{s,c,u}_\theta$, $\Pi^{s,c,u}_{K(\theta)} = \Pi^{s,c,u}_\theta$ and $\Pi^{s,c,u}_{\tilde{K}(\theta)} = \tilde{\Pi}^{s,c,u}_\theta$ in Lemma 6.1 and we use that

$$\|\tilde{A}(\theta) - A(\theta)\|_{\rho,X,Y} \leq \|\mathcal{X}\|_{C^1} \|\tilde{K}(\theta) - K(\theta)\|_{\rho,X}.$$

□

## 7. Approximate Solution of the Cohomology Equation on the Center Subspace

We now come to the solution of the projected equation (45) on the center subspace. The first point which has to be noticed is that by the spectral non-degeneracy assumption 3.3 the center subspace $X^c_\theta$ is finite-dimensional (with dimension $2\ell$). As a consequence, we end up with standard small divisors equations. This is in contrast with other studies of Hamiltonian partial differential equations like the Schrödinger equation for which there is an infinite number of eigenvalues on the imaginary axis (see [6] ) and the KAM theory is more involved. Another aspect of Definition 3.3 is that the formal symplectic structure on $X$ restricts to a standard one on the center bundle. Finally, it has to be noticed that by the finite-dimensionality assumption, all the issues related to unbounded operators do not play a role and indeed, the treatment in this section is very similar to that in the finite dimensional case in [27].

We note that we will not obtain an exact solution in the center component, but we will obtain a function that solves the right side up to a term which is quadratic in the error. This is enough for the iterative methods.

We denote

$$\Delta^c(\theta) = \Pi^c_\theta \Delta K(\theta).$$

The projected linearized equation (45) becomes

$$\partial_\omega \Delta^c(\theta) - (D\mathcal{X}) \circ K \Delta^c(\theta) = -\Pi^c_\theta E(\theta) = -E^c(\theta). \tag{74}$$

We first recall a well-known result by Rüssmann (see [25,67–69]) which allows one to solve small divisor equations along characteristics.

**Proposition 7.1.** *Assume that $\omega \in D_h(\kappa, \nu)$ with $\kappa > 0$ and $\nu \geq \ell - 1$ and that $\mathcal{M}$ is a finite dimensional space. Let $h : D_\rho \supset \mathbb{T}^\ell \to \mathcal{M}$ be a real analytic function with zero average with values in $\mathcal{M}$. Then, for any $0 < \delta < \rho$ there exists a unique analytic solution $v : D_{\rho-\delta} \supset \mathbb{T}^\ell \to \mathcal{M}$ of the linear equation*

$$\sum_{j=1}^{l} \omega_j \frac{\partial v}{\partial \theta_j} = h$$

*having zero average. Moreover, if $h \in \mathcal{A}_{\rho,\mathcal{M}}$, then $v$ satisfies the estimate*

$$\|v\|_{\rho-\delta,\mathcal{M}} \leqq C\kappa\delta^{-\nu}\|h\|_{\rho,\mathcal{M}}, \qquad 0 < \delta < \rho.$$

*The constant $C$ depends on $\nu$ and the dimension of the torus $\ell$.*

As in [31] and [27], we will find an explicit change of variables so that the vector-field $D\mathcal{X} \circ K \Delta^c(\theta)$ becomes a constant coefficient vector-field. Then we will be able to apply the small divisor result as stated in Proposition 7.1 to the cohomology equations (74).

### *7.1. Geometry of the Invariant Tori*

As is well known in KAM theory, in a finite dimensional framework, maximal invariant tori are Lagrangian submanifolds and whiskered tori are isotropic. In our context of an infinite dimensional phase space $X$, the picture is less clear, but nevertheless, thanks to our assumptions (which are satisfied in many models, including those in Section 10 11), one can produce precise geometric conclusions in the cases we consider.

We prove the following lemma on the isotropic character of approximate invariant tori:

**Lemma 7.2.** *Let $K : D_\rho \supset \mathbb{T}^\ell \to \mathcal{M}$, $\rho > 0$, be a real analytic mapping. Define the error in the invariance equation as*

$$E(\theta) := \partial_\omega K(\theta) - \mathcal{X}(K(\theta)).$$

*Let $L(\theta) = DK(\theta)^\perp J_c DK(\theta)$ be the matrix which expresses the form $K^*\Omega$ on the torus in the canonical basis.*

*There exists a constant $C$ depending on $l$, $\nu$ and $\|DK\|_\rho$ such that*

$$\|L\|_{\rho-2\delta, X_\theta^c, X_\theta^c} \leqq C\kappa\delta^{-(\nu+1)}\|E\|_{\rho,Y}, \qquad 0 < \delta < \rho/2.$$

*In particular, if $E = 0$, then*

$$L \equiv 0.$$

**Proof.** By assumption **H3.2** we have that there exists a one-form $\alpha_K$ on the torus $\mathbb{T}^\ell$ such that

$$K^*\Omega = d\alpha_K.$$

In coordinates on $\mathbb{T}^\ell$, $\alpha_K$ writes as

$$\alpha_K = g_K(\theta)d\theta.$$

Hence one has $L(\theta) = Dg_K^\perp(\theta) - Dg_K(\theta)$ and the lemma follows from Cauchy estimates and Proposition 7.1 (see also [27]). □

### *7.2. Basis of the Center Subspace $X_\theta^c$*

We introduce a suitable representation of the center subspace $X_\theta^c$. In [27,31,32] it is shown that the change of variables given by the following matrix:

$$[DK(\theta), J_c^{-1}DK(\theta)N(\theta)], \tag{75}$$

allows one to transform the linearized equations in the center subspace into two cohomology equations with approximately constant coefficients. This phenomenon is called *automatic reducibility*.

The geometric idea behind automatic reducibility is that, in the case that the torus is exactly invariant, since the motion on the torus with the chosen parameterization is a rotation, then the derivatives along the coordinates of the parameterization are preserved by the evolution. Since the symplectic form is preserved, the symplectic conjugates $J_c^{-1} DK(\theta)N(\theta)$ are also preserved (up to a shift in the directions of the tangent). In the case of approximately invariant tori, we have that the above identities hold up to errors that can be bounded by the derivatives of the error of the invariance equation.

The argument presented in the references above works word by word here thanks to the fact that the center subspace $X^c_\theta$ is finite dimensional. We will go over the main points in Section 7.3.

We will start by recalling some symplectic properties.

**7.2.1. Some Symplectic Preliminaries** We prove the following lemma:

**Lemma 7.3.** *The* $2-$*form* $\Omega$ *which is the restriction to the center subspace is non-degenerate in the sense that* $\Omega(u, v) = 0\, \forall u \in X$ *implies that* $v = 0$.

**Proof.** A quick proof would follow from the fact that the symplectic form in the center direction is non-degenerate when $K = 0$. Then, because the non-degeneracy assumptions are open (remember we are assuming that the center direction is finite-dimensional), this follows in a small neighborhood.

The argument to follow gives a more global argument valid in all the center manifold. By the non-degeneracy assumptions 3.3, there exist maps $U^{s,c,u}_\theta(t)$ generating the linearizations on $X^{s,c,u}_\theta$. These maps preserve $\Omega$. Indeed, one has the following: let $u(t)$, $v(t)$ satisfy

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = A(\theta + \omega t)u(t)$$

and

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = A(\theta + \omega t)v(t)$$

where $A(\theta) = J^{-1}\nabla^2 H \circ K(\theta)$, then

$$\Omega(u(t), v(t)) = \Omega(u(0), v(0)).$$

Indeed,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Omega(u(t), v(t)) &= \Omega(\dot{u}(t), v(t)) + \Omega(u(t), \dot{v}(t)) \\
&= < J^{-1}\nabla^2 H \circ K(\theta + \omega t)u(t), Jv(t) > \\
&\quad + < u(t), JJ^{-1}\nabla^2 H \circ K(\theta + \omega t)v(t) > \\
&= - < \nabla^2 H \circ K(\theta + \omega t)u(t), v(t) > + < u(t), \nabla^2 H \circ K(\theta + \omega t)v(t) >,
\end{aligned}$$

since $\nabla^2 H \circ K$ is symmetric. Hence the result.

Therefore, we have for any $u, v \in X^{s,c,u}_\theta$ that

$$\Omega(u, v) = \Omega(U^{s,c,u}_\theta(t)u, U^{s,c,u}_\theta(t)v), \qquad t \in \mathbb{R}^+, \mathbb{R}, \mathbb{R}^-.$$

Using now the estimates in 3.3, we have that the form $\Omega$ satisfies $\Omega(u, v) = 0$ in the following cases:

- $u, v \in X_\theta^s$,
- $u, v \in X_\theta^u$,
- $u \in X_\theta^s \cup X_\theta^u$ and $v \in X_\theta^c$,
- $v \in X_\theta^c$ and $v \in X_\theta^s \cup X_\theta^u$.

This implies that the form $\Omega$ restricted to the center bundle $X_\theta^c$ is non—degenerate and the lemma is proved. □

The form $\Omega$ is then a symplectic form since we assumed that the restriction of the form to $X_\theta^c$ is closed. Denote by $J_c$ the restriction of the operator $J$ on $X_\theta^c$. Finally we define the operator $M(\theta)$ from $\mathbb{R}^\ell$ into $X_\theta^c$:

$$M(\theta) = [DK(\theta), J_c^{-1} DK(\theta)N(\theta)]. \tag{76}$$

Notice that by assumption, $X_\theta^c$ is isomorphic to $Y_\theta^c$. We emphasize the fact that the operator $M(\theta)$ belongs to $X_\theta^c$. Indeed, it is clear from the equation that $DK$ (by just differentiating) belongs to the center space and so is $J_c^{-1} DK(\theta)N(\theta)$ by the fact that we consider the restriction $J_c$ of $J$ to apply to the center.

### *7.3. Normalization Procedure*

Let $W : D_\rho \supset \mathbb{T}^\ell \to X_\theta^c$ be such that

$$\Delta^c(\theta) = M(\theta)W(\theta).$$

From now on, the proof is very similar to the one in [27], and we just sketch the proofs. We refer the reader to [27] for the details. The first lemma provides a reducibility argument for *exact* solutions of (7). We note that since the space $X_\theta^c$ is finite dimensional the symplectic form needs to be defined only in a very weak sense.

**Lemma 7.4.** *Let $K$ be a solution of*

$$\partial_\omega K(\theta) = \mathcal{X}(K(\theta)),$$

*with $M_\theta$ be the matrix in $X_\theta^c$ defined in (76). Recall that $K(\mathbb{T}^\ell)$ is an isotropic manifold.*

*Then, there exists an $\ell \times \ell$-matrix $S(\theta)$ such that*

$$\partial_\omega M(\theta) - A(\theta)M(\theta) = M(\theta)\begin{pmatrix} 0_\ell & S(\theta) \\ 0_\ell & 0_\ell \end{pmatrix}, \tag{77}$$

*where*

$$S(\theta) = N(\theta)DK(\theta)^\top [J_c^{-1}\partial_\omega(DKN) - A(\theta)J_c^{-1}DKN](\theta),$$

*and where we have denoted $A(\theta) = J_c^{-1}D(\nabla H(K))$.*

**Proof.** By differentiating the equation, we clearly have that the first $\ell$ columns of the matrix

$$W(\theta) = A(\theta)M(\theta) - \partial_\omega M(\theta)$$

are zero. Now write

$$W_1(\theta) = A(\theta)J_c^{-1}DK(\theta)N(\theta) - J_c^{-1}\partial_\omega(DK(\theta)N(\theta)).$$

Easy computations show that

$$\begin{aligned} W_1(\theta) = A(\theta)J_c^{-1}DK(\theta)N(\theta) - J_c^{-1}\partial_\omega(DK(\theta))N(\theta) \\ + J_c^{-1}DK(\theta)N(\theta)\partial_\omega(DK^\top(\theta))N(\theta) \\ + J_c^{-1}DK(\theta)N(\theta)DK(\theta)^\top\partial_\omega(DK(\theta))N(\theta), \end{aligned}$$

but since $DK$ and $J_c^{-1}DK(\theta)N(\theta)$ form a basis of the center subspace, one can write

$$W_1 = DK\, S + J_c^{-1}DKN\, T.$$

We will prove that $T = 0$, giving the form of the matrix in the lemma. Multiply the previous equation by $DK(\theta)^\top J_c$, then by the Lagrangian character of $K$, we have

$$DK(\theta)^\top J_c W_1(\theta) = T.$$

Hence, using straightforward computations, we have that the second term plus the fourth term in $DK(\theta)^\top J_c W_1(\theta)$ is zero and the first term plus the third term in $DK(\theta)^\top J_c W_1(\theta)$ is equal to

$$\big(DK^\top D(\nabla H(K))J_c^{-1} + \partial_\omega(DK)^\top\big)DKN.$$

However, using the fact the symplectic form is skew-symmetric, the quantity in parenthesis is just the derivative of the equation, hence it has to be zero.

We now check the expression of the matrix $S$. We multiply by $NDK^\top$ to have

$$S = NDK^\top W_1 = NDK^\top\big(A(\theta)J_c^{-1}DK(\theta)N(\theta) - J_c^{-1}\partial_\omega(DK(\theta)N(\theta))\big).$$

This gives the result. □

The next lemma provides a generalized inverse for the operator $M$.

**Lemma 7.5.** *Let $K$ be a solution of* (7). *Then the matrix $M^\perp J_c M$ is invertible and*

$$(M^\perp J_c M)^{-1} = \begin{pmatrix} N^\top DK^\top J_c^{-1}DK\, N & -\mathrm{Id}_\ell \\ \mathrm{Id}_\ell & 0 \end{pmatrix}.$$

We now establish a similar result for approximate solutions, that is solutions of (7) up to error $E(\theta) = \mathcal{F}_\omega(K)(\theta)$. When $K$ is just an approximate solution, we define

$$(e_1, e_2) = \partial_\omega M(\theta) - A(\theta)M(\theta) - M(\theta)\begin{pmatrix} 0_\ell & S(\theta) \\ 0_\ell & 0_\ell \end{pmatrix}. \tag{78}$$

Using that $\partial_\omega DK(\theta) - A(\theta)DK(\theta) = DE(\theta)$ and the definition of $S$ above mentioned gives $e_1 = DE$ and $e_2 = O(\|E\|_{\rho,Y}, \|DE\|_{\rho,Y})$.

We then get

$$[\partial_\omega M(\theta) - A(\theta)M(\theta)]\xi(\theta) + M(\theta)\partial_\omega \xi(\theta) = -E^c(\theta). \tag{79}$$

For the *approximate* solutions of (7), we have the following lemma:

**Lemma 7.6.** *Assume $\omega$ is Diophantine in the sense of definition 3.1 and $\|E^c\|_{\rho,Y_\theta^c}$ small enough. Then there exist a matrix $B(\theta)$ and vectors $p_1$ and $p_2$ such that, by the change of variables $\Delta^c = M\xi$, the projected equation on the center subspace can be written as*

$$\left[\begin{pmatrix} 0_l & S(\theta) \\ 0_l & 0_l \end{pmatrix} + B(\theta)\right]\xi(\theta) + \partial_\omega \xi(\theta) = p_1(\theta) + p_2(\theta). \tag{80}$$

*The following estimates hold:*

$$\|p_1\|_{\rho,X_\theta^c} \leqq C\|E^c\|_{\rho,Y_\theta^c}, \tag{81}$$

$$\|p_2\|_{\rho-\delta,X_\theta^c} \leqq C\kappa\delta^{-(\nu+1)}\|E^c\|^2_{\rho,Y_\theta^c}, \tag{82}$$

*and*

$$\|B\|_{\rho-2\delta,X_\theta^c} \leqq C\kappa\delta^{-(\nu+1)}\|E^c\|_{\rho,Y_\theta^c}, \tag{83}$$

*where $C$ depends $l$, $\nu$, $\rho$, $\|N\|_\rho$, $\|DK\|_{\rho,Y}$, $|H|_{C^2(B_r)}$. Furthermore the vector $p_1$ has the expression*

$$p_1(\theta) = \begin{pmatrix} -N(\theta)^\top DK(\theta)^\top E^c(\theta) \\ DK(\theta)^\top J_c E^c(\theta) \end{pmatrix}.$$

**Proof.** The proof follows the one in [27] with only minor changes. Notice that, even if the original problem is infinitely dimensional, the center subspace is finitely dimensional and the result depends only on calculations in this center space. We also refer to [27] for the very simple geometric reason on the main calculation.

The main point of Lemma 7.6 are that $B$ is bounded by the error in the invariance and that $p_2$ is quadratic in the error. Hence, the main term to be solved is the equation ignoring the $B$ and $p_2$. We also anticipate that, for exact systems, the average of the second component of $p_1$ is zero (see Lemma Proposition 7.7).

From the previous computations one has

$$(e_1, e_2) = \partial_\omega M(\theta) - A(\theta)M(\theta) - M(\theta)\begin{pmatrix} 0_\ell & S(\theta) \\ 0_\ell & 0_\ell \end{pmatrix}.$$

Hence we have

$$M^\perp J_c\Big[\partial_\omega M(\theta) - A(\theta)M(\theta)\Big]\xi(\theta) = (M^\perp J_c M)\partial_\omega \xi = M^\perp J_c E_c.$$

Thus, by the previous Lemma,

$$\left[\begin{pmatrix} 0_l & S(\theta) \\ 0_l & 0_l \end{pmatrix} + (M^{\perp} J_c M)^{-1}(e_1, e_2)\right] \xi(\theta) + \partial_\omega \xi(\theta) = (M^{\perp} J_c M)^{-1} M^{\perp} J_c E_c. \tag{84}$$

Hence, denoting

$$B(\theta) = (M^{\perp} J_c M)^{-1}(e_1, e_2),$$

direct computations give $p_1$ and $p_2$ and the desired estimates. □

### *7.4. Solutions to the Reduced Equations*

We anticipate that from Lemma 7.6, the terms $B\xi$ and $p_2$ are quadratic in the error. Hence an approximate solution has the form $\xi = (\xi_1, \xi_2)$ and solves

$$\begin{aligned} S(\theta)\xi_2(\theta) - \partial_\omega \xi_1(\theta) &= -N(\theta)^\top DK(\theta)^\top E^c(\theta), \\ \partial_\omega \xi_2(\theta) &= DK(\theta)^\top J_c E^c(\theta). \end{aligned} \tag{85}$$

We prove the following result, providing a solution to equations (85):

**Proposition 7.7.** *There exists a solution* $(\xi_1, \xi_2)$ *of* (85) *with the estimates*

$$\begin{aligned} \|\xi_1\|_{\rho-\delta, X_\theta^c} &\leqq C_1 \kappa \delta^{-\nu} \|E^c\|_{\rho, X_\theta^c}, \\ \|\xi_2\|_{\rho-2\delta, X_\theta^c} &\leqq C_2 \kappa \delta^{-2\nu} \|E^c\|_{\rho, X_\theta^c} \end{aligned}$$

*for any* $\rho \in (0, \delta/2)$ *and where the constants* $C_1, C_2$ *just depend on* $l$, $\nu$, $\rho$, $\|N\|_\rho$, $\|DK\|_{\rho, X_\theta^c}$, $|\text{avg}\,(S)|^{-1}$.

**Proof.** In order to apply Prop. 7.1, one needs to study the average on the torus $\mathbb{T}^\ell$ of $DK(\theta)^\top J_c E^c(\theta)$. To do so, we first consider assumption **H3.1** which gives in coordinates

$$DK^\top J_c DK = Dg^\top - Dg$$

for some function $g$ on $\mathbb{T}^\ell$. Now taking the inner product with $\omega$ and using the equation, one has

$$DK^\top J_c (E + \mathcal{X}(K)) = Dg^\top \cdot \omega - Dg \cdot \omega.$$

Therefore, the average of $DK^\top J_c E$ is the sum of the average of $Dg^\top \cdot \omega - Dg \cdot \omega$ which is zero and the average of $DK^\top J_c \mathcal{X}(K)$. Now notice that

$$DK^\top J_c \mathcal{X}(K) = i_{\mathcal{X} \circ K} K^* \Omega(.),$$

hence its average is zero by assumption **H4**. As a consequence the average on $\mathbb{T}^\ell$ of the right hand side. $DK(\theta)^\top J_c E^c(\theta)$ is zero. Hence an application of Prop. 7.1 gives the solvability in $\xi_2$ with the desired bound. Since the average of $\xi_2$ is free, one uses it and the twist condition to solve in $\xi_1$. This gives the desired result (see [27] for more details). □

## 8. Uniqueness Statement

In this section, we prove the uniqueness part of Theorem 3.5.

We assume that the embeddings $K_1$ and $K_2$ satisfy the hypotheses in Theorem 3.5, in particular that $K_1$ and $K_2$ are solutions of (7). If $\tau \neq 0$, we write $K_1$ for $K_1 \circ T_\tau$, which is also a solution. Therefore $\mathcal{F}_\omega(K_1) = \mathcal{F}_\omega(K_2) = 0$. By Taylor's theorem we can write

$$0 = \mathcal{F}_\omega(K_1) - \mathcal{F}_\omega(K_2) = D_K\mathcal{F}_\omega(K_2)(K_1 - K_2) + \mathcal{R}(K_1, K_2), \tag{86}$$

where

$$\mathcal{R}(K_1, K_2) = \frac{1}{2}\int_0^1 D^2\mathcal{F}_\omega(K_2 + t(K_1 - K_2))(K_1 - K_2)^2 \, \mathrm{d}t.$$

Then, there exists $C > 0$ such that

$$\|\mathcal{R}(K_1, K_2)\|_{\rho,Y} \leqq C\|K_1 - K_2\|^2_{\rho,X}.$$

Hence we end up with the following linearized equation:

$$D_K\mathcal{F}_\omega(K_2)(K_1 - K_2) = -\mathcal{R}(K_1, K_2). \tag{87}$$

We denote $\Delta = K_1 - K_2$. Projecting (87) on the center subspace with $\Pi^c_{K_2(\theta+\omega t)}$, writing $\Delta^c(\theta) = \Pi^c_{K_2(\theta)}\Delta(\theta)$ and making the change of function $\Delta^c(\theta) = M(\theta)W(\theta)$, where $M$ is defined in (76) with $K = K_2$, we now perform the same type of normalization as in Section 7 at arrive at two small divisor equations of the type

$$\begin{aligned} S(\theta)\xi_2(\theta) - \partial_\omega\xi_1(\theta) &= -N(\theta)^\top DK(\theta)^\perp \mathcal{R}(0, 0, K_1, K_2)(\theta)^c, \\ \partial_\omega\xi_2(\theta) &= DK(\theta)^\top J_c\mathcal{R}(0, 0, K_1, K_2)(\theta)^c. \end{aligned} \tag{88}$$

We begin by looking for $\xi_2$. We search for it in the form $\xi_2 = \xi_2^\perp + \operatorname{avg}(\xi_2)$. We have $\|\xi_2^\perp\|_{\rho-\delta} \leqq C\kappa\delta^{-\nu}\|K_1 - K_2\|^2_{\rho,X}$.

The condition on the right-hand side of (88) to have zero average gives $|\operatorname{avg}(\xi_2)| \leq C\kappa\delta^{-\nu}\|K_1 - K_2\|^2_{\rho,X}$. Then

$$\|\xi_1 - \operatorname{avg}(\xi_1)\|_{\rho-2\delta} \leq C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|^2_{\rho,X},$$

but $\operatorname{avg}(\xi_1)$ is free. Then

$$\|\Delta^c - (\operatorname{avg}(\Delta^c)_1, 0)^\top\|_{\rho-2\delta} \leqq C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|^2_{\rho,X}.$$

The next step is done in the same way as in [27]. We quote Lemma 14 of that reference using our notation. It is basically an application of the standard implicit function theorem in finite dimensional spaces.

**Lemma 8.1.** *There exists a constant $C$ such that if $C\|K_1 - K_2\|_{\rho,X} \leqq 1$ then there exists an initial phase $\tau_1 \in \{\tau \in \mathbb{R}^\ell \mid |\tau| < \|K_1 - K_2\|_{\rho,X}\}$ such that*

$$\operatorname{avg}(T_2(\theta)\Pi^c_{K_2(\theta)}(K_1 \circ T_{\tau_1} - K_2)(\theta)) = 0.$$

The proof is just a simple application of the implicit function theorem in $\mathbb{R}^\ell$. (It suffices to compute the derivative with respect to $\tau_1$ at $\tau_1 = 0$. More details in the computation are in [27]).

As a consequence of Lemma 8.1, if $\tau_1$ is as in the statement, then $K \circ T_{\tau_1}$ is a solution of (7) such that for all $\delta \in (0, \rho/2)$ we have the estimate

$$\|W\|_{\rho-2\delta,X} < C\kappa^2\delta^{-2\nu}\|\mathcal{R}\|_\rho^2 \leqq C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_{\rho,X}^2.$$

This leads to, on the center subspace,

$$\|\Pi^c_{K_2(\theta)}(K_1 \circ T_{\tau_1} - K_2)\|_{\rho-2\delta,X} \leqq C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_{\rho,X}^2.$$

Furthermore, taking projections on the hyperbolic subspace, we have that $\Delta^h = \Pi^h_{K_2(\theta)}(K_1 - K_2)$ satisfies the estimate

$$\|\Delta^h\|_{\rho-2\delta,X} < C\|\mathcal{R}\|_{\rho,Y}.$$

All in all, we have proven the estimate for $K_1 \circ T_{\tau_1} - K_2$ (up to a change in the original constants), which is

$$\|K_1 \circ T_{\tau_1} - K_2\|_{\rho-2\delta,X} \leqq C\kappa^2\delta^{-2\nu}\|K_1 - K_2\|_{\rho,X}^2.$$

We are now in position to carry out an argument based on iteration. We can take a sequence $\{\tau_m\}_{m\geqq 1}$ such that $|\tau_1| \leq \|K_1 - K_2\|_{\rho,X}$ and

$$|\tau_m - \tau_{m-1}| \leqq \|K_1 \circ T_{\tau_{m-1}} - K_2\|_{\rho_{m-1},X}, \qquad m \geq 2,$$

and

$$\|K_1 \circ T_{\tau_m} - K_2\|_{\rho_m,X} \leqq C\kappa^2\delta_m^{-2\nu}\|K_1 \circ T_{\tau_{m-1}} - K_2\|_{\rho_{m-1},X}^2,$$

where $\delta_1 = \rho/4$, $\delta_{m+1} = \delta_m/2$ for $m \geq 1$ and $\rho_0 = \rho$, $\rho_m = \rho_0 - \sum_{k=1}^m \delta_k$ for $m \geq 1$. By an induction argument, we end up with

$$\|K_1 \circ T_{\tau_m} - K_2\|_{\rho_m,X} \leqq (C\kappa^2\delta_1^{-2\nu}2^{2\nu}\|K_1 - K_2\|_{\rho_0,X})^{2^m}2^{-2\nu m}.$$

Therefore, under the smallness assumptions on $\|K_1 - K_2\|_{\rho_0,X}$, the sequence $\{\tau_m\}_{m\geqq 1}$ converges and one gets

$$\|K_1 \circ T_{\tau_\infty} - K_2\|_{\rho/2,X} = 0.$$

Since both $K_1 \circ T_{\tau_\infty}$ and $K_2$ are analytic in $D_\rho$ and coincide in $D_{\rho/2}$ we obtain the result.

**Remark 8.1.** The argument here is very similar to the local uniquess argument in [27]. It is patterned on the fact that there esists a left inverse for the linearization modulo the shift; see [80].

There are other arguments based on the Hadamard Three-circle theorem which do not require any iteration, but they seem to require selecting some normaliation. We expect that they can be adapted to the present situation [9].

## 9. Nash–Moser Iteration

In this section, we show that, if the initial error of the approximate invariance equation (8) is small enough, the Newton procedure can be iterated infinitely many times and converges to a solution. This is somewhat standard in KAM theory given the estimates already obtained.

Let $K_0$ be an approximate solution of (7) (that is a solution of the linearized equation with error $E_0$). We define the following sequence of approximate solutions:

$$K_m = K_{m-1} + \Delta K_{m-1}, \qquad m \geq 1,$$

where $\Delta K_{m-1}$ is a solution of

$$D_K \mathcal{F}_\omega(K_{m-1})\Delta K_{m-1} = -E_{m-1}$$

with $E_{m-1}(\theta) = \mathcal{F}_\omega(K_{m-1})(\theta)$. The next lemma provides that the solution at step $m$ improves the solution at step $m-1$ and the norm of the error at step $m$ is bounded in a smaller complex domain by the square of the norm of the error at step $m-1$.

**Proposition 9.1.** *Assume that $K_{m-1} \in ND(\rho_{m-1})$ is an approximate solution of equation (7) and that the following holds:*

$$r_{m-1} = \|K_{m-1} - K_0\|_{\rho_{m-1},X} < r.$$

*If $E_{m-1}$ is small enough such that Proposition 7.6 applies, that is*

$$C\kappa\delta_{m-1}^{-\nu-1}\|E_{m-1}\|_{\rho_{m-1},Y} < 1/2$$

*for some $0 < \delta_{m-1} \leq \rho_{m-1}/3$, then there exists a function $\Delta K_{m-1} \in \mathcal{A}_{\rho_{m-1}-3\delta_{m-1},X}$ for some $0 < \delta_{m-1} < \rho_{m-1}/3$ such that*

$$\|\Delta K_{m-1}\|_{\rho_{m-1}-2\delta_{m-1},X} \leqq \left(C^1_{m-1} + C^2_{m-1}\kappa^2\delta_{m-1}^{-2\nu}\right)\|E_{m-1}\|_{\rho_{m-1},Y}, \tag{89}$$

$$\|D\Delta K_{m-1}\|_{\rho_{m-1}-3\delta_{m-1},X} \leqq \left(C^1_{m-1}\delta_{m-1}^{-1} + C^2_{m-1}\kappa^2\delta_{m-1}^{-(2\nu+1)}\right)\|E_{m-1}\|_{\rho_{m-1},Y}, \tag{90}$$

*where $C^1_{m-1}$, $C^2_{m-1}$ depend only on $\nu$, $l$, $|\mathcal{X}|_{C^1(B_r)}$, $\|DK_{m-1}\|_{\rho_{m-1},X}$, $\|\Pi^s_{K_{m-1}(\theta)}\|_{\rho_{m-1},Y^s_\theta,X}$, $\|\Pi^c_{K_{m-1}(\theta)}\|_{\rho_{m-1},Y^c_\theta,X}$, $\|\Pi^u_{K_{m-1}(\theta)}\|_{\rho_{m-1},Y^u_\theta,X}$, and $|\mathrm{avg}\,(S_{m-1})|^{-1}$. Moreover, if $K_m = K_{m-1} + \Delta K_{m-1}$ and*

$$r_{m-1} + \left(C^1_{m-1} + C^2_{m-1}\kappa^2\delta_{m-1}^{-2\nu}\right)\|E_{m-1}\|_{\rho_{m-1},Y} < r,$$

*then we can redefine $C^1_{m-1}$ and $C^2_{m-1}$ and all previous quantities such that the error $E_m(\theta) = \mathcal{F}_\omega(K_m)(\theta)$ satisfies (defining $\rho_m = \rho_{m-1} - 3\delta_{m-1}$)*

$$\|E_m\|_{\rho_m,Y} \leqq C_{m-1}\kappa^4\delta_{m-1}^{-4\nu}\|E_{m-1}\|^2_{\rho_{m-1},Y}. \tag{91}$$

**Proof.** We have $\Delta K_{m-1}(\theta) = \Pi_\theta^h \Delta K_{m-1}(\theta) + \Pi_\theta^c \Delta K_{m-1}(\theta)$, where $\Pi_\theta^h$ is the projection on the hyperbolic subspace and belong to $\mathcal{L}(Y_\theta^h, X)$. Estimates (48) follow from the previous two sections. The second part of estimate (48) follows from the first line of (48), Cauchy's inequalities and the fact that the projected equations on the hyperbolic subspace are exactly solved. □

Thanks to the previous proposition, one is able to obtain the convergence of the Newton method in a standard way.

The other non-degeneracy conditions can be checked in exactly the same way as described in [31] and we do not repeat the arguments.

**Lemma 9.2.** *If $\|E_{m-1}\|_{\rho_{m-1},Y_\theta^c}$ is small enough, then we have the following:*

- *If $DK_{m-1}^\perp DK_{m-1}$ is invertible with inverse $N_{m-1}$ then*

$$DK_m^\perp DK_m$$

*is invertible with inverse $N_m$ and we have*

$$\|N_m\|_{\rho_m} \leqq \|N_{m-1}\|_{\rho_{m-1}} + C_{m-1}\kappa^2 \delta_{m-1}^{-(2\nu+1)} \|E_{m-1}\|_{\rho_{m-1},Y_\theta^c}.$$

- *If* avg $(S_{m-1})$ *is non-singular then also* avg $(S_m)$ *is and we have the estimate*

$$|\text{avg}\,(S_m)|^{-1} \leqq |\text{avg}\,(S_{m-1})|^{-1} + C'_{m-1}\kappa^2 \delta_{m-1}^{-(2\nu+1)} \|E_{m-1}\|_{\rho_{m-1},Y_\theta^c}.$$

## 10. Construction of Quasi-Periodic Solutions for the Boussinesq Equation

This section is devoted to an application of Theorem 3.5 to a concrete equation that has appeared in the literature.

In Section 10.1, we will verify the formal hypothesis of the general Theorem 3.5. First we will verify the geometric hypothesis, choose the concrete spaces that will play the role of the abstract ones, etc.. In Section 10.5, we will construct approximate solutions that satisfy the quantitative properties. By applying Theorem 3.5, to these approximate solutions, we will obtain Theorem 3.6.

### *10.1. Formal and Geometric Considerations*

The Boussinesq equation has been widely studied in the context of fluid mechanics since the pioneering work [5]. It is the equation (in one dimension) with periodic boundary conditions

$$u_{tt} = \mu u_{xxxx} + u_{xx} + (u^2)_{xx} \quad \text{on } \mathbb{T},\ \ t \in \mathbb{R}, \tag{92}$$

where $\mu > 0$ is a parameter.

We will introduce an additional parameter $\varepsilon$ which will be useful in the sequel as a mnemonic device to perform perturbation theory. Note however that it can be eliminated by rescaling the $u$, considering $v = \varepsilon u$, so that discussing small $\varepsilon$ is equivalent to discussing small amplitude equations.

The equation (92) is ill-posed in any space and one can construct initial data for which there is no existence in any finite interval of time. As we will see later, the non-linear term does not make it well posed in the spaces $X$ we will consider later.

The equation (92) is a 4th order equation in space. Since it is second order in time, it is standard to write it as a first order system

$$\dot{z} = \mathcal{L}_\mu z + \mathcal{N}(z), \tag{93}$$

where

$$\mathcal{L}_\mu = \begin{pmatrix} 0 & 1 \\ \partial_x^2 + \mu \partial_x^4 & 0 \end{pmatrix}$$

and

$$\mathcal{N}(z) = (0, \partial_x^2 u^2).$$

Notice that (93) has the structure we assumed in (5), namely that the evolution operator is the sum of a linear and constant operator and a nonlinear part, which is of lower order than the linear part.

### *10.2. Choice of Spaces*

In this section we present some choices of spaces $X$,$Y$ for which the operators entering in the Boussinesq equation satisfy the assumptions of Theorem 3.5. As indicated in Section 3.6.2, there are several choices and it is advantageous to make different choices for the local uniqueness part and a different one for the existence (local uniqueness results are stronger if one makes them in very general spaces and existence results are stronger if you make them in small spaces). Notice that combining both, we can get automatic bootstraps of regularity. The spaces we consider will have one free parameter.

For $\rho > 0$, we denote

$$D_\rho = \left\{ z \in \mathbb{C}^\ell / \mathbb{Z}^\ell \mid |\operatorname{Im} z_i| < \rho \right\},$$

and denote $H^{\rho,m}(\mathbb{T})$ for $\rho > 0$ and $m \in \mathbb{N}$, the space of analytic functions $f$ in $D_\rho$ such that the quantity

$$\|f\|_{\rho,m}^2 = \sum_{k \in \mathbb{Z}} |f_k|^2 e^{4\pi\rho|k|} (|k|^{2m} + 1)$$

is finite, and where $\{f_k\}_{k\in\mathbb{Z}}$ are the Fourier coefficients of $f$. For any $\rho > 0$ and $m \in \mathbb{N}$, the space $\left(H^{\rho,m}(\mathbb{T}), \|\cdot\|_{\rho,m}\right)$ is a Hilbert space. Furthermore, this scale of Hilbert spaces $H^{\rho,m}(\mathbb{T})$ for $\rho > 0$ and $m > \frac{1}{2}$ is actually a Hilbert algebra for pointwise multiplication, that is for every $u, v \in H^{\rho,m}(\mathbb{T})$ there exists a constant $C$ such that

$$\|u\, v\|_{\rho,m} \leqq C \|u\|_{\rho,m} \|v\|_{\rho,m}.$$

Extending the definition to $\rho = 0$, $H^{0,m}(\mathbb{T})$ is the standard Sobolev space on the torus and for $\rho > 0$, $H^{\rho,m}(\mathbb{T})$ consists of analytic functions on the extended strip $D_\rho$ with some $L^2$-integrability conditions on the derivatives up to order $m$ on the strip $D_\rho$. As already noticed, we are going to construct quasi-periodic solutions in the class of small amplitude solutions for (92).

For the system (93), it is natural to consider the space for $\rho > 0$ and $m > \frac{5}{2}$

$$X_{\rho,m} = H^{\rho,m} \times H^{\rho,m-2}. \tag{94}$$

We note that $\mathcal{L}_\mu$ sends $X_{\rho,m}$ into $X_{\rho,m-2}$, but we observe that this is not really used in Theorem 3.5. By the Banach algebra property of the scale of spaces $H^{\rho,m}(\mathbb{T})$ when $m > 1/2$ and the particular form of the nonlinearity, we have the following proposition (see [26]).

**Proposition 10.1.** *The non linearity $\mathcal{N}$ is analytic from $X_{\rho,m}$ into $X_{\rho,m}$ when $m > 5/2$.*

In the system language, it is useful to think of $\mathcal{L}_\mu$ as an operator of order 2 and of $\mathcal{N}$ as an operator of order 0, hence, in the present case, we can take $Y = X$ in the abstract Theorem 3.5.

**Remark 10.1.** Note that this gives a rigorous proof that the nonlinear evolution is ill-posed. If the non-linear evolution was well-posed in some of the $X_{\rho,m}$ spaces with $m > 5/2$, we could consider the nonlinear evolution as a perturbation of the linear one. Using the usual Duhamel formula of Lipschitz perturbations of semigroups [40], we could conclude that the linear evolution is well posed, which is patently false.

We will be actually considering a subspace of $X$ denoted $X_0$ consisting of functions $z(t) \in X$ such that

$$\int_0^1 \mathrm{d}x\, z(\cdot, x)\, \mathrm{d}x = 0, \tag{95}$$

$$\int_0^1 \mathrm{d}x\, \partial_t z(\cdot, x) = 0, \tag{96}$$

$$z(\cdot, x) = z(\cdot, -x). \tag{97}$$

At the formal level, the subspace $X_0$ is invariant under the equation of (93). In contrast with the normalizations (95) and (96) that can be enforced by a change of variables, (97) is a real restriction. It is possible to develop a theory without (97), but we will not pursue it here.

We now check that the assumptions of Theorem 3.5 are met. The main steps are to verify the formal assumptions of Theorem 3.5 and construct approximate solutions which are non degenerate.

### *10.3. Linearization Around* 0

We first study the eigenvalue problem for $U \in X, \sigma \in \mathbb{C}$

$$\mathcal{L}_\mu U = \sigma U.$$

This leads to the eigenvalue relation

$$\sigma^2 = -4\pi^2 k^2 + 16\pi^4 \mu k^4 = -4\pi^2 k^2 (1 - 4\pi^2 \mu k^2)$$

for $k \in \mathbb{Z}$. By symmetry, we assume that $k \geqq 0$ and the spectrum follows by reflection with respect to the imaginary axis. We have the following lemma:

**Lemma 10.2.** *The operator $\mathcal{L}_\mu$ has discrete spectrum in $X$. Furthermore, we have the following:*

- *The center spectrum of $\mathcal{L}_\mu$ consists of a finite number of eigenvalues. Furthermore, the dimension of the center subspace is even.*
- *The hyperbolic spectrum is well separated from the center spectrum.*

**Proof.** From the equation,

$$\sigma^2 = -4\pi^2 k^2 + 16\pi^4 \mu k^4 = -4\pi^2 k^2 (1 - 4\pi^2 \mu k^2),$$

we deduce easily that the spectrum is discrete in $X$. Furthermore, 0 is not an eigenvalue since we assume $u$ to have average 0. Finally, we notice that when $0 < k^2 < \frac{1}{4\pi^2 \mu}$, one has $\sigma^2 < 0$ and since there is a finite (even) number of values in this set, this leads to the desired result. The separation of the spectrum directly follows from the discreteness of the spectrum. □

We then have the following set of eigenvalues:

$$\operatorname{Spec}(\mathcal{L}_\mu) = \left\{ \pm 2\pi i |k| \sqrt{1 - 4\pi^2 \mu k^2} = \pm \sigma_k(\mu) \right\}_{k \geqq 1}.$$

The center space $X_0^c$ is the eigenspace generated by the eigenfunctions corresponding to the eigenvalues $\sigma_k(\mu)$ for which we have $1 - 4\pi^2 \mu k^2 \geq 0$. The center subspace $X_0^c$ is spanned by the eigenvectors

$$U_k = (u_k, v_k) = (\cos(2\pi k x), \sigma_k(\mu) \cos(2\pi k x))_{k=1,\dots,\ell}.$$

Any element $U$ on the center subspace can be expressed as

$$U = \sum_{k=1}^{\ell} \alpha_k U_k,$$

with the $\alpha_k$ being arbitrary real numbers.

### *10.4. Verifying the Smoothing Properties of the Partial Evolutions of the Linearization Around* 0

We now come to the evolution operators and their smoothing properties. We have

**Lemma 10.3.** *The operator* $\mathcal{L}_\mu$ *generates semi-group operators* $U_\theta^{s,u}(t)$ *in positive and negative times. Furthermore, the following estimates hold:*

$$\|U_\theta^s(t)\|_{X,X} \leqq \frac{C}{t^{\frac{1}{2}}} e^{-Dt}, t > 0$$

*and*

$$\|U_\theta^u(t)\|_{X,X} \leqq \frac{C'}{|t|^{\frac{1}{2}}} e^{D't}, t < 0$$

*for some constants* $C, C', D, D' > 0$.

**Proof.** The proof is given in detail in [26, page 404–405]. It is based on observing that the evolution operator in the (un)stable spaces can be expressed in Fourier series. Since the norms considered are given by the Fourier terms (with different weights), it suffices to estimate the sup of the multipliers times the ratio of the weights. □

Until now, we have considered only the linearization around the equilibrium 0 in $X$. Of course, by the stability theory of the splittings developed in Section 6, the spectral non-degeneracy properties will be satisfied by all the approximate solutions that are small enough in the smooth norms. As we will see, our approximate solutions will be trigonometric polynomials with small coefficients.

### *10.5. Construction of an Approximate Solution*

This section is devoted to the construction of an approximate non-degenerate solution for equation (93). We use a Lindstedt series argument to construct approximate solutions for all *"nonresonant"* values of $\mu$. Then, we will verify the twist non-degeneracy conditions for some values of $\mu$ only.

**Remark 10.2.** For the experts, we note that the analysis is remarkably similar to the perturbative analysis near elliptic fixed points in Hamiltonian systems. We have found useful the treatment in [63, Vol 2]. More modern treatments based on transformation theory are in [29,58,79]. In our case, the transformation theory is problematic (we have to make changes of variables in unbounded operators), hence we will use the more elementary Lindstedt procedure to produce approximate solutions. These approximate solutions can be the starting points of our a-posteriori theorem.

The following result establishes the existence (and some uniqueness which we will not use) of the Lindstedt series under appropriate non-resonance conditions:

**Lemma 10.4.** *Let $\ell$ be as before. For all $N \geq 2$, assume the non-resonance condition to order $N$ given by*

$$F_\mu(k, j) \neq 0, \qquad k \in \mathbb{Z}^\ell, j \in \mathbb{N}, 1 < |k| \leq N, \tag{98}$$

*where*

$$F_\mu(k, j) \equiv \left[(\omega^0 \cdot k)^2 - 2\pi^2(j^2 - 2\mu\pi^2 j^2)\right]$$

*and*

$$\omega^0(k) = 2\pi|k|\sqrt{1 - 4\pi^2\mu|k|^2}.$$

*Then, for all $\mathcal{U}_1$ depending on $\ell$ parameters, there exist $(\omega^1, ..., \omega^N) \in (\mathbb{R}^\ell)^N$ and $(\mathcal{U}_2, ..., \mathcal{U}_N) \in (H^{\rho,m}(\mathbb{T}))^{N-1}$ parameterized by $(A_1^1, ..., A_\ell^1) \in \mathbb{R}^\ell$ for any $\rho > 0$ such that for any $\sigma \geqq 0$,*

$$\|(u_\varepsilon^{[\leq N]})_{tt} - (u_\varepsilon^{[\leq N]})_{xx} - \mu(u_\varepsilon^{[\leq N]})_{xxxx} - ((u_\varepsilon^{[\leq N]})^2)_{xx}.\|_{H^{\rho,m}(\mathbb{T})} \leqq C\varepsilon^{N+1}$$

*for some constant $C > 0$ and*

$$u_\varepsilon^{[\leq N]}(t, x) = \sum_{k=1}^{N} \varepsilon^k \mathcal{U}_k(\omega_\varepsilon^{[\leqq N]} t, x)$$

*where*

$$\omega_\varepsilon^{[\leqq N]} = \omega^0 + \sum_{k=1}^{N} \varepsilon^k \omega^k.$$

*The coefficients $\mathcal{U}_k$ are trigonometric polynomials and can be obtained in such a way that the projection over the kernel of*

$$\mathcal{M}_0 = (\omega^0 \cdot \partial_\theta)^2 - \partial_{xx}^2 - \mu\partial_{xxxx}^4$$

*is zero. Moreover, the normalizations (99) and (95) are satisfied. With such a normalization, they are unique.*

**Remark 10.3.** Note that for fixed $(k, j)$ the expression $F_\mu(k, j)$ is an analytic function of $\mu$. Therefore, the assumption (98) holds except for a finite number of values of $\mu$ in the interval considered. Of course, for small values of $(k, j)$ this is easy to verify explicitly for large ranges of $\mu$. As we will see later, for the application of Theorem 3.5, it suffices to consider just a rather low order $N$ and hence the number of $(k, j)$ considered is just a small multiple of $\ell^2$.

**10.5.1. Generalities on Lindstedt Series and Proof of Lemma 10.4** Before going into the proof itself, we comment a bit on the theory of Lindstedt series. We define the hull function as

$$u_\varepsilon(t, x) = \mathcal{U}_\varepsilon(\omega_\varepsilon t, x),$$

where $\mathcal{U}_\varepsilon : \mathbb{T}^\ell \times \mathbb{T} \mapsto \mathbb{R}$ with $\ell = \frac{\dim X_0^c}{2}$.

There are two versions of the theory: one assuming the symmetry condition for the solutions

$$\mathcal{U}_\varepsilon(\theta, \cdot) = \mathcal{U}_\varepsilon(-\theta, \cdot), \tag{99}$$

and another one without assuming (99).

To avoid making the paper longer, we will only consider symmetric solutions here, but we hope that the case of general solutions will be taken up.

We note that, thanks to the a-posteriori format of the theorem, we only need to produce an approximate solution and verify the non-degeneracy conditions.

The function $\mathcal{U}_\varepsilon$ and the frequency $\omega_\varepsilon$ produce a solution of (93) if and only if they satisfy the equation

$$(\omega_\varepsilon \cdot \partial_\theta)^2 \mathcal{U}_\varepsilon = \partial_{xx}^2 \mathcal{U}_\varepsilon + \mu \partial_{xxxx}^4 \mathcal{U}_\varepsilon + (\mathcal{U}_\varepsilon^2)_{xx}. \tag{100}$$

We emphasize that we are considering now that both $\mathcal{U}_\varepsilon$ and $\omega_\varepsilon$ are unknowns to be determined in (100). As we will see, we will obtain $\mathcal{U}_\varepsilon$ and $\omega_\varepsilon$, depending on $\ell$ free arbitrary parameters.

Following the standard procedure of Lindstedt series, we will consider formal expansions $\mathcal{U}_\varepsilon$ and $\omega_\varepsilon$ in powers of $\varepsilon$. We will impose that finite order truncations to order $N$ satisfy the equation (100) up to an error $C_N|\varepsilon|^{N+1}$. Hence, the series are not meant to converge (in general they will not) but they indicate a sequence of approximate solutions that solve the equation to higher and higher order in $\varepsilon$. We will also verify the other non-degeneracy hypothesis of Theorem 3.5.

We consider the formal sums

$$\begin{aligned} \mathcal{U}_\varepsilon(\theta, x) &\sim \sum_{k=1}^{\infty} \varepsilon^k \mathcal{U}_k(\theta, x) \\ \omega_\varepsilon &\sim \omega^0 + \sum_{k=1}^{\infty} \varepsilon^k \omega^k. \end{aligned} \tag{101}$$

**Remark 10.4.** Notice that the sum for $\mathcal{U}_\varepsilon$ starts with $\varepsilon$ since we have in mind to consider small amplitude solutions of the equation.

The meaning of formal power solutions is that we truncate these sums at order $N$ arbitrary, $N \geqq 1$ and consider

$$\begin{aligned} u_\varepsilon^{[\leqq N]}(\theta, x) &= \sum_{k=1}^{N} \varepsilon^k \mathcal{U}_k(\theta, x) \\ \omega_\varepsilon^{[\leqq N]} &= \omega^0 + \sum_{k=1}^{N} \varepsilon^k \omega^k. \end{aligned}$$

As often happens in Lindstedt series theory, the first terms of the recursion are different from the others. In our case, the first step will allow us to choose solutions of the first step depending on $\ell$ parameters. Once these solutions are chosen, we can obtain all the other solutions in a unique way. We note that the computations are very algorithmic and subsequently can be programmed. The normalization in the last item of Lemma 10.4 is natural in Lindstedt series theory. If one changes the parameters, introducing new parameters $A_i^1 = B_i^1 + \varepsilon \hat{A}_i(B_1^1, \dots, B_\ell^1; \varepsilon)$, one obtains a totally different series, which of course parameterizes the same set of solutions. In any case, we emphasize that for us the main issue is to construct an approximate solution.

**Proof.** We substitute the sums for $\omega_\varepsilon$ and $\mathcal{U}_\varepsilon$ into (100) and identify at all orders.

**Order** 1: We get

$$(\omega_0 \cdot \partial_\theta)^2 \mathcal{U}_1 = \partial^2_{xx}\mathcal{U}_1 + \mu \partial^4_{xxxx}\mathcal{U}_1.$$

We search for solutions of the form $\cos(2\pi\omega_j^0\theta_j)\cos(2\pi jx)$ where $j \in \mathbb{N}$. Therefore the frequencies are given by the relation

$$\omega_j^0 = 2\pi|j|\sqrt{1 - 4\pi^2\mu j^2}.$$

We assume now that $4\pi^2\mu j^2 \neq 1$ and $1 - 4\pi^2\mu j^2 \geqq 0$, which means that $j = 1, \dots, \ell$ where $\ell = \lfloor \sqrt{\frac{1}{2\pi\mu}} \rfloor$.

Now, we get the frequency vector $\omega^0$, given by

$$(\omega^0)_{j=1,\dots,\ell} = \left(2\pi|j|\sqrt{1 - 4\pi^2\mu j^2}\right)_{j=1,\dots,\ell.}. \tag{102}$$

All the solutions of the equation satisfying the symmetry conditions (38), (99) are given by

$$\mathcal{U}_1(\theta, x) = \sum_{j=1}^{\ell} A_j^1 \cos(2\pi\theta_j)\cos(2\pi jx). \tag{103}$$

This is the customary analysis of the linearized equations in normal modes. For future reference, we denote

$$\mathcal{M}_0 = (\omega_0 \cdot \partial_\theta)^2 - \partial^2_{xx} - \mu\partial^4_{xxxx}.$$

We note that the operator $\mathcal{M}_0$ is diagonal on trigonometric polynomials and we have that

$$\mathcal{M}_0 \cos(2\pi k \cdot \theta)\cos(2\pi jx) = F_\mu(k, j)\cos(2\pi k \cdot \theta)\cos(2\pi jx),$$

where

$$F_\mu(k, j) \equiv \left[(\omega^0 \cdot k)^2 - (2\pi)^2(j^2 - \mu(2\pi)^2 j^2)\right].$$

For convenience, we will apply the important **non-resonance condition** to order $N$ introduced in (98). This non-resonance condition is very customary in the study of elliptic fixed point; it says that the basic frequencies are not a combination of each other. The following remark is obvious, but it will be useful for us later:

**Proposition 10.5.** *Under the non-resonance condition* (98)*, the kernel of the operator* $\mathcal{M}_0$ *is precisely* $\omega^0 \cdot \partial_\theta$ *of the span of the solutions* $\mathcal{U}_1$ *obtained before in* (103).

**Order** $m \geqq 2$**:** The equation to be solved expanding (100) to order $m$ has the form

$$\mathcal{M}_0\mathcal{U}_m + 2(\omega^{m-1} \cdot \partial_\theta)(\omega^0 \cdot \partial_\theta)\mathcal{U}_1 = \mathcal{R}_m(\mathcal{U}_1, ..., \mathcal{U}_{m-1}, \omega^0, ..., \omega^{m-2}), \quad (104)$$

where $\mathcal{R}_m$ is polynomial in its arguments and their derivatives (up to order 4). In particular, if $\mathcal{U}_1, ..., \mathcal{U}_{m-1}$ are trigonometric polynomials then so is $\mathcal{R}_m$. It is also easy to see that if $\mathcal{U}_1, ..., \mathcal{U}_{m-1}$ have the symmetry properties (99), so does $\mathcal{R}_m$. Hence, using the addition formula for the products of angles, we can state

$$\mathcal{R}_m = \sum_{k\in\mathbb{Z}^\ell, j\in\mathbb{Z}} C_{k,j}(A_1^0, \ldots, A_\ell^0)\cos(2\pi k \cdot \theta)\cos(2\pi jx).$$

We inductively assume that $\mathcal{U}_1, \ldots, \mathcal{U}_{m-1}$ are trigonometric polynomials and that $\omega^0, \ldots, \omega^{m-2}$ have been found. Then, we will show that we can find $\omega^{m-1}$, $\mathcal{U}_m$ in such a way that the equation (104) is solvable. Furthermore, the solution is unique if we impose the normalization at the end of Lemma 10.4. The equation (104) can be solved by identifying the coefficients of $\cos(2\pi k \cdot \theta)\cos(2\pi jx)$ on both sides.

The analysis of (104) can be summarized as follows: we note that under the non-resonance condition (98) all the terms in $\mathcal{R}_m$ can be separated in a unique way into terms in the kernel of $\mathcal{M}_0$ and the range of $\mathcal{M}_0$. Since we observe that $\mathcal{M}_0\mathcal{U}_m$ belongs to the range of $\mathcal{M}_0$ and $(\omega^{n+1} \cdot \partial_\theta)(\omega^0 \cdot \partial_\theta)\mathcal{U}_1$ belongs to the kernel of $\mathcal{M}_0$, we see that equation (104) is equivalent to

$$\mathcal{M}_0\mathcal{U}_m = \Pi_{Range(\mathcal{M}_0)}\mathcal{R}_m$$

and

$$(\omega^{n+1} \cdot \partial_\theta)(\omega^0 \cdot \partial_\theta)\mathcal{U}_1 = \Pi_{\mathrm{Ker}\,(\mathcal{M}_0)}\mathcal{R}_m.$$

We can see that these equations have unique solutions. Note that one of the consequences of the above analysis is that the solutions are unique. Furthermore, we note that the change in frequencies $\omega^n$ as functions of the amplitudes are caused by the appearance of terms in the kernel of $\mathcal{M}_0$ in $\mathcal{R}_m$. The terms in Ker $(\mathcal{M}_0)$ that lead to a shift in the frequencies are called *resonant terms*.

Since $\mathcal{M}_0$ is diagonal, the terms in the kernel of $\mathcal{M}_0$ are precisely those that are not in the range of $\mathcal{M}_0$. For the terms for which the multiplier $F_\mu(k, j)$ is non zero (that is those terms in the range of $\mathcal{M}_0$), we can invert $\mathcal{M}_0$ and hence obtain

$$\mathcal{U}_m(k, j) = \frac{C_{k,j}}{F_\mu(k, j)}.$$

For the terms that lie in the kernel of $\mathcal{M}_0$, we cannot divide by the multiplier $F(k, j)$ but instead obtain $\omega^{m-1}$ to solve (104). Note that this uses the non-resonance condition so that that the kernel of $\mathcal{M}_0$ is precisely functions that appear in $\mathcal{U}_1$.

Of course, to solve (104), we could add any function in the kernel of $\mathcal{M}_0$. Under the normalization condition, we see that the term to add is uniquely determined to be zero. The evaluation of the norm in the Lemma comes directly from the fact that we are dealing with trigonometric polynomials, hence belonging to any Sobolev space. □

*10.6. Application of Theorem 3.5 to the Approximate Solutions. End of the Proof of Theorem 3.6*

Let $\omega^0$ as in Theorem 3.6 and consider $\mathcal{U}_\varepsilon$ the function constructed in the previous section. Denote

$$K_0(\theta) = \begin{pmatrix} \mathcal{U}_\varepsilon(\theta, .) \\ \omega_\varepsilon \cdot \partial_\theta \mathcal{U}_\varepsilon(\theta, .) \end{pmatrix} \in X_0.$$

.

We will proceed to verify the assumptions of Theorem 3.5 taking as initial conditions of the iteration the results of the Lindstedt series. This will require carrying out explicitly the calculations indicated before to order 3 and verifying that the twist condition is satisfied.

**10.6.1. Smallness Assumption on the Error and Range of $K_0$** Consider $K_0$ as above. Then Lemma 10.4 ensures directly that the smallness assumption in Theorem 3.5 are satisfied with an error smaller than $C_N|\varepsilon|^{N+1}$ for arbitrary large $N$.

Note that this is verified for all values of $\ell$.

**10.6.2. Spectral Non-degeneracy** We check conditions 3.3. For $\varepsilon = 0$, all the conditions in 3.3 are met by the previous discussion. In particular there exists an invariant splitting denoted

$$X_0 = X_0^c \oplus X_0^s \oplus X_0^u. \tag{105}$$

Now, by construction of $K_0$, choosing $\varepsilon$ small enough again and using the perturbation theory of the bundles developed in section 6 (see Lemma 6.2), there exists an invariant splitting for $K_0$ for $\varepsilon$ small enough satisfying all the desired properties and this proves the spectral non-degeneracy conditions 3.3 for $K_0$, together with the suitable estimates.

Note that this is verified for all values of $\ell$.

**10.6.3. Twist Condition** We now check the twist condition in Definition 3.4. Pick a Diophantine frequency $\omega$ as in Theorem 3.6. Recall that the family of perturbative solutions is parameterized by $A_j^1$ for $j = 1, ..., \ell$, the $\ell$ parameters giving $\mathcal{U}_1$. In the system of coordinates given by $(A_1^1, \dots, \theta)$, the twist condition amounts to showing that

$$|\det\Big(\partial_{A_j^1}\omega_i^N\Big)|^{-1} > T_N(\varepsilon) > 0. \tag{106}$$

If we can show that $T_N(\varepsilon) > C|\varepsilon|^a$ for some positive $a, C$, $(1 \le a < N)$ then we claim that we can finish the construction. The crucial remark is that we also have

$$T_{\tilde{N}}(\varepsilon) \ge \tilde{C}|\varepsilon|^a$$

for any $\tilde{N} > N$, since we are only adding higher order terms.

**10.6.4. The Case $\ell = 1$** In this section, we study the case when $\ell = 1$. This case corresponds to periodic orbits and the linearized equations do not need analysis with small divisors. Nevertheless, we note that the result of the existence of periodic orbits is not trivial since the equation is ill-posed; regardless, it can be reduced to a fixed point [12]. Note that the methods of this section give analytic self-contained proofs of the existence of branches and we obtain formulas for the change of frequency with the amplitude. From the point of view of this paper, this section is useful to prepare the analysis of the general case for which we can get sharper results and it can be used only as motivation.

As we will see, $\omega^1 = 0$, so we will have to go to order 3. Let us first consider the case $m = 2$. We have that the equation at order 2 and assuming that $\ell = 1$ writes

$$\mathcal{M}_0\mathcal{U}_2 + 2(\omega^1 \cdot \partial_\theta)(\omega^0 \cdot \partial_\theta)\mathcal{U}_1 = (\mathcal{U}_1^2)_{xx}.$$

We have

$$\mathcal{U}_1^2 = A^2\cos^2(2\pi\theta)\cos^2(2\pi x),$$

Which yields

$$\mathcal{U}_1^2 = \frac{A^2}{4}(1+\cos(4\pi\theta))(1+\cos(4\pi x))$$

and

$$(\mathcal{U}_1^2)_{xx} = -4\pi^2A^2(1+\cos(4\pi\theta))\cos(4\pi x).$$

Since this is not in the range, one has $\omega^1 = 0$. We then go to order $m = 3$, which gives the equation (taking into account that $\omega^1 = 0$)

$$\mathcal{M}_0\mathcal{U}_3 + 2(\omega^0 \cdot \partial_\theta)(\omega^2 \cdot \partial_\theta)\mathcal{U}_1 = 2(\mathcal{U}_1\mathcal{U}_2)_{xx}.$$

From the previous step, one has

$$\mathcal{U}_2 = -4\pi^2A^2\Big(\frac{\cos(4\pi x)\cos(4\pi\theta)}{F(2,2)} + \frac{\cos(4\pi x)}{F(0,2)}\Big).$$

Hence we have

$$(\mathcal{U}_1\mathcal{U}_2)_{xx} = -4\pi^2A^4\Big(-\cos(2\pi x) - 9\cos(6\pi x)\Big)$$
$$\Big(\frac{-(\cos(2\pi\theta)+\cos(6\pi\theta))}{4F(2,2)} + \frac{\cos(2\pi\theta)}{2F(0,2)}\Big).$$

Identifying, according to the discussion before, one gets that $\omega^2$ is given by

$$\omega^2 = CA^4\Big(\frac{1}{4}\frac{1}{F(2,2)} - \frac{1}{2}\frac{1}{F(0,2)}\Big)$$

for some constant $C$. We now check that $\left(\frac{1}{4}\frac{1}{F(2,2)} - \frac{1}{2}\frac{1}{F(0,2)}\right) \neq 0$. We compute

$$F(0,2) - 2F(2,2) = -12 - 8\mu \neq 0,$$

hence $\omega^2 \neq 0$.

As a consequence, one has

$$\omega_\varepsilon^{[\leqq N]} = \omega^0 + \varepsilon^2\omega^2 + h.o.t.,$$

and furthermore, $\omega^2 \neq 0$. Since $\omega^0$ does not depend on $A$, we have that the twist condition writes as

$$\varepsilon^2\Big(\frac{\mathrm{d}\omega^2}{\mathrm{d}A}\Big) + h.o.t..$$

Hence, taking $\tilde{N}$ sufficiently large, we can apply Theorem 4.1 to obtain Theorem 3.6.

*The case* $\ell \geqq 2$ In this section, we consider the case $\ell > 1$, dealing with two or more frequencies. We note that in contrast with the case $\ell = 1$, one has to deal here with small divisors and we need the full force of Theorem 3.5. We fix $\ell \geqq 2$, which corresponds to fixing the range of $\mu$. For simplicity, we will impose the condition

$$\omega^0(\mu)\cdot k \neq 0 \quad k \in \mathbb{Z}^\ell \setminus \{0\}\,, \quad |k| \leqq 4, \tag{107}$$

where we emphasize that $\omega^0$ depends on $\mu$. This condition is not truly necessary but will simplify the calculations. We note that since $\omega^0(\mu)\cdot k$ for a fixed $k$ is analytic in $\mu$ and nontrivial we have that the previous non-resonance condition only fails at most for a finite set of $\mu$. Since the number of checks is small, it is not difficult to compute these values.

The idea of the argument presented here is to study the recurrence relation (104) and show that there are very few resonant terms, that is very few terms that contribute to $\omega^n$, which makes $\omega^n$ easy to analyze. Even if (107) fails it only means that verifying the twist requires analyzing more complicated expressions.

Taking into account the symmetries, we consider

$$\mathcal{U}_1 = \sum_{j=0}^{\ell} A_j \cos(2\pi\omega_j^0(\mu)\theta_j)\cdot\cos(2\pi jx). \tag{108}$$

Notice again that the amplitudes $A_j$ are parameters of our choice. We can think of the Lindstedt series as providing a family of approximate solutions parameterized by the amplitudes $A$.

The recurrence relation for the Lindstedt series writes as

$$\begin{aligned} \mathcal{M}_0 \mathcal{U}_n &= \Pi_{Range(\mathcal{M}_0)} \Big( \textstyle\sum_{j=1}^{n} \mathcal{U}_{n-j} \mathcal{U}_j \Big)_{xx} \\ \omega^{n-1} \partial_\theta (\omega^0 \cdot \partial_\theta) \mathcal{U}_1 &= \Pi_{\mathrm{Ker}\,(\mathcal{M}_0)} \Big( \textstyle\sum_{j=1}^{n} \mathcal{U}_{n-j} \mathcal{U}_j \Big)_{xx}. \end{aligned} \tag{109}$$

It is easy to show by induction that $\mathcal{U}_m$ are of the form

$$\mathcal{U}_m = \sum_{(k,j)\in\tau_m} C^m_{k,j} \cos(2\pi k \cdot \theta) \cos(2\pi j x),$$

where $\tau_m \subset \mathbb{Z}^\ell \times \mathbb{N}$ is a finite set. We refer to $\tau_m$ as the support of $\mathcal{U}^m$ since it represents the indices for which the coefficient is not zero. We observe that the support of a sum of terms is contained in the union. For $\mathcal{U}_1$, we just take

$$\tau_1 = \{((0, ..., 1, ..., 0), j), \, j = 1, ..., \ell\},$$

where the 1 is at the $j^{th}$ position. Proceeding by induction we observe that

$$\begin{aligned} \mathcal{U}_n \mathcal{U}_m &= \sum_{(k,j)\in\tau_m, (\alpha,\beta)\in\tau_n} C^m_{k,j} C^n_{\alpha,\beta} \cos(2\pi k\cdot\theta)\cos(2\pi\alpha\cdot\theta)\cos(2\pi j x)\cos(2\pi\beta x) \\ &= \sum_{(k,j)\in\tau_m, (\alpha,\beta)\in\tau_n} \frac{C^m_{k,j} C^n_{\alpha,\beta}}{4} (\cos(2\pi(k+\alpha)\cdot\theta)\cos(2\pi(k-\alpha)\cdot\theta)) \\ &\quad \cdot (\cos(2\pi(j+\beta)x) + \cos(2\pi(j-\beta)x)). \end{aligned}$$

Hence, for each of the products, we obtain 4 terms corresponding to $k \pm \alpha$ and $j \pm \beta$. Since taking two derivatives in $x$ does not change the support of a function, this motivates the following definition for composition of supports:

$$\begin{aligned} \tau \pm \tilde{\tau} = \{&(k+\alpha, j+\beta), (k+\alpha, j-\beta), (k-\alpha, j+\beta), (k-\alpha, j-\beta) \\ &| \, (k, j) \in \tau, (\alpha, \beta) \in \tilde{\tau}\}. \end{aligned}$$

We therefore obtain from the recursion relation that the support of $\sum_{j=1}^{\ell} \mathcal{U}_{m-j}\mathcal{U}_j$ is contained in

$$\bigcup_{j=1}^{\ell} \tau_{m-j} \pm \tau_j. \tag{110}$$

It is easy to see by induction that $(k, j) \in \tau_m$ implies that $|k| \leqq m$ and $j < m$ where $|k| = |k_1| + ... + |k_\ell|$. The goal of our analysis will be then to show that the supports grow and to demonstrate what resonant terms may appear in (110) that will lead to an $\omega^{n-1}$. The non-resonant terms in (110) will lead to terms in $\mathcal{U}_m$. Hence the support of $\mathcal{U}_m$ will be obtained by removing from (110) the resonant terms. The resonance assumption (107) tells us that the only resonant terms correspond to $k = 0$. We have

**Proposition 10.6.** *Under the non-resonance assumptions* (98) *or* (107) *we have that:*

- *There are no resonant terms at order* 2.

- *The only resonant terms correspond to order* 3 *and are of the form*

$$\Gamma_{j,j'}(\mu)A_jA_{j'}^2\cos(2\pi\omega_jt)\cos(2\pi jx).$$

**Proof.** We see that expanding to order 2, we obtain only indices of the form

$$(\omega_j \pm_1 \omega_{j'}, j \pm_2 j'),$$

where the choices of signs in $\pm_1$, $\pm_2$ are independent of each other. We will obtain a resonance only in case that $\omega_j \pm_1 \omega_{j'} = \omega_{j\pm_2 j'}$, or equivalently, that $\omega_j \pm_1 \omega_{j'} - \omega_{j\pm_2 j'} = 0$.

We note that because the $j$'s are not zero, $\omega_{j\pm_2 j'}$ cannot be equal to at least one of the $j, j'$. By the non-resonance condition (107), we have that this cannot happen. Hence, there are no resonant terms to order 2.

Computing to order 3, we obtain that the terms that appear to order 3 have the indices

$$(\omega_j \pm_1 \omega_{j'} \pm_2 \omega_{j''}, j \pm_3 j' \pm_4 j''),$$

and resonances can only happen when

$$\omega_j \pm_1 \omega_{j'} \pm_2 \omega_{j''} - \omega_{j\pm_3 j' \pm_4 j''} = 0.$$

If $j \pm_3 j' \pm_4 j''$ is different from the other $j, j', j''$, this cannot happen because of (107).

If $j \pm_3 j' \pm_4 j'' = j$, we obtain that $\pm_4 = -$ and $j' = j''$. Analyzing this case further, we see that if $\pm_2 = +$, we obtain $\omega_j \pm_1 2\omega_{j'} - \omega_j = 0$, which is a contradiction with (107). If $\pm_2 = -$, we obtain a case that can indeed happen. By permuting $j, j', j''$, we can always reduce to this case when one of the frequencies is the same.

Therefore, we see that

$$\omega_j^3 = A_j \sum_{j'} \Gamma_{j,j'}(\mu) A_{j'}^2$$

as claimed, where the $\Gamma_{j,j'}$ are algebraic functions of the frequencies obtained by performing algebraic operations on the frequencies. □

Now we can complete the proof of Theorem 3.6. The $\Gamma_{j,j'}$ are analytic functions of $\mu$ and are clearly non-trivial (they are easily computable in the limit $\mu \to \infty$).

Therefore, we see that

$$\frac{\partial \omega_j^{[\leq 3]}}{\partial A_{j'}} = \sum_k \Gamma_{j,k}(\mu)A_k^2 + A_j 2\Gamma_{j,j'}(\mu)A_{j'}$$

and therefore the determinant of the above matrix will be a polynomial on the $A_j$ variables. This polynomial will be homogeneous of degree $2\ell$. The coefficients of this polynomial are analytic functions of $\mu$. Excluding at most a finite number of values of $\mu$, we obtain that the polynomial giving the determinant is not identically zero. Hence, it vanishes only on a set of measure zero of the variables.

If we consider now scaled versions of the variables $A_j = \varepsilon B_j$ for a fixed value of the $B$, we obtain that the determinant is bigger than $C(B)\varepsilon^{2\ell}$.

If we fix $\mu$ and the value of $B$ and carry the Lindstedt series to order $N$, we obtain that the error in the invariance equation is $O(\varepsilon^{N+1})$ uniformly in $B$, the hyperbolicity constants are of order 1 uniformly in $B$ and the twist conditions are $O(\varepsilon^{2\ell}C(B)) + O(\varepsilon^{2\ell+1})$ where $C(B)$ is an analytic function that does not vanish except in an algebraic set of codimension 1.

Applying Theorem 3.5, we can obtain a true solution for all the sets for which $C(B) \geq O(\varepsilon^{N-4\ell-2})$. This corresponds to a set of $B$ whose measure grows to full as $\varepsilon$ decreases.

## 11. Application to the Boussinesq System

In this section, we consider the Boussinesq system of water waves. This system is even more interesting than the Boussinesq equation (see Section 10) since first the system is more "singular" and, therefore, the full power of the *two spaces* approach has to be used, that is one has to take the spaces $X$ and $Y$ such that $X \neq Y$. In this section, we will prove Theorem 3.7, which only considers the existence of periodic solutions. We note that the proof of of the set up is very similar to that of the proof of Theorem 3.6 and indeed similar to that in [26] and it works for any number of frequencies. In this section, we will verify the existence of approximate solutions only in the case of one frequency. We certainly expect that the construction of approximate solutions can be done for any number of frequencies, and thus improve Theorem 3.7, but we will not do it here. As in Remark 3.13, we note that the periodic solutions can be constructed using reduction to the center and Lyapunov orbits or by elementary contraction methods, but the expansions produced in this section give quantitative information.

The system is written as

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & -\partial_x - \mu\partial_{xxx} \\ -\partial_x & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \partial_x(uv) \\ 0 \end{pmatrix}, \tag{111}$$

where $t > 0$ and $x \in \mathbb{T}$.

The elementary linear analysis around the $(0, 0)$ equilibrium can be found in [26]. Recall that the eigenvalues of the linearization around 0 are given by

$$\omega(k) = \pm|k|2\pi i\sqrt{1 - 4\pi^2\mu k^2} \quad k \in \mathbb{Z}. \tag{112}$$

The eigenvectors are given by

$$U_j = (2\pi j \cos(2\pi\theta_j)\cos(2\pi j), \sqrt{(2\pi j)^2 - \mu(2\pi jx)^4}\sin(2\pi\theta_j)\sin(2\pi jx))$$

for $j = 1, \ldots\ell$, where $\ell$ is the smallest integer such that $1 - 4\pi^2\mu k \geqq 0$.

We denote by $\omega^0$ the vector whose components are all the real frequencies that appear as

$$\begin{aligned} \omega^0 &= (\omega(k_1), \omega(k_2), \ldots, \omega(k_\ell)); \\ \{k_1, \ldots k_\ell\} &= \{k \in \mathbb{Z} \,|\, k > 0;\, 1 - 4\pi^2\mu k^2 \geq 0\}. \end{aligned} \tag{113}$$

The following symmetries are, by this equation, formally preserved:

$$\begin{cases} u(t,-x) = u(-t,x) = u(t,x), \\ v(t,-x) = v(-t,x) = -v(t,x). \end{cases} \tag{114}$$

We remind the reader that we take

$$X = H^{\rho,m}(\mathbb{T}) \times H^{\rho,m+1}(\mathbb{T})$$

and

$$Y = H^{\rho,m-1}(\mathbb{T}) \times H^{\rho,m}(\mathbb{T}).$$

We denote by $X_0$ the set of functions in $X$ satisfying the symmetries (114), and also the momentum

$$\int_0^1 u(t,x)\,\mathrm{d}x = 0$$

and

$$\int_0^1 v(t,x)\,\mathrm{d}x = 0.$$

The previous quantities, as in the case of the Boussinesq equation, are preserved by the equation under consideration. The following proposition is proved in [26]:

**Proposition 11.1.** *The nonlinearity $\mathcal{N}(u,v) = (\partial_x(uv), 0)$ is analytic (indeed a polynomial) from $X$ to $Y$.*

Furthermore, one has

**Lemma 11.2.** *For $t > 0$, one has that*

$$\|U_\theta^s(t)\|_{Y,X} \leqq \frac{C}{t^{1/2}} e^{-Dt},$$

*and for $t < 0$ one has that*

$$\|U_\theta^u(t)\|_{Y,X} \leqq \frac{C'}{|t|^{1/2}} e^{D't}$$

*for some $C, C', D, D' > 0$.*

**Proof.** The proof of Lemma 11.2 is rather straightforward. It suffices to observe that the evolution operator is diagonal in a Fourier series. Then, express $\|U_\theta^s(t)\|_X^2$ using the Fourier terms and estimate the resulting sum. Full details for the same spaces are done in [26]. □

**11.0.5. Approximate Solution** We will not repeat the whole discussion which is very close to the one on the Boussinesq equation. Instead, we provide the necessary changes. The strategy is completely parallel to the one for the Boussinesq equation. Define two hull functions

$$u_\varepsilon(t, x) = \mathcal{U}_\varepsilon(\omega_\varepsilon t, x)$$

and

$$v_\varepsilon(t, x) = \mathcal{V}_\varepsilon(\omega_\varepsilon t, x).$$

Once again we consider Lindstedt series in powers of $\varepsilon$.

In a fashion similar to the previous section, we have

**Lemma 11.3.** *Let $\ell$ be as before. For all $N > 1$, there exists $(\omega^1, ..., \omega^N) \in (\mathbb{R}^\ell)^N$, $(\mathcal{U}_1, ..., \mathcal{U}_N) \in (H^{\rho,m}(\mathbb{T}))^N$ and $(\mathcal{V}_1, ..., \mathcal{V}_N) \in (H^{\rho,m-1}(\mathbb{T}))^N$ for some $\rho > 0$ such that*

$$\left\| \partial_t \begin{pmatrix} u_\varepsilon \\ v_\varepsilon \end{pmatrix} - \begin{pmatrix} 0 & -\partial_x - \mu\partial_{xxx} \\ -\partial_x & 0 \end{pmatrix} \begin{pmatrix} u_\varepsilon \\ v_\varepsilon \end{pmatrix} + \begin{pmatrix} \partial_x(u_\varepsilon v_\varepsilon) \\ 0 \end{pmatrix} \right\|_{H^{\rho,m}(\mathbb{T})\times H^{\rho,m-1}(\mathbb{T})} \leqq C\varepsilon^{N+1} \tag{115}$$

*for some constant $C > 0$ and*

$$u_\varepsilon^{[\leqq N]}(t, x) = \sum_{k=1}^{N} \varepsilon^k \mathcal{U}_k(\omega_\varepsilon^{[\leqq N]} t, x),$$

$$v_\varepsilon^{[\leqq N]}(t, x) = \sum_{k=1}^{N} \varepsilon^k \mathcal{V}_k(\omega_\varepsilon^{[\leqq N]} t, x),$$

*where*

$$\omega_\varepsilon^{[\leqq N]} = \omega^0 + \sum_{k=1}^{N} \varepsilon^k \omega^k.$$

*The solutions depend on $\ell$ arbitrary parameters, where $\ell$ is the number of the degrees of freedom of the kernel.*

**Proof.** We develop a general theory, parallel with the one of the Boussinesq equation in the previous section. The main new difficulties are that we are dealing with systems of equations and that the linear operator is not diagonal in an obvious sense. Denote

$$\mathcal{A} = \begin{pmatrix} 0 & -\partial_x - \mu\partial_x^3 \\ -\partial_x & 0 \end{pmatrix}.$$

At general order $m \geqq 2$, we search for solutions of the form

$$\mathcal{U}_m(\theta, x) = \sum_{j\mathbb{Z}, k\mathbb{Z}^\ell} U_{k,j}^m \cos(2\pi k \cdot \theta) \cos(2\pi j x)$$

and

$$\mathcal{V}_m(\theta, x) = \sum_{j\mathbb{Z}, k\mathbb{Z}^\ell} V^m_{k,j} \sin(2\pi k \cdot \theta) \sin(2\pi j x).$$

The previous formulae come from the assumptions of symmetry of the solutions. Denoting $\mathcal{W}_m = (\mathcal{U}_m, \mathcal{V}_m)$, one has

$$\Big(\omega^0 \cdot \partial_\theta - \mathcal{A}\Big)\mathcal{W}_m + \omega^{m-1} \cdot \partial_\theta \mathcal{W}_1 = \mathcal{R}_m(\omega^0, ..., \omega^{m-2}, \mathcal{W}_{m-1}).$$

It is important to notice that the operator $\mathcal{M}_0 = \Big(\omega^0 \cdot \partial_\theta - \mathcal{A}\Big)$ is not self-adjoint in $X$ and does not act as a multiplication in an easy basis of vectors. We then need to understand the range of this operator, its domain is spanned by

$$\Big((\cos(2\pi k \cdot \theta) \cos(2\pi j x), \sin(2\pi k \cdot \theta) \sin(2\pi j x))\Big).$$

The range is then the space of vector functions of the form of linear combinations of the basis

$$\Big((\sin(2\pi k \cdot \theta) \cos(2\pi j x), \cos(2\pi k \cdot \theta) \sin(2\pi j x))\Big).$$

**Order 1** One has

$$\omega^0 \cdot \partial_\theta \begin{pmatrix} \mathcal{U}_1 \\ \mathcal{V}_1 \end{pmatrix} = \begin{pmatrix} -\partial_x \mathcal{V}_1 - \mu \partial_x^3 \mathcal{V}_1 \\ -\partial_x \mathcal{U}_1 \end{pmatrix}. \tag{116}$$

We expand

$$\mathcal{U}_1 = \sum_{j=1}^{\ell} A^1_j \cos(2\pi\theta_j) \cos(2\pi j x)$$

$$\mathcal{V}_1 = \sum_{j=1}^{\ell} B^1_j \sin(2\pi\theta_j) \sin(2\pi j x).$$

As in the case of the Boussinesq equation, this gives, directly, the vector $\omega^0$, and one can take any $A^1_j$, $B^1_j$. For later convenience, we assume

$$A^1_j \neq 0, B^1_j \neq 0, \quad j = 1, \dots, \ell.$$

The rest of the orders are as in the previous section on the Boussinesq equation. □

We now prove Theorem 3.7, that is considering the case $\ell = 1$. This amounts to applying the abstract theorem 3.5. As in Section 10, this is done by checking the twist condition, the rest of the proof being completely parallel. We have first that

$$\mathcal{W}_1 = A \begin{pmatrix} \cos(2\pi\theta) \cos(2\pi x) 2\pi \\ \sin(2\pi\theta) \sin(2\pi x) 2\pi \omega^0 \end{pmatrix}.$$

For simplicity of writing we suppress the harmless parameter $A$.

At order 2, one has

$$\omega^0 \cdot \partial_\theta \begin{pmatrix} \mathcal{U}_2 \\ \mathcal{V}_2 \end{pmatrix} + \omega^1 \cdot \partial_\theta \begin{pmatrix} \mathcal{U}_1 \\ \mathcal{V}_1 \end{pmatrix} = \begin{pmatrix} -\partial_x \mathcal{V}_2 - \mu \partial_x^3 \mathcal{V}_2 + \partial_x(\mathcal{U}_1 \mathcal{V}_1) \\ -\partial_x \mathcal{U}_2 \end{pmatrix}. \tag{117}$$

Furthermore, one has (the map $F(j, k) = F_\mu(k, j)$ is defined as in the previous section)

$$\partial_x(\mathcal{U}_1 \mathcal{V}_1) = \frac{1}{2} \sin(4\pi\theta) \sin(4\pi x).$$

This is never in the range of $\mathcal{M}_0 = \omega^0 \cdot \partial_\theta - \mathcal{A}$. Therefore, we obtain $\omega^1 = 0$. Additionally, one has

$$\begin{aligned} \mathcal{W}_2 = &\frac{1}{F(2,2)} \begin{pmatrix} \frac{1}{2}\cos(4\pi\theta)\cos(4\pi x) \\ \frac{1}{2}\sin(4\pi\theta)\sin(4\pi x)\omega^0 \end{pmatrix} \\ &+ \frac{1}{F(-2,2)} \begin{pmatrix} \frac{1}{2}\cos(4\pi\theta)\cos(4\pi x) \\ -\frac{1}{2}\sin(4\pi\theta)\sin(4\pi x)\omega^0 \end{pmatrix}. \end{aligned}$$

We go now to order 3. We have

$$\mathcal{M}_0 \begin{pmatrix} \mathcal{U}_3 \\ \mathcal{V}_3 \end{pmatrix} + \omega^2 \cdot \partial_\theta \begin{pmatrix} \mathcal{U}_1 \\ \mathcal{V}_1 \end{pmatrix} = \begin{pmatrix} \partial_x(\mathcal{U}_1 \mathcal{V}_2) + \partial_x(\mathcal{U}_2 \mathcal{V}_1) \\ 0 \end{pmatrix}. \tag{118}$$

We have, by lengthy but straightforward computations,

$$\begin{aligned} \mathcal{U}_1 \mathcal{V}_2 = &\frac{1}{8} \frac{1}{F(-2,2)} \Big( \sin(6\pi\theta) - \sin(2\pi\theta) \Big) \Big( \sin(6\pi x) - \sin(2\pi x) \Big) \\ &- \frac{\omega^0}{8F(2,2)} \Big( \sin(6\pi\theta) - \sin(2\pi\theta) \Big) \Big( \sin(6\pi x) - \sin(2\pi x) \Big). \end{aligned}$$

Similarly,

$$\begin{aligned} \mathcal{U}_2 \mathcal{V}_1 = &\frac{1}{8} \frac{1}{F(2,2)} \Big( \sin(6\pi\theta) - \sin(2\pi\theta) \Big) \Big( \sin(6\pi x) - \sin(2\pi x) \Big) \\ &+ \frac{\omega^0}{8F(-2,2)} \Big( \sin(6\pi\theta) - \sin(2\pi\theta) \Big) \Big( \sin(6\pi x) - \sin(2\pi x) \Big). \end{aligned}$$

Hence one has

$$\begin{aligned} &\partial_x(\mathcal{U}_1 \mathcal{V}_2) + \partial_x(\mathcal{U}_2 \mathcal{V}_1) \\ &\quad = \frac{1}{8} \Big( \frac{1}{F(-2,2)} - \frac{\omega^2}{F(2,2)} + \frac{1}{F(-2,2)} - \frac{1}{F(2,2)} \Big) \\ &\quad \Big( 2\pi \sin(2\pi\theta) \cos(2\pi x) \Big) + R(\theta, x), \end{aligned} \tag{119}$$

where $R(\theta, x)$ is a trigonometric polynomial involving higher order frequencies. Since the coefficient

$$\frac{\pi}{4} \Big( \frac{1}{F(-2,2)} - \frac{\omega^0}{F(2,2)} + \frac{1}{F(-2,2)} - \frac{1}{F(2,2)} \Big)$$

is non-zero only on a finite number of values of $\mu$, one deduces that $\omega^2$ is nonzero, hence the twist condition. The rest of the proof follows.

*Acknowledgements.* We thank Prof. Šverák for a very careful and constructive work as an editor.

## References


1. AMBROSE, D.M.: Strong solutions for time-dependent mean field games with non-separable Hamiltonians. *J. Math. Pures Appl. (9)*, **113**, 141–154 2018
2. ABRAHAM, R., ROBBIN, J.: *Transversal Mappings and Flows*. W. A. Benjamin, Inc., New York-Amsterdam, 1967
3. BAMBUSI, D.: On the Darboux theorem for weak symplectic manifolds. *Proc. Amer. Math. Soc.*, **127**(11), 3383–3391 1999
4. BERTI, M.: *Nonlinear Oscillations of Hamiltonian PDEs*. Progress in Nonlinear Differential Equations and their Applications, 74. Birkhäuser Boston, Inc., Boston, MA, 2007
5. BOUSSINESQ, J.: Thorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond. *Jour. de Math. Pures et App. (2)*, **17**, 55–108 1872
6. BOURGAIN, J.: Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. of Math. (2)*, **148**(2), 363–439 1998
7. BOURGAIN, J.: Periodic solutions of nonlinear wave equations. In: *Harmonic Analysis and Partial Differential Equations (Chicago, IL, 1996)*, pages 69–97. Univ. Chicago Press, Chicago, IL, 1999
8. CALLEJA, R., CELLETTI, A., DE LA LLAVE, R.: Whitney regularity, monogenic dependence of solutions in kam theory: an a-posteriori approach. Manuscript 2015
9. CALLEJA, R., DE LA LLAVE, R.: A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification. *Nonlinearity*, **23**(9), 2029–2058 2010
10. CHENG, H., DE LA LLAVE, R.: Stable manifolds to bounded solutions in possibly ill-posed pde's. Preprint 2017
11. CONSTANTIN, P., FOIAS, C., NICOLAENKO, B., TEMAM, R.: *Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations*, volume 70 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1989
12. CASTELLI, R., GAMEIRO, M., LESSARD, J.-P.: Rigorous numerics for ill-posed pdes: Periodic orbits in the boussinesq equation. *Archive for Rational Mechanics and Analysis*, 2017
13. CRAIG, W., GUYENNE, P., NICHOLLS, D. P., SULEM, C.: Hamiltonian long-wave expansions for water waves over a rough bottom. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **461**(2055), 839–873 2005
14. CHUESHOV, I.D.: *Introduction to the Theory of Infinite-Dimensional Dissipative Systems*. AKTA Scientific Publishing House, Kharkiv, 2002
15. CHOW, S.-N., LEIVA, H.: Existence and roughness of the exponential dichotomy for skew-product semiflow in Banach spaces. *J. Differential Equations*, **120**(2), 429–477 1995
16. CHOW, S.-N., LEIVA, H.: Unbounded perturbation of the exponential dichotomy for evolution equations. *J. Differential Equations*, **129**(2), 509–531 1996
17. CHERNOFF, P. R., MARSDEN, J. E.: *Properties of Infinite Dimensional Hamiltonian Systems*. Springer-Verlag, Berlin, 1974. Lecture Notes in Mathematics, Vol. 425
18. CHERRIER, P., MILANI, A.: *Linear and Quasi-linear Evolution Equations in Hilbert Spaces*, volume 135 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012
19. CHEN, M., NGUYEN, N. V., SUN, S.-M.: Existence of traveling-wave solutions to Boussinesq systems. *Differential Integral Equations*, **24**(9-10), 895–908 2011

20. Craig, W.: *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, volume 9 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000
21. Craig, W.: Transformation theory of Hamiltonian PDE and the problem of water waves. In: *Hamiltonian Dynamical Systems and Applications*, NATO Sci. Peace Secur. Ser. B Phys. Biophys., pages 67–83. Springer, Dordrecht, 2008
22. Chepyzhov, V. V., Vishik, M. I.: *Attractors for Equations of Mathematical Physics*, volume 49 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2002
23. Craig, W., Wayne, C. E.: Newton's method and periodic solutions of nonlinear wave equations. *Comm. Pure Appl. Math.*, **46**(11), 1409–1498 1993
24. Craig, W., Wayne, C. E.: Periodic solutions of nonlinear Schrödinger equations and the Nash-Moser method. In: *Hamiltonian Mechanics (Toruń, 1993)*, volume 331 of *NATO Adv. Sci. Inst. Ser. B Phys.*, pages 103–122. Plenum, New York, 1994
25. de la Llave, R.: A tutorial on KAM theory. In: *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001
26. de la Llave, R.: A smooth center manifold theorem which applies to some ill-posed partial differential equations with unbounded nonlinearities. *J. Dynam. Differential Equations*, **21**(3), 371–415 2009
27. de la Llave, R., González, A., Jorba, À., Villanueva, J.: KAM theory without action-angle variables. *Nonlinearity*, **18**(2), 855–895 2005
28. Desolneux-Moulis, N.: Orbites périodiques des systèmes hamiltoniens autonomes (d'après Clarke, Ekeland-Lasry, Moser, Rabinowitz, Weinstein). In: *Bourbaki Seminar, Vol. 1979/80*, volume 842 of *Lecture Notes in Math.*, pages 156–173. Springer, Berlin-New York, 1981
29. Douady, R.: Stabilité ou instabilité des points fixes elliptiques. *Ann. Sci. École Norm. Sup. (4)*, **21**(1), 1–46 1988
30. Eliasson, L.H.: Hamiltonian systems with linear normal form near an invariant torus. In: *Nonlinear Dynamics (Bologna, 1988)*, pages 11–29. World Sci. Publishing, Teaneck, NJ, 1989
31. Fontich, E., de la Llave, R., Sire, Y.: Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions. *J. Differential Equations*, **246**(8), 3136–3213 2009
32. Fontich, E., de la Llave, R., Sire, Y.: A method for the study of whiskered quasi-periodic and almost-periodic solutions in finite and infinite dimensional Hamiltonian systems. *Electron. Res. Announc. Math. Sci.*, **16**, 9–22 2009
33. Fontich, E., de la Llave, R., Sire, Y.: Construction of invariant whiskered tori by a parameterization method. Part II: Quasi-periodic and almost periodic breathers in coupled map lattices. *J. Differential Equations*, **259**(6), 2180–2279 2015
34. Figueras, J., Gameiro, M., Lessard, J.-P., de la Llave, R.: A framework for the numerical computation and a posteriori verification of invariant objects of evolution equations. *SIAM J. Appl. Dyn. Syst.*, **16**(2), 1070–1088 2017
35. Grébert, B., Kappeler, T.: *The Defocusing NLS Equation and Its Normal Form*. EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, 2014
36. Goldstein, J. A.: *Semigroups of Linear Operators and Applications*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1985
37. Graff, S. M.: On the conservation of hyperbolic invariant tori for Hamiltonian systems. *J. Differential Equations*, **15**, 1–69 1974
38. Hale, J. K.: *Asymptotic Behavior of Dissipative Systems*, volume 25 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1988
39. Haro, A., de la Llave, R.: A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: rigorous results. *J. Differential Equations*, **228**(2), 530–579 2006

40. HENRY, D.: *Geometric Theory of Semilinear Parabolic Equations*, volume 840 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1981
41. HALE, J. K., MAGALHÃES, L. T., OLIVA, W. M.: *Dynamics in Infinite Dimensions*, volume 47 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 2002. With an appendix by Krzysztof P. Rybakowski
42. HILLE, E., PHILLIPS, R. S.: *Functional Analysis and Semi-groups*. American Mathematical Society, Providence, R. I., 1974. Third printing of the revised edition of 1957, American Mathematical Society Colloquium Publications, Vol. XXXI
43. HIRSCH, M.W., PUGH, C.C., SHUB, M.: *Invariant Manifolds*. Springer-Verlag, Berlin, 1977. Lecture Notes in Mathematics, Vol. 583
44. KAPPELER, T., PÖSCHEL, J.: *KdV & KAM*, volume 45 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 2003
45. KIRCHGÄSSNER, K., SCHEURLE, J.: On the bounded solutions of a semilinear elliptic equation in a strip. *J. Differential Equations*, **32**(1), 119–148 1979
46. KUKSIN, S. B.: *Nearly Integrable Infinite-Dimensional Hamiltonian Systems*, volume 1556 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1993
47. KUKSIN, S.B.: KAM-Theory for Partial Differential Equations. In *First European Congress of Mathematics, Vol. II (Paris, 1992)*, pages 123–157. Birkhäuser, Basel, 1994
48. KUKSIN, S. B.: *Analysis of Hamiltonian PDEs*, volume 19 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2000
49. KUKSIN, S. B.: Hamiltonian PDEs. In: *Handbook of Dynamical Systems. Vol. 1B*, pages 1087–1133. Elsevier B. V., Amsterdam, 2006. With an appendix by Dario Bambusi
50. LEWICKA, M., MUCHA, P. B.: On the existence of traveling waves in the 3D Boussinesq system. *Comm. Math. Phys.*, **292**(2), 417–429 2009
51. LUQUE, A., VILLANUEVA, J.: A KAM theorem without action-angle variables for elliptic lower dimensional tori. *Nonlinearity*, **24**(4), 1033–1080 2011
52. LI, Y., YI, Y.: Persistence of lower dimensional tori of general types in Hamiltonian systems. *Trans. Amer. Math. Soc.*, **357**(4), 1565–1600 (electronic), 2005
53. MCKEAN, H.P.: Boussinesq's equation on the circle. *Comm. Pure Appl. Math.*, **34**(5), 599–691 1981
54. MIELKE, A.: Reduction of quasilinear elliptic equations in cylindrical domains with applications. *Math. Methods Appl. Sci.*, **10**(1), 51–66 1988
55. MIYADERA, I.: *Nonlinear Semigroups*, volume 109 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1992. Translated from the 1977 Japanese original by Choong Yun Cho
56. MOSER, J.: A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)*, **20**, 499–535 1966
57. MOSER, J.: A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, **20**, 265–315 1966
58. MOSER, J. K.: *Lectures on Hamiltonian systems*. Mem. Amer. Math. Soc. No. 81. Amer. Math. Soc., Providence, R.I., 1968
59. MOSER, J.: Periodic orbits near an equilibrium and a theorem by Alan Weinstein. *Comm. Pure Appl. Math.*, **29**(6), 724–747 1976
60. NELSON, E.: *Topics in Dynamics. I: Flows*. Princeton University Press, Princeton, N.J., 1969
61. NEUBERGER, J.W.: *Sobolev gradients and differential equations*, volume 1670 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2010
62. PAZY, A.: *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983
63. POINCARÉ, H.: *Les méthodes nouvelles de la mécanique céleste*, volume 1, 2, 3. Gauthier-Villars, Paris, 1892–1899

64. PLISS, V. A., SELL, G. R.: Robustness of exponential dichotomies in infinite-dimensional dynamical systems. *J. Dynam. Differential Equations*, **11**(3), 471–513 1999
65. POLÁCIK, P., VALDEBENITO, D. A.: Existence of quasiperiodic solutions of elliptic equations on $\mathbb{R}^{N+1}$ via center manifold and KAM theorems. *J. Differential Equations*, **262**(12), 6109–6164 2017
66. ROBINSON, J. C.: *Infinite-Dimensional Dynamical Systems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2001. An introduction to dissipative parabolic PDEs and the theory of global attractors
67. RÜSSMANN, H.: On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In: *Dynamical Systems, Theory and Applications (Battelle Rencontres, Seattle, Wash., 1974)*, pages 598–624. Lecture Notes in Phys., Vol. 38, Berlin, 1975. Springer
68. RÜSSMANN, H.: On a new proof of Moser's twist mapping theorem. In: *Proceedings of the Fifth Conference on Mathematical Methods in Celestial Mechanics (Oberwolfach, 1975), Part I*, volume 14, pages 19–31, 1976
69. RÜSSMANN, H.: On optimal estimates for the solutions of linear difference equations on the circle. *Celestial Mech.*, **14**(1), 33–37 1976
70. SCHWARTZ, J.: On Nash's implicit functional theorem. *Comm. Pure Appl. Math.*, **13**, 509–530 1960
71. SEVRYUK, M. B.: Partial preservation of frequencies in KAM theory. *Nonlinearity*, **19**(5), 1099–1140 2006
72. SHOWALTER, R.E.: *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997
73. SACKER, R.J., SELL, G.R.: Existence of dichotomies and invariant splittings for linear differential systems. III. *J. Differential Equations*, **22**(2), 497–522 1976
74. SELL, G. R., YOU, Y.: *Dynamics of evolutionary equations*, volume 143 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2002
75. TEMAM, R.: *Infinite-dimensional dynamical systems in mechanics and physics*, volume 68 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1997
76. VANO, J.: A Nash–Moser implicit function theorem with Whitney regularity and applications. *Ph.D. Thesis, University of Texas*, 2002
77. WEINSTEIN, A.: Normal modes for nonlinear Hamiltonian systems. *Invent. Math.*, **20**, 47–57 1973
78. YOU, J.: Perturbations of lower-dimensional tori for Hamiltonian systems. *J. Differential Equations*, **152**(1), 1–29 1999
79. ZEHNDER, E.: Homoclinic points near elliptic fixed points. *Comm. Pure Appl. Math.*, **26**, 131–182 1973
80. ZEHNDER, E.: Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, **28**, 91–140 1975
81. ZEHNDER, E.: Generalized implicit function theorems with applications to some small divisor problems. II. *Comm. Pure Appl. Math.*, **29**(1), 49–111 1976

RAFAEL DE LA LLAVE
School of Mathematics,
Georgia Institute of Technology,
Atlanta,
GA 30332,
USA.
e-mail: rafael.delallave@math.gatech.edu

and

YANNICK SIRE
Johns Hopkins University,
Krieger Hall,
Baltimore,
USA.
e-mail: sire@math.jhu.edu

(*Received April 26, 2016 / Accepted August 2, 2018*)
*Published online August 14, 2018*