SISSA - ISAS
International School for Advanced Studies

Area of Mathematics

KAM for quasi-linear and fully nonlinear perturbations of Airy and KdV equations

Ph.D. Thesis

Supervisor:
Prof. Massimiliano Berti

Candidate:
Riccardo Montalto

Academic Year 2013/2014

Il presente lavoro costituisce la tesi presentata da Riccardo Montalto, sotto la direzione del Prof. Massimiliano Berti, al fine di ottenere l'attestato di ricerca post-universitaria Doctor Philosophiae presso la SISSA, Curriculum in Analisi Matematica, Area di Matematica. Ai sensi dell'art. 1, comma 4, dello Statuto della Sissa pubblicato sulla G.U. no. 36 del 13.02.2012, il predetto attestato é equipollente al titolo di Dottore di Ricerca in Matematica.

Trieste, Anno Accademico 2013-2014.

Acknowledgements

I am extremely grateful to my supervisor, Massimiliano Berti, for having introduced me to the world of research and for his continuous encouragements and precious suggestions.

I wish to express my deep gratitude to Pietro Baldi for his constant support, for having spent a lot of time in working togheter and for having taught me many mathematical tricks.

I warmly thank Thomas Alazard, for having dedicated me a lot of time during my visiting in Paris.

I wish to thank also Luca Biasco, Livia Corsi, Roberto Feola and Michela Procesi for many useful discussions.

Contents

Chapter 1

Introduction

In the last years much progresses have been achieved in the theory of quasi-periodic motions for infinite dimensional Hamiltonian and reversible dynamical systems, that we shall call, in a broad sense, KAM (Kolmogorov-Arnold-Moser) theory for PDEs (including also the Newton-Nash-Moser implicit function theorem approach).

A challenging and open question concerns its possible extension to *quasi-linear* (also called "*strongly nonlinear*" in [56]) and *fully nonlinear* PDEs, namely equations whose nonlinearities contain derivatives of the same order as the linear operator. Besides its mathematical interest, this question is also relevant in view of applications to physical real world nonlinear models, for example in fluid dynamics, water waves and elasticity. This Thesis is a first step in this direction. In particular we develop KAM theory for both quasi-linear and fully nonlinear forced perturbations of the Airy equation

$$u_t + u_{xxx} + \varepsilon f(\omega t, x, u, u_x, u_{xx}, u_{xxx}) = 0\,, \tag{1.0.1}$$

and for quasi-linear Hamiltonian autonomous perturbations of KdV

$$u_t + u_{xxx} - 6uu_x + \mathcal{N}_4(x, u, u_x, u_{xx}, u_{xxx}) = 0\,, \tag{1.0.2}$$

with periodic boundary conditions $x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$.

The main results of this Thesis prove the existence of *Cantor families* of small amplitude, linearly stable, quasi-periodic solutions for both the equations (1.0.1), (1.0.2), under suitable assumptions on the nonlinearities f (see for instance (1.2.6), (1.2.7), the reversibility condition (1.2.15)) and $\mathcal{N}_4$ in (1.3.2). We recall that a quasi-periodic solution is a function

$$u(\omega t, x)\,, \qquad u : \mathbb{T}^\nu \times \mathbb{T} \to \mathbb{R}\,,$$

where the frequency vector $\omega := (\omega_1, \omega_2, \dots, \omega_\nu)$ is rationally independent, namely $\omega \cdot l \neq 0$, for all $l \in \mathbb{Z}^\nu \setminus \{0\}$.

Note that the equation (1.0.1) depends on the frequency vector ω. It will be used as an external parameter in order to impose the non-resonance conditions which naturally appear in KAM theory. On the other hand, (1.0.2) is an autonomous PDE with no external parameters, hence the frequency of the expected quasi-periodic solution is a-priori unknown. A careful bifurcation analysis has to be performed, in order to determine how the frequencies depend on the amplitudes of the quasi-periodic solutions. This is one of the reasons why we have decided to study the forced equation (1.0.1) before the autonomous PDE (1.0.2).

Before describing the main KAM results concerning the equations (1.0.1), (1.0.2), we outline a short history of the KAM and Nash-Moser theory for PDEs, focusing in particular on the results which deal with unbounded perturbations.

1.1 Historical preface

The "KAM for PDEs" theory is a generalization of the original KAM theory for quasi integrable finite dimensional Hamiltonian systems, developed in the sixties by Kolmogorov [52] and Arnold [2] for analytic Hamiltonian systems and then extended by Moser for only differentiable perturbations and for reversible dynamical systems, see [61], [63]. These results prove that, for Hamiltonian systems which are small perturbations of an integrable one, under suitable non-degeneracy conditions on the Hamiltonian, the quasi-periodic orbits form a full-measure set of the phase space. Such quasi-periodic orbits are constructed by means of an iterative scheme. The main difficulty of this procedure is due to the well-known *small-divisors*, namely the numbers $\omega \cdot l$, $l \in \mathbb{Z}^\nu$ (ω is the frequency of oscillation of the solution). The small divisors enter at the denominator of the Fourier coefficients in the Fourier expansion of the approximate solutions defined at each step of the KAM iteration. They can become arbitrarily small, affecting the convergence of the iterative scheme, since, for almost every ω, the set

$$\big\{\omega \cdot l : l \in \mathbb{Z}^\nu\big\}$$

accumulates to 0. This difficulty is overcome by imposing non-resonance *diophantine* conditions of the form

$$|\omega \cdot l| \geq \frac{\gamma}{|l|^\tau}\,, \qquad \forall l \neq 0\,, \qquad \gamma \in (0,1)\,,$$

which are sufficient to prove the convergence of the scheme. Such non-resonance conditions are called *zero-th order Melnikov* conditions.

Later on, KAM theory has been extended by Moser [63], Eliasson [36] and Pöschel [64] for elliptic invariant tori of lower dimension. In these problems, also *first and second order Melnikov* non resonance conditions (see (2.1.14)-(2.1.16)) are required along the KAM iterative scheme.

In the ninetiees, it started the investigation concerning the existence of periodic and quasi-periodic solutions for PDEs. In order to overcome the small divisors difficulty, the two main approaches which have been developed are:

- **normal form KAM methods,**
- **Newton-Nash-Moser implicit function iterative schemes.**

The normal form KAM procedure consists in an iterative super-quadratic scheme, which, by means of infinitely many canonical transformations, brings the Hamiltonian associated to the PDE into another one which has an invariant torus at the origin. This method is an infinite dimensional extension of the KAM theory for lower dimensional elliptic tori in Eliasson [36] and Pöschel [64]. The classical KAM procedure for 1-dimensional PDEs with bounded nonlinear perturbations will be described in Section 2.1.

The small divisors arise at each step of the iteration in solving the so-called *homological equations* (see (2.1.7)). In the usual KAM framework, such equations are constant coefficients linear PDE, which can be solved by imposing the Melnikov non resonance conditions on the frequencies (see (2.1.13)-(2.1.16)). As a consequence of having solved constant coefficients homological equations at each step of the iteration, also the linearized equation at the final KAM torus has constant coefficients (reducible torus). In many cases its Lyapunov exponents are purely imaginary and therefore the KAM quasi-periodic solutions are linearly stable, see the end of Section 2.1.

The KAM theory for PDEs has been developed for the first time in the pioneering works of Kuksin [53] and Wayne [73] for bounded perturbations of parameter dependent 1-dimensional linear Schrödinger and wave equations with Dirichlet boundary conditions. Then Kuksin-Pöschel [57] and Pöschel [65] extended these results for parameter independent nonlinear Schrödinger (NLS) and nonlinear wave (NLW) equations like

$$\mathrm{i}u_t - u_{xx} + mu + f(|u|^2)u = 0\,, \quad m > 0\,, \quad (\mathrm{NLS})\,, \tag{1.1.1}$$

$$u_{tt} - u_{xx} + mu + au^3 + f(u)\,, \quad m > 0\,, \quad a \neq 0\,, \quad f(u) = O(u^5)\,, \quad (\mathrm{NLW})\,, \tag{1.1.2}$$

where f is an analytic nonlinearity (for more references see also the monograph [56]).

Now we describe the Newton-Nash-Moser implicit function theorem approach for Hamiltonian PDEs. In this method, the search of periodic and quasi-periodic solutions is reduced to find zeros of a nonlinear operator by means of an iterative quadratic Newton-type scheme in scales of Banach spaces of analytic or differentiable functions. The typical framework is the following: one has to solve a nonlinear functional equation of the form

$$F(u) = 0\,, \tag{1.1.3}$$

where F is a nonlinear operator acting on a scale of Banach spaces. The approximate solutions are defined iteratively as

$$u_0 := 0\,, \qquad u_{n+1} := u_n + h_{n+1}\,, \qquad h_{n+1} := -S_n F'(u_n)^{-1} F(u_n)\,,$$

where S_n is a suitable smoothing operator which regularizes the approximate solutions at each step (this is strictly required only to deal with spaces of differentiable functions).

The main advantage of the Nash-Moser method is to require only the *first order Melnikov* non-resonance conditions to invert the linearized operator $L_n := F'(u_n)$ at each step of the iteration. These conditions are essentially the minimal assumptions. On the other hand, the main difficulty is that the linear operator L_n is an operator with variable coefficients; it is represented by a matrix which is a small perturbation of a diagonal matrix with arbitrarily small eigenvalues and therefore it is hard to estimate its inverse in high Sobolev norm.

The Newton-Nash-Moser approach has been proposed by Craig-Wayne in [34] (see also the monograph [31]) to prove the existence of periodic solutions of 1-dimensional nonlinear Klein-Gordon and Schrödinger equations with periodic boundary conditions. Later on, still for periodic solutions, it has been extended by Berti-Bolle in [17], [18] for completely resonant nonlinear wave equations both with analytic and differentiable nonlinearities, see also Gentile-Mastropietro-Procesi [41] and the monograph [13].

For quasi-periodic solutions, the Nash-Moser techniques have been considerably extended by Bourgain in [25], [27], [29] for analytic NLS and NLW with convolution potential on $\mathbb{T}^d$. We underline that this approach is especially convenient for PDEs in higher space dimension, because the second order Melnikov conditions (required in the KAM scheme) are violated due to the high multiplicity of the eigenvalues. The techniques of Bourgain have been recently extended by Wang [72] for completely resonant NLS on $\mathbb{T}^d$, and by Berti-Bolle [21], [20] for forced NLS and NLW with a multiplicative potential on $\mathbb{T}^d$ and differentiable nonlinearities (see [19], [42] for previous results about periodic solutions). We mention also the recent paper of Berti-Corsi-Procesi [24] which contains an abstract Nash-Moser implicit function theorem with applications to NLW and NLS on compact Lie groups.

As a consequence of having imposed only the first order Melnikov conditions, this method does not provide information about the linear stability of the quasi-periodic solutions, because the linearized equations have variable coefficients.

Via KAM methods, existence and stability of quasi-periodic solutions with periodic boundary conditions have been proved by Chierchia-You in [30] for 1-dimensional NLW equations. For what concerns PDEs in higher space dimension, the first KAM results have been obtained by Eliasson-Kuksin [39] for NLS with convolution potential on $\mathbb{T}^d$. The second order Melnikov conditions are verified by introducing the notion of "Töplitz-Lipschitz" Hamiltonians. KAM results for completely resonant NLS in any space dimension have been then obtained by Procesi-Procesi [68], see also Geng-You-Xu [40] for $d = 2$.

All the results quoted above, concern PDEs in which the nonlinearity is a bounded nonlinear differential operator of order 0. Now we start to describe KAM and Nash-Moser results for PDEs with unbounded nonlinearities. In this case the usual way to construct a bounded transformation of the phase space at each KAM-step fails. If such a transformation were unbounded (as the perturbation) then along the iteration the order of unboundedness of the transformed vector fields would increase quadratically and the scheme would not converge.

The first KAM results for *unbounded* perturbations have been obtained by Kuksin [55], [56] and then, Kappeler-Pöschel [49] for Hamiltonian, analytic perturbations of KdV

$$u_t + u_{xxx} - 6uu_x + \varepsilon\partial_x f(x,u) = 0\,, \tag{1.1.4}$$

with periodic boundary conditions $x \in \mathbb{T}$. Note that the constant coefficients linear operator is ∂_{xxx} and the nonlinearity contains one space derivative ∂_x. These results, which we describe in Section 2.2, prove the continuation of Cantor families of *finite gap* solutions of KdV. The main issue is that the Hamiltonian vector field generated by the perturbation is unbounded of order 1. In order to overcome this difficulty, the key idea, introduced by Kuksin in [55], is to work with a variable-coefficients normal form (see (2.2.11)). The frequencies of KdV grow as $\sim j^3$, hence for $j \neq k$ (outside the diagonal) the difference $|j^3 - k^3| \geq (j^2 + k^2)/2$, so that KdV gains two derivatives. This smoothing effect of the small divisors is sufficient to produce a bounded transformation of the phase space at each step of the KAM iteration, since the perturbation is of order 1. On the other hand for $j = k$ there is no smoothing effect and therefore such diagonal terms cannot be removed in the homological equations. These angle-dependent terms will be inserted into the normal form. As a consequence, the homological equations have variable coefficients and they can be solved thanks to the "Kuksin's lemma" (see Lemma 2.2.1). Note that such homological equations are scalar and so they are much easier than the variable coefficients functional equations which appear in the Newton-Nash-Moser approach of Craig-Wayne-Bourgain.

The proof given in [55] and [49] works also for Hamiltonian analytic pseudo-differential perturbations of order 2 (in space), like

$$u_t + u_{xxx} - 6uu_x + \varepsilon\partial_x|\partial_x|^{\frac{1}{2}} f(x, |\partial_x|^{\frac{1}{2}} u) = 0 \,, \quad x \in \mathbb{T} \,,$$

using an improved version of the Kuksin's lemma proved by Liu-Yuan in [60] (see also [59]). Then in [60] (see also Zhang-Gao-Yuan [75]) Liu-Yuan applied it to 1-dimensional derivative NLS (DNLS) and Benjamin-Ono equations, where the highest order constant coefficients linear operator is ∂_{xx} and the nonlinearity contains one derivative ∂_x. These methods apply to dispersive PDEs with derivatives like KdV, DNLS, the Duffing oscillator (see Bambusi-Graffi [10]), but not to derivative wave equations (DNLW) which contain first order derivatives ∂_x, ∂_t in the nonlinearity.

For DNLW Bourgain [28] proved the existence of periodic solutions of

$$u_{tt} - u_{xx} + mu + u_t^2 = 0 \,, \quad m > 0 \,, \quad x \in \mathbb{T} \,,$$

extending the Craig-Wayne approach in [34].

KAM theory for DNLW has been recently developed by Berti-Biasco-Procesi in [15] for the Hamiltonian equation

$$u_{tt} - u_{xx} + mu + g(Du) = 0 \,, \quad D := \sqrt{-\partial_{xx} + m} \,, \quad x \in \mathbb{T} \,, \tag{1.1.5}$$

and in [16] for the reversible equation

$$u_{tt} - u_{xx} + mu + g(x, u, u_t, u_x) = 0 \,, \quad x \in \mathbb{T} \,, \tag{1.1.6}$$

assuming the conditions

$$g(x, u, u_x, u_t) = g(x, u, u_x, -u_t) \,, \quad g(x, u, u_x, u_t) = g(-x, u, -u_x, u_t) \,,$$

where in both the equations (1.1.5), (1.1.6), g is analytic. The key ingredient is an asymptotic expansion of the perturbed eigenvalues that is sufficiently accurate to impose the second order Melnikov non-resonance conditions. This is achieved by introducing the notion of "quasi-Töplitz" vector field developed by Procesi-Xu in [69] (see also Grébert-Thomann [43], Procesi-Procesi [68]).

All the aforementioned results concern "semilinear" PDEs, namely equations in which the nonlinearity contains *strictly less* derivatives than the constant coefficients linear differential operator. For quasi-linear or fully nonlinear PDEs (called "*strongly non linear*" in [56]) the perturbative effect is much stronger, and the possibility of extending KAM theory in this context is doubtful, see [49], [31], [60], because of the possible phenomenon of formation of singularities outlined in Lax [58], Klainerman and Majda [51]. For example, Kappeler-Pöschel [49] (remark 3, page 19) wrote:

"It would be interesting to obtain perturbation results which also include terms of higher order, at least in the region where the KdV approximation is valid. However, results of this type are still out of reach, if true at all".

The study of this important issue is at its first steps.

For quasi-linear and fully nonlinear PDEs, the literature concerns, before the results [6], [8], [9] presented in this Thesis, only existence of *periodic* solutions.

We quote the classical bifurcation results of Rabinowitz [70] for fully nonlinear forced wave equations with a small dissipation term

$$u_{tt} - u_{xx} + \alpha u_t + \varepsilon f(t, x, u, u_t, u_x, u_{tt}, u_{tx}, u_{xx}) = 0\,, \quad x \in \mathbb{T}\,.$$

Recently, Baldi, in [3], proved existence of periodic forced vibrations for quasi-linear Kirchhoff equations

$$u_{tt} - \Big(1 + \int_\Omega |\nabla u|^2\, dx\Big)\Delta u = \varepsilon f(\omega t, x)\,, \quad x \in \Omega$$

with Dirichlet boundary conditions $u|_{\partial\Omega} = 0$ and also for periodic boundary conditions $\Omega = \mathbb{T}^d$. Here the quasi-linear perturbation term depends explicitly only on time. Both these results are proved via Nash-Moser methods.

For the water waves equations (see Section 7), which are fully nonlinear PDEs, we mention the pioneering work of Iooss-Plotnikov-Toland [45] about existence of time periodic standing waves, and of Iooss-Plotnikov [46], [47] for 3-dimensional traveling water waves. The key idea is to use diffeomorphisms of the torus $\mathbb{T}^2$ and pseudo-differential operators, in order to conjugate the linearized operator to one with constant coefficients plus a sufficiently smoothing remainder. This is enough to invert the whole linearized operator by Neumann series. Very recently Baldi [4] has further developed the techniques of [45], proving the existence of periodic solutions for fully nonlinear autonomous, reversible Benjamin-Ono equations

$$u_t + \mathcal{H}u_{xx} + \partial_x(u^3) + f(x, u, \mathcal{H}u, u_x, \mathcal{H}u_{xx}) = 0\,, \quad x \in \mathbb{T}\,, \quad \mathcal{H}(e^{\mathrm{i}jx}) = -\mathrm{i}\,\mathrm{sign}(j) e^{\mathrm{i}jx}\,, \quad j \in \mathbb{Z}$$

where $\mathcal{H}$ is the Hilbert transform.

We mention also the recent paper of Alazard and Baldi [1] concerning the existence of periodic standing solutions of the water waves equations with surface tension.

These methods do not work for proving the existence of quasi-periodic solutions and they do not imply either the linear stability of the solutions.

In the remaining part of this introduction, we shall present in detail the results proved in this Thesis about the existence and stability of quasi-periodic solutions of the equations (1.0.1), (1.0.2) and the main ideas of the proofs, see Sections 1.2.1, 1.3.1. To the best of our knowledge, these are the first KAM results for quasi-linear or fully nonlinear PDEs.

1.2 Main results for forced Airy equation

We now present the results announced in [5] and proved in [6]. The details of the proofs will be given in Chapter 4.

We consider quasi-linear or fully nonlinear perturbations of Airy equation, namely

$$u_t + u_{xxx} + \varepsilon f(\omega t, x, u, u_x, u_{xx}, u_{xxx}) = 0\,, \quad x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}\,, \tag{1.2.1}$$

where $\varepsilon > 0$ is a small parameter, the nonlinearity is quasi-periodic in time with diophantine frequency vector

$$\omega = \lambda\bar{\omega} \in \mathbb{R}^\nu\,, \quad \lambda \in \Lambda := \Big[\frac{1}{2}, \frac{3}{2}\Big], \quad |\bar{\omega}\cdot l| \geq \frac{3\gamma_0}{|l|^{\tau_0}} \quad \forall l \in \mathbb{Z}^\nu \setminus \{0\}, \tag{1.2.2}$$

and $f(\varphi, x, z)$, $\varphi \in \mathbb{T}^\nu$, $z := (z_0, z_1, z_2, z_3) \in \mathbb{R}^4$, is a finitely many times differentiable function, namely

$$f \in C^q(\mathbb{T}^\nu \times \mathbb{T} \times \mathbb{R}^4; \mathbb{R}) \tag{1.2.3}$$

for some $q \in \mathbb{N}$ large enough. For simplicity we fix in (1.2.2) the diophantine exponent $\tau_0 := \nu$. The only "external" parameter in (1.2.1) is λ, which is the length of the frequency vector (this corresponds to a time scaling). We consider the following questions:

- *For ε small enough, do there exist quasi-periodic solutions of* (1.2.1) *for positive measure sets of $\lambda \in \Lambda$?*
- *Are these solutions linearly stable?*

Clearly, if $f(\varphi, x, 0)$ is not identically zero, then $u = 0$ is not a solution of (1.2.1) for $\varepsilon \neq 0$. Thus we look for non-trivial $(2\pi)^{\nu+1}$-periodic solutions $u(\varphi, x)$ of the Airy equation

$$\omega \cdot \partial_\varphi u + u_{xxx} + \varepsilon f(\varphi, x, u, u_x, u_{xx}, u_{xxx}) = 0 \tag{1.2.4}$$

in the Sobolev space

$$\begin{aligned} H^s &:= H^s(\mathbb{T}^\nu \times \mathbb{T}; \mathbb{R}) \\ &:= \Big\{ u(\varphi, x) = \sum_{(l,j)\in\mathbb{Z}^\nu\times\mathbb{Z}} u_{l,j}\, e^{\mathrm{i}(l\cdot\varphi + jx)} \in \mathbb{R}, \ \ \bar{u}_{l,j} = u_{-l,-j}\,, \ \ \|u\|_s^2 := \sum_{(l,j)\in\mathbb{Z}^\nu\times\mathbb{Z}} \langle l, j\rangle^{2s} |u_{l,j}|^2 < \infty \Big\} \end{aligned} \tag{1.2.5}$$

where

$$\langle l, j\rangle := \max\{1, |l|, |j|\}.$$

From now on, we fix $s_0 := (\nu + 2)/2 > (\nu + 1)/2$, so that for all $s \geq s_0$ the Sobolev space H^s is a Banach algebra, and it is continuously embedded $H^s(\mathbb{T}^{\nu+1}) \hookrightarrow C(\mathbb{T}^{\nu+1})$.

We need some assumptions on the perturbation $f(\varphi, x, u, u_x, u_{xx}, u_{xxx})$. We suppose that

- Type (F). The *fully nonlinear* perturbation has the form

$$f(\varphi, x, u, u_x, u_{xxx}), \tag{1.2.6}$$

namely it is independent of u_{xx} (note that the dependence on u_{xxx} may be nonlinear). Otherwise, we require that

- Type (Q). The perturbation is *quasi-linear*, namely

$$f = f_0(\varphi, x, u, u_x, u_{xx}) + f_1(\varphi, x, u, u_x, u_{xx})u_{xxx}$$

is affine in u_{xxx}, and it satisfies (naming the variables $z_0 = u$, $z_1 = u_x$, $z_2 = u_{xx}$, $z_3 = u_{xxx}$)

$$\partial_{z_2} f = \alpha(\varphi)\big(\partial^2_{z_3 x} f + z_1 \partial^2_{z_3 z_0} f + z_2 \partial^2_{z_3 z_1} f + z_3 \partial^2_{z_3 z_2} f\big) \tag{1.2.7}$$

for some function $\alpha(\varphi)$ (independent on x).

The Hamiltonian nonlinearities in (1.2.11) satisfy the above assumption (Q), see remark 4.1.2. In comment 3 after Theorem 1.2.5 we explain the reason for assuming either condition (F) or (Q).

The following theorem is an existence result of quasi-periodic solutions.

Theorem 1.2.1. (Existence) *There exist $s := s(\nu) > 0$, $q := q(\nu) \in \mathbb{N}$, such that: For every quasi-linear nonlinearity $f \in C^q$ of the form*

$$f = \partial_x\big(g(\omega t, x, u, u_x, u_{xx})\big) \tag{1.2.8}$$

satisfying the (Q)-condition (1.2.7)*, for all $\varepsilon \in (0, \varepsilon_0)$, where $\varepsilon_0 := \varepsilon_0(f, \nu)$ is small enough, there exists a Cantor set $\mathcal{C}_\varepsilon \subset \Lambda$ of asymptotically full Lebesgue measure, i.e.*

$$|\mathcal{C}_\varepsilon| \to 1 \quad as \quad \varepsilon \to 0, \tag{1.2.9}$$

such that $\forall \lambda \in \mathcal{C}_\varepsilon$ the perturbed equation (1.2.4) *has a solution $u(\varepsilon, \lambda) \in H^s$ with $\|u(\varepsilon, \lambda)\|_s \to 0$ as $\varepsilon \to 0$.*

We may ensure the *linear stability* of the solutions requiring further conditions on the nonlinearity, see Theorem 1.2.5 for the precise statement. The first case is that of *Hamiltonian* equations

$$\partial_t u = X_H(u)\,,$$

$$X_H(u) := \partial_x \nabla_{L^2} H(t, x, u, u_x)\,, \quad H(t, x, u, u_x) := \int_{\mathbb{T}} \frac{u_x^2}{2} + \varepsilon F(\omega t, x, u, u_x)\, dx \tag{1.2.10}$$

which have the form (1.2.1), (1.2.8) with

$$f(\varphi, x, u, u_x, u_{xx}, u_{xxx}) = -\partial_x\big\{(\partial_{z_0} F)(\varphi, x, u, u_x)\big\} + \partial_{xx}\big\{(\partial_{z_1} F)(\varphi, x, u, u_x)\big\}\,. \tag{1.2.11}$$

In this thesis, with a slight abuse of notation, $\int_{\mathbb{T}^d}$ is short for denoting the average $(2\pi)^{-d}\int_{\mathbb{T}^d}$. This notation will be used in all the definitions which will be given, i.e L^2-scalar product, symplectic form, definitions of Hamiltonians etc.

The phase space of (1.2.10) is

$$H^1_0(\mathbb{T}) := \Big\{u(x) \in H^1(\mathbb{T}, \mathbb{R}) \,:\, \int_{\mathbb{T}} u(x)\, dx = 0\Big\} \tag{1.2.12}$$

endowed with the non-degenerate two symplectic form

$$\Omega(u, v) := \int_{\mathbb{T}} (\partial_x^{-1} u)\, v\, dx\,, \quad \forall u, v \in H^1_0(\mathbb{T})\,, \tag{1.2.13}$$

where $\partial_x^{-1} u$ is the periodic primitive of u with zero average (see (4.1.19)).

Notice that the Hamiltonian vector field $X_H(u) := \partial_x \nabla H(u)$ is the unique vector field satisfying the equality

$$dH(u)[h] = (\nabla H(u), h)_{L^2(\mathbb{T})} = \Omega(X_H(u), h)\,, \quad \forall u, h \in H_0^1(\mathbb{T})\,,$$

where for all $u, v \in L^2(\mathbb{T}) := L^2(\mathbb{T}, \mathbb{R})$, we define

$$(u, v)_{L^2(\mathbb{T})} := \int_{\mathbb{T}} u(x)v(x)\,dx = \sum_{j\in\mathbb{Z}} u_j v_{-j}\,, \quad u(x) = \sum_{j\in\mathbb{Z}} u_j e^{\mathrm{i}jx}\,, \quad v(x) = \sum_{j\in\mathbb{Z}} v_j e^{\mathrm{i}jx}\,.$$

We recall also that the Poisson bracket between two Hamiltonians F, $G : H_0^1(\mathbb{T}) \to \mathbb{R}$ are

$$\{F(u), G(u)\} := \Omega(X_F, X_G) = \int_{\mathbb{T}} \nabla F(u)\partial_x \nabla G(u) dx\,. \tag{1.2.14}$$

As proved in remark 4.1.2, the Hamiltonian nonlinearity f in (1.2.11) satisfies also the (Q)-condition (1.2.7). As a consequence, Theorem 1.2.1 implies the existence of quasi-periodic solutions of (1.2.10). In addition, we also prove their linear stability.

Theorem 1.2.2. (Hamiltonian case) *For all Hamiltonian quasi-linear equations* (1.2.10) *the quasi-periodic solution $u(\varepsilon, \lambda)$ found in Theorem 1.2.1 is* LINEARLY STABLE *(see Theorem 1.2.5).*

The stability of the quasi-periodic solutions also follows by the *reversibility* condition

$$f(-\varphi, -x, z_0, -z_1, z_2, -z_3) = -f(\varphi, x, z_0, z_1, z_2, z_3). \tag{1.2.15}$$

Actually (1.2.15) implies that the infinite-dimensional non-autonomous dynamical system

$$u_t = V(t, u), \quad V(t, u) := -u_{xxx} - \varepsilon f(\omega t, x, u, u_x, u_{xx}, u_{xxx})$$

is reversible with respect to the involution

$$S : u(x) \to u(-x), \quad S^2 = I,$$

namely

$$-SV(-t, u) = V(t, Su)\,.$$

In this case it is natural to look for "reversible" solutions of (1.2.4), that is

$$u(\varphi, x) = u(-\varphi, -x)\,. \tag{1.2.16}$$

Theorem 1.2.3. (Reversible case) *There exist $s := s(\nu) > 0$, $q := q(\nu) \in \mathbb{N}$, such that:*
For every nonlinearity $f \in C^q$ that satisfies

(i) the reversibility condition (1.2.15)*,*

and

(ii) either the (F)-condition (1.2.6) *or the (Q)-condition* (1.2.7)*,*

for all $\varepsilon \in (0, \varepsilon_0)$, where $\varepsilon_0 := \varepsilon_0(f, \nu)$ is small enough, there exists a Cantor set $\mathcal{C}_\varepsilon \subset \Lambda$ with Lebesgue measure satisfying (1.2.9)*, such that for all $\lambda \in \mathcal{C}_\varepsilon$ the perturbed Airy equation* (1.2.4) *has a solution $u(\varepsilon, \lambda) \in H^s$ that satisfies* (1.2.16)*, with $\|u(\varepsilon, \lambda)\|_s \to 0$ as $\varepsilon \to 0$. In addition, $u(\varepsilon, \lambda)$ is* LINEARLY STABLE*.*

Let us make some comments on the results.

1. The quasi-periodic solutions of Theorem 1.2.1 could be unstable because the nonlinearity f has no special structure and some eigenvalues of the linearized operator at the solutions could have non zero real part (partially hyperbolic tori). In any case, we *reduce* to *constant* coefficients the linearized operator (Theorem 1.2.4) and we may compute its eigenvalues (i.e. Lyapunov exponents) with any order of accuracy. With further conditions on the nonlinearity—like reversibility or in the Hamiltonian case—the eigenvalues are purely imaginary, and the torus is linearly stable. The present situation is very different with respect to [34], [25]-[29], [21]-[20] and also [45]-[47], [4], where the lack of stability information is due to the fact that the linearized equation has variable coefficients.

2. One cannot expect the existence of quasi-periodic solutions of (1.2.4) for *any* perturbation f. Actually, if $f = m \neq 0$ is a constant, then, integrating (1.2.4) in (φ, x) we find the contradiction $\varepsilon m = 0$. This is a consequence of the fact that

$$\mathrm{Ker}(\omega \cdot \partial_\varphi + \partial_{xxx}) = \mathbb{R} \tag{1.2.17}$$

is non trivial. Both the condition (1.2.8) (which is satisfied by the Hamiltonian nonlinearities) and the reversibility condition (1.2.15) allow to overcome this obstruction, working in a space of functions with zero average. The degeneracy (1.2.17) also reflects in the fact that the solutions of (1.2.4) appear as a 1-dimensional family $c+u_c(\varepsilon, \lambda)$ parametrized by the "average" $c \in \mathbb{R}$. We could also avoid this degeneracy by adding a "mass" term $+mu$ in (1.2.1), but it does not seem to have physical meaning.

3. In Theorem 1.2.1 we have not considered the case in which f is fully nonlinear and satisfies condition (F) in (1.2.6), because any nonlinearity of the form (1.2.8) is automatically quasi-linear (and so the first condition in (1.2.7) holds) and (1.2.6) trivially implies the second condition in (1.2.7) with $\alpha(\varphi) = 0$.

4. The solutions $u \in H^s$ have the same regularity in both variables (φ, x). This functional setting is convenient when using changes of variables that mix the time and space variables, like the composition operators $\mathcal{A}$, $\mathcal{T}$ in Sections 4.1.1, 4.1.4,

5. In the Hamiltonian case (1.2.10), the nonlinearity f in (1.2.11) satisfies the reversibility condition (1.2.15) if and only if $F(-\varphi, -x, z_0, -z_1) = F(\varphi, x, z_0, z_1)$.

Theorems 1.2.1-1.2.3 are based on a Nash-Moser iterative scheme. An essential ingredient in the proof—which also implies the linear stability of the quasi-periodic solutions—is the *reducibility* of the linear operator

$$\mathcal{L} := \mathcal{L}(u) = \omega \cdot \partial_\varphi + (1 + a_3(\varphi, x))\partial_{xxx} + a_2(\varphi, x)\partial_{xx} + a_1(\varphi, x)\partial_x + a_0(\varphi, x) \tag{1.2.18}$$

obtained by linearizing (1.2.4) at any approximate (or exact) solution u, where the coefficients $a_i(\varphi, x)$ are defined in (4.1.2). Let $H^s_x := H^s(\mathbb{T})$ denote the usual Sobolev spaces of functions of $x \in \mathbb{T}$ only.

Theorem 1.2.4. (Reducibility) *There exist $\bar\sigma > 0$, $q \in \mathbb{N}$, depending on ν, such that:*
For every nonlinearity $f \in C^q$ that satisfies the hypotheses of Theorems 1.2.1 or 1.2.3, for all $\varepsilon \in (0, \varepsilon_0)$, where $\varepsilon_0 := \varepsilon_0(f, \nu)$ is small enough, for all u in the ball $\|u\|_{\mathfrak{s}_0+\bar\sigma} \leq 1$, there exists a Cantor like set $\Lambda_\infty(u) \subset \Lambda$ such that, for all $\lambda \in \Lambda_\infty(u)$:
i) for all $s \in (s_0, q - \bar\sigma)$, if $\|u\|_{s+\bar\sigma} < +\infty$ then there exist linear invertible bounded operators W_1, $W_2 : H^s(\mathbb{T}^{\nu+1}) \to H^s(\mathbb{T}^{\nu+1})$ (see (4.2.72)*) with bounded inverse, that semi-conjugate the linear operator $\mathcal{L}(u)$ in* (1.2.18) *to the diagonal operator $\mathcal{L}_\infty$, namely*

$$\mathcal{L}(u) = W_1 \mathcal{L}_\infty W_2^{-1}\,, \quad \mathcal{L}_\infty := \omega \cdot \partial_\varphi + \mathcal{D}_\infty \tag{1.2.19}$$

where

$$\mathcal{D}_\infty := \mathrm{diag}_{j\in\mathbb{Z}}\{\mu_j\}, \quad \mu_j := \mathrm{i}(-m_3 j^3 + m_1 j) + r_j\,, \quad m_3, m_1 \in \mathbb{R}\,, \quad \sup_j |r_j| \leq C\varepsilon\,. \tag{1.2.20}$$

ii) For each $\varphi \in \mathbb{T}^\nu$ the operators W_i are also bounded linear bijections of H^s_x (see notation (3.1.17)*)*

$$W_i(\varphi)\,, W_i^{-1}(\varphi) : H^s_x \to H^s_x\,, \quad i = 1, 2\,.$$

A curve $h(t) = h(t, \cdot) \in H^s_x$ is a solution of the quasi-periodically forced linear equation

$$\partial_t h + (1 + a_3(\omega t, x))\partial_{xxx} h + a_2(\omega t, x)\partial_{xx} h + a_1(\omega t, x)\partial_x h + a_0(\omega t, x) h = 0 \tag{1.2.21}$$

if and only if the transformed curve

$$v(t) := v(t, \cdot) := W_2^{-1}(\omega t)[h(t)] \in H^s_x$$

is a solution of the constant coefficients dynamical system

$$\partial_t v + \mathcal{D}_\infty v = 0\,, \quad \dot v_j = -\mu_j v_j\,, \quad \forall j \in \mathbb{Z}\,. \tag{1.2.22}$$

In the reversible or Hamiltonian case all the $\mu_j \in \mathrm{i}\mathbb{R}$ are purely imaginary.

The operator W_1 differs from W_2 (see (4.2.72)) only for the multiplication by the function ρ in (4.1.26) which comes from the re-parametrization of time of Section 4.1.2. As explained in Section 3.4 this does not affect the dynamical consequence of Theorem 1.2.4-*ii*).

The exponents μ_j can be effectively computed. All the solutions of (1.2.22) are

$$v(t) = \sum_{j\in\mathbb{Z}} v_j(t) e^{\mathrm{i}jx}\,, \quad v_j(t) = e^{-\mu_j t} v_j(0)\,.$$

If the μ_j are purely imaginary—as in the reversible or the Hamiltonian cases—all the solutions of (1.2.22) are almost periodic in time (in general) and the Sobolev norm

$$\|v(t)\|_{H^s_x} = \Big(\sum_{j\in\mathbb{Z}} |v_j(t)|^2 \langle j\rangle^{2s}\Big)^{1/2} = \Big(\sum_{j\in\mathbb{Z}} |v_j(0)|^2 \langle j\rangle^{2s}\Big)^{1/2} = \|v(0)\|_{H^s_x} \tag{1.2.23}$$

is constant in time. As a consequence we have:

Theorem 1.2.5. (Linear stability) *Assume the hypothesis of Theorem 1.2.4 and, in addition, that f is Hamiltonian (see* (1.2.11)*) or it satisfies the reversibility condition* (1.2.15)*. Then, $\forall s \in (\mathfrak{s}_0, q - \bar{\sigma} - s_0)$, $\|u\|_{s+s_0+\bar{\sigma}} < +\infty$, there exists $K_0 > 0$ such that for all $\lambda \in \Lambda_\infty(u)$, $\varepsilon \in (0, \varepsilon_0)$, all the solutions of* (1.2.21) *satisfy*

$$\|h(t)\|_{H^s_x} \le K_0 \|h(0)\|_{H^s_x} \tag{1.2.24}$$

and, for some $\mathtt{a} \in (0,1)$,

$$\|h(0)\|_{H^s_x} - \varepsilon^{\mathtt{a}} K_0 \|h(0)\|_{H^{s+1}_x} \le \|h(t)\|_{H^s_x} \le \|h(0)\|_{H^s_x} + \varepsilon^{\mathtt{a}} K_0 \|h(0)\|_{H^{s+1}_x} \,. \tag{1.2.25}$$

Theorems 1.2.1-1.2.5 are proved in Section 4.3.1 collecting all the informations of Sections 4.1-4.3.

1.2.1 Ideas of the proof

The proof of Theorems 1.2.1-1.2.3 is based on a Nash-Moser iterative scheme in the scale of Sobolev spaces H^s. The main issue concerns the invertibility of the linearized operator $\mathcal{L}$ in (1.2.18), at each step of the iteration, and the proof of the tame estimates (4.3.7) for its right inverse. This information is obtained in Theorem 4.2.3 by conjugating $\mathcal{L}$ to constant coefficients. This is also the key which implies the stability results for the Hamiltonian and reversible nonlinearities, see Theorems 1.2.4-1.2.5.

We now explain the main ideas of the reducibility scheme. The term of $\mathcal{L}$ that produces the strongest perturbative effect to the spectrum (and eigenfunctions) is $a_3(\varphi, x)\partial_{xxx}$, and, then $a_2(\varphi, x)\partial_{xx}$. The usual KAM transformations are not able to deal with these terms. The reason is the following: if in the Homological equation (4.2.41), the operator $\mathcal{R}$ were unbounded of order 3, the solution Ψ defined in (4.2.44), would be unbounded of order 1, thanks to the fact that the small divisors gain two space derivatives (see (4.2.17)). Hence the iterative scheme would not converge in any norm, therefore we adopt the following strategy. First, we conjugate the operator $\mathcal{L}$ in (1.2.18) to a constant coefficients third order differential operator plus a bounded remainder

$$\mathcal{L}_5 = \omega \cdot \partial_\varphi + m_3 \partial_{xxx} + m_1 \partial_x + \mathcal{R}_0, \quad m_3 = 1 + O(\varepsilon),\ m_1 = O(\varepsilon)\,,\ m_1, m_3 \in \mathbb{R}\,, \tag{1.2.26}$$

(see (4.1.56)), via changes of variables induced by diffeomorphisms of the torus, a reparametrization of time, and pseudo-differential operators. This is the goal of Section 4.1. All these transformations could be composed into one map, but we find more convenient to split the regularization procedure into separate steps (Sections 4.1.1-4.1.5), both to highlight the basic ideas, and, especially, in order to derive estimates on the coefficients in Section 4.1.6. Let us make some comments on this procedure.

1. In order to eliminate the space variable dependence of the highest order perturbation $a_3(\varphi, x)\partial_{xxx}$ (see (4.1.20)) we use, in Section 4.1.1, φ-dependent changes of variables of the form

   $$(\mathcal{A}h)(\varphi, x) := h(\varphi, x + \beta(\varphi, x))\,.$$

   These transformations converge pointwise to the identity if $\beta \to 0$ but not in operatorial norm. If β is odd, $\mathcal{A}$ preserves the reversible structure, see remark 4.1.4. On the other hand for the Hamiltonian equation (1.2.10) we use the modified transformation

$$(\mathcal{A}h)(\varphi,x) := (1+\beta_x(\varphi,x))\,h(\varphi,x+\beta(\varphi,x)) = \frac{d}{dx}\big\{(\partial_x{}^{-1}h)(\varphi,x+\beta(\varphi,x))\big\} \qquad (1.2.27)$$

for all $h(\varphi,\cdot) \in H^1_0(\mathbb{T})$. This map is canonical, for each $\varphi \in \mathbb{T}^\nu$, with respect to the KdV-symplectic form (1.2.13), see remark 4.1.3. Thus (1.2.27) preserves the Hamiltonian structure and also eliminates the term of order ∂_{xx}, see remark 4.1.5.

2. In the second step of Section 4.1.2 we eliminate the time dependence of the coefficients of the highest order spatial derivative operator ∂_{xxx} by a quasi-periodic time re-parametrization. This procedure preserves the reversible and the Hamiltonian structure, see remark 4.1.6 and 4.1.7.

3. Assumptions (Q) (see (1.2.7)) or (F) (see (1.2.6)) allow to eliminate terms like $a(\varphi,x)\partial_{xx}$ along this reduction procedure, see (4.1.41). This is possible, by a conjugation with multiplication operators (see (4.1.34)), if (see (4.1.40))

$$\int_{\mathbb{T}} \frac{a_2(\varphi,x)}{1+a_3(\varphi,x)}\,dx = 0\,. \qquad (1.2.28)$$

   If (F) holds, then the coefficient $a_2(\varphi,x)=0$ and (1.2.28) is satisfied. If (Q) holds, then an easy computation shows that $a_2(\varphi,x)=\alpha(\varphi)\,\partial_x a_3(\varphi,x)$ (using the explicit expression of the coefficients in (4.1.2)), and so

$$\int_{\mathbb{T}} \frac{a_2(\varphi,x)}{1+a_3(\varphi,x)}\,dx = \int_{\mathbb{T}} \alpha(\varphi)\,\partial_x\big(\log[1+a_3(\varphi,x)]\big)\,dx = 0\,.$$

   In both cases (Q) and (F), condition (1.2.28) is satisfied.

   In the Hamiltonian case there is no need of this step because the symplectic transformation (1.2.27) also eliminates the term of order ∂_{xx}, see remark 4.1.7.

   We note that without assumptions (Q) or (F) we may always reduce $\mathcal{L}$ to a time dependent operator with $a(\varphi)\partial_{xx}$. If $a(\varphi)$ were a constant, then this term would even simplify the analysis, killing the small divisors. The pathological situation that we avoid by assuming (Q) or (F) is when $a(\varphi)$ changes sign. In such a case, this term acts as a friction when $a(\varphi)<0$ and as an amplifier when $a(\varphi)>0$.

4. In Sections 4.1.4-4.1.5, we are finally able to conjugate the linear operator to another one with a coefficient in front of ∂_x which is constant, i.e. obtaining (1.2.26). In this step we use a transformation of the form $I + w(\varphi,x)\partial_x^{-1}$, see (4.1.49). In the Hamiltonian case we use the symplectic map $e^{\pi_0 w(\varphi,x)\partial_x^{-1}}$, see remark 4.1.13.

5. We can iterate the regularization procedure at any *finite* order $k=0,1,\ldots$, conjugating $\mathcal{L}$ to an operator of the form $\mathfrak{D}+\mathcal{R}$, where

$$\mathfrak{D} = \omega\cdot\partial_\varphi + \mathcal{D}, \quad \mathcal{D} = m_3\partial_x^3 + m_1\partial_x + \ldots + m_{-k}\partial_x^{-k}\,, \quad m_i \in \mathbb{R}\,,$$

has constant coefficients, and the remainder $\mathcal{R}$ is arbitrarily regularizing in space, namely

$$\partial_x^k \circ \mathcal{R} = \text{bounded}\,. \tag{1.2.29}$$

However, one cannot iterate this regularization infinitely many times, because it is not a quadratic scheme, and therefore, because of the small divisors, it does not converge. This regularization procedure is sufficient to prove the invertibility of $\mathcal{L}$, giving tame estimates for the inverse, in the periodic case, but it does not work for quasi-periodic solutions. The reason is the following. In order to use Neumann series, one needs that $\mathfrak{D}^{-1}\mathcal{R} = (\mathfrak{D}^{-1}\partial_x^{-k})(\partial_x^k\mathcal{R})$ is bounded, namely, in view of (1.2.29), that $\mathfrak{D}^{-1}\partial_x^{-k}$ is bounded. In the region where the eigenvalues $(\mathrm{i}\omega\cdot l + \mathcal{D}_j)$ of $\mathfrak{D}$ are small, space and time derivatives are related, $|\omega\cdot l| \sim |j|^3$, where l is the Fourier index of time, j is that of space, and $\mathcal{D}_j = -\mathrm{i}m_3 j^3 + \mathrm{i}m_1 j + \ldots$ are the eigenvalues of $\mathcal{D}$. Imposing the first order Melnikov conditions $|\mathrm{i}\omega\cdot l + \mathcal{D}_j| > \gamma|l|^{-\tau}$, in that region, $(\mathfrak{D}^{-1}\partial_x^{-k})$ has eigenvalues

$$\Big|\frac{1}{(\mathrm{i}\omega\cdot l + \mathcal{D}_j)j^k}\Big| < \frac{|l|^\tau}{\gamma|j|^k} < \frac{C|l|^\tau}{|\omega\cdot l|^{k/3}}\,.$$

In the periodic case, $\omega\in\mathbb{R}$, $l\in\mathbb{Z}$, $|\omega\cdot l| = |\omega||l|$, and this determines the order of regularization that is required by the procedure: $k \geq 3\tau$. In the quasi-periodic case, instead, $|l|$ is not controlled by $|\omega\cdot l|$, and the argument fails.

Once (1.2.26) has been obtained, we implement a quadratic reducibility KAM scheme *à la* Eliasson-Kuksin, in order to diagonalize $\mathcal{L}_5$, namely to conjugate $\mathcal{L}_5$ to the diagonal operator $\mathcal{L}_\infty$ in (1.2.19). Since we work with finite regularity, we perform a Nash-Moser smoothing regularization (time-Fourier truncation). In order to decrease quadratically the size of the perturbation $\mathcal{R}$, we use standard KAM transformations of the form

$$\Phi = I + \Psi\,, \qquad \text{or} \qquad \Phi = \exp(\Psi) \qquad \text{in Hamiltonian case}\,,$$

see Section 4.2.1. At each step of the iteration we have an operator

$$\mathcal{L} = \omega\cdot\partial_\varphi + \mathcal{D} + \mathcal{R}\,,$$

where $\mathcal{D}$ is a diagonal operator with eigenvalues μ_j, $j\in\mathbb{Z}$ and $\mathcal{R}$ is a bounded linear operator small in size. If the operator Ψ solves the homological equation

$$\omega\cdot\partial_\varphi\Psi + [\mathcal{D},\Psi] + \Pi_N\mathcal{R} = [\mathcal{R}]\,, \qquad [\mathcal{R}] := \mathrm{diag}_{j\in\mathbb{Z}}\mathcal{R}_j^j(0)\,, \tag{1.2.30}$$

where Π_N is the time Fourier truncation operator defined in (3.1.18), then

$$\mathcal{L}_+ := \Phi^{-1}\mathcal{L}\Phi = \omega\cdot\partial_\varphi + \mathcal{D}_+ + \mathcal{R}_+\,,$$

where

$$\mathcal{D}_+ := \mathcal{D} + [\mathcal{R}]\,, \qquad \mathcal{R}_+ := \Phi^{-1}\Big(\Pi_N^\perp\mathcal{R} + \mathcal{R}\Psi - \Psi[\mathcal{R}]\Big)\,.$$

The remainder $\mathcal{R}_+$ is the sum of a quadratic function of Ψ, $\mathcal{R}$ and a remainder supported on the high modes. This iterative scheme will converge (see Theorem 4.2.2), since the initial remainder

$\mathcal{R}_0$ in (1.2.26) is a bounded linear operator (of the space variable x) small in size and this property is preserved along the iteration passing from $\mathcal{R}$ to $\mathcal{R}_+$. This is the reason why we have performed the regularization procedure in Sections 4.1.1-4.1.5, before starting with the KAM reducibility scheme. The homological equation (1.2.30) may be solved by imposing the second order Melnikov nonresonance conditions

$$|\mathrm{i}\omega \cdot l + \mu_j(\lambda) - \mu_k(\lambda)| \geq \frac{\gamma |j^3 - k^3|}{\langle l \rangle^\tau}\,, \quad \forall l \in \mathbb{Z}^\nu\,,\ |l| \leq N\,, \quad j, k, \in \mathbb{Z}\,,$$

(see (4.2.17)). We are able to verify that for most parameters $\lambda \in [1/2, 3/2]$, such non resonance conditions are satisfied, thanks to the sharp asymptotic expansion of the eigenvalues

$$\mu_j := \mu_j(\lambda) = \mathrm{i}(-m_3(\lambda) j^3 + m_1(\lambda) j) + r_j(\lambda)\,, \qquad m_3(\lambda) - 1\,, m_1(\lambda) = O(\varepsilon)\,, \quad \sup_{j \in \mathbb{Z}} |r_j(\lambda)| = O(\varepsilon)\,.$$

Remark 1.2.1. *We underline that the goal of the Töplitz-Lipschitz [39], [40], [43] and quasi-Töplitz property [69], [15], [16], [68] is precisely to provide an asymptotic expansion of the perturbed eigenvalues sharp enough to verify the second order Melnikov conditions.*

Note that the above eigenvalues μ_j could not be purely imaginary, i.e. r_j could have a non-zero real part which depends on the nonlinearity (unlike the reversible or Hamiltonian case, where $r_j \in \mathrm{i}\mathbb{R}$). In such a case, the invariant torus could be (partially) hyperbolic. Since we do not control the real part of r_j (i.e. the hyperbolicity may vanish), we perform the measure estimates proving the diophantine lower bounds of the imaginary part of the small divisors.

The final comment concerns the dynamical consequences of Theorem 1.2.4-*ii*). All the above transformations (both the changes of variables of Sections 4.1.1-4.1.5 as well as the KAM matrices of the reducibility scheme) are time-dependent quasi-periodic maps of the phase space (of functions of x only), see Section 3.4. It is thanks to this "Töplitz-in-time" structure that the linear equation (1.2.21) is transformed into the dynamical system (1.2.22) as explained in Section 3.4. Note that in [45] (and also [29], [21],[20]) the analogous transformations have not this Töplitz-in-time structure and stability informations are not obtained.

1.3 Main results for autonomous KdV

In this section we present the result announced in [7] and proved in [8], concerning the existence and the stability of Cantor families of quasi-periodic solutions of Hamiltonian *quasi-linear* perturbations of the KdV equation

$$u_t + u_{xxx} - 6uu_x + \mathcal{N}_4(x, u, u_x, u_{xx}, u_{xxx}) = 0\,, \tag{1.3.1}$$

under periodic boundary conditions $x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$, where

$$\mathcal{N}_4(x, u, u_x, u_{xx}, u_{xxx}) := -\partial_x \big[(\partial_u f)(x, u, u_x) - \partial_x ((\partial_{u_x} f)(x, u, u_x)) \big] \tag{1.3.2}$$

is the most general quasi-linear Hamiltonian (local) nonlinearity. Note that $\mathcal{N}_4$ contains as many derivatives as the linear part ∂_{xxx}. The equation (1.3.1) is the Hamiltonian PDE $u_t = \partial_x \nabla H(u)$ where ∇H denotes the $L^2(\mathbb{T})$ gradient of the Hamiltonian

$$H(u) = \int_{\mathbb{T}} \frac{u_x^2}{2} + u^3 + f(x, u, u_x)\, dx \tag{1.3.3}$$

on the real phase space $H^1_0(\mathbb{T})$ defined in (1.2.12). We assume that the "Hamiltonian density" $f \in C^q(\mathbb{T} \times \mathbb{R} \times \mathbb{R}; \mathbb{R})$ for some q large enough, and that

$$f = f_5(u, u_x) + f_{\geq 6}(x, u, u_x)\,, \tag{1.3.4}$$

where $f_5(u, u_x)$ denotes the homogeneous component of f of degree 5 and $f_{\geq 6}$ collects all the higher order terms. By (1.3.4) the nonlinearity $\mathcal{N}_4$ vanishes of order 4 at $u = 0$ and (1.3.1) may be seen, close to the origin, as a "small" perturbation of the KdV equation

$$u_t + u_{xxx} - 6uu_x = 0\,. \tag{1.3.5}$$

The KdV equation is completely integrable in the strongest possible sense, namely it may be described by *global analytic action angle variables*. This has been proved in a series of works by Kappeler and collaborators, see [11], [12], [48] and also the monograph [49].

A natural question is to know whether the periodic, quasi-periodic or almost periodic solutions of (1.3.5) persist under small perturbations. This is the content of KAM theory.

We prove the existence of small amplitude, linearly stable, quasi-periodic solutions of (1.3.1), see Theorem 1.3.1. Note that (1.3.1) does not depend on external parameters. Moreover the KdV equation (1.3.1) is a *completely resonant* PDE, namely the linearized equation at the origin is the linear Airy equation $u_t + u_{xxx} = 0$, which possesses only the 2π-periodic in time solutions

$$u(t, x) = \sum\nolimits_{j \in \mathbb{Z} \setminus \{0\}} u_j \mathrm{e}^{\mathrm{i} j^3 t} e^{\mathrm{i} j x}\,. \tag{1.3.6}$$

Thus the existence of quasi-periodic solutions of (1.3.1) is a purely nonlinear phenomenon (the diophantine frequencies in (1.3.12) are $O(|\xi|)$-close to integers with $\xi \to 0$) and a perturbation theory is more difficult.

The solutions that we find are localized in Fourier space close to finitely many "tangential sites"

$$S^+ := \{\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu\}\,, \quad S := S^+ \cup (-S^+) = \{\pm j : j \in S^+\}\,, \quad \bar{\jmath}_i \in \mathbb{N} \setminus \{0\}\,, \quad \forall i = 1, \ldots, \nu\,. \tag{1.3.7}$$

The set S is required to be even because the solutions u of (1.3.1) have to be real valued. Moreover, we also assume the following explicit hypotheses on S:

- ($\mathtt{S1}$) $j_1 + j_2 + j_3 \neq 0$ for all $j_1, j_2, j_3 \in S$.
- ($\mathtt{S2}$) $\nexists j_1, \ldots, j_4 \in S$ such that $j_1 + j_2 + j_3 + j_4 \neq 0$, $j_1^3 + j_2^3 + j_3^3 + j_4^3 - (j_1 + j_2 + j_3 + j_4)^3 = 0$.

Theorem 1.3.1. *Given $\nu \in \mathbb{N}$, let $f \in C^q$ (with $q := q(\nu)$ large enough) satisfy* (1.3.4)*. Then, for all the tangential sites S as in* (1.3.7) *satisfying* ($\mathtt{S1}$)-($\mathtt{S2}$)*, the KdV equation* (1.3.1) *possesses small amplitude quasi-periodic solutions with diophantine frequency vector $\omega := \omega(\xi) = (\omega_j)_{j \in S^+} \in \mathbb{R}^\nu$, of the form*

$$u(t, x) = \sum\nolimits_{j \in S^+} 2\sqrt{\xi_j}\, \cos(\omega_j t + jx) + o(\sqrt{|\xi|}), \quad \omega_j := j^3 - 6\xi_j j^{-1}\,, \tag{1.3.8}$$

for a "Cantor-like" set of small amplitudes $\xi \in \mathbb{R}^\nu_+$ with density 1 at $\xi = 0$. The term $o(\sqrt{|\xi|})$ is small in some H^s-Sobolev norm, $s < q$. These quasi-periodic solutions are linearly stable.

This theorem is proved in Chapter 5. Let us make some comments.

1. **(Tangential sites)** The set of tangential sites S satisfying ($\mathtt{S1}$)-($\mathtt{S2}$) can be iteratively constructed in an explicit way, see the end of Section 5.7. After fixing $\{\bar{\jmath}_1, \ldots, \bar{\jmath}_n\}$, in the choice of $\bar{\jmath}_{n+1}$ there are only finitely many forbidden values, while all the other infinitely many values are good choices for $\bar{\jmath}_{n+1}$. In this precise sense the set S is "generic".

2. **(Stability)** The linear stability of the quasi-periodic solutions is discussed after (5.7.41). In a suitable set of symplectic coordinates (ψ, η, w), $\psi \in \mathbb{T}^\nu$, near the invariant torus, the linearized equations at the quasi-periodic solutions assume the form (5.7.41), (5.7.42). Actually there is a complete KAM normal form near the invariant torus (remark 5.4.1), see also [22].

A similar result holds for perturbed (focusing/defocusing) modified KdV equations (m-KdV)

$$u_t + u_{xxx} \pm \partial_x u^3 + \mathcal{N}_4(x, u, u_x, u_{xx}, u_{xxx}) = 0\,, \tag{1.3.9}$$

for Hamiltonian quasi-linear nonlinearity $\mathcal{N}_4$ as in (1.3.2).

Theorem 1.3.2. *Given $\nu \in \mathbb{N}$, let $f \in C^q$ (with $q := q(\nu)$ large enough) satisfy*

$$f(x, u, u_x) = O(|(u, u_x)|^5)\,. \tag{1.3.10}$$

Then, for all the tangential sites S in (1.3.7) *satisfying*

$$\frac{2}{2\nu - 1} \sum_{i=1}^{\nu} \bar{\jmath}_i^2 \notin \mathbb{Z} \tag{1.3.11}$$

the m-KdV equation (1.3.9) *possesses small amplitude quasi-periodic solutions with diophantine frequency vector $\omega := \omega(\xi) = (\omega_j)_{j \in S^+} \in \mathbb{R}^\nu$, of the form*

$$u(t,x) = \sum\nolimits_{j \in S^+} 2\sqrt{\xi_j}\, \cos(\omega_j t + jx) + o(\sqrt{|\xi|}), \quad \omega_j(\xi) = j^3 + O(|\xi|), \quad \forall j \in S^+\,, \tag{1.3.12}$$

for a "Cantor-like" set of small amplitudes $\xi \in \mathbb{R}^\nu_+$ with density 1 *at $\xi = 0$. The term $o(\sqrt{|\xi|})$ is small in some H^s-Sobolev norm, $s < q$. These quasi-periodic solutions are linearly stable. In addition, if the Hamiltonian density $f = f(u, u_x)$ does not depend on x, the theorem holds for any choice of the tangential sites S.*

We describe how to prove this Theorem in Chapter 6. We remark that the m-KdV equation

$$u_t + u_{xxx} \pm \partial_x(u^3) = 0$$

is completely integrable and the defocusing m-KdV admits global analytic action-angle coordinates, see Kappeler-Schaad-Topalov [50].

Notice that the Theorem 1.3.1 for the KdV equation (1.3.1) is more difficult than Theorem 1.3.2 for the m-KdV (1.3.9) because the nonlinearity is quadratic and not cubic.

We make some further comments.

1. It is possible to prove also the existence of quasi-periodic solutions for cubic perturbations of KdV, namely equations of the form

$$u_t + u_{xxx} - 6uu_x + a\partial_x u^3 + \mathcal{N}_4(x, u, u_x, u_{xx}, u_{xxx}) = 0 \,, \quad a \neq 0 \,.$$

2. A relevant point is that the fourth order Birkhoff normal form of KdV and mKdV is completely integrable. The present strategy of the proof — that we describe in detail below — is a rather general approach for constructing small amplitude quasi-periodic solutions of quasi-linear perturbed KdV equations. For example it could be applied to generalized KdV equations with leading nonlinearity u^p, $p \geq 4$, by using the normal form techniques of Procesi-Procesi [67]-[68].

3. A further interesting open question concerns perturbations of the finite gap solutions of KdV.

1.3.1 Ideas of the proof

We now describe the strategy of the proof of Theorem 1.3.1 for the KdV equation (1.3.1). The small changes required to prove Theorem 1.3.2 for the m-KdV equation (1.3.9) are given in Chapter 6.

Weak Birkhoff normal form. We decompose the phase space $H_0^1(\mathbb{T})$ in the symplectic subspaces

$$H_0^1(\mathbb{T}) = H_S \oplus H_S^\perp \,, \qquad H_S := \mathrm{span}\{e^{\mathrm{i}jx} : j \in S\} \,,$$

and according to the above decomposition, we write $u = v + z$, where $v \in H_S$ is called the tangential variable and $z \in H_S^\perp$ is called the normal one. The dynamics of these two components is quite different. The variable v contains the largest oscillations of the quasi-periodic solution (1.3.12), while z remains much closer to the origin. We write the Hamiltonian (1.3.3) as $H = H_2 + H_3 + H_{\geq 5}$, where H_2, H_3 are given in (5.1.2) and $H_{\geq 5}$ is defined in (5.1.1). We perform a "weak" Birkhoff normal form (weak BNF), whose goal is to find an invariant manifold of solutions of the third order approximate KdV equation (1.3.1), on which the dynamics is completely integrable, see Section 5.1. Thus we need to remove-normalize the monomials of the Hamiltonian H which are linear in z (this is the reason why we call this BNF only "weak"). Since the KdV nonlinearity is quadratic, two steps of weak BNF are required. In the first step we remove the cubic terms $O(v^3)$, $O(v^2 z)$, and in the second one we remove-normalize the terms $O(v^4)$, $O(v^3 z)$. The present Birkhoff map is close to the identity up to *finite dimensional* operators, see Proposition 5.1.1. The key advantage is that it modifies $\mathcal{N}_4$ very mildly, only up to finite dimensional operators (see for example Lemma 5.5.1), and thus the spectral analysis of the linearized equations (that we shall perform in Section 5.6) is essentially the same as if we were in the original coordinates.

The weak normal form (5.1.5) does not remove (or normalize) the monomials $O(z^2)$. This could be done. However, we do not perform such stronger normal form (called "partial BNF" in Pöschel [66]) because the corresponding Birkhoff map is close to the identity only up to an operator of order $O(\partial_x^{-1})$, and so it would produce, in the transformed vector field $\mathcal{N}_4$, terms of order ∂_{xx} and ∂_x. A fortiori, we cannot either use the full Birkhoff normal form computed in [49] for KdV, which completely diagonalizes the fourth order terms, because such Birkhoff map is only close to the identity up to a bounded operator. For the same reason, we do not use the global nonlinear Fourier transform in [49] (Birkhoff coordinates), which is close to the Fourier transform up to smoothing operators of order $O(\partial_x^{-1})$.

The weak BNF procedure of Section 5.1 is sufficient to find the first nonlinear (integrable) approximation of the solutions and to extract the "frequency-to-amplitude" modulation (5.2.10).

In Proposition 5.1.1 we also remove the terms $O(v^5)$, $O(v^4 z)$ in order to have sufficiently good approximate solutions so that the Nash-Moser iteration of Section 5.7 will converge. This is necessary for KdV whose nonlinearity is quadratic at the origin. These further steps of Birkhoff normal form are not required if the nonlinearity is yet cubic as for mKdV, see Remark 5.1.1. To this aim, we choose the tangential sites S such that ($\mathtt{S2}$) holds. We also note that we assume (1.3.4) because we use the conservation of momentum up to the homogeneity order 5, see (5.0.9).

Action-angle and rescaling. At this point, in Section 5.2 we introduce the action-angle variables (5.2.1) on the tangential sites and, after the rescaling (5.2.5), the Hamiltonian $\mathcal{H}$ in (5.1.5) transforms into the Hamiltonian

$$H_\varepsilon := \mathcal{N} + P\,, \qquad \mathcal{N} := \alpha(\xi) \cdot y + \frac{1}{2}\big(N(\theta)z, z\big)_{L^2(\mathbb{T})}$$

(see (5.2.9)), where $\alpha(\xi)$ is the frequency-to-amplitude relation defined in (5.2.10).

Note that the coefficients of the normal form $\mathcal{N}$ in (5.2.11) depend on the angles θ, unlike the usual KAM theorems [66], [53], where the whole normal form is reduced to constant coefficients (see Section 2.1). This is because the weak BNF of Section 5.1 did *not* normalize the quadratic terms $O(z^2)$. These terms are dealt with the "linear Birkhoff normal form" (linear BNF) in Sections 5.6.4, 5.6.5. In some sense here the "partial" Birkhoff normal form of [66] is splitted into the weak BNF of Section 5.1 and the linear BNF of Sections 5.6.4, 5.6.5.

The action-angle variables are convenient for proving the stability of the solutions.

The nonlinear functional setting. We look for an embedded invariant torus $i : \mathbb{T}^\nu \to \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$, $\varphi \to i(\varphi) := (\theta(\varphi), y(\varphi), z(\varphi))$ of the Hamiltonian vector field X_{H_ε} filled by quasi-periodic solutions with diophantine frequency ω. Notice that by (5.3.3), the diophantine frequency ω is $O(\varepsilon^2)$ close to the integer vector $\bar\omega$ in (5.0.3), therefore the diophantine constant γ in (5.3.4) satisfies $\gamma = o(\varepsilon^2)$. Actually, in order to find an invariant torus for X_{H_ε}, we look for zeros of the nonlinear operator

$$\mathcal{F}(i, \zeta) := \mathcal{F}(i, \zeta, \omega) := \omega \cdot \partial_\varphi i - X_{H_{\varepsilon,\zeta}}(i)\,, \tag{1.3.13}$$

(see (5.3.6)), where $X_{H_{\varepsilon,\zeta}}$ is the Hamiltonian vector field generated by the modified Hamiltonian $H_{\varepsilon,\zeta} := H_\varepsilon + \zeta \cdot \theta$ with $\zeta \in \mathbb{R}^\nu$. The unknowns in (1.3.13) are the embedded invariant torus i and ζ, the frequency ω plays the role of an "external" parameter. The auxiliary variable $\zeta \in \mathbb{R}^\nu$ is introduced in order to control the average in the y-component of the linearized equation (see (5.4.42), (5.4.43)). By Lemma 5.4.1, if $\mathcal{F}(i, \zeta) = 0$ then $\zeta = 0$ and thus $\varphi \to i(\varphi)$ is an invariant torus for the Hamiltonian vector field X_{H_ε}. The solution of the functional equation $\mathcal{F}(i, \zeta) = 0$, is obtained by a Nash-Moser iterative scheme in Sobolev scales. The key step is to construct (for ω restricted to a suitable Cantor-like set) an approximate inverse (*à la* Zehnder [76]) of the linearized operator $d_{i,\zeta}\mathcal{F}(i_0, \zeta_0)$ in (5.4.1) at any approximate solution (i_0, ζ_0). This means to find a linear operator $\mathbf{T}_0$ such that

$$d_{i,\zeta}\mathcal{F}(i_0, \zeta_0) \circ \mathbf{T}_0 - I = O(\gamma^{-1}\mathcal{F}(i_0, \zeta_0))\,,$$

See Theorem 5.4.1. Note that the operator $\mathbf{T}_0$ is an exact right inverse of the linearized operator, at an exact solution $\mathcal{F}(i_0, \zeta_0) = 0$.

A major difficulty is that the tangential and the normal dynamics near an invariant torus are strongly coupled. This difficulty is overcome by implementing the abstract procedure in Berti-Bolle [22]-[23] developed in order to prove existence of quasi-periodic solutions for autonomous NLW (and NLS) with a multiplicative potential. This approach reduces the search of an approximate inverse for (5.3.6) to the invertibility of a quasi-periodically forced PDE restricted on the normal directions. This method approximately decouples the "tangential" and the "normal" dynamics around an approximate invariant torus, introducing a suitable set of symplectic variables (ψ, η, w) near the torus, see (5.4.26). Note that, in the first line of (5.4.26), ψ is the "natural" angle variable which coordinates the torus, and, in the third line, the normal variable z is only translated by the component $z_0(\psi)$ of the torus. The second line completes this transformation to a symplectic one. The canonicity of this map is proved in Lemma 5.4.5, using the isotropy of the approximate invariant torus i_δ, see Lemma 5.4.4. The change of variable (5.4.26) brings the torus i_δ "at the origin". The advantage is that the second equation in (5.4.42) (which corresponds to the action variables of the torus) can be immediately solved, see (5.4.44). Then it remains to solve the third equation (5.4.45), i.e. to invert the linear operator $\mathcal{L}_\omega$. This is, up to finite dimensional remainders, a quasi-periodic Hamiltonian linear Airy equation perturbed by a variable coefficients differential operator of order $O(\partial_{xxx})$. The exact form of $\mathcal{L}_\omega$ is obtained in Proposition 5.5.1.

Reduction of the linearized operator in the normal directions. In Section 5.6 we conjugate the variable coefficients operator $\mathcal{L}_\omega$ in (5.5.34) to a diagonal operator with constant coefficients which describes infinitely many harmonic oscillators

$$\dot{v}_j + \mu_j^\infty v_j = 0 \,, \quad \mu_j^\infty := \mathrm{i}(-m_3 j^3 + m_1 j) + r_j^\infty \in \mathrm{i}\mathbb{R} \,, \quad j \notin S \,, \tag{1.3.14}$$

where the constants $m_3 - 1$, $m_1 \in \mathbb{R}$ and $\sup_j |r_j^\infty|$ are small, see Theorem 4.2.1. The main perturbative effect to the spectrum (and the eigenfunctions) of $\mathcal{L}_\omega$ is clearly due to the term $a_1(\omega t, x)\partial_{xxx}$ (see (5.5.34)), and it is too strong for the usual reducibility KAM techniques to work directly. The conjugacy of $\mathcal{L}_\omega$ with (1.3.14) is obtained in several steps. The first task (obtained in Sections 5.6.1-5.6.6) is to conjugate $\mathcal{L}_\omega$ to another Hamiltonian operator acting on $H_S^\perp$ with constant coefficients

$$\mathcal{L}_6 := \Pi_S^\perp \Big(\omega \cdot \partial_\varphi + m_3 \partial_{xxx} + m_1 \partial_x \Big) \Pi_S^\perp + R_6 \,, \quad m_1, m_3 \in \mathbb{R} \,, \tag{1.3.15}$$

up to a small bounded remainder $R_6 = O(\partial_x^0)$, see (5.6.113). This expansion of $\mathcal{L}_\omega$ in "decreasing symbols" with constant coefficients is similar to the one explained in Section 1.2.1 for the forced Airy equation and it is somehow in the spirit of the works of Iooss, Plotnikov and Toland [45]-[47] in water waves theory, and Baldi [4] for Benjamin-Ono. It is obtained by transformations which are very different from the usual KAM changes of variables. There are several differences with respect to the forced case:

1. The first step is to eliminate the x-dependence from the coefficient $a_1(\omega t, x)\partial_{xxx}$ of the Hamiltonian operator $\mathcal{L}_\omega$. We cannot use the symplectic transformation $\mathcal{A}$ defined in (5.6.1), used in Section 4.1.1, because $\mathcal{L}_\omega$ acts on the normal subspace $H_S^\perp$ only, and not on the whole

Sobolev space as $\mathcal{L}$ in (4.1.1). We can not use the restricted map $\mathcal{A}_\perp := \Pi_S^\perp \mathcal{A} \Pi_S^\perp$ which is *not* symplectic. In order to find a symplectic diffeomorphism of $H_S^\perp$ near $\mathcal{A}_\perp$, the first observation is to realize $\mathcal{A}$ as the flow map of the time dependent Hamiltonian transport linear PDE (5.6.3). Thus we conjugate $\mathcal{L}_\omega$ with the flow map of the projected Hamiltonian equation (5.6.5). In Lemma 5.6.1 we prove that it differs from $\mathcal{A}_\perp$ up to finite dimensional operators. A technical, but important, fact is that the remainders produced after this conjugation of $\mathcal{L}_\omega$ remain of the finite dimensional form (5.5.7), see Lemma 5.6.2.

This step may be seen as a quantitative application of the Egorov theorem, see [71], which describes how the principal symbol of a pseudo-differential operator (here $a_1(\omega t, x)\partial_{xxx}$) transforms under the flow of a linear hyperbolic PDE (here (5.6.5)).

2. Since the weak BNF procedure of Section 5.1 did not touch the quadratic terms $O(z^2)$, the operator $\mathcal{L}_\omega$ has variable coefficients also at the orders $O(\varepsilon)$ and $O(\varepsilon^2)$, see (5.5.34)-(5.5.35). These terms cannot be reduced to constants by the perturbative scheme, which applies to terms R such that $R\gamma^{-1} \ll 1$ where γ is the diophantine constant of the frequency vector ω. Here, since KdV is completely resonant, such $\gamma = o(\varepsilon^2)$, see (5.3.4). These terms are reduced to constant coefficients in Sections 5.6.4-5.6.5 by means of purely algebraic arguments (linear BNF), which, ultimately, stem from the complete integrability of the fourth order BNF of the KdV equation (1.3.5), see [49].

The order of the transformations of Sections 5.6.1-5.6.7 used to reduce $\mathcal{L}_\omega$ is not accidental. The first two steps in Sections 5.6.1, 5.6.2 reduce to constant coefficients the quasi-linear term $O(\partial_{xxx})$ and eliminate the term $O(\partial_{xx})$, see (5.6.45) (the second transformation is a time quasi-periodic reparametrization of time). Then, in Section 5.6.3, we apply the transformation $\mathcal{T}$ (5.6.64) in such a way that the space average of the coefficient $d_1(\varphi, \cdot)$ in (5.6.65) is constant. This is done in view of the applicability of the descent method in Section 5.6.6. All these transformations are composition operators induced by diffeomorphisms of the torus. Therefore they are well-defined operators of a Sobolev space into itself, but their decay norm is infinite! We perform the transformation $\mathcal{T}$ *before* the linear Birkhoff normal form steps of Sections 5.6.4-5.6.5, because $\mathcal{T}$ is a change of variable that preserves the form (5.5.7) of the remainders (it is not evident after the Birkhoff normal form). The Birkhoff transformations are symplectic maps of the form $I + \varepsilon O(\partial_x^{-1})$. Thanks to this property the coefficient $d_1(\varphi, x)$ obtained in step 5.6.3 is *not* changed by these Birkhoff maps. The transformation in Section 5.6.6 is one step of "descent method" which transforms $d_1(\varphi, x)\partial_x$ into a constant $m_1\partial_x$. It is at this point of the regularization procedure that the assumption (S1) on the tangential sites is used, so that the space average of the function $q_{>2}$ is zero, see Lemma 5.5.5. Actually we only need that the average of the function in (5.5.33) is zero. If $f_5 = 0$ (see (1.3.4)) then (S1) is not required. This completes the task of conjugating $\mathcal{L}_\omega$ to $\mathcal{L}_6$ in (1.3.15).

Finally, in Section 5.6.7 we apply the abstract reducibility Theorem 4.2.2, based on a quadratic KAM scheme, which completely diagonalizes the linearized operator, obtaining (1.3.14). The required smallness condition (5.6.115) for R_6 holds. Indeed the biggest term in R_6 comes from the conjugation of $\varepsilon\partial_x v_\varepsilon(\theta_0(\varphi), y_\delta(\varphi))$ in (5.5.35). The linear BNF procedure of Section 5.6.4 had eliminated its main contribution $\varepsilon\partial_x v_\varepsilon(\varphi, 0)$. It remains $\varepsilon\partial_x\big(v_\varepsilon(\theta_0(\varphi), y_\delta(\varphi)) - v_\varepsilon(\varphi, 0)\big)$ which has size $O(\varepsilon^{7-2b}\gamma^{-1})$ due to the estimate (5.4.4) of the approximate solution. This term enters in the variable coefficients of $d_1(\varphi, x)\partial_x$ and $d_0(\varphi, x)\partial_x^0$. The first one had been reduced to the constant operator $m_1\partial_x$ by the descent method of Section 5.6.6. The latter term is an operator of order $O(\partial_x^0)$ which satisfies (5.6.115). Thus $\mathcal{L}_6$ may be diagonalized by the iterative scheme of Theorem 4.2 in 4.2.2 which requires the smallness condition $O(\varepsilon^{7-2b}\gamma^{-2}) \ll 1$. This is the content of Section 5.6.7.

The Nash-Moser iteration. In Section 5.7 we perform the nonlinear Nash-Moser iteration which finally proves Theorem 5.3.1 and, therefore, Theorem 1.3.1. The optimal smallness condition required for the convergence of the scheme is $\varepsilon\|\mathcal{F}(\varphi, 0, 0)\|_{s_0+\mu}\gamma^{-2} \ll 1$, see (5.7.5). It is verified because $\|X_P(\varphi, 0, 0)\|_s \leq_s \varepsilon^{6-2b}$ (see (5.3.15)), which, in turn, is a consequence of having eliminated the terms $O(v^5), O(v^4 z)$ from the original Hamiltonian (5.1.1), see (5.1.5). This requires the condition ($\mathtt{S2}$).

The Thesis is organized as follows:

- In Chapter 2 we present the classical KAM approach for 1-dimensional Hamiltonian PDEs both for bounded and unbounded perturbations.

  More precisely, in Section 2.1 we describe the classical KAM Theorem of Kuksin [54] and Pöschel [65], for infinite dimensional Hamiltonian systems with bounded nonlinear perturbations. We follow the presentation in [65].

  In Section 2.2 we describe the KAM result obtained by Kuksin [55] and Kappeler and Pöschel [49] for unbounded Hamiltonian perturbations of KdV (equation 2.2.1).

- In Chapter 3 we introduce some definitions and technical tools which will be used in Chapters 4, 5, for the proof of the theorems stated in Sections 1.2, 1.3. We collect the properties of the matrix decay norm for a linear operator (Section 3.1) and we also define real, reversible and Hamiltonian operators, see Sections 3.2, 3.3. In Section 3.4 we give the dynamical interpretation of the reduction procedure of Sections 4.1, 4.2.

- In Chapter 4 we prove Theorems 1.2.1-1.2.5 concerning the existence and the stability of quasi-periodic solutions for the forced quasi-linear and fully nonlinear perturbed Airy equation (1.2.1).

- In Chapter 5 we prove Theorem 1.3.1 concerning the existence and the stability of quasi-periodic small-amplitude solutions of the quasi-linear Hamiltonian autonomous perturbed KdV equation (1.3.1).

- In Chapter 6 we explain how to modify the proof given in Chapter 5 for the quasi-linear KdV equation (1.3.1), to prove Theorem 1.3.2 for quasi-linear Hamiltonian perturbations of the modified KdV equation (1.3.9).

- In Chapter 7 we present some future perspectives.

- In Appendix A we collect classical tame estimates for product, composition of functions and changes of variables in Sobolev spaces.

Chapter 2

Classical KAM results for Hamiltonian PDEs

2.1 The classical KAM approach for Hamiltonian PDEs

As we have already said in Section 1.1, the normal form KAM approach has been introduced by Kuksin [53], [54], Wayne [73], Pöschel [65], [66] for 1-dimensional nonlinear Schrödinger and wave equations with bounded perturbations in the case of Dirichlet boundary conditions.

In the KAM framework it is usual to reduce the search of quasi-periodic solutions of parameter independent equations to the search of invariant tori for parameter dependent Hamiltonians which are small perturbations of a quadratic *normal form*. In the applications to the parameter independent NLS and NLW equations (1.1.1), (1.1.2), the "unperturbed actions" are introduced as parameters thanks to the non-degeneracy of the Birkhoff normal form.

Following [65], we consider a ξ-dependent family of real-valued Hamiltonians

$$H(\theta, I, z, \bar{z}, \xi) := N(I, z, \bar{z}, \xi) + P(\theta, I, z, \bar{z}, \xi)\,, \tag{2.1.1}$$

defined on the phase space

$$\mathcal{P}^{a,p} := \mathbb{T}^\nu \times \mathbb{R}^\nu \times \ell^{a,p} \times \ell^{a,p}\,,$$

where $\ell^{a,p}$ is the complex Hilbert space of sequences $(w_n)_{n\geq 1}$ endowed with the norm $\| \, \|_{a,p}$, where

$$\|w\|_{a,p}^2 := \sum_{n\geq 1} |w_n|^2 |n|^{2p} e^{2an}\,.$$

The *normal form* is

$$N(I, z, \bar{z}, \xi) := \omega(\xi) \cdot I + \Omega(\xi) \cdot z\bar{z}\,, \tag{2.1.2}$$

where

$$\omega(\xi) := (\omega_1(\xi), \ldots, \omega_\nu(\xi)) \in \mathbb{R}^\nu$$

are called the *tangential frequencies* and

$$\Omega(\xi) := \mathrm{diag}_{j\geq 1}\Omega_j(\xi)\,,$$

$\Omega_j \in \mathbb{R}$ for all $j \geq 1$, are called the *normal frequencies*. We use the notation

$$\Omega(\xi) \cdot z\bar{z} := \sum_{j\geq 1} \Omega_j(\xi) z_j \bar{z}_j \, .$$

The perturbation P is analytic in all its variables and the parameters ξ are in Π which is a bounded domain on $\mathbb{R}^\nu$. The symplectic form on the phase space $\mathcal{P}^{a,p}$ is

$$\mathcal{W} := \sum_{i=1}^{\nu} d\theta_i \wedge dI_i - \mathrm{i} \sum_{j\geq 1} dz_j \wedge d\bar{z}_j \, .$$

The Hamiltonian vector field generated by the Hamiltonian $H : \mathcal{P}^{a,p} \to \mathbb{R}$ is defined as

$$X_H(\theta, I, z, \bar{z}) := (\partial_I H, -\partial_\theta H, \mathrm{i}\partial_{\bar{z}} H, -\mathrm{i}\partial_z H) \, ,$$

and the Hamilton equations are

$$\frac{d}{dt}(\theta, I, z, \bar{z}) = X_H(\theta, I, z, \bar{z}) \, .$$

The Poisson bracket between two Hamiltonians $G, F : \mathcal{P}^{a,p} \to \mathbb{R}$ are defined as

$$\{G, F\} := \mathcal{W}(X_G, X_F) = \sum_{i=1}^{\nu} \Big(\partial_{\theta_i} G \partial_{I_i} F - \partial_{I_i} G \partial_{\theta_i} F \Big) - \mathrm{i} \sum_{j\geq 1} \Big(\partial_{z_j} G \partial_{\bar{z}_j} F - \partial_{\bar{z}_j} G \partial_{z_j} F \Big) . \quad (2.1.3)$$

If in (2.1.1) the perturbation $P = 0$, the Hamilton equations become

$$\begin{cases} \dot{\theta} = \omega(\xi) \\ \dot{I} = 0 \\ \dot{z} = \mathrm{i}\Omega(\xi) z \\ \dot{\bar{z}} = -\mathrm{i}\Omega(\xi)\bar{z} \, . \end{cases}$$

This system admits the quasi-periodic solutions

$$\theta(t) = \theta_0 + \omega(\xi) t \, , \qquad I(t) = 0 \, , \qquad z(t) = 0 \, , \qquad \forall \xi \in \Pi \, ,$$

hence the trivial torus $\mathbb{T}^\nu \times \{0\} \times \{0\} \times \{0\}$ is invariant for the Hamiltonian vector field X_N. The purpose of KAM Theory is to prove the persistence of such invariant tori for the perturbed Hamiltonian $H = N + P$. In Kuksin [54] and Pöschel [65] the assumptions are the following.

- **(H1) Non degeneracy**. The map $\xi \to \omega(\xi)$ is a homeomorphism between Π and $\omega(\Pi)$ and it is Lipschitz continuous together with its inverse. Moreover for any $l \in \mathbb{Z}^\nu$, $j, j' \geq 1$ the sets

  $$\mathcal{R}^{(I)}_{lj} := \Big\{ \xi \in \Pi : \omega(\xi) \cdot l \pm \Omega_j(\xi) = 0 \Big\} \, ,$$

  $$\mathcal{R}^{(II)}_{ljj'} := \Big\{ \xi \in \Pi : \omega(\xi) \cdot l + \Omega_j(\xi) \pm \Omega_{j'}(\xi) = 0 \Big\}$$

  have zero Lebesgue measure.

  Note that this hyphothesis is violated in the presence of frequencies with double multiplicity. Indeed if $\Omega_j = \Omega_{j'}$ for some $j \neq j'$, the set $\mathcal{R}^{(II)}_{0jj'}$ coincides with the whole parameter space Π.

- **(H2) Frequency asymptotics**. There exist $d \geq 1$, $\delta < d-1$ such that the *normal frequencies* $\Omega_j(\xi)$ satisfy

$$\Omega_j(\xi) = j^d + \ldots + O(j^\delta), \qquad \forall j \geq 1,$$

where the dots stand for fixed lower order terms in j. More precisely for all $j \geq 1$ there exists $\overline{\Omega}_j = j^d + \ldots$, parameter independent, such that the functions

$$\xi \to j^{-\delta}[\Omega_j(\xi) - \overline{\Omega}_j]$$

are uniformly Lipschitz on Π.

- **(H3) Regularity**. The perturbation P is real analytic on $\mathcal{P}^{a,p}$ and Lipschitz with respect to the parameter $\xi \in \Pi$, moreover the Hamiltonian vector field

$$X_P := (\partial_I P, -\partial_\theta P, \mathrm{i}\partial_{\bar{z}} P, -\mathrm{i}\partial_z P)$$

satisfies

$$X_P : \mathcal{P}^{a,p} \to \mathcal{P}^{a,\bar{p}}, \qquad \begin{cases} \bar{p} \geq p, & d > 1 \\ \bar{p} > p, & d = 1 \end{cases},$$

and $p - \bar{p} \leq \delta < d - 1$. To make this assumption quantitative, we assume that X_P is real analytic on a complex neighbourhood $D(s,r)$ of $\mathbb{T}^\nu \times \{0\} \times \{0\} \times \{0\}$ defined as follows:

$$D(s,r) := \{|\mathrm{Im}\theta| \leq s\} \times \{|I| \leq r^2\} \times \{\|z\|_{a,p} + \|\bar{z}\|_{a,p} \leq r\}.$$

For $W = (\theta, I, z, \bar{z})$, we introduce the weighted norm

$$\|W\|_r := |\theta| + r^{-2}|I| + r^{-1}\|z\|_{a,p} + r^{-1}\|\bar{z}\|_{a,p},$$

and correspondingly we define

$$\|X_P\|_{r,D(s,r)} := \sup_{D(s,r)} \|X_P\|_r,$$

$$\|X_P\|^{\sup}_{r,D(s,r)} := \sup_{\xi\in\Pi} \|X_P(\cdot,\xi)\|_{r,D(s,r)},$$

$$\|X_P\|^{\mathrm{lip}}_{r,D(s,r)} := \sup_{\xi_1\neq\xi_2} \frac{\|X_P(\cdot,\xi_1) - X_P(\cdot,\xi_2)\|_{r,D(s,r)}}{|\xi_1 - \xi_2|}$$

and the weighted norm

$$\|X_P\|^{\mathrm{Lip}(\gamma)}_{r,D(s,r)} := \|X_P\|^{\sup}_{r,D(s,r)} + \gamma\|X_P\|^{\mathrm{lip}}_{r,D(s,r)}, \qquad \gamma \in (0,1).$$

We remark that the hyphothesis **(H3)** fails in the case of Hamiltonian systems with unbounded perturbations. In the next Section 2.2, we will explain how to deal with unbounded perturbations of order 1, in order to develop KAM theory for the semilinear KdV (1.1.4).

Now let us state the KAM theorem of Kuksin [54] and Pöschel [65].

Theorem 2.1.1. *Let us suppose that the Hamiltonian $H = N + P$ satisfies the assumptions* **(H1)**-**(H3)**, *and let*

$$\varepsilon := \|X_P\|^{\mathrm{Lip}(\gamma)}_{r,D(s,r)}\,.$$

There exists a small constant $\delta := \delta(s, \nu) > 0$, such that for

$$\varepsilon\gamma^{-1} \leq \delta\,,$$

there exists a Cantor set $\Pi_\infty \subset \Pi$, a Lipschitz continuous family of embedded tori

$$\Phi_\infty : \mathbb{T}^\nu \times \Pi_\infty \to \mathcal{P}^{a,\bar{p}}\,,$$

and a Lipschitz map $\omega_\infty : \Pi_\infty \to \mathbb{R}^\nu$ such that for all $\xi \in \Pi_\infty$, $\Phi_\infty(\cdot, \xi)$ is an invariant torus with frequency $\omega_\infty(\xi)$ for the Hamiltonian $H(\cdot, \xi) = N(\cdot, \xi) + P(\cdot, \xi)$. The frequency map ω_∞ satisfies

$$|\omega_\infty - \omega|^{\mathrm{Lip}(\gamma)} \leq C\varepsilon\,.$$

The map Φ_∞ is real analytic on the complex neighbourhood $|Im\theta| < \frac{s}{2}$ and it satisfies

$$\|\Phi_\infty - \Phi_0\|^{\mathrm{Lip}(\gamma)}_r \leq C\varepsilon\gamma^{-1}\,,$$

where Φ_0 is the trivial embedding of the torus $\mathbb{T}^\nu \times \{0\} \times \{0\} \times \{0\}$ into the phase space $\mathcal{P}^{a,p}$. The Cantor set Π_∞ satisfies

$$|\Pi \setminus \Pi_\infty| \leq C\gamma\,.$$

We give a short outline of the proof, following Pöschel [65]. The idea is to construct iteratively a sequence of symplectic transformations $\Phi_1, \Phi_2, \ldots, \Phi_k, \ldots$, which transform the Hamiltonian (2.1.1) into another Hamiltonian which has an invariant torus at the origin. At the k-th step of the iteration, we have a Hamiltonian

$$H_k = N_k + P_k\,,$$

which is a perturbation P_k of a normal form N_k as in (2.1.2). We look for a symplectic transformation Φ_k such that

$$H_k \circ \Phi_k = N_{k+1} + P_{k+1}\,,$$

where N_{k+1} has still the form (2.1.2) and

$$X_{P_{k+1}} \approx X^{\alpha}_{P_k}\,, \qquad \text{for some} \quad \alpha > 1\,,$$

which implies that the iterative scheme converges super-exponentially fast.

Let us describe the step of this iteration in more details. To simplify notations we drop the index k and we write $+$ instead of $k + 1$. First we write

$$H = N + P = N + R + (P - R)\,, \tag{2.1.4}$$

where R is a truncation in the Taylor expansion of P defined as

$$\begin{aligned} R :=& P^{000}(\theta) + P^{010}(\theta) \cdot I + \langle P^{010}(\theta), z\rangle + \langle P^{001}(\theta), \bar{z}\rangle + \langle P^{020}(\theta)z, z\rangle \\ &+ \langle P^{011}(\theta)z, \bar{z}\rangle + \langle P^{002}\bar{z}, \bar{z}\rangle\,, \end{aligned} \tag{2.1.5}$$

with

$$\langle P^{010}(\theta), z\rangle := \sum_{j\geq 1} P^{010}_j(\theta) z_j\,, \qquad \langle P^{001}(\theta), \bar z\rangle := \sum_{j\geq 1} P^{001}_j(\theta) \bar z_j\,,$$

$$\langle P^{020}(\theta) z, z\rangle := \sum_{j,j'\geq 1} P^{020}_{jj'}(\theta) z_j z_{j'}\,, \quad \langle P^{011}(\theta) z, \bar z\rangle := \sum_{j,j'\geq 1} P^{011}_{jj'}(\theta) z_j \bar z_{j'}\,,$$

$$\langle P^{002}\bar z, \bar z\rangle := \sum_{j,j'\geq 1} P^{002}_{jj'}(\theta) \bar z_j \bar z_{j'}\,.$$

The purpose is to remove-normalize the term R. The reason is that if $R = 0$ the torus $\mathbb{T}^\nu \times \{0\} \times \{0\} \times \{0\}$ is invariant for H. To do this, we look for a symplectic transformation Φ as the time-1 flow map $(\Phi^t_F)_{|t=1}$ of a Hamiltonian F of the form (2.1.5), namely

$$\begin{aligned} F :=& F^{000}(\theta) + F^{010}(\theta)\cdot I + \langle F^{010}(\theta), z\rangle + \langle F^{001}(\theta), \bar z\rangle + \langle F^{020}(\theta) z, z\rangle \\ &+ \langle F^{011}(\theta) z, \bar z\rangle + \langle F^{002}\bar z, \bar z\rangle\,. \end{aligned} \tag{2.1.6}$$

Using the Lie expansion

$$H\circ\Phi^t_F := \exp(\mathrm{ad}_F(H)) = \sum_{n\geq 0} \frac{\mathrm{ad}^n_F(H) t^n}{n!}\,, \quad \mathrm{ad}_F(H) := \{H, F\}\,,$$

we get

$$H\circ\Phi = N\circ\Phi + R\circ\Phi + (P-R)\circ\Phi = N + R + \{N, F\} + Q + (P-R)\circ\Phi\,,$$

where Q is a quadratic term in R and F. The point is to normalize the *linear terms*, namely to solve the *homological equation*

$$\{N, F\} + R = [R]\,, \tag{2.1.7}$$

where

$$[R] = \widehat{e} + \widehat{\omega}\cdot I + \widehat{\Omega}\cdot z\bar z$$

$$\widehat{e} := \langle P^{000}\rangle_\theta\,, \quad \widehat{\omega} := \langle P^{100}\rangle_\theta\,, \quad \widehat{\Omega} := \mathrm{diag}_{j\geq 1}\langle P^{011}_{jj}\rangle_\theta\,,$$

and for any function $f(\theta)$, $\theta\in\mathbb{T}^\nu$, the notation $\langle f\rangle_\theta$ stands for the average of f on the torus $\mathbb{T}^\nu$. According to (2.1.3), (2.1.5), (2.1.6), the equation (2.1.7) leads to solve

$$\omega\cdot\partial_\theta F^{000}(\theta) + P^{000}(\theta) = \widehat{e}\,, \quad \omega\cdot\partial_\theta F^{100}(\theta) + P^{100}(\theta) = \widehat{\omega}\,, \tag{2.1.8}$$

$$(\omega\cdot\partial_\theta + \mathrm{i}\Omega_j) F^{010}_j(\theta) + P^{010}_j(\theta) = 0\,, \quad (\omega\cdot\partial_\theta - \mathrm{i}\Omega_j) F^{001}_j(\theta) + P^{001}_j(\theta) = 0\,, \ \ \forall j\geq 1\,, \tag{2.1.9}$$

$$(\omega\cdot\partial_\theta + \mathrm{i}\Omega_j + \mathrm{i}\Omega_{j'}) F^{020}_{jj'}(\theta) + P^{020}_{jj'}(\theta) = 0\,, \quad (\omega\cdot\partial_\theta - \mathrm{i}\Omega_j - \mathrm{i}\Omega_{j'}) F^{002}_{jj'}(\theta) + P^{002}_{jj'}(\theta) = 0\,, \tag{2.1.10}$$

for all $j, j'\geq 1$,

$$(\omega\cdot\partial_\theta + \mathrm{i}\Omega_j - \mathrm{i}\Omega_{j'}) F^{011}_{jj'}(\theta) + P^{011}_{jj'}(\theta) = 0\,, \qquad \forall j\neq j'\,, \tag{2.1.11}$$

$$\omega\cdot\partial_\theta F^{011}_{jj}(\theta) + P^{011}_{jj}(\theta) = \langle P^{011}_{jj}\rangle_\theta\,, \quad \forall j\geq 1\,. \tag{2.1.12}$$

The equations (2.1.8)-(2.1.12) are *constant coefficients* partial differential equations on the torus $\mathbb{T}^\nu$. They can be solved by expanding the coefficients of F and P in Fourier series with respect to θ and imposing the following non-resonance conditions:

- **Zero-th order Melnikov conditions**

$$|\omega \cdot l| \geq \frac{\gamma}{|l|^\tau}\,, \qquad \forall l \in \mathbb{Z}^\nu \setminus \{0\}\,, \tag{2.1.13}$$

to solve the equations in (2.1.8), (2.1.12).

- **First order Melnikov conditions**

$$|\omega \cdot l \pm \Omega_j| \geq \frac{\gamma}{|l|^\tau}\,, \qquad \forall l \in \mathbb{Z}^\nu\,,\ j \geq 1\,, \tag{2.1.14}$$

to solve the equations in (2.1.9).

- **Second order Melnikov conditions**

$$|\omega \cdot l + \Omega_j + \Omega_{j'}| \geq \frac{\gamma}{|l|^\tau}\,, \quad \forall l \in \mathbb{Z}^\nu\,,\ j, j' \geq 1 \tag{2.1.15}$$

to solve the equations in (2.1.10) and

$$|\omega \cdot l + \Omega_j - \Omega_{j'}| \geq \frac{\gamma}{|l|^\tau}\,, \qquad \forall (l, j, j') \neq (0, j, j) \tag{2.1.16}$$

to solve the equation in (2.1.11).

Since the Hamiltonian vector field X_P is a bounded nonlinear operator of the phase space (see Assumption **(H3)**), the solution F of the Homological equation (2.1.7) generates a Hamiltonian vector field X_F which is bounded too, in particular it satisfies the estimate

$$\|X_F\|^{\mathrm{Lip}(\gamma)}_{r, D(s-\sigma, r)} \leq C\gamma^{-1}\sigma^{-C(\tau,\nu)}\|X_P\|^{\mathrm{Lip}(\gamma)}_{r, D(s,r)}\,, \qquad \forall\ 0 < \sigma \leq s\,.$$

This implies that the Hamiltonian flow Φ_F^t generated by the Hamiltonian vector field X_F is a bounded symplectic transformation of the phase space $\mathcal{P}^{a,p}$ *onto itself*. This is the core of the iterative procedure.

Actually the previous approach proves also the existence of a KAM normal form nearby the invariant tori found in Theorem 2.1.1. Indeed it is possible to prove that the normal forms

$$N_k = \omega_k(\xi) \cdot I + \Omega_k(\xi) \cdot z\bar{z}$$

converge on the analytic domain $D(s/2, r/2)$ to a normal form

$$N_\infty := \omega_\infty(\xi) \cdot I + \Omega_\infty(\xi) \cdot z\bar{z}\,,$$

where the frequencies ω_∞, Ω_∞ satisfy the Melnikov non-resonance conditions (2.1.13)-(2.1.16). The transformations Φ_k converge to a symplectic analytic map

$$\Phi_\infty : D(s/2, r/2) \times \Pi_\infty \to D(s, r)\,.$$

Moreover the transformed Hamiltonian $H_\infty : D(s/2, r/2) \times \Pi_\infty \to \mathbb{R}$ is

$$H_\infty := H \circ \Phi_\infty = N_\infty + P_\infty\,, \tag{2.1.17}$$

where the perturbation P_∞ does not contain the monomials of the type (2.1.5) in its Taylor expansion, namely according to the splitting (2.1.4), (2.1.5), one has $R_\infty = 0$. For the Hamiltonian (2.1.17) the trivial torus $\mathbb{T}^\nu \times \{0\} \times \{0\} \times \{0\}$ is an invariant torus with frequency $\omega_\infty(\xi)$, for all $\xi \in \Pi_\infty$.

Remark 2.1.1. *In the KAM formulation of Berti-Biasco in [14], the Cantor set Π_∞ is defined (in a more convenient way) only in terms of the final frequencies $(\omega_\infty, \Omega_\infty)$ and not inductively at each step of the iteration.*

(Linear stability) The existence of the KAM normal form H_∞ in (2.1.17) implies the linear stability of the invariant KAM torus. Indeed the linearized equation at the trivial torus $\varphi \to (\varphi, 0, 0, 0)$ has the form

$$\begin{cases} \dot\theta = A(\omega_\infty t)I + B(\omega_\infty t)u \\ \dot I = 0 \\ \dot u = \mathrm{i}\Big(\boldsymbol{\Omega}_\infty u + C(\omega_\infty t)I\Big)\,, \end{cases} \tag{2.1.18}$$

where

$$u := (z, \bar z)\,, \qquad \boldsymbol{\Omega}_\infty := \begin{pmatrix} \Omega_\infty & 0 \\ 0 & -\Omega_\infty \end{pmatrix},$$

$$A(\omega_\infty t) : \mathbb{R}^\nu \to \mathbb{R}^\nu\,, \quad B(\omega_\infty t) : \ell^{a,p} \times \ell^{a,p} \to \mathbb{R}^\nu \quad C(\omega_\infty t) : \mathbb{R}^\nu \to \ell^{a,p} \times \ell^{a,p}\,.$$

Let $Z_1(\varphi)$, $Z_2(\varphi)$, $Z_3(\varphi)$ satisfy

$$\omega_\infty \cdot \partial_\varphi Z_1(\varphi) = \mathrm{i}\boldsymbol{\Omega}_\infty Z_1(\varphi) + \mathrm{i}C(\varphi)\,, \tag{2.1.19}$$

$$\omega_\infty \cdot \partial_\varphi Z_2(\varphi) = -\mathrm{i}Z_2(\varphi)\boldsymbol{\Omega}_\infty + A(\varphi)\,, \tag{2.1.20}$$

$$\omega_\infty \cdot \partial_\varphi Z_3(\varphi) = A(\varphi)Z_1(\varphi) + B(\varphi) - \mathbf{B}\,, \tag{2.1.21}$$

where $\mathbf{B}$ is the constant coefficients $\nu \times \nu$ matrix defined as

$$\mathbf{B} := \int_{\mathbb{T}^\nu} A(\varphi)Z_1(\varphi) + B(\varphi)\, d\varphi\,.$$

Notice that the equations (2.1.19)-(2.1.21) may be solved, since the frequencies $(\omega_\infty, \Omega_\infty)$ satisfy the Melnikov conditions (2.1.13), (2.1.14).

Under the change of coordinates

$$\begin{pmatrix} \theta \\ I \\ u \end{pmatrix} \to \begin{pmatrix} \theta + Z_3(\omega_\infty t)I + Z_2(\omega_\infty t)u \\ I \\ u + Z_1(\omega_\infty t)I \end{pmatrix}$$

(see Eliasson-Kuksin [37], Section 1.7) the linearized system (2.1.18) is reduced to the constant coefficients linear system

$$\begin{cases} \dot\theta = \mathbf{B}I \\ \dot I = 0 \\ \dot u = \mathrm{i}\boldsymbol{\Omega}_\infty u\,, \end{cases}$$

whose solutions are

$$\theta(t) = \theta_0 + (\mathbf{B}I_0)t\,, \quad I(t) = I_0\,, \quad u(t) = \exp(\mathrm{i}\boldsymbol{\Omega}_\infty t)u_0\,, \quad \forall t \in \mathbb{R}\,.$$

The linear stability follows, since the actions $I(t)$ remain constants and the normal variables $u(t)$ do not increase their norm $\|\ \|_{a,p}$.

2.2 KAM for unbounded perturbations

In this section we describe the KAM result proved in Kuksin [55] and Kappeler and Pöschel [49] for analytic Hamiltonian perturbations of the KdV equation

$$u_t + u_{xxx} - 6uu_x + \varepsilon\partial_x f(x,u) = 0\,, \qquad x \in \mathbb{T}\,, \tag{2.2.1}$$

where $\varepsilon > 0$ is a small parameter. In the above results the authors develop a perturbation theory for large *finite gap* solutions of KdV.

The equation (2.2.1) is a Hamiltonian PDE of the form $\partial_t u = \partial_x \nabla H(u)$, with Hamiltonian

$$H := H_{KdV} + \varepsilon K\,, \tag{2.2.2}$$

where

$$H_{KdV}(u) := \int_{\mathbb{T}} \frac{u_x^2}{2} + u^3\,dx\,, \quad K(u) := \int_{\mathbb{T}} F(x,u)\,dx\,, \qquad \partial_u F(x,u) = -f(x,u) \tag{2.2.3}$$

defined on the phase space $H^1_0(\mathbb{T})$.

For $\varepsilon = 0$ the equation (2.2.1) reduces to the KdV equation

$$u_t + u_{xxx} - 6uu_x = 0\,.$$

The KdV equation is completely integrable and it may be described by *global analytic action-angle coordinates*, see Theorem 2.2.1. In order to give a precise statement, we introduce some notations. For all $s \geq 0$, we define the real Hilbert spaces

$$\ell^s := \Big\{(x_1, x_2, \ldots) : \quad x_j \in \mathbb{R} \qquad \forall j\,, \qquad \|x\|_s^2 := \sum_{j\geq 1} |j|^{2s}|x_j|^2 < +\infty\Big\}\,, \tag{2.2.4}$$

and we consider the product $\ell^s \times \ell^s$ endowed with the symplectic form

$$\mathbf{w} := \sum_{j\geq 1} dx_j \wedge dy_j\,. \tag{2.2.5}$$

For all $s \geq 0$, we define

$$H^s_0(\mathbb{T},\mathbb{R}) := \Big\{u \in H^s(\mathbb{T},\mathbb{R}) : \int_{\mathbb{T}} u(x)\,dx = 0\Big\}$$

and we use the notation $L^2_0(\mathbb{T}) := H^0_0(\mathbb{T})$.

The following theorem has been first proved in [11], [12]. A different proof has been presented in [48]. Here we report the statement of Theorem 1.1 in Kappeler-Pöschel [49].

Theorem 2.2.1 (**Global Birkhoff coordinates**). *There exists a symplectic diffeomorphism*

$$\Phi : \ell^{\frac12} \times \ell^{\frac12} \to L^2_0(\mathbb{T})$$

which satisfies the following properties:

- Φ *is analytic togheter with its inverse, and it preserves the symplectic forms* (1.2.13)*,* (2.2.5)*.*

- *For any $s \geq 0$,*
$$\Phi : \ell^{s+\frac{1}{2}} \times \ell^{s+\frac{1}{2}} \to H^s_0(\mathbb{T}),$$
is invertible and analytic togheter with its inverse.

- *The transformed Hamiltonian*
$$\mathcal{H}_{KdV}(x,y) := (H_{KdV} \circ \Phi)(x,y)\,, \qquad (x,y) \in \ell^{s+\frac{1}{2}} \times \ell^{s+\frac{1}{2}}\,, \quad x = (x_n)_{n\geq 1}\,, \quad y = (y_n)_{n \geq 1}\,,$$
depends only on $x_n^2 + y_n^2$, for all $n \geq 1$.

As a consequence the transformed Hamiltonian $\mathcal{H}_{KdV}$ depends only on the actions
$$I := (I_n)_{n \geq 1}\,, \qquad I_n := \frac{x_n^2 + y_n^2}{2}\,, \quad \forall n \geq 1\,,$$
namely
$$\mathcal{H}_{KdV} = \mathcal{H}_{KdV}(I_1, I_2, \ldots)\,.$$
The Hamilton equations generated by $\mathcal{H}_{KdV}$ are
$$\begin{cases} \dot{x}_n = \omega_n(I) y_n \\ \dot{y}_n = -\omega_n(I) x_n\,, \end{cases} \qquad \omega_n(I) := \partial_{I_n} \mathcal{H}_{KdV}(I)\,, \qquad \forall n \geq 1\,,$$
whose solutions are
$$x_n(t) = \sqrt{2 I_{n,0}} \cos(\theta_{n,0} + \omega_n(I_0) t)\,, \qquad y_n(t) = \sqrt{2 I_{n,0}} \sin(\theta_{n,0} + \omega_n(I_0) t)$$
for all $n \geq 1$ and $I_0 := (I_{n,0})_{n \geq 1}$. This implies that all the solutions of the KdV equation are *periodic*, *quasi periodic* or *almost periodic* in time. The quasi-periodic solutions of KdV are called *finite gap* solutions. In the remaining part of this Chapter, we will explain (following the presentation of Kappeler and Pöschel in [49]) how to prove the persistence of such quasi-periodic solutions for the perturbed KdV equation (2.2.1).

Remark 2.2.1 (Frequency map). *It is proved in [49]-Section 15, that the frequency map*
$$I \to \omega(I) := (\omega_n(I))_{n \geq 1}\,, \quad \omega : \mathbb{P}\ell^1_3 \to \ell^\infty_{-3}\,,$$
where
$$\ell^1_3 := \Big\{ I := (I_n)_{n \geq 1} : \|I\|_{\ell^1_3} := \sum_{n \geq 1} |n|^3 |I_n| < +\infty \Big\}\,, \quad \mathbb{P}\ell^1_3 := \Big\{ I \in \ell^3_1 : I_n > 0\,, \quad \forall n \geq 1 \Big\}\,,$$
$$\ell^\infty_{-\beta} := \Big\{ I := (I_n)_{n \geq 1} : \|I\|_{\ell^\infty_{-3}} := \sup_{n \geq 1} |n|^{-\beta} |I_n| < +\infty \Big\} \quad \forall \beta \in \mathbb{R}\,.$$
Moreover
$$\omega = \lambda + \widetilde{\omega}\,, \quad \lambda := (n^3)_{n \in \mathbb{N}}\,, \tag{2.2.6}$$
(we restricted to the phase space $H^1_0(\mathbb{T})$ and so $c = 0$ in the definition of λ_n given in Corollary 15.2 of [49]) and the map
$$\widetilde{\omega} : \mathbb{P}\ell^1_3 \to \ell^\infty_{-1}$$

is real analytic (see Theorem 15.4 in [49]).

In addition, close to the origin $I = 0$, the frequency map has the following expansion

$$\omega_n(I) = n^3 - 6I_n + O(I^2)\,, \tag{2.2.7}$$

where $O(I^2)$ denotes the quadratic terms in the actions (see also Corollary 1.5 in [49]).

Under the Birkhoff map Φ, the Hamiltonian H in (2.2.2) transforms into the Hamiltonian

$$\mathcal{H}_{KdV} + \varepsilon \mathcal{K}\,, \tag{2.2.8}$$

where $\mathcal{H}_{KdV}$ is defined in Theorem 2.2.1 and $\mathcal{K} := K \circ \Phi$.

Let $A \subset \mathbb{N}$ be a finite set with cardinality $|A| = \nu$. We introduce the action-angle variables

$$\begin{cases} x_n := \sqrt{2(\xi_n + \mathbf{y}_n)}\cos(\theta_n) \\ y_n := \sqrt{2(\xi_n + \mathbf{y}_n)}\sin(\theta_n) \end{cases}, \qquad n \in A$$

$$z_n := \frac{x_n + \mathrm{i} y_n}{\sqrt{2}}\,, \qquad n \in \mathbb{N} \setminus A\,,$$

where for all $n \in A$, $\xi_n > 0$ and $|\mathbf{y}_n| < \xi_n$. In this coordinates the Hamiltonian $\mathcal{H}_{KdV}$ becomes

$$\mathcal{H}_{KdV} = \mathcal{H}_{KdV}(\xi + \mathbf{y}, z\bar{z})\,,$$

where $\xi := (\xi_n)_{n\in A}$, $\mathbf{y} := (\mathbf{y}_n)_{n \in A}$ and $z\bar{z} := (z_n \bar{z}_n)_{n \in \mathbb{N}\setminus A}$. Expanding in Taylor series we get

$$\mathcal{H}_{KdV}(\xi + \mathbf{y}, z\bar{z}) = \mathcal{H}_{KdV}(\xi, 0) + \sum_{n \in A} \omega_n(\xi)\mathbf{y}_n + \sum_{n \in \mathbb{N}\setminus A} \omega_n(\xi) z_n \bar{z}_n + Q(\mathbf{y}, z\bar{z})\,,$$

where Q stands for the Taylor remainder of second order in $\mathbf{y}$, $z\bar{z}$ and

$$\omega_n(\xi) := \partial_{I_n} \mathcal{H}_{KdV}(\xi, 0)\,, \qquad \forall n \geq 1\,.$$

Defining

$$\omega(\xi) := (\omega_n(\xi))_{n \in A}\,, \qquad \Omega(\xi) := (\omega_n(\xi))_{n \in \mathbb{N}\setminus A}\,,$$

the Hamiltonian (2.2.8) becomes in the coordinates

$$(\theta, \mathbf{y}, z, \bar{z}) \in \mathcal{U}^s := \mathbb{T}^\nu \times \mathbb{R}^\nu \times \ell^s_{\mathbb{C}} \times \ell^s_{\mathbb{C}} \tag{2.2.9}$$

($\ell^s_{\mathbb{C}}$ is the complexification of the space ℓ^s defined in (2.2.4))

$$H := N + P\,, \quad N := e(\xi) + \omega(\xi) \cdot \mathbf{y} + \Omega(\xi) \cdot z\bar{z}\,, \quad P := Q + \varepsilon \mathcal{K}\,. \tag{2.2.10}$$

Hence one is reduced to study the ξ-dependent family of Hamiltonians (2.2.10) which are (close to the origin $\mathbf{y} = 0$, $z = 0$) small perturbations of isochronous normal forms. Kappeler and Pöschel applied to such Hamiltonian the abstract KAM Theorem 16.1 in [49]. They assume the hyphotheses **(H1)** and **(H2)** (with $d > 1$) of Section 2.1 and they modify the hyphothesis **(H3)** in order to deal with unbounded perturbations of order 1 (see Assumption C, Page 136 in [49]).

As a consequence of the results recalled in Remark 2.2.1, the frequency map

$$\xi \to \omega(\xi),$$

is real analytic. Moreover, in Proposition 15.5-[49], it is proved to be non-degenerate, i.e.

$$\det[\partial_\xi \omega] \neq 0 .$$

This follows by the expansion (2.2.7) near the origin and analiticity.

The asymptotic expansion of the normal frequencies $\Omega(\xi) := (\omega_n(\xi))_{n\in\mathbb{N}\setminus A}$ may be proved thanks to the expansion (2.2.6), thus assumptions **(H1)**, **(H2)** (with $d = 3$) of Section 2.1 are verified (the proof is given in [49] Page 142).

The main issue in the KAM proof is that the Hamiltonian vector field associated to the perturbation in (2.2.10) is unbounded of order 1, namely

$$X_P : \mathcal{U}^s \to \mathcal{U}^{s-1}$$

(we recall (2.2.9)). Let us describe how it is possible to overcome this difficulty. The key idea, introduced by Kuksin in [55], is to work at each step of the KAM iteration, with a variable coefficients normal form, more precisely at the generic step of the KAM iteration we deal with a Hamiltonian

$$H = N + P ,$$

where the normal form

$$N := \omega(\xi) \cdot I + \Omega(\theta, \xi) \cdot z\bar{z} , \quad \Omega(\theta, \xi) \cdot z\bar{z} := \sum_{n\in\mathbb{N}} \Omega_n(\theta, \xi) z_n \bar{z}_n \tag{2.2.11}$$

depends on the angle θ. The reason is the following: since the normal frequencies grow asimptotically as $\Omega_j(\xi) \sim j^3$, we are able to impose non resonance conditions of the form

$$|\omega \cdot l + \Omega_j| \geq \frac{\gamma |j|^3}{|l|^\tau} , \quad \forall l \in \mathbb{Z}^\nu , \quad j \geq 1 , \tag{2.2.12}$$

$$|\omega \cdot l + \Omega_j - \Omega_{j'}| \geq \frac{\gamma (j^2 + j'^2)}{|l|^\tau} , \quad \forall (l, j, j') \neq (0, j, j) , \tag{2.2.13}$$

$$|\omega \cdot l + \Omega_j + \Omega_{j'})| \geq \frac{\gamma (j^2 + j'^2)}{|l|^\tau} , \quad \forall l \in \mathbb{Z}^\nu , \quad j, j' \geq 1 , \tag{2.2.14}$$

hence the coefficients of F in the equations (2.1.8)-(2.1.11) are bounded because P is unbounded of order 1 and the small divisors gain at least two space-derivatives thanks to the non resonance conditions above. The only problem is in the equation (2.1.12). To solve this equation, we require the zero-th order Melnikov conditions which do not give any smoothing effect, hence

$$P^{011}_{jj} = O(j) \quad \text{implies} \quad F^{011}_{jj} = O(j)$$

and thus

$$\text{diag}_j F^{011}_{jj}(\theta) : \ell^s_{\mathbb{C}} \to \ell^{s-1}_{\mathbb{C}} .$$

This would produce a vector field X_F unbounded of order 1. As a consequence, the diagonal angle-dependent terms $P^{011}_{jj}(\theta)$, $j \geq 1$ cannot be removed in the homological equation and they will give a contribution to the normal form.

The fact that the normal form has variable coefficients implies that the homological equations are non-constant coefficients linear equations. The coefficients of the Hamiltonian F in the homological equation (2.1.7) has to satisfy a first order non-constant coefficients linear PDE on the torus $\mathbb{T}^\nu$ of the form

$$\mathrm{i}\omega \cdot \partial_\theta u + \lambda u + b(\theta) u = f \,, \tag{2.2.15}$$

where $b, f : \mathbb{T}^\nu \to \mathbb{R}$ are given functions and $u : \mathbb{T}^\nu \to \mathbb{R}$ is the scalar unknown. These kind of equations can be solved thanks to the Kuksin's Lemma (see Chapter 5 in [49]).

Lemma 2.2.1 (Kuksin's Lemma). *Let us assume that*

- *there are some constants $\gamma, c_0 > 0$, $\tau > \nu$ such that $|\lambda| \geq \gamma c_0$ and the frequency vector $\omega \in \mathbb{R}^\nu$ satisfies the non resonance conditions*

$$|\omega \cdot l| \geq \frac{\gamma}{|l|^\tau} \,, \quad |\omega \cdot l + \lambda| \geq \frac{\gamma c_0}{|l|^\tau} \,, \qquad \forall l \in \mathbb{Z}^\nu \setminus \{0\} \,,$$

- *the function $b = \sum_{l \neq 0} b_l e^{\mathrm{i} l \cdot \theta}$ is analytic on the strip $D(s) := \{|\mathrm{Im}\theta| \leq s\}$ around the torus $\mathbb{T}^\nu$ and*

$$\|b\|_{s,\tau} := \sum_{l \in \mathbb{Z}^\nu} |l|^\tau e^{|l|s} |b_l| \leq \gamma \delta$$

  for some $\delta > 0$,

- *the function f is analytic on $D(s)$ and*

$$\|f\|_s := \|f\|_{s,0} < +\infty \,,$$

then for $\delta/(\gamma c_0)$ small enough there exists a solution u, analytic on the strip $D(s - \sigma)$, $0 < \sigma \leq s$ of the equation (2.2.15) *satisfying the estimate*

$$\|u\|_{s-\sigma} \leq \frac{C}{\gamma c_0 \sigma^{\tau+\nu+3}} \|f\|_s \,.$$

The above method may be adapted also to deal with Hamiltonian perturbations unbounded of order 2 like the equation

$$u_t + u_{xxx} - 6uu_x + \varepsilon \partial_x |\partial_x|^{\frac{1}{2}} f(x, |\partial_x|^{\frac{1}{2}} u) = 0 \,, \quad x \in \mathbb{T} \,.$$

The intuitive reason is that if

$$X_P : \mathcal{U}^s \to \mathcal{U}^{s-2},$$

using that the small divisors (2.2.12)-(2.2.14) gain two space derivatives, the solution F of the equations (2.1.8)-(2.1.11) produces a bounded vector field

$$X_F : \mathcal{U}^s \to \mathcal{U}^s \,.$$

The rigorous proof is based on the improved version of the Kuksin's Lemma proved by Liu-Yuan in [60].

This method does not certainly work for unbounded perturbations of order 3. Actually, in this case, the solution of the homological equations generates an unbounded vector field of order 1

$$X_F : \mathcal{U}^s \to \mathcal{U}^{s-1}$$

which does not define a flow on the same phase space $\mathcal{U}^s$ (we recall (2.2.9)).

In order to deal with unbounded perturbations of order 3, a sharper perturbative analysis is required. This is the content of the Chapters 4 and 5.

Chapter 3

Functional setting

In this Chapter we introduce some notations, definitions and technical tools which will be used along the proofs of Chapters 4, 5, 6.

Norms. Along this Thesis, we shall use the notation

$$\|u\|_s := \|u\|_{H^s(\mathbb{T}^{\nu+1})} := \|u\|_{H^s_{\varphi,x}} \tag{3.0.1}$$

to denote the Sobolev norm of functions $u = u(\varphi, x)$ in the Sobolev space $H^s(\mathbb{T}^{\nu+1})$. We shall denote by $\| \ \|_{H^s_x}$ the Sobolev norm in the phase space of functions $u := u(x) \in H^s(\mathbb{T})$. Moreover $\| \ \|_{H^s_\varphi}$ will denote the Sobolev norm of scalar functions, like the Fourier components $u_j(\varphi)$.

For a function $f : \Lambda_o \to E$, $\lambda \mapsto f(\lambda)$, where $(E, \| \ \|_E)$ is a Banach space and Λ_o is a subset of $\mathbb{R}^\nu$, we define the sup-norm and the Lipschitz semi-norm

$$\|f\|^{\sup}_E := \|f\|^{\sup}_{E,\Lambda_o} := \sup_{\lambda \in \Lambda_o} \|f(\lambda)\|_E \, , \quad \|f\|^{\mathrm{lip}}_E := \|f\|^{\mathrm{lip}}_{E,\Lambda_o} := \sup_{\substack{\lambda_1,\lambda_2 \in \Lambda_o \\ \lambda_1 \neq \lambda_2}} \frac{\|f(\lambda_1) - f(\lambda_2)\|_E}{|\lambda_1 - \lambda_2|} \, , \tag{3.0.2}$$

and, for $\gamma > 0$, the Lipschitz norm

$$\|f\|^{\mathrm{Lip}(\gamma)}_E := \|f\|^{\mathrm{Lip}(\gamma)}_{E,\Lambda_o} := \|f\|^{\sup}_E + \gamma \|f\|^{\mathrm{lip}}_E \, . \tag{3.0.3}$$

If $E = H^s$ we simply denote $\|f\|^{\mathrm{Lip}(\gamma)}_{H^s} := \|f\|^{\mathrm{Lip}(\gamma)}_s$. We fix $s_0 := (\nu+2)/2$ so that $H^{s_0}(\mathbb{T}^{\nu+1}) \hookrightarrow L^\infty(\mathbb{T}^{\nu+1})$ and the spaces $H^s(\mathbb{T}^{\nu+1})$, $s > s_0$, are an algebra. As a notation, we write

$$a \leq_s b \quad \Longleftrightarrow \quad a \leq C(s) b$$

for some constant $C(s)$ and for $s = s_0 := (\nu+2)/2$ we only write $a \lessdot b$. More in general the notation $a \lessdot b$ means $a \leq Cb$ where the constant C may depend on the data of the problems (1.2.1) and (1.3.1), namely the nonlinearities f and $\mathcal{N}_4$, the number ν of frequencies, the diophantine exponent $\tau > 0$ in the non-resonance conditions in (4.2.6), (5.6.120). Also the small constants δ in the sequel depend on the data of the problems.

3.1 Matrices with off-diagonal decay

Let $b \in \mathbb{N}$ and consider the exponential basis $\{e_i : i \in \mathbb{Z}^b\}$ of $L^2(\mathbb{T}^b)$, so that $L^2(\mathbb{T}^b)$ is the vector space $\{u = \sum u_i e_i, \sum |u_i|^2 < \infty\}$. Any linear operator $A : L^2(\mathbb{T}^b) \to L^2(\mathbb{T}^b)$ can be represented by the infinite dimensional matrix

$$(A_i^{i'})_{i,i'\in\mathbb{Z}^b}, \quad A_i^{i'} := (Ae_{i'}, e_i)_{L^2(\mathbb{T}^b)}, \quad Au = \sum_{i,i'} A_i^{i'} u_{i'} e_i.$$

We now define the s-norm (introduced in [21]) of an infinite dimensional matrix.

Definition 3.1.1. *The s-decay norm of an infinite dimensional matrix $A := (A_{i_1}^{i_2})_{i_1,i_2\in\mathbb{Z}^b}$ is*

$$|A|_s^2 := \sum_{i\in\mathbb{Z}^b} \langle i\rangle^{2s} \Big(\sup_{i_1-i_2=i} |A_{i_1}^{i_2}| \Big)^2 . \tag{3.1.1}$$

For parameter dependent matrices $A := A(\lambda)$, $\lambda \in \Lambda_o \subseteq \mathbb{R}^\nu$, the definitions (3.0.2) *and* (3.0.3) *become*

$$|A|_s^{\sup} := \sup_{\lambda\in\Lambda_o} |A(\lambda)|_s \, , \quad |A|_s^{\rm lip} := \sup_{\lambda_1\neq\lambda_2} \frac{|A(\lambda_1)-A(\lambda_2)|_s}{|\lambda_1-\lambda_2|} \, , \quad |A|_s^{{\rm Lip}(\gamma)} := |A|_s^{\sup} + \gamma |A|_s^{\rm lip} \, . \tag{3.1.2}$$

Clearly, the matrix decay norm (3.1.1) is increasing with respect to the index s, namely

$$|A|_s \le |A|_{s'} \, , \quad \forall s < s'.$$

The s-norm is designed to estimate the polynomial off-diagonal decay of matrices, actually it implies

$$|A_{i_1}^{i_2}| \le \frac{|A|_s}{\langle i_1 - i_2\rangle^s} \, , \quad \forall i_1, i_2 \in \mathbb{Z}^b \, ,$$

and, on the diagonal elements,

$$|A_i^i| \le |A|_0 \, , \quad |A_i^i|^{\rm lip} \le |A|_0^{\rm lip} \, . \tag{3.1.3}$$

We now list some properties of the matrix decay norm proved in [21].

Lemma 3.1.1. (Multiplication operator) *Let $p = \sum_i p_i e_i \in H^s(\mathbb{T}^b)$. The multiplication operator $h \mapsto ph$ is represented by the Töplitz matrix $T_i^{i'} = p_{i-i'}$ and*

$$|T|_s = \|p\|_s. \tag{3.1.4}$$

Moreover, if $p = p(\lambda)$ is a Lipschitz family of functions,

$$|T|_s^{{\rm Lip}(\gamma)} = \|p\|_s^{{\rm Lip}(\gamma)} \, . \tag{3.1.5}$$

The s-norm satisfies classical algebra and interpolation inequalities.

Lemma 3.1.2. (Interpolation) *For all $s \ge s_0 > b/2$ there are $C(s) \ge C(s_0) \ge 1$ such that*

$$|AB|_s \le C(s)|A|_{s_0}|B|_s + C(s_0)|A|_s|B|_{s_0} \, . \tag{3.1.6}$$

In particular, the algebra property holds

$$|AB|_s \le C(s)|A|_s|B|_s \, . \tag{3.1.7}$$

If $A = A(\lambda)$ and $B = B(\lambda)$ depend in a Lipschitz way on the parameter $\lambda \in \Lambda_o \subset \mathbb{R}$, then

$$|AB|_s^{{\rm Lip}(\gamma)} \le C(s)|A|_s^{{\rm Lip}(\gamma)}|B|_s^{{\rm Lip}(\gamma)} \, , \tag{3.1.8}$$

$$|AB|_s^{{\rm Lip}(\gamma)} \le C(s)|A|_s^{{\rm Lip}(\gamma)}|B|_{s_0}^{{\rm Lip}(\gamma)} + C(s_0)|A|_{s_0}^{{\rm Lip}(\gamma)}|B|_s^{{\rm Lip}(\gamma)}. \tag{3.1.9}$$

For all $n \geq 1$, using (3.1.7) with $s = s_0$, we get

$$|A^n|_{s_0} \leq [C(s_0)]^{n-1}|A|_{s_0}^n \qquad \text{and} \qquad |A^n|_s \leq n[C(s_0)|A|_{s_0}]^{n-1}C(s)|A|_s\,,\ \forall s \geq s_0\,. \tag{3.1.10}$$

Moreover (3.1.9) implies that (3.1.10) also holds for Lipschitz norms $|\ |_s^{\mathrm{Lip}(\gamma)}$.

The s-decay norm controls the Sobolev norm, also for Lipschitz families:

$$\begin{aligned} \|Ah\|_s \leq& C(s)\big(|A|_{s_0}\|h\|_s + |A|_s\|h\|_{s_0}\big), \\ \|Ah\|_s^{\mathrm{Lip}(\gamma)} \leq& C(s)\big(|A|_{s_0}^{\mathrm{Lip}(\gamma)}\|h\|_s^{\mathrm{Lip}(\gamma)} + |A|_s^{\mathrm{Lip}(\gamma)}\|h\|_{s_0}^{\mathrm{Lip}(\gamma)}\big). \end{aligned} \tag{3.1.11}$$

Lemma 3.1.3. *Let $\Phi = I + \Psi$ with $\Psi := \Psi(\lambda)$, depending in a Lipschitz way on the parameter $\lambda \in \Lambda_o \subset \mathbb{R}$, such that $C(s_0)|\Psi|_{s_0}^{\mathrm{Lip}(\gamma)} \leq 1/2$. Then Φ is invertible and, for all $s \geq s_0 > b/2$,*

$$|\Phi^{-1} - I|_s \leq C(s)|\Psi|_s\,, \quad |\Phi^{-1}|_{s_0}^{\mathrm{Lip}(\gamma)} \leq 2\,, \quad |\Phi^{-1} - I|_s^{\mathrm{Lip}(\gamma)} \leq C(s)|\Psi|_s^{\mathrm{Lip}(\gamma)}\,. \tag{3.1.12}$$

If $\Phi_i = I + \Psi_i$, $i = 1, 2$, satisfy $C(s_0)|\Psi_i|_{s_0}^{\mathrm{Lip}(\gamma)} \leq 1/2$, then

$$|\Phi_2^{-1} - \Phi_1^{-1}|_s \leq C(s)\big(|\Psi_2 - \Psi_1|_s + \big(|\Psi_1|_s + |\Psi_2|_s\big)|\Psi_2 - \Psi_1|_{s_0}\big)\,. \tag{3.1.13}$$

Proof. Estimates (3.1.12) follow by Neumann series and (3.1.10). To prove (3.1.13), observe that

$$\Phi_2^{-1} - \Phi_1^{-1} = \Phi_1^{-1}(\Phi_1 - \Phi_2)\Phi_2^{-1} = \Phi_1^{-1}(\Psi_1 - \Psi_2)\Phi_2^{-1}$$

and use (3.1.6), (3.1.12). □

Töplitz-in-time matrices

Let now $b := \nu + 1$ and

$$e_i(\varphi, x) := e^{\mathrm{i}(l\cdot\varphi + jx)}, \quad i := (l, j) \in \mathbb{Z}^b, \quad l \in \mathbb{Z}^\nu, \quad j \in \mathbb{Z}\,.$$

An important sub-algebra of matrices is formed by the matrices Töplitz in time defined by

$$A_{(l_1,j_1)}^{(l_2,j_2)} := A_{j_1}^{j_2}(l_1 - l_2)\,, \tag{3.1.14}$$

whose decay norm (3.1.1) is

$$|A|_s^2 = \sum_{j\in\mathbb{Z}, l\in\mathbb{Z}^\nu} \sup_{j_1 - j_2 = j} |A_{j_1}^{j_2}(l)|^2\langle l, j\rangle^{2s}\,. \tag{3.1.15}$$

These matrices are identified with the φ-dependent family of operators

$$A(\varphi) := \big(A_{j_1}^{j_2}(\varphi)\big)_{j_1,j_2\in\mathbb{Z}}\,, \quad A_{j_1}^{j_2}(\varphi) := \sum_{l\in\mathbb{Z}^\nu} A_{j_1}^{j_2}(l)e^{\mathrm{i}l\cdot\varphi} \tag{3.1.16}$$

which act on functions of the x-variable as

$$A(\varphi) : h(x) = \sum_{j\in\mathbb{Z}} h_j e^{\mathrm{i}jx} \mapsto A(\varphi)h(x) = \sum_{j_1,j_2\in\mathbb{Z}} A_{j_1}^{j_2}(\varphi)h_{j_2}e^{\mathrm{i}j_1 x}\,. \tag{3.1.17}$$

We still denote by $|A(\varphi)|_s$ the s-decay norm of the matrix in (3.1.16).

Lemma 3.1.4. *Let A be a Töplitz matrix as in* (3.1.14)*, and $s_0 := (\nu+2)/2$ (as defined above). Then*

$$|A(\varphi)|_s \leq C(s_0)|A|_{s+s_0}\,, \quad \forall \varphi \in \mathbb{T}^\nu\,.$$

Proof. For all $\varphi \in \mathbb{T}^\nu$ we have

$$\begin{aligned} |A(\varphi)|_s^2 \quad &:= \quad \sum_{j\in\mathbb{Z}} \langle j\rangle^{2s} \sup_{j_1-j_2=j} |A_{j_1}^{j_2}(\varphi)|^2 \lessdot \sum_{j\in\mathbb{Z}} \langle j\rangle^{2s} \sup_{j_1-j_2=j} \sum_{l\in\mathbb{Z}^\nu} |A_{j_1}^{j_2}(l)|^2 \langle l\rangle^{2s_0} \\ &\lessdot \quad \sum_{j\in\mathbb{Z}} \sup_{j_1-j_2=j} \sum_{l\in\mathbb{Z}^\nu} |A_{j_1}^{j_2}(l)|^2 \langle l,j\rangle^{2(s+s_0)} \lessdot \sum_{j\in\mathbb{Z}, l\in\mathbb{Z}^\nu} \sup_{j_1-j_2=j} |A_{j_1}^{j_2}(l)|^2 \langle l,j\rangle^{2(s+s_0)} \\ &\overset{(3.1.15)}{\lessdot} \quad |A|_{s+s_0}^2, \end{aligned}$$

whence the lemma follows. ☐

Given $N \in \mathbb{N}$, we define the smoothing operator Π_N as

$$\big(\Pi_N A\big)_{(l_1,j_1)}^{(l_2,j_2)} := \begin{cases} A_{(l_1,j_1)}^{(l_2,j_2)} & \text{if } |l_1-l_2| \leq N \\ 0 & \text{otherwise.} \end{cases} \tag{3.1.18}$$

Lemma 3.1.5. *The operator $\Pi_N^\perp := I - \Pi_N$ satisfies*

$$|\Pi_N^\perp A|_s \leq N^{-\beta}|A|_{s+\beta}\,, \quad |\Pi_N^\perp A|_s^{\mathrm{Lip}(\gamma)} \leq N^{-\beta}|A|_{s+\beta}^{\mathrm{Lip}(\gamma)}\,, \quad \beta \geq 0, \tag{3.1.19}$$

where in the second inequality $A := A(\lambda)$ is a Lipschitz family $\lambda \in \Lambda$.

3.2 Real and Reversible operators

We consider the space of *real* functions

$$Z := \{u(\varphi,x) = \overline{u(\varphi,x)}\}, \tag{3.2.1}$$

and of *even* (in space-time), respectively *odd*, functions

$$X := \{u(\varphi,x) = u(-\varphi,-x)\}\,, \quad Y := \{u(\varphi,x) = -u(-\varphi,-x)\}\,. \tag{3.2.2}$$

Definition 3.2.1. *An operator R is*

1. REAL *if $R : Z \to Z$*
2. REVERSIBLE *if $R : X \to Y$*
3. REVERSIBILITY-PRESERVING *if $R : X \to X$, $R : Y \to Y$.*

The composition of a reversible and a reversibility-preserving operator is reversible.

The above properties may be characterized in terms of matrix elements.

Lemma 3.2.1. *We have*

$$R : X \to Y \iff R_{-k}^{-j}(-l) = -R_k^j(l)\,, \qquad R : X \to X \iff R_{-k}^{-j}(-l) = R_k^j(l)\,,$$

$$R : Z \to Z \quad \iff \quad \overline{R_k^j(l)} = R_{-k}^{-j}(-l)\,.$$

3.3 Linear time-dependent Hamiltonian systems and Hamiltonian operators

In this Section we give some definitions and properties of the linear time-dependent Hamiltonian systems which will be used in Sections 4.1, 4.2.

Definition 3.3.1. *A time dependent linear vector field $X(t) : H_0^1(\mathbb{T}) \to H_0^1(\mathbb{T})$ is* HAMILTONIAN *if $X(t) = \partial_x G(t)$ for some real linear operator $G(t)$ which is self-adjoint with respect to the L^2 scalar product. The vector field $X(t)$ is generated by the quadratic Hamiltonian*

$$H(t,h) := \frac{1}{2}\big(G(t)h\,,\,h\big)_{L^2(\mathbb{T})} = \frac{1}{2}\int_{\mathbb{T}} G(t)[h]\,h\,dx\,, \qquad h \in H_0^1(\mathbb{T})\,.$$

If $G(t) = G(\omega t)$ is quasi-periodic in time, we say that the associated operator $\omega \cdot \partial_\varphi - \partial_x G(\varphi)$ (see (3.4.4)*) is Hamiltonian.*

Definition 3.3.2. *A linear operator $A : H_0^1(\mathbb{T}) \to H_0^1(\mathbb{T})$ is* SYMPLECTIC *if*

$$\Omega(Au, Av) = \Omega(u,v)\,, \quad \forall u,v \in H_0^1(\mathbb{T})\,, \tag{3.3.1}$$

where the symplectic 2-form Ω is defined in (1.2.13)*. Equivalently $A^T \partial_x^{-1} A = \partial_x^{-1}$.*

If $A(\varphi)$, $\forall \varphi \in \mathbb{T}^\nu$, is a family of symplectic maps we say that the operator A defined by $Ah(\varphi,x) := A(\varphi)h(\varphi,x)$, acting on the functions $h : \mathbb{T}^{\nu+1} \to \mathbb{R}$, is symplectic.

Under a time dependent family of symplectic transformations $u = \Phi(t)v$ the linear Hamiltonian equation

$$u_t = \partial_x G(t) u \quad \text{with Hamiltonian} \quad H(t,u) := \tfrac{1}{2}\big(G(t)u,u\big)_{L^2}$$

transforms into the equation

$$v_t = \partial_x E(t) v, \quad E(t) := \Phi(t)^T G(t)\Phi(t) - \Phi(t)^T \partial_x^{-1}\Phi_t(t)$$

with Hamiltonian

$$K(t,v) = \tfrac{1}{2}\big(G(t)\Phi(t)v, \Phi(t)v\big)_{L^2} - \tfrac{1}{2}\big(\partial_x^{-1}\Phi_t(t)v, \Phi(t)v\big)_{L^2}\,. \tag{3.3.2}$$

Note that $E(t)$ is self-adjoint with respect to the L^2 scalar product because $\Phi^T\partial_x^{-1}\Phi_t + \Phi_t^T\partial_x^{-1}\Phi = 0$. If the operators $G(t)$, $\Phi(t)$ are quasi-periodic in time, The Hamiltonian operator $\omega \cdot \partial_\varphi - \partial_x G(\varphi)$ transforms into the operator $\omega \cdot \partial_\varphi - \partial_x E(\varphi)$, which is still Hamiltonian, according to the definition 3.3.1.

3.4 Dynamical reducibility

All the transformations that we construct in Sections 4.1 and 4.2 act on functions $u(\varphi,x)$ (of time and space). They can also be seen as:

(a) transformations of the phase space H_x^s that depend quasi-periodically on time (Sections 4.1.1, 4.1.3-4.1.5 and 4.2);

(*b*) quasi-periodic reparametrizations of time (Section 4.1.2).

This observation allows to interpret the conjugacy procedure from a dynamical point of view.

Consider a quasi-periodic linear dynamical system

$$\partial_t u = L(\omega t)u. \tag{3.4.1}$$

We want to describe how (3.4.1) changes under the action of a transformation of type (a) or (b).

Let $A(\omega t)$ be of type (a), and let $u = A(\omega t)v$. Then (3.4.1) is transformed into the linear system

$$\partial_t v = L_+(\omega t)v \quad \text{where} \quad L_+(\omega t) = A(\omega t)^{-1}L(\omega t)A(\omega t) - A(\omega t)^{-1}\partial_t[A(\omega t)]\,. \tag{3.4.2}$$

The transformation $A(\omega t)$ may be regarded to act on functions $u(\varphi, x)$ as

$$(\tilde{A}u)(\varphi, x) := \big(A(\varphi)u(\varphi, \cdot)\big)(x) := A(\varphi)u(\varphi, x) \tag{3.4.3}$$

and one can check that $(\tilde{A}^{-1}u)(\varphi, x) = A^{-1}(\varphi)u(\varphi, x)$. The operator associated to (3.4.1) (on quasi-periodic functions)

$$\mathcal{L} := \omega \cdot \partial_\varphi - L(\varphi) \tag{3.4.4}$$

transforms under the action of $\tilde{A}$ into

$$\tilde{A}^{-1}\mathcal{L}\tilde{A} = \omega \cdot \partial_\varphi - L_+(\varphi),$$

which is exactly the linear system in (3.4.2), acting on quasi-periodic functions.

Now consider a transformation of type (b), namely a change of the time variable

$$\tau := t + \alpha(\omega t) \;\Leftrightarrow\; t = \tau + \tilde{\alpha}(\omega\tau); \quad (Bv)(t) := v(t + \alpha(\omega t)), \quad (B^{-1}u)(\tau) = u(\tau + \tilde{\alpha}(\omega\tau)), \tag{3.4.5}$$

where $\alpha = \alpha(\varphi)$, $\varphi \in \mathbb{T}^\nu$, is a 2π-periodic function of ν variables (in other words, $t \mapsto t + \alpha(\omega t)$ is the diffeomorphism of $\mathbb{R}$ induced by the transformation B). If $u(t)$ is a solution of (3.4.1), then $v(\tau)$, defined by $u = Bv$, solves

$$\partial_\tau v(\tau) = L_+(\omega\tau)v(\tau)\,, \quad L_+(\omega\tau) := \Big(\frac{L(\omega t)}{1 + (\omega \cdot \partial_\varphi \alpha)(\omega t)}\Big)_{|t=\tau+\tilde{\alpha}(\omega\tau)}\,. \tag{3.4.6}$$

We may regard the associated transformation on quasi-periodic functions defined by

$$(\tilde{B}h)(\varphi, x) := h(\varphi + \omega\alpha(\varphi), x)\,, \quad (\tilde{B}^{-1}h)(\varphi, x) := h(\varphi + \omega\tilde{\alpha}(\varphi), x)\,,$$

as in step 4.1.2, where we calculate

$$B^{-1}\mathcal{L}B = \rho(\varphi)\mathcal{L}_+\,, \quad \rho(\varphi) := B^{-1}(1 + \omega \cdot \partial_\varphi \alpha)\,,$$

$$\mathcal{L}_+ = \omega \cdot \partial_\varphi - L_+(\varphi)\,, \quad L_+(\varphi) := \frac{1}{\rho(\varphi)}L(\varphi + \omega\tilde{\alpha}(\varphi))\,. \tag{3.4.7}$$

(3.4.7) is nothing but the linear system (3.4.6), acting on quasi-periodic functions.

Chapter 4

KAM for quasi-linear and fully nonlinear forced perturbations of the Airy equation

In this Chapter we prove the Theorems (1.2.1)-(1.2.5). As we explained in Section 1.2.1, the core of the proof is the analysis of the linearized operator in (1.2.18). In Section 4.1 we perform a regularization procedure (using changes of variables induced by diffeomorphisms of the torus and pseudo-differential operators), which conjugates the linearized operator $\mathcal{L}$ to the operator $\mathcal{L}_5$ defined in (4.1.56), which is a diagonal operator plus a bounded remainder term. In Section 4.2, we perform a Nash-Moser KAM reducibility scheme in Sobolev class, which conjugates the operator $\mathcal{L}_5$ to the diagonal operator $\mathcal{L}_\infty$ defined in (4.2.7).

In Section 4.2.2 we construct a right inverse for the Linearized operator $\mathcal{L}$ satisfying tame estimates (see Theorem (4.2.3)) and finally in Section 4.3 we implement The Nash-Moser scheme, in order to construct a solution for the problem (1.2.4), concluding the proofs of Theorems (1.2.1)-(1.2.5).

4.1 Regularization of the linearized operator

Our existence proof is based on a Nash-Moser iterative scheme. The main step concerns the invertibility of the linearized operator (see (1.2.18))

$$\mathcal{L}h = \mathcal{L}(\lambda, u, \varepsilon)h := \omega \cdot \partial_\varphi h + (1 + a_3)\partial_{xxx}h + a_2\partial_{xx}h + a_1\partial_x h + a_0 h \tag{4.1.1}$$

obtained linearizing (1.2.4) at any approximate (or exact) solution u. The coefficients $a_i = a_i(\varphi, x) = a_i(u, \varepsilon)(\varphi, x)$ are periodic functions of (φ, x), depending on u, ε. They are explicitly obtained from the partial derivatives of $\varepsilon f(\varphi, x, z)$ as

$$a_i(\varphi, x) = \varepsilon(\partial_{z_i} f)\big(\varphi, x, u(\varphi, x), u_x(\varphi, x), u_{xx}(\varphi, x), u_{xxx}(\varphi, x)\big), \quad i = 0, 1, 2, 3. \tag{4.1.2}$$

The operator $\mathcal{L}$ depends on λ because $\omega = \lambda\bar{\omega}$. Since ε is a (small) fixed parameter, we simply write $\mathcal{L}(\lambda, u)$ instead of $\mathcal{L}(\lambda, u, \varepsilon)$, and $a_i(u)$ instead of $a_i(u, \varepsilon)$. We emphasize that the coefficients a_i do not depend explicitly on the parameter λ (they depend on λ only through $u(\lambda)$).

In the Hamiltonian case (1.2.11) the linearized operator (4.1.1) has the form

$$\mathcal{L}h = \omega \cdot \partial_\varphi h + \partial_x \Big(\partial_x \big\{ A_1(\varphi, x) \partial_x h \big\} - A_0(\varphi, x) h \Big)$$

where

$$A_1(\varphi, x) := 1 + \varepsilon (\partial_{z_1 z_1} F)(\varphi, x, u, u_x)\,, \quad A_0(\varphi, x) := -\varepsilon \partial_x \{ (\partial_{z_0 z_1} F)(\varphi, x, u, u_x) \} + \varepsilon (\partial_{z_0 z_0} F)(\varphi, x, u, u_x)$$

and it is generated by the quadratic Hamiltonian

$$H_L(\varphi, h) := \frac{1}{2} \int_{\mathbb{T}} \Big(A_0(\varphi, x) h^2 + A_1(\varphi, x) h_x^2 \Big)\, dx\,, \quad h \in H_0^1\,.$$

Remark 4.1.1. *In the reversible case, i.e. the nonlinearity f satisfies* (1.2.15) *and $u \in X$ (see* (3.2.2), (1.2.16)*) the coefficients a_i satisfy the parity*

$$a_3, a_1 \in X, \quad a_2, a_0 \in Y, \tag{4.1.3}$$

and $\mathcal{L}$ maps X into Y, namely $\mathcal{L}$ is reversible, see Definition 3.2.1.

Remark 4.1.2. *In the Hamiltonian case* (1.2.11)*, assumption (Q)-*(1.2.7) *is automatically satisfied (with $\alpha(\varphi) = 2$) because*

$$f(\varphi, x, u, u_x, u_{xx}, u_{xxx}) = a(\varphi, x, u, u_x) + b(\varphi, x, u, u_x) u_{xx} + c(\varphi, x, u, u_x) u_{xx}^2 + d(\varphi, x, u, u_x) u_{xxx}$$

where

$$b = 2(\partial^3_{z_1 z_1 x} F) + 2 z_1 (\partial^3_{z_1 z_1 z_0} F), \qquad c = \partial^3_{z_1} F, \qquad d = \partial^2_{z_1} F,$$

and so

$$\partial_{z_2} f = b + 2 z_2 c = 2(d_x + z_1 d_{z_0} + z_2 d_{z_1}) = 2\big(\partial^2_{z_3 x} f + z_1 \partial^2_{z_3 z_0} f + z_2 \partial^2_{z_3 z_1} f + z_3 \partial^2_{z_3 z_2} f \big)\,.$$

The coefficients a_i, together with their derivative $\partial_u a_i(u)[h]$ with respect to u in the direction h, satisfy tame estimates:

Lemma 4.1.1. *Let $f \in C^q$, see* (1.2.3)*. For all $s_0 \le s \le q - 2$, $\|u\|_{s_0+3} \le 1$, we have, for all $i = 0, 1, 2, 3$,*

$$\|a_i(u)\|_s \le \varepsilon\, C(s) \big(1 + \|u\|_{s+3} \big), \tag{4.1.4}$$

$$\|\partial_u a_i(u)[h]\|_s \le \varepsilon\, C(s) \big(\|h\|_{s+3} + \|u\|_{s+3} \|h\|_{s_0+3} \big)\,. \tag{4.1.5}$$

If, moreover, $\lambda \mapsto u(\lambda) \in H^s$ is a Lipschitz family satisfying $\|u\|^{\mathrm{Lip}(\gamma)}_{s_0+3} \le 1$ (see (3.0.3)*), then*

$$\|a_i\|_s^{\mathrm{Lip}(\gamma)} \le \varepsilon\, C(s) \big(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+3} \big)\,. \tag{4.1.6}$$

Proof. The tame estimate (4.1.4) follows by Lemma A.0.8(i) applied to the function $\partial_{z_i} f$, $i = 0, \ldots, 3$, which is valid for $s + 1 \le q$. The tame bound (4.1.5) for

$$\partial_u a_i(u)[h] \stackrel{(4.1.2)}{=} \varepsilon \sum_{k=0}^{3} (\partial^2_{z_k z_i} f)\big(\varphi, x, u, u_x, u_{xx}, u_{xxx}\big)\, \partial_x^k h, \quad i = 0, \ldots, 3,$$

follows by (A.0.5) and applying Lemma A.0.8(i) to the functions $\partial^2_{z_k z_i} f$, which gives

$$\|(\partial^2_{z_k z_i} f)\big(\varphi, x, u, u_x, u_{xx}, u_{xxx}\big)\|_s \le C(s) \|f\|_{C^{s+2}} (1 + \|u\|_{s+3}),$$

for $s + 2 \le q$. The Lipschitz bound (4.1.6) follows similarly. $\square$

4.1.1 Step 1. Change of the space variable

We consider a φ-dependent family of diffeomorphisms of the 1-dimensional torus $\mathbb{T}$ of the form

$$y = x + \beta(\varphi, x), \tag{4.1.7}$$

where β is a (small) real-valued function, 2π periodic in all its arguments. The change of variables (4.1.7) induces on the space of functions the linear operator

$$(\mathcal{A}h)(\varphi, x) := h(\varphi, x + \beta(\varphi, x)). \tag{4.1.8}$$

The operator $\mathcal{A}$ is invertible, with inverse

$$(\mathcal{A}^{-1}v)(\varphi, y) = v(\varphi, y + \tilde\beta(\varphi, y)), \tag{4.1.9}$$

where $y \mapsto y + \tilde\beta(\varphi, y)$ is the inverse diffeomorphism of (4.1.7), namely

$$x = y + \tilde\beta(\varphi, y) \quad \Longleftrightarrow \quad y = x + \beta(\varphi, x). \tag{4.1.10}$$

Remark 4.1.3. *In the Hamiltonian case* (1.2.11) *we use, instead of* (4.1.8)*, the modified change of variable* (1.2.27) *which is symplectic, for each* $\varphi \in \mathbb{T}^\nu$*. Indeed, setting* $U := \partial_x^{-1} u$ *(and neglecting to write the* φ*-dependence)*

$$\begin{aligned}
\Omega(\mathcal{A}u, \mathcal{A}v) &= \int_{\mathbb{T}} \partial_x^{-1}\Big(\partial_x\big\{U(x+\beta(x))\big\}\Big)\,(1+\beta_x(x))v(x+\beta(x))\,dx \\
&= \int_{\mathbb{T}} U(x+\beta(x))(1+\beta_x(x))v(x+\beta(x))dx - c\int_{\mathbb{T}}(1+\beta_x(x))v(x+\beta(x))dx \\
&= \int_{\mathbb{T}} U(y)v(y)dy = \Omega(u,v)\,, \quad v \in H_0^1\,,
\end{aligned}$$

where c *is the average of* $U(x+\beta(x))$ *in* $\mathbb{T}$*. The inverse operator of* (1.2.27) *is* $(\mathcal{A}^{-1}v)(\varphi, y) = (1+\tilde\beta_y(\varphi, y))v(y+\tilde\beta(\varphi, y))$ *which is also symplectic.*

Now we calculate the conjugate $\mathcal{A}^{-1}\mathcal{L}\mathcal{A}$ of the linearized operator $\mathcal{L}$ in (4.1.1) with $\mathcal{A}$ in (4.1.8).

The conjugate $\mathcal{A}^{-1}a\mathcal{A}$ of any multiplication operator $a : h(\varphi, x) \mapsto a(\varphi, x)h(\varphi, x)$ is the multiplication operator $(\mathcal{A}^{-1}a)$ that maps $v(\varphi, y) \mapsto (\mathcal{A}^{-1}a)(\varphi, y)\, v(\varphi, y)$. By conjugation, the differential operators become

$$\begin{aligned}
\mathcal{A}^{-1}\omega\cdot\partial_\varphi\mathcal{A} &= \omega\cdot\partial_\varphi + \{\mathcal{A}^{-1}(\omega\cdot\partial_\varphi\beta)\}\,\partial_y, \\
\mathcal{A}^{-1}\partial_x\mathcal{A} &= \{\mathcal{A}^{-1}(1+\beta_x)\}\,\partial_y, \\
\mathcal{A}^{-1}\partial_{xx}\mathcal{A} &= \{\mathcal{A}^{-1}(1+\beta_x)^2\}\,\partial_{yy} + \{\mathcal{A}^{-1}(\beta_{xx})\}\,\partial_y, \\
\mathcal{A}^{-1}\partial_{xxx}\mathcal{A} &= \{\mathcal{A}^{-1}(1+\beta_x)^3\}\,\partial_{yyy} + \{3\mathcal{A}^{-1}[(1+\beta_x)\beta_{xx}]\}\,\partial_{yy} + \{\mathcal{A}^{-1}(\beta_{xxx})\}\,\partial_y,
\end{aligned}$$

where all the coefficients $\{A^{-1}(\ldots)\}$ are periodic functions of (φ, y). Thus (recall (4.1.1))

$$\mathcal{L}_1 := \mathcal{A}^{-1}\mathcal{L}\mathcal{A} = \omega\cdot\partial_\varphi + b_3(\varphi, y)\partial_{yyy} + b_2(\varphi, y)\partial_{yy} + b_1(\varphi, y)\partial_y + b_0(\varphi, y) \tag{4.1.11}$$

where

$$b_3 = \mathcal{A}^{-1}[(1+a_3)(1+\beta_x)^3], \qquad b_1 = \mathcal{A}^{-1}[\omega\cdot\partial_\varphi\beta + (1+a_3)\beta_{xxx} + a_2\beta_{xx} + a_1(1+\beta_x)], \quad (4.1.12)$$

$$b_0 = \mathcal{A}^{-1}(a_0), \qquad b_2 = \mathcal{A}^{-1}[(1+a_3)3(1+\beta_x)\beta_{xx} + a_2(1+\beta_x)^2]. \quad (4.1.13)$$

We look for $\beta(\varphi,x)$ such that the coefficient $b_3(\varphi,y)$ of the highest order derivative ∂_{yyy} in (4.1.11) does not depend on y, namely

$$b_3(\varphi,y) \stackrel{(4.1.12)}{=} \mathcal{A}^{-1}[(1+a_3)(1+\beta_x)^3](\varphi,y) = b(\varphi) \quad (4.1.14)$$

for some function $b(\varphi)$ of φ only. Since $\mathcal{A}$ changes only the space variable, $\mathcal{A}b = b$ for every function $b(\varphi)$ that is independent on y. Hence (4.1.14) is equivalent to

$$\big(1 + a_3(\varphi,x)\big)\big(1+\beta_x(\varphi,x)\big)^3 = b(\varphi), \quad (4.1.15)$$

namely

$$\beta_x = \rho_0, \qquad \rho_0(\varphi,x) := b(\varphi)^{1/3}\big(1+a_3(\varphi,x)\big)^{-1/3} - 1. \quad (4.1.16)$$

The equation (4.1.16) has a solution β, periodic in x, if and only if $\int_{\mathbb{T}} \rho_0(\varphi,x)\,dx = 0$. This condition uniquely determines

$$b(\varphi) = \left(\int_{\mathbb{T}} \big(1+a_3(\varphi,x)\big)^{-\frac{1}{3}}\,dx \right)^{-3} \quad (4.1.17)$$

(we recall that $\int_{\mathbb{T}}$ stands for $(2\pi)^{-1}\int_{\mathbb{T}}$). Then we fix the solution (with zero average) of (4.1.16),

$$\beta(\varphi,x) := (\partial_x^{-1}\rho_0)(\varphi,x)\,, \quad (4.1.18)$$

where ∂_x^{-1} is defined by linearity as

$$\partial_x^{-1} e^{\mathrm{i}jx} := \frac{e^{\mathrm{i}jx}}{\mathrm{i}j} \quad \forall j \in \mathbb{Z}\setminus\{0\}, \qquad \partial_x^{-1}1 = 0. \quad (4.1.19)$$

In other words, $\partial_x^{-1}h$ is the primitive of h with zero average in x.

With this choice of β, we get (see (4.1.11), (4.1.14))

$$\mathcal{L}_1 = \mathcal{A}^{-1}\mathcal{L}\mathcal{A} = \omega\cdot\partial_\varphi + b_3(\varphi)\partial_{yyy} + b_2(\varphi,y)\partial_{yy} + b_1(\varphi,y)\partial_y + b_0(\varphi,y), \quad (4.1.20)$$

where $b_3(\varphi) := b(\varphi)$ is defined in (4.1.17).

Remark 4.1.4. *In the reversible case, $\beta \in Y$ because $a_3 \in X$, see* (4.1.3)*. Therefore the operator A in* (4.1.8)*, as well as $\mathcal{A}^{-1}$ in* (4.1.9)*, maps $X \to X$ and $Y \to Y$, namely it is reversibility-preserving, see Definition 3.2.1. By* (4.1.3) *the coefficients of $\mathcal{L}_1$ (see* (4.1.12), (4.1.13)*) have parity*

$$b_3, b_1 \in X, \qquad b_2, b_0 \in Y, \quad (4.1.21)$$

and $\mathcal{L}_1$ maps $X \to Y$, namely it is reversible.

Remark 4.1.5. *In the Hamiltonian case* (1.2.11) *the resulting operator $\mathcal{L}_1$ in* (4.1.20) *is Hamiltonian and $b_2(\varphi,y) = 2\partial_y b_3(\varphi) \equiv 0$. Actually, by* (3.3.2)*, the corresponding Hamiltonian has the form*

$$K(\varphi,v) = \frac{1}{2}\int_{\mathbb{T}} b_3(\varphi)v_y^2 + B_0(\varphi,y)v^2\,dy\,, \quad (4.1.22)$$

for some function $B_0(\varphi,y)$.

4.1.2 Step 2. Time reparametrization

The goal of this Section is to make constant the coefficient of the highest order spatial derivative operator ∂_{yyy} of $\mathcal{L}_1$ in (4.1.20), by a quasi-periodic reparametrization of time. We consider a diffeomorphism of the torus $\mathbb{T}^\nu$ of the form

$$\varphi \mapsto \varphi + \omega\alpha(\varphi), \quad \varphi \in \mathbb{T}^\nu, \quad \alpha(\varphi) \in \mathbb{R}\,, \tag{4.1.23}$$

where α is a (small) *real* valued function, 2π-periodic in all its arguments. The induced linear operator on the space of functions is

$$(Bh)(\varphi, y) := h\big(\varphi + \omega\alpha(\varphi), \, y\big) \tag{4.1.24}$$

whose inverse is

$$(B^{-1}v)(\vartheta, y) := v\big(\vartheta + \omega\tilde{\alpha}(\vartheta), \, y\big) \tag{4.1.25}$$

where $\varphi = \vartheta + \omega\tilde{\alpha}(\vartheta)$ is the inverse diffeomorphism of $\vartheta = \varphi + \omega\alpha(\varphi)$. By conjugation, the differential operators become

$$B^{-1}\omega \cdot \partial_\varphi B = \rho(\vartheta)\, \omega \cdot \partial_\vartheta, \quad B^{-1}\partial_y B = \partial_y, \quad \rho := B^{-1}(1 + \omega \cdot \partial_\varphi \alpha). \tag{4.1.26}$$

Thus, see (4.1.20),

$$B^{-1}\mathcal{L}_1 B = \rho\, \omega \cdot \partial_\vartheta + \{B^{-1}b_3\}\, \partial_{yyy} + \{B^{-1}b_2\}\, \partial_{yy} + \{B^{-1}b_1\}\, \partial_y + \{B^{-1}b_0\}. \tag{4.1.27}$$

We look for $\alpha(\varphi)$ such that the (variable) coefficients of the highest order derivatives ($\omega \cdot \partial_\vartheta$ and ∂_{yyy}) are proportional, namely

$$\{B^{-1}b_3\}(\vartheta) = m_3\rho(\vartheta) = m_3\{B^{-1}(1 + \omega \cdot \partial_\varphi\alpha)\}(\vartheta) \tag{4.1.28}$$

for some constant $m_3 \in \mathbb{R}$. Since B is invertible, this is equivalent to require that

$$b_3(\varphi) = m_3\big(1 + \omega \cdot \partial_\varphi\alpha(\varphi)\big). \tag{4.1.29}$$

Integrating on $\mathbb{T}^\nu$ determines the value of the constant m_3,

$$m_3 := \int_{\mathbb{T}^\nu} b_3(\varphi)\, d\varphi. \tag{4.1.30}$$

Thus we choose the unique solution of (4.1.29) with zero average

$$\alpha(\varphi) := \frac{1}{m_3}\, (\omega \cdot \partial_\varphi)^{-1}(b_3 - m_3)(\varphi) \tag{4.1.31}$$

where $(\omega \cdot \partial_\varphi)^{-1}$ is defined by linearity

$$(\omega \cdot \partial_\varphi)^{-1} e^{\mathrm{i} l \cdot \varphi} := \frac{e^{\mathrm{i} l \cdot \varphi}}{\mathrm{i}\omega \cdot l}\,, \ l \neq 0\,, \quad (\omega \cdot \partial_\varphi)^{-1} 1 = 0\,.$$

With this choice of α we get (see (4.1.27), (4.1.28))

$$B^{-1}\mathcal{L}_1 B = \rho\, \mathcal{L}_2, \qquad \mathcal{L}_2 := \omega \cdot \partial_\vartheta + m_3\, \partial_{yyy} + c_2(\vartheta, y)\, \partial_{yy} + c_1(\vartheta, y)\, \partial_y + c_0(\vartheta, y), \tag{4.1.32}$$

where

$$c_i := \frac{B^{-1}b_i}{\rho}\,, \quad i = 0, 1, 2. \tag{4.1.33}$$

Remark 4.1.6. *In the reversible case, α is odd because b_3 is even (see (4.1.21)), and B is reversibility preserving. Since ρ (defined in (4.1.26)) is even, the coefficients $c_3, c_1 \in X$, $c_2, c_0 \in Y$ and $\mathcal{L}_2 : X \to Y$ is reversible.*

Remark 4.1.7. *In the Hamiltonian case, the operator $\mathcal{L}_2$ is still Hamiltonian (the new Hamiltonian is the old one at the new time, divided by the factor ρ). The coefficient $c_2(\vartheta, y) \equiv 0$ because $b_2 \equiv 0$, see remark 4.1.5.*

4.1.3 Step 3. Descent method: step zero

The aim of this Section is to eliminate the term of order ∂_{yy} from $\mathcal{L}_2$ in (4.1.32).

Consider the multiplication operator

$$\mathcal{M}h := v(\vartheta, y)h \tag{4.1.34}$$

where the function v is periodic in all its arguments. Calculate the difference

$$\mathcal{L}_2 \mathcal{M} - \mathcal{M}\,(\omega \cdot \partial_\vartheta + m_3 \partial_{yyy}) = T_2 \partial_{yy} + T_1 \partial_y + T_0, \tag{4.1.35}$$

where

$$T_2 := 3m_3 v_y + c_2 v, \quad T_1 := 3m_3 v_{yy} + 2c_2 v_y + c_1 v, \quad T_0 := \omega \cdot \partial_\vartheta v + m_3 v_{yyy} + c_2 v_{yy} + c_1 v_y + c_0 v. \tag{4.1.36}$$

To eliminate the factor T_2, we need

$$3m_3 v_y + c_2 v = 0. \tag{4.1.37}$$

Equation (4.1.37) has the periodic solution

$$v(\vartheta, y) = \exp\Big\{ -\frac{1}{3m_3}\,(\partial_y^{-1} c_2)(\vartheta, y) \Big\} \tag{4.1.38}$$

provided that

$$\int_{\mathbb{T}} c_2(\vartheta, y)\, dy = 0. \tag{4.1.39}$$

Let us prove (4.1.39). By (4.1.33), (4.1.26), for each $\vartheta = \varphi + \omega\alpha(\varphi)$ we get

$$\int_{\mathbb{T}} c_2(\vartheta, y)\, dy = \frac{1}{\{B^{-1}(1 + \omega \cdot \partial_\varphi \alpha)\}(\vartheta)} \int_{\mathbb{T}} (B^{-1} b_2)(\vartheta, y)\, dy = \frac{1}{1 + \omega \cdot \partial_\varphi \alpha(\varphi)} \int_{\mathbb{T}} b_2(\varphi, y)\, dy.$$

By the definition (4.1.13) of b_2 and changing variable $y = x + \beta(\varphi, x)$ in the integral (recall (4.1.8))

$$\begin{aligned} \int_{\mathbb{T}} b_2(\varphi, y)\, dy &\overset{(4.1.13)}{=} \int_{\mathbb{T}} \Big((1 + a_3) 3 (1 + \beta_x) \beta_{xx} + a_2 (1 + \beta_x)^2 \Big) (1 + \beta_x)\, dx \\ &\overset{(4.1.15)}{=} b(\varphi) \Big\{ 3 \int_{\mathbb{T}} \frac{\beta_{xx}(\varphi, x)}{1 + \beta_x(\varphi, x)}\, dx + \int_{\mathbb{T}} \frac{a_2(\varphi, x)}{1 + a_3(\varphi, x)}\, dx \Big\}. \end{aligned} \tag{4.1.40}$$

The first integral in (4.1.40) is zero because $\beta_{xx}/(1 + \beta_x) = \partial_x \log(1 + \beta_x)$. The second one is zero because of assumptions (Q)-(1.2.7) or (F)-(1.2.6), see (1.2.28). As a consequence (4.1.39) is proved, and (4.1.37) has the periodic solution v defined in (4.1.38). Note that v is close to 1 for ε small. Hence the multiplication operator $\mathcal{M}$ defined in (4.1.34) is invertible and $\mathcal{M}^{-1}$ is the multiplication operator for $1/v$. By (4.1.35) and since $T_2 = 0$, we deduce

$$\mathcal{L}_3 := \mathcal{M}^{-1} \mathcal{L}_2 \mathcal{M} = \omega \cdot \partial_\vartheta + m_3 \partial_{yyy} + d_1(\vartheta, y) \partial_y + d_0(\vartheta, y), \qquad d_i := \frac{T_i}{v}, \quad i = 0, 1. \tag{4.1.41}$$

Remark 4.1.8. *In the reversible case, since c_2 is odd (see Remark 4.1.6) the function v is even, then $\mathcal{M}$, $\mathcal{M}^{-1}$ are reversibility preserving and by* (4.1.36) *and* (4.1.41) *$d_1 \in X$ and $d_0 \in Y$, which implies that $\mathcal{L}_3 : X \to Y$.*

Remark 4.1.9. *In the Hamiltonian case, there is no need to perform this step because $c_2 \equiv 0$, see remark 4.1.7.*

4.1.4 Step 4. Change of space variable (translation)

Consider the change of the space variable

$$z = y + p(\vartheta)$$

which induces the operators

$$\mathcal{T}h(\vartheta, y) := h(\vartheta, y + p(\vartheta)), \quad \mathcal{T}^{-1}v(\vartheta, z) := v(\vartheta, z - p(\vartheta)). \tag{4.1.42}$$

The differential operators become

$$\mathcal{T}^{-1}\omega \cdot \partial_\vartheta \mathcal{T} = \omega \cdot \partial_\vartheta + \{\omega \cdot \partial_\vartheta p(\vartheta)\}\, \partial_z, \qquad \mathcal{T}^{-1}\partial_y \mathcal{T} = \partial_z.$$

Thus, by (4.1.41),

$$\mathcal{L}_4 := \mathcal{T}^{-1}\mathcal{L}_3\mathcal{T} = \omega \cdot \partial_\vartheta + m_3 \partial_{zzz} + e_1(\vartheta, z)\, \partial_z + e_0(\vartheta, z)$$

where

$$e_1(\vartheta, z) := \omega \cdot \partial_\vartheta p(\vartheta) + (\mathcal{T}^{-1}d_1)(\vartheta, z), \quad e_0(\vartheta, z) := (\mathcal{T}^{-1}d_0)(\vartheta, z). \tag{4.1.43}$$

Now we look for $p(\vartheta)$ such that the average

$$\int_{\mathbb{T}} e_1(\vartheta, z)\, dz = m_1\,, \quad \forall \vartheta \in \mathbb{T}^\nu\,, \tag{4.1.44}$$

for some constant $m_1 \in \mathbb{R}$ (independent of ϑ). Equation (4.1.44) is equivalent to

$$\omega \cdot \partial_\vartheta p = m_1 - \int_{\mathbb{T}} d_1(\vartheta, y)\, dy =: V(\vartheta). \tag{4.1.45}$$

The equation (4.1.45) has a periodic solution $p(\vartheta)$ if and only if $\int_{\mathbb{T}^\nu} V(\vartheta)\, d\vartheta = 0$. Hence we have to define

$$m_1 := \int_{\mathbb{T}^{\nu+1}} d_1(\vartheta, y)\, d\vartheta dy \tag{4.1.46}$$

and

$$p(\vartheta) := (\omega \cdot \partial_\vartheta)^{-1} V(\vartheta)\,. \tag{4.1.47}$$

With this choice of p, after renaming the space-time variables $z = x$ and $\vartheta = \varphi$, we have

$$\mathcal{L}_4 = \omega \cdot \partial_\varphi + m_3 \partial_{xxx} + e_1(\varphi, x)\, \partial_x + e_0(\varphi, x), \qquad \int_{\mathbb{T}} e_1(\varphi, x)\, dx = m_1\,, \ \ \forall \varphi \in \mathbb{T}^\nu\,. \tag{4.1.48}$$

Remark 4.1.10. *By* (4.1.45), (4.1.47) *and since $d_1 \in X$ (see remark 4.1.8), the function p is odd. Then $\mathcal{T}$ and $\mathcal{T}^{-1}$ defined in* (4.1.42) *are reversibility preserving and the coefficients e_1, e_0 defined in* (4.1.43) *satisfy $e_1 \in X$, $e_0 \in Y$. Hence $\mathcal{L}_4 : X \to Y$ is reversible.*

Remark 4.1.11. *In the Hamiltonian case the operator $\mathcal{L}_4$ is Hamiltonian, because the operator $\mathcal{T}$ in* (4.1.42) *is symplectic (it is a particular case of the change of variables* (1.2.27) *with $\beta(\varphi, x) = p(\varphi)$).*

4.1.5 Step 5. Descent method: conjugation by pseudo-differential operators

The goal of this Section is to conjugate $\mathcal{L}_4$ in (4.1.48) to an operator of the form $\omega \cdot \partial_\varphi + m_3 \partial_{xxx} + m_1 \partial_x + \mathcal{R}$ where the constants m_3, m_1 are defined in (4.1.30), (4.1.46), and $\mathcal{R}$ is a pseudo-differential operator of order 0.

Consider an operator of the form

$$\mathcal{S} := I + w(\varphi, x)\partial_x^{-1} \tag{4.1.49}$$

where $w : \mathbb{T}^{\nu+1} \to \mathbb{R}$ and the operator ∂_x^{-1} is defined in (4.1.19). Note that

$$\partial_x \partial_x^{-1} = \partial_x^{-1}\partial_x = \pi_0 \,, \quad \pi_0(u) := u - \int_{\mathbb{T}} u(x)\, dx \,. \tag{4.1.50}$$

A direct computation shows that the difference

$$\mathcal{L}_4 \mathcal{S} - \mathcal{S}(\omega \cdot \partial_\varphi + m_3 \partial_{xxx} + m_1 \partial_x) = r_1 \partial_x + r_0 + r_{-1}\partial_x^{-1} \tag{4.1.51}$$

where (using $\partial_x \pi_0 = \pi_0 \partial_x = \partial_x$, $\partial_x^{-1}\partial_{xxx} = \partial_{xx}$)

$$r_1 := 3m_3 w_x + e_1(\varphi, x) - m_1 \tag{4.1.52}$$

$$r_0 := e_0 + \big(3m_3 w_{xx} + e_1 w - m_1 w\big)\pi_0 \tag{4.1.53}$$

$$r_{-1} := \omega \cdot \partial_\varphi w + m_3 w_{xxx} + e_1 w_x + e_0 w \,. \tag{4.1.54}$$

We look for a periodic function $w(\varphi, x)$ such that $r_1 = 0$. By (4.1.52) and (4.1.44) we take

$$w = \frac{1}{3m_3}\partial_x^{-1}[m_1 - e_1]. \tag{4.1.55}$$

For ε small enough the operator $\mathcal{S}$ is invertible and we obtain, by (4.1.51),

$$\mathcal{L}_5 := \mathcal{S}^{-1}\mathcal{L}_4\mathcal{S} = \omega \cdot \partial_\varphi + m_3\partial_{xxx} + m_1\partial_x + \mathcal{R}, \qquad \mathcal{R} := \mathcal{S}^{-1}(r_0 + r_{-1}\partial_x^{-1}). \tag{4.1.56}$$

Remark 4.1.12. *In the reversible case, the function $w \in Y$, because $e_1 \in X$, see remark 4.1.10. Then $\mathcal{S}$, $\mathcal{S}^{-1}$ are reversibility preserving. By* (4.1.53) *and* (4.1.54)*, $r_0 \in Y$ and $r_{-1} \in X$. Then the operators $\mathcal{R}, \mathcal{L}_5$ defined in* (4.1.56) *are reversible, namely $\mathcal{R}, \mathcal{L}_5 : X \to Y$.*

Remark 4.1.13. *In the Hamiltonian case, we consider, instead of* (4.1.49)*, the modified operator*

$$\mathcal{S} := e^{\pi_0 w(\varphi,x)\partial_x^{-1}} := I + \pi_0 w(\varphi, x)\partial_x^{-1} + \ldots \tag{4.1.57}$$

which, for each $\varphi \in \mathbb{T}^\nu$, is symplectic. Actually $\mathcal{S}$ is the time one flow map of the Hamiltonian vector field $\pi_0 w(\varphi, x)\partial_x^{-1}$ which is generated by the Hamiltonian

$$H_{\mathcal{S}}(\varphi, u) := -\frac{1}{2}\int_{\mathbb{T}} w(\varphi, x)\big(\partial_x^{-1}u\big)^2 dx \,, \quad u \in H_0^1 \,.$$

The corresponding $\mathcal{L}_5$ in (4.1.56) *is Hamiltonian. Note that the operators* (4.1.57) *and* (4.1.49) *differ only for pseudo-differential smoothing operators of order $O(\partial_x^{-2})$ and of smaller size $O(w^2) = O(\varepsilon^2)$.*

4.1.6 Estimates on $\mathcal{L}_5$

Summarizing the steps performed in the previous Sections 4.1.1-4.1.5, we have (semi)-conjugated the operator $\mathcal{L}$ defined in (4.1.1) to the operator $\mathcal{L}_5$ defined in (4.1.56), namely

$$\mathcal{L} = \Phi_1 \mathcal{L}_5 \Phi_2^{-1}, \qquad \Phi_1 := \mathcal{A}\mathcal{B}\rho\mathcal{M}\mathcal{T}\mathcal{S}, \quad \Phi_2 := \mathcal{A}\mathcal{B}\mathcal{M}\mathcal{T}\mathcal{S} \tag{4.1.58}$$

(where ρ means the multiplication operator for the function ρ defined in (4.1.26)).

In the next lemma we give tame estimates for $\mathcal{L}_5$ and Φ_1, Φ_2. We define the constants

$$\sigma := 2\tau_0 + 2\nu + 17, \quad \sigma' := 2\tau_0 + \nu + 14 \tag{4.1.59}$$

where τ_0 is defined in (1.2.2) and ν is the number of frequencies.

Lemma 4.1.2. *Let $f \in C^q$, see (1.2.3), and $s_0 \le s \le q - \sigma$. There exists $\delta > 0$ such that, if $\varepsilon\gamma_0^{-1} < \delta$ (the constant γ_0 is defined in (1.2.2)), then, for all*

$$\|u\|_{s_0+\sigma} \le 1\,, \tag{4.1.60}$$

(i) the transformations Φ_1, Φ_2 defined in (4.1.58) are invertible operators of $H^s(\mathbb{T}^{\nu+1})$, and satisfy

$$\|\Phi_i h\|_s + \|\Phi_i^{-1} h\|_s \le C(s)\big(\|h\|_s + \|u\|_{s+\sigma}\|h\|_{s_0}\big), \tag{4.1.61}$$

for $i = 1, 2$. Moreover, if $u(\lambda)$, $h(\lambda)$ are Lipschitz families with

$$\|u\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \le 1, \tag{4.1.62}$$

then

$$\|\Phi_i h\|_s^{\mathrm{Lip}(\gamma)} + \|\Phi_i^{-1} h\|_s^{\mathrm{Lip}(\gamma)} \le C(s)\big(\|h\|_{s+3}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|h\|_{s_0+3}^{\mathrm{Lip}(\gamma)}\big), \quad i = 1, 2. \tag{4.1.63}$$

(ii) The constant coefficients m_3, m_1 of $\mathcal{L}_5$ defined in (4.1.56) satisfy

$$|m_3 - 1| + |m_1| \le \varepsilon C\,, \tag{4.1.64}$$

$$|\partial_u m_3(u)[h]| + |\partial_u m_1(u)[h]| \le \varepsilon C \|h\|_\sigma\,. \tag{4.1.65}$$

Moreover, if $u(\lambda)$ is a Lipschitz family satisfying (4.1.62), then

$$|m_3 - 1|^{\mathrm{Lip}(\gamma)} + |m_1|^{\mathrm{Lip}(\gamma)} \le \varepsilon C. \tag{4.1.66}$$

(iii) The operator $\mathcal{R}$ defined in (4.1.56) satisfies:

$$|\mathcal{R}|_s \le \varepsilon C(s)(1 + \|u\|_{s+\sigma}), \tag{4.1.67}$$

$$|\partial_u \mathcal{R}(u)[h]\,|_s \le \varepsilon C(s)\big(\|h\|_{s+\sigma'} + \|u\|_{s+\sigma}\|h\|_{s_0+\sigma'}\big)\,, \tag{4.1.68}$$

where $\sigma > \sigma'$ are defined in (4.1.59). Moreover, if $u(\lambda)$ is a Lipschitz family satisfying (4.1.62), then

$$|\mathcal{R}|_s^{\mathrm{Lip}(\gamma)} \le \varepsilon C(s)(1 + \|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}). \tag{4.1.69}$$

Finally, in the reversible case, the maps Φ_i, Φ_i^{-1}, $i = 1, 2$ are reversibility preserving and $\mathcal{R}, \mathcal{L}_5 : X \to Y$ are reversible. In the Hamiltonian case the operator $\mathcal{L}_5$ is Hamiltonian.

Proof. The proof is elementary. It is based on a repeated use of the tame estimates of the Lemmata of the Appendix A. For convenience, we split it into many points. We remind that $s_0 := (\nu+2)/2$ is fixed.

Estimates in Step 1.

1. — We prove that $b_3 = b$ defined in (4.1.17) satisfies the tame estimates

$$\|b_3 - 1\|_s \le \varepsilon\, C(s)\big(1 + \|u\|_{s+3}\big), \tag{4.1.70}$$

$$\|\partial_u b_3(u)[h]\|_s \le \varepsilon\, C(s)\big(\|h\|_{s+3} + \|u\|_{s+3}\|h\|_{s_0+3}\big), \tag{4.1.71}$$

$$\|b_3 - 1\|_s^{\mathrm{Lip}(\gamma)} \le \varepsilon\, C(s)\big(1 + \|u\|_{s+3}^{\mathrm{Lip}(\gamma)}\big). \tag{4.1.72}$$

Proof of (4.1.70). Write $b_3 = b$ (see (4.1.17)) as

$$b_3 - 1 = \psi\big(M[g(a_3) - g(0)]\,\big) - \psi(0), \tag{4.1.73}$$

where

$$\psi(t) := (1+t)^{-3}, \quad Mh := \int_{\mathbb{T}} h\, dx, \quad g(t) := (1+t)^{-\frac{1}{3}}.$$

Thus, for ε small,

$$\|b_3 - 1\|_s \le C(s)\|M[g(a_3) - g(0)]\,\|_s \le C(s)\|g(a_3) - g(0)\|_s \le C(s)\|a_3\|_s.$$

In the first inequality we have applied Lemma A.0.8(i) to the function ψ, with $u = 0$, $p = 0$, $h = M[g(a_3) - g(0)]$. In the second inequality we have used the trivial fact that $\|Mh\|_s \le \|h\|_s$ for all h. In the third inequality we have applied again Lemma A.0.8(i) to the function g, with $u = 0$, $p = 0$, $h = a_3$. Finally we estimate a_3 by (4.1.4) with $s_0 = s_0$, which holds for $s + 2 \le q$.
Proof of (4.1.71). Using (4.1.73), the derivative of b_3 with respect to u in the direction h is

$$\partial_u b_3(u)[h] = \psi'\big(M[g(a_3) - g(0)]\big)\, M\big(g'(a_3)\partial_u a_3[h]\,\big).$$

Then use (A.0.5), Lemma A.0.8(i) applied to the functions ψ' and g', and (4.1.5).
Proof of (4.1.72). It follows from (4.1.70), (4.1.71) and Lemma A.0.9.

2. — Using the definition (4.1.16) of ρ_0, estimates (4.1.70), (4.1.71), (4.1.72) for b_3 and estimates (4.1.4), (4.1.5), (4.1.6) for a_3, one proves that ρ_0 also satisfies the same estimates (4.1.70), (4.1.71), (4.1.72) as $(b_3 - 1)$. Since $\beta = \partial_x^{-1}\rho_0$ (see (4.1.18)), by Lemma A.0.7(i) we get

$$|\beta|_{s,\infty} \le C(s)\|\beta\|_{s+s_0} \le C(s)\|\rho_0\|_{s+s_0} \le \varepsilon\, C(s)\big(1 + \|u\|_{s+s_0+3}\big), \tag{4.1.74}$$

and, with the same chain of inequalities,

$$|\partial_u \beta(u)[h]|_{s,\infty} \le \varepsilon\, C(s)\big(\|h\|_{s+s_0+3} + \|u\|_{s+s_0+3}\|h\|_{s_0+3}\big)\,. \tag{4.1.75}$$

Then Lemma A.0.9 implies

$$|\beta|_{s,\infty}^{\mathrm{Lip}(\gamma)} \le \varepsilon\, C(s)\big(1 + \|u\|_{s+s_0+3}^{\mathrm{Lip}(\gamma)}\big), \tag{4.1.76}$$

for all $s + s_0 + 3 \le q$. Note that $x \mapsto x + \beta(\varphi, x)$ is a well-defined diffeomorphism if $|\beta|_{1,\infty} \le 1/2$, and, by (5.6.35), this condition is satisfied provided $\varepsilon\, C\big(1 + \|u\|_{s_0+4}\big) \le 1/2$.

Let $(\varphi, y) \mapsto (\varphi, y + \tilde\beta(\varphi, y))$ be the inverse diffeomorphism of $(\varphi, x) \mapsto (\varphi, x + \beta(\varphi, x))$. By Lemma A.0.10($i$) on the torus $\mathbb{T}^{\nu+1}$, $\tilde\beta$ satisfies

$$|\tilde\beta|_{s,\infty} \le C|\beta|_{s,\infty} \overset{(5.6.35)}{\le} \varepsilon\, C(s)\big(1 + \|u\|_{s+3+s_0}\big). \tag{4.1.77}$$

Writing explicitly the dependence on u, we have $\tilde\beta(\varphi, y; u) + \beta\big(\varphi,\, y + \tilde\beta(\varphi, y; u); u\big) = 0$. Differentiating the last equality with respect to u in the direction h gives

$$(\partial_u \tilde\beta)[h] = -\mathcal{A}^{-1}\Big(\frac{\partial_u \beta[h]}{1 + \beta_x}\Big),$$

therefore, applying Lemma A.0.10(iii) to deal with $\mathcal{A}^{-1}$, (A.0.6) for the product $(\partial_u\beta[h])(1+\beta_x)^{-1}$, the estimates (5.6.35), (4.1.75), (4.1.76) for β, and (A.0.2) (with $a_0 = s_0 + 3$, $b_0 = s_0 + 4$, $p = 1$, $q = s - 1$), we obtain (for $s + s_0 + 4 \le q$)

$$|\partial_u \tilde\beta(u)[h]|_{s,\infty} \le \varepsilon\, C(s)\big(\|h\|_{s+3+s_0} + \|u\|_{s+4+s_0}\|h\|_{3+s_0}\big)\,. \tag{4.1.78}$$

Then, using Lemma A.0.9 with $p = 4 + s_0$, the bounds (4.1.77), (4.1.78) imply

$$|\tilde\beta|^{\mathrm{Lip}(\gamma)}_{s,\infty} \le \varepsilon\, C(s)\big(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+4+s_0}\big). \tag{4.1.79}$$

3. — Estimates of $\mathcal{A}(u)$ and $\mathcal{A}(u)^{-1}$. By (A.0.16), (5.6.35) and (4.1.77),

$$\|\mathcal{A}(u)h\|_s + \|\mathcal{A}(u)^{-1}h\|_s \le\, C(s)\big(\|h\|_s + \|u\|_{s+s_0+3}\|h\|_1\big). \tag{4.1.80}$$

Moreover, by (A.0.18), (4.1.76) and (4.1.79),

$$\|\mathcal{A}(u)h\|^{\mathrm{Lip}(\gamma)}_s + \|\mathcal{A}(u)^{-1}h\|^{\mathrm{Lip}(\gamma)}_s \le C(s)\big(\|h\|^{\mathrm{Lip}(\gamma)}_{s+1} + \|u\|^{\mathrm{Lip}(\gamma)}_{s+s_0+4}\|h\|^{\mathrm{Lip}(\gamma)}_2\big). \tag{4.1.81}$$

Since $\mathcal{A}(u)g(\varphi, x) = g(\varphi, x{+}\beta(\varphi, x; u))$, the derivative of $\mathcal{A}(u)g$ with respect to u in the direction h is the product $\partial_u\big(\mathcal{A}(u)g\big)[h] \;=\; (\mathcal{A}(u)g_x)\,\partial_u\beta(u)[h]$. Then, by (A.0.7), (4.1.75) and (4.1.80),

$$\|\partial_u(\mathcal{A}(u)g)[h]\,\|_s \le \varepsilon C(s)\Big(\|g\|_{s+1}\|h\|_{s_0+3} + \|g\|_2\|h\|_{s+s_0+3} + \|u\|_{s+s_0+3}\|g\|_2\|h\|_{s_0+3}\Big). \tag{4.1.82}$$

Similarly $\partial_u(\mathcal{A}(u)^{-1}g)[h] = (\mathcal{A}(u)^{-1}g_x)\,\partial_u\tilde\beta(u)[h]$, therefore (A.0.7), (4.1.78), (4.1.80) imply that

$$\|\partial_u(\mathcal{A}^{-1}(u)g)[h]\,\|_s \le \varepsilon C(s)\Big(\|g\|_{s+1}\|h\|_{s_0+3} + \|g\|_2\|h\|_{s+s_0+3} + \|u\|_{s+s_0+4}\|g\|_2\|h\|_{s_0+3}\Big). \tag{4.1.83}$$

4. — The coefficients b_0, b_1, b_2 are given in (4.1.12), (4.1.13). By (A.0.7), (4.1.80), (4.1.62), (5.6.35) and (4.1.4),

$$\|b_i\|_s \le \varepsilon C(s)(1 + \|u\|_{s+s_0+6}), \quad i = 0, 1, 2. \tag{4.1.84}$$

Moreover, in analogous way, by (A.0.7), (4.1.81), (4.1.62), (4.1.76) and (4.1.6),

$$\|b_i\|^{\mathrm{Lip}(\gamma)}_s \le \varepsilon C(s)(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+s_0+7}), \quad i = 0, 1, 2. \tag{4.1.85}$$

Now we estimate the derivative with respect to u of b_1. The estimates for b_0 and b_2 are analogous. By (4.1.12) we write $b_1(u) = \mathcal{A}(u)^{-1}b_1^*(u)$ where $b_1^* := \omega\cdot\partial_\varphi\beta + (1 + a_3)\beta_{xxx} + a_2\beta_{xx} + a_1(1 + \beta_x)$. The bounds (4.1.5), (4.1.75), (5.6.35), (4.1.62), and (A.0.7) imply that

$$\|\partial_u b_1^*(u)[h]\|_s \le \varepsilon C(s)\big(\|h\|_{s+s_0+6} + \|u\|_{s+s_0+6}\|h\|_{s_0+6}\big)\,. \tag{4.1.86}$$

Now,

$$\partial_u b_1(u)[h] = \partial_u\big(\mathcal{A}(u)^{-1} b_1^*(u)\big)[h] = (\partial_u \mathcal{A}(u)^{-1})(b_1^*(u))[h] + \mathcal{A}(u)^{-1}(\partial_u b_1^*(u)[h]). \tag{4.1.87}$$

Then (A.0.5), (4.1.87), (4.1.80), (4.1.83), (A.0.2) (with $a_0 = s_0 + 4$, $\beta_0 = s_0 + 6$, $p = s - 1$, $q = 1$) (4.1.86) imply

$$\begin{aligned} \|\partial_u \mathcal{A}(u)^{-1}(b_1^*(u))[h]\|_s &\le \varepsilon C(s)\big(\|h\|_{s+s_0+3} + \|u\|_{s+s_0+7}\|h\|_{s_0+3}\big) && (4.1.88)\\ \|\mathcal{A}(u)^{-1}\partial_u b_1^*(u)[h]\|_s &\le \varepsilon C(s)\big(\|h\|_{s+s_0+6} + \|u\|_{s+s_0+6}\|h\|_{s_0+6}\big). && (4.1.89)\end{aligned}$$

Finally (4.1.87), (4.1.88) and (4.1.89) imply

$$\|\partial_u b_1(u)[h]\|_s \le \varepsilon C(s)\big(\|h\|_{s+s_0+6} + \|u\|_{s+s_0+7}\|h\|_{s_0+6}\big), \tag{4.1.90}$$

which holds for all $s + s_0 + 7 \le q$.

Estimates in Step 2.

5. — We prove that the coefficient m_3, defined in (4.1.30), satisfies the following estimates:

$$\begin{aligned} |m_3 - 1|\,, |m_3 - 1|^{\mathrm{Lip}(\gamma)} &\le \varepsilon C && (4.1.91)\\ |\partial_u m_3(u)[h]| &\le \varepsilon C\|h\|_{s_0+3}. && (4.1.92)\end{aligned}$$

Using (4.1.30) (4.1.70), (4.1.62)

$$|m_3 - 1| \le \int_{\mathbb{T}^\nu} |b_3 - 1|\, d\varphi \le C\|b_3 - 1\|_{s_0} \le \varepsilon C.$$

Similarly we get the Lipschitz part of (4.1.91). The estimate (4.1.92) follows by (4.1.71), since

$$|\partial_u m_3(u)[h]\,| \le \int_{\mathbb{T}^\nu} |\partial_u b_3(u)[h]|\, d\varphi \le C\|\partial_u b_3(u)[h]\|_{s_0} \le \varepsilon C\|h\|_{s_0+3}.$$

6. — Estimates of α. The function $\alpha(\varphi)$, defined in (4.1.31), satisfies

$$\begin{aligned} |\alpha|_{s,\infty} &\le \varepsilon\gamma_0^{-1}\, C(s)\big(1 + \|u\|_{s+\tau_0+s_0+3}\big) && (4.1.93)\\ |\alpha|^{\mathrm{Lip}(\gamma)}_{s,\infty} &\le \varepsilon\gamma_0^{-1}\, C(s)\big(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+\tau_0+s_0+3}\big) && (4.1.94)\\ |\partial_u\alpha(u)[h]|_{s,\infty} &\le \varepsilon\gamma_0^{-1}\, C(s)\big(\|h\|_{s+\tau_0+s_0+3} + \|u\|_{s+\tau_0+s_0+3}\|h\|_{s_0+3}\big). && (4.1.95)\end{aligned}$$

Remember that $\omega = \lambda\bar\omega$, and $|\bar\omega \cdot l| \ge 3\gamma_0|l|^{-\tau_0}$, $\forall l \ne 0$, see (1.2.2). By (4.1.70) and (4.1.91),

$$|\alpha|_{s,\infty} \le \|\alpha\|_{s+s_0} \le C\gamma_0^{-1}\|b_3 - m_3\|_{s+s_0+\tau_0} \le C(s)\gamma_0^{-1}\varepsilon(1 + \|u\|_{s+\tau_0+s_0+3})$$

proving (4.1.93). Then (4.1.94) holds similarly using (4.1.72) and $(\omega\cdot\partial_\varphi)^{-1} = \lambda^{-1}\,(\bar\omega\cdot\partial_\varphi)^{-1}$. Differentiating formula (4.1.31) with respect to u in the direction h gives

$$\partial_u\alpha(u)[h] = (\lambda\bar\omega\cdot\partial_\varphi)^{-1}\Big(\frac{\partial_u b_3(u)[h] m_3 - b_3\partial_u m_3(u)[h]}{m_3^2}\Big)$$

then, the standard Sobolev embedding, (4.1.70), (4.1.71), (4.1.91), (4.1.92) imply (4.1.95). Estimates (4.1.94) and (4.1.95) hold for $s + \tau_0 + s_0 + 3 \le q$. Note that (4.1.23) is a well-defined diffeomorphism if $|\alpha|_{1,\infty} \le 1/2$, and, by (4.1.94), this holds by (4.1.60).

7. — Estimates of $\tilde\alpha$. Let $\vartheta \to \vartheta + \omega\tilde\alpha(\vartheta)$ be the inverse change of variable of (4.1.23). The following estimates hold:

$$
\begin{aligned}
|\tilde\alpha|_{s,\infty} &\leq \varepsilon\gamma_0^{-1} C(s)\big(1 + \|u\|_{s+\tau_0+s_0+3}\big) && (4.1.96)\\
|\tilde\alpha|_{s,\infty}^{\mathrm{Lip}(\gamma)} &\leq \varepsilon\gamma_0^{-1} C(s)\big(1 + \|u\|_{s+\tau_0+s_0+4}^{\mathrm{Lip}(\gamma)}\big) && (4.1.97)\\
|\partial_u\tilde\alpha(u)[h]|_{s,\infty} &\leq \varepsilon\gamma_0^{-1} C(s)\big(\|h\|_{s+\tau_0+s_0+3} + \|u\|_{s+\tau_0+s_0+4}\|h\|_{\tau_0+s_0+3}\big). && (4.1.98)
\end{aligned}
$$

The bounds (4.1.96), (4.1.97) follow by (A.0.14), (4.1.93), and (A.0.15), (4.1.94), respectively. To estimate the partial derivative of $\tilde\alpha$ with respect to u we differentiate the identity $\tilde\alpha(\vartheta;u) + \alpha(\vartheta + \omega\tilde\alpha(\vartheta;u);u) = 0$, which gives

$$
\partial_u\tilde\alpha(u)[h] = -B^{-1}\Big(\frac{\partial_u\alpha[h]}{1+\omega\cdot\partial_\varphi\alpha}\Big).
$$

Then applying Lemma A.0.10(iii) to deal with B^{-1}, (A.0.6) for the product $\partial_u\alpha[h]\,(1+\omega\cdot\partial_\varphi\alpha)^{-1}$, and estimates (4.1.94), (4.1.95), (A.0.2), we obtain (4.1.98).

8. — The transformations $B(u)$ and $B(u)^{-1}$, defined in (4.1.24) resp. (4.1.25), satisfy the following estimates:

$$
\begin{aligned}
\|B(u)h\|_s + \|B(u)^{-1}h\|_s &\leq C(s)\big(\|h\|_s + \|u\|_{s+\tau_0+s_0+3}\|h\|_1\big) && (4.1.99)\\
\|B(u)h\|_s^{\mathrm{Lip}(\gamma)} + \|B(u)^{-1}h\|_s^{\mathrm{Lip}(\gamma)} &\leq C(s)\big(\|h\|_{s+1}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\tau_0+s_0+4}^{\mathrm{Lip}(\gamma)}\|h\|_2^{\mathrm{Lip}(\gamma)}\big) && (4.1.100)\\
\|\partial_u(B(u)g)[h]\|_s &\leq C(s)\big(\|g\|_{s+1}\|h\|_{\sigma_0} + \|g\|_1\|h\|_{s+\sigma_0}\\
&\qquad + \|u\|_{s+\sigma_0}\|g\|_2\|h\|_{\sigma_0}\big) && (4.1.101)\\
\|\partial_u(B(u)^{-1}g)[h]\|_s &\leq C(s)\big(\|g\|_{s+1}\|h\|_{\sigma_0} + \|g\|_1\|h\|_{s+\sigma_0}\\
&\qquad + \|u\|_{s+\sigma_0+1}\|g\|_2\|h\|_{\sigma_0}\big) && (4.1.102)
\end{aligned}
$$

where $\sigma_0 := \tau_0 + s_0 + 3$. Estimates (4.1.99) and (4.1.100) follow by Lemma A.0.10(ii) and (4.1.93), (4.1.96), (4.1.94), (4.1.97). The derivative of $B(u)g$ with respect to u in the direction h is the product fz where $f := B(u)(\omega\cdot\partial_\varphi g)$ and $z := \partial_u\alpha(u)[h]$. By (A.0.7), $\|fz\|_s \leq C(s)(\|f\|_s|z|_{L^\infty} + \|f\|_0|z|_{s,\infty})$. Then (4.1.95), (4.1.99) imply (4.1.101). In analogous way, (4.1.98) and (4.1.99) give (4.1.102).

9. — estimates of ρ. The function ρ defined in (4.1.26), namely $\rho = 1 + B^{-1}(\omega\cdot\partial_\varphi\alpha)$, satisfies

$$
\begin{aligned}
|\rho - 1|_{s,\infty} &\leq \varepsilon\gamma_0^{-1} C(s)(1 + \|u\|_{s+\tau_0+s_0+4}) && (4.1.103)\\
|\rho - 1|_{s,\infty}^{\mathrm{Lip}(\gamma)} &\leq \varepsilon\gamma_0^{-1} C(s)(1 + \|u\|_{s+\tau_0+s_0+5}^{\mathrm{Lip}(\gamma)}) && (4.1.104)\\
\|\partial_u\rho(u)[h]\,\|_s &\leq \varepsilon\gamma_0^{-1} C(s)\big(\|h\|_{s+\tau_0+s_0+4} + \|u\|_{s+\tau_0+s_0+5}\|h\|_{\tau_0+s_0+4}\big). && (4.1.105)
\end{aligned}
$$

The bound (4.1.103) follows by (4.1.26), (A.0.19), (4.1.93), (4.1.60). Similarly (4.1.104) follows by (A.0.20), (4.1.94) and (4.1.62). Differentiating (4.1.26) with respect to u in the direction h we obtain

$$
\partial_u\rho(u)[h] \;=\; \partial_u B(u)^{-1}(\omega\cdot\partial_\varphi\alpha)[h] + B(u)^{-1}\big(\omega\cdot\partial_\varphi(\partial_u\alpha(u)[h])\big).
$$

By (4.1.102), (4.1.93), and (4.1.60), we get

$$\|\partial_u B(u)^{-1}(\omega\cdot\partial_\varphi\alpha)[h]\|_s \le \varepsilon\gamma_0^{-1}C(s)\big(\|h\|_{s+\tau_0+s_0+3}+\|u\|_{s+\tau_0+s_0+5}\|h\|_{\tau_0+s_0+3}\big). \tag{4.1.106}$$

Using (4.1.99), (4.1.95), (4.1.60), and applying (A.0.2), one has

$$\|B(u)^{-1}\big(\omega\cdot\partial_\varphi(\partial_u\alpha(u)[h])\big)\|_s \le \varepsilon\gamma_0^{-1}C(s)\big(\|h\|_{s+\tau_0+s_0+4}+\|u\|_{s+\tau_0+s_0+4}\|h\|_{\tau_0+s_0+4}\big)\,. \tag{4.1.107}$$

Then (4.1.106) and (4.1.107) imply (4.1.105), for all $s+\tau_0+s_0+5\le q$.

10. — The coefficients c_0, c_1, c_2 defined in (4.1.33) satisfy the following estimates: for $i=0,1,2$, $s\ge s_0$,

$$\begin{aligned}
\|c_i\|_s &\le \varepsilon C(s)\big(1+\|u\|_{s+\tau_0+s_0+6}\big), && (4.1.108)\\
\|c_i\|_s^{\mathrm{Lip}(\gamma)} &\le \varepsilon C(s)\big(1+\|u\|^{\mathrm{Lip}(\gamma)}_{s+\tau_0+s_0+7}\big), && (4.1.109)\\
\|\partial_u c_i[h]\|_s &\le \varepsilon C(s)\big(\|h\|_{s+\tau_0+s_0+6}+\|u\|_{s+\tau_0+s_0+7}\|h\|_{\tau_0+2s_0+6}\big)\,. && (4.1.110)
\end{aligned}$$

The definition of c_i in (4.1.33), (A.0.7), (4.1.60), (4.1.99), (4.1.103), (4.1.84) and $\varepsilon\gamma_0^{-1}<1$, imply (4.1.108). Similarly (4.1.62), (4.1.100), (4.1.104) and (4.1.85) imply (4.1.109). Finally (4.1.110) follows from differentiating the formula of $c_i(u)$ and using (4.1.60), (4.1.84), (4.1.102), (4.1.99), (A.0.5)-(A.0.7), (4.1.103), (4.1.105).

Estimates in the step 3.

11. — The function v defined in (4.1.38) satisfies the following estimates:

$$\begin{aligned}
\|v-1\|_s &\le \varepsilon C(s)\big(1+\|u\|_{s+\tau_0+s_0+6}\big) && (4.1.111)\\
\|v-1\|_s^{\mathrm{Lip}(\gamma)} &\le \varepsilon C(s)\big(1+\|u\|^{\mathrm{Lip}(\gamma)}_{s+\tau_0+s_0+7}\big) && (4.1.112)\\
\|\partial_u v[h]\|_s &\le \varepsilon C(s)\big(\|h\|_{s+\tau_0+s_0+6}+\|u\|_{s+\tau_0+s_0+7}\|h\|_{\tau_0+2s_0+6}\big) && (4.1.113)
\end{aligned}$$

In order to prove (4.1.111) we apply the Lemma A.0.8(i) with $f(t):=\exp(t)$ (and $u=0$, $p=0$):

$$\|v-1\|_s=\Big\|f\Big(-\frac{\partial_y^{-1}c_2}{3m_3}\Big)-f(0)\Big\|_s \overset{(4.1.91)}{\le} C\|c_2\|_s \overset{(4.1.108)}{\le} \varepsilon C(s)\big(1+\|u\|_{s+\tau_0+s_0+6}\big)\,.$$

Similarly (4.1.112) follows. Differentiating formula (4.1.38) we get

$$\partial_u v[h]=-f'\Big(-\frac{\partial_y^{-1}c_2}{3m_3}\Big)\left\{\frac{1}{3m_3}\partial_u\Big(\partial_y^{-1}c_2\Big)[h]-\frac{\partial_y^{-1}c_2\partial_u m_3[h]}{3m_3^2}\right\}.$$

Then using (4.1.60), (A.0.5), Lemma A.0.8(i) applied to $f'=f$, and the estimates (4.1.108), (4.1.110), (4.1.91) and (4.1.92) we get (4.1.113).

12. — The multiplication operator $\mathcal{M}$ defined in (4.1.34) and its inverse $\mathcal{M}^{-1}$ (which is the multiplication operator by v^{-1}) both satisfy

$$\begin{aligned}
\|\mathcal{M}^{\pm1}h\|_s &\le C(s)\big(\|h\|_s+\|u\|_{s+\tilde\sigma}\|h\|_{s_0}\big), && (4.1.114)\\
\|\mathcal{M}^{\pm1}h\|_s^{\mathrm{Lip}(\gamma)} &\le C(s)\big(\|h\|_s^{\mathrm{Lip}(\gamma)}+\|u\|^{\mathrm{Lip}(\gamma)}_{s+\tilde\sigma+1}\|h\|^{\mathrm{Lip}(\gamma)}_{s_0}\big), && (4.1.115)\\
\|\partial_u\mathcal{M}^{\pm1}(u)g[h]\|_s &\le \varepsilon C(s)\big(\|g\|_s\|h\|_{s_0+\tilde\sigma}+\|g\|_{s_0}\|h\|_{s+\tilde\sigma}+\|u\|_{s+\tilde\sigma+1}\|g\|_{s_0}\|h\|_{s_0+\tilde\sigma}\big), && (4.1.116)
\end{aligned}$$

with $\tilde\sigma := \tau_0 + s_0 + 6$.

The inequalities (4.1.114)-(4.1.116) follow by (4.1.60), (4.1.62), (A.0.5), (4.1.111)-(4.1.113).

13. — The coefficients d_1, d_0, defined in (4.1.41), satisfy, for $i = 0, 1$

$$\|d_i\|_s \leq \varepsilon C(s)(1 + \|u\|_{s+\tau_0+s_0+9}), \tag{4.1.117}$$
$$\|d_i\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon C(s)(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+\tau_0+s_0+10}), \tag{4.1.118}$$
$$\|\partial_u d_i(u)[h]\|_s \leq \varepsilon C(s)\big(\|h\|_{s+\tau_0+s_0+9} + \|u\|_{s+\tau_0+s_0+10}\|h\|_{\tau_0+2s_0+9}\big), \tag{4.1.119}$$

by (A.0.5), (4.1.60), (4.1.62), (4.1.108)-(4.1.110) and (4.1.111)-(4.1.113).

Estimates in the Step 4.

14. — The constant m_1 defined in (4.1.46) satisfies

$$|m_1| + |m_1|^{\mathrm{Lip}(\gamma)} \leq \varepsilon C, \quad |\partial_u m_1(u)[h]| \leq \varepsilon C\|h\|_{\tau_0+2s_0+9}\,, \tag{4.1.120}$$

by (4.1.62), (4.1.117)-(4.1.119).

15. — The function $p(\vartheta)$ defined in (4.1.47) satisfies the following estimates:

$$|p|_{s,\infty} \leq \varepsilon\gamma_0^{-1}C(s)(1 + \|u\|_{s+2\tau_0+2s_0+9}) \tag{4.1.121}$$
$$|p|^{\mathrm{Lip}(\gamma)}_{s,\infty} \leq \varepsilon\gamma_0^{-1}C(s)(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+2\tau_0+2s_0+10}) \tag{4.1.122}$$
$$|\partial_u p(u)[h]|_{s,\infty} \leq \varepsilon\gamma_0^{-1}C(s)\big(\|h\|_{s+2\tau_0+2s_0+9} + \|u\|_{s+2\tau_0+2s_0+10}\|h\|_{\tau_0+2s_0+9}\big). \tag{4.1.123}$$

which follow by (4.1.117)-(4.1.119) and (4.1.120) applying the same argument used in the proof of (4.1.94).

16. — The operators $\mathcal{T}$, $\mathcal{T}^{-1}$ defined in (4.1.42) satisfy

$$\|\mathcal{T}^{\pm1}h\|_s \leq C(s)\big(\|h\|_s + \|u\|_{s+\bar\sigma}\|h\|_1\big) \tag{4.1.124}$$
$$\|\mathcal{T}^{\pm1}h\|_s^{\mathrm{Lip}(\gamma)} \leq C(s)\big(\|h\|^{\mathrm{Lip}(\gamma)}_{s+1} + \|u\|^{\mathrm{Lip}(\gamma)}_{s+\bar\sigma+1}\|h\|_2^{\mathrm{Lip}(\gamma)}\big) \tag{4.1.125}$$
$$\begin{aligned}\|\partial_u(\mathcal{T}^{\pm1}(u)g)[h]\|_s \leq\ & \varepsilon\gamma_0^{-1}C(s)\big(\|g\|_{s+1}\|h\|_{\bar\sigma} + \|g\|_1\|h\|_{s+\bar\sigma} \\ & + \|u\|_{s+\bar\sigma+1}\|g\|_2\|h\|_{\bar\sigma}\big),\end{aligned} \tag{4.1.126}$$

with $\bar\sigma := 2\tau_0 + 2s_0 + 9$. The estimates (4.1.124) and (4.1.125) follow by (A.0.16), (A.0.18) and using (5.6.74) and (4.1.122). The derivative $\partial_u(\mathcal{T}(u)g)[h]$ is the product $(\mathcal{T}(u)g_y)\,\partial_u p(u)[h]$. Hence (A.0.7), (4.1.124) and (4.1.123) imply (4.1.126).

17. — The coefficients e_0, e_1, defined in (4.1.43), satisfy the following estimates: for $i = 0, 1$

$$\|e_i\|_s \leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2s_0+9}), \tag{4.1.127}$$
$$\|e_i\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon C(s)(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+2\tau_0+2s_0+10}), \tag{4.1.128}$$
$$\|\partial_u e_i(u)[h]\|_s \leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2s_0+9} + \|u\|_{s+2\tau_0+2s_0+10}\|h\|_{2\tau_0+2s_0+9}\big)\,. \tag{4.1.129}$$

The estimates (4.1.127), (4.1.128) follow by (4.1.60), (4.1.62), (4.1.45), (4.1.117), (4.1.118), (4.1.124) and (4.1.125). The estimate (4.1.129) follows differentiating the formulae of e_0 and e_1 in (4.1.43), and applying (4.1.117), (4.1.119), (4.1.124) and (4.1.126).

Estimates in the Step 5.

18. — The function w defined in (4.1.55) satisfies the following estimates:

$$\begin{aligned} \|w\|_s &\leq \varepsilon C(s)(1+\|u\|_{s+2\tau_0+2s_0+9}) && (4.1.130)\\ \|w\|_s^{\mathrm{Lip}(\gamma)} &\leq \varepsilon C(s)(1+\|u\|^{\mathrm{Lip}(\gamma)}_{s+2\tau_0+2s_0+10}) && (4.1.131)\\ \|\partial_u w(u)[h]\|_s &\leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2s_0+9}+\|u\|_{s+2\tau_0+2s_0+10}\|h\|_{2\tau_0+2s_0+9}\big) && (4.1.132)\end{aligned}$$

which follow by (4.1.91), (4.1.92), (4.1.120), (4.1.127)-(4.1.129), (4.1.60), (4.1.62).

19. — The operator $\mathcal{S}=I+w\partial_x^{-1}$, defined in (4.1.49), and its inverse $\mathcal{S}^{-1}$ both satisfy the following estimates (where the s-decay norm $|\cdot|_s$ is defined in (3.1.1)):

$$\begin{aligned} |\mathcal{S}^{\pm1}-I|_s &\leq \varepsilon C(s)(1+\|u\|_{s+2\tau_0+2s_0+9}), && (4.1.133)\\ |\mathcal{S}^{\pm1}-I|_s^{\mathrm{Lip}(\gamma)} &\leq \varepsilon C(s)(1+\|u\|^{\mathrm{Lip}(\gamma)}_{s+2\tau_0+2s_0+10}), && (4.1.134)\\ \big|\partial_u\mathcal{S}^{\pm1}(u)[h]\big|_s &\leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2s_0+9}+\|u\|_{s+2\tau_0+2s_0+10}\|h\|_{2\tau_0+3s_0+9}\big). && (4.1.135)\end{aligned}$$

Thus (4.1.133)-(4.1.135) for $\mathcal{S}$ follow by (4.1.130)-(4.1.132) and the fact that the matrix decay norm $|\partial_x^{-1}|_s\leq 1$, $s\geq 0$, using (3.1.4), (3.1.5), (3.1.7), (3.1.8). The operator $\mathcal{S}^{-1}$ satisfies the same bounds (4.1.133)-(4.1.134) by Lemma 3.1.3, which may be applied thanks to (4.1.133), (4.1.60), (4.1.62) and ε small enough.

Finally (4.1.135) for $\mathcal{S}^{-1}$ follows by

$$\partial_u\mathcal{S}^{-1}(u)[h]=-\mathcal{S}^{-1}(u)\,\partial_u\mathcal{S}(u)[h]\,\mathcal{S}^{-1}(u)\,,$$

and (3.1.6), (4.1.133) for $\mathcal{S}^{-1}$, and (4.1.135) for $\mathcal{S}$.

20. — The operatpr $\mathcal{R}$, defined in (4.1.56) , where r_0, r_{-1} are defined in (4.1.53), (4.1.54), satisfies the following estimates:

$$\begin{aligned} \big|\mathcal{R}\big|_s &\leq \varepsilon C(s)(1+\|u\|_{s+2\tau_0+2s_0+12}) && (4.1.136)\\ \big|\mathcal{R}\big|_s^{\mathrm{Lip}(\gamma)} &\leq \varepsilon C(s)(1+\|u\|^{\mathrm{Lip}(\gamma)}_{s+2\tau_0+2s_0+13}) && (4.1.137)\\ \big|\partial_u\mathcal{R}(u)[h]\big|_s &\leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2s_0+12}+\|u\|_{s+2\tau_0+2s_0+13}\|h\|_{2\tau_0+3s_0+12}\big). && (4.1.138)\end{aligned}$$

Let $T:=r_0+r_{-1}\partial_x^{-1}$. By (3.1.4), (3.1.5), (A.0.5), (4.1.130), (4.1.131), (4.1.127), (4.1.128), (4.1.120), (4.1.91), and using the trivial fact that $|\partial_x^{-1}|_s\leq 1$ and $|\pi_0|_s\leq 1$ for all $s\geq 0$, we get

$$\begin{aligned} \big|T\big|_s &\leq \varepsilon C(s)(1+\|u\|_{s+2\tau_0+2s_0+12}) && (4.1.139)\\ \big|T\big|_s^{\mathrm{Lip}(\gamma)} &\leq \varepsilon C(s)(1+\|u\|^{\mathrm{Lip}(\gamma)}_{s+2\tau_0+2s_0+13}). && (4.1.140)\end{aligned}$$

Differentiating T with respect to u, and using (3.1.4), (A.0.5), (4.1.132), (4.1.129), (4.1.120), (4.1.91) and (4.1.92), one has

$$\big|\partial_u T(u)[h]\big|_s\leq\varepsilon C(s)\big(\|h\|_{s+2\tau_0+2s_0+12}+\|u\|_{s+2\tau_0+2s_0+13}\|h\|_{2\tau_0+3s_0+12}\big). \tag{4.1.141}$$

Finally (3.1.6), (3.1.9) (4.1.133)-(4.1.135), (4.1.139)-(4.1.141) imply the estimates (4.1.136)-(4.1.138).

21. — Using Lemma A.0.11, (4.1.60) and all the previous estimates on $\mathcal{A}, B, \rho, \mathcal{M}, \mathcal{T}, \mathcal{S}$, the operators $\Phi_1 = \mathcal{A}B\rho\mathcal{M}\mathcal{T}\mathcal{S}$ and $\Phi_2 = \mathcal{A}B\mathcal{M}\mathcal{T}\mathcal{S}$, defined in (4.1.58), satisfy (4.1.61) (note that $\sigma > 2\tau_0 + 2s_0 + 9$). Finally, if the condition (4.1.62) holds, we get the estimate (4.1.63).

The other estimates (4.1.64)-(4.1.69) follow by (4.1.91), (4.1.92), (4.1.120), (4.1.136)-(4.1.138). The proof of the Lemma is complete.

□

In the same way we get the following lemma.

Lemma 4.1.3. *In the same hypotheses of Lemma 4.1.2, for all $\varphi \in \mathbb{T}^\nu$, the operators $\mathcal{A}(\varphi)$, $\mathcal{M}(\varphi)$, $\mathcal{T}(\varphi)$, $\mathcal{S}(\varphi)$ are invertible operators of the phase space $H^s_x := H^s(\mathbb{T})$, with*

$$\|\mathcal{A}^{\pm1}(\varphi)h\|_{H^s_x} \le C(s)\big(\|h\|_{H^s_x} + \|u\|_{s+s_0+3}\|h\|_{H^1_x}\big), \tag{4.1.142}$$

$$\|(\mathcal{A}^{\pm1}(\varphi) - I)h\|_{H^s_x} \le \varepsilon C(s)\big(\|h\|_{H^{s+1}_x} + \|u\|_{s+s_0+3}\|h\|_{H^2_x}\big), \tag{4.1.143}$$

$$\|(\mathcal{M}(\varphi)\mathcal{T}(\varphi)\mathcal{S}(\varphi))^{\pm1}h\|_{H^s_x} \le C(s)\big(\|h\|_{H^s_x} + \|u\|_{s+\sigma}\|h\|_{H^1_x}\big), \tag{4.1.144}$$

$$\|((\mathcal{M}(\varphi)\mathcal{T}(\varphi)\mathcal{S}(\varphi))^{\pm1} - I)h\|_{H^s_x} \le \varepsilon\gamma_0^{-1}C(s)\big(\|h\|_{H^{s+1}_x} + \|u\|_{s+\sigma}\|h\|_{H^1_x}\big). \tag{4.1.145}$$

Proof. For each fixed $\varphi \in \mathbb{T}^\nu$, $\mathcal{A}(\varphi)h(x) := h(x+\beta(\varphi, x))$. Apply (A.0.16) to the change of variable $\mathbb{T} \to \mathbb{T}$, $x \mapsto x + \beta(\varphi, x)$:

$$\|\mathcal{A}(\varphi)h\|_{H^s_x} \le C(s)\big(\|h\|_{H^s_x} + |\beta(\varphi,\cdot)|_{W^{s,\infty}(\mathbb{T})}\|h\|_{H^1_x}\big).$$

Since $|\beta(\varphi,\cdot)|_{W^{s,\infty}(\mathbb{T})} \le |\beta|_{s,\infty}$ for all $\varphi \in \mathbb{T}^\nu$, by (5.6.35) we deduce (4.1.142). Using (A.0.17), (4.1.60), and (5.6.35),

$$\|(\mathcal{A}(\varphi) - I)h\|_{H^s_x} \le_s |\beta|_{L^\infty}\|h\|_{H^{s+1}_x} + |\beta|_{s,\infty}\|h\|_{H^2_x} \le_s \varepsilon\big(\|h\|_{H^{s+1}_x} + \|u\|_{s+s_0+3}\|h\|_{H^2_x}\big).$$

By (4.1.77), estimates (4.1.142) and (4.1.143) also hold for

$$\mathcal{A}(\varphi)^{-1} = \mathcal{A}^{-1}(\varphi) : h(y) \mapsto h(y + \tilde\beta(\varphi, y)).$$

The multiplication operator $\mathcal{M}(\varphi) : H^s_x \to H^s_x$, $\mathcal{M}(\varphi)h := v(\varphi,\cdot)h$ satisfies

$$\begin{aligned}\|(\mathcal{M}(\varphi) - I)h\|_{H^s_x} &= \|(v(\varphi,\cdot) - 1)h\|_{H^s_x} \le_s \|v(\varphi,\cdot) - 1\|_{H^s_x}\|h\|_{H^1_x} + \|v(\varphi,\cdot) - 1\|_{H^1_x}\|h\|_{H^s_x} \\ &\le_s \|v - 1\|_{s+s_0}\|h\|_{H^1_x} + \|v - 1\|_{1+s_0}\|h\|_{H^s_x} \le_s \varepsilon\big(\|h\|_{H^s_x} + \|u\|_{s+\tau_0+2s_0+6}\|h\|_{H^1_x}\big)\end{aligned} \tag{4.1.146}$$

by (A.0.5), (3.1.4), Lemma 3.1.4, (4.1.111) and (4.1.60). The same estimate also holds for $\mathcal{M}(\varphi)^{-1} = \mathcal{M}^{-1}(\varphi)$, which is the multiplication operator by $v^{-1}(\varphi,\cdot)$. The operators $\mathcal{T}^{\pm1}(\varphi)h(x) = h(x \pm p(\varphi))$ satisfy

$$\|\mathcal{T}^{\pm1}(\varphi)h\|_{H^s_x} = \|h\|_{H^s_x}, \quad \|(\mathcal{T}^{\pm1}(\varphi) - I)h\|_{H^s_x} \le \varepsilon\gamma_0^{-1}C\|h\|_{H^{s+1}_x}, \tag{4.1.147}$$

by (A.0.17), (4.1.60), (5.6.74) and by the fact that $p(\varphi)$ is independent on the space variable.

By (3.1.11), (4.1.133), (4.1.60) and Lemma 3.1.4, the operator $\mathcal{S}(\varphi) = I + w(\varphi,\cdot)\partial_x^{-1}$ and its inverse satisfy

$$\|(\mathcal{S}^{\pm1}(\varphi) - I)h\|_{H^s_x} \le_s \varepsilon\big(\|h\|_{H^s_x} + \|u\|_{s+2\tau_0+3s_0+9}\|h\|_{H^1_x}\big). \tag{4.1.148}$$

Collecting estimates (4.1.146), (4.1.147), (4.1.148) we get (4.1.144) and (4.1.145) and the proof is concluded.

□

4.2 Reduction of the linearized operator to constant coefficients

The goal of this Section is to diagonalize the linear operator $\mathcal{L}_5$ obtained in (4.1.56), and therefore to complete the reduction of $\mathcal{L}$ in (4.1.1) into constant coefficients. For $\tau > \tau_0$ (see (1.2.2)) we define the constant

$$\beta := 7\tau + 6 \,. \tag{4.2.1}$$

Theorem 4.2.1. *Let $f \in C^q$, see (1.2.3). Let $\gamma \in (0,1)$ and $s_0 \le s \le q - \sigma - \beta$ where σ is defined in (4.1.59), and β in (4.2.1). Let $u(\lambda)$ be a family of functions depending on the parameter $\lambda \in \Lambda_o \subset \Lambda := [1/2, 3/2]$ in a Lipschitz way, with*

$$\|u\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma+\beta,\Lambda_o} \le 1. \tag{4.2.2}$$

Then there exist δ_0, C (depending on the data of the problem) such that, if

$$\varepsilon\gamma^{-1} \le \delta_0 \,, \tag{4.2.3}$$

then:

(i) **(Eigenvalues)** $\forall \lambda \in \Lambda$ *there exists a sequence*

$$\mu_j^\infty(\lambda) := \mu_j^\infty(\lambda, u) = \tilde\mu_j^0(\lambda) + r_j^\infty(\lambda) \,, \ \tilde\mu_j^0(\lambda) := \mathrm{i}\big(-\tilde m_3(\lambda) j^3 + \tilde m_1(\lambda) j\big) \,, \ j \in \mathbb{Z} \,, \tag{4.2.4}$$

where $\tilde m_3, \tilde m_1$ coincide with the coefficients of $\mathcal{L}_5$ in (4.1.56) for all $\lambda \in \Lambda_o$, and the corrections r_j^∞ satisfy

$$|\tilde m_3 - 1|^{\mathrm{Lip}(\gamma)} + |\tilde m_1|^{\mathrm{Lip}(\gamma)} + |r_j^\infty|^{\mathrm{Lip}(\gamma)}_\Lambda \le \varepsilon C \,, \quad \forall j \in \mathbb{Z} \,. \tag{4.2.5}$$

Moreover, in the reversible case (i.e. (1.2.15) holds) or Hamiltonian case (i.e. (1.2.11) holds), all the eigenvalues μ_j^∞ are purely imaginary.

(ii) **(Conjugacy)**. *For all λ in*

$$\Lambda_\infty^{2\gamma} := \Lambda_\infty^{2\gamma}(u) := \Big\{ \lambda \in \Lambda_o \, : \, |\mathrm{i}\lambda\bar\omega \cdot l + \mu_j^\infty(\lambda) - \mu_k^\infty(\lambda)| \ge 2\gamma |j^3 - k^3| \langle l \rangle^{-\tau}, \quad \forall l \in \mathbb{Z}^\nu, \, j,k \in \mathbb{Z} \Big\} \tag{4.2.6}$$

there is a bounded, invertible linear operator $\Phi_\infty(\lambda) : H^s \to H^s$, with bounded inverse $\Phi_\infty^{-1}(\lambda)$, that conjugates $\mathcal{L}_5$ in (4.1.56) to constant coefficients, namely

$$\mathcal{L}_\infty(\lambda) := \Phi_\infty^{-1}(\lambda) \circ \mathcal{L}_5(\lambda) \circ \Phi_\infty(\lambda) = \lambda\bar\omega \cdot \partial_\varphi + \mathcal{D}_\infty(\lambda), \quad \mathcal{D}_\infty(\lambda) := \mathrm{diag}_{j\in\mathbb{Z}} \mu_j^\infty(\lambda) \,. \tag{4.2.7}$$

The transformations $\Phi_\infty, \Phi_\infty^{-1}$ are close to the identity in matrix decay norm, with estimates

$$|\Phi_\infty(\lambda) - I|^{\mathrm{Lip}(\gamma)}_{s,\Lambda_\infty^{2\gamma}} + |\Phi_\infty^{-1}(\lambda) - I|^{\mathrm{Lip}(\gamma)}_{s,\Lambda_\infty^{2\gamma}} \le \varepsilon\gamma^{-1} C(s) \big(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+\sigma+\beta,\Lambda_o}\big). \tag{4.2.8}$$

For all $\varphi \in \mathbb{T}^\nu$, the operator $\Phi_\infty(\varphi) : H^s_x \to H^s_x$ is invertible (where $H^s_x := H^s(\mathbb{T})$) with inverse $(\Phi_\infty(\varphi))^{-1} = \Phi_\infty^{-1}(\varphi)$, and

$$\|(\Phi_\infty^{\pm 1}(\varphi) - I)h\|_{H^s_x} \le \varepsilon\gamma^{-1} C(s) \big(\|h\|_{H^s_x} + \|u\|_{s+\sigma+\beta+s_0} \|h\|_{H^1_x} \big). \tag{4.2.9}$$

In the reversible case $\Phi_\infty, \Phi_\infty^{-1} : X \to X$, $Y \to Y$ are reversibility preserving, and $\mathcal{L}_\infty : X \to Y$ is reversible. In the Hamiltonian case the final $\mathcal{L}_\infty$ is Hamiltonian.

An important point of Theorem 4.2.1 is to require *only* the bound (4.2.2) for the low norm of u, but it provides the estimate for $\Phi_\infty^{\pm 1} - I$ in (4.2.8) also for the higher norms $|\cdot|_s$, depending also on the high norms of u. From Theorem 4.2.1 we shall deduce tame estimates for the inverse linearized operators in Theorem 4.2.3.

Note also that the set $\Lambda_\infty^{2\gamma}$ in (4.2.6) depends only of the final eigenvalues, and it is not defined inductively as in usual KAM theorems. This characterization of the set of parameters which fulfill all the required Melnikov non-resonance conditions (at any step of the iteration) was first observed in [17], [14] in an analytic setting. Theorem 4.2.1 extends this property also in a differentiable setting. A main advantage of this formulation is that it allows to discuss the measure estimates only once and not inductively: the Cantor set $\Lambda_\infty^{2\gamma}$ in (4.2.6) could be empty (actually its measure $|\Lambda_\infty^{2\gamma}| = 1 - O(\gamma)$ as $\gamma \to 0$) but the functions $\mu_j^\infty(\lambda)$ are anyway well defined for all $\lambda \in \Lambda$, see (4.2.4). In particular we shall perform the measure estimates only along the nonlinear iteration, see Section 4.3.

Theorem 4.2.1 is deduced from the following iterative Nash-Moser reducibility theorem for a linear operator of the form

$$\mathcal{L}_0 = \omega \cdot \partial_\varphi + \mathcal{D}_0 + \mathcal{R}_0 \,, \tag{4.2.10}$$

where $\omega = \lambda \bar{\omega}$,

$$\mathcal{D}_0 := m_3(\lambda, u(\lambda))\partial_{xxx} + m_1(\lambda, u(\lambda))\partial_x \,, \quad \mathcal{R}_0(\lambda, u(\lambda)) := \mathcal{R}(\lambda, u(\lambda)) \,, \tag{4.2.11}$$

the $m_3(\lambda, u(\lambda)), m_1(\lambda, u(\lambda)) \in \mathbb{R}$ and $u(\lambda)$ is defined for $\lambda \in \Lambda_o \subset \Lambda$. Clearly $\mathcal{L}_5$ in (4.1.56) has the form (4.2.10). Define

$$N_{-1} := 1 \,, \quad N_\nu := N_0^{\chi^\nu} \ \forall \nu \geq 0 \,, \quad \chi := 3/2 \tag{4.2.12}$$

(then $N_{\nu+1} = N_\nu^\chi$, $\forall \nu \geq 0$) and

$$\alpha := 7\tau + 4, \quad \sigma_2 := \sigma + \beta \tag{4.2.13}$$

where σ is defined in (4.1.59) and β is defined in (4.2.1).

Theorem 4.2.2. (KAM reducibility) *Let $q > \sigma + s_0 + \beta$. There exist $C_0 > 0$, $N_0 \in \mathbb{N}$ large, such that, if*

$$N_0^{C_0} |\mathcal{R}_0|_{s_0+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1} \leq 1, \tag{4.2.14}$$

then, for all $\nu \geq 0$:

$(\mathbf{S1})_\nu$ *There exists an operator*

$$\mathcal{L}_\nu := \omega \cdot \partial_\varphi + \mathcal{D}_\nu + \mathcal{R}_\nu \quad \textit{where} \quad \mathcal{D}_\nu = \mathrm{diag}_{j \in \mathbb{Z}} \{\mu_j^\nu(\lambda)\} \tag{4.2.15}$$

$$\mu_j^\nu(\lambda) = \mu_j^0(\lambda) + r_j^\nu(\lambda), \quad \mu_j^0(\lambda) := -\mathrm{i}\big(m_3(\lambda, u(\lambda)) j^3 - m_1(\lambda, u(\lambda)) j\big), \quad j \in \mathbb{Z} \,, \tag{4.2.16}$$

defined for all $\lambda \in \Lambda_\nu^\gamma(u)$, where $\Lambda_0^\gamma(u) := \Lambda_o$ (is the domain of u), and, for $\nu \geq 1$,

$$\Lambda_\nu^\gamma := \Lambda_\nu^\gamma(u) := \Big\{ \lambda \in \Lambda_{\nu-1}^\gamma : \Big| \mathrm{i}\omega \cdot l + \mu_j^{\nu-1}(\lambda) - \mu_k^{\nu-1}(\lambda) \Big| \geq \gamma \frac{|j^3 - k^3|}{\langle l \rangle^\tau} \quad \forall \, |l| \leq N_{\nu-1}, \ j, k \in \mathbb{Z} \Big\}. \tag{4.2.17}$$

For $\nu \geq 0$, $r_j^\nu = \overline{r_{-j}^\nu}$, equivalently $\mu_j^\nu = \overline{\mu_{-j}^\nu}$, and

$$|r_j^\nu|^{\mathrm{Lip}(\gamma)} := |r_j^\nu|^{\mathrm{Lip}(\gamma)}_{\Lambda_\nu^\gamma} \leq \varepsilon C\,. \tag{4.2.18}$$

The remainder $\mathcal{R}_\nu$ is real (Definition 3.2.1) and, $\forall s \in [s_0, q - \sigma - \beta]$,

$$|\mathcal{R}_\nu|_s^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_0|_{s+\beta}^{\mathrm{Lip}(\gamma)} N_{\nu-1}^{-\alpha}\,, \quad |\mathcal{R}_\nu|_{s+\beta}^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_0|_{s+\beta}^{\mathrm{Lip}(\gamma)} N_{\nu-1}\,. \tag{4.2.19}$$

Moreover, for $\nu \geq 1$,

$$\mathcal{L}_\nu = \Phi_{\nu-1}^{-1} \mathcal{L}_{\nu-1} \Phi_{\nu-1}\,, \quad \Phi_{\nu-1} := I + \Psi_{\nu-1}\,, \tag{4.2.20}$$

where the map $\Psi_{\nu-1}$ is real, Töplitz in time $\Psi_{\nu-1} := \Psi_{\nu-1}(\varphi)$ (see (3.1.16)*), and satisfies*

$$|\Psi_{\nu-1}|_s^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_0|_{s+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1} N_{\nu-1}^{2\tau+1} N_{\nu-2}^{-\alpha}\,. \tag{4.2.21}$$

In the reversible case, $\mathcal{R}_\nu : X \to Y$, $\Psi_{\nu-1}, \Phi_{\nu-1}, \Phi_{\nu-1}^{-1}$ are reversibility preserving. Moreover, all the $\mu_j^\nu(\lambda)$ are purely imaginary and $\mu_j^\nu = -\mu_{-j}^\nu$, $\forall j \in \mathbb{Z}$.

(S2)$_\nu$ *For all $j \in \mathbb{Z}$, there exist Lipschitz extensions $\widetilde{\mu}_j^\nu(\cdot) : \Lambda \to \mathbb{R}$ of $\mu_j^\nu(\cdot) : \Lambda_\nu^\gamma \to \mathbb{R}$ satisfying, for $\nu \geq 1$,*

$$|\widetilde{\mu}_j^\nu - \widetilde{\mu}_j^{\nu-1}|^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_{\nu-1}|_{s_0}^{\mathrm{Lip}(\gamma)}\,. \tag{4.2.22}$$

(S3)$_\nu$ *Let $u_1(\lambda)$, $u_2(\lambda)$, be Lipschitz families of Sobolev functions, defined for $\lambda \in \Lambda_o$ and such that conditions* (4.2.2), (4.2.14) *hold with $\mathcal{R}_0 := \mathcal{R}_0(u_i)$, $i = 1, 2$, see* (4.2.11)*.*

Then, for $\nu \geq 0$, $\forall \lambda \in \Lambda_\nu^{\gamma_1}(u_1) \cap \Lambda_\nu^{\gamma_2}(u_2)$, with $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$,

$$|\mathcal{R}_\nu(u_2) - \mathcal{R}_\nu(u_1)|_{s_0} \leq \varepsilon N_{\nu-1}^{-\alpha} \|u_1 - u_2\|_{s_0+\sigma_2}\,, \ |\mathcal{R}_\nu(u_2) - \mathcal{R}_\nu(u_1)|_{s_0+\beta} \leq \varepsilon N_{\nu-1} \|u_1 - u_2\|_{s_0+\sigma_2}\,. \tag{4.2.23}$$

Moreover, for $\nu \geq 1$, $\forall s \in [s_0, s_0 + \beta]$, $\forall j \in \mathbb{Z}$,

$$\big|\big(r_j^\nu(u_2) - r_j^\nu(u_1)\big) - \big(r_j^{\nu-1}(u_2) - r_j^{\nu-1}(u_1)\big)\big| \leq |\mathcal{R}_{\nu-1}(u_2) - \mathcal{R}_{\nu-1}(u_1)|_{s_0}\,, \tag{4.2.24}$$

$$|r_j^\nu(u_2) - r_j^\nu(u_1)| \leq \varepsilon C \|u_1 - u_2\|_{s_0+\sigma_2}\,. \tag{4.2.25}$$

(S4)$_\nu$ *Let u_1, u_2 like in* **(S3)**$_\nu$ *and $0 < \rho < \gamma/2$. For all $\nu \geq 0$ such that*

$$\varepsilon C N_{\nu-1}^{\tau} \|u_1 - u_2\|_{s_0+\sigma_2}^{\sup} \leq \rho \quad \Longrightarrow \quad \Lambda_\nu^\gamma(u_1) \subseteq \Lambda_\nu^{\gamma-\rho}(u_2)\,. \tag{4.2.26}$$

Remark 4.2.1. *In the Hamiltonian case $\Psi_{\nu-1}$ is Hamiltonian and, instead of* (4.2.20) *we consider the symplectic map*

$$\Phi_{\nu-1} := \exp(\Psi_{\nu-1})\,. \tag{4.2.27}$$

The corresponding operators $\mathcal{L}_\nu$, $\mathcal{R}_\nu$ are Hamiltonian. Note that the operators (4.2.27) *and* (4.2.20) *differ for an operator of order $\Psi_{\nu-1}^2$.*

The proof of Theorem 4.2.2 is postponed in Subsection 4.2.1. We first give some consequences.

Corollary 4.2.1. (KAM transformation) *$\forall \lambda \in \cap_{\nu\geq 0}\Lambda_\nu^\gamma$ the sequence*

$$\widetilde{\Phi}_\nu := \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu \tag{4.2.28}$$

converges in $|\cdot|_s^{\mathrm{Lip}(\gamma)}$ to an operator Φ_∞ and

$$|\Phi_\infty - I|_s^{\mathrm{Lip}(\gamma)} + \left|\Phi_\infty^{-1} - I\right|_s^{\mathrm{Lip}(\gamma)} \leq C(s)\, |\mathcal{R}_0|_{s+\beta}^{\mathrm{Lip}(\gamma)}\, \gamma^{-1}\,. \tag{4.2.29}$$

In the reversible case Φ_∞ and Φ_∞^{-1} are reversibility preserving.

Proof. To simplify notations we write $|\cdot|_s$ for $|\cdot|_s^{\mathrm{Lip}(\gamma)}$. For all $\nu \geq 0$ we have $\widetilde{\Phi}_{\nu+1} = \widetilde{\Phi}_\nu \circ \Phi_{\nu+1} = \widetilde{\Phi}_\nu + \widetilde{\Phi}_\nu \Psi_{\nu+1}$ (see (4.2.20)) and so

$$|\widetilde{\Phi}_{\nu+1}|_{s_0} \overset{(3.1.8)}{\leq} |\widetilde{\Phi}_\nu|_{s_0} + C|\widetilde{\Phi}_\nu|_{s_0}\, |\Psi_{\nu+1}|_{s_0} \overset{(4.2.21)}{\leq} |\widetilde{\Phi}_\nu|_{s_0}(1+\varepsilon_\nu) \tag{4.2.30}$$

where $\varepsilon_\nu := C'|\mathcal{R}_0|_{s_0+\beta}^{\mathrm{Lip}(\gamma)}\gamma^{-1}N_{\nu+1}^{2\tau+1}N_\nu^{-\alpha}$. Iterating (4.2.30) we get, for all ν,

$$|\widetilde{\Phi}_{\nu+1}|_{s_0} \leq |\widetilde{\Phi}_0|_{s_0}\Pi_{\nu\geq 0}(1+\varepsilon_\nu) \leq |\Phi_0|_{s_0} e^{C|\mathcal{R}_0|_{s_0+\beta}^{\mathrm{Lip}(\gamma)}\gamma^{-1}} \leq 2 \tag{4.2.31}$$

using (4.2.21) (with $\nu = 1$, $s = s_0$) to estimate $|\Phi_0|_{s_0}$ and (4.2.14). The high norm of $\widetilde{\Phi}_{\nu+1} = \widetilde{\Phi}_\nu + \widetilde{\Phi}_\nu \Psi_{\nu+1}$ is estimated by (3.1.9), (4.2.31) (for $\widetilde{\Phi}_\nu$), as

$$\begin{aligned}
|\widetilde{\Phi}_{\nu+1}|_s \quad & \leq \quad |\widetilde{\Phi}_\nu|_s(1 + C(s)\,|\Psi_{\nu+1}|_{s_0}) + C(s)\,|\Psi_{\nu+1}|_s \\
& \overset{(4.2.21),(4.2.13)}{\leq} \quad |\widetilde{\Phi}_\nu|_s(1+\varepsilon_\nu^{(0)}) + \varepsilon_\nu^{(s)}\,,\ \varepsilon_\nu^{(0)} := |\mathcal{R}_0|_{s_0+\beta}\gamma^{-1}N_\nu^{-1}\,,\ \varepsilon_\nu^{(s)} := |\mathcal{R}_0|_{s+\beta}\gamma^{-1}N_\nu^{-1}\,.
\end{aligned}$$

Iterating the above inequality and, using $\Pi_{j\geq 0}(1+\varepsilon_j^{(0)}) \leq 2$, we get

$$|\widetilde{\Phi}_{\nu+1}|_s \leq_s \sum_{j=0}^{\infty} \varepsilon_j^{(s)} + |\widetilde{\Phi}_0|_s \leq C(s)\big(1 + |\mathcal{R}_0|_{s+\beta}\gamma^{-1}\big) \tag{4.2.32}$$

using $|\Phi_0|_s \leq 1 + C(s)|\mathcal{R}_0|_{s+\beta}\gamma^{-1}$. Finally, the $\widetilde{\Phi}_j$ a Cauchy sequence in norm $|\cdot|_s$ because

$$\begin{aligned}
|\widetilde{\Phi}_{\nu+m} - \widetilde{\Phi}_\nu|_s \quad & \leq \quad \sum_{j=\nu}^{\nu+m-1} |\widetilde{\Phi}_{j+1} - \widetilde{\Phi}_j|_s \overset{(3.1.9)}{\leq_s} \sum_{j=\nu}^{\nu+m-1} \Big(|\widetilde{\Phi}_j|_s|\Psi_{j+1}|_{s_0} + |\widetilde{\Phi}_j|_{s_0}|\Psi_{j+1}|_s\Big) \\
& \overset{(4.2.32),(4.2.21),(4.2.31),(4.2.14)}{\leq_s} \quad \sum_{j\geq\nu} |\mathcal{R}_0|_{s+\beta}\,\gamma^{-1}N_j^{-1} \leq_s |\mathcal{R}_0|_{s+\beta}\,\gamma^{-1}N_\nu^{-1}\,.
\end{aligned} \tag{4.2.33}$$

Hence $\widetilde{\Phi}_\nu \overset{|\cdot|_s}{\to} \Phi_\infty$. The bound for $\Phi_\infty - I$ in (4.2.29) follows by (4.2.33) with $m = \infty$, $\nu = 0$ and $|\widetilde{\Phi}_0 - I|_s = |\Psi_0|_s \lessdot \gamma^{-1}|\mathcal{R}_0|_{s+\beta}$. Then the estimate for $\Phi_\infty^{-1} - I$ follows by (3.1.12).

In the reversible case all the Φ_ν are reversibility preserving and so $\widetilde{\Phi}_\nu$, Φ_∞ are reversibility preserving. $\square$

Remark 4.2.2. *In the Hamiltonian case, the transformation $\widetilde{\Phi}_\nu$ in (4.2.28) is symplectic, because Φ_ν is symplectic for all ν (see Remark 4.2.1). Therefore Φ_∞ is also symplectic.*

Let us define for all $j \in \mathbb{Z}$

$$\mu_j^\infty(\lambda) = \lim_{\nu \to +\infty} \widetilde{\mu}_j^\nu(\lambda) = \tilde{\mu}_j^0 + r_j^\infty(\lambda), \quad r_j^\infty(\lambda) := \lim_{\nu \to +\infty} \tilde{r}_j^\nu(\lambda) \quad \forall \lambda \in \Lambda.$$

It could happen that $\Lambda_{\nu_0}^\gamma = \emptyset$ (see (4.2.17)) for some ν_0. In such a case the iterative process of Theorem 4.2.2 stops after finitely many steps. However, we can always set $\widetilde{\mu}_j^\nu := \widetilde{\mu}_j^{\nu_0}$, $\forall \nu \geq \nu_0$, and the functions $\mu_j^\infty : \Lambda \to \mathbb{R}$ are always well defined.

Corollary 4.2.2. (Final eigenvalues) *For all $\nu \in \mathbb{N}$, $j \in \mathbb{Z}$*

$$|\mu_j^\infty - \widetilde{\mu}_j^\nu|_\Lambda^{\mathrm{Lip}(\gamma)} = |r_j^\infty - \widetilde{r}_j^\nu|_\Lambda^{\mathrm{Lip}(\gamma)} \leq C\, |\mathcal{R}_0|_{s_0+\beta}^{\mathrm{Lip}(\gamma)}\, N_{\nu-1}^{-\alpha}\,, \quad |\mu_j^\infty - \widetilde{\mu}_j^0|_\Lambda^{\mathrm{Lip}(\gamma)} = |r_j^\infty|_\Lambda^{\mathrm{Lip}(\gamma)} \leq C\, |\mathcal{R}_0|_{s_0+\beta}^{\mathrm{Lip}(\gamma)}\,. \tag{4.2.34}$$

Proof. The bound (4.2.34) follows by (4.2.22) and (4.2.19) by summing the telescopic series. □

Lemma 4.2.1. (Cantor set)

$$\Lambda_\infty^{2\gamma} \subset \cap_{\nu \geq 0} \Lambda_\nu^\gamma\,. \tag{4.2.35}$$

Proof. Let $\lambda \in \Lambda_\infty^{2\gamma}$. By definition $\Lambda_\infty^{2\gamma} \subset \Lambda_0^\gamma := \Lambda_o$. Then for all $\nu > 0$, $|l| \leq N_\nu$, $j \neq k$

$$\begin{aligned} \left|\mathrm{i}\omega \cdot l + \mu_j^\nu - \mu_k^\nu\right| &\geq \left|\mathrm{i}\omega \cdot l + \mu_j^\infty - \mu_k^\infty\right| - \left|\mu_j^\nu - \mu_j^\infty\right| - \left|\mu_k^\nu - \mu_k^\infty\right| \\ &\overset{(4.2.6),(4.2.34)}{\geq} 2\gamma \left|j^3 - k^3\right| \langle l \rangle^{-\tau} - 2C|\mathcal{R}_0|_{s_0+\beta} N_{\nu-1}^{-\alpha} \geq \gamma \left|j^3 - k^3\right| \langle l \rangle^{-\tau} \end{aligned}$$

because $\gamma |j^3 - k^3| \langle l \rangle^{-\tau} \geq \gamma N_\nu^{-\tau} \overset{(4.2.14)}{\geq} 2C|\mathcal{R}_0|_{s_0+\beta} N_{\nu-1}^{-\alpha}$. □

Lemma 4.2.2. *For all $\lambda \in \Lambda_\infty^{2\gamma}(u)$,*

$$\mu_j^\infty(\lambda) = \overline{\mu_{-j}^\infty(\lambda)}, \quad r_j^\infty(\lambda) = \overline{r_{-j}^\infty(\lambda)}\,, \tag{4.2.36}$$

and in the reversible case

$$\mu_j^\infty(\lambda) = -\mu_{-j}^\infty(\lambda), \quad r_j^\infty(\lambda) = -r_{-j}^\infty(\lambda)\,. \tag{4.2.37}$$

Actually in the reversible case $\mu_j^\infty(\lambda)$ are purely imaginary for all $\lambda \in \Lambda$.

Proof. Formula (4.2.36) and (4.2.37) follow because, for all $\lambda \in \Lambda_\infty^{2\gamma} \subseteq \cap_{\nu \geq 0} \Lambda_\nu^\gamma$ (see (4.2.35)), we have $\mu_j^\nu = \overline{\mu_{-j}^\nu}$, $r_j^\nu = \overline{r_{-j}^\nu}$, and, in the reversible case, the μ_j^ν are purely imaginary and $\mu_j^\nu = -\mu_{-j}^\nu$, $r_j^\nu = -r_{-j}^\nu$. The final statement follows because, in the reversible case, the $\mu_j^\nu(\lambda) \in \mathrm{i}\mathbb{R}$ as well as its extension $\widetilde{\mu}_j^\nu(\lambda)$. □

Remark 4.2.3. *In the reversible case,* (4.2.37) *imply that $\mu_0^\infty = r_0^\infty = 0$.*

Proof of Theorem 4.2.1. We apply Theorem 4.2.2 to the linear operator $\mathcal{L}_0 := \mathcal{L}_5$ in (4.1.56), where $\mathcal{R}_0 = \mathcal{R}$ defined in (4.2.11) satisfies

$$|\mathcal{R}_0|_{s_0+\beta}^{\mathrm{Lip}(\gamma)} \overset{(4.1.69)}{\leq} \varepsilon C(s_0 + \beta)\Big(1 + \|u\|_{s_0+\sigma+\beta}^{\mathrm{Lip}(\gamma)}\Big) \overset{(4.2.2)}{\leq} 2\varepsilon C(s_0 + \beta)\,. \tag{4.2.38}$$

Then the smallness condition (4.2.14) is implied by (4.2.3) taking $\delta_0 := \delta_0(\nu)$ small enough.

For all $\lambda \in \Lambda^{2\gamma}_{\infty} \subset \cap_{\nu \geq 0} \Lambda^{\gamma}_{\nu}$ (see (4.2.35)), the operators

$$\mathcal{L}_{\nu} \overset{(4.2.15)}{=} \omega \cdot \partial_{\varphi} + \mathcal{D}_{\nu} + \mathcal{R}_{\nu} \overset{|\cdot|^{\mathrm{Lip}(\gamma)}_s}{\longrightarrow} \omega \cdot \partial_{\varphi} + \mathcal{D}_{\infty} =: \mathcal{L}_{\infty} , \quad \mathcal{D}_{\infty} := \mathrm{diag}_{j \in \mathbb{Z}} \mu^{\infty}_j \tag{4.2.39}$$

because

$$|\mathcal{D}_{\nu} - \mathcal{D}_{\infty}|^{\mathrm{Lip}(\gamma)}_s = \sup_{j \in \mathbb{Z}} \left| \mu^{\nu}_j - \mu^{\infty}_j \right|^{\mathrm{Lip}(\gamma)} \overset{(4.2.34)}{\leq} C \, |\mathcal{R}_0|^{\mathrm{Lip}(\gamma)}_{s_0+\beta} \, N^{-\alpha}_{\nu-1}, \quad |\mathcal{R}_{\nu}|^{\mathrm{Lip}(\gamma)}_s \overset{(4.2.19)}{\leq} |\mathcal{R}_0|^{\mathrm{Lip}(\gamma)}_{s+\beta} \, N^{-\alpha}_{\nu-1} \, .$$

Applying (4.2.20) iteratively we get $\mathcal{L}_{\nu} = \widetilde{\Phi}^{-1}_{\nu-1} \mathcal{L}_0 \widetilde{\Phi}_{\nu-1}$ where $\widetilde{\Phi}_{\nu-1}$ is defined by (4.2.28) and $\widetilde{\Phi}_{\nu-1} \to \Phi_{\infty}$ in $| \ |_s$ (Corollary 4.2.1). Passing to the limit we deduce (4.2.7). Moreover (4.2.34) and (4.2.38) imply (4.2.5). Then (4.2.29), (4.1.69) (applied to $\mathcal{R}_0 = \mathcal{R}$) imply (4.2.8).

Estimate (4.2.9) follows from (3.1.11) (in $H^s_x(\mathbb{T})$), Lemma 3.1.4, and the bound (4.2.8).

In the reversible case, since Φ_{∞}, Φ^{-1}_{∞} are reversibility preserving (see Corollary 4.2.1), and $\mathcal{L}_0$ is reversible (see Remark 4.1.12 and Lemma 4.1.2), we get that $\mathcal{L}_{\infty}$ is reversible too. The eigenvalues μ^{∞}_j are purely imaginary by Lemma 4.2.2.

In the Hamiltonian case, $\mathcal{L}_0 \equiv \mathcal{L}_5$ is Hamiltonian, Φ_{∞} is symplectic, and therefore $\mathcal{L}_{\infty} = \Phi^{-1}_{\infty} \mathcal{L}_5 \Phi_{\infty}$ (see (4.2.7)) is Hamiltonian, namely $\mathcal{D}_{\infty}$ has the structure $\mathcal{D}_{\infty} = \partial_x \mathcal{B}$, where $\mathcal{B} = \mathrm{diag}_{j \neq 0} \{b_j\}$ is self-adjoint. This means that $b_j \in \mathbb{R}$, and therefore $\mu^{\infty}_j = \mathrm{i} j b_j$ are all purely imaginary. ■

4.2.1 Proof of Theorem 4.2.2

Proof of $(\mathbf{Si})_0$, $i = 1, \ldots, 4$. Properties (4.2.15)-(4.2.19) in $(\mathbf{S1})_0$ hold by (4.2.10)-(4.2.11) with μ^0_j defined in (4.2.16) and $r^0_j(\lambda) = 0$ (for (4.2.19) recall that $N_{-1} := 1$, see (4.2.12)). Moreover, since m_1, m_3 are real functions, μ^0_j are purely imaginary, $\mu^0_j = \overline{\mu^0_{-j}}$ and $\mu^0_j = -\mu^0_{-j}$. In the reversible case, remark 4.1.12 implies that $\mathcal{R}_0 := \mathcal{R}$, $\mathcal{L}_0 := \mathcal{L}_5$ are reversible operators. Then there is nothing else to verify.

$(\mathbf{S2})_0$ holds extending from $\Lambda^{\gamma}_0 := \Lambda_o$ to Λ the eigenvalues $\mu^0_j(\lambda)$, namely extending the functions $m_1(\lambda)$, $m_3(\lambda)$ to $\tilde{m}_1(\lambda)$, $\tilde{m}_3(\lambda)$, preserving the sup norm and the Lipschitz semi-norm, by Kirszbraun theorem, see e.g. [65]-Lemma A.2, or [57].

$(\mathbf{S3})_0$ follows by (4.1.68), for $s = s_0, s_0 + \beta$, and (4.2.2), (4.2.13).

$(\mathbf{S4})_0$ is trivial because, by definition, $\Lambda^{\gamma}_0(u_1) = \Lambda_o = \Lambda^{\gamma-\rho}_0(u_2)$.

The reducibility step

We now describe the generic inductive step, showing how to define $\mathcal{L}_{\nu+1}$ (and Φ_{ν}, Ψ_{ν}, etc). To simplify notations, in this Section we drop the index ν and we write $+$ for $\nu + 1$. We have

$$\begin{aligned} \mathcal{L}\Phi h &= \omega \cdot \partial_{\varphi}(\Phi(h)) + \mathcal{D}\Phi h + \mathcal{R}\Phi h \\ &= \omega \cdot \partial_{\varphi} h + \Psi \omega \cdot \partial_{\varphi} h + (\omega \cdot \partial_{\varphi} \Psi) h + \mathcal{D} h + \mathcal{D}\Psi h + \mathcal{R} h + \mathcal{R}\Psi h \\ &= \Phi\Big(\omega \cdot \partial_{\varphi} h + \mathcal{D} h\Big) + \Big(\omega \cdot \partial_{\varphi} \Psi + [\mathcal{D}, \Psi] + \Pi_N \mathcal{R}\Big) h + \Big(\Pi^{\perp}_N \mathcal{R} + \mathcal{R}\Psi\Big) h \end{aligned} \tag{4.2.40}$$

where $[\mathcal{D}, \Psi] := \mathcal{D}\Psi - \Psi\mathcal{D}$ and $\Pi_N \mathcal{R}$ is defined in (3.1.18).

Remark 4.2.4. *The application of the smoothing operator Π_N is necessary since we are performing a differentiable Nash-Moser scheme. Note also that Π_N regularizes only in time (see* (3.1.18)*) because the loss of derivatives of the inverse operator is only in φ (see* (4.2.44) *and the bound on the small divisors* (4.2.17)*).*

We look for a solution of the *homological* equation

$$\omega \cdot \partial_\varphi \Psi + [\mathcal{D}, \Psi] + \Pi_N \mathcal{R} = [\mathcal{R}] \qquad \text{where} \qquad [\mathcal{R}] := \mathrm{diag}_{j\in\mathbb{Z}} \mathcal{R}_j^j(0)\,. \tag{4.2.41}$$

Lemma 4.2.3. (Homological equation) *For all $\lambda \in \Lambda_{\nu+1}^\gamma$, (see* (4.2.17)*) there exists a unique solution $\Psi := \Psi(\varphi)$ of the homological equation* (4.2.41)*. The map Ψ satisfies*

$$|\Psi|_s^{\mathrm{Lip}(\gamma)} \leq C N^{2\tau+1} \gamma^{-1} |\mathcal{R}|_s^{\mathrm{Lip}(\gamma)}\,. \tag{4.2.42}$$

Moreover if $\gamma/2 \leq \gamma_1, \gamma_2 \leq 2\gamma$ and if $u_1(\lambda)$, $u_2(\lambda)$ are Lipschitz functions, then $\forall s \in [s_0, s_0+\beta]$, $\lambda \in \Lambda_{\nu+1}^{\gamma_1}(u_1) \cap \Lambda_{\nu+1}^{\gamma_2}(u_2)$

$$|\Delta_{12}\Psi|_s \leq C N^{2\tau+1}\gamma^{-1}\Big(|\mathcal{R}(u_2)|_s \|u_1 - u_2\|_{s_0+\sigma_2} + |\Delta_{12}\mathcal{R}|_s \Big) \tag{4.2.43}$$

where we define $\Delta_{12}\Psi := \Psi(u_1) - \Psi(u_2)$.

In the reversible case, Ψ is reversibility-preserving.

Proof. Since $\mathcal{D} := \mathrm{diag}_{j\in\mathbb{Z}}(\mu_j)$ we have $[\mathcal{D},\Psi]_j^k = (\mu_j - \mu_k)\Psi_j^k(\varphi)$ and (4.2.41) amounts to

$$\omega\cdot\partial_\varphi \Psi_j^k(\varphi) + (\mu_j - \mu_k)\Psi_j^k(\varphi) + \mathcal{R}_j^k(\varphi) = [\mathcal{R}]_j^k\,, \quad \forall j,k \in \mathbb{Z}\,,$$

whose solutions are $\Psi_j^k(\varphi) = \sum_{l\in\mathbb{Z}^\nu} \Psi_j^k(l) e^{\mathrm{i}l\cdot\varphi}$ with coefficients

$$\Psi_j^k(l) := \begin{cases} \dfrac{\mathcal{R}_j^k(l)}{\delta_{ljk}(\lambda)} & \text{if } (j-k, l) \neq (0,0) \ \text{ and } \ |l| \leq N\,, \\ 0 & \text{otherwise,} \end{cases} \tag{4.2.44}$$

where

$$\delta_{ljk}(\lambda) := \mathrm{i}\omega\cdot l + \mu_j - \mu_k\,.$$

Note that, for all $\lambda \in \Lambda_{\nu+1}^\gamma$, by (4.2.17) and (1.2.2), if $j \neq k$ or $l \neq 0$ the divisors $\delta_{ljk}(\lambda) \neq 0$. Recalling the definition of the s-norm in (3.1.1) we deduce by (4.2.44), (4.2.17), (1.2.2), that

$$|\Psi|_s \leq \gamma^{-1} N^\tau |\mathcal{R}|_s\,, \quad \forall \lambda \in \Lambda_{\nu+1}^\gamma\,. \tag{4.2.45}$$

For $\lambda_1, \lambda_2 \in \Lambda_{\nu+1}^\gamma$,

$$|\Psi_j^k(l)(\lambda_1) - \Psi_j^k(l)(\lambda_2)| \leq \frac{|\mathcal{R}_j^k(l)(\lambda_1) - \mathcal{R}_j^k(l)(\lambda_2)|}{|\delta_{ljk}(\lambda_1)|} + |\mathcal{R}_j^k(l)(\lambda_2)| \frac{|\delta_{ljk}(\lambda_1) - \delta_{ljk}(\lambda_2)|}{|\delta_{ljk}(\lambda_1)||\delta_{ljk}(\lambda_2)|} \tag{4.2.46}$$

and, since $\omega = \lambda\bar\omega$,

$$\begin{aligned} |\delta_{ljk}(\lambda_1) - \delta_{ljk}(\lambda_2)| &\overset{(4.2.44)}{=} |(\lambda_1 - \lambda_2)\bar\omega\cdot l + (\mu_j - \mu_k)(\lambda_1) - (\mu_j-\mu_k)(\lambda_2)| \qquad (4.2.47) \\ &\overset{(4.2.16)}{\leq} |\lambda_1-\lambda_2||\bar\omega\cdot l| + |m_3(\lambda_1) - m_3(\lambda_2)||j^3 - k^3| + |m_1(\lambda_1) - m_1(\lambda_2)||j-k| \\ &\quad + |r_j(\lambda_1) - r_j(\lambda_2)| + |r_k(\lambda_1) - r_k(\lambda_2)| \\ &\lessdot |\lambda_1 - \lambda_2|\Big(|l| + \varepsilon\gamma^{-1}|j^3 - k^3| + \varepsilon\gamma^{-1}|j-k| + \varepsilon\gamma^{-1}\Big) \qquad (4.2.48) \end{aligned}$$

because

$$\gamma |m_3|^{\mathrm{lip}} = \gamma |m_3 - 1|^{\mathrm{lip}} \le |m_3 - 1|^{\mathrm{Lip}(\gamma)} \le \varepsilon C, \quad |m_1|^{\mathrm{Lip}(\gamma)} \le \varepsilon C, \quad |r_j|^{\mathrm{Lip}(\gamma)} \le \varepsilon C \quad \forall j \in \mathbb{Z}.$$

Hence, for $j \neq k$, $\varepsilon\gamma^{-1} \le 1$,

$$\begin{aligned} \frac{|\delta_{ljk}(\lambda_1) - \delta_{ljk}(\lambda_2)|}{|\delta_{ljk}(\lambda_1)||\delta_{ljk}(\lambda_2)|} &\overset{(4.2.48),(4.2.17)}{\lessdot} |\lambda_1 - \lambda_2| \Big(|l| + |j^3 - k^3| \Big) \frac{\langle l \rangle^{2\tau}}{\gamma^2 \, |j^3 - k^3|^2} \\ &\lessdot |\lambda_1 - \lambda_2| N^{2\tau+1} \gamma^{-2} \end{aligned} \tag{4.2.49}$$

for $|l| \le N$. Finally, recalling (3.1.1), the bounds (4.2.46), (4.2.49) and (4.2.45) imply (4.2.42). Now we prove (4.2.43). By (4.2.44), for any $\lambda \in \Lambda^{\gamma_1}_{\nu+1}(u_1) \cap \Lambda^{\gamma_2}_{\nu+1}(u_2)$, $l \in \mathbb{Z}^\nu$, $j \neq k$, we get

$$\Delta_{12} \Psi_j^k(l) = \frac{\Delta_{12} \mathcal{R}_j^k(l)}{\delta_{ljk}(u_1)} - \mathcal{R}_j^k(l)(u_2) \frac{\Delta_{12}\delta_{ljk}}{\delta_{ljk}(u_1)\delta_{ljk}(u_2)} \tag{4.2.50}$$

where

$$\begin{aligned} |\Delta_{12}\delta_{ljk}| &= |\Delta_{12}(\mu_j - \mu_k)| \le |\Delta_{12} m_3| \, |j^3 - k^3| + |\Delta_{12} m_1| \, |j - k| + |\Delta_{12} r_j| + |\Delta_{12} r_k| \\ &\overset{(4.1.65),(4.2.25)}{\lessdot} \varepsilon |j^3 - k^3| \|u_1 - u_2\|_{s_0+\sigma_2} \, . \end{aligned} \tag{4.2.51}$$

Then (4.2.50), (4.2.51), $\varepsilon\gamma^{-1} \le 1$, $\gamma_1^{-1}, \gamma_2^{-1} \le \gamma^{-1}$ imply

$$|\Delta_{12} \Psi_j^k(l)| \lessdot N^{2\tau} \gamma^{-1} \Big(|\Delta_{12} \mathcal{R}_j^k(l)| + |\mathcal{R}_j^k(l)(u_2)| \|u_1 - u_2\|_{s_0+\sigma_2} \Big)$$

and so (4.2.43) (in fact, (4.2.43) holds with 2τ instead of $2\tau + 1$).

In the reversible case $\mathrm{i}\omega \cdot l + \mu_j - \mu_k \in \mathrm{i}\mathbb{R}$, $\overline{\mu_{-j}} = \mu_j$ and $\mu_{-j} = -\mu_j$. Hence Lemma 3.2.1 and (4.2.44) imply

$$\overline{\Psi_{-j}^{-k}(-l)} = \frac{\overline{\mathcal{R}_{-j}^{-k}(-l)}}{-\mathrm{i}\omega \cdot (-l) + \overline{\mu_{-j}} - \overline{\mu_{-k}}} = \frac{\mathcal{R}_j^k(l)}{\mathrm{i}\omega \cdot l + \mu_j - \mu_k} = \Psi_j^k(l)$$

and so Ψ is real, again by Lemma 3.2.1. Moreover, since $\mathcal{R} : X \to Y$,

$$\Psi_{-j}^{-k}(-l) = \frac{\mathcal{R}_{-j}^{-k}(-l)}{\mathrm{i}\omega \cdot (-l) + \mu_{-j} - \mu_{-k}} = \frac{-\mathcal{R}_j^k(l)}{\mathrm{i}\omega \cdot (-l) - \mu_j + \mu_k} = \Psi_j^k(l)$$

which implies $\Psi : X \to X$ by Lemma 3.2.1. Similarly we get $\Psi : Y \to Y$. $\square$

Remark 4.2.5. *In the Hamiltonian case $\mathcal{R}$ is Hamiltonian and the solution Ψ in* (4.2.44) *of the homological equation is Hamiltonian, because $\overline{\delta_{l,j,k}} = \delta_{-l,k,j}$ and, in terms of matrix elements, an operator $G(\varphi)$ is self-adjoint if and only if $\overline{G_j^k(l)} = G_k^j(-l)$.*

Let Ψ be the solution of the homological equation (4.2.41) which has been constructed in Lemma 4.2.3. By Lemma 3.1.3, if $C(s_0)|\Psi|_{s_0} < 1/2$ then $\Phi := I + \Psi$ is invertible and by (4.2.40) (and (4.2.41)) we deduce that

$$\mathcal{L}_+ := \Phi^{-1} \mathcal{L} \Phi = \omega \cdot \partial_\varphi + \mathcal{D}_+ + \mathcal{R}_+ \, , \tag{4.2.52}$$

where

$$\mathcal{D}_+ := \mathcal{D} + [\mathcal{R}] \, , \quad \mathcal{R}_+ := \Phi^{-1} \Big(\Pi_N^\perp \mathcal{R} + \mathcal{R}\Psi - \Psi[\mathcal{R}] \Big). \tag{4.2.53}$$

Note that $\mathcal{L}_+$ has the same form of $\mathcal{L}$, but the remainder $\mathcal{R}_+$ is the sum of a quadratic function of $\Psi, \mathcal{R}$ and a remainder supported on high modes.

Lemma 4.2.4. (New diagonal part). *The eigenvalues of*

$$\mathcal{D}_+ = \mathrm{diag}_{j\in\mathbb{Z}}\{\mu_j^+(\lambda)\}, \quad \textit{where } \ \mu_j^+ := \mu_j + \mathcal{R}_j^j(0) = \mu_j^0 + r_j + \mathcal{R}_j^j(0) = \mu_j^0 + r_j^+, \quad r_j^+ := r_j + \mathcal{R}_j^j(0),$$

satisfy $\mu_j^+ = \overline{\mu_{-j}^+}$ *and*

$$|\mu_j^+ - \mu_j|^{\mathrm{lip}} = |r_j^+ - r_j|^{\mathrm{lip}} = |\mathcal{R}_j^j(0)|^{\mathrm{lip}} \leq |\mathcal{R}|_{s_0}^{\mathrm{lip}}, \quad \forall j \in \mathbb{Z}. \tag{4.2.54}$$

Moreover if $u_1(\lambda)$, $u_2(\lambda)$ *are Lipschitz functions, then for all* $\lambda \in \Lambda_\nu^{\gamma_1}(u_1) \cap \Lambda_\nu^{\gamma_2}(u_2)$

$$|\Delta_{12} r_j^+ - \Delta_{12} r_j| \leq |\Delta_{12}\mathcal{R}|_{s_0}. \tag{4.2.55}$$

In the reversible case, all the μ_j^+ *are purely imaginary and satisfy* $\mu_j^+ = -\mu_{-j}^+$ *for all* $j \in \mathbb{Z}$.

Proof. The estimates (4.2.54)-(4.2.55) follow using (3.1.3) because $|\mathcal{R}_j^j(0)|^{\mathrm{lip}} = |\mathcal{R}_{(l,j)}^{(l,j)}|^{\mathrm{lip}} \leq |\mathcal{R}|_0^{\mathrm{lip}} \leq |\mathcal{R}|_{s_0}^{\mathrm{lip}}$ and

$$|\Delta_{12} r_j^+ - \Delta_{12} r_j| = |\Delta_{12}\mathcal{R}_j^j(0)| = |\Delta_{12}\mathcal{R}_{(l,j)}^{(l,j)}| \leq |\Delta_{12}\mathcal{R}|_0 \leq |\Delta_{12}\mathcal{R}|_{s_0}.$$

Since $\mathcal{R}$ is real, by Lemma 3.2.1,

$$\mathcal{R}_j^k(l) = \overline{\mathcal{R}_{-j}^{-k}(-l)} \qquad \Longrightarrow \qquad \mathcal{R}_j^j(0) = \overline{\mathcal{R}_{-j}^{-j}(0)}$$

and so $\mu_j^+ = \overline{\mu_{-j}^+}$. If $\mathcal{R}$ is also reversible, by Lemma 3.2.1,

$$\mathcal{R}_j^k(l) = -\mathcal{R}_{-j}^{-k}(-l), \quad \mathcal{R}_j^k(l) = \overline{\mathcal{R}_{-j}^{-k}(-l)} = -\overline{\mathcal{R}_j^k(l)}.$$

We deduce that $\mathcal{R}_j^j(0) = -\mathcal{R}_{-j}^{-j}(0)$, $\mathcal{R}_j^j(0) \in \mathrm{i}\mathbb{R}$ and therefore, $\mu_j^+ = -\mu_{-j}^+$ and $\mu_j^+ \in \mathrm{i}\mathbb{R}$. □

Remark 4.2.6. *In the Hamiltonian case,* $\mathcal{D}_\nu$ *is Hamiltonian, namely* $\mathcal{D}_\nu = \partial_x \mathcal{B}$ *where* $\mathcal{B} = \mathrm{diag}_{j\neq 0}\{b_j\}$ *is self-adjoint. This means that* $b_j \in \mathbb{R}$, *and therefore all* $\mu_j^\nu = \mathrm{i} j b_j$ *are purely imaginary.*

The iteration

Let $\nu \geq 0$, and suppose that the statements $(\mathbf{Si})_\nu$ are true. We prove $(\mathbf{Si})_{\nu+1}$, $i = 1, \ldots, 4$. To simplify notations we write $|\cdot|_s$ instead of $|\cdot|_s^{\mathrm{Lip}(\gamma)}$.

Proof of $(\mathbf{S1})_{\nu+1}$. By $(\mathbf{S1})_\nu$, the eigenvalues μ_j^ν are defined on Λ_ν^γ. Therefore the set $\Lambda_{\nu+1}^\gamma$ is well-defined. By Lemma 4.2.3, for all $\lambda \in \Lambda_{\nu+1}^\gamma$ there exists a real solution Ψ_ν of the homological equation (4.2.41) which satisfies, $\forall s \in [s_0, q - \sigma - \beta]$,

$$|\Psi_\nu|_s \overset{(4.2.42)}{\lessdot} N_\nu^{2\tau+1} |\mathcal{R}_\nu|_s \gamma^{-1} \overset{(4.2.19)}{\lessdot} |\mathcal{R}_0|_{s+\beta} \gamma^{-1} N_\nu^{2\tau+1} N_{\nu-1}^{-\alpha} \tag{4.2.56}$$

which is (4.2.21) at the step $\nu + 1$. In particular, for $s = s_0$,

$$C(s_0) |\Psi_\nu|_{s_0} \overset{(4.2.56)}{\leq} C(s_0) |\mathcal{R}_0|_{s_0+\beta} \gamma^{-1} N_\nu^{2\tau+1} N_{\nu-1}^{-\alpha} \overset{(4.2.14)}{\leq} 1/2 \tag{4.2.57}$$

for N_0 large enough. Then the map $\Phi_\nu := I + \Psi_\nu$ is invertible and, by (3.1.12),

$$\left|\Phi_\nu^{-1}\right|_{s_0} \le 2\,, \quad \left|\Phi_\nu^{-1}\right|_s \le 1 + C(s)|\Psi_\nu|_s\,. \tag{4.2.58}$$

Hence (4.2.52)-(4.2.53) imply $\mathcal{L}_{\nu+1} := \Phi_\nu^{-1}\mathcal{L}_\nu\Phi_\nu = \omega\cdot\partial_\varphi + \mathcal{D}_{\nu+1} + \mathcal{R}_{\nu+1}$ where (see Lemma 4.2.4)

$$\mathcal{D}_{\nu+1} := \mathcal{D}_\nu + [\mathcal{R}_\nu] = \mathrm{diag}_{j\in\mathbb{Z}}(\mu_j^{\nu+1})\,, \quad \mu_j^{\nu+1} := \mu_j^\nu + (\mathcal{R}_\nu)_j^j(0)\,, \tag{4.2.59}$$

with $\mu_j^{\nu+1} = \overline{\mu_{-j}^{\nu+1}}$ and

$$\mathcal{R}_{\nu+1} := \Phi_\nu^{-1}H_\nu, \quad H_\nu := \Pi_{N_\nu}^\perp\mathcal{R}_\nu + \mathcal{R}_\nu\Psi_\nu - \Psi_\nu[\mathcal{R}_\nu]\,. \tag{4.2.60}$$

In the reversible case, $\mathcal{R}_\nu : X \to Y$, therefore, by Lemma 4.2.3, Ψ_ν, Φ_ν, Φ_ν^{-1} are reversibility preserving, and then, by formula (4.2.60), also $\mathcal{R}_{\nu+1} : X \to Y$.

Let us prove the estimates (4.2.19) for $\mathcal{R}_{\nu+1}$. For all $s\in[s_0, q-\sigma-\beta]$ we have

$$\begin{aligned}
|\mathcal{R}_{\nu+1}|_s &\overset{(4.2.60),(3.1.9)}{\le_s} |\Phi_\nu^{-1}|_{s_0}\Big(|\Pi_{N_\nu}^\perp\mathcal{R}_\nu|_s + |\mathcal{R}_\nu|_s|\Psi_\nu|_{s_0} + |\mathcal{R}_\nu|_{s_0}|\Psi_\nu|_s\Big) + |\Phi_\nu^{-1}|_s\Big(|\Pi_{N_\nu}^\perp\mathcal{R}_\nu|_{s_0} + |\mathcal{R}_\nu|_{s_0}|\Psi_\nu|_{s_0}\Big)\\
&\overset{(4.2.58)}{\le_s} 2\Big(|\Pi_{N_\nu}^\perp\mathcal{R}_\nu|_s + |\mathcal{R}_\nu|_s|\Psi_\nu|_{s_0} + |\mathcal{R}_\nu|_{s_0}|\Psi_\nu|_s\Big) + (1+|\Psi_\nu|_s)\Big(|\Pi_{N_\nu}^\perp\mathcal{R}_\nu|_{s_0} + |\mathcal{R}_\nu|_{s_0}|\Psi_\nu|_{s_0}\Big)\\
&\overset{(4.2.57)}{\le_s} |\Pi_{N_\nu}^\perp\mathcal{R}_\nu|_s + |\mathcal{R}_\nu|_s|\Psi_\nu|_{s_0} + |\mathcal{R}_\nu|_{s_0}|\Psi_\nu|_s\\
&\overset{(4.2.42)}{\le_s} |\Pi_{N_\nu}^\perp\mathcal{R}_\nu|_s + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|_s|\mathcal{R}_\nu|_{s_0}\,.
\end{aligned} \tag{4.2.61}$$

Hence (4.2.61) and (3.1.19) imply

$$|\mathcal{R}_{\nu+1}|_s \le_s N_\nu^{-\beta}|\mathcal{R}_\nu|_{s+\beta} + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|_s|\mathcal{R}_\nu|_{s_0} \tag{4.2.62}$$

which shows that the iterative scheme is quadratic plus a super-exponentially small term. In particular

$$|\mathcal{R}_{\nu+1}|_s \overset{(4.2.62),(4.2.19)}{\le_s} N_\nu^{-\beta}|\mathcal{R}_0|_{s+\beta}N_{\nu-1} + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_0|_{s+\beta}|\mathcal{R}_0|_{s_0+\beta}N_{\nu-1}^{-2\alpha} \overset{(4.2.1),(4.2.13),(4.2.14)}{\le} |\mathcal{R}_0|_{s+\beta}N_\nu^{-\alpha}$$

($\chi = 3/2$) which is the first inequality of (4.2.19) at the step $\nu+1$. The next key step is to control the divergence of the high norm $|\mathcal{R}_{\nu+1}|_{s+\beta}$. By (4.2.61) (with $s+\beta$ instead of s) we get

$$|\mathcal{R}_{\nu+1}|_{s+\beta} \le_{s+\beta} |\mathcal{R}_\nu|_{s+\beta} + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|_{s+\beta}|\mathcal{R}_\nu|_{s_0} \tag{4.2.63}$$

(the difference with respect to (4.2.62) is that we do not apply to $|\Pi_{N_\nu}^\perp\mathcal{R}_\nu|_{s+\beta}$ any smoothing). Then (4.2.63), (4.2.19), (4.2.14), (4.2.13) imply the inequality

$$|\mathcal{R}_{\nu+1}|_{s+\beta} \le C(s+\beta)|\mathcal{R}_\nu|_{s+\beta},$$

whence, iterating,

$$|\mathcal{R}_{\nu+1}|_{s+\beta} \le N_\nu|\mathcal{R}_0|_{s+\beta}$$

for $N_0 := N_0(s,\beta)$ large enough, which is the second inequality of (4.2.19) with index $\nu+1$.

By Lemma 4.2.4 the eigenvalues $\mu_j^{\nu+1} := \mu_j^0 + r_j^{\nu+1}$, defined on $\Lambda_{\nu+1}^{\gamma}$, satisfy $\mu_j^{\nu+1} = \overline{\mu_{-j}^{\nu+1}}$, and, in the reversible case, the $\mu_j^{\nu+1}$ are purely imaginary and $\mu_j^{\nu+1} = -\mu_{-j}^{\nu+1}$.

It remains only to prove (4.2.18) for $\nu+1$, which is proved below.

PROOF OF $(\mathbf{S2})_{\nu+1}$. By (4.2.54),

$$|\mu_j^{\nu+1} - \mu_j^{\nu}|^{\mathrm{Lip}(\gamma)} = |r_j^{\nu+1} - r_j^{\nu}|^{\mathrm{Lip}(\gamma)} \le |\mathcal{R}_\nu|_{s_0}^{\mathrm{Lip}(\gamma)} \overset{(4.2.19)}{\le} |\mathcal{R}_0|_{s_0+\beta}^{\mathrm{Lip}(\gamma)} N_{\nu-1}^{-\alpha}\,. \tag{4.2.64}$$

By Kirszbraun theorem, we extend the function $\mu_j^{\nu+1} - \mu_j^{\nu} = r_j^{\nu+1} - r_j^{\nu}$ to the whole Λ, still satisfying (4.2.64). In this way we define $\tilde{\mu}_j^{\nu+1}$. Finally (4.2.18) follows summing all the terms in (4.2.64) and using (4.1.69).

PROOF OF $(\mathbf{S3})_{\nu+1}$. Set, for brevity,

$$\mathcal{R}_\nu^i := \mathcal{R}_\nu(u_i), \quad \Psi_{\nu-1}^i := \Psi_{\nu-1}(u_i), \quad \Phi_{\nu-1}^i := \Phi_{\nu-1}(u_i), \quad H_{\nu-1}^i := H_{\nu-1}(u_i)\,, \quad i := 1, 2\,,$$

which are all operators defined for $\lambda \in \Lambda_\nu^{\gamma_1}(u_1) \cap \Lambda_\nu^{\gamma_2}(u_2)$. By Lemma 4.2.3 one can construct $\Psi_\nu^i := \Psi_\nu(u_i)$, $\Phi_\nu^i := \Phi_\nu(u_i)$, $i = 1, 2$, for all $\lambda \in \Lambda_{\nu+1}^{\gamma_1}(u_1) \cap \Lambda_{\nu+1}^{\gamma_2}(u_2)$. One has

$$\begin{aligned} |\Delta_{12}\Psi_\nu|_{s_0} &\overset{(4.2.43)}{\lessdot} N_\nu^{2\tau+1}\gamma^{-1}\Big(|\mathcal{R}_\nu(u_2)|_{s_0}\|u_2 - u_1\|_{s_0+\sigma_2} + |\Delta_{12}\mathcal{R}_\nu|_{s_0}\Big) \\ &\overset{(4.2.19),(4.2.23)}{\lessdot} N_\nu^{2\tau+1}N_{\nu-1}^{-\alpha}\gamma^{-1}\big(|\mathcal{R}_0|_{s_0+\beta} + \varepsilon\big)\|u_2 - u_1\|_{s_0+\sigma_2} \\ &\overset{(4.1.69),(4.2.2)}{\lessdot} N_\nu^{2\tau+1}N_{\nu-1}^{-\alpha}\varepsilon\gamma^{-1}\|u_2 - u_1\|_{s_0+\sigma_2} \le \|u_2 - u_1\|_{s_0+\sigma_2}. \end{aligned} \tag{4.2.65}$$

for $\varepsilon\gamma^{-1}$ small (and (4.2.13)). By (3.1.13), applied to $\Phi := \Phi_\nu$, and (4.2.65), we get

$$|\Delta_{12}\Phi_\nu^{-1}|_s \le_s \big(|\Psi_\nu^1|_s + |\Psi_\nu^2|_s\big)\|u_1 - u_2\|_{s_0+\sigma_2} + |\Delta_{12}\Psi_\nu|_s \tag{4.2.66}$$

which implies for $s = s_0$, and using (4.2.21), (4.2.14), (4.2.65)

$$|\Delta_{12}\Phi_\nu^{-1}|_{s_0} \lessdot \|u_1 - u_2\|_{s_0+\sigma_2}. \tag{4.2.67}$$

Let us prove the estimates (4.2.23) for $\Delta_{12}\mathcal{R}_{\nu+1}$, which is defined on $\lambda \in \Lambda_{\nu+1}^{\gamma_1}(u_1) \cap \Lambda_{\nu+1}^{\gamma_2}(u_2)$. For all $s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$, using the interpolation (3.1.6) and (4.2.60),

$$\begin{aligned} |\Delta_{12}\mathcal{R}_{\nu+1}|_s \le_s & |\Delta_{12}\Phi_\nu^{-1}|_s |H_\nu^1|_{s_0} + |\Delta_{12}\Phi_\nu^{-1}|_{s_0}|H_\nu^1|_s + |(\Phi_\nu^2)^{-1}|_s|\Delta_{12}H_\nu|_{s_0} \\ & + |(\Phi_\nu^2)^{-1}|_{s_0}|\Delta_{12}H_\nu|_s\,. \end{aligned} \tag{4.2.68}$$

We estimate the above terms separately. Set for brevity $A_s^\nu := |\mathcal{R}_\nu(u_1)|_s + |\mathcal{R}_\nu(u_2)|_s$. By (4.2.60) and (3.1.6),

$$\begin{aligned} |\Delta_{12}H_\nu|_s \quad \le_s \quad & \Big|\Pi_{N_\nu}^{\perp}\Delta_{12}\mathcal{R}_\nu\Big|_s + |\Delta_{12}\Psi_\nu|_s|\mathcal{R}_\nu^1|_{s_0} + |\Delta_{12}\Psi_\nu|_{s_0}|\mathcal{R}_\nu^1|_s + |\Psi_\nu^2|_s|\Delta_{12}\mathcal{R}_\nu|_{s_0} + |\Psi_\nu^2|_{s_0}|\Delta_{12}\mathcal{R}_\nu|_s \\ \overset{(4.2.42),(4.2.43)}{\le_s} \quad & \Big|\Pi_{N_\nu}^{\perp}\Delta_{12}\mathcal{R}_\nu\Big|_s + N_\nu^{2\tau+1}\gamma^{-1}A_{s_0}^\nu A_s^\nu\|u_1 - u_2\|_{s_0+\sigma_2} \\ & + N_\nu^{2\tau+1}\gamma^{-1}A_s^\nu|\Delta_{12}\mathcal{R}_\nu|_{s_0} + N_\nu^{2\tau+1}\gamma^{-1}A_{s_0}^\nu|\Delta_{12}\mathcal{R}_\nu|_s\,. \end{aligned} \tag{4.2.69}$$

Estimating the four terms in the right hand side of (4.2.68) in the same way, using (4.2.66), (4.2.60), (4.2.42), (4.2.43), (4.2.21), (4.2.67), (4.2.58), (4.2.69), (4.2.19), we deduce

$$
\begin{aligned}
|\Delta_{12}\mathcal{R}_{\nu+1}|_s \quad \leq_s \quad & |\Pi^{\perp}_{N_\nu}\Delta_{12}\mathcal{R}_\nu|_s + N_\nu^{2\tau+1}\gamma^{-1}A_s^\nu A_{s_0}^\nu \|u_1-u_2\|_{s_0+\sigma_2} \\
& + N_\nu^{2\tau+1}\gamma^{-1}A_s^\nu|\Delta_{12}\mathcal{R}_\nu|_{s_0} + N_\nu^{2\tau+1}\gamma^{-1}A_{s_0}^\nu|\Delta_{12}\mathcal{R}_\nu|_s \, . \qquad (4.2.70)
\end{aligned}
$$

Specializing (4.2.70) for $s = s_0$ and using (4.1.69), (3.1.19), (4.2.19), (4.2.23), we deduce

$$
|\Delta_{12}\mathcal{R}_{\nu+1}|_{s_0} \leq C(\varepsilon N_{\nu-1}N_\nu^{-\beta} + N_\nu^{2\tau+1}N_{\nu-1}^{-2\alpha}\varepsilon^2\gamma^{-1})\|u_1-u_2\|_{s_0+\sigma_2} \leq \varepsilon N_\nu^{-\alpha}\|u_1-u_2\|_{s_0+\sigma_2}
$$

for N_0 large and $\varepsilon\gamma^{-1}$ small. Next by (4.2.70) with $s = s_0+\beta$

$$
\begin{aligned}
|\Delta_{12}\mathcal{R}_\nu|_{s_0+\beta} \quad & \overset{(4.2.19),(4.2.23),(4.2.14)}{\leq_{s_0+\beta}} \quad A^\nu_{s_0+\beta}\|u_1-u_2\|_{s_0+\sigma_2} + |\Delta_{12}\mathcal{R}_\nu|_{s_0+\beta} \\
& \overset{(4.2.19)(4.2.23)}{\leq} \quad C(s_0+\beta)\varepsilon N_{\nu-1}\|u_1-u_2\|_{s_0+\sigma_2} \leq \varepsilon N_\nu \|u_1-u_2\|_{s_0+\sigma_2}
\end{aligned}
$$

for N_0 large enough. Finally note that (4.2.24) is nothing but (4.2.55).

Proof of $(\mathbf{S4})_{\nu+1}$. We have to prove that, if $C\varepsilon N_\nu^\tau\|u_1-u_2\|_{s_0+\sigma_2} \leq \rho$, then

$$
\lambda \in \Lambda^\gamma_{\nu+1}(u_1) \quad \Longrightarrow \quad \lambda \in \Lambda^{\gamma-\rho}_{\nu+1}(u_2) \, .
$$

Let $\lambda \in \Lambda^\gamma_{\nu+1}(u_1)$. Definition (4.2.17) and $(\mathbf{S4})_\nu$ (see (4.2.26)) imply that $\Lambda^\gamma_{\nu+1}(u_1) \subseteq \Lambda^\gamma_\nu(u_1) \subseteq \Lambda^{\gamma-\rho}_\nu(u_2)$. Hence $\lambda \in \Lambda^{\gamma-\rho}_\nu(u_2) \subset \Lambda^{\gamma/2}_\nu(u_2)$. Then, by $(\mathbf{S1})_\nu$, the eigenvalues $\mu_j^\nu(\lambda, u_2(\lambda))$ are well defined. Now (4.2.16) and the estimates (4.1.65), (4.2.25) (which holds because $\lambda \in \Lambda^\gamma_\nu(u_1) \cap \Lambda^{\gamma/2}_\nu(u_2)$) imply that

$$
\begin{aligned}
|(\mu_j^\nu - \mu_k^\nu)(\lambda, u_2(\lambda)) - (\mu_j^\nu - \mu_k^\nu)(\lambda, u_1(\lambda))| \quad \leq \quad & |(\mu_j^0 - \mu_k^0)(\lambda, u_2(\lambda)) - (\mu_j^0 - \mu_k^0)(\lambda, u_1(\lambda))| \\
& + 2\sup_{j\in\mathbb{Z}} |r_j^\nu(\lambda, u_2(\lambda)) - r_j^\nu(\lambda, u_1(\lambda))| \\
\leq \quad & \varepsilon C|j^3-k^3| \|u_2-u_1\|^{\sup}_{s_0+\sigma_2} \, . \qquad (4.2.71)
\end{aligned}
$$

Then we conclude that for all $|l| \leq N_\nu$, $j \neq k$, using the definition of $\Lambda^\gamma_{\nu+1}(u_1)$ (which is (4.2.17) with $\nu+1$ instead of ν) and (4.2.71),

$$
\begin{aligned}
|\mathrm{i}\omega\cdot l + \mu_j^\nu(u_2) - \mu_k^\nu(u_2)| \quad \geq \quad & |\mathrm{i}\omega\cdot l + \mu_j^\nu(u_1) - \mu_k^\nu(u_1)| - |(\mu_j^\nu - \mu_k^\nu)(u_2) - (\mu_j^\nu - \mu_k^\nu)(u_1)| \\
\geq \quad & \gamma|j^3-k^3|\langle l\rangle^{-\tau} - C\varepsilon|j^3-k^3|\|u_1-u_2\|_{s_0+\sigma_2} \\
\geq \quad & (\gamma-\rho)|j^3-k^3|\langle l\rangle^{-\tau}
\end{aligned}
$$

provided $C\varepsilon N_\nu^\tau\|u_1-u_2\|_{s_0+\sigma_2} \leq \rho$. Hence $\lambda \in \Lambda^{\gamma-\rho}_{\nu+1}(u_2)$. This proves (4.2.26) at the step $\nu+1$.

4.2.2 Inversion of $\mathcal{L}(u)$

In (4.1.58) we have conjugated the linearized operator $\mathcal{L}$ to $\mathcal{L}_5$ defined in (4.1.56), namely $\mathcal{L} = \Phi_1\mathcal{L}_5\Phi_2^{-1}$. In Theorem 4.2.1 we have conjugated the operator $\mathcal{L}_5$ to the diagonal operator $\mathcal{L}_\infty$ in (4.2.7), namely $\mathcal{L}_5 = \Phi_\infty\mathcal{L}_\infty\Phi_\infty^{-1}$. As a consequence

$$
\mathcal{L} = W_1\mathcal{L}_\infty W_2^{-1}, \quad W_i := \Phi_i\Phi_\infty, \quad \Phi_1 := \mathcal{A}\mathcal{B}\rho\mathcal{M}\mathcal{T}\mathcal{S}, \quad \Phi_2 := \mathcal{A}\mathcal{B}\mathcal{M}\mathcal{T}\mathcal{S} \, . \qquad (4.2.72)
$$

We first prove that W_1, W_2 and their inverses are linear bijections of H^s. We take

$$
\gamma \leq \gamma_0/2 \, , \quad \tau \geq \tau_0 \, . \qquad (4.2.73)
$$

Lemma 4.2.5. *Let $s_0 \le s \le q - \sigma - \beta - 3$ where β is defined in* (4.2.1) *and σ in* (4.1.59). *Let $u := u(\lambda)$ satisfy $\|u\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma+\beta+3} \le 1$, and $\varepsilon\gamma^{-1} \le \delta$ be small enough. Then W_i, $i = 1, 2$, satisfy, $\forall \lambda \in \Lambda^{2\gamma}_\infty(u)$,*

$$\|W_i h\|_s + \big\|W_i^{-1} h\big\|_s \le C(s)\big(\|h\|_s + \|u\|_{s+\sigma+\beta} \|h\|_{s_0} \big)\,, \tag{4.2.74}$$

$$\|W_i h\|^{\mathrm{Lip}(\gamma)}_s + \big\|W_i^{-1} h\big\|^{\mathrm{Lip}(\gamma)}_s \le C(s)\big(\|h\|^{\mathrm{Lip}(\gamma)}_{s+3} + \|u\|^{\mathrm{Lip}(\gamma)}_{s+\sigma+\beta+3} \|h\|^{\mathrm{Lip}(\gamma)}_{s_0+3} \big)\,. \tag{4.2.75}$$

In the reversible case (i.e. (1.2.15) *holds), W_i, W_i^{-1}, $i = 1, 2$ are reversibility-preserving.*

Proof. The bound (4.2.74), resp. (4.2.75), follows by (4.2.8), (4.1.61), resp. (4.1.63), (3.1.11) and Lemma A.0.11. In the reversible case $W_i^{\pm 1}$ are reversibility preserving because $\Phi_i^{\pm 1}$, $\Phi_\infty^{\pm 1}$ are reversibility preserving. □

By (4.2.72) we are reduced to show that, $\forall \lambda \in \Lambda^{2\gamma}_\infty(u)$, the operator

$$\mathcal{L}_\infty := \mathrm{diag}_{j \in \mathbb{Z}} \{ \mathrm{i}\lambda\bar\omega \cdot l + \mu^\infty_j(\lambda) \}\,, \quad \mu^\infty_j(\lambda) = -\mathrm{i}\big(m_3(\lambda) j^3 - m_1(\lambda) j\big) + r^\infty_j(\lambda)$$

is invertible, assuming (1.2.8) or the reversibility condition (1.2.15).

We introduce the following notation:

$$\Pi_C u := \frac{1}{(2\pi)^{\nu+1}} \int_{\mathbb{T}^{\nu+1}} u(\varphi, x)\, d\varphi dx, \quad \mathbb{P}u := u - \Pi_C u, \quad H^s_{00} := \{ u \in H^s(\mathbb{T}^{\nu+1}) : \Pi_C u = 0 \}. \tag{4.2.76}$$

If (1.2.8) holds, then the linearized operator $\mathcal{L}$ in (4.1.1) satisfies

$$\mathcal{L} : H^{s+3} \to H^s_{00} \tag{4.2.77}$$

(for $s_0 \le s \le q - 1$). In the reversible case (1.2.15)

$$\mathcal{L} : X \cap H^{s+3} \to Y \cap H^s \subset H^s_{00}\,. \tag{4.2.78}$$

Lemma 4.2.6. *Assume either* (1.2.8) *or the reversibility condition* (1.2.15). *Then the eigenvalue*

$$\mu^\infty_0(\lambda) = r^\infty_0(\lambda) = 0\,, \quad \forall \lambda \in \Lambda^{2\gamma}_\infty(u)\,. \tag{4.2.79}$$

Proof. Assume (1.2.8). If $r^\infty_0 \neq 0$ then there exists a solution of $\mathcal{L}_\infty w = 1$, which is $w = 1/r^\infty_0$. Therefore, by (4.2.72),

$$\mathcal{L} W_2 [1/r^\infty_0] = \mathcal{L} W_2 w = W_1 \mathcal{L}_\infty w = W_1[1]$$

which is a contradiction because $\Pi_C W_1[1] \neq 0$, for $\varepsilon\gamma^{-1}$ small enough, but the average $\Pi_C \mathcal{L} W_2[1/r^\infty_0] = 0$ by (4.2.77). In the reversible case $r^\infty_0 = 0$ was proved in remark 4.2.3. □

As a consequence of (4.2.79), the definition of $\Lambda^{2\gamma}_\infty$ in (4.2.6) (just specializing (4.2.6) with $k = 0$), and (1.2.2) (with γ and τ as in (4.2.73)), we deduce also the *first* order Melnikov non-resonance conditions

$$\forall \lambda \in \Lambda^{2\gamma}_\infty\,, \qquad \big| \mathrm{i}\lambda\bar\omega \cdot l + \mu^\infty_j(\lambda) \big| \ge 2\gamma \frac{\langle j \rangle^3}{\langle l \rangle^\tau}, \quad \forall (l, j) \neq (0, 0)\,. \tag{4.2.80}$$

Lemma 4.2.7. (Invertibility of $\mathcal{L}_\infty$) *For all $\lambda \in \Lambda^{2\gamma}_\infty(u)$, for all $g \in H^s_{00}$ the equation $\mathcal{L}_\infty w = g$ has the unique solution with zero average*

$$\mathcal{L}_\infty^{-1}\, g(\varphi, x) := \sum_{(l,j)\neq(0,0)} \frac{g_{lj}}{\mathrm{i}\lambda\bar\omega \cdot l + \mu_j^\infty(\lambda)}\, e^{\mathrm{i}(l\cdot\varphi + jx)}. \tag{4.2.81}$$

For all Lipschitz family $g := g(\lambda) \in H^s_{00}$ we have

$$\left\|\mathcal{L}_\infty^{-1} g\right\|_s^{\mathrm{Lip}(\gamma)} \le C\gamma^{-1} \|g\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)}\,. \tag{4.2.82}$$

In the reversible case, if $g \in Y$ then $\mathcal{L}_\infty^{-1} g \in X$.

Proof. For all $\lambda \in \Lambda^{2\gamma}_\infty(u)$, by (4.2.80), formula (4.2.81) is well defined and

$$\left\|\mathcal{L}_\infty^{-1}(\lambda) g(\lambda)\right\|_s \lessdot \gamma^{-1} \|g(\lambda)\|_{s+\tau}\,. \tag{4.2.83}$$

Now we prove the Lipschitz estimate. For $\lambda_1, \lambda_2 \in \Lambda^{2\gamma}_\infty(u)$

$$\mathcal{L}_\infty^{-1}(\lambda_1) g(\lambda_1) - \mathcal{L}_\infty^{-1}(\lambda_2) g(\lambda_2) = \mathcal{L}_\infty^{-1}(\lambda_1)[g(\lambda_1) - g(\lambda_2)] + \big(\mathcal{L}_\infty^{-1}(\lambda_1) - \mathcal{L}_\infty^{-1}(\lambda_2)\big) g(\lambda_2)\,. \tag{4.2.84}$$

By (4.2.83)

$$\gamma \|\mathcal{L}_\infty^{-1}(\lambda_1)[g(\lambda_1) - g(\lambda_2)]\|_s \lessdot \|g(\lambda_1) - g(\lambda_2)\|_{s+\tau} \le \gamma^{-1} \|g\|_{s+\tau}^{\mathrm{Lip}(\gamma)} |\lambda_1 - \lambda_2|\,. \tag{4.2.85}$$

Now we estimate the second term of (4.2.84). We simplify notations writing $g := g(\lambda_2)$ and $\delta_{lj} := \mathrm{i}\lambda\bar\omega \cdot l + \mu_j^\infty$.

$$\big(\mathcal{L}_\infty^{-1}(\lambda_1) - \mathcal{L}_\infty^{-1}(\lambda_2)\big) g = \sum_{(l,j)\neq(0,0)} \frac{\delta_{lj}(\lambda_2) - \delta_{lj}(\lambda_1)}{\delta_{lj}(\lambda_1)\delta_{lj}(\lambda_2)}\, g_{lj} e^{\mathrm{i}(l\cdot\varphi + jx)}\,. \tag{4.2.86}$$

The bound (4.2.5) imply $|\mu_j^\infty|^{\mathrm{lip}} \lessdot \varepsilon\gamma^{-1}|j|^3 \lessdot |j|^3$ and, using also (4.2.80),

$$\gamma \frac{|\delta_{lj}(\lambda_2) - \delta_{lj}(\lambda_1)|}{|\delta_{lj}(\lambda_1)||\delta_{lj}(\lambda_2)|} \ \lessdot\ \frac{(|l| + |j|^3)\langle l\rangle^{2\tau}}{\gamma\langle j\rangle^6} |\lambda_2 - \lambda_1| \lessdot \langle l\rangle^{2\tau+1}\gamma^{-1}|\lambda_2 - \lambda_1|\,. \tag{4.2.87}$$

Then (4.2.86) and (4.2.87) imply $\gamma\|(\mathcal{L}_\infty^{-1}(\lambda_2) - \mathcal{L}_\infty^{-1}(\lambda_1))g\|_s \lessdot \gamma^{-1}\|g\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)}|\lambda_2 - \lambda_1|$ that, finally, with (4.2.83), (4.2.85), prove (4.2.82). The last statement follows by the property (4.2.37). $\square$

In order to solve the equation $\mathcal{L}h = f$ we first prove the following lemma.

Lemma 4.2.8. *Let $s_0 + \tau + 3 \le s \le q - \sigma - \beta - 3$. Under the assumption* (1.2.8) *we have*

$$W_1(H^s_{00}) = H^s_{00}\,, \quad W_1^{-1}(H^s_{00}) = H^s_{00}\,. \tag{4.2.88}$$

Proof. It is sufficient to prove that $W_1(H^s_{00}) = H^s_{00}$ because the second equality of (4.2.88) follows applying the isomorphism W_1^{-1}. Let us give the proof of the inclusion

$$W_1(H^s_{00}) \subseteq H^s_{00} \tag{4.2.89}$$

(which is essentially algebraic). For any $g \in H^s_{00}$, let $w(\varphi, x) := \mathcal{L}_\infty^{-1} g \in H^{s-\tau}_{00}$ defined in (4.2.81). Then $h := W_2 w \in H^{s-\tau}$ satisfies

$$\mathcal{L}h \overset{(4.2.72)}{=} W_1 \mathcal{L}_\infty W_2^{-1} h = W_1 \mathcal{L}_\infty w = W_1 g \,.$$

By (4.2.77) we deduce that $W_1 g = \mathcal{L}h \in H^{s-\tau-3}_{00}$. Since $W_1 g \in H^s$ by Lemma 4.2.5, we conclude $W_1 g \in H^s \cap H^{s-\tau-3}_{00} = H^s_{00}$. The proof of (4.2.89) is complete.

It remains to prove that $H^s_{00} \setminus W_1(H^s_{00}) = \emptyset$. By contradiction, let $f \in H^s_{00} \setminus W_1(H^s_{00})$. Let $g := W_1^{-1} f \in H^s$ by Lemma 4.2.5. Since $W_1 g = f \notin W_1(H^s_{00})$, it follows that $g \notin H^s_{00}$ (otherwise it contradicts (4.2.89)), namely $c := \Pi_C g \neq 0$. Decomposing $g = c + \mathbb{P}g$ (recall (4.2.76)) and applying W_1, we get $W_1 g = c W_1[1] + W_1 \mathbb{P} g$. Hence

$$W_1[1] = c^{-1}(W_1 g - W_1 \mathbb{P} g) \in H^s_{00}$$

because $W_1 g = f \in H^s_{00}$ and $W_1 \mathbb{P} g \in W_1(H^s_{00}) \subseteq H^s_{00}$ by (4.2.89). However, $\Pi_C W_1[1] \neq 0$, a contradiction. $\square$

Remark 4.2.7. *In the Hamiltonian case (which always satisfies* (1.2.8)*), the $W_i(\varphi)$ are maps of (a subspace of) H^1_0 so that Lemma 4.2.8 is automatic, and there is no need of Lemma 4.2.6.*

We may now prove the main result of Sections 4.1 and 4.2.

Theorem 4.2.3. (Right inverse of $\mathcal{L}$) *Let*

$$\tau_1 := 2\tau + 7, \quad \mu := 4\tau + \sigma + \beta + 14 \,, \tag{4.2.90}$$

where σ, β are defined in (4.1.59), (4.2.1) *respectively. Let $u(\lambda)$, $\lambda \in \Lambda_o \subseteq \Lambda$, be a Lipschitz family with*

$$\|u\|^{\mathrm{Lip}(\gamma)}_{s_0+\mu} \leq 1 \,. \tag{4.2.91}$$

Then there exists δ (depending on the data of the problem) such that if

$$\varepsilon\gamma^{-1} \leq \delta \,,$$

and condition (1.2.8)*, resp. the reversibility condition* (1.2.15)*, holds, then for all $\lambda \in \Lambda^{2\gamma}_\infty(u)$ defined in* (4.2.6)*, the linearized operator $\mathcal{L} := \mathcal{L}(\lambda, u(\lambda))$ (see* (4.1.1)*) admits a right inverse on H^s_{00}, resp. $Y \cap H^s$. More precisely, for $\mathfrak{s}_0 \leq s \leq q - \mu$, for all Lipschitz family $f(\lambda) \in H^s_{00}$, resp. $Y \cap H^s$, the function*

$$h := \mathcal{L}^{-1} f := W_2 \mathcal{L}_\infty^{-1} W_1^{-1} f \tag{4.2.92}$$

is a solution of $\mathcal{L}h = f$. In the reversible case, $\mathcal{L}^{-1} f \in X$. Moreover

$$\|\mathcal{L}^{-1} f\|^{\mathrm{Lip}(\gamma)}_s \leq C(s)\gamma^{-1}\Big(\|f\|^{\mathrm{Lip}(\gamma)}_{s+\tau_1} + \|u\|^{\mathrm{Lip}(\gamma)}_{s+\mu}\|f\|^{\mathrm{Lip}(\gamma)}_{s_0}\Big) \,. \tag{4.2.93}$$

Proof. Given $f \in H^s_{00}$, resp. $f \in Y \cap H^s$, with s like in Lemma 4.2.8, the equation $\mathcal{L}h = f$ can be solved for h because $\Pi_C f = 0$. Indeed, by (4.2.72), the equation $\mathcal{L}h = f$ is equivalent to $\mathcal{L}_\infty W_2^{-1} h = W_1^{-1} f$ where $W_1^{-1} f \in H^s_{00}$ by Lemma 4.2.8, resp. $W_1^{-1} f \in Y \cap H^s$ being W_1^{-1}

reversibility-preserving (Lemma 4.2.5). As a consequence, by Lemma 4.2.7, all the solutions of $\mathcal{L}h = f$ are

$$h = cW_2[1] + W_2\mathcal{L}_\infty^{-1}W_1^{-1}f, \quad c \in \mathbb{R}\,. \tag{4.2.94}$$

The solution (4.2.92) is the one with $c = 0$. In the reversible case, the fact that $\mathcal{L}^{-1}f \in X$ follows by (4.2.92) and the fact that W_i, W_i^{-1} are reversibility-preserving and $\mathcal{L}_\infty^{-1} : Y \to X$, see Lemma 4.2.7.

Finally (4.2.75), (4.2.82), (4.2.91) imply

$$\|\mathcal{L}^{-1}f\|_s^{\mathrm{Lip}(\gamma)} \le C(s)\gamma^{-1}\big(\|f\|_{s+2\tau+7}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+2\tau+\sigma+\beta+7}^{\mathrm{Lip}(\gamma)}\|f\|_{s_0+2\tau+7}^{\mathrm{Lip}(\gamma)}\big)$$

and (4.2.93) follows using (A.0.2) with $b_0 = s_0$, $a_0 := s_0 + 2\tau + \sigma + \beta + 7$, $q = 2\tau + 7$, $p = s - s_0$. $\square$

In the next Section we apply Theorem 4.2.3 to deduce tame estimates for the inverse linearized operators at any step of the Nash-Moser scheme. The approximate solutions along the iteration will satisfy (4.2.91).

4.3 The Nash-Moser iteration

We define the finite-dimensional subspaces of trigonometric polynomials

$$H_n := \Big\{u \in L^2(\mathbb{T}^{\nu+1}) : u(\varphi, x) = \sum_{|(l,j)| \le N_n} u_{lj}e^{\mathrm{i}(l\cdot\varphi + jx)}\Big\}$$

where $N_n := N_0^{\chi^n}$ (see (4.2.12)) and the corresponding orthogonal projectors

$$\Pi_n := \Pi_{N_n} : L^2(\mathbb{T}^{\nu+1}) \to H_n\,, \quad \Pi_n^\perp := I - \Pi_n\,.$$

The following smoothing properties hold: for all $\alpha, s \ge 0$,

$$\|\Pi_n u\|_{s+\alpha}^{\mathrm{Lip}(\gamma)} \le N_n^\alpha \|u\|_s^{\mathrm{Lip}(\gamma)}, \ \ \forall u(\lambda) \in H^s\,; \quad \|\Pi_n^\perp u\|_s^{\mathrm{Lip}(\gamma)} \le N_n^{-\alpha}\|u\|_{s+\alpha}^{\mathrm{Lip}(\gamma)}, \ \ \forall u(\lambda) \in H^{s+\alpha}, \tag{4.3.1}$$

where the function $u(\lambda)$ depends on the parameter λ in a Lipschitz way. The bounds (4.3.1) are the classical smoothing estimates for truncated Fourier series, which also hold with the norm $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$ defined in (3.0.3).

Let

$$F(u) := F(\lambda, u) := \lambda\bar\omega \cdot \partial_\varphi u + u_{xxx} + \varepsilon f(\varphi, x, u, u_x, u_{xx}, u_{xxx})\,. \tag{4.3.2}$$

We define the constants

$$\kappa := 28 + 6\mu, \qquad \beta_1 := 50 + 11\mu, \tag{4.3.3}$$

where μ is the loss of regularity in (4.2.90).

Theorem 4.3.1. (Nash-Moser) *Assume that $f \in C^q$, $q \ge s_0 + \mu + \beta_1$, satisfies the assumptions of Theorem 1.2.1 or Theorem 1.2.3. Let $0 < \gamma \le \min\{\gamma_0, 1/48\}$, $\tau > \nu + 1$. Then there exist $\delta > 0$, $C_* > 0$, $N_0 \in \mathbb{N}$ (that may depend also on τ) such that, if $\varepsilon\gamma^{-1} < \delta$, then, for all $n \ge 0$:*

$(\mathcal{P}1)_n$ *there exists a function* $u_n : \mathcal{G}_n \subseteq \Lambda \to H_n$, $\lambda \mapsto u_n(\lambda)$, *with* $\|u_n\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \le 1$, $u_0 := 0$, *where* $\mathcal{G}_n$ *are Cantor like subsets of* $\Lambda := [1/2, 3/2]$ *defined inductively by:* $\mathcal{G}_0 := \Lambda$,

$$\begin{aligned} \mathcal{G}_{n+1} \;\; := \;\; & \Big\{ \lambda \in \mathcal{G}_n \; : \; |\mathrm{i}\omega \cdot l + \mu_j^\infty(u_n) - \mu_k^\infty(u_n)| \ge \frac{2\gamma_n |j^3 - k^3|}{\langle l \rangle^\tau} \,, \\ & \quad \forall j, k \in \mathbb{Z}, \; l \in \mathbb{Z}^\nu \Big\} \,, \end{aligned} \tag{4.3.4}$$

where $\gamma_n := \gamma(1 + 2^{-n})$. *In the reversible case, namely* (1.2.15) *holds, then* $u_n(\lambda) \in X$.

The difference $h_n := u_n - u_{n-1}$, *where, for convenience,* $h_0 := 0$, *satisfy*

$$\|h_n\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \le C_* \varepsilon \gamma^{-1} N_n^{-\sigma_1} \,, \quad \sigma_1 := 18 + 2\mu \,. \tag{4.3.5}$$

$(\mathcal{P}2)_n$ $\|F(u_n)\|_{s_0}^{\mathrm{Lip}(\gamma)} \le C_* \varepsilon N_n^{-\kappa}$.

$(\mathcal{P}3)_n$ (High norms). $\|u_n\|_{s_0+\beta_1}^{\mathrm{Lip}(\gamma)} \le C_* \varepsilon \gamma^{-1} N_n^{\kappa}$ *and* $\|F(u_n)\|_{s_0+\beta_1}^{\mathrm{Lip}(\gamma)} \le C_* \varepsilon N_n^{\kappa}$.

$(\mathcal{P}4)_n$ (Measure). *The measure of the Cantor like sets satisfy*

$$|\mathcal{G}_0 \setminus \mathcal{G}_1| \le C_* \gamma \,, \quad \big|\mathcal{G}_n \setminus \mathcal{G}_{n+1}\big| \le \gamma C_* N_n^{-1} \,, \; n \ge 1. \tag{4.3.6}$$

All the Lip norms are defined on $\mathcal{G}_n$.

Proof. The proof of Theorem 4.3.1 is split into several steps. For simplicity, we denote $\| \, \|^{\mathrm{Lip}}$ by $\| \, \|$.

Step 1: *prove* $(\mathcal{P}1, 2, 3)_0$. $(\mathcal{P}1)_0$ and the first inequality of $(\mathcal{P}3)_0$ are trivial because $u_0 = h_0 = 0$. $(\mathcal{P}2)_0$ and the second inequality of $(\mathcal{P}3)_0$ follow with $C_* \ge \max\{\|f(0)\|_{s_0} N_0^{\kappa}, \, \|f(0)\|_{s_0+\beta_1} N_0^{-\kappa}\}$.

Step 2: *assume that* $(\mathcal{P}1, 2, 3)_n$ *hold for some* $n \ge 0$, *and prove* $(\mathcal{P}1, 2, 3)_{n+1}$. By $(\mathcal{P}1)_n$ we know that $\|u_n\|_{s_0+\mu} \le 1$, namely condition (4.2.91) is satisfied. Hence, for $\varepsilon\gamma^{-1}$ small enough, Theorem 4.2.3 applies. Then, for all $\lambda \in \mathcal{G}_{n+1}$ defined in (4.3.4), the linearized operator

$$\mathcal{L}_n(\lambda) := \mathcal{L}(\lambda, u_n(\lambda)) = F'(\lambda, u_n(\lambda))$$

(see (4.1.1)) admits a right inverse for all $h \in H_{00}^s$, if condition (1.2.8) holds, respectively for $h \in Y \cap H^s$ if the reversibility condition (1.2.15) holds. Moreover (4.2.93) gives the estimates

$$\|\mathcal{L}_n^{-1} h\|_s \le_s \gamma^{-1} \Big(\|h\|_{s+\tau_1} + \|u_n\|_{s+\mu} \|h\|_{s_0} \Big) \,, \quad \forall h(\lambda), \tag{4.3.7}$$

$$\|\mathcal{L}_n^{-1} h\|_{s_0} \le \gamma^{-1} N_{n+1}^{\tau_1} \|h\|_{s_0} \,, \quad \forall h(\lambda) \in H_{n+1} \,, \tag{4.3.8}$$

(use (4.3.1) and $\|u_n\|_{s_0+\mu} \le 1$), for all Lipschitz map $h(\lambda)$. Then, for all $\lambda \in \mathcal{G}_{n+1}$, we define

$$u_{n+1} := u_n + h_{n+1} \in H_{n+1} \,, \quad h_{n+1} := -\Pi_{n+1} \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) \,, \tag{4.3.9}$$

which is well defined because, if condition (1.2.8) holds then $\Pi_{n+1} F(u_n) \in H_{00}^s$, and, respectively, if (1.2.15) holds, then $\Pi_{n+1} F(u_n) \in Y \cap H^s$ (hence in both cases $\mathcal{L}_n^{-1} \Pi_{n+1} F(u_n)$ exists). Note also that in the reversible case $h_{n+1} \in X$ and so $u_{n+1} \in X$.

Recalling (4.3.2) and that $\mathcal{L}_n := F'(u_n)$, we write

$$F(u_{n+1}) = F(u_n) + \mathcal{L}_n h_{n+1} + \varepsilon Q(u_n, h_{n+1}) \tag{4.3.10}$$

where

$$Q(u_n, h_{n+1}) := \mathcal{N}(u_n + h_{n+1}) - \mathcal{N}(u_n) - \mathcal{N}'(u_n)h_{n+1}, \quad \mathcal{N}(u) := f(\varphi, x, u, u_x, u_{xx}, u_{xxx}).$$

With this definition,

$$F(u) = L_\omega u + \varepsilon \mathcal{N}(u), \quad F'(u)h = L_\omega h + \varepsilon \mathcal{N}'(u)h, \quad L_\omega := \omega \cdot \partial_\varphi + \partial_{xxx}.$$

By (4.3.10) and (4.3.9) we have

$$\begin{aligned}
F(u_{n+1}) &= F(u_n) - \mathcal{L}_n \Pi_{n+1} \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + \varepsilon Q(u_n, h_{n+1}) \\
&= \Pi_{n+1}^\perp F(u_n) + \mathcal{L}_n \Pi_{n+1}^\perp \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + \varepsilon Q(u_n, h_{n+1}) \\
&= \Pi_{n+1}^\perp F(u_n) + \Pi_{n+1}^\perp \mathcal{L}_n \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + [\mathcal{L}_n, \Pi_{n+1}^\perp] \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + \varepsilon Q(u_n, h_{n+1}) \\
&= \Pi_{n+1}^\perp F(u_n) + \varepsilon [\mathcal{N}'(u_n), \Pi_{n+1}^\perp] \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + \varepsilon Q(u_n, h_{n+1})
\end{aligned} \tag{4.3.11}$$

where we have gained an extra ε from the commutator

$$[\mathcal{L}_n, \Pi_{n+1}^\perp] = [L_\omega + \varepsilon \mathcal{N}'(u_n), \Pi_{n+1}^\perp] = \varepsilon [\mathcal{N}'(u_n), \Pi_{n+1}^\perp] .$$

Lemma 4.3.1. *Set*

$$U_n := \|u_n\|_{s_0+\beta_1} + \gamma^{-1} \|F(u_n)\|_{s_0+\beta_1} , \qquad w_n := \gamma^{-1} \|F(u_n)\|_{s_0} . \tag{4.3.12}$$

There exists $C_0 := C(\tau_1, \mu, \nu, \beta_1) > 0$ such that

$$w_{n+1} \le C_0 N_{n+1}^{-\beta_1+\mu'} U_n (1 + w_n) + C_0 N_{n+1}^{6+2\mu} w_n^2, \qquad U_{n+1} \le C_0 N_{n+1}^{9+2\mu} (1 + w_n)^2 \, U_n . \tag{4.3.13}$$

Proof. The operators $\mathcal{N}'(u_n)$ and $Q(u_n, \cdot)$ satisfy the following tame estimates:

$$\|Q(u_n, h)\|_s \le_s \|h\|_{s_0+3} \Big(\|h\|_{s+3} + \|u_n\|_{s+3} \|h\|_{s_0+3} \Big) \quad \forall h(\lambda), \tag{4.3.14}$$

$$\|Q(u_n, h)\|_{s_0} \le N_{n+1}^6 \|h\|_{s_0}^2 \quad \forall h(\lambda) \in H_{n+1}, \tag{4.3.15}$$

$$\|\mathcal{N}'(u_n)h\|_s \le_s \|h\|_{s+3} + \|u_n\|_{s+3} \|h\|_{s_0+3} \quad \forall h(\lambda), \tag{4.3.16}$$

where $h(\lambda)$ depends on the parameter λ in a Lipschitz way. The bounds (4.3.14) and (4.3.16) follow by Lemma A.0.8(i) and Lemma A.0.9. (4.3.15) is simply (4.3.14) at $s = s_0$, using that $\|u_n\|_{s_0+3} \le 1$, $u_n, h_{n+1} \in H_{n+1}$ and the smoothing (4.3.1).

By (4.3.7) and (4.3.16), the term (in (4.3.11)) $R_n := [\mathcal{N}'(u_n), \Pi_{n+1}^\perp] \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n)$ satisfies, using also that $u_n \in H_n$ and (4.3.1),

$$\|R_n\|_s \le_s \gamma^{-1} N_{n+1}^{\mu'} \Big(\|F(u_n)\|_s + \|u_n\|_s \|F(u_n)\|_{s_0} \Big), \quad \mu' := 3 + \mu, \tag{4.3.17}$$

$$\|R_n\|_{s_0} \le_{s_0+\beta_1} \gamma^{-1} N_{n+1}^{-\beta_1+\mu'} \Big(\|F(u_n)\|_{s_0+\beta_1} + \|u_n\|_{s_0+\beta_1} \|F(u_n)\|_{s_0} \Big), \tag{4.3.18}$$

because $\mu \geq \tau_1 + 3$. In proving (4.3.17) and (4.3.18), we have simply estimated $\mathcal{N}'(u_n)\Pi_{n+1}^{\perp}$ and $\Pi_{n+1}^{\perp}\mathcal{N}'(u_n)$ separately, without using the commutator structure.

From the definition (4.3.9) of h_{n+1}, using (4.3.7), (4.3.8) and (4.3.1), we get

$$\|h_{n+1}\|_{s_0+\beta_1} \leq_{s_0+\beta_1} \gamma^{-1} N_{n+1}^{\mu} \Big(\|F(u_n)\|_{s_0+\beta_1} + \|u_n\|_{s_0+\beta_1} \|F(u_n)\|_{s_0} \Big), \tag{4.3.19}$$

$$\|h_{n+1}\|_{s_0} \leq_{s_0} \gamma^{-1} N_{n+1}^{\mu} \|F(u_n)\|_{s_0} \tag{4.3.20}$$

because $\mu \geq \tau_1$. Then

$$\begin{aligned} \|u_{n+1}\|_{s_0+\beta_1} &\overset{(4.3.9)}{\leq} \|u_n\|_{s_0+\beta_1} + \|h_{n+1}\|_{s_0+\beta_1} \\ &\overset{(4.3.19)}{\leq}{}_{s_0+\beta_1} \|u_n\|_{s_0+\beta_1} \Big(1 + \gamma^{-1} N_{n+1}^{\mu} \|F(u_n)\|_{s_0} \Big) \qquad (4.3.21) \\ &\quad + \gamma^{-1} N_{n+1}^{\mu} \|F(u_n)\|_{s_0+\beta_1}. \qquad (4.3.22) \end{aligned}$$

Formula (4.3.11) for $F(u_{n+1})$, and (4.3.18), (4.3.15), (4.3.20), $\varepsilon\gamma^{-1} \leq 1$, (4.3.1), imply

$$\begin{aligned} \|F(u_{n+1})\|_{s_0} \leq_{s_0+\beta_1} & N_{n+1}^{-\beta_1+\mu'} \Big(\|F(u_n)\|_{s_0+\beta_1} + \|u_n\|_{s_0+\beta_1} \|F(u_n)\|_{s_0} \Big) \\ & + \varepsilon \gamma^{-2} N_{n+1}^{6+2\mu} \|F(u_n)\|_{s_0}^2. \end{aligned} \tag{4.3.23}$$

Similarly, using the "high norm" estimates (4.3.17), (4.3.14), (4.3.19), (4.3.20), $\varepsilon\gamma^{-1} \leq 1$ and (4.3.1),

$$\begin{aligned} \|F(u_{n+1})\|_{s_0+\beta_1} \leq_{s_0+\beta_1} & N_{n+1}^{\mu'} \Big(\|F(u_n)\|_{s_0+\beta_1} + \|u_n\|_{s_0+\beta_1} \|F(u_n)\|_{s_0} \Big) \\ & + N_{n+1}^{9+2\mu} \gamma^{-1} \|F(u_n)\|_{s_0} \Big(\|F(u_n)\|_{s_0+\beta_1} + \|u_n\|_{s_0+\beta_1} \|F(u_n)\|_{s_0} \Big). \end{aligned} \tag{4.3.24}$$

By (4.3.22), (4.3.23) and (4.3.24) we deduce (4.3.13). □

By $(\mathcal{P}2)_n$ we deduce, for $\varepsilon\gamma^{-1}$ small, that (recall the definition on w_n in (4.3.12))

$$w_n \leq \varepsilon \gamma^{-1} C_* N_n^{-\kappa} \leq 1, \tag{4.3.25}$$

Then, by the second inequality in (4.3.13), (4.3.25), $(\mathcal{P}3)_n$ (recall the definition on U_n in (4.3.12)) and the choice of κ in (4.3.3), we deduce $U_{n+1} \leq C_* \varepsilon \gamma^{-1} N_{n+1}^{\kappa}$, for N_0 large enough. This proves $(\mathcal{P}3)_{n+1}$.

Next, by the first inequality in (4.3.13), (4.3.25), $(\mathcal{P}2)_n$ (recall the definition on w_n in (4.3.12)) and (4.3.3), we deduce $w_{n+1} \leq C_* \varepsilon \gamma^{-1} N_{n+1}^{\kappa}$, for N_0 large, $\varepsilon\gamma^{-1}$ small. This proves $(\mathcal{P}2)_{n+1}$.

The bound (4.3.5) at the step $n+1$ follows by (4.3.20) and $(\mathcal{P}2)_n$ (and (4.3.3)). Then

$$\|u_{n+1}\|_{s_0+\mu} \leq \|u_0\|_{s_0+\mu} + \sum_{k=1}^{n+1} \|h_k\|_{s_0+\mu} \leq \sum_{k=1}^{\infty} C_* \varepsilon \gamma^{-1} N_k^{-\sigma_1} \leq 1$$

for $\varepsilon\gamma^{-1}$ small enough. As a consequence $(\mathcal{P}1,2,3)_{n+1}$ hold.

STEP 3: *prove* $(\mathcal{P}4)_n$, $n \geq 0$. For all $n \geq 0$,

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} = \bigcup_{l \in \mathbb{Z}^\nu, j,k \in \mathbb{Z}} R_{ljk}(u_n) \tag{4.3.26}$$

where

$$R_{ljk}(u_n) := \left\{ \lambda \in \mathcal{G}_n : \left| \mathrm{i}\lambda\bar{\omega}\cdot l + \mu_j^\infty(\lambda, u_n(\lambda)) - \mu_k^\infty(\lambda, u_n(\lambda)) \right| < \frac{2\gamma_n |j^3 - k^3|}{\langle l \rangle^{-\tau}} \right\} . \tag{4.3.27}$$

Notice that, by the definition (4.3.27), $R_{ljk}(u_n) = \emptyset$ for $j = k$. Then we can suppose in the sequel that $j \neq k$. We divide the estimate into some lemmata.

Lemma 4.3.2. *For $\varepsilon\gamma^{-1}$ small enough, for all $n \geq 0$, $|l| \leq N_n$,*

$$R_{ljk}(u_n) \subseteq R_{ljk}(u_{n-1}). \tag{4.3.28}$$

Proof. We claim that, for all $j, k \in \mathbb{Z}$,

$$|(\mu_j^\infty - \mu_k^\infty)(u_n) - (\mu_j^\infty - \mu_k^\infty)(u_{n-1})| \leq C\varepsilon |j^3 - k^3| N_n^{-\alpha} , \quad \forall \lambda \in \mathcal{G}_n , \tag{4.3.29}$$

where $\mu_j^\infty(u_n) := \mu_j^\infty(\lambda, u_n(\lambda))$ and α is defined in (4.2.13). Before proving (4.3.29) we show how it implies (4.3.28). For all $j \neq k$, $|l| \leq N_n$, $\lambda \in \mathcal{G}_n$, by (4.3.29)

$$\begin{aligned} |\mathrm{i}\lambda\bar{\omega}\cdot l + \mu_j^\infty(u_n) - \mu_k^\infty(u_n)| &\geq |\mathrm{i}\lambda\bar{\omega}\cdot l + \mu_j^\infty(u_{n-1}) - \mu_k^\infty(u_{n-1})| - |(\mu_j^\infty - \mu_k^\infty)(u_n) - (\mu_j^\infty - \mu_k^\infty)(u_{n-1})| \\ &\geq 2\gamma_{n-1}|j^3 - k^3|\langle l \rangle^{-\tau} - C\varepsilon|j^3 - k^3| N_n^{-\alpha} \geq 2\gamma_n |j^3 - k^3| \langle l \rangle^{-\tau} \end{aligned}$$

for $C\varepsilon\gamma^{-1} N_n^{\tau-\alpha}\, 2^{n+1} \leq 1$ (recall that $\gamma_n := \gamma(1 + 2^{-n})$), which implies (4.3.28).
PROOF OF (4.3.29). By (4.2.4),

$$\begin{aligned} (\mu_j^\infty - \mu_k^\infty)(u_n) - (\mu_j^\infty - \mu_k^\infty)(u_{n-1}) = &-\mathrm{i}\big[m_3(u_n) - m_3(u_{n-1})\big](j^3 - k^3) + \mathrm{i}\big[m_1(u_n) - m_1(u_{n-1})\big](j-k) \\ &+ r_j^\infty(u_n) - r_j^\infty(u_{n-1}) - \big(r_k^\infty(u_n) - r_k^\infty(u_{n-1})\big) \end{aligned} \tag{4.3.30}$$

where $m_3(u_n) := m_3(\lambda, u_n(\lambda))$ and similarly for m_1, r_j^∞. We first apply Theorem 4.2.2-$(\mathbf{S4})_\nu$ with $\nu = n+1$, $\gamma = \gamma_{n-1}$, $\gamma - \rho = \gamma_n$, and u_1, u_2, replaced, respectively, by u_{n-1}, u_n, in order to conclude that

$$\Lambda_{n+1}^{\gamma_{n-1}}(u_{n-1}) \subseteq \Lambda_{n+1}^{\gamma_n}(u_n) . \tag{4.3.31}$$

The smallness condition in (4.2.26) is satisfied because $\sigma_2 < \mu$ (see definitions (4.2.13), (4.2.90)) and so

$$\varepsilon C N_n^\tau \|u_n - u_{n-1}\|_{s_0+\sigma_2} \leq \varepsilon C N_n^\tau \|u_n - u_{n-1}\|_{s_0+\mu} \overset{(4.3.5)}{\leq} \varepsilon^2\gamma^{-1} C C_* N_n^{\tau-\sigma_1} \leq \gamma_{n-1} - \gamma_n =: \rho = \gamma 2^{-n}$$

for $\varepsilon\gamma^{-1}$ small enough, because $\sigma_1 > \tau$ (see (4.3.5), (4.2.90)). Then, by the definitions (4.3.4) and (4.2.6), we have

$$\mathcal{G}_n := \mathcal{G}_{n-1} \cap \Lambda_\infty^{2\gamma_{n-1}}(u_{n-1}) \overset{(4.2.35)}{\subseteq} \bigcap_{\nu \geq 0} \Lambda_\nu^{\gamma_{n-1}}(u_{n-1}) \subset \Lambda_{n+1}^{\gamma_{n-1}}(u_{n-1}) \overset{(4.3.31)}{\subseteq} \Lambda_{n+1}^{\gamma_n}(u_n).$$

Next, for all $\lambda \in \mathcal{G}_n \subset \Lambda_{n+1}^{\gamma_{n-1}}(u_{n-1}) \cap \Lambda_{n+1}^{\gamma_n}(u_n)$ both $r_j^{n+1}(u_{n-1})$ and $r_j^{n+1}(u_n)$ are well defined, and we deduce by Theorem 4.2.2-$(\mathbf{S3})_\nu$ with $\nu = n+1$, that

$$|r_j^{n+1}(u_n) - r_j^{n+1}(u_{n-1})| \overset{(4.2.25)}{\lessdot} \varepsilon \|u_{n-1} - u_n\|_{s_0+\sigma_2} . \tag{4.3.32}$$

Moreover (4.2.34) (with $\nu = n+1$) and (4.1.67) imply that

$$\begin{aligned} |r_j^\infty(u_{n-1}) - r_j^{n+1}(u_{n-1})| + |r_j^\infty(u_n) - r_j^{n+1}(u_n)| &\lessdot \varepsilon(1 + \|u_{n-1}\|_{s_0+\beta+\sigma} + \|u_n\|_{s_0+\beta+\sigma})N_n^{-\alpha} \\ &\lessdot \varepsilon N_n^{-\alpha} \end{aligned} \tag{4.3.33}$$

because $\sigma + \beta < \mu$ and $\|u_{n-1}\|_{s_0+\mu} + \|u_n\|_{s_0+\mu} \le 2$ by $(\mathbf{S1})_{n-1}$ and $(\mathbf{S1})_n$. Therefore, for all $\lambda \in \mathcal{G}_n$, $\forall j \in \mathbb{Z}$,

$$\begin{aligned} \left|r_j^\infty(u_n) - r_j^\infty(u_{n-1})\right| &\le \left|r_j^{n+1}(u_n) - r_j^{n+1}(u_{n-1})\right| + |r_j^\infty(u_n) - r_j^{n+1}(u_n)| + |r_j^\infty(u_{n-1}) - r_j^{n+1}(u_{n-1})| \\ &\overset{(4.3.32),(4.3.33)}{\lessdot} \varepsilon\|u_n - u_{n-1}\|_{s_0+\sigma_2} + \varepsilon N_n^{-\alpha} \overset{(4.3.5)}{\lessdot} \varepsilon N_n^{-\alpha} \end{aligned} \tag{4.3.34}$$

because $\sigma_1 > \alpha$ (see (4.2.13), (4.3.5)). Finally (4.3.30), (4.3.34), (4.1.65), $\|u_n\|_{s_0+\mu} \le 1$, imply (4.3.29). $\square$

By definition, $R_{ljk}(u_n) \subset \mathcal{G}_n$ (see (4.3.27)) and, by (4.3.28), for all $|l| \le N_n$, we have $R_{ljk}(u_n) \subseteq R_{ljk}(u_{n-1})$. On the other hand $R_{ljk}(u_{n-1}) \cap \mathcal{G}_n = \emptyset$, see (4.3.4). As a consequence, $\forall |l| \le N_n$, $R_{ljk}(u_n) = \emptyset$, and

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} \overset{(4.3.26)}{\subseteq} \bigcup_{|l|>N_n, j,k\in\mathbb{Z}} R_{ljk}(u_n)\,, \quad \forall n \ge 1. \tag{4.3.35}$$

Lemma 4.3.3. *Let $n \ge 0$. If $R_{ljk}(u_n) \ne \emptyset$, then $|j^3 - k^3| \le 8|\bar\omega \cdot l|$.*

Proof. If $R_{ljk}(u_n) \ne \emptyset$ then there exists $\lambda \in \Lambda$ such that $|\mathrm{i}\lambda\bar\omega \cdot l + \mu_j^\infty(\lambda, u_n(\lambda)) - \mu_k^\infty(\lambda, u_n(\lambda))| < 2\gamma_n|j^3 - k^3|\langle l\rangle^{-\tau}$ and, therefore,

$$|\mu_j^\infty(\lambda, u_n(\lambda)) - \mu_k^\infty(\lambda, u_n(\lambda))| < 2\gamma_n|j^3 - k^3|\langle l\rangle^{-\tau} + 2|\bar\omega \cdot l|. \tag{4.3.36}$$

Moreover, by (4.2.4), (4.1.64), (4.2.5), for ε small enough,

$$|\mu_j^\infty - \mu_k^\infty| \ge |m_3||j^3 - k^3| - |m_1||j-k| - |r_j^\infty| - |r_k^\infty| \ge \frac{1}{2}|j^3 - k^3| - C\varepsilon|j-k| - C\varepsilon \ge \frac{1}{3}|j^3 - k^3| \tag{4.3.37}$$

if $j \ne k$. Since $\gamma_n \le 2\gamma$ for all $n \ge 0$, $\gamma \le 1/48$, by (4.3.36) and (4.3.37) we get

$$2|\bar\omega \cdot l| \ge \Big(\frac{1}{3} - \frac{4\gamma}{\langle l\rangle^\tau}\Big)|j^3 - k^3| \ \ge \frac{1}{4}|j^3 - k^3|$$

proving the Lemma. $\square$

Lemma 4.3.4. *For all $n \ge 0$,*

$$|R_{ljk}(u_n)| \le C\gamma\,\langle l\rangle^{-\tau}\,. \tag{4.3.38}$$

Proof. Consider the function $\phi : \Lambda \to \mathbb{C}$ defined by

$$\begin{aligned} \phi(\lambda) &:= \mathrm{i}\lambda\bar\omega \cdot l + \mu_j^\infty(\lambda) - \mu_k^\infty(\lambda) \\ &\overset{(4.2.4)}{=} \mathrm{i}\lambda\bar\omega \cdot l - \mathrm{i}\tilde m_3(\lambda)(j^3 - k^3) + \mathrm{i}\tilde m_1(\lambda)(j-k) + r_j^\infty(\lambda) - r_k^\infty(\lambda) \end{aligned}$$

where $\tilde{m}_3(\lambda)$, $\tilde{m}_1(\lambda)$, $r_j^\infty(\lambda)$, $\mu_j^\infty(\lambda)$, are defined for all $\lambda \in \Lambda$ and satisfy (4.2.5) by $\|u_n\|_{s_0+\mu,\mathcal{G}_n}^{\mathrm{Lip}(\gamma)} \leq 1$ (see $(\mathcal{P}1)_n$). Recalling $|\cdot|^{\mathrm{lip}} \leq \gamma^{-1}|\cdot|^{\mathrm{Lip}(\gamma)}$ and using (4.2.5)

$$|\mu_j^\infty - \mu_k^\infty|^{\mathrm{lip}} \leq |\tilde{m}_3|^{\mathrm{lip}}|j^3 - k^3| + |\tilde{m}_1|^{\mathrm{lip}}|j-k| + |r_j^\infty|^{\mathrm{lip}} + |r_k^\infty|^{\mathrm{lip}} \leq C\varepsilon\gamma^{-1}|j^3-k^3|\,. \tag{4.3.39}$$

Moreover Lemma 4.3.3 implies that, $\forall \lambda_1, \lambda_2 \in \Lambda$,

$$|\phi(\lambda_1)-\phi(\lambda_2)| \geq \big(|\bar{\omega}\cdot l| - |\mu_j^\infty - \mu_k^\infty|^{\mathrm{lip}}\big)|\lambda_1-\lambda_2| \overset{(4.3.39)}{\geq} \big(\frac{1}{8} - C\varepsilon\gamma^{-1}\big)|j^3-k^3||\lambda_1-\lambda_2| \geq \frac{|j^3-k^3|}{9}|\lambda_1-\lambda_2|$$

for $\varepsilon\gamma^{-1}$ small enough. Hence

$$|R_{ljk}(u_n)| \leq \frac{4\gamma_n|j^3-k^3|}{\langle l\rangle^\tau}\frac{9}{|j^3-k^3|} \leq \frac{72\gamma}{\langle l\rangle^\tau}\,,$$

which is (4.3.38). $\square$

Now we prove $(\mathcal{P}4)_0$. We observe that, for each fixed l, all the indices j,k such that $R_{ljk}(0) \neq \emptyset$ are confined in the ball $j^2+k^2 \leq 16|\bar{\omega}||l|$, because

$$|j^3-k^3| = |j-k||j^2+jk+k^2| \geq j^2+k^2-|jk| \geq \frac{1}{2}(j^2+k^2)\,, \quad \forall j,k \in \mathbb{Z},\ j \neq k,$$

and $|j^3-k^3| \leq 8|\bar{\omega}||l|$ by Lemma 4.3.3. As a consequence

$$|\mathcal{G}_0 \setminus \mathcal{G}_1| \overset{(4.3.26)}{=} \Big|\bigcup_{l,j,k} R_{ljk}(0)\Big| \leq \sum_{l\in\mathbb{Z}^\nu}\sum_{j^2+k^2\leq 16|\bar{\omega}||l|} |R_{ljk}(0)| \overset{(4.3.38)}{\lessdot} \sum_{l\in\mathbb{Z}^\nu}\gamma\langle l\rangle^{-\tau+1} = C\gamma$$

if $\tau > \nu+1$. Thus the first estimate in (4.3.6) is proved, taking a larger C_* if necessary.

Finally, $(\mathcal{P}4)_n$ for $n \geq 1$, follows by

$$\begin{aligned}|\mathcal{G}_n \setminus \mathcal{G}_{n+1}| &\overset{(4.3.35)}{\leq} \sum_{|l|>N_n\,|j|,|k|\leq C|l|^{1/2}} |R_{ljk}(u_n)| \overset{(4.3.38)}{\lessdot} \sum_{|l|>N_n\,|j|,|k|\leq C|l|^{1/2}} \gamma\langle l\rangle^{-\tau} \\ &\lessdot \sum_{|l|>N_n}\gamma\langle l\rangle^{-\tau+1} \lessdot \gamma N_n^{-\tau+\nu} \leq C\gamma N_n^{-1}\end{aligned}$$

and (4.3.6) is proved. The proof of Theorem 4.3.1 is complete. $\square$

4.3.1 Proof of Theorems 1.2.1, 1.2.2, 1.2.3, 1.2.4 and 1.2.5

Proof of Theorems 1.2.1, 1.2.2, 1.2.3. Assume that $f \in C^q$ satisfies the assumptions in Theorem 1.2.1 or in Theorem 1.2.3 with a smoothness exponent $q := q(\nu) \geq s_0+\mu+\beta_1$ which depends only on ν once we have fixed $\tau := \nu+2$ (recall that $s_0 := (\nu+2)/2$, β_1 is defined in (4.3.3) and μ in (4.2.90)).

For $\gamma = \varepsilon^a$, $a \in (0,1)$ the smallness condition $\varepsilon\gamma^{-1} = \varepsilon^{1-a} < \delta$ of Theorem 4.3.1 is satisfied. Hence on the Cantor set $\mathcal{G}_\infty := \cap_{n\geq 0}\mathcal{G}_n$, the sequence $u_n(\lambda)$ is well defined and converges in norm $\|\cdot\|_{s_0+\mu,\mathcal{G}_\infty}^{\mathrm{Lip}(\gamma)}$ (see (4.3.5)) to a solution $u_\infty(\lambda)$ of

$$F(\lambda, u_\infty(\lambda)) = 0 \quad \text{with} \quad \sup_{\lambda\in\mathcal{G}_\infty}\|u_\infty(\lambda)\|_{s_0+\mu} \leq C\varepsilon\gamma^{-1} = C\varepsilon^{1-a}\,,$$

namely $u_\infty(\lambda)$ is a solution of the perturbed equation (1.2.4) with $\omega = \lambda\bar\omega$. Moreover, by (4.3.6), the measure of the complementary set satisfies

$$|\Lambda \setminus \mathcal{G}_\infty| \le \sum_{n\ge 0} |\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \le C\gamma + \sum_{n\ge 1} \gamma C N_n^{-1} \le C\gamma = C\varepsilon^a ,$$

proving (1.2.9). The proof of Theorem 1.2.1 is complete. In order to finish the proof of Theorems 1.2.2 or 1.2.3, it remains to prove the linear stability of the solution, namely Theorem 1.2.5.

Proof of Theorem 1.2.4. Part (i) follows by (4.2.72), Lemma 4.2.5, Theorem 4.2.1 (applied to the solution $u_\infty(\lambda)$) with the exponents $\bar\sigma := \sigma + \beta + 3$, $\Lambda_\infty(u) := \Lambda_\infty^{2\gamma}(u)$, see (4.2.6). Part (ii) follows by the dynamical interpretation of the conjugation procedure, as explained in Section 3.4. Explicitely, in Sections 4.1 and 4.2, we have proved that

$$\mathcal{L} = \mathcal{A} B \rho W \mathcal{L}_\infty W^{-1} B^{-1} \mathcal{A}^{-1}, \quad W := \mathcal{M}\mathcal{T}\mathcal{S}\Phi_\infty .$$

By the arguments in Section 3.4 we deduce that a curve $h(t)$ in the phase space H^s_x is a solution of the dynamical system (1.2.21) if and only if the transformed curve

$$v(t) := W^{-1}(\omega t) B^{-1} \mathcal{A}^{-1}(\omega t) h(t) \tag{4.3.40}$$

(see notation (3.1.17), Lemma 4.1.3, (4.2.9)) is a solution of the constant coefficients dynamical system (1.2.22).

Proof of Theorem 1.2.5. If all μ_j are purely imaginary, the Sobolev norm of the solution $v(t)$ of (1.2.22) is constant in time, see (1.2.23). We now show that also the Sobolev norm of the solution $h(t)$ in (4.3.40) does not grow in time. For each $t \in \mathbb{R}$, $\mathcal{A}(\omega t)$ and $W(\omega t)$ are transformations of the phase space H^s_x that depend quasi-periodically on time, and satisfy, by (4.1.142), (4.1.144), (4.2.9),

$$\|\mathcal{A}^{\pm 1}(\omega t) g\|_{H^s_x} + \|W^{\pm 1}(\omega t) g\|_{H^s_x} \le C(s) \|g\|_{H^s_x} , \quad \forall t \in \mathbb{R}, \ \forall g = g(x) \in H^s_x, \tag{4.3.41}$$

where the constant $C(s)$ depends on $\|u\|_{s+\sigma+\beta+s_0} < +\infty$. Moreover, the transformation B is a quasi-periodic reparametrization of the time variable (see (3.4.5)), namely

$$B f(t) = f(\psi(t)) = f(\tau), \quad B^{-1} f(\tau) = f(\psi^{-1}(\tau)) = f(t) \quad \forall f : \mathbb{R} \to H^s_x, \tag{4.3.42}$$

where $\tau = \psi(t) := t + \alpha(\omega t)$, $t = \psi^{-1}(\tau) = \tau + \tilde\alpha(\omega\tau)$ and α, $\tilde\alpha$ are defined in Section 4.1.2. Thus

$$\begin{aligned}
\|h(t)\|_{H^s_x} &\overset{(4.3.40)}{=} \|\mathcal{A}(\omega t) B W(\omega t) v(t)\|_{H^s_x} \overset{(4.3.41)}{\le} C(s) \|B W(\omega t) v(t)\|_{H^s_x} \overset{(4.3.42)}{=} C(s) \|W(\omega\tau) v(\tau)\|_{H^s_x} \\
&\overset{(4.3.41)}{\le} C(s) \|v(\tau)\|_{H^s_x} \overset{(1.2.23)}{=} C(s) \|v(\tau_0)\|_{H^s_x} \overset{(4.3.40)}{=} C(s) \|W^{-1}(\omega\tau_0) B^{-1} \mathcal{A}^{-1}(\omega\tau_0) h(\tau_0)\|_{H^s_x} \\
&\overset{(4.3.41)}{\le} C(s) \|B^{-1} \mathcal{A}^{-1}(\omega\tau_0) h(\tau_0)\|_{H^s_x} \overset{(4.3.42)}{=} C(s) \|\mathcal{A}^{-1}(0) h(0)\|_{H^s_x} \overset{(4.3.41)}{\le} C(s) \|h(0)\|_{H^s_x}
\end{aligned}$$

having chosen $\tau_0 := \psi(0) = \alpha(0)$ (in the reversible case, α is an odd function, and so $\alpha(0) = 0$). Hence (1.2.24) is proved. To prove (1.2.25), we collect the estimates (4.1.143), (4.1.145), (4.2.9) into

$$\|(\mathcal{A}^{\pm 1}(\omega t) - I) g\|_{H^s_x} + \|(W^{\pm 1}(\omega t) - I) g\|_{H^s_x} \le \varepsilon\gamma^{-1} C(s) \|g\|_{H^{s+1}_x} , \quad \forall t \in \mathbb{R}, \ \forall g \in H^s_x, \tag{4.3.43}$$

where the constant $C(s)$ depends on $\|u\|_{s+\sigma+\beta+s_0}$. Thus

$$
\begin{aligned}
\|h(t)\|_{H^s_x} &\overset{(4.3.40)}{=} \|\mathcal{A}(\omega t)BW(\omega t)v(t)\|_{H^s_x} \le \|BW(\omega t)v(t)\|_{H^s_x} + \|(\mathcal{A}(\omega t)-I)BW(\omega t)v(t)\|_{H^s_x} \\
&\overset{(4.3.42)(4.3.43)}{\le} \|W(\omega\tau)v(\tau)\|_{H^s_x} + \varepsilon\gamma^{-1}C(s)\|BW(\omega t)v(t)\|_{H^{s+1}_x} \\
&\overset{(4.3.42)}{=} \|W(\omega\tau)v(\tau)\|_{H^s_x} + \varepsilon\gamma^{-1}C(s)\|W(\omega\tau)v(\tau)\|_{H^{s+1}_x} \\
&\overset{(4.3.41)}{\le} \|v(\tau)\|_{H^s_x} + \|(W(\omega\tau)-I)v(\tau)\|_{H^s_x} + \varepsilon\gamma^{-1}C(s)\|v(\tau)\|_{H^{s+1}_x} \\
&\overset{(4.3.43)}{\le} \|v(\tau)\|_{H^s_x} + \varepsilon\gamma^{-1}C(s)\|v(\tau)\|_{H^{s+1}_x} \overset{(1.2.23)}{=} \|v(\tau_0)\|_{H^s_x} + \varepsilon\gamma^{-1}C(s)\|v(\tau_0)\|_{H^{s+1}_x} \\
&\overset{(4.3.40)}{=} \|W^{-1}(\omega\tau_0)B^{-1}\mathcal{A}^{-1}(\omega\tau_0)h(\tau_0)\|_{H^s_x} + \varepsilon\gamma^{-1}C(s)\|W^{-1}(\omega\tau_0)B^{-1}\mathcal{A}^{-1}(\omega\tau_0)h(\tau_0)\|_{H^{s+1}_x} .
\end{aligned}
$$

Applying the same chain of inequalities at $\tau = \tau_0$, $t = 0$, we get that the last term is

$$
\le \|h(0)\|_{H^s_x} + \varepsilon\gamma^{-1}C(s)\|h(0)\|_{H^{s+1}_x} ,
$$

proving the second inequality in (1.2.25) with $\mathtt{a} := 1 - a$. The first one follows similarly.

Chapter 5

KAM for autonomous quasi-linear Hamiltonian perturbations of KdV

This Chapter is devoted to the proof of Theorem 1.3.1, stated in Section 1.3. First, in Section 5.1 we perform three steps of weak Birkhoff normal form, in order to determine the frequency-to-amplitude modulation (see (5.2.10)) and to provide a sufficiently good approximation of the solution for the convergence of the Nash-Moser scheme.

In Sections 5.2, 5.3 we introduce the *action-angle* variables (see (5.2.1)) and we reduce the problem of finding quasi periodic solutions of the equation (1.3.1) to the searching of invariant tori for the Hamiltonian H_ε in (5.2.11). The existence of invariant tori for H_ε is stated in Theorem 5.3.1, which is proved in the remaining Sections of the Chapter (Sections 5.4-5.7).

In Section 5.4, we describe the construction of the approximate inverse for the linearized operator (5.4.1). As we explained in Section 1.3, this procedure is inspired to [22], and it reduces the search of an approximate inverse for (5.4.1) to the inversion of the linear operator $\mathcal{L}_\omega$ in (5.4.45), acting on the space of the normal variables $H_S^\perp$. In Section 5.5, Proposition 5.5.1, we prove that $\mathcal{L}_\omega$ has the form (5.5.34). Then in Section 5.6 we reduce $\mathcal{L}_\omega$ to constant coefficients, semi-conjugating it to the operator $\mathcal{L}_\infty$ in (5.6.121) and in Theorem 5.6.2 we prove the invertibility on $\mathcal{L}_\omega$ and the tame estimates (5.6.124) for its inverse.

Finally in Section 5.7 we implement the Nash-Moser scheme for the nonlinear operator $\mathcal{F}$ in (5.3.6). This concludes the proof of Theorem 5.3.1.

Now we introduce some notations and we recall some well-known definitions, which will be used along this Chapter.

Tangential and normal variables. Let $\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu \geq 1$ be ν distinct integers, and $S^+ := \{\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu\}$. Let S be the symmetric set in (1.3.7), and $S^c := \{j \in \mathbb{Z}\setminus\{0\} : j \notin S\}$ its complementary set in $\mathbb{Z} \setminus \{0\}$. We decompose the phase space as

$$H_0^1(\mathbb{T}) := H_S \oplus H_S^\perp\,, \quad H_S := \mathrm{span}\{e^{\mathrm{i}jx} : j \in S\}, \quad H_S^\perp := \big\{u = \sum_{j \in S^c} u_j e^{\mathrm{i}jx} \in H_0^1(\mathbb{T})\big\}, \tag{5.0.1}$$

and we denote by Π_S, $\Pi_S^\perp$ the corresponding orthogonal projectors. Accordingly we decompose

$$u = v + z, \qquad v = \Pi_S u := \sum\nolimits_{j \in S} u_j\, e^{\mathrm{i}jx}, \quad z = \Pi_S^\perp u := \sum\nolimits_{j \in S^c} u_j\, e^{\mathrm{i}jx}\,, \tag{5.0.2}$$

where v is called the *tangential* variable and z the *normal* one. We shall sometimes identify $v \equiv (v_j)_{j \in S}$ and $z \equiv (z_j)_{j \in S^c}$. The subspaces H_S and $H_S^\perp$ are *symplectic*. The dynamics of these two components is quite different. On H_S we shall introduce the action-angle variables, see (5.2.1). The linear frequencies of oscillations on the tangential sites are

$$\bar{\omega} := (\bar{\jmath}_1^3, \dots, \bar{\jmath}_\nu^3) \in \mathbb{N}^\nu. \tag{5.0.3}$$

We shall also denote

$$H^s_{S^\perp}(\mathbb{T}^{\nu+1}) := \big\{ u \in H^s(\mathbb{T}^{\nu+1}) \,:\, u(\varphi, \cdot) \in H_S^\perp \ \forall \varphi \in \mathbb{T}^\nu \big\}\,, \tag{5.0.4}$$
$$H^s_S(\mathbb{T}^{\nu+1}) := \big\{ u \in H^s(\mathbb{T}^{\nu+1}) \,:\, u(\varphi, \cdot) \in H_S \ \forall \varphi \in \mathbb{T}^\nu \big\}\,. \tag{5.0.5}$$

Symplectic transformations. A map

$$\Phi : H^1_0(\mathbb{T}) \to H^1_0(\mathbb{T})$$

is symplectic if it preserves the 2-form Ω in (1.2.13), namely

$$\Omega(D\Phi(u)[h_1], D\Phi(u)[h_2]) = \Omega(h_1, h_2)\,, \qquad \forall u, h_1, h_2 \in H^1_0(\mathbb{T})\,,$$

which is equivalent to say that

$$D\Phi(u)^T \partial_x^{-1} D\Phi(u) = \partial_x^{-1}\,, \qquad \forall u \in H^1_0(\mathbb{T})\,.$$

It is well known that the symplectic maps preserve the Hamiltonian structure. This means that given a Hamiltonian $H : H^1_0(\mathbb{T}) \to \mathbb{R}$ and a symplectic map $\Phi : H^1_0(\mathbb{T}) \to H^1_0(\mathbb{T})$, the push-forward of the Hamiltonian vector field X_H,

$$\Phi^* X_H(u) := D\Phi(u)^{-1} X_H(\Phi(u))$$

is the Hamiltonian vector field X_K generated by the transformed Hamiltonian $K := H \circ \Phi$.

Given a Hamiltonian $F : H^1_0(\mathbb{T}) \to \mathbb{R}$, the time flow map Φ^t_F generated by the Hamiltonian vector field X_F, namely the flow of the PDE

$$\partial_t u = X_F(u)\,,$$

is a symplectic transformation, moreover for all function $H : H^1_0(\mathbb{T}) \to \mathbb{R}$ we have the Lie expansion

$$H \circ \Phi^t_F = \sum_{n \geq 0} \frac{\mathrm{ad}^n_F(H) t^n}{n!}\,, \qquad \mathrm{ad}_F(H) := \{H\,,\, F\}\,,$$

where the Poisson bracket $\{H\,,\,F\}$ is defined in (1.2.14).

Fourier series representation. It is convenient to regard the equation (1.3.1) also in the Fourier representation

$$u(x) = \sum\nolimits_{j \in \mathbb{Z}\setminus\{0\}} u_j e^{\mathrm{i} jx}\,, \qquad u(x) \longleftrightarrow u := (u_j)_{j \in \mathbb{Z}\setminus\{0\}}\,, \quad u_{-j} = \overline{u}_j\,, \tag{5.0.6}$$

where the Fourier indices $j \in \mathbb{Z} \setminus \{0\}$ by the definition of the phase space $H_0^1(\mathbb{T})$ and $u_{-j} = \overline{u}_j$ because $u(x)$ is real-valued. The symplectic structure writes

$$\Omega = \frac{1}{2} \sum_{j \neq 0} \frac{1}{\mathrm{i}j} du_j \wedge du_{-j} = \sum_{j \geq 1} \frac{1}{\mathrm{i}j} du_j \wedge du_{-j} \,, \qquad \Omega(u,v) = \sum_{j \neq 0} \frac{1}{\mathrm{i}j} u_j v_{-j} = \sum_{j \neq 0} \frac{1}{\mathrm{i}j} u_j \overline{v}_j \,, \tag{5.0.7}$$

the Hamiltonian vector field X_H and the Poisson bracket $\{F, G\}$ are

$$[X_H(u)]_j = \mathrm{i}j(\partial_{u_{-j}} H)(u) \,, \ \forall j \neq 0 \,, \quad \{F(u), G(u)\} = -\sum\nolimits_{j \neq 0} \mathrm{i}j(\partial_{u_{-j}} F)(u)(\partial_{u_j} G)(u) \,. \tag{5.0.8}$$

Conservation of momentum. A Hamiltonian

$$H(u) = \sum_{j_1, \ldots, j_n \in \mathbb{Z} \setminus \{0\}} H_{j_1, \ldots, j_n} u_{j_1} \ldots u_{j_n}, \quad u(x) = \sum_{j \in \mathbb{Z} \setminus \{0\}} u_j e^{\mathrm{i}jx}, \tag{5.0.9}$$

homogeneous of degree n, *preserves the momentum* if the coefficients $H_{j_1, \ldots, j_n}$ are zero for $j_1 + \ldots + j_n \neq 0$, so that the sum in (5.0.9) is restricted to integers such that $j_1 + \ldots + j_n = 0$. Equivalently, H preserves the momentum if $\{H, M\} = 0$, where M is the momentum $M(u) := \int_{\mathbb{T}} u^2 dx = \sum_{j \in \mathbb{Z} \setminus \{0\}} u_j u_{-j}$. The homogeneous components of degree ≤ 5 of the KdV Hamiltonian H in (1.3.3) preserve the momentum because, by (1.3.4), the homogeneous component f_5 of degree 5 does not depend on the space variable x.

5.1 Weak Birkhoff normal form

The Hamiltonian of the perturbed KdV equation (1.3.1) is $H = H_2 + H_3 + H_{\geq 5}$ (see (1.3.3)) where

$$H_2(u) := \frac{1}{2} \int_{\mathbb{T}} u_x^2 \, dx \,, \quad H_3(u) := \int_{\mathbb{T}} u^3 dx \,, \quad H_{\geq 5}(u) := \int_{\mathbb{T}} f(x, u, u_x) dx \,, \tag{5.1.1}$$

and f satisfies (1.3.4). According to the splitting (5.0.2) $u = v + z$, $v \in H_S$, $z \in H_S^{\perp}$, we have

$$H_2(u) = \int_{\mathbb{T}} \frac{v_x^2}{2} \, dx + \int_{\mathbb{T}} \frac{z_x^2}{2} \, dx, \quad H_3(u) = \int_{\mathbb{T}} v^3 dx + 3 \int_{\mathbb{T}} v^2 z dx + 3 \int_{\mathbb{T}} v z^2 dx + \int_{\mathbb{T}} z^3 dx \,. \tag{5.1.2}$$

For a finite-dimensional space

$$E := E_C := \mathrm{span}\{e^{\mathrm{i}jx} : 0 < |j| \leq C\}, \quad C > 0, \tag{5.1.3}$$

let Π_E denote the corresponding L^2-projector on E.

The notation $R(v^{k-q} z^q)$ indicates a homogeneous polynomial of degree k in (v, z) of the form

$$R(v^{k-q} z^q) = M[\, \underbrace{v, \ldots, v}_{(k-q) \text{ times}} \,, \underbrace{z, \ldots, z}_{q \text{ times}}], \qquad M = k\text{-linear} \,.$$

Proposition 5.1.1. (Weak Birkhoff normal form) *Assume Hypothesis* (S2). *Then there exists an analytic invertible symplectic transformation of the phase space* $\Phi_B : H_0^1(\mathbb{T}) \to H_0^1(\mathbb{T})$ *of the form*

$$\Phi_B(u) = u + \Psi(u), \quad \Psi(u) = \Pi_E \Psi(\Pi_E u), \tag{5.1.4}$$

where E is a finite-dimensional space as in (5.1.3)*, such that the transformed Hamiltonian is*

$$\mathcal{H} := H \circ \Phi_B = H_2 + \mathcal{H}_3 + \mathcal{H}_4 + \mathcal{H}_5 + \mathcal{H}_{\geq 6}\,, \tag{5.1.5}$$

where H_2 is defined in (5.1.1)*,*

$$\mathcal{H}_3 := \int_{\mathbb{T}} z^3\,dx + 3\int_{\mathbb{T}} vz^2\,dx\,, \quad \mathcal{H}_4 := -\frac{3}{2}\sum_{j\in S}\frac{|u_j|^4}{j^2} + \mathcal{H}_{4,2} + \mathcal{H}_{4,3}\,, \quad \mathcal{H}_5 := \sum_{q=2}^{5} R(v^{5-q}z^q)\,, \tag{5.1.6}$$

$$\mathcal{H}_{4,2} := 6\int_{\mathbb{T}} vz\Pi_S\big((\partial_x^{-1}v)(\partial_x^{-1}z)\big)\,dx + 3\int_{\mathbb{T}} z^2\pi_0(\partial_x^{-1}v)^2\,dx\,, \quad \mathcal{H}_{4,3} := R(vz^3)\,, \tag{5.1.7}$$

and $\mathcal{H}_{\geq 6}$ collects all the terms of order at least six in (v,z).

The rest of this section is devoted to the proof of Proposition 5.1.1.

First, we remove the cubic terms $\int_{\mathbb{T}} v^3 + 3\int_{\mathbb{T}} v^2 z$ from the Hamiltonian H_3 defined in (5.1.2). In the Fourier coordinates (5.0.6), we have

$$H_2 = \frac{1}{2}\sum_{j\neq 0} j^2|u_j|^2, \quad H_3 = \sum_{j_1+j_2+j_3=0} u_{j_1}u_{j_2}u_{j_3}\,. \tag{5.1.8}$$

We look for a symplectic transformation $\Phi^{(3)}$ of the phase space which eliminates the monomials $u_{j_1}u_{j_2}u_{j_3}$ of H_3 with at most *one* index outside S. Note that, by the relation $j_1+j_2+j_3=0$, they are *finitely* many. We look for $\Phi^{(3)} := (\Phi^t_{F^{(3)}})_{|t=1}$ as the time-1 flow map generated by the Hamiltonian vector field $X_{F^{(3)}}$, with an auxiliary Hamiltonian of the form

$$F^{(3)}(u) := \sum_{j_1+j_2+j_3=0} F^{(3)}_{j_1j_2j_3}u_{j_1}u_{j_2}u_{j_3}\,.$$

The transformed Hamiltonian is

$$\begin{aligned} &H^{(3)} := H\circ\Phi^{(3)} = H_2 + H_3^{(3)} + H_4^{(3)} + H_{\geq 5}^{(3)}\,,\\ &H_3^{(3)} = H_3 + \{H_2, F^{(3)}\}, \quad H_4^{(3)} = \frac{1}{2}\{\{H_2,F^{(3)}\},F^{(3)}\} + \{H_3,F^{(3)}\}, \end{aligned} \tag{5.1.9}$$

where $H_{\geq 5}^{(3)}$ collects all the terms of order at least five in (u,u_x). By (5.1.8) and (5.0.8) we calculate

$$H_3^{(3)} = \sum_{j_1+j_2+j_3=0}\big\{1 - \mathrm{i}(j_1^3+j_2^3+j_3^3)F^{(3)}_{j_1j_2j_3}\big\}\,u_{j_1}u_{j_2}u_{j_3}\,.$$

Hence, in order to eliminate the monomials with at most one index outside S, we choose

$$F^{(3)}_{j_1j_2j_3} := \begin{cases} \dfrac{1}{\mathrm{i}(j_1^3+j_2^3+j_3^3)} & \text{if } (j_1,j_2,j_3)\in\mathcal{A}\,,\\ 0 & \text{otherwise,}\end{cases} \tag{5.1.10}$$

where $\mathcal{A} := \big\{(j_1,j_2,j_3)\in(\mathbb{Z}\setminus\{0\})^3 : j_1+j_2+j_3=0,\ j_1^3+j_2^3+j_3^3\neq 0$, and at least 2 among j_1,j_2,j_3 belong to $S\big\}$. Note that

$$\mathcal{A} = \big\{(j_1,j_2,j_3)\in(\mathbb{Z}\setminus\{0\})^3 : j_1+j_2+j_3=0, \text{ and at least 2 among } j_1,j_2,j_3 \text{ belong to } S\big\} \tag{5.1.11}$$

because of the elementary relation

$$j_1 + j_2 + j_3 = 0 \quad \Rightarrow \quad j_1^3 + j_2^3 + j_3^3 = 3j_1j_2j_3 \neq 0 \tag{5.1.12}$$

being $j_1, j_2, j_3 \in \mathbb{Z} \setminus \{0\}$. Also note that $\mathcal{A}$ is a finite set, actually $\mathcal{A} \subseteq [-2C_S, 2C_S]^3$ where the tangential sites $S \subseteq [-C_S, C_S]$. As a consequence, the Hamiltonian vector field $X_{F^{(3)}}$ has finite rank and vanishes outside the finite dimensional subspace $E := E_{2C_S}$ (see (5.1.3)), namely

$$X_{F^{(3)}}(u) = \Pi_E X_{F^{(3)}}(\Pi_E u)\,.$$

Hence its flow $\Phi^{(3)} : H_0^1(\mathbb{T}) \to H_0^1(\mathbb{T})$ has the form (5.1.4) and it is analytic.

By construction, all the monomials of H_3 with at least two indices outside S are not modified by the transformation $\Phi^{(3)}$. Hence (see (5.1.2)) we have

$$H_3^{(3)} = \int_{\mathbb{T}} z^3\, dx + 3\int_{\mathbb{T}} vz^2\, dx\,. \tag{5.1.13}$$

We now compute the fourth order term $H_4^{(3)} = \sum_{i=0}^4 H_{4,i}^{(3)}$ in (5.1.9), where $H_{4,i}^{(3)}$ is of type $R(v^{4-i}z^i)$.

Lemma 5.1.1. *One has (recall the definition* (4.1.50) *of* π_0*)*

$$H_{4,0}^{(3)} := \frac{3}{2}\int_{\mathbb{T}} v^2\pi_0[(\partial_x^{-1}v)^2]dx\,, \quad H_{4,2}^{(3)} := 6\int_{\mathbb{T}} vz\Pi_S\big((\partial_x^{-1}v)(\partial_x^{-1}z)\big)\, dx + 3\int_{\mathbb{T}} z^2\pi_0[(\partial_x^{-1}v)^2]dx\,. \tag{5.1.14}$$

Proof. We write $H_3 = H_{3,\le 1} + H_3^{(3)}$ where $H_{3,\le 1}(u) := \int_{\mathbb{T}} v^3 dx + 3\int_{\mathbb{T}} v^2 z\, dx$. Then, by (5.1.9), we get

$$H_4^{(3)} = \frac{1}{2}\{H_{3,\le 1}\,, F^{(3)}\} + \{H_3^{(3)}, F^{(3)}\}\,. \tag{5.1.15}$$

By (5.1.10), (5.1.12), the auxiliary Hamiltonian may be written as

$$F^{(3)}(u) = -\frac{1}{3}\sum_{(j_1,j_2,j_3)\in\mathcal{A}} \frac{u_{j_1}u_{j_2}u_{j_3}}{(\mathrm{i}j_1)(\mathrm{i}j_2)(\mathrm{i}j_3)} = -\frac{1}{3}\int_{\mathbb{T}}(\partial_x^{-1}v)^3dx - \int_{\mathbb{T}}(\partial_x^{-1}v)^2(\partial_x^{-1}z)dx\,.$$

Hence, using that the projectors Π_S, $\Pi_S^\perp$ are self-adjoint and ∂_x^{-1} is skew-selfadjoint,

$$\nabla F^{(3)}(u) = \partial_x^{-1}\big\{(\partial_x^{-1}v)^2 + 2\Pi_S\big[(\partial_x^{-1}v)(\partial_x^{-1}z)\big]\big\} \tag{5.1.16}$$

(we have used that $\partial_x^{-1}\pi_0 = \partial_x^{-1}$ be the definition of ∂_x^{-1}). Recalling the Poisson bracket definition (1.2.14), using that $\nabla H_{3,\le 1}(u) = 3v^2 + 6\Pi_S(vz)$ and (5.1.16), we get

$$\begin{aligned}\{H_{3,\le 1}, F^{(3)}\} &= \int_{\mathbb{T}}\big\{3v^2 + 6\Pi_S(vz)\big\}\pi_0\big\{(\partial_x^{-1}v)^2 + 2\Pi_S\big[(\partial_x^{-1}v)(\partial_x^{-1}z)\big]\big\}\, dx \\ &= 3\int_{\mathbb{T}} v^2\pi_0(\partial_x^{-1}v)^2\, dx + 12\int_{\mathbb{T}}\Pi_S(vz)\Pi_S[(\partial_x^{-1}v)(\partial_x^{-1}z)]\, dx + R(v^3z)\,.\end{aligned} \tag{5.1.17}$$

Similarly, since $\nabla H_3^{(3)}(u) = 3z^2 + 6\Pi_S^\perp(vz)$,

$$\{H_3^{(3)}, F^{(3)}\} = 3\int_{\mathbb{T}} z^2\pi_0(\partial_x^{-1}v)^2\, dx + R(v^3z) + R(vz^3)\,. \tag{5.1.18}$$

The lemma follows by (5.1.15), (5.1.17), (5.1.18). $\square$

We now construct a symplectic map $\Phi^{(4)}$ such that the Hamiltonian system obtained transforming $H_2 + H_3^{(3)} + H_4^{(3)}$ possesses the invariant subspace H_S (see (5.0.1)) and its dynamics on H_S is integrable and non-isocronous. Hence we have to eliminate the term $H_{4,1}^{(3)}$ (which is linear in z), and to normalize $H_{4,0}^{(3)}$ (which is independent of z). We need the following elementary lemma (Lemma 13.4 in [49]).

Lemma 5.1.2. *Let $j_1, j_2, j_3, j_4 \in \mathbb{Z}$ such that $j_1 + j_2 + j_3 + j_4 = 0$. Then*

$$j_1^3 + j_2^3 + j_3^3 + j_4^3 = -3(j_1 + j_2)(j_1 + j_3)(j_2 + j_3).$$

Lemma 5.1.3. *There exists a symplectic transformation $\Phi^{(4)}$ of the form* (5.1.4) *such that*

$$H^{(4)} := H^{(3)} \circ \Phi^{(4)} = H_2 + H_3^{(3)} + H_4^{(4)} + H_{\geq 5}^{(4)} , \qquad H_4^{(4)} := -\frac{3}{2} \sum_{j \in S} \frac{|u_j|^4}{j^2} + H_{4,2}^{(3)} + H_{4,3}^{(3)} , \tag{5.1.19}$$

where $H_3^{(3)}$ is defined in (5.1.13), *$H_{4,2}^{(3)}$ in* (5.1.14), *$H_{4,3}^{(3)} = R(vz^3)$ and $H_{\geq 5}^{(4)}$ collects all the terms of degree at least five in (u, u_x).*

Proof. We look for a map $\Phi^{(4)} := (\Phi^t_{F^{(4)}})_{|t=1}$ which is the time 1-flow map of an auxiliary Hamiltonian

$$F^{(4)}(u) := \sum_{\substack{j_1+j_2+j_3+j_4=0 \\ \text{at least 3 indices are in } S}} F^{(4)}_{j_1 j_2 j_3 j_4} u_{j_1} u_{j_2} u_{j_3} u_{j_4}$$

with the same form of the Hamiltonian $H_{4,0}^{(3)} + H_{4,1}^{(3)}$. The transformed Hamiltonian is

$$H^{(4)} := H^{(3)} \circ \Phi^{(4)} = H_2 + H_3^{(3)} + H_4^{(4)} + H_{\geq 5}^{(4)}, \quad H_4^{(4)} = \{H_2, F^{(4)}\} + H_4^{(3)}, \tag{5.1.20}$$

where $H_{\geq 5}^{(4)}$ collects all the terms of order at least five. We write $H_4^{(4)} = \sum_{i=0}^4 H_{4,i}^{(4)}$ where each $H_{4,i}^{(4)}$ if of type $R(v^{4-i}z^i)$. We choose the coefficients

$$F^{(4)}_{j_1 j_2 j_3 j_4} := \begin{cases} \dfrac{H^{(3)}_{j_1 j_2 j_3 j_4}}{\mathrm{i}(j_1^3 + j_2^3 + j_3^3 + j_4^3)} & \text{if } (j_1, j_2, j_3, j_4) \in \mathcal{A}_4 , \\ 0 & \text{otherwise,} \end{cases} \tag{5.1.21}$$

where

$$\mathcal{A}_4 := \big\{ (j_1, j_2, j_3, j_4) \in (\mathbb{Z} \setminus \{0\})^4 : j_1 + j_2 + j_3 + j_4 = 0, \, j_1^3 + j_2^3 + j_3^3 + j_4^3 \neq 0, \\ \text{and at most one among } j_1, j_2, j_3, j_4 \text{ outside } S \big\} .$$

By this definition $H_{4,1}^{(4)} = 0$ because there exist no integers $j_1, j_2, j_3 \in S$, $j_4 \in S^c$ satisfying $j_1 + j_2 + j_3 + j_4 = 0$, $j_1^3 + j_2^3 + j_3^3 + j_4^3 = 0$, by Lemma 5.1.2 and the fact that S is symmetric. By construction, the terms $H_{4,i}^{(4)} = H_{4,i}^{(3)}$, $i = 2, 3, 4$, are not changed by $\Phi^{(4)}$. Finally, by (5.1.14)

$$H_{4,0}^{(4)} = \frac{3}{2} \sum_{\substack{j_1, j_2, j_3, j_4 \in S \\ j_1+j_2+j_3+j_4=0 \\ j_1^3+j_2^3+j_3^3+j_4^3=0 \\ j_1+j_2 \, , \, j_3+j_4 \neq 0}} \frac{1}{(\mathrm{i}j_3)(\mathrm{i}j_4)} u_{j_1} u_{j_2} u_{j_3} u_{j_4} . \tag{5.1.22}$$

If $j_1+j_2+j_3+j_4=0$ and $j_1^3+j_2^3+j_3^3+j_4^3=0$, then $(j_1+j_2)(j_1+j_3)(j_2+j_3)=0$ by Lemma 5.1.2. We develop the sum in (5.1.22) with respect to the first index j_1. Since $j_1+j_2\neq 0$ the possible cases are:

$$(i)\ \big\{j_2\neq -j_1,\ j_3=-j_1,\ j_4=-j_2\big\} \qquad \text{or} \qquad (ii)\ \big\{j_2\neq -j_1,\ j_3\neq -j_1,\ j_3=-j_2,\ j_4=-j_1\big\}.$$

Hence, using $u_{-j}=\bar u_j$ (recall (5.0.6)), and since S is symmetric, we have

$$\begin{aligned}\sum_{(i)}\frac{1}{j_3j_4}u_{j_1}u_{j_2}u_{j_3}u_{j_4} &= \sum_{j_1,j_2\in S, j_2\neq -j_1}\frac{|u_{j_1}|^2|u_{j_2}|^2}{j_1j_2}\\ &= \sum_{j,j'\in S}\frac{|u_j|^2|u_{j'}|^2}{jj'}+\sum_{j\in S}\frac{|u_j|^4}{j^2}=\sum_{j\in S}\frac{|u_j|^4}{j^2}\,,\end{aligned}\tag{5.1.23}$$

and in the second case (ii)

$$\sum_{(ii)}\frac{1}{j_3j_4}u_{j_1}u_{j_2}u_{j_3}u_{j_4}=\sum_{j_1,j_2,j_2\neq\pm j_1}\frac{1}{j_1j_2}u_{j_1}u_{j_2}u_{-j_2}u_{-j_1}=\sum_{j\in S}\frac{1}{j}|u_j|^2\Big(\sum_{j_2\neq\pm j}\frac{1}{j_2}|u_{j_2}|^2\Big)=0\,.\tag{5.1.24}$$

Then (5.1.19) follows by (5.1.22), (5.1.23), (5.1.24). □

Note that the Hamiltonian $H_2+H_3^{(3)}+H_4^{(4)}$ (see (5.1.19)) possesses the invariant subspace $\{z=0\}$ and the system restricted to $\{z=0\}$ is completely integrable and non-isochronous (actually it is formed by ν decoupled rotators). We shall construct quasi-periodic solutions which bifurcate from this invariant manifold.

In order to enter in a perturbative regime, we have to eliminate further monomials of $H^{(4)}$ in (5.1.19). The minimal requirement for the convergence of the nonlinear Nash-Moser iteration is to eliminate the monomials $R(v^5)$ and $R(v^4z)$. Here we need the choice of the sites of Hypothesis (S2).

Remark 5.1.1. *In the KAM theorems [56], [66] (and [68], [72]), as well as for the perturbed mKdV equations* (1.3.9)*, these further steps of Birkhoff normal form are not required because the nonlinearity of the original PDE is yet cubic. A difficulty of KdV is that the nonlinearity is quadratic.*

We spell out Hypothesis (S2) as follows:

- ($\mathtt{S2}_0$). There is no choice of 5 integers $j_1,\dots,j_5\in S$ such that

$$j_1+\ldots+j_5=0\,,\quad j_1^3+\ldots+j_5^3=0\,.\tag{5.1.25}$$

- ($\mathtt{S2}_1$). There is no choice of 4 integers $j_1,\dots,j_4$ in S and an integer in the complementary set $j_5\in S^c:=(\mathbb{Z}\setminus\{0\})\setminus S$ such that (5.1.25) holds.

The homogeneous component of degree 5 of $H^{(4)}$ is

$$H_5^{(4)}(u)=\sum_{j_1+\ldots+j_5=0}H^{(4)}_{j_1,\ldots,j_5}u_{j_1}\ldots u_{j_5}\,.$$

We want to remove from $H_5^{(4)}$ the terms with at most one index among $j_1, \ldots, j_5$ outside S. We consider the auxiliary Hamiltonian

$$F^{(5)} = \sum_{\substack{j_1+\ldots+j_5=0 \\ \text{at most one index outside } S}} F^{(5)}_{j_1\ldots j_5} u_{j_1} \ldots u_{j_5}\,, \quad F^{(5)}_{j_1\ldots j_5} := \frac{H^{(5)}_{j_1\ldots j_5}}{\mathrm{i}(j_1^3 + \ldots + j_5^3)}\,. \tag{5.1.26}$$

By Hypotheses $(\mathtt{S2})_0, (\mathtt{S2})_1$, if $j_1+\ldots+j_5 = 0$ with at most one index outside S then $j_1^3+\ldots+j_5^3 \neq 0$ and $F^{(5)}$ is well defined. Let $\Phi^{(5)}$ be the time 1-flow generated by $X_{F^{(5)}}$. The new Hamiltonian is

$$H^{(5)} := H^{(4)} \circ \Phi^{(5)} = H_2 + H_3^{(3)} + H_4^{(4)} + \{H_2, F^{(5)}\} + H_5^{(4)} + H_{\geq 6}^{(5)} \tag{5.1.27}$$

where, by (5.1.26),

$$H_5^{(5)} := \{H_2, F^{(5)}\} + H_5^{(4)} = \sum\nolimits_{q=2}^{5} R(v^{5-q} z^q)\,.$$

Renaming $\mathcal{H} := H^{(5)}$, namely $\mathcal{H}_n := H_n^{(n)}$, $n = 3, 4, 5$, and setting $\Phi_B := \Phi^{(3)} \circ \Phi^{(4)} \circ \Phi^{(5)}$, formula (5.1.5) follows.

The homogeneous component $H_5^{(4)}$ preserves the momentum, see (5.0.9). Hence $F^{(5)}$ also preserves the momentum. As a consequence, also $H_k^{(5)}$, $k \leq 5$, preserve the momentum.

Finally, since $F^{(5)}$ is Fourier-supported on a finite set, the transformation $\Phi^{(5)}$ is of type (5.1.4) (and analytic), and therefore also the composition Φ_B is of type (5.1.4) (and analytic).

5.2 Action-angle variables

We now introduce action-angle variables on the tangential directions by the change of coordinates

$$\begin{cases} u_j := \sqrt{\xi_j + |j| y_j}\, e^{\mathrm{i}\theta_j}, & \text{if } j \in S\,, \\ u_j := z_j, & \text{if } j \in S^c\,, \end{cases} \tag{5.2.1}$$

where (recall $u_{-j} = \overline{u}_j$)

$$\xi_{-j} = \xi_j\,, \quad \xi_j > 0\,, \quad y_{-j} = y_j\,, \quad \theta_{-j} = -\theta_j\,, \quad \theta_j, y_j \in \mathbb{R}\,, \quad \forall j \in S\,. \tag{5.2.2}$$

For the tangential sites $S^+ := \{\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu\}$ we shall also denote $\theta_{\bar{\jmath}_i} := \theta_i$, $y_{\bar{\jmath}_i} := y_i$, $\xi_{\bar{\jmath}_i} := \xi_i$, $i = 1, \ldots \nu$.

The symplectic 2-form Ω in (5.0.7) (i.e. (1.2.13)) becomes

$$\mathcal{W} := \sum_{i=1}^{\nu} d\theta_i \wedge dy_i + \frac{1}{2} \sum_{j \in S^c \setminus \{0\}} \frac{1}{\mathrm{i}j} dz_j \wedge dz_{-j} = \Big(\sum_{i=1}^{\nu} d\theta_i \wedge dy_i\Big) \oplus \Omega_{S^\perp} = d\Lambda \tag{5.2.3}$$

where $\Omega_{S^\perp}$ denotes the restriction of Ω to $H_S^\perp$ (see (5.0.1)) and Λ is the contact 1-form on $\mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$ defined by $\Lambda_{(\theta,y,z)} : \mathbb{R}^\nu \times \mathbb{R}^\nu \times H_S^\perp \to \mathbb{R}$,

$$\Lambda_{(\theta,y,z)}[\widehat{\theta}, \widehat{y}, \widehat{z}] := -y \cdot \widehat{\theta} + \frac{1}{2}(\partial_x^{-1} z, \widehat{z})_{L^2(\mathbb{T})}\,. \tag{5.2.4}$$

Instead of working in a shrinking neighborhood of the origin, it is a convenient devise to rescale the "unperturbed actions" ξ and the action-angle variables as

$$\xi \mapsto \varepsilon^2 \xi\,, \quad y \mapsto \varepsilon^{2b} y\,, \quad z \mapsto \varepsilon^b z\,. \tag{5.2.5}$$

Then the symplectic 2-form in (5.2.3) transforms into $\varepsilon^{2b}\mathcal{W}$. Hence the Hamiltonian system generated by $\mathcal{H}$ in (5.1.5) transforms into the new Hamiltonian system

$$\dot\theta = \partial_y H_\varepsilon(\theta, y, z)\,,\ \dot y = -\partial_\theta H_\varepsilon(\theta, y, z)\,,\ z_t = \partial_x \nabla_z H_\varepsilon(\theta, y, z)\,, \quad H_\varepsilon := \varepsilon^{-2b}\mathcal{H} \circ A_\varepsilon \tag{5.2.6}$$

where

$$A_\varepsilon(\theta, y, z) := \varepsilon v_\varepsilon(\theta, y) + \varepsilon^b z := \varepsilon \sum\nolimits_{j \in S} \sqrt{\xi_j + \varepsilon^{2(b-1)}|j| y_j}\, e^{\mathrm{i}\theta_j} e^{\mathrm{i}jx} + \varepsilon^b z\,. \tag{5.2.7}$$

We shall still denote by $X_{H_\varepsilon} = (\partial_y H_\varepsilon, -\partial_\theta H_\varepsilon, \partial_x \nabla_z H_\varepsilon)$ the Hamiltonian vector field in the variables $(\theta, y, z) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$.

We now write explicitly the Hamiltonian $H_\varepsilon(\theta, y, z)$ in (5.2.6). The quadratic Hamiltonian H_2 in (5.1.1) transforms into

$$\varepsilon^{-2b} H_2 \circ A_\varepsilon = const + \sum\nolimits_{j \in S^+} j^3 y_j + \frac{1}{2}\int_{\mathbb{T}} z_x^2\, dx\,, \tag{5.2.8}$$

and, recalling (5.1.6), (5.1.7), the Hamiltonian $\mathcal{H}$ in (5.1.5) transforms into (shortly writing $v_\varepsilon := v_\varepsilon(\theta, y)$)

$$\begin{aligned} H_\varepsilon(\theta, y, z) = {} & e(\xi) + \alpha(\xi) \cdot y + \frac{1}{2}\int_{\mathbb{T}} z_x^2 dx + \varepsilon^b \int_{\mathbb{T}} z^3 dx + 3\varepsilon \int_{\mathbb{T}} v_\varepsilon z^2 dx \\ & + \varepsilon^2 \Big\{ 6 \int_{\mathbb{T}} v_\varepsilon z \Pi_S\big((\partial_x^{-1} v_\varepsilon)(\partial_x^{-1} z)\big)\, dx + 3 \int_{\mathbb{T}} z^2 \pi_0 (\partial_x^{-1} v_\varepsilon)^2\, dx \Big\} - \frac{3}{2}\varepsilon^{2b} \sum\nolimits_{j \in S} y_j^2 \\ & + \varepsilon^{b+1} R(v_\varepsilon z^3) + \varepsilon^3 R(v_\varepsilon^3 z^2) + \varepsilon^{2+b} \sum_{q=3}^{5} \varepsilon^{(q-3)(b-1)} R(v_\varepsilon^{5-q} z^q) + \varepsilon^{-2b}\mathcal{H}_{\ge 6}(\varepsilon v_\varepsilon + \varepsilon^b z) \end{aligned} \tag{5.2.9}$$

where $e(\xi)$ is a constant, and the frequency-amplitude map is

$$\alpha(\xi) := \bar\omega + \varepsilon^2 \mathbb{A}\xi\,, \quad \mathbb{A} := -6\, \mathrm{diag}\{1/j\}_{j \in S^+}\,. \tag{5.2.10}$$

We write the Hamiltonian in (5.2.9) as

$$H_\varepsilon = \mathcal{N} + P\,, \quad \mathcal{N}(\theta, y, z) = \alpha(\xi) \cdot y + \frac{1}{2}\big(N(\theta) z, z\big)_{L^2(\mathbb{T})}\,, \tag{5.2.11}$$

where

$$\begin{aligned} \frac{1}{2}\big(N(\theta) z, z\big)_{L^2(\mathbb{T})} := {} & \frac{1}{2}\big((\partial_z \nabla H_\varepsilon)(\theta, 0, 0)[z], z\big)_{L^2(\mathbb{T})} = \frac{1}{2}\int_{\mathbb{T}} z_x^2 dx + 3\varepsilon \int_{\mathbb{T}} v_\varepsilon(\theta, 0) z^2 dx \\ & + \varepsilon^2 \Big\{ 6 \int_{\mathbb{T}} v_\varepsilon(\theta, 0) z \Pi_S\big((\partial_x^{-1} v_\varepsilon(\theta, 0))(\partial_x^{-1} z)\big) dx + 3 \int_{\mathbb{T}} z^2 \pi_0 (\partial_x^{-1} v_\varepsilon(\theta, 0))^2 dx \Big\} + \ldots \end{aligned} \tag{5.2.12}$$

and $P := H_\varepsilon - \mathcal{N}$.

5.3 The nonlinear functional setting

We look for an embedded invariant torus

$$i : \mathbb{T}^\nu \to \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp, \quad \varphi \mapsto i(\varphi) := (\theta(\varphi), y(\varphi), z(\varphi)) \tag{5.3.1}$$

of the Hamiltonian vector field X_{H_ε} filled by quasi-periodic solutions with diophantine frequency ω. We require that ω belongs to the set

$$\Omega_\varepsilon := \alpha([1,2]^\nu) = \{\alpha(\xi) : \xi \in [1,2]^\nu\} \tag{5.3.2}$$

where α is the diffeomorphism (5.2.10), and, in the Hamiltonian H_ε in (5.2.11), we choose

$$\xi = \alpha^{-1}(\omega) = \varepsilon^{-2}\mathbb{A}^{-1}(\omega - \bar{\omega}) . \tag{5.3.3}$$

Since any $\omega \in \Omega_\varepsilon$ is ε^2-close to the integer vector $\bar{\omega}$ (see (5.0.3)), we require that the constant γ in the diophantine inequality

$$|\omega \cdot l| \geq \gamma \langle l \rangle^{-\tau} , \quad \forall l \in \mathbb{Z}^\nu \setminus \{0\} , \quad \text{satisfies} \ \ \gamma = \varepsilon^{2+a} \quad \text{for some } a > 0 . \tag{5.3.4}$$

We remark that the definition of γ in (5.3.4) is slightly stronger than the minimal condition, which is $\gamma \leq c\varepsilon^2$ with c small enough. In addition to (5.3.4) we shall also require that ω satisfies the first and second order Melnikov-non-resonance conditions (5.6.120).

We look for an embedded invariant torus of the modified Hamiltonian vector field $X_{H_{\varepsilon,\zeta}} = X_{H_\varepsilon} + (0, \zeta, 0)$ which is generated by the Hamiltonian

$$H_{\varepsilon,\zeta}(\theta, y, z) := H_\varepsilon(\theta, y, z) + \zeta \cdot \theta , \quad \zeta \in \mathbb{R}^\nu . \tag{5.3.5}$$

Note that $X_{H_{\varepsilon,\zeta}}$ is periodic in θ (unlike $H_{\varepsilon,\zeta}$). It turns out that an invariant torus for $X_{H_{\varepsilon,\zeta}}$ is actually invariant for X_{H_ε}, see Lemma 5.4.1. We introduce the parameter $\zeta \in \mathbb{R}^\nu$ in order to control the average in the y-component of the linearized equations. Thus we look for zeros of the nonlinear operator

$$\begin{aligned}
\mathcal{F}(i, \zeta) &:= \mathcal{F}(i, \zeta, \omega, \varepsilon) := \mathcal{D}_\omega i(\varphi) - X_{H_{\varepsilon,\zeta}}(i(\varphi)) \\
&= \mathcal{D}_\omega i(\varphi) - X_{\mathcal{N}}(i(\varphi)) - X_P(i(\varphi)) + (0, \zeta, 0) \qquad (5.3.6)\\
&:= \begin{pmatrix} \mathcal{D}_\omega \theta(\varphi) - \partial_y H_\varepsilon(i(\varphi)) \\ \mathcal{D}_\omega y(\varphi) + \partial_\theta H_\varepsilon(i(\varphi)) + \zeta \\ \mathcal{D}_\omega z(\varphi) - \partial_x \nabla_z H_\varepsilon(i(\varphi)) \end{pmatrix} = \begin{pmatrix} \mathcal{D}_\omega \Theta(\varphi) - \partial_y P(i(\varphi)) \\ \mathcal{D}_\omega y(\varphi) + \frac{1}{2}\partial_\theta (N(\theta(\varphi)) z(\varphi), z(\varphi))_{L^2(\mathbb{T})} + \partial_\theta P(i(\varphi)) + \zeta \\ \mathcal{D}_\omega z(\varphi) - \partial_x N(\theta(\varphi)) z(\varphi) - \partial_x \nabla_z P(i(\varphi)) \end{pmatrix}
\end{aligned}$$

where $\Theta(\varphi) := \theta(\varphi) - \varphi$ is $(2\pi)^\nu$-periodic and we use the short notation

$$\mathcal{D}_\omega := \omega \cdot \partial_\varphi . \tag{5.3.7}$$

The Sobolev norm of the periodic component of the embedded torus

$$\mathfrak{I}(\varphi) := i(\varphi) - (\varphi, 0, 0) := (\Theta(\varphi), y(\varphi), z(\varphi)) , \quad \Theta(\varphi) := \theta(\varphi) - \varphi , \tag{5.3.8}$$

is

$$\|\mathfrak{I}\|_s := \|\Theta\|_{H^s_\varphi} + \|y\|_{H^s_\varphi} + \|z\|_s \tag{5.3.9}$$

where $\|z\|_s := \|z\|_{H^s_{\varphi,x}}$ is defined in (3.0.1). We link the rescaling (5.2.5) with the diophantine constant $\gamma = \varepsilon^{2+a}$ by choosing

$$\gamma = \varepsilon^{2b}, \qquad b = 1 + (a/2). \tag{5.3.10}$$

Other choices are possible, see Remark 5.3.1.

Theorem 5.3.1. *Let the tangential sites S in* (1.3.7) *satisfy* (S1), (S2). *Then, for all $\varepsilon \in (0, \varepsilon_0)$, where ε_0 is small enough, there exists a Cantor-like set $\mathcal{C}_\varepsilon \subset \Omega_\varepsilon$, with asympotically full measure as $\varepsilon \to 0$, namely*

$$\lim_{\varepsilon \to 0} \frac{|\mathcal{C}_\varepsilon|}{|\Omega_\varepsilon|} = 1, \tag{5.3.11}$$

such that, for all $\omega \in \mathcal{C}_\varepsilon$, there exists a solution $i_\infty(\varphi) := i_\infty(\omega, \varepsilon)(\varphi)$ of $\mathcal{D}_\omega i_\infty(\varphi) - X_{H_\varepsilon}(i_\infty(\varphi)) = 0$. Hence the embedded torus $\varphi \mapsto i_\infty(\varphi)$ is invariant for the Hamiltonian vector field $X_{H_\varepsilon(\cdot,\xi)}$ with ξ as in (5.3.3), *and it is filled by quasi-periodic solutions with frequency ω. The torus i_∞ satisfies*

$$\|i_\infty(\varphi) - (\varphi, 0, 0)\|^{\mathrm{Lip}(\gamma)}_{s_0+\mu} = O(\varepsilon^{6-2b}\gamma^{-1}) \tag{5.3.12}$$

for some $\mu := \mu(\nu) > 0$. Moreover, the torus i_∞ is LINEARLY STABLE.

Theorem 5.3.1 is proved in Sections 5.4-5.7. It implies Theorem 1.3.1 where the ξ_j in (1.3.12) are $\varepsilon^2\xi_j$, $\xi_j \in [1, 2]$, in (5.3.3). By (5.3.12), going back to the variables before the rescaling (5.2.5), we get $\Theta_\infty = O(\varepsilon^{6-2b}\gamma^{-1})$, $y_\infty = O(\varepsilon^6\gamma^{-1})$, $z_\infty = O(\varepsilon^{6-b}\gamma^{-1})$, which, as $b \to 1^+$, tend to the expected optimal estimates.

Remark 5.3.1. *There are other possible ways to link the rescaling* (5.2.5) *with the diophantine constant $\gamma = \varepsilon^{2+a}$. The choice $\gamma > \varepsilon^{2b}$ reduces to study perturbations of an isochronous system (as in [53], [56], [66]), and it is convenient to introduce $\xi(\omega)$ as a variable. The case $\varepsilon^{2b} > \gamma$, in particular $b = 1$, has to be dealt with a perturbation approach of a non-isochronous system à la Arnold-Kolmogorov.*

We now give the tame estimates for the composition operator induced by the Hamiltonian vector fields $X_\mathcal{N}$ and X_P in (5.3.6), that we shall use in the next Sections.

We first estimate the composition operator induced by $v_\varepsilon(\theta, y)$ defined in (5.2.7). Since the functions $y \mapsto \sqrt{\xi + \varepsilon^{2(b-1)}|j|y}$, $\theta \mapsto e^{\mathrm{i}\theta}$ are analytic for ε small enough and $|y| \le C$, the composition Lemma A.0.8 implies that, for all $\Theta, y \in H^s(\mathbb{T}^\nu, \mathbb{R}^\nu)$, $\|\Theta\|_{s_0}, \|y\|_{s_0} \le 1$, setting $\theta(\varphi) := \varphi + \Theta(\varphi)$, $\|v_\varepsilon(\theta(\varphi), y(\varphi))\|_s \le_s 1 + \|\Theta\|_s + \|y\|_s$. Hence, using also (5.3.3), the map A_ε in (5.2.7) satisfies, for all $\|\mathfrak{I}\|^{\mathrm{Lip}(\gamma)}_{s_0} \le 1$ (see (5.3.8))

$$\|A_\varepsilon(\theta(\varphi), y(\varphi), z(\varphi))\|^{\mathrm{Lip}(\gamma)}_s \le_s \varepsilon(1 + \|\mathfrak{I}\|^{\mathrm{Lip}(\gamma)}_s). \tag{5.3.13}$$

We now give tame estimates for the Hamiltonian vector fields $X_\mathcal{N}$, X_P, X_{H_ε}, see (5.2.11)-(5.2.12).

Lemma 5.3.1. *Let $\mathfrak{I}(\varphi)$ in (5.3.8) satisfy $\|\mathfrak{I}\|_{s_0+3}^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^{6-2b}\gamma^{-1}$. Then*

$$\|\partial_y P(i)\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4 + \varepsilon^{2b}\|\mathfrak{I}\|_{s+1}^{\mathrm{Lip}(\gamma)}, \qquad \|\partial_\theta P(i)\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^{6-2b}(1 + \|\mathfrak{I}\|_{s+1}^{\mathrm{Lip}(\gamma)}) \tag{5.3.14}$$

$$\|\nabla_z P(i)\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^{5-b} + \varepsilon^{6-b}\gamma^{-1}\|\mathfrak{I}\|_{s+1}^{\mathrm{Lip}(\gamma)}, \qquad \|X_P(i)\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^{6-2b} + \varepsilon^{2b}\|\mathfrak{I}\|_{s+3}^{\mathrm{Lip}(\gamma)} \tag{5.3.15}$$

$$\|\partial_\theta\partial_y P(i)\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4 + \varepsilon^5\gamma^{-1}\|\mathfrak{I}\|_{s+2}^{\mathrm{Lip}(\gamma)}, \qquad \|\partial_y \nabla_z P(i)\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^{b+3} + \varepsilon^{2b-1}\|\mathfrak{I}\|_{s+2}^{\mathrm{Lip}(\gamma)} \tag{5.3.16}$$

$$\|\partial_{yy} P(i) + 3\varepsilon^{2b} I_{\mathbb{R}^\nu}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^{2+2b} + \varepsilon^{2b+3}\gamma^{-1}\|\mathfrak{I}\|_{s+2}^{\mathrm{Lip}(\gamma)} \tag{5.3.17}$$

and, for all $\widehat{\imath} := (\widehat{\Theta}, \widehat{y}, \widehat{z})$,

$$\|\partial_y d_i X_P(i)[\widehat{\imath}]\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^{2b-1}\big(\|\widehat{\imath}\|_{s+3}^{\mathrm{Lip}(\gamma)} + \|\mathfrak{I}\|_{s+3}^{\mathrm{Lip}(\gamma)}\|\widehat{\imath}\|_{s_0+3}^{\mathrm{Lip}(\gamma)}\big) \tag{5.3.18}$$

$$\|d_i X_{H_\varepsilon}(i)[\widehat{\imath}] + (0, 0, \partial_{xxx}\hat{z})\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon\big(\|\widehat{\imath}\|_{s+3}^{\mathrm{Lip}(\gamma)} + \|\mathfrak{I}\|_{s+3}^{\mathrm{Lip}(\gamma)}\|\widehat{\imath}\|_{s_0+3}^{\mathrm{Lip}(\gamma)}\big) \tag{5.3.19}$$

$$\|d_i^2 X_{H_\varepsilon}(i)[\widehat{\imath}, \widehat{\imath}]\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon\Big(\|\widehat{\imath}\|_{s+3}^{\mathrm{Lip}(\gamma)}\|\widehat{\imath}\|_{s_0+3}^{\mathrm{Lip}(\gamma)} + \|\mathfrak{I}\|_{s+3}^{\mathrm{Lip}(\gamma)}(\|\widehat{\imath}\|_{s_0+3}^{\mathrm{Lip}(\gamma)})^2\Big). \tag{5.3.20}$$

In the sequel we will also use that, by the diophantine condition (5.3.4), the operator $\mathcal{D}_\omega^{-1}$ (see (5.3.7)) is defined for all functions u with zero φ-average, and satisfies

$$\|\mathcal{D}_\omega^{-1} u\|_s \leq C\gamma^{-1}\|u\|_{s+\tau}, \quad \|\mathcal{D}_\omega^{-1} u\|_s^{\mathrm{Lip}(\gamma)} \leq C\gamma^{-1}\|u\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)}. \tag{5.3.21}$$

5.4 Approximate inverse

In order to implement a convergent Nash-Moser scheme that leads to a solution of $\mathcal{F}(i, \zeta) = 0$ our aim is to construct an *approximate right inverse* (which satisfies tame estimates) of the linearized operator

$$d_{i,\zeta}\mathcal{F}(i_0, \zeta_0)[\widehat{\imath}, \widehat{\zeta}] = d_{i,\zeta}\mathcal{F}(i_0)[\widehat{\imath}, \widehat{\zeta}] = \mathcal{D}_\omega \widehat{\imath} - d_i X_{H_\varepsilon}(i_0(\varphi))[\widehat{\imath}] + (0, \widehat{\zeta}, 0), \tag{5.4.1}$$

see Theorem 5.4.1. Note that $d_{i,\zeta}\mathcal{F}(i_0, \zeta_0) = d_{i,\zeta}\mathcal{F}(i_0)$ is independent of ζ_0 (see (5.3.6)).

The notion of approximate right inverse is introduced in [76]. It denotes a linear operator which is an *exact* right inverse at a solution (i_0, ζ_0) of $\mathcal{F}(i_0, \zeta_0) = 0$. We want to implement the general strategy in [22]-[23] which reduces the search of an approximate right inverse of (5.4.1) to the search of a right inverse of the linearized operator $\mathcal{L}_\omega$ (see (5.4.45)) on the normal directions only.

It is well known that an invariant torus i_0 with diophantine flow is isotropic (see e.g. [22]), namely the pull-back 1-form $i_0^*\Lambda$ is closed, where Λ is the contact 1-form in (5.2.4). This is tantamount to say that the 2-form $\mathcal{W}$ (see (5.2.3)) vanishes on the torus $i_0(\mathbb{T}^\nu)$ (i.e. $\mathcal{W}$ vanishes on the tangent space at each point $i_0(\varphi)$ of the manifold $i_0(\mathbb{T}^\nu)$), because $i_0^*\mathcal{W} = i_0^* d\Lambda = d i_0^*\Lambda$. For an "approximately invariant" torus i_0 the 1-form $i_0^*\Lambda$ is only "approximately closed". In order to make this statement quantitative we consider

$$i_0^*\Lambda = \sum\nolimits_{k=1}^{\nu} a_k(\varphi) d\varphi_k, \quad a_k(\varphi) := -\big([\partial_\varphi \theta_0(\varphi)]^T y_0(\varphi)\big)_k + \frac{1}{2}(\partial_{\varphi_k} z_0(\varphi), \partial_x^{-1} z_0(\varphi))_{L^2(\mathbb{T})} \tag{5.4.2}$$

and we quantify how small is

$$i_0^*\mathcal{W} = d\, i_0^*\Lambda = \sum\nolimits_{1 \leq k < j \leq \nu} A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j, \quad A_{kj}(\varphi) := \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi). \tag{5.4.3}$$

Along this Section we will always assume the following hypothesis (which will be verified at each step of the Nash-Moser iteration):

- Assumption. The map $\omega \mapsto i_0(\omega)$ is a Lipschitz function defined on some subset $\Omega_o \subset \Omega_\varepsilon$, where Ω_ε is defined in (5.3.2), and, for some $\mu := \mu(\tau,\nu) > 0$,

$$\|\mathfrak{I}_0\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \le C\varepsilon^{6-2b}\gamma^{-1}, \quad \|Z\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \le C\varepsilon^{6-2b}, \quad \gamma = \varepsilon^{2+a}, \tag{5.4.4}$$
$$b := 1 + (a/2), \quad a \in (0, 1/6),$$

where $\mathfrak{I}_0(\varphi) := i_0(\varphi) - (\varphi, 0, 0)$, and

$$Z(\varphi) := (Z_1, Z_2, Z_3)(\varphi) := \mathcal{F}(i_0, \zeta_0)(\varphi) = \omega \cdot \partial_\varphi i_0(\varphi) - X_{H_{\varepsilon,\zeta_0}}(i_0(\varphi)). \tag{5.4.5}$$

Lemma 5.4.1. $|\zeta_0|^{\mathrm{Lip}(\gamma)} \le C\|Z\|_{s_0}^{\mathrm{Lip}(\gamma)}$. *If $\mathcal{F}(i_0,\zeta_0) = 0$ then $\zeta_0 = 0$, namely the torus i_0 is invariant for X_{H_ε}.*

Proof. The proof is given in Lemma 3-[22]. We give a more direct proof.

Let

$$i_{\psi_0}(\varphi) := (\theta_{\psi_0}(\varphi), y_{\psi_0}(\varphi), z_{\psi_0}(\varphi)) := i_0(\psi_0 + \varphi)$$

be the translated torus embedding, for all $\psi_0 \in \mathbb{T}^\nu$. Since the Hamiltonian H_ε in (5.2.11) is autonomous, the restricted action functional

$$\mathbf{\Phi}(\psi_0) := \int_{\mathbb{T}^\nu} \Big\{ y_{\psi_0}(\varphi) \cdot \mathcal{D}_\omega \theta_{\psi_0}(\varphi) - \frac{1}{2}\big(\partial_x^{-1} z_{\psi_0}(\varphi), \mathcal{D}_\omega z_{\psi_0}(\varphi)\big)_{L^2(\mathbb{T})} - H_\varepsilon(i_{\psi_0}(\varphi)) \Big\}\, d\varphi$$

is constant, namely $\mathbf{\Phi}(\psi_0) = \mathbf{\Phi}(0)$, for all $\psi_0 \in \mathbb{T}^\nu$. Hence $\partial_{\psi_0}\mathbf{\Phi}(\psi_0) = 0$, for all $\psi_0 \in \mathbb{T}^\nu$, in particular differentiating at $\psi_0 = 0$ we get

$$\begin{aligned}
\partial_{\psi_0}\mathbf{\Phi}(0)[\widehat{\psi}] = &+ \int_{\mathbb{T}^\nu} \partial_\varphi y_0(\varphi)[\widehat{\psi}] \cdot \mathcal{D}_\omega \theta_0(\varphi)\, d\varphi + \int_{\mathbb{T}^\nu} y_0(\varphi) \cdot \mathcal{D}_\omega \partial_\varphi \theta_0(\varphi)[\widehat{\psi}]\, d\varphi \\
&- \frac{1}{2}\int_{\mathbb{T}^\nu} \big(\partial_x^{-1}\partial_\varphi z_0(\varphi)[\widehat{\psi}], \mathcal{D}_\omega z_0(\varphi)\big)_{L^2(\mathbb{T})}\, d\varphi - \frac{1}{2}\int_{\mathbb{T}^\nu} \big(\partial_x^{-1} z_0(\varphi), \mathcal{D}_\omega \partial_\varphi z_0(\varphi)[\widehat{\psi}]\big)_{L^2(\mathbb{T})}\, d\varphi \\
&- \int_{\mathbb{T}^\nu} \Big\{ \partial_\theta H_\varepsilon(i_0(\varphi)) \cdot \partial_\varphi \theta_0(\varphi)[\widehat{\psi}] + \partial_y H_\varepsilon(i_0(\varphi)) \cdot \partial_\varphi y_0(\varphi)[\widehat{\psi}] \Big\}\, d\varphi \\
&- \int_{\mathbb{T}^\nu} \big(\nabla_z H_\varepsilon(i_0(\varphi)), \partial_\varphi z_0(\varphi)[\widehat{\psi}]\big)_{L^2(\mathbb{T})}\, d\varphi\,.
\end{aligned} \tag{5.4.6}$$

By (5.3.6), (5.4.5) one has

$$\int_{\mathbb{T}^\nu} \big(\mathcal{D}_\omega \theta_0(\varphi) - \partial_y H_\varepsilon(i_0(\varphi))\big) \cdot \partial_\varphi y_0(\varphi)[\widehat{\psi}]\, d\varphi = \int_{\mathbb{T}^\nu} [\partial_\varphi y_0(\varphi)]^T Z_1(\varphi)\, d\varphi \cdot \widehat{\psi}\,, \tag{5.4.7}$$

$$\begin{aligned}
\int_{\mathbb{T}^\nu} y_0(\varphi) \cdot \mathcal{D}_\omega \partial_\varphi \theta_0(\varphi)[\widehat{\psi}] - \partial_\theta H_\varepsilon(i_0(\varphi)) \cdot \partial_\varphi \theta_0(\varphi)[\widehat{\psi}]\, d\varphi &= -\int_{\mathbb{T}^\nu} \big(\mathcal{D}_\omega y_0(\varphi) + \partial_\theta H_\varepsilon(i_0(\varphi))\big) \cdot \partial_\varphi \theta_0(\varphi)[\widehat{\psi}] \\
&= -\int_{\mathbb{T}^\nu} [\partial_\varphi \theta_0(\varphi)]^T Z_2(\varphi)\, d\varphi \cdot \widehat{\psi} \\
&\quad + \zeta_0 \cdot \widehat{\psi}
\end{aligned} \tag{5.4.8}$$

where in the last equality we have used that

$$\int_{\mathbb{T}^\nu} [\partial_\varphi \theta_0(\varphi)]^T\, d\varphi = I + \int_{\mathbb{T}^\nu} [\partial_\varphi \Theta_0(\varphi)]^T\, d\varphi = I\,,$$

since $\Theta_0(\varphi)$ is a 2π-periodic in all its components ($I : \mathbb{R}^\nu \to \mathbb{R}^\nu$ is the identity).

Moreover since $(\partial_x^{-1})^T = -\partial_x^{-1}$ and integrating by parts in $d\varphi$, we get

$$
\begin{aligned}
-\frac{1}{2}\int_{\mathbb{T}^\nu} \big(\partial_x^{-1}\partial_\varphi z_0(\varphi)[\widehat{\psi}], \mathcal{D}_\omega z_0(\varphi)\big)_{L^2(\mathbb{T})} &+ \big(\partial_x^{-1} z_0(\varphi), \mathcal{D}_\omega \partial_\varphi z_0(\varphi)[\widehat{\psi}]\big)_{L^2(\mathbb{T})}\, d\varphi \\
&= \int_{\mathbb{T}^\nu} \big(\partial_x^{-1}\mathcal{D}_\omega z_0(\varphi), \partial_\varphi z_0(\varphi)[\widehat{\psi}]\big)_{L^2(\mathbb{T})}\, d\varphi \\
&= \int_{\mathbb{T}^\nu} [\partial_\varphi z_0(\varphi)]^T \nabla_z H_\varepsilon(i_0(\varphi))\, d\varphi \cdot \widehat{\psi} \\
&+ \int_{\mathbb{T}^\nu} [\partial_\varphi z_0(\varphi)]^T \partial_x^{-1} Z_3(\varphi)\, d\varphi \cdot \widehat{\psi}\,.
\end{aligned} \tag{5.4.9}
$$

Hence by (5.4.6)-(5.4.9), $\partial_{\psi_0}\mathbf{\Phi}(0) = 0$ we get

$$
\zeta_0 = \int_{\mathbb{T}^\nu} -[\partial_\varphi y_0(\varphi)]^T Z_1(\varphi) + [\partial_\varphi \theta_0(\varphi)]^T Z_2(\varphi) - [\partial_\varphi z_0(\varphi)]^T \partial_x^{-1} Z_3(\varphi)\, d\varphi
$$

and the lemma follows by (5.4.4) and usual algebra estimate. □

We now quantify the size of $i_0^*\mathcal{W}$ in terms of Z.

Lemma 5.4.2. *The coefficients $A_{kj}(\varphi)$ in* (5.4.3) *satisfy*

$$
\|A_{kj}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \gamma^{-1}\big(\|Z\|_{s+2\tau+2}^{\mathrm{Lip}(\gamma)} \|\mathfrak{I}_0\|_{s_0+1}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+1}^{\mathrm{Lip}(\gamma)} \|\mathfrak{I}_0\|_{s+2\tau+2}^{\mathrm{Lip}(\gamma)}\big)\,. \tag{5.4.10}
$$

Proof. We estimate the coefficients of the Lie derivative $L_\omega(i_0^*\mathcal{W}) := \sum_{k<j} \mathcal{D}_\omega A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j$. Denoting by $\underline{e}_k$ the k-th versor of $\mathbb{R}^\nu$ we have

$$
\mathcal{D}_\omega A_{kj} = L_\omega(i_0^*\mathcal{W})(\varphi)[\underline{e}_k, \underline{e}_j] = \mathcal{W}\big(\partial_\varphi Z(\varphi)\underline{e}_k, \partial_\varphi i_0(\varphi)\underline{e}_j\big) + \mathcal{W}\big(\partial_\varphi i_0(\varphi)\underline{e}_k, \partial_\varphi Z(\varphi)\underline{e}_j\big)
$$

(see Lemma 5 in [22]). Hence

$$
\|\mathcal{D}_\omega A_{kj}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|Z\|_{s+1}^{\mathrm{Lip}(\gamma)} \|\mathfrak{I}_0\|_{s_0+1}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+1}^{\mathrm{Lip}(\gamma)} \|\mathfrak{I}_0\|_{s+1}^{\mathrm{Lip}(\gamma)}\,. \tag{5.4.11}
$$

The bound (5.4.10) follows applying $\mathcal{D}_\omega^{-1}$ and using (5.4.3), (5.3.21). □

As in [22] we first modify the approximate torus i_0 to obtain an isotropic torus i_δ which is still approximately invariant.

First we report some basic facts about differential 1-forms on the torus $\mathbb{T}^\nu$. We regard a 1-form $a := \sum_{k=1}^\nu a_k(\varphi)\, d\varphi$ equivalently as the vector field

$$
\vec{a}(\varphi) := (a_1(\varphi), \dots, a_\nu(\varphi))\,, \qquad \varphi \in \mathbb{T}^\nu\,.
$$

We denote the Laplacian $\Delta_\varphi := \sum_{k=1}^\nu \partial_{\varphi_k}^2$. The following Lemma is proved in [22]-Lemma 4.

Lemma 5.4.3 (Helmotz decomposition). *Let $\vec{a}$ be a smooth vector field on the torus $\mathbb{T}^\nu$, then*

$$
\vec{a} = \nabla U + \vec{c} + \rho\,, \tag{5.4.12}
$$

where

$$
U : \mathbb{T}^\nu \to \mathbb{R}\,, \qquad \vec{c} \in \mathbb{R}^\nu\,,
$$

and $\rho(\varphi) := (\rho_1(\varphi), \dots, \rho_\nu(\varphi))$, $\varphi \in \mathbb{T}^\nu$ *is a smooth vector field on* $\mathbb{T}^\nu$ *satisfying*

$$\mathrm{div}\rho := \sum_{k=1}^{\nu} \partial_{\varphi_k} \rho_k = 0\,, \qquad \int_{\mathbb{T}^\nu} \rho(\varphi)\, d\varphi = 0\,.$$

The above decomposition of the vector field $\vec{a}$ *is unique if we impose that* $\int_{\mathbb{T}^\nu} U(\varphi)\, d\varphi = 0$. *We have* $U := \Delta_\varphi^{-1}(\mathrm{div}\vec{a})$, *for all* $j = 1, \dots, \nu$, *the* j*-th component* ρ_j *of the vector field* ρ *satisfies*

$$\rho_j(\varphi) = \Delta_\varphi^{-1}\Big(\sum_{k=1}^{\nu} \partial_{\varphi_k} A_{kj}(\varphi)\Big)\,, \quad A_{kj} := \partial_{\varphi_k} a_j - \partial_{\varphi_j} a_k\,, \tag{5.4.13}$$

and the j*-th component* c_j *of vector* $\vec{c}$ *satisfies*

$$c_j := \int_{\mathbb{T}^\nu} a_j(\varphi)\, d\varphi\,, \quad \forall j = 1, \dots, \nu\,. \tag{5.4.14}$$

Proof. Note that

$$\mathrm{div}(\nabla U - \vec{a}) = 0 \quad \Longleftrightarrow \quad \Delta_\varphi U = \mathrm{div}\vec{a}\,.$$

Since $\mathrm{div}\vec{a}$ has zero average, the above equation has the solution

$$U := \Delta_\varphi^{-1}(\mathrm{div}\vec{a})\,,$$

and the decomposition (5.4.12) holds, defining

$$\rho := \vec{a} - \nabla U - \vec{c}. \tag{5.4.15}$$

Moreover

$$\int_{\mathbb{T}^\nu} \rho(\varphi)\, d\varphi = \int_{\mathbb{T}^\nu} \vec{a}(\varphi)\, d\varphi - \vec{c} = 0\,,$$

choosing $\vec{c} = (c_1, \dots, c_\nu)$ as in (5.4.14).

Now let us prove (5.4.13). By (5.4.15) we have

$$\partial_{\varphi_k}\rho_j - \partial_{\varphi_j}\rho_k = \partial_{\varphi_k} a_j - \partial_{\varphi_j} a_k =: A_{kj}\,,$$

since by the Schwartz Theorem $\partial_{\varphi_k\varphi_j} U = \partial_{\varphi_j\varphi_k} U$. For all $j = 1, \dots, \nu$ we differentiate with respect to φ_k and we get

$$\partial_{\varphi_k\varphi_k}\rho_j - \partial_{\varphi_j\varphi_k}\rho_k = \partial_{\varphi_k} A_{kj}\,.$$

Summing over $k = 1, \dots, \nu$ we get

$$\Delta_\varphi \rho_j - \sum_{k=1}^{\nu} \partial_{\varphi_j\varphi_k}\rho_k = \sum_{k=1}^{\nu} \partial_{\varphi_k} A_{kj}\,,$$

but since

$$\sum_{k=1}^{\nu} \partial_{\varphi_j\varphi_k}\rho_k = \partial_{\varphi_j} \sum_{k=1}^{\nu} \partial_{\varphi_k}\rho_k = \partial_{\varphi_j}\mathrm{div}\rho = 0\,,$$

we get for all $j = 1, \dots, \nu$

$$\Delta_\varphi \rho_j = \sum_{k=1}^{\nu} \partial_{\varphi_k} A_{kj}\,,$$

which implies (5.4.13), since the right hand side of the above equation has zero average.

□

By lemma 5.4.3 we deduce immediately the following corollary.

Corollary 5.4.1. *Let $a := \sum_{k=1}^{\nu} a_k(\varphi)\, d\varphi_k$ be a smooth differential 1-form on the torus $\mathbb{T}^\nu$, and let $\rho(\varphi) := (\rho_1(\varphi), \dots, \rho_\nu(\varphi))$ be defined by* (5.4.13)*. Then the 1-form $a - \sum_{k=1}^{\nu} \rho_k(\varphi)\, d\varphi_k$ is closed.*

In the next Lemma we show how to modify the approximate invariant torus i_0, in order to obtain an isotropic approximate invariant torus i_δ.

Lemma 5.4.4. (Isotropic torus) *The torus $i_\delta(\varphi) := (\theta_0(\varphi), y_\delta(\varphi), z_0(\varphi))$ defined by*

$$y_\delta := y_0 + [\partial_\varphi \theta_0(\varphi)]^{-T} \rho(\varphi)\,, \qquad \rho_j(\varphi) := \Delta_\varphi^{-1} \sum\nolimits_{k=1}^{\nu} \partial_{\varphi_j} A_{kj}(\varphi) \tag{5.4.16}$$

is isotropic. If (5.4.4) *holds, then, for some $\sigma := \sigma(\nu, \tau)$,*

$$\|y_\delta - y_0\|_s^{\mathrm{Lip}(\gamma)} \leq_s \gamma^{-1} \big(\|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \|\mathfrak{I}_0\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \|\mathfrak{I}_0\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \big)\,, \tag{5.4.17}$$

$$\|\mathcal{F}(i_\delta, \zeta_0)\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)} \|\mathfrak{I}_0\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \tag{5.4.18}$$

$$\|\partial_i [i_\delta][\widehat{\imath}]\|_s \leq_s \|\widehat{\imath}\|_s + \|\mathfrak{I}_0\|_{s+\sigma} \|\widehat{\imath}\|_s\,. \tag{5.4.19}$$

We denote equivalently the differential by ∂_i or d_i. Moreover we denote by $\sigma := \sigma(\nu, \tau)$ possibly different (larger) "loss of derivatives" constants.

Proof. In this proof we write $\| \ \|_s$ to denote $\| \ \|_s^{\mathrm{Lip}(\gamma)}$. By definitions (5.4.2), (5.4.3), (5.4.16) we have

$$i_\delta^* \Lambda = i_0^* \Lambda - \sum_{k=1}^{\nu} \rho_k(\varphi)\, d\varphi_k\,.$$

Applying corollary 5.4.1, we get that $i_\delta^* \Lambda$ is closed, which implies the isotropy of the embedded torus $\varphi \to i_\delta(\varphi)$. The estimate (5.4.17) follows by (5.4.16), (5.4.10), (5.4.4) and the tame bound for the inverse $\|[\partial_\varphi \theta_0]^{-T}\|_s \leq_s 1 + \|\mathfrak{I}_0\|_{s+1}$. It remains to estimate the difference (see (5.3.6) and note that $X_{\mathcal{N}}$ does not depend on y)

$$\mathcal{F}(i_\delta, \zeta_0) - \mathcal{F}(i_0, \zeta_0) = \begin{pmatrix} 0 \\ \mathcal{D}_\omega (y_\delta - y_0) \\ 0 \end{pmatrix} + X_P(i_\delta) - X_P(i_0). \tag{5.4.20}$$

Using (5.3.16), (5.3.17), we get $\|\partial_y X_P(i)\|_s \leq_s \varepsilon^{2b} + \varepsilon^{2b-1} \|\mathfrak{I}\|_{s+3}$. Hence (5.4.17), (5.4.4) imply

$$\|X_P(i_\delta) - X_P(i_0)\|_s \leq_s \|\mathfrak{I}_0\|_{s_0+\sigma} \|Z\|_{s+\sigma} + \|\mathfrak{I}_0\|_{s+\sigma} \|Z\|_{s_0+\sigma}\,. \tag{5.4.21}$$

Differentiating (5.4.16) we have

$$\mathcal{D}_\omega (y_\delta - y_0) = [\partial_\varphi \theta_0(\varphi)]^{-T} \mathcal{D}_\omega \rho(\varphi) + (\mathcal{D}_\omega [\partial_\varphi \theta_0(\varphi)]^{-T}) \rho(\varphi) \tag{5.4.22}$$

and $\mathcal{D}_\omega \rho_j(\varphi) = \Delta_\varphi^{-1} \sum_{k=1}^{\nu} \partial_{\varphi_j} \mathcal{D}_\omega A_{kj}(\varphi)$. Using (5.4.11), we deduce that

$$\|[\partial_\varphi \theta_0]^{-T} \mathcal{D}_\omega \rho\|_s \leq_s \|Z\|_{s+1} \|\mathfrak{I}_0\|_{s_0+1} + \|Z\|_{s_0+1} \|\mathfrak{I}_0\|_{s+1}\,. \tag{5.4.23}$$

To estimate the second term in (5.4.22), we differentiate $Z_1(\varphi) = \mathcal{D}_\omega \theta_0(\varphi) - \omega - (\partial_y P)(i_0(\varphi))$ (which is the first component in (5.3.6)) with respect to φ. We get $\mathcal{D}_\omega \partial_\varphi \theta_0(\varphi) = \partial_\varphi (\partial_y P)(i_0(\varphi)) + \partial_\varphi Z_1(\varphi)$. Then, by (5.3.14),

$$\|\mathcal{D}_\omega [\partial_\varphi \theta_0]^T\|_s \leq_s \varepsilon^4 + \varepsilon^{2b} \|\mathfrak{I}_0\|_{s+2} + \|Z\|_{s+1} \,. \tag{5.4.24}$$

Since $\mathcal{D}_\omega [\partial_\varphi \theta_0(\varphi)]^{-T} = -[\partial_\varphi \theta_0(\varphi)]^{-T} \big(\mathcal{D}_\omega [\partial_\varphi \theta_0(\varphi)]^T\big) [\partial_\varphi \theta_0(\varphi)]^{-T}$, the bounds (5.4.24), (5.4.10), (5.4.4) imply

$$\|(\mathcal{D}_\omega [\partial_\varphi \theta_0]^{-T}) \rho\|_s \leq_s \varepsilon^{6-2b} \gamma^{-1} \big(\|Z\|_{s+\sigma} \|\mathfrak{I}_0\|_{s_0+\sigma} + \|Z\|_{s_0+\sigma} \|\mathfrak{I}_0\|_{s+\sigma}\big) \,. \tag{5.4.25}$$

In conclusion (5.4.20), (5.4.21), (5.4.22), (5.4.23), (5.4.25) imply (5.4.18). The bound (5.4.19) follows by (5.4.16), (5.4.3), (5.4.2), (5.4.4). $\square$

Note that there is no γ^{-1} in the right hand side of (5.4.18). It turns out that an approximate inverse of $d_{i,\zeta} \mathcal{F}(i_\delta)$ is an approximate inverse of $d_{i,\zeta} \mathcal{F}(i_0)$ as well. In order to find an approximate inverse of the linearized operator $d_{i,\zeta} \mathcal{F}(i_\delta)$ we introduce a suitable set of symplectic coordinates nearby the isotropic torus i_δ. We consider the map $G_\delta : (\psi, \eta, w) \to (\theta, y, z)$ of the phase space $\mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$ defined by

$$\begin{pmatrix} \theta \\ y \\ z \end{pmatrix} := G_\delta \begin{pmatrix} \psi \\ \eta \\ w \end{pmatrix} := \begin{pmatrix} \theta_0(\psi) \\ y_\delta(\psi) + [\partial_\psi \theta_0(\psi)]^{-T} \eta + \big[(\partial_\theta \tilde{z}_0)(\theta_0(\psi))\big]^T \partial_x^{-1} w \\ z_0(\psi) + w \end{pmatrix} \tag{5.4.26}$$

where $\tilde{z}_0(\theta) := z_0(\theta_0^{-1}(\theta))$. Now we prove that G_δ is symplectic, using the isotropy of the embedded torus i_δ (Lemma 5.4.4). This proof is given in [22]-Lemma 2.

Lemma 5.4.5. *The transformation G_δ defined in* (5.4.26) *is symplectic.*

Proof. We may regard G_δ as the composition $G_\delta = G_2 \circ G_1$ of the diffeomorphisms

$$G_1 : \begin{pmatrix} \psi \\ \eta \\ w \end{pmatrix} \to \begin{pmatrix} \theta_0(\psi) \\ [\partial_\psi \theta_0(\psi)]^{-T} \eta \\ w \end{pmatrix} \tag{5.4.27}$$

and

$$G_2 : \begin{pmatrix} \theta \\ y \\ z \end{pmatrix} \to \begin{pmatrix} \theta \\ \widetilde{y}_\delta(\theta) + y + [\partial_\theta \tilde{z}_0(\theta)]^T \partial_x^{-1} z \\ \tilde{z}_0(\theta) + z \end{pmatrix} , \tag{5.4.28}$$

where $\widetilde{y}_\delta := y_\delta \circ \theta_0^{-1}$, $\widetilde{z}_0 := z_0 \circ \theta_0^{-1}$. We claim that G_1 and G_2 are symplectic and the lemma follows. To prove this we show that $G_i^* \Lambda = \Lambda$, $i = 1, 2$, where Λ is the contact 1-form defined in (5.2.4). Since $d\Lambda = \mathcal{W}$ and the differential commutes with the pull back, we get

$$G_i^* \mathcal{W} = G_i^* d\Lambda = d(G_i^* \Lambda) = d\Lambda = \mathcal{W} \,,$$

which implies that G_1 and G_2 are symplectic.

G_1 IS SYMPLECTIC. By the definition of the contact 1-form Λ in (5.2.4), one has

$$(G_1^*\Lambda)_{(\psi,\eta,w)}[\widehat{\theta},\widehat{\eta},\widehat{w}] = -[\partial_\psi\theta_0(\psi)]^{-T}\eta\cdot\partial_\psi\theta_0(\psi)[\widehat{\psi}] + \frac{1}{2}\big(\partial_x^{-1}w,\widehat{w}\big)_{L^2(\mathbb{T})}$$
$$= -\eta\cdot\widehat{\psi} + \frac{1}{2}\big(\partial_x^{-1}w,\widehat{w}\big)_{L^2(\mathbb{T})},$$

hence $G_1^*\Lambda = \Lambda$ and G_1 is symplectic.

G_2 IS SYMPLECTIC. We calculate

$$(G_2^*\Lambda)_{(\theta,y,z)}[\widehat{\theta},\widehat{y},\widehat{z}] = -\Big(\widetilde{y}_\delta(\theta) + y + [\partial_\theta\tilde{z}_0(\theta)]^T\partial_x^{-1}z\Big)\cdot\widehat{\theta}$$
$$+\frac{1}{2}\big(\partial_x^{-1}\{\tilde{z}_0(\theta)+z\},\widehat{z}+\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}]\big)_{L^2(\mathbb{T})},$$

therefore

$$(G_2^*\Lambda-\Lambda)_{(\theta,y,z)}[\widehat{\theta},\widehat{y},\widehat{z}] = -\Big(\widetilde{y}_\delta(\theta) + [\partial_\theta\tilde{z}_0(\theta)]^T\partial_x^{-1}z\Big)\cdot\widehat{\theta} + \frac{1}{2}\big(\partial_x^{-1}\tilde{z}_0(\theta),\widehat{z}\big)_{L^2(\mathbb{T})}$$
$$+\frac{1}{2}\big(\partial_x^{-1}\tilde{z}_0(\theta),\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}]\big)_{L^2(\mathbb{T})} + \frac{1}{2}\big(\partial_x^{-1}z,\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}]\big)_{L^2(\mathbb{T})}. \tag{5.4.29}$$

Since

$$-\big([\partial_\theta\tilde{z}_0(\theta)]^T\partial_x^{-1}z\big)\cdot\widehat{\theta}+\frac{1}{2}\big(\partial_x^{-1}z,\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}]\big)_{L^2(\mathbb{T})} = -\frac{1}{2}\big(\partial_x^{-1}z,\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}]\big)_{L^2(\mathbb{T})} = \frac{1}{2}\big(\partial_x^{-1}\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}],z\big)_{L^2(\mathbb{T})},$$

(using $(\partial_x^{-1})^T = -\partial_x^{-1}$) by (5.4.29) we get

$$(G_2^*\Lambda-\Lambda)_{(\theta,y,z)}[\widehat{\theta},\widehat{y},\widehat{z}] = -\widetilde{y}_\delta(\theta)\cdot\widehat{\theta} + \frac{1}{2}\big(\partial_x^{-1}\tilde{z}_0(\theta),\widehat{z}\big)_{L^2(\mathbb{T})} + \frac{1}{2}\big(\partial_x^{-1}\tilde{z}_0(\theta),\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}]\big)_{L^2(\mathbb{T})}$$
$$+\frac{1}{2}\big(\partial_x^{-1}\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}],z\big)_{L^2(\mathbb{T})}. \tag{5.4.30}$$

Note that the 1-form

$$(\widehat{\theta},\widehat{y},\widehat{z}) \to \big(\partial_x^{-1}\tilde{z}_0(\theta),\widehat{z}\big)_{L^2(\mathbb{T})} + \big(\partial_x^{-1}\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}],z\big)_{L^2(\mathbb{T})} = d\Big(\big(\partial_x^{-1}\tilde{z}_0(\theta),z\big)_{L^2(\mathbb{T})}\Big)[\widehat{\theta},\widehat{y},\widehat{z}] \tag{5.4.31}$$

is exact. Moreover

$$-\widetilde{y}_\delta(\theta)\cdot\widehat{\theta} + \frac{1}{2}\big(\partial_x^{-1}\tilde{z}_0(\theta),\partial_\theta\tilde{z}_0(\theta)[\widehat{\theta}]\big)_{L^2(\mathbb{T})} = (j^*\Lambda)_\theta[\widehat{\theta}], \tag{5.4.32}$$

where

$$j := i_\delta\circ\theta_0^{-1} : \theta \to (\theta,\widetilde{y}_\delta(\theta),\widetilde{z}_0(\theta)).$$

Then (5.4.30)-(5.4.32) imply

$$(G_2^*\Lambda-\Lambda)_{(\theta,y,z)} = \Pi_{\mathbb{T}^\nu}^*(j^*\Lambda)_{(\theta,y,z)} + \frac{1}{2}d\Big(\big(\partial_x^{-1}\tilde{z}_0(\theta),z\big)_{L^2(\mathbb{T})}\Big),$$

where

$$\Pi_{\mathbb{T}^\nu} : \mathbb{T}^\nu\times\mathbb{R}^\nu\times H_S^\perp \to \mathbb{T}^\nu$$

is the canonical projection. Since the torus $j(\mathbb{T}^\nu) = i_\delta(\mathbb{T}^\nu)$ is isotropic (see Lemma 5.4.4), the 1-form $j^*\Lambda$ on $\mathbb{T}^\nu$ is closed and the lemma follows. □

In the new coordinates, i_δ is the trivial embedded torus $(\psi,\eta,w)=(\psi,0,0)$. The transformed Hamiltonian $K:=K(\psi,\eta,w,\zeta_0)$ is (recall (5.3.5))

$$
\begin{aligned}
K := H_{\varepsilon,\zeta_0}\circ G_\delta = \theta_0(\psi)\cdot\zeta_0 + K_{00}(\psi) + K_{10}(\psi)\cdot\eta + (K_{01}(\psi),w)_{L^2(\mathbb{T})} + \frac{1}{2}K_{20}(\psi)\eta\cdot\eta \\
+ \big(K_{11}(\psi)\eta,w\big)_{L^2(\mathbb{T})} + \frac{1}{2}\big(K_{02}(\psi)w,w\big)_{L^2(\mathbb{T})} + K_{\geq 3}(\psi,\eta,w)
\end{aligned}
\tag{5.4.33}
$$

where $K_{\geq 3}$ collects the terms at least cubic in the variables (η,w). At any fixed ψ, the Taylor coefficient $K_{00}(\psi)\in\mathbb{R}$, $K_{10}(\psi)\in\mathbb{R}^\nu$, $K_{01}(\psi)\in H_S^\perp$ (it is a function of $x\in\mathbb{T}$), $K_{20}(\psi)$ is a $\nu\times\nu$ real matrix, $K_{02}(\psi)$ is a linear self-adjoint operator of $H_S^\perp$ and $K_{11}(\psi):\mathbb{R}^\nu\to H_S^\perp$. Note that the above Taylor coefficients do not depend on the parameter ζ_0.

The Hamilton equations associated to (5.4.33) are

$$
\begin{cases}
\dot\psi = K_{10}(\psi) + K_{20}(\psi)\eta + K_{11}^T(\psi)w + \partial_\eta K_{\geq 3}(\psi,\eta,w) \\
\dot\eta = -[\partial_\psi\theta_0(\psi)]^T\zeta_0 - \partial_\psi K_{00}(\psi) - [\partial_\psi K_{10}(\psi)]^T\eta - [\partial_\psi K_{01}(\psi)]^T w \\
\qquad -\partial_\psi\big(\frac{1}{2}K_{20}(\psi)\eta\cdot\eta + (K_{11}(\psi)\eta,w)_{L^2(\mathbb{T})} + \frac{1}{2}(K_{02}(\psi)w,w)_{L^2(\mathbb{T})} + K_{\geq 3}(\psi,\eta,w)\big) \\
\dot w = \partial_x\big(K_{01}(\psi) + K_{11}(\psi)\eta + K_{02}(\psi)w + \nabla_w K_{\geq 3}(\psi,\eta,w)\big)
\end{cases}
\tag{5.4.34}
$$

where $[\partial_\psi K_{10}(\psi)]^T$ is the $\nu\times\nu$ transposed matrix and $[\partial_\psi K_{01}(\psi)]^T$, $K_{11}^T(\psi):H_S^\perp\to\mathbb{R}^\nu$ are defined by the duality relation $(\partial_\psi K_{01}(\psi)[\hat\psi],w)_{L^2} = \hat\psi\cdot[\partial_\psi K_{01}(\psi)]^T w$, $\forall\hat\psi\in\mathbb{R}^\nu, w\in H_S^\perp$, and similarly for K_{11}. Explicitly, for all $w\in H_S^\perp$, and denoting $\underline{e}_k$ the k-th versor of $\mathbb{R}^\nu$,

$$
K_{11}^T(\psi)w = \sum\nolimits_{k=1}^{\nu}\big(K_{11}^T(\psi)w\cdot\underline{e}_k\big)\underline{e}_k = \sum\nolimits_{k=1}^{\nu}\big(w,K_{11}(\psi)\underline{e}_k\big)_{L^2(\mathbb{T})}\underline{e}_k \in\mathbb{R}^\nu\,. \tag{5.4.35}
$$

In the next lemma we estimate the coefficients K_{00}, K_{10}, K_{01} in the Taylor expansion (5.4.33). Note that on an exact solution we have $Z=0$ and therefore $K_{00}(\psi)=\text{const}$, $K_{10}=\omega$ and $K_{01}=0$.

Lemma 5.4.6. *Assume* (5.4.4). *Then there is* $\sigma:=\sigma(\tau,\nu)$ *such that*

$$
\|\partial_\psi K_{00}\|_s^{\mathrm{Lip}(\gamma)} + \|K_{10}-\omega\|_s^{\mathrm{Lip}(\gamma)} + \|K_{01}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|Z\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|\mathfrak{I}_0\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\,.
$$

Proof. Let $\mathcal{F}(i_\delta,\zeta_0):=Z_\delta:=(Z_{1,\delta},Z_{2,\delta},Z_{3,\delta})$. By a direct calculation as in [22] (using (5.4.33), (5.3.6))

$$
\begin{aligned}
\partial_\psi K_{00}(\psi) &= -[\partial_\psi\theta_0(\psi)]^T\big(\zeta_0 - Z_{2,\delta} - [\partial_\psi y_\delta][\partial_\psi\theta_0]^{-1}Z_{1,\delta} + [(\partial_\theta\tilde z_0)(\theta_0(\psi))]^T\partial_x^{-1}Z_{3,\delta} \\
&\qquad + [(\partial_\theta\tilde z_0)(\theta_0(\psi))]^T\partial_x^{-1}\partial_\psi z_0(\psi)[\partial_\psi\theta_0(\psi)]^{-1}Z_{1,\delta}\big)\,, \\
K_{10}(\psi) &= \omega - [\partial_\psi\theta_0(\psi)]^{-1}Z_{1,\delta}(\psi)\,, \\
K_{01}(\psi) &= -\partial_x^{-1}Z_{3,\delta} + \partial_x^{-1}\partial_\psi z_0(\psi)[\partial_\psi\theta_0(\psi)]^{-1}Z_{1,\delta}(\psi)\,.
\end{aligned}
$$

Then (5.4.4), (5.4.17), (5.4.18) and Lemma 5.4.1 (using Lemma A.0.10) imply the lemma. ☐

Remark 5.4.1. *If* $\mathcal{F}(i_0,\zeta_0)=0$ *then* $\zeta_0=0$ *by Lemma 5.4.1, and Lemma 5.4.6 implies that* (5.4.33) *simplifies to* $K = const + \omega\cdot\eta + \frac{1}{2}K_{20}(\psi)\eta\cdot\eta + \big(K_{11}(\psi)\eta,w\big)_{L^2(\mathbb{T})} + \frac{1}{2}\big(K_{02}(\psi)w,w\big)_{L^2(\mathbb{T})} + K_{\geq 3}$.

We now estimate K_{20},K_{11} in (5.4.33). The norm of K_{20} is the sum of the norms of its matrix entries.

Lemma 5.4.7. *Assume* (5.4.4). *Then*

$$\|K_{20} + 3\varepsilon^{2b} I\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^{2b+2} + \varepsilon^{2b}\|\mathfrak{I}_0\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \varepsilon^3\gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \tag{5.4.36}$$

$$\begin{aligned}\|K_{11}\eta\|_s^{\mathrm{Lip}(\gamma)} &\leq_s \varepsilon^5\gamma^{-1}\|\eta\|_s^{\mathrm{Lip}(\gamma)} \\ &\quad + \varepsilon^{2b-1}(\|\mathfrak{I}_0\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)})\|\eta\|_{s_0}^{\mathrm{Lip}(\gamma)}\end{aligned} \tag{5.4.37}$$

$$\begin{aligned}\|K_{11}^T w\|_s^{\mathrm{Lip}(\gamma)} &\leq_s \varepsilon^5\gamma^{-1}\|w\|_{s+2}^{\mathrm{Lip}(\gamma)} \\ &\quad + \varepsilon^{2b-1}(\|\mathfrak{I}_0\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}\|Z\|_{s+\sigma}^{\mathrm{Lip}(\gamma)})\|w\|_{s_0+2}^{\mathrm{Lip}(\gamma)}.\end{aligned} \tag{5.4.38}$$

In particular $\|K_{20} + 3\varepsilon^{2b} I\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^6\gamma^{-1}$, *and*

$$\|K_{11}\eta\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^5\gamma^{-1}\|\eta\|_{s_0}^{\mathrm{Lip}(\gamma)}, \quad \|K_{11}^T w\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^5\gamma^{-1}\|w\|_{s_0}^{\mathrm{Lip}(\gamma)}.$$

Proof. To shorten the notation, in this proof we write $\| \ \|_s$ for $\| \ \|_s^{\mathrm{Lip}(\gamma)}$. We have

$$K_{20}(\varphi) = [\partial_\varphi\theta_0(\varphi)]^{-1}\partial_{yy}H_\varepsilon(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T} = [\partial_\varphi\theta_0(\varphi)]^{-1}\partial_{yy}P(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T}.$$

Then (5.3.17), (5.4.4), (5.4.17) imply (5.4.36). Now (see also [22])

$$\begin{aligned}K_{11}(\varphi) &= \partial_y\nabla_z H_\varepsilon(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T} - \partial_x^{-1}(\partial_\theta\tilde z_0)(\theta_0(\varphi))(\partial_{yy}H_\varepsilon)(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T} \\ &\overset{(5.2.11)}{=} \partial_y\nabla_z P(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T} - \partial_x^{-1}(\partial_\theta\tilde z_0)(\theta_0(\varphi))(\partial_{yy}P)(i_\delta(\varphi))[\partial_\varphi\theta_0(\varphi)]^{-T},\end{aligned}$$

therefore, using (5.3.16), (5.3.17), (5.4.4), we deduce (5.4.37). The bound (5.4.38) for K_{11}^T follows by (5.4.35). □

Under the linear change of variables

$$DG_\delta(\varphi,0,0)\begin{pmatrix}\widehat\psi \\ \widehat\eta \\ \widehat w\end{pmatrix} := \begin{pmatrix}\partial_\psi\theta_0(\varphi) & 0 & 0 \\ \partial_\psi y_\delta(\varphi) & [\partial_\psi\theta_0(\varphi)]^{-T} & -[(\partial_\theta\tilde z_0)(\theta_0(\varphi))]^T\partial_x^{-1} \\ \partial_\psi z_0(\varphi) & 0 & I\end{pmatrix}\begin{pmatrix}\widehat\psi \\ \widehat\eta \\ \widehat w\end{pmatrix} \tag{5.4.39}$$

the linearized operator $d_{i,\zeta}\mathcal{F}(i_\delta)$ transforms (approximately, see (5.4.59)) into the operator obtained linearizing (5.4.34) at $(\psi,\eta,w,\zeta) = (\varphi,0,0,\zeta_0)$ (with $\partial_t \rightsquigarrow \mathcal{D}_\omega$), namely

$$\begin{pmatrix}\mathcal{D}_\omega\widehat\psi - \partial_\psi K_{10}(\varphi)[\widehat\psi\,] - K_{20}(\varphi)\widehat\eta - K_{11}^T(\varphi)\widehat w \\ \mathcal{D}_\omega\widehat\eta + [\partial_\psi\theta_0(\varphi)]^T\widehat\zeta + \partial_\psi[\partial_\psi\theta_0(\varphi)]^T[\widehat\psi,\zeta_0] + \partial_{\psi\psi}K_{00}(\varphi)[\widehat\psi] + [\partial_\psi K_{10}(\varphi)]^T\widehat\eta + [\partial_\psi K_{01}(\varphi)]^T\widehat w \\ \mathcal{D}_\omega\widehat w - \partial_x\{\partial_\psi K_{01}(\varphi)[\widehat\psi] + K_{11}(\varphi)\widehat\eta + K_{02}(\varphi)\widehat w\}\end{pmatrix}. \tag{5.4.40}$$

We now estimate the induced composition operator.

Lemma 5.4.8. *Assume* (5.4.4) *and let* $\widehat\imath := (\widehat\psi,\widehat\eta,\widehat w)$. *Then*

$$\|DG_\delta(\varphi,0,0)[\widehat\imath]\|_s + \|DG_\delta(\varphi,0,0)^{-1}[\widehat\imath]\|_s \leq_s \|\widehat\imath\|_s + (\|\mathfrak{I}_0\|_{s+\sigma} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}\|Z\|_{s+\sigma})\|\widehat\imath\|_{s_0}, \tag{5.4.41}$$

$$\|D^2G_\delta(\varphi,0,0)[\widehat\imath_1,\widehat\imath_2]\|_s \leq_s \|\widehat\imath_1\|_s\|\widehat\imath_2\|_{s_0} + \|\widehat\imath_1\|_{s_0}\|\widehat\imath_2\|_s + (\|\mathfrak{I}_0\|_{s+\sigma} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}\|Z\|_{s+\sigma})\|\widehat\imath_1\|_{s_0}\|\widehat\imath_2\|_{s_0}$$

for some $\sigma := \sigma(\nu,\tau)$. *Moreover the same estimates hold if we replace the norm* $\| \ \|_s$ *with* $\| \ \|_s^{\mathrm{Lip}(\gamma)}$.

Proof. The estimate (5.4.41) for $DG_\delta(\varphi,0,0)$ follows by (5.4.39) and (5.4.17). By (5.4.4), $\|(DG_\delta(\varphi,0,0)-I)\widehat{\imath}\|_{s_0} \le C\varepsilon^{6-2b}\gamma^{-1}\|\widehat{\imath}\|_{s_0} \le \|\widehat{\imath}\|_{s_0}/2$. Therefore $DG_\delta(\varphi,0,0)$ is invertible and, by Neumann series, the inverse satisfies (5.4.41). The bound for D^2G_δ follows by differentiating DG_δ. □

In order to construct an approximate inverse of (5.4.40) it is sufficient to solve the equation

$$\mathbb{D}[\widehat{\psi},\widehat{\eta},\widehat{w},\widehat{\zeta}] := \begin{pmatrix} \mathcal{D}_\omega\widehat{\psi} - K_{20}(\varphi)\widehat{\eta} - K_{11}^T(\varphi)\widehat{w} \\ \mathcal{D}_\omega\widehat{\eta} + [\partial_\psi\theta_0(\varphi)]^T\widehat{\zeta} \\ \mathcal{D}_\omega\widehat{w} - \partial_x K_{11}(\varphi)\widehat{\eta} - \partial_x K_{02}(\varphi)\widehat{w} \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \tag{5.4.42}$$

which is obtained by neglecting in (5.4.40) the terms $\partial_\psi K_{10}$, $\partial_{\psi\psi}K_{00}$, $\partial_\psi K_{00}$, $\partial_\psi K_{01}$ and $\partial_\psi[\partial_\psi\theta_0(\varphi)]^T[\cdot,\zeta_0]$ (which are naught at a solution by Lemmata 5.4.6 and 5.4.1).

First we solve the second equation in (5.4.42), namely $\mathcal{D}_\omega\widehat{\eta} = g_2 - [\partial_\psi\theta_0(\varphi)]^T\widehat{\zeta}$. We choose $\widehat{\zeta}$ so that the φ-average of the right hand side is zero, namely

$$\widehat{\zeta} = \langle g_2\rangle \tag{5.4.43}$$

(we denote $\langle g\rangle := \fint_{\mathbb{T}^\nu} g(\varphi)d\varphi$). Note that the φ-averaged matrix $\langle[\partial_\psi\theta_0]^T\rangle = \langle I + [\partial_\psi\Theta_0]^T\rangle = I$ because $\theta_0(\varphi) = \varphi + \Theta_0(\varphi)$ and $\Theta_0(\varphi)$ is a periodic function. Therefore

$$\widehat{\eta} := \mathcal{D}_\omega^{-1}\big(g_2 - [\partial_\psi\theta_0(\varphi)]^T\langle g_2\rangle\big) + \langle\widehat{\eta}\rangle\,, \quad \langle\widehat{\eta}\rangle \in \mathbb{R}^\nu\,, \tag{5.4.44}$$

where the average $\langle\widehat{\eta}\rangle$ will be fixed below. Then we consider the third equation

$$\mathcal{L}_\omega\widehat{w} = g_3 + \partial_x K_{11}(\varphi)\widehat{\eta}\,, \quad \mathcal{L}_\omega := \omega\cdot\partial_\varphi - \partial_x K_{02}(\varphi)\,. \tag{5.4.45}$$

- Inversion assumption. *There exists a set $\Omega_\infty \subset \Omega_o$ such that for all $\omega \in \Omega_\infty$, for every function $g \in H^{s+\mu}_{S^\perp}(\mathbb{T}^{\nu+1})$ there exists a solution $h := \mathcal{L}_\omega^{-1}g \in H^s_{S^\perp}(\mathbb{T}^{\nu+1})$ of the linear equation $\mathcal{L}_\omega h = g$ which satisfies*

$$\|\mathcal{L}_\omega^{-1}g\|_s^{\mathrm{Lip}(\gamma)} \le C(s)\gamma^{-1}\big(\|g\|_{s+\mu}^{\mathrm{Lip}(\gamma)} + \varepsilon\gamma^{-1}\{\|\mathfrak{I}_0\|_{s+\mu}^{\mathrm{Lip}(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)}\|Z\|_{s+\mu}^{\mathrm{Lip}(\gamma)}\}\|g\|_{s_0}^{\mathrm{Lip}(\gamma)}\big) \tag{5.4.46}$$

for some $\mu := \mu(\tau,\nu) > 0$.

Remark 5.4.2. *The term $\varepsilon\gamma^{-1}\{\|\mathfrak{I}_0\|_{s+\mu}^{\mathrm{Lip}(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)}\|Z\|_{s+\mu}^{\mathrm{Lip}(\gamma)}\}$ arises because the remainder R_6 in Section 5.6.6 contains the term $\varepsilon(\|\Theta_0\|_{s+\mu}^{\mathrm{Lip}(\gamma)} + \|y_\delta\|_{s+\mu}^{\mathrm{Lip}(\gamma)}) \le \varepsilon\|\mathfrak{I}_\delta\|_{s+\mu}^{\mathrm{Lip}(\gamma)}$, see Lemma 5.6.14.*

By the above assumption there exists a solution

$$\widehat{w} := \mathcal{L}_\omega^{-1}[g_3 + \partial_x K_{11}(\varphi)\widehat{\eta}] \tag{5.4.47}$$

of (5.4.45). Finally, we solve the first equation in (5.4.42), which, substituting (5.4.44), (5.4.47), becomes

$$\mathcal{D}_\omega\widehat{\psi} = g_1 + M_1(\varphi)\langle\widehat{\eta}\rangle + M_2(\varphi)g_2 + M_3(\varphi)g_3 - M_2(\varphi)[\partial_\psi\theta_0]^T\langle g_2\rangle\,, \tag{5.4.48}$$

where

$$M_1(\varphi) := K_{20}(\varphi) + K_{11}^T(\varphi)\mathcal{L}_\omega^{-1}\partial_x K_{11}(\varphi)\,, \quad M_2(\varphi) := M_1(\varphi)\mathcal{D}_\omega^{-1}\,, \quad M_3(\varphi) := K_{11}^T(\varphi)\mathcal{L}_\omega^{-1}\,. \tag{5.4.49}$$

In order to solve the equation (5.4.48) we have to choose $\langle \widehat{\eta} \rangle$ such that the right hand side in (5.4.48) has zero average. By Lemma 5.4.7 and (5.4.4), the φ-averaged matrix $\langle M_1 \rangle = -3\varepsilon^{2b} I + O(\varepsilon^{10}\gamma^{-3})$. Therefore, for ε small, $\langle M_1 \rangle$ is invertible and $\langle M_1 \rangle^{-1} = O(\varepsilon^{-2b}) = O(\gamma^{-1})$ (recall (5.3.10)). Thus we define

$$\langle \widehat{\eta} \rangle := -\langle M_1 \rangle^{-1} [\langle g_1 \rangle + \langle M_2 g_2 \rangle + \langle M_3 g_3 \rangle - \langle M_2 [\partial_\psi \theta_0]^T \rangle \langle g_2 \rangle]. \tag{5.4.50}$$

With this choice of $\langle \widehat{\eta} \rangle$ the equation (5.4.48) has the solution

$$\widehat{\psi} := \mathcal{D}_\omega^{-1} [g_1 + M_1(\varphi) \langle \widehat{\eta} \rangle + M_2(\varphi) g_2 + M_3(\varphi) g_3 - M_2(\varphi) [\partial_\psi \theta_0]^T \langle g_2 \rangle]. \tag{5.4.51}$$

In conclusion, we have constructed a solution $(\widehat{\psi}, \widehat{\eta}, \widehat{w}, \widehat{\zeta})$ of the linear system (5.4.42).

Proposition 5.4.1. *Assume* (5.4.4) *and* (5.4.46). *Then,* $\forall \omega \in \Omega_\infty$, $\forall g := (g_1, g_2, g_3)$, *the system* (5.4.42) *has a solution* $\mathbb{D}^{-1} g := (\widehat{\psi}, \widehat{\eta}, \widehat{w}, \widehat{\zeta})$ *where* $(\widehat{\psi}, \widehat{\eta}, \widehat{w}, \widehat{\zeta})$ *are defined in* (5.4.51), (5.4.44), (5.4.50), (5.4.47), (5.4.43) *satisfying*

$$\begin{aligned} \|\mathbb{D}^{-1} g\|_s^{\mathrm{Lip}(\gamma)} \leq_s \gamma^{-1} \|g\|_{s+\mu}^{\mathrm{Lip}(\gamma)} \\ + \varepsilon \gamma^{-2} \big\{ \|\mathfrak{I}_0\|_{s+\mu}^{\mathrm{Lip}(\gamma)} + \gamma^{-1} \|\mathfrak{I}_0\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \|\mathcal{F}(i_0, \zeta_0)\|_{s+\mu}^{\mathrm{Lip}(\gamma)} \big\} \|g\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)}. \end{aligned} \tag{5.4.52}$$

Proof. Recalling (5.4.49), by Lemma 5.4.7, (5.4.46), (5.4.4) we get $\|M_2 h\|_{s_0} + \|M_3 h\|_{s_0} \leq C \|h\|_{s_0+\sigma}$. Then, by (5.4.50) and $\langle M_1 \rangle^{-1} = O(\varepsilon^{-2b}) = O(\gamma^{-1})$, we deduce $|\langle \widehat{\eta} \rangle|^{\mathrm{Lip}(\gamma)} \leq C \gamma^{-1} \|g\|_{s_0+\sigma}^{\mathrm{Lip}(\gamma)}$ and (5.4.44), (5.3.21) imply $\|\widehat{\eta}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \gamma^{-1} \big(\|g\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} + \|\mathfrak{I}_0\|_{s+\sigma} \|g\|_{s_0}^{\mathrm{Lip}(\gamma)} \big)$. The bound (5.4.52) is sharp for $\widehat{w}$ because $\mathcal{L}_\omega^{-1} g_3$ in (5.4.47) is estimated using (5.4.46). Finally $\widehat{\psi}$ satisfies (5.4.52) using (5.4.51), (5.4.49), (5.4.46), (5.3.21) and Lemma 5.4.7. □

Finally we prove that the operator

$$\mathbf{T}_0 := (D\widetilde{G}_\delta)(\varphi, 0, 0) \circ \mathbb{D}^{-1} \circ (DG_\delta)(\varphi, 0, 0)^{-1} \tag{5.4.53}$$

is an approximate right inverse for $d_{i,\zeta} \mathcal{F}(i_0)$ where $\widetilde{G}_\delta(\psi, \eta, w, \zeta) := \big(G_\delta(\psi, \eta, w), \zeta \big)$ is the identity on the ζ-component. We denote the norm $\|(\psi, \eta, w, \zeta)\|_s^{\mathrm{Lip}(\gamma)} := \max\{ \|(\psi, \eta, w)\|_s^{\mathrm{Lip}(\gamma)}, |\zeta|^{\mathrm{Lip}(\gamma)} \}$.

Theorem 5.4.1. (Approximate inverse) *Assume* (5.4.4) *and the inversion assumption* (5.4.46). *Then there exists* $\mu := \mu(\tau, \nu) > 0$ *such that, for all* $\omega \in \Omega_\infty$, *for all* $g := (g_1, g_2, g_3)$, *the operator* $\mathbf{T}_0$ *defined in* (5.4.53) *satisfies*

$$\begin{aligned} \|\mathbf{T}_0 g\|_s^{\mathrm{Lip}(\gamma)} \leq_s \gamma^{-1} \|g\|_{s+\mu}^{\mathrm{Lip}(\gamma)} \\ + \varepsilon \gamma^{-2} \big\{ \|\mathfrak{I}_0\|_{s+\mu}^{\mathrm{Lip}(\gamma)} + \gamma^{-1} \|\mathfrak{I}_0\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \|\mathcal{F}(i_0, \zeta_0)\|_{s+\mu}^{\mathrm{Lip}(\gamma)} \big\} \|g\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)}. \end{aligned} \tag{5.4.54}$$

It is an approximate inverse of $d_{i,\zeta} \mathcal{F}(i_0)$, *namely*

$$\begin{aligned} & \|(d_{i,\zeta} \mathcal{F}(i_0) \circ \mathbf{T}_0 - I) g\|_s^{\mathrm{Lip}(\gamma)} \\ & \leq_s \gamma^{-1} \Big(\|\mathcal{F}(i_0, \zeta_0)\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \|g\|_{s+\mu}^{\mathrm{Lip}(\gamma)} + \big\{ \|\mathcal{F}(i_0, \zeta_0)\|_{s+\mu}^{\mathrm{Lip}(\gamma)} + \varepsilon \gamma^{-1} \|\mathcal{F}(i_0, \zeta_0)\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \|\mathfrak{I}_0\|_{s+\mu}^{\mathrm{Lip}(\gamma)} \big\} \|g\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \Big). \end{aligned} \tag{5.4.55}$$

Proof. We denote $\| \ \|_s$ instead of $\| \ \|_s^{\mathrm{Lip}(\gamma)}$. The bound (5.4.54) follows from (5.4.53), (5.4.52), (5.4.41). By (5.3.6), since $X_{\mathcal{N}}$ does not depend on y, and i_δ differs from i_0 only for the y component, we have

$$\begin{aligned} d_{i,\zeta}\mathcal{F}(i_0)[\widehat{\imath},\widehat{\zeta}\,] - d_{i,\zeta}\mathcal{F}(i_\delta)[\widehat{\imath},\widehat{\zeta}\,] &= d_i X_P(i_\delta)[\widehat{\imath}\,] - d_i X_P(i_0)[\widehat{\imath}\,] \\ &= \int_0^1 \partial_y d_i X_P(\theta_0, y_0 + s(y_\delta - y_0), z_0)[y_\delta - y_0, \widehat{\imath}\,]ds =: \mathcal{E}_0[\widehat{\imath},\widehat{\zeta}\,]\,. \end{aligned} \tag{5.4.56}$$

By (5.3.18), (5.4.17), (5.4.4), we estimate

$$\|\mathcal{E}_0[\widehat{\imath},\widehat{\zeta}\,]\|_s \leq_s \|Z\|_{s_0+\sigma}\|\widehat{\imath}\|_{s+\sigma} + \|Z\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma} + \varepsilon^{2b-1}\gamma^{-1}\|Z\|_{s_0+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}\|\mathfrak{I}_0\|_{s+\sigma} \tag{5.4.57}$$

where $Z := \mathcal{F}(i_0, \zeta_0)$ (recall (5.4.5)). Note that $\mathcal{E}_0[\widehat{\imath},\widehat{\zeta}]$ is, in fact, independent of $\widehat{\zeta}$. Denote the set of variables $(\psi, \eta, w) =: \mathtt{u}$. Under the transformation G_δ, the nonlinear operator $\mathcal{F}$ in (5.3.6) transforms into

$$\mathcal{F}(G_\delta(\mathtt{u}(\varphi)), \zeta) = DG_\delta(\mathtt{u}(\varphi))\big(\mathcal{D}_\omega \mathtt{u}(\varphi) - X_K(\mathtt{u}(\varphi), \zeta)\big)\,, \quad K = H_{\varepsilon,\zeta} \circ G_\delta\,, \tag{5.4.58}$$

see (5.4.34). Differentiating (5.4.58) at the trivial torus $\mathtt{u}_\delta(\varphi) = G_\delta^{-1}(i_\delta)(\varphi) = (\varphi, 0, 0)$, at $\zeta = \zeta_0$, in the directions $(\widehat{\mathtt{u}}, \widehat{\zeta}) = (DG_\delta(\mathtt{u}_\delta)^{-1}[\widehat{\imath}], \widehat{\zeta}) = D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}[\widehat{\imath}, \widehat{\zeta}\,]$, we get

$$d_{i,\zeta}\mathcal{F}(i_\delta)[\widehat{\imath},\widehat{\zeta}\,] = DG_\delta(\mathtt{u}_\delta)\big(\mathcal{D}_\omega\widehat{\mathtt{u}} - d_{\mathtt{u},\zeta}X_K(\mathtt{u}_\delta, \zeta_0)[\widehat{\mathtt{u}},\widehat{\zeta}\,]\big) + \mathcal{E}_1[\widehat{\imath},\widehat{\zeta}\,]\,, \tag{5.4.59}$$

$$\mathcal{E}_1[\widehat{\imath},\widehat{\zeta}\,] := D^2G_\delta(\mathtt{u}_\delta)\big[DG_\delta(\mathtt{u}_\delta)^{-1}\mathcal{F}(i_\delta, \zeta_0),\, DG_\delta(\mathtt{u}_\delta)^{-1}[\widehat{\imath}\,]\big]\,, \tag{5.4.60}$$

where $d_{\mathtt{u},\zeta}X_K(\mathtt{u}_\delta, \zeta_0)$ is expanded in (5.4.40). In fact, $\mathcal{E}_1$ is independent of $\widehat{\zeta}$. We split

$$\mathcal{D}_\omega\widehat{\mathtt{u}} - d_{\mathtt{u},\zeta}X_K(\mathtt{u}_\delta, \zeta_0)[\widehat{\mathtt{u}},\widehat{\zeta}] = \mathbb{D}[\widehat{\mathtt{u}},\widehat{\zeta}\,] + R_Z[\widehat{\mathtt{u}},\widehat{\zeta}\,],$$

where $\mathbb{D}[\widehat{\mathtt{u}},\widehat{\zeta}]$ is defined in (5.4.42) and

$$R_Z[\widehat{\psi},\widehat{\eta},\widehat{w},\widehat{\zeta}] := \begin{pmatrix} -\partial_\psi K_{10}(\varphi)[\widehat{\psi}] \\ \partial_\psi[\partial_\psi\theta_0(\varphi)]^T[\widehat{\psi},\zeta_0] + \partial_{\psi\psi}K_{00}(\varphi)[\widehat{\psi}] + [\partial_\psi K_{10}(\varphi)]^T\widehat{\eta} + [\partial_\psi K_{01}(\varphi)]^T\widehat{w} \\ -\partial_x\{\partial_\psi K_{01}(\varphi)[\widehat{\psi}]\} \end{pmatrix} \tag{5.4.61}$$

(R_Z is independent of $\widehat{\zeta}$). By (5.4.56) and (5.4.59),

$$d_{i,\zeta}\mathcal{F}(i_0) = DG_\delta(\mathtt{u}_\delta) \circ \mathbb{D} \circ D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1} + \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2\,, \quad \mathcal{E}_2 := DG_\delta(\mathtt{u}_\delta) \circ R_Z \circ D\widetilde{G}_\delta(\mathtt{u}_\delta)^{-1}\,. \tag{5.4.62}$$

By Lemmata 5.4.6, 5.4.8, 5.4.1, and (5.4.18), (5.4.4), the terms $\mathcal{E}_1, \mathcal{E}_2$ (see (5.4.60), (5.4.62), (5.4.61)) satisfy the same bound (5.4.57) as $\mathcal{E}_0$ (in fact even better). Thus the sum $\mathcal{E} := \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2$ satisfies (5.4.57). Applying $\mathbf{T}_0$ defined in (5.4.53) to the right in (5.4.62), since $\mathbb{D} \circ \mathbb{D}^{-1} = I$ (see Proposition 5.4.1), we get $d_{i,\zeta}\mathcal{F}(i_0) \circ \mathbf{T}_0 - I = \mathcal{E} \circ \mathbf{T}_0$. Then (5.4.55) follows from (5.4.54) and the bound (5.4.57) for $\mathcal{E}$. $\square$

5.5 The linearized operator in the normal directions

The goal of this Section is to write an explicit expression of the linearized operator $\mathcal{L}_\omega$ defined in (5.4.45), see Proposition 5.5.1. To this aim, we compute $\frac{1}{2}(K_{02}(\psi)w, w)_{L^2(\mathbb{T})}$, $w \in H_S^\perp$, which collects all the components of $(H_\varepsilon \circ G_\delta)(\psi, 0, w)$ that are quadratic in w, see (5.4.33).

We first prove some preliminary lemmata.

Lemma 5.5.1. *Let H be a Hamiltonian of class $C^2(H_0^1(\mathbb{T}), \mathbb{R})$ and consider a map $\Phi(u) := u + \Psi(u)$ satisfying $\Psi(u) = \Pi_E \Psi(\Pi_E u)$, for all u, where E is a finite dimensional subspace as in (5.1.3). Then*

$$\partial_u \big[\nabla(H \circ \Phi)\big](u)[h] = (\partial_u \nabla H)(\Phi(u))[h] + \mathcal{R}(u)[h]\,, \tag{5.5.1}$$

where $\mathcal{R}(u)$ has the "finite dimensional" form

$$\mathcal{R}(u)[h] = \sum\nolimits_{|j| \le C} \big(h, g_j(u)\big)_{L^2(\mathbb{T})} \chi_j(u) \tag{5.5.2}$$

with $\chi_j(u) = e^{\mathrm{i}jx}$ or $g_j(u) = e^{\mathrm{i}jx}$. The remainder $\mathcal{R}(u) = \mathcal{R}_0(u) + \mathcal{R}_1(u) + \mathcal{R}_2(u)$ with

$$\begin{aligned} &\mathcal{R}_0(u) := (\partial_u \nabla H)(\Phi(u)) \partial_u \Psi(u), \qquad \mathcal{R}_1(u) := [\partial_u \{\Psi'(u)^T\}][\cdot, \nabla H(\Phi(u))], \\ &\mathcal{R}_2(u) := [\partial_u \Psi(u)]^T (\partial_u \nabla H)(\Phi(u)) \partial_u \Phi(u). \end{aligned} \tag{5.5.3}$$

Proof. By a direct calculation,

$$\nabla(H \circ \Phi)(u) = [\Phi'(u)]^T \nabla H(\Phi(u)) = \nabla H(\Phi(u)) + [\Psi'(u)]^T \nabla H(\Phi(u)) \tag{5.5.4}$$

where $\Phi'(u) := (\partial_u \Phi)(u)$ and $[\ ]^T$ denotes the transpose with respect to the L^2 scalar product. Differentiating (5.5.4), we get (5.5.1) and (5.5.3).

Let us show that each $\mathcal{R}_m$ has the form (5.5.2). We have

$$\Psi'(u) = \Pi_E \Psi'(\Pi_E u) \Pi_E\,, \quad [\Psi'(u)]^T = \Pi_E [\Psi'(\Pi_E u)]^T \Pi_E\,. \tag{5.5.5}$$

Hence, setting $A := (\partial_u \nabla H)(\Phi(u)) \Pi_E \Psi'(\Pi_E u)$, we get

$$\mathcal{R}_0(u)[h] = A[\Pi_E h] = \sum\nolimits_{|j| \le C} h_j A(e^{\mathrm{i}jx}) = \sum\nolimits_{|j| \le C} (h, g_j)_{L^2(\mathbb{T})} \chi_j$$

with $g_j := e^{\mathrm{i}jx}$, $\chi_j := A(e^{\mathrm{i}jx})$. Similarly, using (5.5.5), and setting $A := [\Psi'(\Pi_E u)]^T \Pi_E (\partial_u \nabla H)(\Phi(u)) \Phi'(u)$, we get

$$\mathcal{R}_2(u)[h] = \Pi_E[Ah] = \sum\nolimits_{|j| \le C} (Ah, e^{\mathrm{i}jx})_{L^2(\mathbb{T})} e^{\mathrm{i}jx} = \sum\nolimits_{|j| \le C} (h, A^T e^{\mathrm{i}jx})_{L^2(\mathbb{T})} e^{\mathrm{i}jx}\,,$$

which has the form (5.5.2) with $g_j := A^T(e^{\mathrm{i}jx})$ and $\chi_j := e^{\mathrm{i}jx}$. Differentiating the second equality in (5.5.5), we see that

$$\mathcal{R}_1(u)[h] = \Pi_E[Ah]\,, \quad Ah := \partial_u \{\Psi'(\Pi_E u)^T\}[\Pi_E h, \Pi_E (\nabla H)(\Phi(u))]\,,$$

which has the same form of $\mathcal{R}_2$ and so (5.5.2). □

Lemma 5.5.2. *Let* $H(u) := \int_{\mathbb{T}} f(u)X(u)dx$ *where* $X(u) = \Pi_E X(\Pi_E u)$ *and* $f(u)(x) := f(u(x))$ *is the composition operator for a function of class* C^2. *Then*

$$(\partial_u \nabla H)(u)[h] = f''(u)X(u)\,h + \mathcal{R}(u)[h] \tag{5.5.6}$$

where $\mathcal{R}(u)$ *has the form* (5.5.2) *with* $\chi_j(u) = e^{\mathrm{i}jx}$ *or* $g_j(u) = e^{\mathrm{i}jx}$.

Proof. A direct calculation proves that $\nabla H(u) = f'(u)X(u)+X'(u)^T[f(u)]$, and (5.5.6) follows with $\mathcal{R}(u)[h] = f'(u)X'(u)[h]+ \partial_u\{X'(u)^T\}[h, f(u)]+ X'(u)^T[f'(u)h]$, which has the form (5.5.2). □

We conclude this Section with a technical lemma used from the end of Section 5.6.3 about the decay norms of "finite dimensional operators". Note that operators of the form (5.5.7) (that will appear in Section 5.6.1) reduce to those in (5.5.2) when the functions $g_j(\tau)$, $\chi_j(\tau)$ are independent of τ

Lemma 5.5.3. *Let* $\mathcal{R}$ *be an operator of the form*

$$\mathcal{R}h = \sum_{|j|\le C} \int_0^1 \big(h\,,\, g_j(\tau)\big)_{L^2(\mathbb{T})}\chi_j(\tau)\,d\tau\,, \tag{5.5.7}$$

where the functions $g_j(\tau)$, $\chi_j(\tau) \in H^s$, $\tau \in [0,1]$ *depend in a Lipschitz way on the parameter* ω. *Then its matrix* s*-decay norm (see* (3.1.1)*-*(3.1.2)*) satisfies*

$$|\mathcal{R}|_s^{\mathrm{Lip}(\gamma)} \le_s \sum\nolimits_{|j|\le C} \sup\nolimits_{\tau\in[0,1]}\big\{\|\chi_j(\tau)\|_s^{\mathrm{Lip}(\gamma)}\|g_j(\tau)\|_{s_0}^{\mathrm{Lip}(\gamma)} + \|\chi_j(\tau)\|_{s_0}^{\mathrm{Lip}(\gamma)}\|g_j(\tau)\|_s^{\mathrm{Lip}(\gamma)}\big\}\,.$$

Proof. For each $\tau \in [0,1]$, the operator $h \mapsto (h, g_j(\tau))\chi_j(\tau)$ is the composition $\chi_j(\tau)\circ\Pi_0\circ g_j(\tau)$ of the multiplication operators for $g_j(\tau), \chi_j(\tau)$ and $h \mapsto \Pi_0 h := \int_{\mathbb{T}} h dx$. Hence the lemma follows by the interpolation estimate (3.1.9) and (3.1.5). □

5.5.1 Composition with the map G_δ

In the sequel we shall use that $\mathfrak{I}_\delta := \mathfrak{I}_\delta(\varphi;\omega) := i_\delta(\varphi;\omega) - (\varphi,0,0)$ satisfies, by Lemma 5.4.4 and (5.4.4),

$$\|\mathfrak{I}_\delta\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \le C\varepsilon^{6-2b}\gamma^{-1}\,. \tag{5.5.8}$$

We now study the Hamiltonian $K := H_\varepsilon\circ G_\delta = \varepsilon^{-2b}\mathcal{H}\circ A_\varepsilon\circ G_\delta$ defined in (5.4.33), (5.2.6).

Recalling (5.2.7) and (5.4.26) the map $A_\varepsilon\circ G_\delta$ has the form

$$A_\varepsilon\circ G_\delta(\psi,\eta,w) = \varepsilon\sum_{j\in S}\sqrt{\xi_j + \varepsilon^{2(b-1)}|j|[y_\delta(\psi)+L_1(\psi)\eta+L_2(\psi)w]_j}\,e^{\mathrm{i}[\theta_0(\psi)]_j}e^{\mathrm{i}jx} + \varepsilon^b(z_0(\psi)+w) \tag{5.5.9}$$

where

$$L_1(\psi) := [\partial_\psi\theta_0(\psi)]^{-T}\,,\quad L_2(\psi) := \big[(\partial_\theta\tilde z_0)(\theta_0(\psi))\big]^T\partial_x^{-1}\,. \tag{5.5.10}$$

By Taylor's formula, we develop (5.5.9) in w at $\eta = 0$, $w = 0$, and we get $A_\varepsilon\circ G_\delta(\psi,0,w) = T_\delta(\psi)+T_1(\psi)w+T_2(\psi)[w,w]+ T_{\ge3}(\psi,w)$, where

$$T_\delta(\psi) := (A_\varepsilon\circ G_\delta)(\psi,0,0) = \varepsilon v_\delta(\psi)+\varepsilon^b z_0(\psi)\,,\quad v_\delta(\psi) := \sum_{j\in S}\sqrt{\xi_j+\varepsilon^{2(b-1)}|j|[y_\delta(\psi)]_j}\,e^{\mathrm{i}[\theta_0(\psi)]_j}e^{\mathrm{i}jx} \tag{5.5.11}$$

is the approximate isotropic torus in phase space (it corresponds to i_δ in Lemma 5.4.4),

$$T_1(\psi)w = \varepsilon \sum_{j\in S} \frac{\varepsilon^{2(b-1)}|j|[L_2(\psi)w]_j \, e^{\mathrm{i}[\theta_0(\psi)]_j}}{2\sqrt{\xi_j + \varepsilon^{2(b-1)}|j|[y_\delta(\psi)]_j}} \, e^{\mathrm{i}jx} + \varepsilon^b w =: \varepsilon^{2b-1} U_1(\psi)w + \varepsilon^b w \tag{5.5.12}$$

$$T_2(\psi)[w,w] = -\varepsilon \sum_{j\in S} \frac{\varepsilon^{4(b-1)} j^2 [L_2(\psi)w]_j^2 \, e^{\mathrm{i}[\theta_0(\psi)]_j}}{8\{\xi_j + \varepsilon^{2(b-1)}|j|[y_\delta(\psi)]_j\}^{3/2}} \, e^{\mathrm{i}jx} =: \varepsilon^{4b-3} U_2(\psi)[w,w] \tag{5.5.13}$$

and $T_{\geq 3}(\psi, w)$ collects all the terms of order at least cubic in w. In the notation of (5.2.7), the function $v_\delta(\psi)$ in (5.5.11) is $v_\delta(\psi) = v_\varepsilon(\theta_0(\psi), y_\delta(\psi))$. The terms $U_1, U_2 = O(1)$ in ε. Moreover, using that $L_2(\psi)$ in (5.5.10) vanishes as $z_0 = 0$, they satisfy

$$\|U_1 w\|_s \leq \|\mathfrak{I}_\delta\|_s \|w\|_{s_0} + \|\mathfrak{I}_\delta\|_{s_0} \|w\|_s \,, \quad \|U_2[w,w]\|_s \leq \|\mathfrak{I}_\delta\|_s \|\mathfrak{I}_\delta\|_{s_0} \|w\|_{s_0}^2 + \|\mathfrak{I}_\delta\|_{s_0}^2 \|w\|_{s_0} \|w\|_s \tag{5.5.14}$$

and also in the $\| \ \|_s^{\mathrm{Lip}(\gamma)}$-norm.

By Taylor's formula $\mathcal{H}(u+h) = \mathcal{H}(u) + ((\nabla \mathcal{H})(u), h)_{L^2(\mathbb{T})} + \frac{1}{2}((\partial_u \nabla \mathcal{H})(u)[h], h)_{L^2(\mathbb{T})} + O(h^3)$. Specifying at $u = T_\delta(\psi)$ and $h = T_1(\psi)w + T_2(\psi)[w,w] + T_{\geq 3}(\psi, w)$, we obtain that the sum of all the components of $K = \varepsilon^{-2b}(\mathcal{H} \circ A_\varepsilon \circ G_\delta)(\psi, 0, w)$ that are quadratic in w is

$$\frac{1}{2}(K_{02} w, w)_{L^2(\mathbb{T})} = \varepsilon^{-2b}((\nabla \mathcal{H})(T_\delta), T_2[w,w])_{L^2(\mathbb{T})} + \varepsilon^{-2b}\frac{1}{2}((\partial_u \nabla \mathcal{H})(T_\delta)[T_1 w], T_1 w)_{L^2(\mathbb{T})} \,.$$

Inserting the expressions (5.5.12), (5.5.13) we get

$$\begin{aligned} K_{02}(\psi)w = {} & (\partial_u \nabla \mathcal{H})(T_\delta)[w] + 2\varepsilon^{b-1}(\partial_u \nabla \mathcal{H})(T_\delta)[U_1 w] + \varepsilon^{2(b-1)} U_1^T (\partial_u \nabla \mathcal{H})(T_\delta)[U_1 w] \\ & + 2\varepsilon^{2b-3} U_2[w, \cdot]^T (\nabla \mathcal{H})(T_\delta). \end{aligned} \tag{5.5.15}$$

Lemma 5.5.4.

$$(K_{02}(\psi)w, w)_{L^2(\mathbb{T})} = ((\partial_u \nabla \mathcal{H})(T_\delta)[w], w)_{L^2(\mathbb{T})} + (R(\psi)w, w)_{L^2(\mathbb{T})} \tag{5.5.16}$$

where $R(\psi)w$ has the "finite dimensional" form

$$R(\psi)w = \sum\nolimits_{|j|\leq C} \big(w, g_j(\psi)\big)_{L^2(\mathbb{T})} \chi_j(\psi) \tag{5.5.17}$$

where, for some $\sigma := \sigma(\nu, \tau) > 0$,

$$\|g_j\|_s^{\mathrm{Lip}(\gamma)} \|\chi_j\|_{s_0}^{\mathrm{Lip}(\gamma)} + \|g_j\|_{s_0}^{\mathrm{Lip}(\gamma)} \|\chi_j\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^{b+1} \|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \tag{5.5.18}$$

$$\begin{aligned} \|\partial_i g_j[\widehat{\imath}]\|_s \|\chi_j\|_{s_0} + \|\partial_i g_j[\widehat{\imath}]\|_{s_0} \|\chi_j\|_s + \|g_j\|_{s_0} \|\partial_i \chi_j[\widehat{\imath}]\|_s + \|g_j\|_s \|\partial_i \chi_j[\widehat{\imath}]\|_{s_0} \leq_s {} & \varepsilon^{b+1} \|\widehat{\imath}\|_{s+\sigma} \\ & + \varepsilon^{2b-1} \|\mathfrak{I}_\delta\|_{s+\sigma} \|\widehat{\imath}\|_{s_0+\sigma} \,, \end{aligned} \tag{5.5.19}$$

and, as usual, $i = (\theta, y, z)$ (see (5.3.1)), $\widehat{\imath} = (\widehat{\theta}, \widehat{y}, \widehat{z})$.

Proof. Since $U_1 = \Pi_S U_1$ and $U_2 = \Pi_S U_2$, the last three terms in (5.5.15) have all the form (5.5.17) (argue as in Lemma 5.5.1). We now prove that they are also small in size.

The contributions in (5.5.15) from H_2 are better analyzed by the expression

$$\varepsilon^{-2b} H_2 \circ A_\varepsilon \circ G_\delta(\psi,\eta,w) = const + \sum_{j\in S^+} j^3 \big[y_\delta(\psi) + L_1(\psi)\eta + L_2(\psi)w\big]_j + \frac{1}{2}\int_{\mathbb{T}} (z_0(\psi)+w)_x^2\,dx$$

which follows by (5.2.8), (5.4.26), (5.5.10). Hence the only contribution to $(K_{02}w,w)$ is $\int_{\mathbb{T}} w_x^2\,dx$. Now we consider the cubic term $\mathcal{H}_3$ in (5.1.6). A direct calculation shows that for $u = v+z$, $\nabla\mathcal{H}_3(u) = 3z^2 + 6\Pi_S^\perp(vz)$, and $\partial_u\nabla\mathcal{H}_3(u)[U_1 w] = 6\Pi_S^\perp(zU_1w)$ (since $U_1 w \in H_S$). Therefore

$$\nabla\mathcal{H}_3(T_\delta) = 3\varepsilon^{2b} z_0^2 + 6\varepsilon^{b+1}\Pi_S^\perp(v_\delta z_0)\,, \quad \partial_u\nabla\mathcal{H}_3(T_\delta)[U_1 w] = 6\varepsilon^b \Pi_S^\perp(z_0\,U_1 w)\,. \tag{5.5.20}$$

By (5.5.20) one has $((\partial_u\nabla\mathcal{H}_3)(T_\delta)[U_1w], U_1w)_{L^2(\mathbb{T})} = 0$, and since also $U_2 = \Pi_S U_2$,

$$\varepsilon^{b-1}\partial_u\nabla\mathcal{H}_3(T_\delta)[U_1w] + \varepsilon^{2b-3}U_2[w,\cdot]^T\nabla\mathcal{H}_3(T_\delta) = 6\varepsilon^{2b-1}\Pi_S^\perp(z_0U_1w) + 3\varepsilon^{4b-3}U_2[w,\cdot]^T z_0^2\,. \tag{5.5.21}$$

These terms have the form (5.5.17) and, using (5.5.14), (5.4.4), they satisfy (5.5.18).

Finally we consider all the terms which arise from $\mathcal{H}_{\geq 4} = O(u^4)$. The operators $\varepsilon^{b-1}\partial_u\nabla\mathcal{H}_{\geq 4}(T_\delta)U_1$, $\varepsilon^{2(b-1)}U_1^T(\partial_u\nabla\mathcal{H}_{\geq 4})(T_\delta)U_1$, $\varepsilon^{2b-3}U_2^T\nabla\mathcal{H}_{\geq 4}(T_\delta)$ have the form (5.5.17) and, using $\|T_\delta\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon(1+\|\mathfrak{I}_\delta\|_s^{\mathrm{Lip}(\gamma)})$, (5.5.14), (5.4.4), the bound (5.5.18) holds. Notice that the biggest term is $\varepsilon^{b-1}\partial_u\nabla\mathcal{H}_{\geq 4}(T_\delta)U_1$.

By (5.4.19) and using explicit formulae (5.5.10)-(5.5.13) we get estimate (5.5.19). $\square$

The conclusion of this Section is that, after the composition with the action-angle variables, the rescaling (5.2.5), and the transformation G_δ, the linearized operator to analyze is $H_S^\perp \ni w \mapsto (\partial_u\nabla\mathcal{H})(T_\delta)[w]$, up to finite dimensional operators which have the form (5.5.17) and size (5.5.18).

5.5.2 The linearized operator in the normal directions

In view of (5.5.16) we now compute $((\partial_u\nabla\mathcal{H})(T_\delta)[w],w)_{L^2(\mathbb{T})}$, $w\in H_S^\perp$, where $\mathcal{H} = H\circ\Phi_B$ and Φ_B is the Birkhoff map of Proposition 5.1.1. It is convenient to estimate separately the terms in

$$\mathcal{H} = H\circ\Phi_B = (H_2+H_3)\circ\Phi_B + H_{\geq 5}\circ\Phi_B \tag{5.5.22}$$

where $H_2, H_3, H_{\geq 5}$ are defined in (5.1.1).

We first consider $H_{\geq 5}\circ\Phi_B$. By (5.1.1) we get $\nabla H_{\geq 5}(u) = \pi_0[(\partial_u f)(x,u,u_x)] - \partial_x\{(\partial_{u_x} f)(x,u,u_x)\}$, see (4.1.50). Since the Birkhoff transformation Φ_B has the form (5.1.4), Lemma 5.5.1 (at $u = T_\delta$, see (5.5.11)) implies that

$$\begin{aligned}\partial_u\nabla(H_{\geq 5}\circ\Phi_B)(T_\delta)[h] &= (\partial_u\nabla H_{\geq 5})(\Phi_B(T_\delta))[h] + \mathcal{R}_{H_{\geq 5}}(T_\delta)[h] \\ &= \partial_x(r_1(T_\delta)\partial_x h) + r_0(T_\delta)h + \mathcal{R}_{H_{\geq 5}}(T_\delta)[h]\end{aligned} \tag{5.5.23}$$

where the multiplicative functions $r_0(T_\delta)$, $r_1(T_\delta)$ are

$$r_0(T_\delta) := \sigma_0(\Phi_B(T_\delta)), \qquad \sigma_0(u) := (\partial_{uu} f)(x,u,u_x) - \partial_x\{(\partial_{uu_x} f)(x,u,u_x)\}, \tag{5.5.24}$$

$$r_1(T_\delta) := \sigma_1(\Phi_B(T_\delta)), \qquad \sigma_1(u) := -(\partial_{u_xu_x} f)(x,u,u_x), \tag{5.5.25}$$

the remainder $\mathcal{R}_{H_{\geq 5}}(u)$ has the form (5.5.2) with $\chi_j = e^{\mathrm{i}jx}$ or $g_j = e^{\mathrm{i}jx}$ and, using (5.5.3), it satisfies, for some $\sigma := \sigma(\nu, \tau) > 0$,

$$\|g_j\|_s^{\mathrm{Lip}(\gamma)} \|\chi_j\|_{s_0}^{\mathrm{Lip}(\gamma)} + \|g_j\|_{s_0}^{\mathrm{Lip}(\gamma)} \|\chi_j\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4 (1 + \|\mathfrak{I}_\delta\|_{s+2}^{\mathrm{Lip}(\gamma)})$$
$$\|\partial_i g_j[\widehat{\imath}]\|_s \|\chi_j\|_{s_0} + \|\partial_i g_j[\widehat{\imath}]\|_{s_0} \|\chi_j\|_s + \|g_j\|_{s_0} \|\partial_i \chi_j[\widehat{\imath}]\|_s + \|g_j\|_s \|\partial_i \chi_j[\widehat{\imath}]\|_{s_0} \leq_s \varepsilon^4 (\|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+2} \|\widehat{\imath}\|_{s_0+2}).$$

Now we consider the contributions from $(H_2 + H_3) \circ \Phi_B$. By Lemma 5.5.1 and the expressions of H_2, H_3 in (5.1.1) we deduce that

$$\partial_u \nabla (H_2 \circ \Phi_B)(T_\delta)[h] = -\partial_{xx} h + \mathcal{R}_{H_2}(T_\delta)[h], \quad \partial_u \nabla (H_3 \circ \Phi_B)(T_\delta)[h] = 6\Phi_B(T_\delta) h + \mathcal{R}_{H_3}(T_\delta)[h],$$

where $\Phi_B(T_\delta)$ is a function with zero space average, because $\Phi_B : H_0^1(\mathbb{T}) \to H_0^1(\mathbb{T})$ (Proposition 5.1.1) and $\mathcal{R}_{H_2}(u)$, $\mathcal{R}_{H_3}(u)$ have the form (5.5.2). By (5.5.3), the size $(\mathcal{R}_{H_2} + \mathcal{R}_{H_3})(T_\delta) = O(\varepsilon)$. We expand

$$(\mathcal{R}_{H_2} + \mathcal{R}_{H_3})(T_\delta) = \varepsilon \mathcal{R}_1 + \varepsilon^2 \mathcal{R}_2 + \tilde{\mathcal{R}}_{>2},$$

where $\tilde{\mathcal{R}}_{>2}$ has size $o(\varepsilon^2)$, and we get, $\forall h \in H_S^\perp$,

$$\Pi_S^\perp \partial_u \nabla ((H_2 + H_3) \circ \Phi_B)(T_\delta)[h] = -\partial_{xx} h + \Pi_S^\perp (6\Phi_B(T_\delta) h) + \Pi_S^\perp (\varepsilon \mathcal{R}_1 + \varepsilon^2 \mathcal{R}_2 + \tilde{\mathcal{R}}_{>2})[h]. \quad (5.5.26)$$

We also develop the function $\Phi_B(T_\delta)$ is powers of ε. Expand $\Phi_B(u) = u + \Psi_2(u) + \Psi_{\geq 3}(u)$, where $\Psi_2(u)$ is quadratic, $\Psi_{\geq 3}(u) = O(u^3)$, and both map $H_0^1(\mathbb{T}) \to H_0^1(\mathbb{T})$. At $u = T_\delta = \varepsilon v_\delta + \varepsilon^b z_0$ we get

$$\Phi_B(T_\delta) = T_\delta + \Psi_2(T_\delta) + \Psi_{\geq 3}(T_\delta) = \varepsilon v_\delta + \varepsilon^2 \Psi_2(v_\delta) + \tilde{q} \quad (5.5.27)$$

where $\tilde{q} := \varepsilon^b z_0 + \Psi_2(T_\delta) - \varepsilon^2 \Psi_2(v_\delta) + \Psi_{\geq 3}(T_\delta)$ has zero space average and it satisfies

$$\|\tilde{q}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^3 + \varepsilon^b \|\mathfrak{I}_\delta\|_s^{\mathrm{Lip}(\gamma)}, \quad \|\partial_i \tilde{q}[\widehat{\imath}]\|_s \leq_s \varepsilon^b \big(\|\widehat{\imath}\|_s + \|\mathfrak{I}_\delta\|_s \|\widehat{\imath}\|_{s_0} \big).$$

In particular, its low norm $\|\tilde{q}\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq_{s_0} \varepsilon^{6-b} \gamma^{-1} = o(\varepsilon^2)$.

We need an exact expression of the terms of order ε and ε^2 in (5.5.26). We compare the Hamiltonian (5.1.5) with (5.5.22), noting that $(H_{\geq 5} \circ \Phi_B)(u) = O(u^5)$ because f satisfies (1.3.4) and $\Phi_B(u) = O(u)$. Therefore

$$(H_2 + H_3) \circ \Phi_B = H_2 + \mathcal{H}_3 + \mathcal{H}_4 + O(u^5),$$

and the homogeneous terms of $(H_2 + H_3) \circ \Phi_B$ of degree $2, 3, 4$ in u are H_2, $\mathcal{H}_3$, $\mathcal{H}_4$ respectively. As a consequence, the terms of order ε and ε^2 in (5.5.26) (both in the function $\Phi_B(T_\delta)$ and in the remainders $\mathcal{R}_1, \mathcal{R}_2$) come only from $H_2 + \mathcal{H}_3 + \mathcal{H}_4$. Actually they come from H_2, $\mathcal{H}_3$ and $\mathcal{H}_{4,2}$ (see (5.1.6), (5.1.7)) because, at $u = T_\delta = \varepsilon v_\delta + \varepsilon^b z_0$, for all $h \in H_S^\perp$,

$$\Pi_S^\perp (\partial_u \nabla \mathcal{H}_4)(T_\delta)[h] = \Pi_S^\perp (\partial_u \nabla \mathcal{H}_{4,2})(T_\delta)[h] + o(\varepsilon^2).$$

A direct calculation based on the expressions (5.1.6), (5.1.7) shows that, for all $h \in H_S^\perp$,

$$\begin{aligned} \Pi_S^\perp (\partial_u \nabla (H_2 + \mathcal{H}_3 + \mathcal{H}_4))(T_\delta)[h] = -\partial_{xx} h + 6\varepsilon \Pi_S^\perp (v_\delta h) + 6\varepsilon^b \Pi_S^\perp (z_0 h) + \varepsilon^2 \Pi_S^\perp \big\{ 6\pi_0 [(\partial_x^{-1} v_\delta)^2] h \\ + 6 v_\delta \Pi_S [(\partial_x^{-1} v_\delta)(\partial_x^{-1} h)] - 6\partial_x^{-1} \{(\partial_x^{-1} v_\delta) \Pi_S [v_\delta h]\} \big\} + o(\varepsilon^2). \end{aligned} \quad (5.5.28)$$

Thus, comparing the terms of order $\varepsilon, \varepsilon^2$ in (5.5.26) (using (5.5.27)) with those in (5.5.28) we deduce that the operators $\mathcal{R}_1, \mathcal{R}_2$ and the function $\Psi_2(v_\delta)$ are

$$\mathcal{R}_1 = 0, \quad \mathcal{R}_2[h] = 6v_\delta \Pi_S\big[(\partial_x^{-1} v_\delta)(\partial_x^{-1} h)\big] - 6\partial_x^{-1}\{(\partial_x^{-1} v_\delta)\Pi_S[v_\delta h]\}\,, \quad \Psi_2(v_\delta) = \pi_0[(\partial_x^{-1} v_\delta)^2]. \tag{5.5.29}$$

In conclusion, by (5.5.22), (5.5.26), (5.5.23), (5.5.27), (5.5.29), we get, for all $h \in H_{S^\perp}$,

$$\begin{aligned}\Pi_S^\perp \partial_u \nabla \mathcal{H}(T_\delta)[h] = & -\partial_{xx} h + \Pi_S^\perp\big[\big(\varepsilon 6 v_\delta + \varepsilon^2 6\pi_0[(\partial_x^{-1} v_\delta)^2] + q_{>2} + p_{\geq 4}\big)h\big] \\ & + \Pi_S^\perp \partial_x(r_1(T_\delta)\partial_x h) + \varepsilon^2 \Pi_S^\perp \mathcal{R}_2[h] + \Pi_S^\perp \mathcal{R}_{>2}[h]\end{aligned} \tag{5.5.30}$$

where r_1 is defined in (5.5.24), $\mathcal{R}_2$ in (5.5.29), the remainder $\mathcal{R}_{>2} := \tilde{\mathcal{R}}_{>2} + \mathcal{R}_{H_{\geq 5}}(T_\delta)$ and the functions (using also (5.5.24), (5.5.25), (1.3.4)),

$$q_{>2} := 6\tilde{q} + \varepsilon^3\big((\partial_{uu} f_5)(v_\delta, (v_\delta)_x) - \partial_x\{(\partial_{uu_x} f_5)(v_\delta, (v_\delta)_x)\}\big) \tag{5.5.31}$$

$$p_{\geq 4} := r_0(T_\delta) - \varepsilon^3\big[(\partial_{uu} f_5)(v_\delta, (v_\delta)_x) - \partial_x\{(\partial_{uu_x} f_5)(v_\delta, (v_\delta)_x)\}\big]\,. \tag{5.5.32}$$

Lemma 5.5.5. $\int_{\mathbb{T}} q_{>2}\, dx = 0$.

Proof. We already observed that $\tilde{q}$ has zero x-average as well as the derivative $\partial_x\{(\partial_{uu_x} f_5)(v, v_x)\}$. Finally

$$(\partial_{uu} f_5)(v, v_x) = \sum_{j_1, j_2, j_3 \in S} c_{j_1 j_2 j_3} v_{j_1} v_{j_2} v_{j_3} e^{\mathrm{i}(j_1+j_2+j_3)x}\,, \quad v := \sum_{j \in S} v_j e^{\mathrm{i} j x} \tag{5.5.33}$$

for some coefficient $c_{j_1 j_2 j_3}$, and therefore it has zero average by hypothesis ($\mathtt{S1}$). $\square$

By Lemma 5.5.4 and the results of this Section (in particular (5.5.30)) we deduce:

Proposition 5.5.1. *Assume* (5.5.8). *Then the Hamiltonian operator $\mathcal{L}_\omega$ has the form,* $\forall h \in H^s_{S^\perp}(\mathbb{T}^{\nu+1})$,

$$\mathcal{L}_\omega h := \omega\cdot\partial_\varphi h - \partial_x K_{02} h = \Pi_S^\perp\big(\omega\cdot\partial_\varphi h + \partial_{xx}(a_1 \partial_x h) + \partial_x(a_0 h) - \varepsilon^2 \partial_x \mathcal{R}_2 h - \partial_x \mathcal{R}_* h\big) \tag{5.5.34}$$

where $\mathcal{R}_2$ is defined in (5.5.29), *$\mathcal{R}_* := \mathcal{R}_{>2} + R(\psi)$ (with $R(\psi)$ defined in Lemma 5.5.4), the functions*

$$a_1 := 1 - r_1(T_\delta)\,, \quad a_0 := -(\varepsilon p_1 + \varepsilon^2 p_2 + q_{>2} + p_{\geq 4})\,, \quad p_1 := 6 v_\delta\,, \quad p_2 := 6\pi_0[(\partial_x^{-1} v_\delta)^2]\,, \tag{5.5.35}$$

the function $q_{>2}$ is defined in (5.5.31) *and satisfies $\int_{\mathbb{T}} q_{>2}\, dx = 0$, the function $p_{\geq 4}$ is defined in* (5.5.32), *r_1 in* (5.5.25), *T_δ and v_δ in* (5.5.11). *For $p_k = p_1, p_2$,*

$$\|p_k\|_s^{\mathrm{Lip}(\gamma)} \leq_s 1 + \|\mathfrak{I}_\delta\|_s^{\mathrm{Lip}(\gamma)}, \qquad \|\partial_i p_k[\hat{\imath}]\|_s \leq_s \|\hat{\imath}\|_{s+1} + \|\mathfrak{I}_\delta\|_{s+1}\|\hat{\imath}\|_{s_0+1}, \tag{5.5.36}$$

$$\|q_{>2}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^3 + \varepsilon^b \|\mathfrak{I}_\delta\|_s^{\mathrm{Lip}(\gamma)}\,, \qquad \|\partial_i q_{>2}[\hat{\imath}]\|_s \leq_s \varepsilon^b\big(\|\hat{\imath}\|_{s+1} + \|\mathfrak{I}_\delta\|_{s+1}\|\hat{\imath}\|_{s_0+1}\big), \tag{5.5.37}$$

$$\|a_1 - 1\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^3\big(1 + \|\mathfrak{I}_\delta\|_{s+1}^{\mathrm{Lip}(\gamma)}\big)\,, \qquad \|\partial_i a_1[\hat{\imath}]\|_s \leq_s \varepsilon^3\big(\|\hat{\imath}\|_{s+1} + \|\mathfrak{I}_\delta\|_{s+1}\|\hat{\imath}\|_{s_0+1}\big)\,, \tag{5.5.38}$$

$$\|p_{\geq 4}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4 + \varepsilon^{b+2}\|\mathfrak{I}_\delta\|_{s+2}^{\mathrm{Lip}(\gamma)}\,, \qquad \|\partial_i p_{\geq 4}[\hat{\imath}]\|_s \leq_s \varepsilon^{b+2}\big(\|\hat{\imath}\|_{s+2} + \|\mathfrak{I}_\delta\|_{s+2}\|\hat{\imath}\|_{s_0+2}\big)\,, \tag{5.5.39}$$

where $\mathfrak{I}_\delta(\varphi) := (\theta_0(\varphi) - \varphi, y_\delta(\varphi), z_0(\varphi))$ *corresponds to* T_δ*. The remainder* $\mathcal{R}_2$ *has the form* (5.5.2) *with*

$$\|g_j\|_s^{\mathrm{Lip}(\gamma)} + \|\chi_j\|_s^{\mathrm{Lip}(\gamma)} \leq_s 1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \quad \|\partial_i g_j[\widehat{\imath}]\|_s + \|\partial_i \chi_j[\widehat{\imath}]\|_s \leq_s \|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma} \tag{5.5.40}$$

and also $\mathcal{R}_*$ *has the form* (5.5.2) *with*

$$\|g_j^*\|_s^{\mathrm{Lip}(\gamma)}\|\chi_j^*\|_{s_0}^{\mathrm{Lip}(\gamma)} + \|g_j^*\|_{s_0}^{\mathrm{Lip}(\gamma)}\|\chi_j^*\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^3 + \varepsilon^{b+1}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \tag{5.5.41}$$

$$\begin{aligned}\|\partial_i g_j^*[\widehat{\imath}]\|_s\|\chi_j^*\|_{s_0} + \|\partial_i g_j^*[\widehat{\imath}]\|_{s_0}\|\chi_j^*\|_s + \|g_j^*\|_{s_0}\|\partial_i\chi_j^*[\widehat{\imath}]\|_s + \|g_j^*\|_s\|\partial_i\chi_j^*[\widehat{\imath}]\|_{s_0} &\leq_s \varepsilon^{b+1}\|\widehat{\imath}\|_{s+\sigma} \\ &+ \varepsilon^{2b-1}\|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}.\end{aligned} \tag{5.5.42}$$

The bounds (5.5.40), (5.5.41) imply, by Lemma 5.5.3, estimates for the s-decay norms of $\mathcal{R}_2$ and $\mathcal{R}_*$. The linearized operator $\mathcal{L}_\omega := \mathcal{L}_\omega(\omega, i_\delta(\omega))$ depends on the parameter ω both directly and also through the dependence on the torus $i_\delta(\omega)$. We have estimated also the partial derivative ∂_i with respect to the variables i (see (5.3.1)) in order to control, along the nonlinear Nash-Moser iteration, the Lipschitz variation of the eigenvalues of $\mathcal{L}_\omega$ with respect to ω and the approximate solution i_δ.

5.6 Reduction of the linearized operator in the normal directions

The goal of this Section is to conjugate the Hamiltonian operator $\mathcal{L}_\omega$ in (5.5.34) to the diagonal operator $\mathcal{L}_\infty$ defined in (5.6.121). The proof is obtained applying different kind of symplectic transformations. We shall always assume (5.5.8).

5.6.1 Change of the space variable

The first task is to conjugate $\mathcal{L}_\omega$ in (5.5.34) to $\mathcal{L}_1$ in (5.6.31), which has the coefficient of ∂_{xxx} independent on the space variable. We look for a φ-dependent family of *symplectic* diffeomorphisms $\Phi(\varphi)$ of $H_S^\perp$ which differ from

$$\mathcal{A}_\perp := \Pi_S^\perp \mathcal{A}\Pi_S^\perp, \quad (\mathcal{A}h)(\varphi, x) := (1 + \beta_x(\varphi, x))h(\varphi, x + \beta(\varphi, x)), \tag{5.6.1}$$

up to a small "finite dimensional" remainder, see (5.6.6). Each $\mathcal{A}(\varphi)$ is a symplectic map of the phase space, see Remark 4.1.3. If $\|\beta\|_{W^{1,\infty}} < 1/2$ then $\mathcal{A}$ is invertible, see Lemma A.0.10, and its inverse and adjoint maps are

$$(\mathcal{A}^{-1}h)(\varphi, y) := (1 + \tilde\beta_y(\varphi, y))h(\varphi, y + \tilde\beta(\varphi, y)), \quad (\mathcal{A}^T h)(\varphi, y) = h(\varphi, y + \tilde\beta(\varphi, y)) \tag{5.6.2}$$

where $x = y + \tilde\beta(\varphi, y)$ is the inverse diffeomorphism (of $\mathbb{T}$) of $y = x + \beta(\varphi, x)$.

The restricted maps $\mathcal{A}_\perp(\varphi) : H_S^\perp \to H_S^\perp$ are not symplectic. In order to find a symplectic diffeomorphism near $\mathcal{A}_\perp(\varphi)$, the first observation is that each $\mathcal{A}(\varphi)$ can be seen as the time 1-flow of a time dependent Hamiltonian PDE. Indeed $\mathcal{A}(\varphi)$ (for simplicity we skip the dependence on φ) is homotopic to the identity via the path of symplectic diffeomorphisms

$$u \mapsto (1 + \tau\beta_x)u(x + \tau\beta(x)), \quad \tau \in [0, 1],$$

which is the trajectory solution of the time dependent, linear Hamiltonian PDE

$$\partial_\tau u = \partial_x(b(\tau,x)u)\,, \quad b(\tau,x) := \frac{\beta(x)}{1+\tau\beta_x(x)}\,, \tag{5.6.3}$$

with value $u(x)$ at $\tau = 0$ and $\mathcal{A}u = (1+\beta_x(x))u(x+\beta(x))$ at $\tau = 1$. The equation (5.6.3) is a *transport* equation. Its associated charactheristic ODE is

$$\frac{d}{d\tau}x = -b(\tau,x)\,. \tag{5.6.4}$$

We denote its flow by $\gamma^{\tau_0,\tau}$, namely $\gamma^{\tau_0,\tau}(y)$ is the solution of (5.6.4) with $\gamma^{\tau_0,\tau_0}(y) = y$. Each $\gamma^{\tau_0,\tau}$ is a diffeomorphism of the torus $\mathbb{T}$.

Remark 5.6.1. *Let $y \mapsto y+\tilde\beta(\tau,y)$ be the inverse diffeomorpshim of $x \mapsto x+\tau\beta(x)$. Differentiating the identity $\tilde\beta(\tau,y)+\tau\beta(y+\tilde\beta(\tau,y)) = 0$ with respect to τ it results that $\gamma^\tau(y) := \gamma^{0,\tau}(y) = y+\tilde\beta(\tau,y)$.*

Then we define a symplectic map Φ of $H_S^\perp$ as the time-1 flow of the Hamiltonian PDE

$$\partial_\tau u = \Pi_S^\perp\partial_x(b(\tau,x)u) = \partial_x(b(\tau,x)u) - \Pi_S\partial_x(b(\tau,x)u)\,, \quad u \in H_S^\perp\,. \tag{5.6.5}$$

Note that $\Pi_S^\perp\partial_x(b(\tau,x)u)$ is the Hamiltonian vector field generated by $\frac{1}{2}\int_{\mathbb{T}} b(\tau,x)u^2dx$ restricted to $H_S^\perp$. We denote by $\Phi^{\tau_0,\tau}$ the flow of (5.6.5), namely $\Phi^{\tau_0,\tau}(u_0)$ is the solution of (5.6.5) with initial condition $\Phi^{\tau_0,\tau_0}(u_0) = u_0$. The flow is well defined in Sobolev spaces $H^s_{S^\perp}(\mathbb{T})$ for $b(\tau,x)$ is smooth enough (standard theory of linear hyperbolic PDEs, see e.g. Section 0.8 in [71]). It is natural to expect that the difference between the flow map $\Phi := \Phi^{0,1}$ and $\mathcal{A}_\perp$ is a "finite-dimensional" remainder of the size of β.

Lemma 5.6.1. *For $\|\beta\|_{W^{s_0+1,\infty}}$ small, there exists an invertible symplectic transformation $\Phi = \mathcal{A}_\perp + \mathcal{R}_\Phi$ of $H^s_{S^\perp}$, where $\mathcal{A}_\perp$ is defined in* (5.6.1) *and $\mathcal{R}_\Phi$ is a "finite-dimensional" remainder*

$$\mathcal{R}_\Phi h = \sum_{j\in S}\int_0^1 (h, g_j(\tau))_{L^2(\mathbb{T})}\chi_j(\tau)d\tau + \sum_{j\in S}\big(h,\psi_j\big)_{L^2(\mathbb{T})}e^{\mathrm{i}jx} \tag{5.6.6}$$

for some functions $\chi_j(\tau), g_j(\tau), \psi_j \in H^s$ satisfying

$$\|\psi_j\|_s\,,\ \|g_j(\tau)\|_s \le_s \|\beta\|_{W^{s+2,\infty}}\,, \quad \|\chi_j(\tau)\|_s \le_s 1+\|\beta\|_{W^{s+1,\infty}}\,, \quad \forall\tau\in[0,1]\,. \tag{5.6.7}$$

Furthermore, the following tame estimates holds

$$\|\Phi^{\pm1}h\|_s \le_s \|h\|_s + \|\beta\|_{W^{s+2,\infty}}\|h\|_{s_0}\,, \quad \forall h \in H^s_{S^\perp}\,. \tag{5.6.8}$$

Proof. Let $w(\tau,x) := (\Phi^\tau u_0)(x)$ denote the solution of (5.6.5) with initial condition $\Phi^0(w) = u_0 \in H_S^\perp$. The difference

$$(\mathcal{A}_\perp - \Phi)u_0 = \Pi_S^\perp\mathcal{A}u_0 - w(1,\cdot) = \mathcal{A}u_0 - w(1,\cdot) - \Pi_S\mathcal{A}u_0\,, \quad \forall u_0 \in H_S^\perp\,, \tag{5.6.9}$$

and

$$\Pi_S\mathcal{A}u_0 = \Pi_S(\mathcal{A}-I)\Pi_S^\perp u_0 = \sum_{j\in S}\big(u_0\,,\,\psi_j\big)_{L^2(\mathbb{T})}e^{\mathrm{i}jx}\,, \quad \psi_j := (\mathcal{A}^T - I)e^{\mathrm{i}jx}\,. \tag{5.6.10}$$

We claim that the difference

$$\mathcal{A}u_0 - w(1,x) = (1+\beta_x(x)) \int_0^1 (1+\tau\beta_x(x))^{-1} \big[\Pi_S \partial_x (b(\tau) w(\tau))\big](\gamma^\tau(x+\beta(x)))\, d\tau \tag{5.6.11}$$

where $\gamma^\tau(y) := \gamma^{0,\tau}(y)$ is the flow of (5.6.4). Indeed the solution $w(\tau,x)$ of (5.6.5) satisfies

$$\partial_\tau \{ w(\tau, \gamma^\tau(y)) \} = b_x(\tau, \gamma^\tau(y)) w(\tau, \gamma^\tau(y)) - \big[\Pi_S \partial_x (b(\tau) w(\tau))\big](\gamma^\tau(y)) \,.$$

Then, by the variation of constant formula, we find

$$w(\tau, \gamma^\tau(y)) = e^{\int_0^\tau b_x(s,\gamma^s(y))\, ds} \Big(u_0(y) - \int_0^\tau e^{-\int_0^s b_x(\zeta, \gamma^\zeta(y))\, d\zeta} \big[\Pi_S \partial_x (b(s) w(s))\big](\gamma^s(y))\, ds \Big) \,.$$

Since $\partial_y \gamma^\tau(y)$ solves the variational equation $\partial_\tau(\partial_y \gamma^\tau(y)) = -b_x(\tau, \gamma^\tau(y))(\partial_y \gamma^\tau(y))$ with $\partial_y \gamma^0(y) = 1$ we have that

$$e^{\int_0^\tau b_x(s,\gamma^s(y)) ds} = \big(\partial_y \gamma^\tau(y)\big)^{-1} = 1 + \tau\beta_x(x) \tag{5.6.12}$$

by remark 5.6.1, and so we derive the expression

$$w(\tau,x) = (1+\tau\beta_x(x)) \Big\{ u_0(x+\tau\beta(x)) - \int_0^\tau (1+s\beta_x(x))^{-1} \big[\Pi_S \partial_x (b(s) w(s))\big](\gamma^s(x+\tau\beta(x)))\, ds \Big\} \,.$$

Evaluating at $\tau = 1$, formula (5.6.11) follows. Next, we develop (recall $w(\tau) = \Phi^\tau(u_0)$)

$$[\Pi_S \partial_x (b(\tau) w(\tau))](x) = \sum_{j\in S} \big(u_0, g_j(\tau)\big)_{L^2(\mathbb{T})} e^{\mathrm{i}jx} \,, \quad g_j(\tau) := -(\Phi^\tau)^T [b(\tau) \partial_x e^{\mathrm{i}jx}] \,, \tag{5.6.13}$$

and (5.6.11) becomes

$$\mathcal{A}u_0 - w(1,\cdot) = -\int_0^1 \sum_{j\in S} \big(u_0 \,,\, g_j(\tau)\big)_{L^2(\mathbb{T})} \chi_j(\tau,\cdot)\, d\tau \,, \tag{5.6.14}$$

where

$$\chi_j(\tau,x) := -(1+\beta_x(x))(1+\tau\beta_x(x))^{-1} e^{\mathrm{i}j\gamma^\tau(x+\beta(x))} \,. \tag{5.6.15}$$

By (5.6.9), (5.6.10), (5.6.11), (5.6.14) we deduce that $\Phi = \mathcal{A}_\perp + \mathcal{R}_\Phi$ as in (5.6.6).

We now prove the estimates (5.6.7). Each function ψ_j in (5.6.10) satisfies $\|\psi_j\|_s \le_s \|\beta\|_{W^{s,\infty}}$, see (5.6.2). The bound $\|\chi_j(\tau)\|_s \le_s 1 + \|\beta\|_{W^{s+1,\infty}}$ follows by (5.6.15). The tame estimates for $g_j(\tau)$ defined in (5.6.13) are more difficult because require tame estimates for the adjoint $(\Phi^\tau)^T$, $\forall \tau \in [0,1]$. The adjoint of the flow map can be represented as the flow map of the "adjoint" PDE

$$\partial_\tau z = \Pi_S^\perp \{ b(\tau,x) \partial_x \Pi_S^\perp z \} = b(\tau,x)\partial_x z - \Pi_S (b(\tau,x)\partial_x z) \,, \quad z \in H_S^\perp \,, \tag{5.6.16}$$

where $-\Pi_S^\perp b(\tau,x)\partial_x$ is the L^2-adjoint of the Hamiltonian vector field in (5.6.5). We denote by $\Psi^{\tau_0,\tau}$ the flow of (5.6.16), namely $\Psi^{\tau_0,\tau}(v)$ is the solution of (5.6.16) with $\Psi^{\tau_0,\tau_0}(v) = v$. Since the derivative $\partial_\tau (\Phi^\tau(u_0), \Psi^{\tau_0,\tau}(v))_{L^2(\mathbb{T})} = 0$, $\forall \tau$, we deduce that $(\Phi^{\tau_0}(u_0), \Psi^{\tau_0,\tau_0}(v))_{L^2(\mathbb{T})} = (\Phi^0(u_0), \Psi^{\tau_0,0}(v))_{L^2(\mathbb{T})}$, namely (recall that $\Psi^{\tau_0,\tau_0}(v) = v$) the adjoint

$$(\Phi^{\tau_0})^T = \Psi^{\tau_0,0} \,, \quad \forall \tau_0 \in [0,1] \,. \tag{5.6.17}$$

Thus it is sufficient to prove tame estimates for the flow $\Psi^{\tau_0,\tau}$. We first provide a useful expression for the solution $z(\tau,x) := \Psi^{\tau_0,\tau}(v)$ of (5.6.16), obtained by the methods of characteristics. Let $\gamma^{\tau_0,\tau}(y)$ be the flow of (5.6.4). Since $\partial_\tau z(\tau,\gamma^{\tau_0,\tau}(y)) = -[\Pi_S(b(\tau)\partial_x z(\tau)](\gamma^{\tau_0,\tau}(y))$ we get

$$z(\tau,\gamma^{\tau_0,\tau}(y)) = v(y) + \int_\tau^{\tau_0} [\Pi_S(b(s)\partial_x z(s)](\gamma^{\tau_0,s}(y))\, ds\,, \quad \forall \tau \in [0,1]\,.$$

Denoting by $y = x + \sigma(\tau,x)$ the inverse diffeomorphism of $x = \gamma^{\tau_0,\tau}(y) = y + \tilde\sigma(\tau,y)$, we get

$$\begin{aligned}
\Psi^{\tau_0,\tau}(v) = z(\tau,x) &= v(x+\sigma(\tau,x)) + \int_\tau^{\tau_0} [\Pi_S(b(s)\partial_x z(s)](\gamma^{\tau_0,s}(x+\sigma(\tau,x)))\, ds \\
&= v(x+\sigma(\tau,x)) + \int_\tau^{\tau_0} \sum_{j\in S} (z(s),p_j(s))\kappa_j(s,x)\, ds \\
&= v(x+\sigma(\tau,x)) + \mathcal{R}_\tau v\,,
\end{aligned} \tag{5.6.18}$$

where $p_j(s) := -\partial_x(b(s)e^{\mathrm{i}jx})$, $\kappa_j(s,x) := e^{\mathrm{i}j\gamma^{\tau_0,s}(x+\sigma(\tau,x))}$ and

$$(\mathcal{R}_\tau v)(x) := \int_\tau^{\tau_0} \sum_{j\in S} (\Psi^{\tau_0,s}(v),p_j(s))_{L^2(\mathbb{T})}\kappa_j(s,x)\, ds\,.$$

Since $\|\sigma(\tau,\cdot)\|_{W^{s,\infty}}, \|\tilde\sigma(\tau,\cdot)\|_{W^{s,\infty}} \le_s \|\beta\|_{W^{s+1,\infty}}$ (recall also (5.6.3)), we derive $\|p_j\|_s \le_s \|\beta\|_{W^{s+2,\infty}}$, $\|\kappa_j\|_s \le_s 1 + \|\beta\|_{W^{s+1,\infty}}$ and $\|v(x+\sigma(\tau,x))\|_s \le_s \|v\|_s + \|\beta\|_{W^{s+1,\infty}}\|v\|_{s_0}$, $\forall \tau \in [0,1]$. Moreover

$$\|\mathcal{R}_\tau v\|_s \le_s \sup_{\tau\in[0,1]}\|\Psi^{\tau_0,\tau}(v)\|_s \|\beta\|_{W^{s_0+2,\infty}} + \sup_{\tau\in[0,1]}\|\Psi^{\tau_0,\tau}(v)\|_{s_0}\|\beta\|_{W^{s+2,\infty}}\,.$$

Therefore, for all $\tau \in [0,1]$,

$$\begin{aligned}
\|\Psi^{\tau_0,\tau} v\|_s \le_s{}& \|v\|_s + \|\beta\|_{W^{s+1,\infty}}\|v\|_{s_0} \\
&+ \sup_{\tau\in[0,1]}\big\{\|\Psi^{\tau_0,\tau}v\|_s\|\beta\|_{W^{s_0+2,\infty}} + \|\Psi^{\tau_0,\tau}v\|_{s_0}\|\beta\|_{W^{s+2,\infty}}\big\}\,.
\end{aligned} \tag{5.6.19}$$

For $s = s_0$ it implies

$$\sup_{\tau\in[0,1]}\|\Psi^{\tau_0,\tau}(v)\|_{s_0} \le_{s_0} \|v\|_{s_0}(1 + \|\beta\|_{W^{s_0+1,\infty}}) + \sup_{\tau\in[0,1]}\|\Psi^{\tau_0,\tau}(v)\|_{s_0}\|\beta\|_{W^{s_0+2,\infty}}$$

and so, for $\|\beta\|_{W^{s_0+2,\infty}} \le c(s_0)$ small enough,

$$\sup_{\tau\in[0,1]}\|\Psi^{\tau_0,\tau}(v)\|_{s_0} \le_{s_0} \|v\|_{s_0}\,. \tag{5.6.20}$$

Finally (5.6.19), (5.6.20) imply the tame estimate

$$\sup_{\tau\in[0,1]}\|\Psi^{\tau_0,\tau}(v)\|_s \le_s \|v\|_s + \|\beta\|_{W^{s+2,\infty}}\|v\|_{s_0}\,. \tag{5.6.21}$$

By (5.6.17) and (5.6.21) we deduce the bound (5.6.7) for g_j defined in (5.6.13). The tame estimate (5.6.8) for Φ follows by that of $\mathcal{A}$ and (5.6.7) (use Lemma A.0.10). The estimate for Φ^{-1} follows in the same way because $\Phi^{-1} = \Phi^{1,0}$ is the backward flow. $\square$

We conjugate $\mathcal{L}_\omega$ in (5.5.34) via the symplectic map $\Phi = \mathcal{A}_\perp + \mathcal{R}_\Phi$ of Lemma 5.6.1. We compute (split $\Pi_S^\perp = I - \Pi_S$)

$$\mathcal{L}_\omega \Phi = \Phi \mathcal{D}_\omega + \Pi_S^\perp \mathcal{A}\big(b_3 \partial_{yyy} + b_2 \partial_{yy} + b_1 \partial_y + b_0\big)\Pi_S^\perp + \mathcal{R}_I \,, \tag{5.6.22}$$

where the coefficients are

$$b_3(\varphi, y) := \mathcal{A}^T[a_1(1+\beta_x)^3] \qquad b_2(\varphi, y) := \mathcal{A}^T\big[2(a_1)_x(1+\beta_x)^2 + 6a_1\beta_{xx}(1+\beta_x)\big] \tag{5.6.23}$$

$$\begin{aligned} b_1(\varphi, y) := \mathcal{A}^T\Big[&(\mathcal{D}_\omega \beta) + 3a_1 \frac{\beta_{xx}^2}{1+\beta_x} + 4a_1\beta_{xxx} + 6(a_1)_x\beta_{xx} \\ &+ (a_1)_{xx}(1+\beta_x) + a_0(1+\beta_x)\Big] \end{aligned} \tag{5.6.24}$$

$$b_0(\varphi, y) := \mathcal{A}^T\Big[\frac{(\mathcal{D}_\omega \beta_x)}{1+\beta_x} + a_1 \frac{\beta_{xxxx}}{1+\beta_x} + 2(a_1)_x \frac{\beta_{xxx}}{1+\beta_x} + (a_1)_{xx}\frac{\beta_{xx}}{1+\beta_x} + a_0 \frac{\beta_{xx}}{1+\beta_x} + (a_0)_x\Big] \tag{5.6.25}$$

and the remainder

$$\begin{aligned} \mathcal{R}_I := &-\Pi_S^\perp \partial_x(\varepsilon^2 \mathcal{R}_2 + \mathcal{R}_*)\mathcal{A}_\perp - \Pi_S^\perp\big(a_1\partial_{xxx} + 2(a_1)_x\partial_{xx} + ((a_1)_{xx} + a_0)\partial_x + (a_0)_x\big)\Pi_S \mathcal{A}\Pi_S^\perp \\ &+ [\mathcal{D}_\omega, \mathcal{R}_\Phi] + (\mathcal{L}_\omega - \mathcal{D}_\omega)\mathcal{R}_\Phi \,. \end{aligned} \tag{5.6.26}$$

The commutator $[\mathcal{D}_\omega, \mathcal{R}_\Phi]$ has the form (5.6.6) with $\mathcal{D}_\omega g_j$ or $\mathcal{D}_\omega \chi_j$, $\mathcal{D}_\omega \psi_j$ instead of χ_j, g_j, ψ_j respectively. Also the last term $(\mathcal{L}_\omega - \mathcal{D}_\omega)\mathcal{R}_\Phi$ in (5.6.26) has the form (5.6.6) (note that $\mathcal{L}_\omega - \mathcal{D}_\omega$ does not contain derivatives with respect to φ). By (5.6.22), and decomposing $I = \Pi_S + \Pi_S^\perp$, we get

$$\mathcal{L}_\omega \Phi = \Phi(\mathcal{D}_\omega + b_3\partial_{yyy} + b_2\partial_{yy} + b_1\partial_y + b_0)\Pi_S^\perp + \mathcal{R}_{II} \,, \tag{5.6.27}$$

$$\mathcal{R}_{II} := \big\{\Pi_S^\perp(\mathcal{A} - I)\Pi_S - \mathcal{R}_\Phi\big\}(b_3\partial_{yyy} + b_2\partial_{yy} + b_1\partial_y + b_0)\Pi_S^\perp + \mathcal{R}_I \,. \tag{5.6.28}$$

Now we choose the function $\beta = \beta(\varphi, x)$ such that

$$a_1(\varphi, x)(1 + \beta_x(\varphi, x))^3 = b_3(\varphi) \tag{5.6.29}$$

so that the coefficient b_3 in (5.6.23) depends only on φ (note that $\mathcal{A}^T[b_3(\varphi)] = b_3(\varphi)$). The only solution of (5.6.29) with zero space average is

$$\beta := \partial_x^{-1}\rho_0, \quad \rho_0 := b_3(\varphi)^{1/3}(a_1(\varphi, x))^{-1/3} - 1, \quad b_3(\varphi) := \Big(\int_{\mathbb{T}} (a_1(\varphi, x))^{-1/3} dx\Big)^{-3}. \tag{5.6.30}$$

Applying the symplectic map Φ^{-1} in (5.6.27) we obtain the Hamiltonian operator (see Definition 3.3.1)

$$\mathcal{L}_1 := \Phi^{-1}\mathcal{L}_\omega \Phi = \Pi_S^\perp\big(\omega \cdot \partial_\varphi + b_3(\varphi)\partial_{yyy} + b_1\partial_y + b_0\big)\Pi_S^\perp + \mathfrak{R}_1 \tag{5.6.31}$$

where $\mathfrak{R}_1 := \Phi^{-1}\mathcal{R}_{II}$. We used that, by the Hamiltonian nature of $\mathcal{L}_1$, the coefficient $b_2 = 2(b_3)_y$ and so, by the choice (5.6.30), we have $b_2 = 2(b_3)_y = 0$. In the next Lemma we analyse the structure of the remainder $\mathfrak{R}_1$.

Lemma 5.6.2. *The operator $\mathfrak{R}_1$ has the form (5.5.7).*

Proof. The remainders $\mathcal{R}_I$ and $\mathcal{R}_{II}$ have the form (5.5.7). Indeed $\mathcal{R}_2, \mathcal{R}_*$ in (5.6.26) have the form (5.5.2) (see Proposition 5.5.1) and the term $\Pi_S \mathcal{A} w = \sum_{j \in S} (\mathcal{A}^T e^{\mathrm{i} jx}, w)_{L^2(\mathbb{T})} e^{\mathrm{i} jx}$ has the same form. By (5.6.6), the terms of $\mathcal{R}_I$, $\mathcal{R}_{II}$ which involves the operator $\mathcal{R}_\Phi$ have the form (5.5.7). All the operations involved preserve this structure: if $R_\tau w = \chi(\tau)(w, g(\tau))_{L^2(\mathbb{T})}$, $\tau \in [0,1]$, then

$$R_\tau \Pi_S^\perp w = \chi(\tau)(\Pi_S^\perp g(\tau), w)_{L^2(\mathbb{T})}\,, \quad R_\tau \mathcal{A} w = \chi(\tau)(\mathcal{A}^T g(\tau), w)_{L^2(\mathbb{T})}\,, \quad \partial_x R_\tau w = \chi_x(\tau)(g(\tau), w)_{L^2(\mathbb{T})}\,,$$
$$\Pi_S^\perp R_\tau w = (\Pi_S^\perp \chi(\tau))(g(\tau), w)_{L^2(\mathbb{T})}\,, \; \mathcal{A} R_\tau w = (\mathcal{A}\chi(\tau))(g(\tau), w)_{L^2(\mathbb{T})}\,, \; \Phi^{-1} R_\tau w = (\Phi^{-1}\chi(\tau))(g(\tau), w)_{L^2(\mathbb{T})}$$

(the last equality holds because $\Phi^{-1}(f(\varphi) w) = f(\varphi)\Phi^{-1}(w)$ for all function $f(\varphi)$). Hence $\mathfrak{R}_1$ has the form (5.5.7) where $\chi_j(\tau) \in H_S^\perp$ for all $\tau \in [0,1]$. □

We now put in evidence the terms of order $\varepsilon, \varepsilon^2, \ldots$, in b_1, b_0, $\mathfrak{R}_1$, recalling that $a_1 - 1 = O(\varepsilon^3)$ (see (5.5.38)), $a_0 = O(\varepsilon)$ (see (5.5.35)-(5.5.39)), and $\beta = O(\varepsilon^3)$ (proved below in (5.6.35)). We expand b_1 in (5.6.24) as

$$b_1 = -\varepsilon p_1 - \varepsilon^2 p_2 - q_{>2} + \mathcal{D}_\omega \beta + 4\beta_{xxx} + (a_1)_{xx} + b_{1,\geq 4} \tag{5.6.32}$$

where $b_{1,\geq 4} = O(\varepsilon^4)$ is defined by difference (the precise estimate is in Lemma 5.6.3).

Remark 5.6.2. *The function $\mathcal{D}_\omega \beta$ has zero average in x by* (5.6.30) *as well as* $(a_1)_{xx}, \beta_{xxx}$.

Similarly, we expand b_0 in (5.6.25) as

$$b_0 = -\varepsilon (p_1)_x - \varepsilon^2 (p_2)_x - (q_{>2})_x + \mathcal{D}_\omega \beta_x + \beta_{xxxx} + b_{0,\geq 4} \tag{5.6.33}$$

where $b_{0,\geq 4} = O(\varepsilon^4)$ is defined by difference.

Using the equalities (5.6.28), (5.6.26) and $\Pi_S \mathcal{A} \Pi_S^\perp = \Pi_S(\mathcal{A} - I)\Pi_S^\perp$ we get

$$\mathfrak{R}_1 := \Phi^{-1}\mathcal{R}_{II} = -\varepsilon^2 \Pi_S^\perp \partial_x \mathcal{R}_2 + \mathcal{R}_* \tag{5.6.34}$$

where $\mathcal{R}_2$ is defined in (5.5.29) and we have renamed $\mathcal{R}_*$ the term of order $o(\varepsilon^2)$ in $\mathfrak{R}_1$. The remainder $\mathcal{R}_*$ in (5.6.34) has the form (5.5.7).

Lemma 5.6.3. *There is $\sigma = \sigma(\tau, \nu) > 0$ such that*

$$\|\beta\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^3 (1 + \|\mathfrak{I}_\delta\|_{s+1}^{\mathrm{Lip}(\gamma)}), \qquad \|\partial_i \beta[\hat\imath]\|_s \leq_s \varepsilon^3 \big(\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}\big)\,, \tag{5.6.35}$$
$$\|b_3 - 1\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4 + \varepsilon^{b+2}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \qquad \|\partial_i b_3[\hat\imath]\|_s \leq_s \varepsilon^{b+2}\big(\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}\big) \tag{5.6.36}$$
$$\|b_{1,\geq 4}\|_s^{\mathrm{Lip}(\gamma)} + \|b_{0,\geq 4}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4 + \varepsilon^{b+2}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \tag{5.6.37}$$
$$\|\partial_i b_{1,\geq 4}[\hat\imath]\|_s + \|\partial_i b_{0,\geq 4}[\hat\imath]\|_s \leq_s \varepsilon^{b+2}\big(\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}\big). \tag{5.6.38}$$

The transformations Φ, Φ^{-1} satisfy

$$\|\Phi^{\pm 1} h\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|h\|_{s+1}^{\mathrm{Lip}(\gamma)} + \|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|h\|_{s_0+1}^{\mathrm{Lip}(\gamma)} \tag{5.6.39}$$
$$\|\partial_i(\Phi^{\pm 1} h)[\hat\imath]\|_s \leq_s \|h\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma} + \|h\|_{s_0+\sigma}\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|h\|_{s_0+\sigma}\|\hat\imath\|_{s_0+\sigma}\,. \tag{5.6.40}$$

Moreover the remainder $\mathcal{R}_*$ has the form (5.5.7), *where the functions $\chi_j(\tau)$, $g_j(\tau)$ satisfy the estimates* (5.5.41)-(5.5.42) *uniformly in $\tau \in [0,1]$.*

Proof. The estimates (5.6.35) follow by (5.6.30), (5.5.38), and the usual interpolation and tame estimates in Lemmata A.0.8-A.0.10 (and Lemma 5.3.13) and (5.5.8). For the estimates of b_3, by (5.6.30) and (5.5.35) we consider the function r_1 defined in (5.5.25). Recalling also (5.1.4) and (5.5.11), the function

$$r_1(T_\delta) = \varepsilon^3(\partial_{u_x u_x} f_5)(v_\delta, (v_\delta)_x) + r_{1,\geq 4}\,, \quad r_{1,\geq 4} := r_1(T_\delta) - \varepsilon^3(\partial_{u_x u_x} f_5)(v_\delta, (v_\delta)_x)\,.$$

Hypothesis (S1) implies, as in the proof of Lemma 5.5.5, that the space average $\fint_{\mathbb{T}}(\partial_{u_x u_x} f_5)(v_\delta, (v_\delta)_x)dx = 0$. Hence the bound (5.6.36) for $b_3 - 1$ follows. For the estimates on Φ, Φ^{-1} we apply Lemma 5.6.1 and the estimate (5.6.35) for β . We estimate the remainder $\mathcal{R}_*$ in (5.6.34), using (5.6.26), (5.6.28) and (5.5.41)-(5.5.42). $\square$

5.6.2 Reparametrization of time

The goal of this Section is to make constant the coefficient of the highest order spatial derivative operator ∂_{yyy}, by a quasi-periodic reparametrization of time. We consider the change of variable

$$(Bw)(\varphi, y) := w(\varphi + \omega\alpha(\varphi), y), \qquad (B^{-1}h)(\vartheta, y) := h(\vartheta + \omega\tilde{\alpha}(\vartheta), y)\,,$$

where $\varphi = \vartheta + \omega\tilde{\alpha}(\vartheta)$ is the inverse diffeomorphism of $\vartheta = \varphi + \omega\alpha(\varphi)$ in $\mathbb{T}^\nu$. By conjugation, the differential operators become

$$B^{-1}\omega \cdot \partial_\varphi B = \rho(\vartheta)\,\omega \cdot \partial_\vartheta, \quad B^{-1}\partial_y B = \partial_y, \quad \rho := B^{-1}(1 + \omega \cdot \partial_\varphi \alpha). \tag{5.6.41}$$

By (5.6.31), using also that B and B^{-1} commute with $\Pi_S^\perp$, we get

$$B^{-1}\mathcal{L}_1 B = \Pi_S^\perp[\rho\omega \cdot \partial_\vartheta + (B^{-1}b_3)\partial_{yyy} + (B^{-1}b_1)\partial_y + (B^{-1}b_0)]\Pi_S^\perp + B^{-1}\mathfrak{R}_1 B. \tag{5.6.42}$$

We choose α such that

$$(B^{-1}b_3)(\vartheta) = m_3\rho(\vartheta)\,, \quad m_3 \in \mathbb{R}\,, \quad \text{namely} \quad b_3(\varphi) = m_3(1 + \omega \cdot \partial_\varphi \alpha(\varphi)) \tag{5.6.43}$$

(recall (5.6.41)). The unique solution with zero average of (5.6.43) is

$$\alpha(\varphi) := \frac{1}{m_3}(\omega \cdot \partial_\varphi)^{-1}(b_3 - m_3)(\varphi), \qquad m_3 := \fint_{\mathbb{T}^\nu} b_3(\varphi)d\varphi\,. \tag{5.6.44}$$

Hence, by (5.6.42),

$$B^{-1}\mathcal{L}_1 B = \rho\mathcal{L}_2\,, \qquad \mathcal{L}_2 := \Pi_S^\perp(\omega \cdot \partial_\vartheta + m_3\partial_{yyy} + c_1\partial_y + c_0)\Pi_S^\perp + \mathfrak{R}_2 \tag{5.6.45}$$

$$c_1 := \rho^{-1}(B^{-1}b_1)\,, \qquad c_0 := \rho^{-1}(B^{-1}b_0)\,, \qquad \mathfrak{R}_2 := \rho^{-1}B^{-1}\mathfrak{R}_1 B\,. \tag{5.6.46}$$

The transformed operator $\mathcal{L}_2$ in (5.6.45) is still Hamiltonian, since the reparametrization of time preserves the Hamiltonian structure, see Remark 4.1.7.

We now put in evidence the terms of order $\varepsilon, \varepsilon^2, \ldots$ in c_1, c_0. To this aim, we anticipate the following estimates: $\rho(\vartheta) = 1 + O(\varepsilon^4)$, $\alpha = O(\varepsilon^4\gamma^{-1})$, $m_3 = 1 + O(\varepsilon^4)$, $B^{-1} - I = O(\alpha)$ (in low norm), which are proved in Lemma 5.6.4 below. Then, by (5.6.32)-(5.6.33), we expand the functions c_1, c_0 in (5.6.46) as

$$c_1 = -\varepsilon p_1 - \varepsilon^2 p_2 - B^{-1}q_{>2} + \varepsilon(p_1 - B^{-1}p_1) + \varepsilon^2(p_2 - B^{-1}p_2) + \mathcal{D}_\omega\beta + 4\beta_{xxx} + (a_1)_{xx} + c_{1,\geq 4}\,, \tag{5.6.47}$$

$$
\begin{aligned}
c_0 = -\varepsilon(p_1)_x - \varepsilon^2(p_2)_x - (B^{-1}q_{>2})_x + \varepsilon(p_1 - B^{-1}p_1)_x + \varepsilon^2(p_2 - B^{-1}p_2)_x + (\mathcal{D}_\omega\beta)_x \\
+ \beta_{xxxx} + c_{0,\geq 4}\,, \qquad (5.6.48)
\end{aligned}
$$

where $c_{1,\geq 4}, c_{0,\geq 4} = O(\varepsilon^4)$ are defined by difference.

Remark 5.6.3. *The functions $\varepsilon(p_1 - B^{-1}p_1) = O(\varepsilon^5\gamma^{-1})$ and $\varepsilon^2(p_2 - B^{-1}p_2) = O(\varepsilon^6\gamma^{-1})$, see* (5.6.53)*. For the reducibility scheme, the terms of order ∂_x^0 with size $O(\varepsilon^5\gamma^{-1})$ are perturbative, since $\varepsilon^5\gamma^{-2} \ll 1$.*

The remainder $\mathfrak{R}_2$ in (5.6.46) has still the form (5.5.7) and, by (5.6.34),

$$
\mathfrak{R}_2 := -\rho^{-1}B^{-1}\mathfrak{R}_1 B = -\varepsilon^2\Pi_S^\perp\partial_x\mathcal{R}_2 + \mathcal{R}_* \tag{5.6.49}
$$

where $\mathcal{R}_2$ is defined in (5.5.29) and we have renamed $\mathcal{R}_*$ the term of order $o(\varepsilon^2)$ in $\mathfrak{R}_2$.

Lemma 5.6.4. *There is $\sigma = \sigma(\nu,\tau) > 0$ (possibly larger than σ in Lemma 5.6.3) such that*

$$
|m_3 - 1|^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^4, \qquad |\partial_i m_3[\widehat{\imath}]| \leq C\varepsilon^{b+2}\|\widehat{\imath}\|_{s_0+\sigma} \tag{5.6.50}
$$

$$
\|\alpha\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4\gamma^{-1} + \varepsilon^{b+2}\gamma^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \qquad \|\partial_i\alpha[\widehat{\imath}]\|_s \leq_s \varepsilon^{b+2}\gamma^{-1}\big(\|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}\big)\,, \tag{5.6.51}
$$

$$
\|\rho - 1\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4 + \varepsilon^{b+2}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \qquad \|\partial_i\rho[\widehat{\imath}]\|_s \leq_s \varepsilon^{b+2}\big(\|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}\big) \tag{5.6.52}
$$

$$
\|p_k - B^{-1}p_k\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4\gamma^{-1} + \varepsilon^{b+2}\gamma^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \quad k = 1,2 \tag{5.6.53}
$$

$$
\|\partial_i(p_k - B^{-1}p_k)[\widehat{\imath}]\|_s \leq_s \varepsilon^{b+2}\gamma^{-1}\big(\|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}\big) \tag{5.6.54}
$$

$$
\|B^{-1}q_{>2}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^3 + \varepsilon^b\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \tag{5.6.55}
$$

$$
\|\partial_i(B^{-1}q_{>2})[\widehat{\imath}]\|_s \leq_s \varepsilon^b\big(\|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\widehat{\imath}\|_{s_0+\sigma}\big)\,. \tag{5.6.56}
$$

The terms $c_{1,\geq 4}, c_{0,\geq 4}$ satisfy the bounds (5.6.37)-(5.6.38)*. The transformations B, B^{-1} satisfy the estimates* (5.6.39), (5.6.40)*. The remainder $\mathcal{R}_*$ has the form* (5.5.7)*, and the functions $g_j(\tau)$, $\chi_j(\tau)$ satisfy the estimates* (5.5.41)-(5.5.42) *for all $\tau \in [0,1]$.*

Proof. (5.6.50) follows from (5.6.44),(5.6.36). The estimate $\|\alpha\|_s \leq_s \varepsilon^4\gamma^{-1} + \varepsilon^{b+2}\gamma^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma}$ and the inequality for $\partial_i\alpha$ in (5.6.51) follow by (5.6.44),(5.6.36),(5.6.50). For the first bound in (5.6.51) we also differentiate (5.6.44) with respect to the parameter ω. The estimates for ρ follow from $\rho - 1 = B^{-1}(b_3 - m_3)/m_3$. $\square$

5.6.3 Translation of the space variable

In view of the next linear Birkhoff normal form steps (whose goal is to eliminate the terms of size ε and ε^2), in the expressions (5.6.47), (5.6.48) we split $p_1 = \bar{p}_1 + (p_1 - \bar{p}_1)$, $p_2 = \bar{p}_2 + (p_2 - \bar{p}_2)$ (see (5.5.35)), where

$$
\bar{p}_1 := 6\bar{v}, \qquad \bar{p}_2 := 6\pi_0[(\partial_x^{-1}\bar{v})^2], \qquad \bar{v}(\varphi,x) := \sum\nolimits_{j\in S}\sqrt{\xi_j}e^{\mathrm{i}\ell(j)\cdot\varphi}e^{\mathrm{i}jx}, \tag{5.6.57}
$$

and $\ell : S \to \mathbb{Z}^\nu$ is the odd injective map (see (1.3.7))

$$
\ell : S \to \mathbb{Z}^\nu, \quad \ell(\bar{\jmath}_i) := e_i, \quad \ell(-\bar{\jmath}_i) := -\ell(\bar{\jmath}_i) = -e_i, \quad i = 1,\ldots,\nu, \tag{5.6.58}
$$

denoting by $e_i = (0,\ldots,1,\ldots,0)$ the i-th vector of the canonical basis of $\mathbb{R}^\nu$.

Remark 5.6.4. *All the functions $\bar p_1$, $\bar p_2$, $p_1 - \bar p_1$, $p_2 - \bar p_2$ have zero average in x.*

We write the variable coefficients c_1, c_0 of the operator $\mathcal{L}_2$ in (5.6.45) (see (5.6.47), (5.6.48)) as

$$c_1 = -\varepsilon \bar p_1 - \varepsilon^2 \bar p_2 + q_{c_1} + c_{1,\geq 4}\,, \qquad c_0 = -\varepsilon(\bar p_1)_x - \varepsilon^2(\bar p_2)_x + q_{c_0} + c_{0,\geq 4}\,, \tag{5.6.59}$$

where we define

$$q_{c_1} := q + 4\beta_{xxx} + (a_1)_{xx}\,, \quad q_{c_0} := q_x + \beta_{xxxx}, \tag{5.6.60}$$

$$q := \varepsilon(p_1 - B^{-1}p_1) + \varepsilon(\bar p_1 - p_1) + \varepsilon^2(p_2 - B^{-1}p_2) + \varepsilon^2(\bar p_2 - p_2) - B^{-1}q_{>2} + \mathcal{D}_\omega\beta\,. \tag{5.6.61}$$

Remark 5.6.5. *The functions q_{c_1}, q_{c_0} have zero average in x (see Remarks 5.6.4, 5.6.2 and Lemma 5.5.5).*

Lemma 5.6.5. *The functions $\bar p_k - p_k$, $k = 1, 2$ and q_{c_m}, $m = 0, 1$, satisfy*

$$\|\bar p_k - p_k\|_s^{\mathrm{Lip}(\gamma)} \leq_s \|\mathfrak{I}_\delta\|_s^{\mathrm{Lip}(\gamma)}, \qquad \|\partial_i(\bar p_k - p_k)[\hat\imath]\|_s \leq_s \|\hat\imath\|_s + \|\mathfrak{I}_\delta\|_s\|\hat\imath\|_{s_0}\,, \tag{5.6.62}$$

$$\|q_{c_m}\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^5\gamma^{-1} + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\,, \quad \|\partial_i q_{c_m}[\hat\imath]\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon\big(\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}\big)\,. \tag{5.6.63}$$

Proof. The bound (5.6.62) follows from (5.6.57), (5.5.35), (5.5.11), (5.5.8). Then use (5.6.62), (5.6.53)-(5.6.56), (5.6.35), (5.5.38) to prove (5.6.63). The biggest term comes from $\varepsilon(\bar p_1 - p_1)$. □

We now apply the transformation $\mathcal{T}$ defined in (5.6.64) whose goal is to remove the space average from the coefficient in front of ∂_y.

Consider the change of the space variable $z = y + p(\vartheta)$ which induces on $H^s_{S^\perp}(\mathbb{T}^{\nu+1})$ the operators

$$(\mathcal{T}w)(\vartheta, y) := w(\vartheta, y + p(\vartheta))\,, \quad (\mathcal{T}^{-1}h)(\vartheta, z) = h(\vartheta, z - p(\vartheta)) \tag{5.6.64}$$

(which are a particular case of those used in Section 5.6.1). The differential operator becomes $\mathcal{T}^{-1}\omega\cdot\partial_\vartheta\mathcal{T} = \omega\cdot\partial_\vartheta + \{\omega\cdot\partial_\vartheta p(\vartheta)\}\partial_z$, $\mathcal{T}^{-1}\partial_y\mathcal{T} = \partial_z$. Since $\mathcal{T}, \mathcal{T}^{-1}$ commute with $\Pi_S^\perp$, we get

$$\mathcal{L}_3 := \mathcal{T}^{-1}\mathcal{L}_2\mathcal{T} = \Pi_S^\perp\big(\omega\cdot\partial_\vartheta + m_3\partial_{zzz} + d_1\partial_z + d_0\big)\Pi_S^\perp + \mathfrak{R}_3\,, \tag{5.6.65}$$

$$d_1 := (\mathcal{T}^{-1}c_1) + \omega\cdot\partial_\vartheta p\,, \qquad d_0 := \mathcal{T}^{-1}c_0\,, \qquad \mathfrak{R}_3 := \mathcal{T}^{-1}\mathfrak{R}_2\mathcal{T}. \tag{5.6.66}$$

We choose

$$m_1 := \int_{\mathbb{T}^{\nu+1}} c_1 d\vartheta dy\,, \quad p := (\omega\cdot\partial_\vartheta)^{-1}\Big(m_1 - \int_{\mathbb{T}} c_1 dy\Big)\,, \tag{5.6.67}$$

so that $\int_{\mathbb{T}} d_1(\vartheta, z)\, dz = m_1$ for all $\vartheta \in \mathbb{T}^\nu$. Note that, by (5.6.59),

$$\int_{\mathbb{T}} c_1(\vartheta, y)\, dy = \int_{\mathbb{T}} c_{1,\geq 4}(\vartheta, y)\, dy\,, \quad \omega\cdot\partial_\vartheta p(\vartheta) = m_1 - \int_{\mathbb{T}} c_{1,\geq 4}(\vartheta, y)\, dy \tag{5.6.68}$$

because $\bar p_1, \bar p_2, q_{c_1}$ have all zero space-average. Also note that $\mathfrak{R}_3$ has the form (5.5.7). Since $\mathcal{T}$ is symplectic, the operator $\mathcal{L}_3$ in (5.6.65) is Hamiltonian.

Remark 5.6.6. *We require Hypothesis* (S1) *so that the function $q_{>2}$ has zero space average (see Lemma 5.5.5). If $q_{>2}$ did not have zero average, then p in* (5.6.67) *would have size $O(\varepsilon^3\gamma^{-1})$ (see* (5.5.31)*) and, since $\mathcal{T}^{-1} - I = O(\varepsilon^3\gamma^{-1})$, the function $\tilde d_0$ in* (5.6.71) *would satisfy $\tilde d_0 = O(\varepsilon^4\gamma^{-1})$. Therefore it would remain a term of order ∂_x^0 which is not perturbative for the reducibility scheme of Section 5.6.7.*

We put in evidence the terms of size $\varepsilon, \varepsilon^2$ in d_0, d_1, $\mathfrak{R}_3$. Recalling (5.6.66), (5.6.59), we split

$$d_1 = -\varepsilon \bar{p}_1 - \varepsilon^2 \bar{p}_2 + \tilde{d}_1 \,, \quad d_0 = -\varepsilon(\bar{p}_1)_x - \varepsilon^2(\bar{p}_2)_x + \tilde{d}_0 \,, \quad \mathfrak{R}_3 = -\varepsilon^2 \Pi_S^\perp \partial_x \bar{\mathcal{R}}_2 + \widetilde{\mathcal{R}}_* \tag{5.6.69}$$

where $\bar{\mathcal{R}}_2$ is obtained replacing v_δ with $\bar{v}$ in $\mathcal{R}_2$ (see (5.5.29)), and

$$\tilde{d}_1 := \varepsilon(\bar{p}_1 - \mathcal{T}^{-1}\bar{p}_1) + \varepsilon^2(\bar{p}_2 - \mathcal{T}^{-1}\bar{p}_2) + \mathcal{T}^{-1}(q_{c_1} + c_{1,\geq 4}) + \omega \cdot \partial_\vartheta p, \tag{5.6.70}$$

$$\tilde{d}_0 := \varepsilon(\bar{p}_1 - \mathcal{T}^{-1}\bar{p}_1)_x + \varepsilon^2(\bar{p}_2 - \mathcal{T}^{-1}\bar{p}_2)_x + \mathcal{T}^{-1}(q_{c_0} + c_{0,\geq 4}), \tag{5.6.71}$$

$$\widetilde{\mathcal{R}}_* := \mathcal{T}^{-1}\mathcal{R}_*\mathcal{T} + \varepsilon^2 \Pi_S^\perp \partial_x(\mathcal{R}_2 - \mathcal{T}^{-1}\mathcal{R}_2\mathcal{T}) + \varepsilon^2 \Pi_S^\perp \partial_x(\bar{\mathcal{R}}_2 - \mathcal{R}_2), \tag{5.6.72}$$

and $\mathcal{R}_*$ is defined in (5.6.49). We have also used that $\mathcal{T}^{-1}$ commutes with ∂_x and with $\Pi_S^\perp$.

Remark 5.6.7. *The space average* $\int_{\mathbb{T}} \tilde{d}_1(\vartheta, z)\, dz = \int_{\mathbb{T}} d_1(\vartheta, z)\, dz = m_1$ *for all* $\vartheta \in \mathbb{T}^\nu$.

Lemma 5.6.6. *There is* $\sigma := \sigma(\nu, \tau) > 0$ *(possibly larger than in Lemma 5.6.4) such that*

$$|m_1|^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^4, \qquad |\partial_i m_1[\hat{\imath}]| \leq C\varepsilon^{b+2}\|\hat{\imath}\|_{s_0+\sigma} \tag{5.6.73}$$

$$\|p\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4\gamma^{-1} + \varepsilon^{b+2}\gamma^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \quad \|\partial_i p[\hat{\imath}]\|_s \leq_s \varepsilon^{b+2}\gamma^{-1}\big(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}\big)\,, \tag{5.6.74}$$

$$\|\tilde{d}_k\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^5\gamma^{-1} + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \qquad \|\partial_i \tilde{d}_k[\hat{\imath}]\|_s \leq_s \varepsilon\big(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}\big) \tag{5.6.75}$$

for $k = 0, 1$. *Moreover the matrix* s*-decay norm (see* (3.1.1)*)*

$$|\widetilde{\mathcal{R}}_*|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^3 + \varepsilon^2\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}, \quad |\partial_i \widetilde{\mathcal{R}}_*[\hat{\imath}]|_s \leq_s \varepsilon^2\|\hat{\imath}\|_{s+\sigma} + \varepsilon^{2b-1}\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}\,. \tag{5.6.76}$$

The transformations $\mathcal{T}$, $\mathcal{T}^{-1}$ *satisfy* (5.6.39), (5.6.40).

Proof. The estimates (5.6.73), (5.6.74) follow by (5.6.67),(5.6.59),(5.6.68), and the bounds for $c_{1,\geq 4}, c_{0,\geq 4}$ in Lemma 5.6.4. The estimates (5.6.75) follow similarly by (5.6.63), (5.6.68), (5.6.74). The estimates (5.6.76) follow because $\mathcal{T}^{-1}\mathcal{R}_*\mathcal{T}$ satisfies the bounds (5.5.41) like $\mathcal{R}_*$ does (use Lemma 5.5.3 and (5.6.74)) and $|\varepsilon^2\Pi_S^\perp\partial_x(\bar{\mathcal{R}}_2 - \mathcal{R}_2)|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^2\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}$. □

It is sufficient to estimate $\widetilde{\mathcal{R}}_*$ (which has the form (5.5.7)) only in the s-decay norm (see (5.6.76)) because the next transformations will preserve it. Such norms are used in the reducibility scheme of Section 5.6.7.

5.6.4 Linear Birkhoff normal form. Step 1

Now we eliminate the terms of order ε and ε^2 of $\mathcal{L}_3$. This step is different from the reducibility steps that we shall perform in Section 5.6.7, because the diophantine constant $\gamma = o(\varepsilon^2)$ (see (5.3.4)) and so terms $O(\varepsilon), O(\varepsilon^2)$ are not perturbative. This reduction is possible thanks to the special form of the terms $\varepsilon\mathcal{B}_1$, $\varepsilon^2\mathcal{B}_2$ defined in (5.6.77): the harmonics of $\varepsilon\mathcal{B}_1$, and $\varepsilon^2 T$ in (5.6.93), which correspond to a possible small divisor are naught, see Corollary 5.6.1, and Lemma 5.6.11. In this Section we eliminate the term $\varepsilon\mathcal{B}_1$. In Section 5.6.5 we eliminate the terms of order ε^2.

Note that, since the previous transformations Φ, B, $\mathcal{T}$ are $O(\varepsilon^4\gamma^{-1})$-close to the identity, the terms of order ε and ε^2 in $\mathcal{L}_3$ are the same as in the original linearized operator.

We first collect all the terms of order ε and ε^2 in the operator $\mathcal{L}_3$ defined in (5.6.65). By (5.6.69), (5.5.29), (5.6.57) we have, renaming $\vartheta = \varphi$, $z = x$,

$$\mathcal{L}_3 = \Pi_S^\perp \big(\omega \cdot \partial_\varphi + m_3 \partial_{xxx} + \varepsilon \mathcal{B}_1 + \varepsilon^2 \mathcal{B}_2 + \tilde d_1 \partial_x + \tilde d_0\big) \Pi_S^\perp + \widetilde{\mathcal{R}}_*$$

where $\tilde d_1$, $\tilde d_0$, $\widetilde{\mathcal{R}}_*$ are defined in (5.6.70)-(5.6.72) and (recall also (4.1.50))

$$\mathcal{B}_1 h := -6 \partial_x (\bar v h), \quad \mathcal{B}_2 h := -6 \partial_x \{\bar v \Pi_S[(\partial_x^{-1} \bar v)\, \partial_x^{-1} h] + h \pi_0[(\partial_x^{-1} \bar v)^2]\} + 6 \pi_0 \{(\partial_x^{-1} \bar v) \Pi_S [\bar v h]\}. \tag{5.6.77}$$

Note that $\mathcal{B}_1$ and $\mathcal{B}_2$ are the linear Hamiltonian vector fields of $H_S^\perp$ generated, respectively, by the Hamiltonian $z \mapsto 3 \int_{\mathbb{T}} v z^2$ in (5.1.6), and the fourth order Birkhoff Hamiltonian $\mathcal{H}_{4,2}$ in (5.1.7) at $v = \bar v$.

We transform $\mathcal{L}_3$ by a symplectic operator $\Phi_1 : H^s_{S^\perp}(\mathbb{T}^{\nu+1}) \to H^s_{S^\perp}(\mathbb{T}^{\nu+1})$ of the form

$$\Phi_1 := \exp(\varepsilon A_1) = I_{H_S^\perp} + \varepsilon A_1 + \varepsilon^2 \frac{A_1^2}{2} + \varepsilon^3 \widehat A_1, \quad \widehat A_1 := \sum\nolimits_{k \geq 3} \frac{\varepsilon^{k-3}}{k!} A_1^k\,, \tag{5.6.78}$$

where $A_1(\varphi) h = \sum_{j, j' \in S^c} (A_1)_j^{j'}(\varphi) h_{j'} e^{\mathrm{i} j x}$ is a Hamiltonian vector field. The map Φ_1 is symplectic, because it is the time-1 flow of a Hamiltonian vector field. Therefore

$$\begin{aligned} & \mathcal{L}_3 \Phi_1 - \Phi_1 \Pi_S^\perp (\mathcal{D}_\omega + m_3 \partial_{xxx}) \Pi_S^\perp \qquad\qquad (5.6.79)\\ & = \Pi_S^\perp (\varepsilon \{\mathcal{D}_\omega A_1 + m_3 [\partial_{xxx}, A_1] + \mathcal{B}_1\} + \varepsilon^2 \{\mathcal{B}_1 A_1 + \mathcal{B}_2 + \frac{1}{2} m_3 [\partial_{xxx}, A_1^2] + \frac{1}{2} (\mathcal{D}_\omega A_1^2)\} + \tilde d_1 \partial_x + R_3) \Pi_S^\perp \end{aligned}$$

where

$$\begin{aligned} R_3 := {} & \tilde d_1 \partial_x (\Phi_1 - I) + \tilde d_0 \Phi_1 + \widetilde{\mathcal{R}}_* \Phi_1 + \varepsilon^2 \mathcal{B}_2 (\Phi_1 - I) \\ & + \varepsilon^3 \big\{\mathcal{D}_\omega \widehat A_1 + m_3 [\partial_{xxx}, \widehat A_1] + \frac{1}{2} \mathcal{B}_1 A_1^2 + \varepsilon \mathcal{B}_1 \widehat A_1\big\}\,. \end{aligned} \tag{5.6.80}$$

Remark 5.6.8. *R_3 has no longer the form (5.5.7). However $R_3 = O(\partial_x^0)$ because $A_1 = O(\partial_x^{-1})$ (see Lemma 5.6.9), and therefore $\Phi_1 - I_{H_S^\perp} = O(\partial_x^{-1})$. Moreover the matrix decay norm of R_3 is $o(\varepsilon^2)$.*

In order to eliminate the order ε from (5.6.79), we choose

$$(A_1)_j^{j'}(l) := \begin{cases} -\dfrac{(\mathcal{B}_1)_j^{j'}(l)}{\mathrm{i}(\omega \cdot l + m_3 (j'^3 - j^3))} & \text{if } \bar\omega \cdot l + j'^3 - j^3 \neq 0\,, \\ 0 & \text{otherwise,} \end{cases} \qquad j, j' \in S^c, \ l \in \mathbb{Z}^\nu. \tag{5.6.81}$$

This definition is well posed. Indeed, by (5.6.77) and (5.6.57),

$$(\mathcal{B}_1)_j^{j'}(l) := \begin{cases} -6 \mathrm{i} j \sqrt{\xi_{j-j'}} & \text{if } j - j' \in S\,, \quad l = \ell(j - j') \\ 0 & \text{otherwise.} \end{cases} \tag{5.6.82}$$

In particular $(\mathcal{B}_1)_j^{j'}(l) = 0$ unless $|l| \leq 1$. Thus, for $\bar\omega \cdot l + j'^3 - j^3 \neq 0$, the denominators in (5.6.81) satisfy

$$\begin{aligned} |\omega \cdot l + m_3 (j'^3 - j^3)| & = |m_3 (\bar\omega \cdot l + j'^3 - j^3) + (\omega - m_3 \bar\omega) \cdot l| \\ & \geq |m_3| |\bar\omega \cdot l + j'^3 - j^3| - |\omega - m_3 \bar\omega| |l| \geq 1/2\,, \quad \forall |l| \leq 1\,, \end{aligned} \tag{5.6.83}$$

for ε small, because the non zero integer $|\bar{\omega}\cdot l + j'^3 - j^3| \geq 1$, (5.6.50), and $\omega = \bar{\omega} + O(\varepsilon^2)$.

A_1 defined in (5.6.81) is a Hamiltonian vector field as $\mathcal{B}_1$.

Remark 5.6.9. *This is a general fact: the denominators $\delta_{l,j,k} := \mathrm{i}(\omega \cdot l + m_3(k^3 - j^3))$ satisfy $\overline{\delta_{l,j,k}} = \delta_{-l,k,j}$ and an operator $G(\varphi)$ is self-adjoint if and only if its matrix elements satisfy $\overline{G_j^k(l)} = G_k^j(-l)$. In a more intrinsic way, we could solve the homological equation of this Birkhoff step directly for the Hamiltonian function whose flow generates Φ_1.*

Lemma 5.6.7. *If $j, j' \in S^c$, $j - j' \in S$, $l = \ell(j - j')$, then $\bar{\omega}\cdot l + j'^3 - j^3 = 3jj'(j' - j) \neq 0$.*

Proof. We have $\bar{\omega}\cdot l = \bar{\omega}\cdot \ell(j - j') = (j - j')^3$ because $j - j' \in S$ (see (5.0.3) and (5.6.58)). Note that $j, j' \neq 0$ because $j, j' \in S^c$, and $j - j' \neq 0$ because $j - j' \in S$. ☐

Corollary 5.6.1. *Let $j, j' \in S^c$. If $\bar{\omega}\cdot l + j'^3 - j^3 = 0$ then $(\mathcal{B}_1)_j^{j'}(l) = 0$.*

Proof. If $(\mathcal{B}_1)_j^{j'}(l) \neq 0$ then $j - j' \in S, l = \ell(j - j')$ by (5.6.82). Hence $\bar{\omega}\cdot l + j'^3 - j^3 \neq 0$ by Lemma 5.6.7. ☐

By (5.6.81) and the previous corollary, the term of order ε in (5.6.79) is

$$\Pi_S^{\perp}\big(\mathcal{D}_\omega A_1 + m_3[\partial_{xxx}, A_1] + \mathcal{B}_1\big)\Pi_S^{\perp} = 0\,. \tag{5.6.84}$$

We now estimate the transformation A_1.

Lemma 5.6.8. (i) *For all $l \in \mathbb{Z}^\nu$, $j, j' \in S^c$,*

$$|(A_1)_j^{j'}(l)| \leq C(|j| + |j'|)^{-1}\,, \quad |(A_1)_j^{j'}(l)|^{\mathrm{lip}} \leq \varepsilon^{-2}(|j| + |j'|)^{-1}\,. \tag{5.6.85}$$

(ii) $(A_1)_j^{j'}(l) = 0$ *for all $l \in \mathbb{Z}^\nu$, $j, j' \in S^c$ such that $|j - j'| > C_S$, where $C_S := \max\{|j| : j \in S\}$.*

Proof. (i) We already noted that $(A_1)_j^{j'}(l) = 0$, $\forall |l| > 1$. Since $|\omega| \leq |\bar{\omega}| + 1$, one has, for $|l| \leq 1$, $j \neq j'$,

$$|\omega\cdot l + m_3(j'^3 - j^3)| \geq |m_3||j'^3 - j^3| - |\omega \cdot l| \geq \frac{1}{4}(j'^2 + j^2) - |\omega| \geq \frac{1}{8}(j'^2 + j^2)\,, \quad \forall (j'^2 + j^2) \geq C,$$

for some constant $C > 0$. Moreover, recalling that also (5.6.83) holds, we deduce that for $j \neq j'$,

$$(A_1)_j^{j'}(l) \neq 0 \quad \Rightarrow \quad |\omega\cdot l + m_3(j'^3 - j^3)| \geq c(|j| + |j'|)^2\,. \tag{5.6.86}$$

On the other hand, if $j = j'$, $j \in S^c$, the matrix $(A_1)_j^j(l) = 0$, $\forall l \in \mathbb{Z}^\nu$, because $(\mathcal{B}_1)_j^j(l) = 0$ by (5.6.82) (recall that $0 \notin S$). Hence (5.6.86) holds for all j, j'. By (5.6.81), (5.6.86), (5.6.82) we deduce the first bound in (5.6.85). The Lipschitz bound follows similarly (use also $|j - j'| \leq C_S$). (ii) follows by (5.6.81)-(5.6.82). ☐

The previous lemma means that $A = O(|\partial_x|^{-1})$. More precisely we deduce that

Lemma 5.6.9. *$|A_1\partial_x|_s^{\mathrm{Lip}(\gamma)} + |\partial_x A_1|_s^{\mathrm{Lip}(\gamma)} \leq C(s)$.*

Proof. Recalling the definition of the (space-time) matrix norm in (3.1.15), since $(A_1)_{j_1}^{j_2}(l) = 0$ outside the set of indices $|l| \le 1, |j_1 - j_2| \le C_S$, we have

$$|\partial_x A_1|_s^2 = \sum_{|l|\le 1,\, |j|\le C_S} \Big(\sup_{j_1 - j_2 = j} |j_1| |(A_1)_{j_1}^{j_2}(l)| \Big)^2 \langle l, j\rangle^{2s} \le C(s)$$

by Lemma 5.6.8. The estimates for $|A_1 \partial_x|_s$ and the Lipschitz bounds follow similarly. $\square$

It follows that the symplectic map Φ_1 in (5.6.78) is invertible for ε small, with inverse

$$\Phi_1^{-1} = \exp(-\varepsilon A_1) = I_{H_S^\perp} + \varepsilon \check{A}_1 \,, \ \check{A}_1 := \sum\nolimits_{n\ge 1} \frac{\varepsilon^{n-1}}{n!} (-A_1)^n \,,$$
$$|\check{A}_1 \partial_x|_s^{\mathrm{Lip}(\gamma)} + |\partial_x \check{A}_1|_s^{\mathrm{Lip}(\gamma)} \le C(s) \,. \tag{5.6.87}$$

Since A_1 solves the homological equation (5.6.84), the ε-term in (5.6.79) is zero, and, with a straightforward calculation, the ε^2-term simplifies to $\mathcal{B}_2 + \frac{1}{2}[\mathcal{B}_1, A_1]$. We obtain the Hamiltonian operator

$$\mathcal{L}_4 := \Phi_1^{-1} \mathcal{L}_3 \Phi_1 = \Pi_S^\perp (\mathcal{D}_\omega + m_3 \partial_{xxx} + \tilde{d}_1 \partial_x + \varepsilon^2 \{\mathcal{B}_2 + \tfrac{1}{2}[\mathcal{B}_1, A_1]\} + \tilde{R}_4) \Pi_S^\perp \tag{5.6.88}$$
$$\tilde{R}_4 := (\Phi_1^{-1} - I) \Pi_S^\perp [\varepsilon^2 (\mathcal{B}_2 + \tfrac{1}{2}[\mathcal{B}_1, A_1]) + \tilde{d}_1 \partial_x] + \Phi_1^{-1} \Pi_S^\perp R_3 \,. \tag{5.6.89}$$

We split A_1 defined in (5.6.81), (5.6.82) into $A_1 = \bar{A}_1 + \widetilde{A}_1$ where, for all $j, j' \in S^c$, $l \in \mathbb{Z}^\nu$,

$$(\bar{A}_1)_j^{j'}(l) := \frac{6j\sqrt{\xi_{j-j'}}}{\bar{\omega} \cdot l + j'^3 - j^3} \qquad \text{if } \ \bar{\omega} \cdot l + j'^3 - j^3 \ne 0, \ \ j - j' \in S, \ \ l = \ell(j - j'), \tag{5.6.90}$$

and $(\bar{A}_1)_j^{j'}(l) := 0$ otherwise. By Lemma 5.6.7, for all $j, j' \in S^c$, $l \in \mathbb{Z}^\nu$, $(\bar{A}_1)_j^{j'}(l) = \frac{2\sqrt{\xi_{j-j'}}}{j'(j'-j)}$ if $j - j' \in S$, $l = \ell(j - j')$, and $(\bar{A}_1)_j^{j'}(l) = 0$ otherwise, namely (recall the definition of $\bar{v}$ in (5.6.57))

$$\bar{A}_1 h = 2\Pi_S^\perp [(\partial_x^{-1} \bar{v})(\partial_x^{-1} h)] \,, \quad \forall h \in H^s_{S^\perp}(\mathbb{T}^{\nu+1}) \,. \tag{5.6.91}$$

The difference is

$$(\widetilde{A}_1)_j^{j'}(l) = (A_1 - \bar{A}_1)_j^{j'}(l) = -\frac{6j\sqrt{\xi_{j-j'}} \{(\omega - \bar{\omega}) \cdot l + (m_3 - 1)(j'^3 - j^3)\}}{\big(\omega \cdot l + m_3(j'^3 - j^3)\big)\big(\bar{\omega} \cdot l + j'^3 - j^3\big)} \tag{5.6.92}$$

for $j, j' \in S^c$, $j - j' \in S$, $l = \ell(j - j')$, and $(\widetilde{A}_1)_j^{j'}(l) = 0$ otherwise. Then, by (5.6.88),

$$\mathcal{L}_4 = \Pi_S^\perp \big(\mathcal{D}_\omega + m_3 \partial_{xxx} + \tilde{d}_1 \partial_x + \varepsilon^2 T + R_4 \big) \Pi_S^\perp \,, \tag{5.6.93}$$

where

$$T := \mathcal{B}_2 + \frac{1}{2}[\mathcal{B}_1, \bar{A}_1] \,, \qquad R_4 := \frac{\varepsilon^2}{2}[\mathcal{B}_1, \widetilde{A}_1] + \tilde{R}_4 \,. \tag{5.6.94}$$

The operator T is Hamiltonian as $\mathcal{B}_2$, $\mathcal{B}_1$, $\bar{A}_1$ (the commutator of two Hamiltonian vector fields is Hamiltonian).

Lemma 5.6.10. *There is $\sigma = \sigma(\nu, \tau) > 0$ (possibly larger than in Lemma 5.6.6) such that*

$$|R_4|_s^{\mathrm{Lip}(\gamma)} \le_s \varepsilon^5 \gamma^{-1} + \varepsilon \|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)} \,, \quad |\partial_i R_4[\widehat{\imath}]|_s \le_s \varepsilon \big(\|\widehat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma} \|\widehat{\imath}\|_{s_0+\sigma} \big) \,. \tag{5.6.95}$$

Proof. We first estimate $[\mathcal{B}_1, \widetilde{A}_1] = (\mathcal{B}_1\partial_x^{-1})(\partial_x\widetilde{A}_1) - (\widetilde{A}_1\partial_x)(\partial_x^{-1}\mathcal{B}_1)$. By (5.6.92), $|\omega - \bar\omega| \leq C\varepsilon^2$ (as $\omega \in \Omega_\varepsilon$ in (5.3.2)) and (5.6.50), arguing as in Lemmata 5.6.8, 5.6.9, we deduce that $|\widetilde{A}_1\partial_x|_s^{\mathrm{Lip}(\gamma)} + |\partial_x\widetilde{A}_1|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^2$. By (5.6.77) the norm $|\mathcal{B}_1\partial_x^{-1}|_s^{\mathrm{Lip}(\gamma)} + |\partial_x^{-1}\mathcal{B}_1|^{\mathrm{Lip}(\gamma)} \leq C(s)$. Hence $\varepsilon^2|[\mathcal{B}_1, \widetilde{A}_1]|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^4$. Finally (5.6.94), (5.6.89), (5.6.87), (5.6.80), (5.6.75), (5.6.76), and the interpolation estimate (3.1.9) imply (5.6.95). □

5.6.5 Linear Birkhoff normal form. Step 2

The goal of this Section is to remove the term $\varepsilon^2 T$ from the operator $\mathcal{L}_4$ defined in (5.6.93). We conjugate the Hamiltonian operator $\mathcal{L}_4$ via a symplectic map

$$\Phi_2 := \exp(\varepsilon^2 A_2) = I_{H_S^\perp} + \varepsilon^2 A_2 + \varepsilon^4 \widehat{A}_2\,, \quad \widehat{A}_2 := \sum\nolimits_{k\geq 2} \frac{\varepsilon^{2(k-2)}}{k!} A_2^k \tag{5.6.96}$$

where $A_2(\varphi) = \sum_{j,j'\in S^c} (A_2)_j^{j'}(\varphi) h_{j'} e^{\mathrm{i}jx}$ is a Hamiltonian vector field. We compute

$$\mathcal{L}_4\Phi_2 - \Phi_2\Pi_S^\perp\big(\mathcal{D}_\omega + m_3\partial_{xxx}\big)\Pi_S^\perp = \Pi_S^\perp(\varepsilon^2\{\mathcal{D}_\omega A_2 + m_3[\partial_{xxx}, A_2] + T\} + \tilde d_1\partial_x + \tilde R_5)\Pi_S^\perp\,, \tag{5.6.97}$$

$$\tilde R_5 := \Pi_S^\perp\{\varepsilon^4((\mathcal{D}_\omega \widehat{A}_2) + m_3[\partial_{xxx}, \widehat{A}_2]) + (\tilde d_1\partial_x + \varepsilon^2 T)(\Phi_2 - I) + R_4\Phi_2\}\Pi_S^\perp\,. \tag{5.6.98}$$

We define

$$(A_2)_j^{j'}(l) := -\frac{T_j^{j'}(l)}{\mathrm{i}(\omega\cdot l + m_3(j'^3 - j^3))} \quad \text{if } \bar\omega\cdot l + j'^3 - j^3 \neq 0; \qquad (A_2)_j^{j'}(l) := 0 \quad \text{otherwise.} \tag{5.6.99}$$

This definition is well posed. Indeed, by (5.6.94), (5.6.82), (5.6.90), (5.6.77), the matrix entries $T_j^{j'}(l) = 0$ for all $|j - j'| > 2C_S$, $l \in \mathbb{Z}^\nu$, where $C_S := \max\{|j|, j \in S\}$. Also $T_j^{j'}(l) = 0$ for all $j, j' \in S^c$, $|l| > 2$ (see also (5.6.100), (5.6.103), (5.6.104) below). Thus, arguing as in (5.6.83), if $\bar\omega \cdot l + j'^3 - j^3 \neq 0$, then $|\omega \cdot l + m_3(j'^3 - j^3)| \geq 1/2$. The operator A_2 is a Hamiltonian vector field because T is Hamiltonian and by Remark 5.6.9.

Now we prove that the Birkhoff map Φ_2 removes completely the term $\varepsilon^2 T$.

Lemma 5.6.11. *Let $j, j' \in S^c$. If $\bar\omega \cdot l + j'^3 - j^3 = 0$, then $T_j^{j'}(l) = 0$.*

Proof. By (5.6.77), (5.6.91) we get $\mathcal{B}_1\bar A_1 h = -12\partial_x\{\bar v\Pi_S^\perp[(\partial_x^{-1}\bar v)(\partial_x^{-1}h)]\}$,

$$\bar A_1\mathcal{B}_1 h = -12\Pi_S^\perp[(\partial_x^{-1}\bar v)\Pi_S^\perp(\bar v h)]\,, \quad \forall h \in H^s_{S^\perp}\,,$$

whence, recalling (5.6.57), for all $j, j' \in S^c$, $l \in \mathbb{Z}^\nu$,

$$([\mathcal{B}_1, \bar A_1])_j^{j'}(l) = 12\mathrm{i} \sum_{\substack{j_1,j_2\in S,\, j_1+j_2=j-j' \\ j'+j_2\in S^c,\, \ell(j_1)+\ell(j_2)=l}} \frac{jj_1 - j'j_2}{j'j_1j_2}\sqrt{\xi_{j_1}\xi_{j_2}}\,, \tag{5.6.100}$$

If $([\mathcal{B}_1, \bar A_1])_j^{j'}(l) \neq 0$ there are $j_1, j_2 \in S$ such that $j_1 + j_2 = j - j'$, $j' + j_2 \in S^c$, $\ell(j_1) + \ell(j_2) = l$. Then

$$\bar\omega \cdot l + j'^3 - j^3 = \bar\omega\cdot\ell(j_1) + \bar\omega\cdot\ell(j_2) + j'^3 - j^3 \overset{(5.6.58)}{=} j_1^3 + j_2^3 + j'^3 - j^3\,. \tag{5.6.101}$$

Thus, if $\bar{\omega}\cdot l + j'^3 - j^3 = 0$, Lemma 5.1.2 implies $(j_1+j_2)(j_1+j')(j_2+j') = 0$. Now j_1+j', $j_2+j' \neq 0$ because $j_1, j_2 \in S$, $j' \in S^c$ and S is symmetric. Hence $j_1 + j_2 = 0$, which implies $j = j'$ and $l = 0$ (the map ℓ in (5.6.58) is odd). In conclusion, if $\bar{\omega}\cdot l + j'^3 - j^3 = 0$, the only nonzero matrix entry $([\mathcal{B}_1, \bar{A}_1])_j^{j'}(l)$ is

$$([\mathcal{B}_1, \bar{A}_1])_j^{j}(0) \overset{(5.6.100)}{=} 24\mathrm{i} \sum_{j_2 \in S,\, j_2+j \in S^c} \xi_{j_2} j_2^{-1} . \tag{5.6.102}$$

Now we consider $\mathcal{B}_2$ in (5.6.77). Split $\mathcal{B}_2 = B_1 + B_2 + B_3$, where $B_1 h := -6\partial_x\{\bar{v}\Pi_S[(\partial_x^{-1}\bar{v})\partial_x^{-1}h]\}$, $B_2 h := -6\partial_x\{h\pi_0[(\partial_x^{-1}\bar{v})^2]\}$, $B_3 h := 6\pi_0\{\Pi_S(\bar{v}h)\partial_x^{-1}\bar{v}\}$. Their Fourier matrix representation is

$$(B_1)_j^{j'}(l) = 6\mathrm{i}j \sum_{\substack{j_1,j_2\in S,\, j_1+j'\in S \\ j_1+j_2=j-j',\, \ell(j_1)+\ell(j_2)=l}} \frac{\sqrt{\xi_{j_1}\xi_{j_2}}}{j_1 j'} , \qquad (B_2)_j^{j'}(l) = 6\mathrm{i}j \sum_{\substack{j_1,j_2\in S,\, j_1+j_2\neq 0 \\ j_1+j_2=j-j',\, \ell(j_1)+\ell(j_2)=l}} \frac{\sqrt{\xi_{j_1}\xi_{j_2}}}{j_1 j_2} , \tag{5.6.103}$$

$$(B_3)_j^{j'}(l) = 6 \sum_{\substack{j_1,j_2\in S,\, j_1+j'\in S \\ j_1+j_2=j-j',\, \ell(j_1)+\ell(j_2)=l}} \frac{\sqrt{\xi_{j_1}\xi_{j_2}}}{\mathrm{i}j_2} , \qquad j, j' \in S^c,\ l \in \mathbb{Z}^\nu . \tag{5.6.104}$$

We study the terms B_1, B_2, B_3 separately. If $(B_1)_j^{j'}(l) \neq 0$, there are $j_1, j_2 \in S$ such that $j_1 + j_2 = j - j'$, $j_1 + j' \in S$, $l = \ell(j_1) + \ell(j_2)$ and (5.6.101) holds. Thus, if $\bar{\omega}\cdot l + j'^3 - j^3 = 0$, Lemma 5.1.2 implies $(j_1+j_2)(j_1+j')(j_2+j') = 0$, and, since $j' \in S^c$ and S is symmetric, the only possibility is $j_1 + j_2 = 0$. Hence $j = j'$, $l = 0$. In conclusion, if $\bar{\omega}\cdot l + j'^3 - j^3 = 0$, the only nonzero matrix element $(B_1)_j^{j'}(l)$ is

$$(B_1)_j^{j}(0) = 6\mathrm{i} \sum_{j_1 \in S,\, j_1+j \in S} \xi_{j_1} j_1^{-1} . \tag{5.6.105}$$

By the same arguments, if $(B_2)_j^{j'}(l) \neq 0$ and $\bar{\omega}\cdot l + j'^3 - j^3 = 0$ we find $(j_1+j_2)(j_1+j')(j_2+j') = 0$, which is impossible because also $j_1 + j_2 \neq 0$. Finally, arguing as for B_1, if $\bar{\omega}\cdot l + j'^3 - j^3 = 0$, then the only nonzero matrix element $(B_3)_j^{j'}(l)$ is

$$(B_3)_j^{j}(0) = 6\mathrm{i} \sum_{j_1 \in S,\, j_1+j \in S} \xi_{j_1} j_1^{-1} . \tag{5.6.106}$$

From (5.6.102), (5.6.105), (5.6.106) we deduce that, if $\bar{\omega}\cdot l + j'^3 - j^3 = 0$, then the only non zero elements $(\frac{1}{2}[\mathcal{B}_1, \bar{A}_1] + B_1 + B_3)_j^{j'}(l)$ must be for $(l, j, j') = (0, j, j)$. In this case, we get

$$\frac{1}{2}([\mathcal{B}_1, \bar{A}_1])_j^{j}(0) + (B_1)_j^{j}(0) + (B_3)_j^{j}(0) = 12\mathrm{i} \sum_{\substack{j_1\in S \\ j_1+j\in S^c}} \frac{\xi_{j_1}}{j_1} + 12\mathrm{i} \sum_{\substack{j_1\in S \\ j_1+j\in S}} \frac{\xi_{j_1}}{j_1} = 12\mathrm{i} \sum_{j_1 \in S} \frac{\xi_{j_1}}{j_1} = 0 \tag{5.6.107}$$

because the case $j_1 + j = 0$ is impossible ($j_1 \in S$, $j' \in S^c$ and S is symmetric), and the function $S \ni j_1 \to \xi_{j_1}/j_1 \in \mathbb{R}$ is odd. The lemma follows by (5.6.94), (5.6.107). $\square$

The choice of A_2 in (5.6.99) and Lemma 5.6.11 imply that

$$\Pi_S^\perp\big(\mathcal{D}_\omega A_2 + m_3[\partial_{xxx}, A_2] + T\big)\Pi_S^\perp = 0 . \tag{5.6.108}$$

Lemma 5.6.12. $|\partial_x A_2|_s^{\mathrm{Lip}(\gamma)} + |A_2\partial_x|_s^{\mathrm{Lip}(\gamma)} \leq C(s)$.

Proof. First we prove that the diagonal elements $T_j^j(l) = 0$ for all $l \in \mathbb{Z}^\nu$. For $l = 0$, we have already proved that $T_j^j(0) = 0$ (apply Lemma 5.6.11 with $j = j'$, $l = 0$). Moreover, in each term $[\mathcal{B}_1, \bar{A}_1]$, B_1, B_2, B_3 (see (5.6.100), (5.6.103), (5.6.104)) the sum is over $j_1 + j_2 = j - j'$, $l = \ell(j_1) + \ell(j_2)$. If $j = j'$, then $j_1 + j_2 = 0$, and $l = 0$. Thus $T_j^j(l) = T_j^j(0) = 0$. For the off-diagonal terms $j \neq j'$ we argue as in Lemmata 5.6.8, 5.6.9, using that all the denominators $|\omega \cdot l + m_3(j'^3 - j^3)| \geq c(|j| + |j'|)^2$. $\square$

For ε small, the map Φ_2 in (5.6.96) is invertible and $\Phi_2 = \exp(-\varepsilon^2 A_2)$. Therefore (5.6.97), (5.6.108) imply

$$\mathcal{L}_5 := \Phi_2^{-1}\mathcal{L}_4\Phi_2 = \Pi_S^\perp(\mathcal{D}_\omega + m_3\partal_{xxx} + \tilde{d}_1\partial_x + R_5)\Pi_S^\perp\,, \tag{5.6.109}$$

$$R_5 := (\Phi_2^{-1} - I)\Pi_S^\perp\tilde{d}_1\partial_x + \Phi_2^{-1}\Pi_S^\perp\tilde{R}_5\,. \tag{5.6.110}$$

Since A_2 is a Hamiltonian vector field, the map Φ_2 is symplectic and so $\mathcal{L}_5$ is Hamiltonian.

Lemma 5.6.13. *R_5 satisfies the same estimates* (5.6.95) *as R_4 (with a possibly larger σ).*

Proof. Use (5.6.110), Lemma 5.6.12, (5.6.75), (5.6.98), (5.6.95) and the interpolation inequalities (3.1.5), (3.1.9). $\square$

5.6.6 Descent method

The goal of this Section is to transform $\mathcal{L}_5$ in (5.6.109) so that the coefficient of ∂_x becomes constant. We conjugate $\mathcal{L}_5$ via a symplectic map of the form

$$\mathcal{S} := \exp(\Pi_S^\perp(w\partial_x^{-1}))\Pi_S^\perp = \Pi_S^\perp\big(I + w\partial_x^{-1}\big)\Pi_S^\perp + \widehat{\mathcal{S}}\,, \quad \widehat{\mathcal{S}} := \sum\nolimits_{k\geq 2}\frac{1}{k!}[\Pi_S^\perp(w\partial_x^{-1})]^k\Pi_S^\perp\,, \tag{5.6.111}$$

where $w : \mathbb{T}^{\nu+1} \to \mathbb{R}$ is a function. Note that $\Pi_S^\perp(w\partial_x^{-1})\Pi_S^\perp$ is the Hamiltonian vector field generated by $-\frac{1}{2}\int_{\mathbb{T}} w(\partial_x^{-1}h)^2\,dx$, $h \in H_S^\perp$. Recalling (4.1.50), we calculate

$$\mathcal{L}_5\mathcal{S} - \mathcal{S}\Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx} + m_1\partial_x)\Pi_S^\perp = \Pi_S^\perp(3m_3w_x + \tilde{d}_1 - m_1)\partial_x\Pi_S^\perp + \tilde{R}_6\,, \tag{5.6.112}$$

$$\begin{aligned}\tilde{R}_6 := \Pi_S^\perp\{&(3m_3w_{xx} + \tilde{d}_1\Pi_S^\perp w - m_1 w)\pi_0 + ((\mathcal{D}_\omega w) + m_3w_{xxx} + \tilde{d}_1\Pi_S^\perp w_x)\partial_x^{-1} + (\mathcal{D}_\omega\widehat{\mathcal{S}})\\ &+ m_3[\partial_{xxx}, \widehat{\mathcal{S}}] + \tilde{d}_1\partial_x\widehat{\mathcal{S}} - m_1\widehat{\mathcal{S}}\partial_x + R_5\mathcal{S}\}\Pi_S^\perp\end{aligned}$$

where $\tilde{R}_6$ collects all the terms of order at most ∂_x^0. By Remark 5.6.7, we solve $3m_3w_x + \tilde{d}_1 - m_1 = 0$ by choosing $w := -(3m_3)^{-1}\partial_x^{-1}(\tilde{d}_1 - m_1)$. For ε small, the operator $\mathcal{S}$ is invertible and, by (5.6.112),

$$\mathcal{L}_6 := \mathcal{S}^{-1}\mathcal{L}_5\mathcal{S} = \Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx} + m_1\partial_x)\Pi_S^\perp + R_6\,, \qquad R_6 := \mathcal{S}^{-1}\tilde{R}_6\,. \tag{5.6.113}$$

Since $\mathcal{S}$ is symplectic, $\mathcal{L}_6$ is Hamiltonian (recall Definition 3.3.1).

Lemma 5.6.14. *There is $\sigma = \sigma(\nu, \tau) > 0$ (possibly larger than in Lemma 5.6.13) such that*

$$|\mathcal{S}^{\pm 1} - I|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^5\gamma^{-1} + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\,, \quad |\partial_i\mathcal{S}^{\pm 1}[\hat{\imath}]|_s \leq_s \varepsilon(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}).$$

The remainder R_6 satisfies the same estimates (5.6.95) *as R_4.*

Proof. By (5.6.75),(5.6.73),(5.6.50), $\|w\|_s^{\mathrm{Lip}(\gamma)} \leq_s \varepsilon^5\gamma^{-1} + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}$, and the lemma follows by (5.6.111). Since $\widehat{\mathcal{S}} = O(\partial_x^{-2})$ the commutator $[\partial_{xxx}, \widehat{\mathcal{S}}] = O(\partial_x^0)$ and $|[\partial_{xxx}, \widehat{\mathcal{S}}]|_s^{\mathrm{Lip}(\gamma)} \leq_s \|w\|_{s_0+3}^{\mathrm{Lip}(\gamma)}\|w\|_{s+3}^{\mathrm{Lip}(\gamma)}$. $\square$

5.6.7 KAM reducibility and inversion of $\mathcal{L}_\omega$

The coefficients m_3, m_1 of the operator $\mathcal{L}_6$ in (5.6.113) are constants, and the remainder R_6 is a bounded operator of order ∂_x^0 with small matrix decay norm, see (5.6.116). Then we can diagonalize $\mathcal{L}_6$ by applying the iterative KAM reducibility Theorem 4.2.2 along the sequence of scales

$$N_n := N_0^{\chi^n}, \quad n = 0, 1, 2, \ldots, \quad \chi := 3/2, \quad N_0 > 0 . \tag{5.6.114}$$

In Section 5.7, the initial N_0 will (slightly) increase to infinity as $\varepsilon \to 0$, see (5.7.5). The required smallness condition (4.2.14) (written in the present notations) is

$$N_0^{C_0} |R_6|_{s_0+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1} \leq 1 \tag{5.6.115}$$

where $\beta := 7\tau + 6$ (see (4.2.1)), τ is the diophantine exponent in (5.3.4) and (5.6.120), and the constant $C_0 := C_0(\tau, \nu) > 0$ is fixed in Theorem 4.2.2. By Lemma 5.6.14, the remainder R_6 satisfies the bound (5.6.95), and using (5.5.8) we get (recall (5.3.10))

$$|R_6|_{s_0+\beta}^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^{7-2b}\gamma^{-1} = C\varepsilon^{3-2a}, \qquad |R_6|_{s_0+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1} \leq C\varepsilon^{1-3a} . \tag{5.6.116}$$

We use that μ in (5.5.8) is assumed to satisfy $\mu \geq \sigma + \beta$ where $\sigma := \sigma(\tau, \nu)$ is given in Lemma 5.6.14.

Theorem 5.6.1. (Reducibility) *Assume that $\omega \mapsto i_\delta(\omega)$ is a Lipschitz function defined on some subset $\Omega_o \subset \Omega_\varepsilon$ (recall* (5.3.2)*), satisfying* (5.5.8) *with $\mu \geq \sigma + \beta$ where $\sigma := \sigma(\tau, \nu)$ is given in Lemma 5.6.14 and $\beta := 7\tau + 6$. Then there exists $\delta_0 \in (0, 1)$ such that, if*

$$N_0^{C_0} \varepsilon^{7-2b} \gamma^{-2} = N_0^{C_0} \varepsilon^{1-3a} \leq \delta_0 , \quad \gamma := \varepsilon^{2+a} , \quad a \in (0, 1/6) , \tag{5.6.117}$$

then:

(i) **(Eigenvalues)**. *For all $\omega \in \Omega_\varepsilon$ there exists a sequence*

$$\mu_j^\infty(\omega) := \mu_j^\infty(\omega, i_\delta(\omega)) := \mathrm{i}\big(- \tilde{m}_3(\omega) j^3 + \tilde{m}_1(\omega) j\big) + r_j^\infty(\omega), \quad j \in S^c , \tag{5.6.118}$$

where $\tilde{m}_3, \tilde{m}_1$ coincide with the coefficients m_3, m_1 of $\mathcal{L}_6$ in (5.6.113) *for all $\omega \in \Omega_o$, and*

$$|\tilde{m}_3 - 1|^{\mathrm{Lip}(\gamma)} + |\tilde{m}_1|^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^4 , \quad |r_j^\infty|^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^{3-2a} , \quad \forall j \in S^c , \tag{5.6.119}$$

for some $C > 0$. All the eigenvalues μ_j^∞ are purely imaginary. We define, for convenience, $\mu_0^\infty(\omega) := 0$.

(ii) **(Conjugacy)**. *For all ω in the set*

$$\Omega_\infty^{2\gamma} := \Omega_\infty^{2\gamma}(i_\delta) := \Big\{ \omega \in \Omega_o : \, |\mathrm{i}\omega \cdot l + \mu_j^\infty(\omega) - \mu_k^\infty(\omega)| \geq \frac{2\gamma |j^3 - k^3|}{\langle l \rangle^\tau} , \\ \forall l \in \mathbb{Z}^\nu, \, j, k \in S^c \cup \{0\} \Big\} \tag{5.6.120}$$

there is a real, bounded, invertible linear operator $\Phi_\infty(\omega) : H^s_{S^\perp}(\mathbb{T}^{\nu+1}) \to H^s_{S^\perp}(\mathbb{T}^{\nu+1})$, with bounded inverse $\Phi_\infty^{-1}(\omega)$, that conjugates $\mathcal{L}_6$ in (5.6.113) *to constant coefficients, namely*

$$\mathcal{L}_\infty(\omega) := \Phi_\infty^{-1}(\omega) \circ \mathcal{L}_6(\omega) \circ \Phi_\infty(\omega) = \omega \cdot \partial_\varphi + \mathcal{D}_\infty(\omega), \quad \mathcal{D}_\infty(\omega) := \mathrm{diag}_{j \in S^c} \{\mu_j^\infty(\omega)\} . \tag{5.6.121}$$

The transformations $\Phi_\infty, \Phi_\infty^{-1}$ are close to the identity in matrix decay norm, with

$$|\Phi_\infty - I|^{\mathrm{Lip}(\gamma)}_{s,\Omega^{2\gamma}_\infty} + |\Phi_\infty^{-1} - I|^{\mathrm{Lip}(\gamma)}_{s,\Omega^{2\gamma}_\infty} \leq_s \varepsilon^5\gamma^{-2} + \varepsilon\gamma^{-1}\|\mathfrak{I}_\delta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma}. \tag{5.6.122}$$

Moreover $\Phi_\infty, \Phi_\infty^{-1}$ are symplectic, and $\mathcal{L}_\infty$ is a Hamiltonian operator.

Proof. The proof is the same as the one of Theorem 4.2.1, which is based on Theorem 4.2.2, Corollaries 4.2.1, 4.2.34, and Lemmata 4.2.1, 4.2.2. A difference is that here $\omega \in \mathbb{R}^\nu$, for the forced Airy equation, the parameter $\lambda \in \mathbb{R}$ is one-dimensional. The proof is the same because Kirszbraun's Theorem on Lipschitz extension of functions also holds in $\mathbb{R}^\nu$ (see, e.g., Lemma A.2 in [65]). The bound (5.6.122) follows by Corollary 4.2.1 and the estimate of R_6 in Lemma 5.6.14. We also use the estimates (5.6.50), (5.6.73) for $\partial_i m_3$, $\partial_i m_1$. Another difference is that here the sites $j \in S^c \subset \mathbb{Z} \setminus \{0\}$ unlike in Chapter 4 where $j \in \mathbb{Z}$. We have defined $\mu_0^\infty := 0$ so that also the first Melnikov conditions (5.6.123) are included in the definition of $\Omega^{2\gamma}_\infty$. □

Remark 5.6.10. *Theorem 4.2.2 also provides the Lipschitz dependence of the (approximate) eigenvalues μ_j^n with respect to the unknown $i_0(\varphi)$, which is used for the measure estimate Lemma 5.7.2.*

All the parameters $\omega \in \Omega^{2\gamma}_\infty$ satisfy (specialize (5.6.120) for $k = 0$)

$$|\mathrm{i}\omega \cdot l + \mu_j^\infty(\omega)| \geq 2\gamma |j|^3 \langle l \rangle^{-\tau}, \quad \forall l \in \mathbb{Z}^\nu, \ j \in S^c, \tag{5.6.123}$$

and the diagonal operator $\mathcal{L}_\infty$ is invertible.

In the following theorem we finally verify the inversion assumption (5.4.46) for $\mathcal{L}_\omega$.

Theorem 5.6.2. (Inversion of $\mathcal{L}_\omega$) *Assume the hypotheses of Theorem 5.6.1 and* (5.6.117). *Then there exists $\sigma_1 := \sigma_1(\tau, \nu) > 0$ such that, $\forall \omega \in \Omega^{2\gamma}_\infty(i_\delta)$ (see* (5.6.120)*), for any function $g \in H^{s+\sigma_1}_{S^\perp}(\mathbb{T}^{\nu+1})$ the equation $\mathcal{L}_\omega h = g$ has a solution $h = \mathcal{L}_\omega^{-1} g \in H^s_{S^\perp}(\mathbb{T}^{\nu+1})$, satisfying*

$$\begin{aligned}\|\mathcal{L}_\omega^{-1} g\|^{\mathrm{Lip}(\gamma)}_s &\leq_s \gamma^{-1}\big(\|g\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_1} + \varepsilon\gamma^{-1}\|\mathfrak{I}_\delta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_1}\|g\|^{\mathrm{Lip}(\gamma)}_{s_0}\big) \\ &\leq_s \gamma^{-1}\big(\|g\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_1} + \varepsilon\gamma^{-1}\big\{\|\mathfrak{I}_0\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_1+\sigma} + \gamma^{-1}\|\mathfrak{I}_0\|^{\mathrm{Lip}(\gamma)}_{s_0+\sigma}\|Z\|^{\mathrm{Lip}(\gamma)}_{s+\sigma_1+\sigma}\big\}\|g\|^{\mathrm{Lip}(\gamma)}_{s_0}\big).\end{aligned} \tag{5.6.124}$$

Proof. Collecting Theorem 5.6.1 with the results of Sections 5.6.1-5.6.6, we have obtained the (semi)-conjugation of the operator $\mathcal{L}_\omega$ (defined in (5.5.34)) to $\mathcal{L}_\infty$ (defined in (5.6.121)), namely

$$\mathcal{L}_\omega = \mathcal{M}_1 \mathcal{L}_\infty \mathcal{M}_2^{-1}, \qquad \mathcal{M}_1 := \Phi B \rho \mathcal{T} \Phi_1 \Phi_2 \mathcal{S} \Phi_\infty, \quad \mathcal{M}_2 := \Phi B \mathcal{T} \Phi_1 \Phi_2 \mathcal{S} \Phi_\infty, \tag{5.6.125}$$

where ρ means the multiplication operator by the function ρ defined in (5.6.41). By (5.6.123) and Lemma 4.2.7 we deduce that $\|\mathcal{L}_\infty^{-1} g\|^{\mathrm{Lip}(\gamma)}_s \leq_s \gamma^{-1}\|g\|^{\mathrm{Lip}(\gamma)}_{s+2\tau+1}$. In order to estimate $\mathcal{M}_2, \mathcal{M}_1^{-1}$, we recall that the composition of tame maps is tame, see Lemma A.0.11. Now, Φ, Φ^{-1} are estimated in Lemma 5.6.3, B, B^{-1} and ρ in Lemma 5.6.4, $\mathcal{T}, \mathcal{T}^{-1}$ in Lemma 5.6.6. The decay norms $|\Phi_1|^{\mathrm{Lip}(\gamma)}_s$, $|\Phi_1^{-1}|^{\mathrm{Lip}(\gamma)}_s$, $|\Phi_2|^{\mathrm{Lip}(\gamma)}_s$, $|\Phi_2^{-1}|^{\mathrm{Lip}(\gamma)}_s \leq C(s)$ by Lemmata 5.6.9, 5.6.12. The decay norm of $\mathcal{S}, \mathcal{S}^{-1}$ is estimated in Lemma 5.6.14, and $\Phi_\infty, \Phi_\infty^{-1}$ in (5.6.122). The decay norm controls the Sobolev norm by (3.1.11). Thus, by (5.6.125),

$$\|\mathcal{M}_2 h\|^{\mathrm{Lip}(\gamma)}_s + \|\mathcal{M}_1^{-1} h\|^{\mathrm{Lip}(\gamma)}_s \leq_s \|h\|^{\mathrm{Lip}(\gamma)}_{s+3} + \varepsilon\gamma^{-1}\|\mathfrak{I}_\delta\|^{\mathrm{Lip}(\gamma)}_{s+\sigma+3}\|h\|^{\mathrm{Lip}(\gamma)}_{s_0},$$

and (5.6.124) follows. The last inequality in (5.6.124) follows by (5.4.17) and (5.4.4). □

5.7 The Nash-Moser nonlinear iteration

In this Section we prove Theorem 5.3.1. It will be a consequence of the Nash-Moser Theorem 5.7.1 below.

Consider the finite-dimensional subspaces

$$E_n := \big\{ \mathfrak{I}(\varphi) = (\Theta, y, z)(\varphi) : \ \Theta = \Pi_n \Theta, \ y = \Pi_n y, \ z = \Pi_n z \big\}$$

where $N_n := N_0^{\chi^n}$ are introduced in (5.6.114), and Π_n are the projectors (which, with a small abuse of notation, we denote with the same symbol)

$$\Pi_n \Theta(\varphi) := \sum_{|l|<N_n} \Theta_l e^{\mathrm{i} l\cdot\varphi}, \quad \Pi_n y(\varphi) := \sum_{|l|<N_n} y_l e^{\mathrm{i} l\cdot\varphi}, \quad \text{where } \Theta(\varphi) = \sum_{l\in\mathbb{Z}^\nu} \Theta_l e^{\mathrm{i} l\cdot\varphi}, \quad y(\varphi) = \sum_{l\in\mathbb{Z}^\nu} y_l e^{\mathrm{i} l\cdot\varphi},$$

$$\Pi_n z(\varphi, x) := \sum_{|(l,j)|<N_n} z_{lj} e^{\mathrm{i}(l\cdot\varphi + jx)}, \quad \text{where } z(\varphi, x) = \sum_{l\in\mathbb{Z}^\nu, j\in S^c} z_{lj} e^{\mathrm{i}(l\cdot\varphi+jx)}. \tag{5.7.1}$$

We define $\Pi_n^\perp := I - \Pi_n$. The classical smoothing properties hold: for all $\alpha, s \geq 0$,

$$\|\Pi_n \mathfrak{I}\|_{s+\alpha}^{\mathrm{Lip}(\gamma)} \leq N_n^\alpha \|\mathfrak{I}\|_s^{\mathrm{Lip}(\gamma)}, \ \forall \mathfrak{I}(\omega) \in H^s, \quad \|\Pi_n^\perp \mathfrak{I}\|_s^{\mathrm{Lip}(\gamma)} \leq N_n^{-\alpha} \|\mathfrak{I}\|_{s+\alpha}^{\mathrm{Lip}(\gamma)}, \ \forall \mathfrak{I}(\omega) \in H^{s+\alpha}. \tag{5.7.2}$$

We define the constants

$$\mu_1 := 3\mu + 9, \qquad \alpha := 3\mu_1 + 1, \qquad \alpha_1 := (\alpha - 3\mu)/2, \tag{5.7.3}$$

$$\kappa := 3\big(\mu_1 + \rho^{-1}\big) + 1, \qquad \beta_1 := 6\mu_1 + 3\rho^{-1} + 3, \qquad 0 < \rho < \frac{1-3a}{C_1(1+a)}, \tag{5.7.4}$$

where $\mu := \mu(\tau, \nu)$ is the "loss of regularity" defined in Theorem 5.4.1 (see (5.4.54)) and C_1 is fixed below.

Theorem 5.7.1. (Nash-Moser) *Assume that $f \in C^q$ with $q > S := s_0 + \beta_1 + \mu + 3$. Let $\tau \geq \nu + 2$. Then there exist $C_1 > \max\{\mu_1 + \alpha, C_0\}$ (where $C_0 := C_0(\tau, \nu)$ is the one in Theorem 5.6.1), $\delta_0 := \delta_0(\tau, \nu) > 0$ such that, if*

$$N_0^{C_1} \varepsilon^{b_*+1} \gamma^{-2} < \delta_0, \quad \gamma := \varepsilon^{2+a} = \varepsilon^{2b}, \quad N_0 := (\varepsilon\gamma^{-1})^\rho, \quad b_* := 6 - 2b, \tag{5.7.5}$$

then, for all $n \geq 0$:

$(\mathcal{P}1)_n$ *there exists a function $(\mathfrak{I}_n, \zeta_n) : \mathcal{G}_n \subseteq \Omega_\varepsilon \to E_{n-1} \times \mathbb{R}^\nu$, $\omega \mapsto (\mathfrak{I}_n(\omega), \zeta_n(\omega))$, $(\mathfrak{I}_0, \zeta_0) := 0$, $E_{-1} := \{0\}$, satisfying $|\zeta_n|^{\mathrm{Lip}(\gamma)} \leq C\|\mathcal{F}(U_n)\|_{s_0}^{\mathrm{Lip}(\gamma)}$,*

$$\|\mathfrak{I}_n\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon^{b_*} \gamma^{-1}, \quad \|\mathcal{F}(U_n)\|_{s_0+\mu+3}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon^{b_*}, \tag{5.7.6}$$

where $U_n := (i_n, \zeta_n)$ with $i_n(\varphi) = (\varphi, 0, 0) + \mathfrak{I}_n(\varphi)$. The sets $\mathcal{G}_n$ are defined inductively by:

$$\mathcal{G}_0 := \big\{ \omega \in \Omega_\varepsilon : |\omega \cdot l| \geq 2\gamma \langle l \rangle^{-\tau}, \forall l \in \mathbb{Z}^\nu \setminus \{0\} \big\},$$

$$\mathcal{G}_{n+1} := \Big\{ \omega \in \mathcal{G}_n : |\mathrm{i}\omega\cdot l + \mu_j^\infty(i_n) - \mu_k^\infty(i_n)| \geq \frac{2\gamma_n |j^3 - k^3|}{\langle l \rangle^\tau}, \forall j, k \in S^c \cup \{0\}, l \in \mathbb{Z}^\nu \Big\}, \tag{5.7.7}$$

where $\gamma_n := \gamma(1+2^{-n})$ and $\mu_j^\infty(\omega) := \mu_j^\infty(\omega, i_n(\omega))$ are defined in (5.6.118) (and $\mu_0^\infty(\omega) = 0$). The differences $\widehat{\mathfrak{I}}_n := \mathfrak{I}_n - \mathfrak{I}_{n-1}$ (where we set $\widehat{\mathfrak{I}}_0 := 0$) is defined on $\mathcal{G}_n$, and satisfy

$$\|\widehat{\mathfrak{I}}_1\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon^{b_*} \gamma^{-1}, \quad \|\widehat{\mathfrak{I}}_n\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon^{b_*} \gamma^{-1} N_{n-1}^{-\alpha_1}, \quad \forall n > 1. \tag{5.7.8}$$

$(\mathcal{P}2)_n$ $\|\mathcal{F}(U_n)\|_{s_0}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon^{b_*} N_{n-1}^{-\alpha}$ *where we set* $N_{-1} := 1$.

$(\mathcal{P}3)_n$ (High norms). $\|\mathfrak{I}_n\|_{s_0+\beta_1}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon^{b_*} \gamma^{-1} N_{n-1}^{\kappa}$ *and* $\|\mathcal{F}(U_n)\|_{s_0+\beta_1}^{\mathrm{Lip}(\gamma)} \leq C_* \varepsilon^{b_*} N_{n-1}^{\kappa}$.

$(\mathcal{P}4)_n$ (Measure). *The measure of the "Cantor-like" sets* $\mathcal{G}_n$ *satisfies*

$$|\Omega_\varepsilon \setminus \mathcal{G}_0| \leq C_* \varepsilon^{2(\nu-1)} \gamma \,, \quad \big|\mathcal{G}_n \setminus \mathcal{G}_{n+1}\big| \leq C_* \varepsilon^{2(\nu-1)} \gamma N_{n-1}^{-1} \,. \tag{5.7.9}$$

All the Lip norms are defined on $\mathcal{G}_n$*, namely* $\| \ \|_s^{\mathrm{Lip}(\gamma)} = \| \ \|_{s,\mathcal{G}_n}^{\mathrm{Lip}(\gamma)}$.

Proof. To simplify notations, in this proof we denote $\| \ \|^{\mathrm{Lip}(\gamma)}$ by $\| \ \|$. We first prove $(\mathcal{P}1,2,3)_n$.

Step 1: *Proof of* $(\mathcal{P}1,2,3)_0$. Recalling (5.3.6) we have $\|\mathcal{F}(U_0)\|_s = \|\mathcal{F}(\varphi,0,0,0)\|_s = \|X_P(\varphi,0,0)\|_s \leq_s \varepsilon^{6-2b}$ by (5.3.15). Hence (recall that $b_* = 6-2b$) the smallness conditions in $(\mathcal{P}1)_0$-$(\mathcal{P}3)_0$ hold taking $C_* := C_*(s_0+\beta_1)$ large enough.

Step 2: *Assume that* $(\mathcal{P}1,2,3)_n$ *hold for some* $n \geq 0$*, and prove* $(\mathcal{P}1,2,3)_{n+1}$. By (5.7.5) and (5.7.4),

$$N_0^{C_1} \varepsilon^{b_*+1} \gamma^{-2} = N_0^{C_1} \varepsilon^{1-3a} = \varepsilon^{1-3a-\rho C_1(1+a)} < \delta_0$$

for ε small enough, and the smallness condition (5.6.117) holds. Moreover (5.7.6) imply (5.4.4) (and so (5.5.8)) and Theorem 5.6.2 applies. Hence the operator $\mathcal{L}_\omega := \mathcal{L}_\omega(\omega, i_n(\omega))$ defined in (5.4.45) is invertible for all $\omega \in \mathcal{G}_{n+1}$ and the last estimate in (5.6.124) holds. This means that the assumption (5.4.46) of Theorem 5.4.1 is verified with $\Omega_\infty = \mathcal{G}_{n+1}$. By Theorem 5.4.1 there exists an approximate inverse $\mathbf{T}_n(\omega) := \mathbf{T}_0(\omega, i_n(\omega))$ of the linearized operator $L_n(\omega) := d_{i,\zeta}\mathcal{F}(\omega, i_n(\omega))$, satisfying (5.4.54). Thus, using also (5.7.5), (5.7.2), (5.7.6),

$$\|\mathbf{T}_n g\|_s \leq_s \gamma^{-1}\big(\|g\|_{s+\mu} + \varepsilon\gamma^{-1}\{\|\mathfrak{I}_n\|_{s+\mu} + \gamma^{-1}\|\mathfrak{I}_n\|_{s_0+\mu}\|\mathcal{F}(U_n)\|_{s+\mu}\}\|g\|_{s_0+\mu}\big) \tag{5.7.10}$$

$$\|\mathbf{T}_n g\|_{s_0} \leq_{s_0} \gamma^{-1}\|g\|_{s_0+\mu} \tag{5.7.11}$$

and, by (5.4.55), using also (5.7.6), (5.7.5), (5.7.2),

$$\begin{aligned}\|\big(L_n \circ \mathbf{T}_n - I\big)g\|_s \leq_s \gamma^{-1}\big(&\|\mathcal{F}(U_n)\|_{s_0+\mu}\|g\|_{s+\mu} + \|\mathcal{F}(U_n)\|_{s+\mu}\|g\|_{s_0+\mu} \\ &+ \varepsilon\gamma^{-1}\|\mathfrak{I}_n\|_{s+\mu}\|\mathcal{F}(U_n)\|_{s_0+\mu}\|g\|_{s_0+\mu}\big)\,,\end{aligned} \tag{5.7.12}$$

$$\begin{aligned}\|\big(L_n \circ \mathbf{T}_n - I\big)g\|_{s_0} &\leq_{s_0} \gamma^{-1}\|\mathcal{F}(U_n)\|_{s_0+\mu}\|g\|_{s_0+\mu} \\ &\leq_{s_0} \gamma^{-1}\big(\|\Pi_n\mathcal{F}(U_n)\|_{s_0+\mu} + \|\Pi_n^\perp\mathcal{F}(U_n)\|_{s_0+\mu}\big)\|g\|_{s_0+\mu} \\ &\leq_{s_0} N_n^{\mu}\gamma^{-1}\big(\|\mathcal{F}(U_n)\|_{s_0} + N_n^{-\beta_1}\|\mathcal{F}(U_n)\|_{s_0+\beta_1}\big)\|g\|_{s_0+\mu}\,.\end{aligned} \tag{5.7.13}$$

Then, for all $\omega \in \mathcal{G}_{n+1}$, $n \geq 0$, we define

$$U_{n+1} := U_n + H_{n+1}\,, \quad H_{n+1} := (\widehat{\mathfrak{I}}_{n+1}, \widehat{\zeta}_{n+1}) := -\widetilde{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) \in E_n \times \mathbb{R}^\nu\,, \tag{5.7.14}$$

where $\widetilde{\Pi}_n(\mathfrak{I},\zeta) := (\Pi_n\mathfrak{I}, \zeta)$ with Π_n in (5.7.1). Since $L_n := d_{i,\zeta}\mathcal{F}(i_n)$, we write $\mathcal{F}(U_{n+1}) = \mathcal{F}(U_n) + L_n H_{n+1} + Q_n$, where

$$Q_n := Q(U_n, H_{n+1})\,, \quad Q(U_n, H) := \mathcal{F}(U_n + H) - \mathcal{F}(U_n) - L_n H\,, \quad H \in E_n \times \mathbb{R}^\nu. \tag{5.7.15}$$

Then, by the definition of H_{n+1} in (5.7.14), and writing $\widetilde{\Pi}_n^\perp(\mathfrak{I},\zeta) := (\Pi_n^\perp \mathfrak{I}, 0)$, we have

$$\begin{aligned}
\mathcal{F}(U_{n+1}) &= \mathcal{F}(U_n) - L_n \widetilde{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n = \mathcal{F}(U_n) - L_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + L_n \widetilde{\Pi}_n^\perp \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n \\
&= \mathcal{F}(U_n) - \Pi_n L_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + (L_n \widetilde{\Pi}_n^\perp - \Pi_n^\perp L_n) \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n \\
&= \Pi_n^\perp \mathcal{F}(U_n) + R_n + Q_n + Q_n' \qquad (5.7.16)
\end{aligned}$$

where

$$R_n := (L_n \widetilde{\Pi}_n^\perp - \Pi_n^\perp L_n) \mathbf{T}_n \Pi_n \mathcal{F}(U_n) \,, \qquad Q_n' := -\Pi_n (L_n \mathbf{T}_n - I) \Pi_n \mathcal{F}(U_n) \,. \qquad (5.7.17)$$

Lemma 5.7.1. *Define*

$$w_n := \varepsilon \gamma^{-2} \|\mathcal{F}(U_n)\|_{s_0} \,, \quad B_n := \varepsilon \gamma^{-1} \|\mathfrak{I}_n\|_{s_0+\beta_1} + \varepsilon \gamma^{-2} \|\mathcal{F}(U_n)\|_{s_0+\beta_1} \,. \qquad (5.7.18)$$

Then there exists $K := K(s_0, \beta_1) > 0$ such that, for all $n \geq 0$, setting $\mu_1 := 3\mu + 9$ (see (5.7.3)),

$$w_{n+1} \leq K N_n^{\mu_1 + \frac{1}{\rho} - \beta_1} B_n + K N_n^{\mu_1} w_n^2 \,, \qquad B_{n+1} \leq K N_n^{\mu_1 + \frac{1}{\rho}} B_n \,. \qquad (5.7.19)$$

Proof. We estimate separately the terms Q_n in (5.7.15) and Q_n', R_n in (5.7.17).
Estimate of Q_n. By (5.7.15), (5.3.6), (5.3.20) and (5.7.6), (5.7.2), we have the quadratic estimates

$$\begin{aligned}
\|Q(U_n, H)\|_s &\leq_s \varepsilon \big(\|\widehat{\mathfrak{I}}\|_{s+3} \|\widehat{\mathfrak{I}}\|_{s_0+3} + \|\mathfrak{I}_n\|_{s+3} \|\widehat{\mathfrak{I}}\|_{s_0+3}^2 \big) \qquad (5.7.20) \\
\|Q(U_n, H)\|_{s_0} &\leq_{s_0} \varepsilon N_n^6 \|\widehat{\mathfrak{I}}\|_{s_0}^2 \,, \quad \forall \widehat{\mathfrak{I}} \in E_n \,. \qquad (5.7.21)
\end{aligned}$$

Now by the definition of H_{n+1} in (5.7.14) and (5.7.2), (5.7.10), (5.7.11), (5.7.6), we get

$$\begin{aligned}
\|\widehat{\mathfrak{I}}_{n+1}\|_{s_0+\beta_1} &\leq_{s_0+\beta_1} N_n^\mu \big(\gamma^{-1} \|\mathcal{F}(U_n)\|_{s_0+\beta_1} + \varepsilon \gamma^{-2} \|\mathcal{F}(U_n)\|_{s_0+\mu} \{ \|\mathfrak{I}_n\|_{s_0+\beta_1} + \gamma^{-1} \|\mathcal{F}(U_n)\|_{s_0+\beta_1} \} \big) \\
&\leq_{s_0+\beta} N_n^\mu \big(\gamma^{-1} \|\mathcal{F}(U_n)\|_{s_0+\beta_1} + \|\mathfrak{I}_n\|_{s_0+\beta_1} \big) \,, \qquad (5.7.22) \\
\|\widehat{\mathfrak{I}}_{n+1}\|_{s_0} &\leq_{s_0} \gamma^{-1} N_n^\mu \|\mathcal{F}(U_n)\|_{s_0} \,. \qquad (5.7.23)
\end{aligned}$$

Then the term Q_n in (5.7.15) satisfies, by (5.7.20), (5.7.21), (5.7.22), (5.7.23), (5.7.5), (5.7.6), $(\mathcal{P}2)_n$, (5.7.3),

$$\begin{aligned}
\|Q_n\|_{s_0+\beta_1} &\leq_{s_0+\beta_1} N_n^{2\mu+9} \gamma \big(\gamma^{-1} \|\mathcal{F}(U_n)\|_{s_0+\beta_1} + \|\mathfrak{I}_n\|_{s_0+\beta_1} \big) \,, \qquad (5.7.24) \\
\|Q_n\|_{s_0} &\leq_{s_0} N_n^{2\mu+6} \varepsilon \gamma^{-2} \|\mathcal{F}(U_n)\|_{s_0}^2 \,. \qquad (5.7.25)
\end{aligned}$$

Estimate of Q_n'. The bounds (5.7.12), (5.7.13), (5.7.2), (5.7.3), (5.7.6) imply

$$\begin{aligned}
\|Q_n'\|_{s_0+\beta_1} &\leq_{s_0+\beta_1} N_n^{2\mu} \big(\|\mathcal{F}(U_n)\|_{s_0+\beta_1} + \|\mathfrak{I}_n\|_{s_0+\beta_1} \|\mathcal{F}(U_n)\|_{s_0} \big) \,, \qquad (5.7.26) \\
\|Q_n'\|_{s_0} &\leq_{s_0} \gamma^{-1} N_n^{2\mu} \big(\|\mathcal{F}(U_n)\|_{s_0} + N_n^{-\beta_1} \|\mathcal{F}(U_n)\|_{s_0+\beta_1} \big) \|\mathcal{F}(U_n)\|_{s_0} \,. \qquad (5.7.27)
\end{aligned}$$

Estimate of R_n. For $H := (\widehat{\mathfrak{I}}, \widehat{\zeta})$ we have $(L_n \widetilde{\Pi}_n^\perp - \Pi_n^\perp L_n) H = [\bar{D}_n, \Pi_n^\perp] \widehat{\mathfrak{I}} = [\Pi_n, \bar{D}_n] \widehat{\mathfrak{I}}$ where $\bar{D}_n := d_i X_{H_\varepsilon}(i_n) + (0, 0, \partial_{xxx})$. Thus Lemma 5.3.1, (5.7.6), (5.7.2) and (5.3.19) imply

$$\begin{aligned}
\|(L_n \widetilde{\Pi}_n^\perp - \Pi_n^\perp L_n) H\|_{s_0} &\leq_{s_0+\beta_1} \varepsilon N_n^{-\beta_1+\mu+3} \big(\|\widehat{\mathfrak{I}}\|_{s_0+\beta_1-\mu} + \|\mathfrak{I}_n\|_{s_0+\beta_1-\mu} \|\widehat{\mathfrak{I}}\|_{s_0+3} \big) \,, \qquad (5.7.28) \\
\|(L_n \widetilde{\Pi}_n^\perp - \Pi_n^\perp L_n) H\|_{s_0+\beta_1} &\leq_s \varepsilon N_n^{\mu+3} \big(\|\widehat{\mathfrak{I}}\|_{s_0+\beta_1-\mu} + \|\mathfrak{I}_n\|_{s_0+\beta_1-\mu} \|\widehat{\mathfrak{I}}\|_{s_0+3} \big) \,. \qquad (5.7.29)
\end{aligned}$$

Hence, applying (5.7.10), (5.7.28), (5.7.29), (5.7.5), (5.7.6), (5.7.2), the term R_n defined in (5.7.17) satisfies

$$\|R_n\|_{s_0} \leq_{s_0+\beta_1} N_n^{\mu+6-\beta_1}(\varepsilon\gamma^{-1}\|\mathcal{F}(U_n)\|_{s_0+\beta_1} + \varepsilon\|\mathfrak{I}_n\|_{s_0+\beta_1}), \tag{5.7.30}$$

$$\|R_n\|_{s_0+\beta_1} \leq_{s_0+\beta_1} N_n^{\mu+6}(\varepsilon\gamma^{-1}\|\mathcal{F}(U_n)\|_{s_0+\beta_1} + \varepsilon\|\mathfrak{I}_n\|_{s_0+\beta_1}). \tag{5.7.31}$$

Estimate of $\mathcal{F}(U_{n+1})$. By (5.7.16) and (5.7.24), (5.7.25), (5.7.26), (5.7.27), (5.7.30), (5.7.31), (5.7.5), (5.7.6), we get

$$\|\mathcal{F}(U_{n+1})\|_{s_0} \leq_{s_0+\beta_1} N_n^{\mu_1-\beta_1}(\varepsilon\gamma^{-1}\|\mathcal{F}(U_n)\|_{s_0+\beta_1} + \varepsilon\|\mathfrak{I}_n\|_{s_0+\beta_1}) + N_n^{\mu_1}\varepsilon\gamma^{-2}\|\mathcal{F}(U_n)\|_{s_0}^2, \tag{5.7.32}$$

$$\|\mathcal{F}(U_{n+1})\|_{s_0+\beta_1} \leq_{s_0+\beta_1} N_n^{\mu_1}(\varepsilon\gamma^{-1}\|\mathcal{F}(U_n)\|_{s_0+\beta_1} + \varepsilon\|\mathfrak{I}_n\|_{s_0+\beta_1}), \tag{5.7.33}$$

where $\mu_1 := 3\mu + 9$.
Estimate of $\mathfrak{I}_{n+1}$. Using (5.7.22) the term $\mathfrak{I}_{n+1} = \mathfrak{I}_n + \widehat{\mathfrak{I}}_{n+1}$ is bounded by

$$\|\mathfrak{I}_{n+1}\|_{s_0+\beta_1} \leq_{s_0+\beta_1} N_n^{\mu}(\|\mathfrak{I}_n\|_{s_0+\beta_1} + \gamma^{-1}\|\mathcal{F}(U_n)\|_{s_0+\beta_1}). \tag{5.7.34}$$

Finally, recalling (5.7.18), the inequalities (5.7.19) follow by (5.7.32)-(5.7.34), (5.7.6) and $\varepsilon\gamma^{-1} = N_0^{1/\rho} \leq N_n^{1/\rho}$. □

Proof of $(\mathcal{P}3)_{n+1}$. By (5.7.19) and $(\mathcal{P}3)_n$,

$$B_{n+1} \leq K N_n^{\mu_1+\frac{1}{\rho}} B_n \leq 2C_* K \varepsilon^{b_*+1}\gamma^{-2} N_n^{\mu_1+\frac{1}{\rho}} N_{n-1}^{\kappa} \leq C_*\varepsilon^{b_*+1}\gamma^{-2}N_n^{\kappa}, \tag{5.7.35}$$

provided $2KN_n^{\mu_1+\frac{1}{\rho}-\kappa}N_{n-1}^{\kappa} \leq 1$, $\forall n \geq 0$. This inequality holds by (5.7.4), taking N_0 large enough (i.e ε small enough). By (5.7.18), the bound $B_{n+1} \leq C_*\varepsilon^{b_*+1}\gamma^{-2}N_n^{\kappa}$ implies $(\mathcal{P}3)_{n+1}$.

Proof of $(\mathcal{P}2)_{n+1}$. Using (5.7.19), (5.7.18) and $(\mathcal{P}2)_n, (\mathcal{P}3)_n$, we get

$$w_{n+1} \leq KN_n^{\mu_1+\frac{1}{\rho}-\beta_1}B_n + KN_n^{\mu_1}w_n^2 \leq KN_n^{\mu_1+\frac{1}{\rho}-\beta_1} 2C_*\varepsilon^{b_*+1}\gamma^{-2}N_{n-1}^{\kappa} + KN_n^{\mu_1}(C_*\varepsilon^{b_*+1}\gamma^{-2}N_{n-1}^{-\alpha})^2$$

which is $\leq C_*\varepsilon^{b_*+1}\gamma^{-2}N_n^{-\alpha}$ provided that

$$4KN_n^{\mu_1+\frac{1}{\rho}-\beta_1+\alpha}N_{n-1}^{\kappa} \leq 1, \quad 2KC_*\varepsilon^{b_*+1}\gamma^{-2}N_n^{\mu_1+\alpha}N_{n-1}^{-2\alpha} \leq 1, \quad \forall n \geq 0. \tag{5.7.36}$$

The inequalities in (5.7.36) hold by (5.7.3)-(5.7.4), (5.7.5), $C_1 > \mu_1 + \alpha$, taking δ_0 in (5.7.5) small enough. By (5.7.18), the inequality $w_{n+1} \leq C_*\varepsilon^{b_*+1}\gamma^{-2}N_n^{-\alpha}$ implies $(\mathcal{P}2)_{n+1}$.

Proof of $(\mathcal{P}1)_{n+1}$. The bound (5.7.8) for $\widehat{\mathfrak{I}}_1$ follows by (5.7.14), (5.7.10) (for $s = s_0 + \mu$) and $\|\mathcal{F}(U_0)\|_{s_0+2\mu} = \|\mathcal{F}(\varphi, 0, 0, 0)\|_{s_0+2\mu} \leq_{s_0+2\mu} \varepsilon^{b_*}$. The bound (5.7.8) for $\widehat{\mathfrak{I}}_{n+1}$ follows by (5.7.2), (5.7.23), $(\mathcal{P}2)_n$, (5.7.3). It remains to prove that (5.7.6) holds at the step $n+1$. We have

$$\|\mathfrak{I}_{n+1}\|_{s_0+\mu} \leq \sum_{k=1}^{n+1}\|\widehat{\mathfrak{I}}_k\|_{s_0+\mu} \leq C_*\varepsilon^{b_*}\gamma^{-1}\sum_{k\geq 1}N_{k-1}^{-\alpha_1} \leq C_*\varepsilon^{b_*}\gamma^{-1} \tag{5.7.37}$$

for N_0 large enough, i.e. ε small. Moreover, using (5.7.2), $(\mathcal{P}2)_{n+1}$, $(\mathcal{P}3)_{n+1}$, (5.7.3), we get

$$\begin{aligned}\|\mathcal{F}(U_{n+1})\|_{s_0+\mu+3} &\leq N_n^{\mu+3}\|\mathcal{F}(U_{n+1})\|_{s_0} + N_n^{\mu+3-\beta_1}\|\mathcal{F}(U_{n+1})\|_{s_0+\beta_1} \\ &\leq C_*\varepsilon^{b_*}N_n^{\mu+3-\alpha} + C_*\varepsilon^{b_*}N_n^{\mu+3-\beta_1+\kappa} \leq C_*\varepsilon^{b_*},\end{aligned}$$

which is the second inequality in (5.7.6) at the step $n+1$. The bound $|\zeta_{n+1}|^{\mathrm{Lip}(\gamma)} \leq C\|\mathcal{F}(U_{n+1})\|_{s_0}^{\mathrm{Lip}(\gamma)}$ is a consequence of Lemma 5.4.1 (it is not inductive).

STEP 3: *Prove* $(\mathcal{P}4)_n$ *for all* $n \geq 0$. For all $n \geq 0$,

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} = \bigcup_{l \in \mathbb{Z}^\nu,\, j,k \in S^c \cup \{0\}} R_{ljk}(i_n) \tag{5.7.38}$$

where

$$R_{ljk}(i_n) := \left\{\omega \in \mathcal{G}_n \,:\, |\mathrm{i}\omega \cdot l + \mu_j^\infty(i_n) - \mu_k^\infty(i_n)| < 2\gamma_n |j^3 - k^3|\, \langle l \rangle^{-\tau} \right\}. \tag{5.7.39}$$

Notice that $R_{ljk}(i_n) = \emptyset$ if $j = k$, so that we suppose in the sequel that $j \neq k$.

Lemma 5.7.2. *For all* $n \geq 1$, $|l| \leq N_{n-1}$, *the set* $R_{ljk}(i_n) \subseteq R_{ljk}(i_{n-1})$.

Proof. Like Lemma 4.3.2 (with ω in the role of $\lambda\bar{\omega}$, and N_{n-1} instead of N_n). ☐

By definition, $R_{ljk}(i_n) \subseteq \mathcal{G}_n$ (see (5.7.39)) and Lemma 5.7.2 implies that, for all $n \geq 1$, $|l| \leq N_{n-1}$, the set $R_{ljk}(i_n) \subseteq R_{ljk}(i_{n-1})$. On the other hand $R_{ljk}(i_{n-1}) \cap \mathcal{G}_n = \emptyset$ (see (5.7.7)). As a consequence, for all $|l| \leq N_{n-1}$, $R_{ljk}(i_n) = \emptyset$ and, by (5.7.38),

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} \subseteq \bigcup_{|l| > N_{n-1},\, j,k \in S^c \cup \{0\}} R_{ljk}(i_n) \qquad \forall n \geq 1. \tag{5.7.40}$$

Lemma 5.7.3. *Let* $n \geq 0$. *If* $R_{ljk}(i_n) \neq \emptyset$ *then* $|l| \geq C|j^3 - k^3| \geq \frac{1}{2}C(j^2 + k^2)$ *for some* $C > 0$.

Proof. Like Lemma 4.3.3. The only difference is that ω is not constrained to a fixed direction. Note also that $|j^3 - k^3| \geq (j^2 + k^2)/2$, $\forall j \neq k$. ☐

Lemma 5.7.4. *For all* $n \geq 0$, *the measure* $|R_{ljk}(i_n)| \leq C\varepsilon^{2(\nu-1)}\gamma\langle l \rangle^{-\tau}$.

Proof. Defining

$$\phi(\omega) := \mathrm{i}\omega \cdot l + \mu_j^\infty(\omega) - \mu_k^\infty(\omega),$$

where $\mu_j^\infty(\omega) := \mu_j^\infty(\omega, i_n(\omega))$ for all $j \in S^c$, we can write

$$R_{ljk}(i_n) = \left\{\omega \in \mathcal{G}_n : |\phi(\omega)| < 2\gamma_n |j^3 - k^3| \langle l \rangle^{-\tau}\right\}.$$

Let us write

$$\omega = \widehat{l}s + v, \qquad \widehat{l} := \frac{l}{|l|}, \qquad v \in \mathbb{R}^\nu, \quad v \cdot l = 0$$

and let us define

$$\psi(s) := \phi(\widehat{l}s + v).$$

Using Lemma 5.7.3, the estimate (5.6.119), we get (for ε small enough)

$$|\psi(s_1) - \psi(s_2)| \geq \frac{|l|}{2}|s_1 - s_2|,$$

which implies

$$\left|\left\{s : \widehat{l}s + v \in R_{ljk}(i_n)\right\}\right| \leq C\gamma\langle l \rangle^{-\tau}.$$

Hence by Fubini's Theorem (using that $\mathcal{G}_n \subseteq \Omega_\varepsilon$ for all $n \geq 0$)

$$|R_{ljk}(i_n)| \lessdot \mathrm{diam}(\mathcal{G}_n)^{\nu-1}\gamma\langle l\rangle^{-\tau} \lessdot \mathrm{diam}(\Omega_\varepsilon)^{\nu-1}\gamma\langle l\rangle^{-\tau} \leq C\varepsilon^{2(\nu-1)}\gamma\langle l\rangle^{-\tau}$$

and the lemma is proved. $\square$

By (5.7.38) and Lemmata 5.7.3, 5.7.4 we get

$$|\mathcal{G}_0 \setminus \mathcal{G}_1| \leq \sum_{l\in\mathbb{Z}^\nu, |j|,|k|\leq C|l|^{1/2}} |R_{ljk}(i_0)| \leq \sum_{l\in\mathbb{Z}^\nu} \frac{C\varepsilon^{2(\nu-1)}\gamma}{\langle l\rangle^{\tau-1}} \leq C'\varepsilon^{2(\nu-1)}\gamma\,.$$

For $n \geq 1$, by (5.7.40),

$$|\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq \sum_{|l|>N_{n-1}, |j|,|k|\leq C|l|^{1/2}} |R_{ljk}(i_n)| \leq \sum_{|l|>N_{n-1}} \frac{C\varepsilon^{2(\nu-1)}\gamma}{\langle l\rangle^{\tau-1}} \leq C'\varepsilon^{2(\nu-1)}\gamma N_{n-1}^{-1}$$

because $\tau \geq \nu + 2$. The estimate $|\Omega_\varepsilon \setminus \mathcal{G}_0| \leq C\varepsilon^{2(\nu-1)}\gamma$ is elementary. Thus (5.7.9) is proved. $\square$

Proof of Theorem 5.3.1 concluded. Theorem 5.7.1 implies that the sequence $(\mathfrak{I}_n, \zeta_n)$ is well defined for $\omega \in \mathcal{G}_\infty := \cap_{n\geq 0}\mathcal{G}_n$, that $\mathfrak{I}_n$ is a Cauchy sequence in $\| \ \|^{\mathrm{Lip}(\gamma)}_{s_0+\mu,\mathcal{G}_\infty}$, see (5.7.8), and $|\zeta_n|^{\mathrm{Lip}(\gamma)} \to 0$. Therefore $\mathfrak{I}_n$ converges to a limit $\mathfrak{I}_\infty$ in norm $\| \ \|^{\mathrm{Lip}(\gamma)}_{s_0+\mu,\mathcal{G}_\infty}$ and, by $(\mathcal{P}2)_n$, for all $\omega \in \mathcal{G}_\infty$, $i_\infty(\varphi) := (\varphi, 0, 0) + \mathfrak{I}_\infty(\varphi)$, is a solution of

$$\mathcal{F}(i_\infty, 0) = 0 \quad \text{with} \quad \|\mathfrak{I}_\infty\|^{\mathrm{Lip}(\gamma)}_{s_0+\mu,\mathcal{G}_\infty} \leq C\varepsilon^{6-2b}\gamma^{-1}$$

by (5.7.6) (recall that $b_* := 6 - 2b$). Therefore $\varphi \mapsto i_\infty(\varphi)$ is an invariant torus for the Hamiltonian vector field X_{H_ε} (see (5.3.5)). By (5.7.9),

$$|\Omega_\varepsilon \setminus \mathcal{G}_\infty| \leq |\Omega_\varepsilon \setminus \mathcal{G}_0| + \sum_{n\geq 0} |\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq 2C_*\varepsilon^{2(\nu-1)}\gamma + C_*\varepsilon^{2(\nu-1)}\gamma \sum_{n\geq 1} N_{n-1}^{-1} \leq C\varepsilon^{2(\nu-1)}\gamma\,.$$

The set Ω_ε in (5.3.2) has measure $|\Omega_\varepsilon| = O(\varepsilon^{2\nu})$. Hence $|\Omega_\varepsilon \setminus \mathcal{G}_\infty|/|\Omega_\varepsilon| \to 0$ as $\varepsilon \to 0$ because $\gamma = o(\varepsilon^2)$, and therefore the measure of $\mathcal{C}_\varepsilon := \mathcal{G}_\infty$ satisfies (5.3.11).

In order to complete the proof of Theorem 5.3.1 we show the linear stability of the solution $i_\infty(\omega t)$. By Section 5.4 the system obtained linearizing the Hamiltonian vector field X_{H_ε} at a quasi-periodic solution $i_\infty(\omega t)$ is conjugated to the linear Hamiltonian system

$$\begin{cases} \dot{\psi} & = K_{20}(\omega t)\eta + K_{11}^T(\omega t)w \\ \dot{\eta} & = 0 \\ \dot{w} - \partial_x K_{02}(\omega t)w & = \partial_x K_{11}(\omega t)\eta \end{cases} \tag{5.7.41}$$

(recall that the torus i_∞ is isotropic and the transformed nonlinear Hamiltonian system is (5.4.34) where $K_{00}, K_{10}, K_{01} = 0$, see Remark 5.4.1). In Section 5.6 we have proved the reducibility of the linear system $\dot{w} - \partial_x K_{02}(\omega t)w$, conjugating the last equation in (5.7.41) to a diagonal system

$$\dot{v}_j + \mu_j^\infty v_j = f_j(\omega t)\,, \quad j \in S^c\,, \quad \mu_j^\infty \in \mathrm{i}\mathbb{R}\,, \tag{5.7.42}$$

see (5.6.121), and $f(\varphi,x)=\sum_{j\in S^c} f_j(\varphi)e^{\mathrm{i}jx}\in H^s_{S^\perp}(\mathbb{T}^{\nu+1})$. Thus (5.7.41) is stable. Indeed the actions $\eta(t)=\eta_0\in\mathbb{R}$, $\forall t\in\mathbb{R}$. Moreover the solutions of the non-homogeneous equation (5.7.42) are

$$v_j(t)=c_j e^{\mu_j^\infty t}+\tilde v_j(t)\,,\quad \text{where}\quad \tilde v_j(t):=\sum_{l\in\mathbb{Z}^\nu}\frac{f_{jl}\,e^{\mathrm{i}\omega\cdot lt}}{\mathrm{i}\omega\cdot l+\mu_j^\infty}$$

is a quasi-periodic solution (recall that the first Melnikov conditions (5.6.123) hold at a solution). As a consequence (recall also $\mu_j^\infty\in\mathrm{i}\mathbb{R}$) the Sobolev norm of the solution of (5.7.42) with initial condition $v(0)=\sum_{j\in S^c} v_j(0)e^{\mathrm{i}jx}\in H^{s_0}(\mathbb{T})$, $s_0<s$, does not increase in time. $\square$

Construction of the set S of tangential sites. We finally prove that, for any $\nu\geq 1$, the set S in (1.3.7) satisfying ($\mathtt{S}1$)-($\mathtt{S}2$) can be constructed inductively with only a *finite* number of restriction at any step of the induction.

First, fix any integer $\bar\jmath_1\geq 1$. Then the set $J_1:=\{\pm\bar\jmath_1\}$ trivially satisfies ($\mathtt{S}1$)-($\mathtt{S}2$). Then, assume that we have fixed n distinct positive integers $\bar\jmath_1,\ldots,\bar\jmath_n$, $n\geq 1$, such that the set $J_n:=\{\pm\bar\jmath_1,\ldots,\pm\bar\jmath_n\}$ satisfies ($\mathtt{S}1$)-($\mathtt{S}2$). We describe how to choose another positive integer $\bar\jmath_{n+1}$, which is different from all $j\in J_n$, such that $J_{n+1}:=J_n\cup\{\pm\bar\jmath_{n+1}\}$ also satisfies ($\mathtt{S}1$), ($\mathtt{S}2$).

Let us begin with analyzing ($\mathtt{S}1$). A set of 3 elements $j_1,j_2,j_3\in J_{n+1}$ can be of these types: (i) all "old" elements $j_1,j_2,j_3\in J_n$; (ii) two "old" elements $j_1,j_2\in J_n$ and one "new" element $j_3=\sigma_3\bar\jmath_{n+1}$, $\sigma_3=\pm1$; (iii) one "old" element $j_1\in J_n$ and two "new" elements $j_2=\sigma_2\bar\jmath_{n+1}$, $j_3=\sigma_3\bar\jmath_{n+1}$, with $\sigma_2,\sigma_3=\pm1$; (iv) all "new" elements $j_i=\sigma_i\bar\jmath_{n+1}$, $\sigma_i=\pm1$, $i=1,2,3$.

In case (i), the sum $j_1+j_2+j_3$ is nonzero by inductive assumption. In case (ii), $j_1+j_2+j_3$ is nonzero provided $\bar\jmath_{n+1}\notin\{j_1+j_2: j_1,j_2\in J_n\}$, which is a finite set. In case (iii), for $\sigma_2+\sigma_3=0$ the sum $j_1+j_2+j_3=j_1$ is trivially nonzero because $0\notin J_n$, while, for $\sigma_2+\sigma_3\neq0$, the sum $j_1+j_2+j_3=j_1+(\sigma_2+\sigma_3)\bar\jmath_{n+1}\neq0$ if $\bar\jmath_{n+1}\notin\{\frac12 j: j\in J_n\}$, which is a finite set. In case (iv), the sum $j_1+j_2+j_3=(\sigma_1+\sigma_2+\sigma_3)\bar\jmath_{n+1}\neq0$ because $\bar\jmath_{n+1}\geq1$ and $\sigma_1+\sigma_2+\sigma_3\in\{\pm1,\pm3\}$.

Now we study ($\mathtt{S}2$) for the set J_{n+1}. Denote, in short, $b:=j_1^3+j_2^3+j_3^3+j_4^3-(j_1+j_2+j_3+j_4)^3$.

A set of 4 elements $j_1,j_2,j_3,j_4\in J_{n+1}$ can be of 5 types: (i) all "old" elements $j_1,j_2,j_3,j_4\in J_n$; (ii) three "old" elements $j_1,j_2,j_3\in J_n$ and one "new" element $j_4=\sigma_4\bar\jmath_{n+1}$, $\sigma_4=\pm1$; (iii) two "old" element $j_1,j_2\in J_n$ and two "new" elements $j_3=\sigma_3\bar\jmath_{n+1}$, $j_4=\sigma_4\bar\jmath_{n+1}$, with $\sigma_3,\sigma_4=\pm1$; (iv) one "old" element $j_1\in J_n$ and three "new" elements $j_i=\sigma_i\bar\jmath_{n+1}$, $\sigma_i=\pm1$, $i=2,3,4$; (v) all "new" elements $j_i=\sigma_i\bar\jmath_{n+1}$, $\sigma_i=\pm1$, $i=1,2,3,4$.

In case (i), $b\neq0$ by inductive assumption.

In case (ii), assume that $j_1+j_2+j_3+j_4\neq0$, and calculate

$$b=-3(j_1+j_2+j_3)\bar\jmath_{n+1}^2-3(j_1+j_2+j_3)^2\sigma_4\bar\jmath_{n+1}+[j_1^3+j_2^3+j_3^3-(j_1+j_2+j_3)^3]\ =:p_{j_1,j_2,j_3,\sigma_4}(\bar\jmath_{n+1}).$$

This is nonzero provided $p_{j_1,j_2,j_3,\sigma_4}(\bar\jmath_{n+1})\neq0$ for all $j_1,j_2,j_3\in J_n$, $\sigma_4=\pm1$. The polynomial p_{j_1,j_2,j_3,σ_4} is never identically zero because either the leading coefficient $-3(j_1+j_2+j_3)\neq0$ (and, if one uses ($\mathtt{S}_3$), this is always the case), or, if $j_1+j_2+j_3=0$, then $j_1^3+j_2^3+j_3^3\neq0$ by (5.1.12) (using also that $0\notin J_n$).

In case (iii), assume that $j_1+\ldots+j_4=j_1+j_2+(\sigma_3+\sigma_4)\bar\jmath_{n+1}\neq0$, and calculate

$$b=-3\alpha\bar\jmath_{n+1}^3-3\alpha^2(j_1+j_2)\bar\jmath_{n+1}^2-3(j_1+j_2)^2\alpha\bar\jmath_{n+1}-j_1j_2(j_1+j_2)=:q_{j_1,j_2,\alpha}(\bar\jmath_{n+1}),$$

where $\alpha := \sigma_3 + \sigma_4$. We impose that $q_{j_1,j_2,\alpha}(\bar{\jmath}_{n+1}) \neq 0$ for all $j_1, j_2 \in J_n$, $\alpha \in \{\pm 2, 0\}$. The polynomial $q_{j_1,j_2,\alpha}$ is never identically zero because either the leading coefficient $-3\alpha \neq 0$, or, for $\alpha = 0$, the constant term $-j_1 j_2 (j_1 + j_2) \neq 0$ (recall that $0 \notin J_n$ and $j_1 + j_2 + \alpha \bar{\jmath}_{n+1} \neq 0$).

In case (iv), assume that $j_1 + \ldots + j_4 = j_1 + \alpha \bar{\jmath}_{n+1} \neq 0$, where $\alpha := \sigma_2 + \sigma_3 + \sigma_4 \in \{\pm 1, \pm 3\}$, and calculate

$$b = \alpha \bar{\jmath}_{n+1} r_{j_1,\alpha}(\bar{\jmath}_{n+1}), \quad r_{j_1,\alpha}(x) := (1 - \alpha^2) x^2 - 3\alpha j_1 x - 3 j_1^2.$$

The polynomial $r_{j_1,\alpha}$ is never identically zero because $j_1 \neq 0$. We impose $r_{j_1,\alpha}(\bar{\jmath}_{n+1}) \neq 0$ for all $j_1 \in J_n$, $\alpha \in \{\pm 1, \pm 3\}$.

In case (v), assume that $j_1 + \ldots + j_4 = \alpha \bar{\jmath}_{n+1} \neq 0$, with $\alpha := \sigma_1 + \ldots + \sigma_4 \neq 0$, and calculate $b = \alpha (1 - \alpha^2) \bar{\jmath}_{n+1}^3$. This is nonzero because $\bar{\jmath}_{n+1} \geq 1$ and $\alpha \in \{\pm 2, \pm 4\}$.

We have proved that, in choosing $\bar{\jmath}_{n+1}$, there are only finitely many integers to avoid.

Chapter 6

Quasi-linear perturbations of m-KdV

In this Chapter we describe how to prove Theorem 1.3.2 for the m-KdV equation (1.3.9). All the details of the proof are given in [9]. The strategy is the same as the one developed in Chapter 5 to prove Theorem 1.3.1 for the KdV equation (1.3.1). We describe below the main differences. To simplify notations we deal with the focusing m-KdV, namely the equation (1.3.9) with the sign $+$ in front of the term $\partial_x(u^3)$. The arguments are analogous in the defocusing case (we look for small amplitude solutions).

The Hamiltonian of the focusing perturbed m-KdV equation may be written as $H = H_2 + H_4 + H_{\geq 5}$, where

$$H_2(u) := \frac{1}{2}\int_{\mathbb{T}} u_x^2\, dx\,, \quad H_3(u) := -\frac{1}{4}\int_{\mathbb{T}} u^4 dx\,, \quad H_{\geq 5}(u) := \int_{\mathbb{T}} f(x,u,u_x)dx\,, \tag{6.0.1}$$

and f satisfies (1.3.10). According to the splitting (5.0.2) $u = v + z$, $v \in H_S$, $z \in H_S^\perp$, the Hamiltonian H_4 becomes

$$H_4 = -\frac{1}{4}\int_{\mathbb{T}} v^4 dx - \int_{\mathbb{T}} v^3 z dx - \frac{3}{2}\int_{\mathbb{T}} v^2 z^2 dx - \int_{\mathbb{T}} v z^3 dx - \frac{1}{4}\int_{\mathbb{T}} z^4 dx\,. \tag{6.0.2}$$

For the cubic nonlinearity it is sufficient to perform only one step of weak Birkhoff normal form in order to remove-normalize the terms of H_4 which are linear in z.

Theorem 6.0.2 (Birkhoff normal form). *There exists an analytic invertible symplectic transformation of the phase space $\Phi : H_0^1(\mathbb{T}) \to H_0^1(\mathbb{T})$ of the form*

$$\Phi(u) = u + \Psi(u)\,, \quad \Psi(u) = \Pi_E \Psi(\Pi_E u)\,, \tag{6.0.3}$$

where E is a finite-dimensional space as in (5.1.3), such that the transformed Hamiltonian is

$$\mathcal{H} := H \circ \Phi = H_2 + \mathcal{H}_4 + \mathcal{H}_{\geq 5}\,, \tag{6.0.4}$$

with

$$\mathcal{H}_4 = \frac{3}{4}\Big(\sum_{j \in S} |u_j|^4 - \sum_{j,j' \in S} |u_j|^2 |u_{j'}|^2\Big) - \frac{3}{2}\int_{\mathbb{T}} v^2 z^2 dx - \int_{\mathbb{T}} v\, z^3 dx - \frac{1}{4}\int_{\mathbb{T}} z^4 dx\,. \tag{6.0.5}$$

and $\mathcal{H}_{\geq 5}$ collects all the terms of order at least five in u.

This theorem may be proved following the same method used for Proposition 5.1.1.

Remark 6.0.1. *In the case in which the Hamiltonian density f does not depend on x, since $\|u\|^2_{L^2(\mathbb{T})}$ is a prime integral, the Hamiltonian system generated by H is equivalent to the Hamiltonian system generated by*

$$H + aG\,, \quad a \in \mathbb{R} \qquad \textit{where} \qquad G(u) := \|u\|^2_{L^2(\mathbb{T})} := \Big(\int_{\mathbb{T}} u^2\,dx \Big)^2\,,$$

hence we deal with the Hamiltonian $H + aG$. Choosing $a = 3/4$, the order 4 of the transformed Hamiltonian under the Birkhoff map Φ becomes

$$\begin{aligned}\mathcal{H}_4 = \frac{3}{4}\sum_{j\in S}|u_j|^4 - \frac{3}{2}\int_{\mathbb{T}} v^2 z^2 dx + \frac{3}{2}\Big(\int_{\mathbb{T}} v^2 dx\Big)\Big(\int_{\mathbb{T}} z^2 dx\Big) - \int_{\mathbb{T}} z^3 v dx \\ - \int_{\mathbb{T}} \frac{z^4}{4} dx + \frac{3}{4}\Big(\int_{\mathbb{T}} z^2 dx\Big)^2 .\end{aligned} \tag{6.0.6}$$

Introducing the action-angle variables (5.2.1) and after the rescaling (5.2.5), with $b := 4/3$, the Hamiltonian $\mathcal{H}$ in (6.0.4) transforms into the Hamiltonian

$$H_\varepsilon := \alpha(\xi)\cdot y + \frac{1}{2}\big(N(\theta)z, z\big)_{L^2(\mathbb{T})} + \varepsilon^{\frac{7}{3}} P(\theta, y, z)\,,$$

where the frequency-to-amplitude modulation is

$$\alpha(\xi) := \bar\omega + \varepsilon^2 \mathbb{A}\xi\,, \tag{6.0.7}$$

$\bar\omega$ is defined in (5.0.3) and $\mathbb{A}$ is the $(\nu\times\nu)$-matrix

$$\mathbb{A} := -3D_S\mathbb{A}_0\,, \qquad D_S := \mathrm{diag}_{j\in S^+} j\,, \qquad \mathbb{A}_0 := 2U - Id\,, \tag{6.0.8}$$

denoting Id the identity matrix on $\mathbb{R}^\nu$ and U the $(\nu\times\nu)$-matrix with all elements equal to 1. A direct calculation shows that the matrix $\mathbb{A}$ is invertible with inverse

$$\mathbb{A}^{-1} = \frac{1}{3}D_S^{-1} - \frac{2}{3(2\nu-1)}UD_S^{-1}\,. \tag{6.0.9}$$

Remark 6.0.2. *If the Hamiltonian density f does not depend on x, by (6.0.6) it turns out that the matrix*

$$\mathbb{A} = 3D_S$$

is diagonal, since the Birkhoff normal form restricted to the tangential sites S is diagonal.

We look for embedded invariant tori $i : \varphi \to i(\varphi) := (\theta(\varphi), y(\varphi), z(\varphi))$ for the Hamiltonian vector field X_{H_ε} filled by quasi-periodic solutions with diophantine frequency ω as in (5.3.4). We require that the diophantine constant γ satisfies $\varepsilon^{\frac{7}{3}}\gamma^{-1} \ll 1$, so that we are reduced to study perturbations of an isochronous system. In this case, it is convenient to introduce ξ as a variable (see Remark 5.3.1) and to look for zeros of the nonlinear operator

$$\mathcal{F}(i,\xi,\zeta) := \omega\cdot\partial_\varphi i(\varphi) - X_{H_{\varepsilon,\zeta}}(i(\varphi),\xi,\zeta)\,, \qquad H_{\varepsilon,\zeta} := H_\varepsilon + \zeta\cdot\theta\,. \tag{6.0.10}$$

The unknowns of the problem are the embedded invariant torus i and ξ, ζ. The variable ζ has the same role as in Chapter 5 and the variable ξ allows to control the average in the θ-component of the linearized equation. The details of these approach are given in [22]. The existence of invariant tori for the Hamiltonian vector field X_{H_ε} is stated in the following theorem.

Theorem 6.0.3. *Let the tangential sites S in* (1.3.7) *satisfy* (1.3.11)*. Then, for all $\varepsilon \in (0, \varepsilon_0)$, where ε_0 is small enough, there exists a Cantor-like set $\mathcal{C}_\varepsilon \subset \Omega_\varepsilon$, with asympotically full measure as $\varepsilon \to 0$, namely*

$$\lim_{\varepsilon \to 0} \frac{|\mathcal{C}_\varepsilon|}{|\Omega_\varepsilon|} = 1 \,, \tag{6.0.11}$$

such that, for all $\omega \in \mathcal{C}_\varepsilon$, there exists

$$\xi_\infty(\omega, \varepsilon) = \xi(\omega, \varepsilon) + O(\varepsilon^{\frac{1}{3}}) \,, \tag{6.0.12}$$

where $\xi = \xi(\omega, \varepsilon)$ is such that $\alpha(\xi) = \omega$, and a solution $i_\infty(\varphi) := i_\infty(\omega, \varepsilon)(\varphi)$ of $\mathcal{D}_\omega i_\infty(\varphi) - X_{H_\varepsilon}(i_\infty(\varphi), \xi_\infty) = 0$. Hence the embedded torus $\varphi \mapsto i_\infty(\varphi)$ is invariant for the Hamiltonian vector field $X_{H_\varepsilon(\cdot, \xi_\infty)}$ and it is filled by quasi-periodic solutions with frequency ω. The torus i_∞ satisfies

$$\|i_\infty(\varphi) - (\varphi, 0, 0)\|_{s_0+\mu}^{\mathrm{Lip}(\gamma)} = O(\varepsilon^{\frac{7}{3}} \gamma^{-1}) \tag{6.0.13}$$

for some $\mu := \mu(\nu) > 0$. Moreover, the torus i_∞ is LINEARLY STABLE.

This Theorem may be deduced by a Nash-Moser scheme with approximate inverse in Sobolev class. As we explained in Section 5.4, the problem of finding an approximate inverse for the linearized operator

$$d_{i,\xi,\zeta} \mathcal{F}(i_0, \xi_0, \zeta_0)[\widehat{\imath}, \widehat{\xi}, \widehat{\zeta}] = \omega \cdot \partial_\varphi \widehat{\imath} - \partial_{i,\xi} X_{H_\varepsilon}(i_0, \xi_0)[\widehat{\imath}, \widehat{\xi}] + (0, \widehat{\zeta}, 0) \,,$$

is reduced to invert the linearized operator in the normal directions $\mathcal{L}_\omega$ (see (5.4.45)).

Following the strategy of Section 5.5 one can prove that the linearized operator in the normal directions has the form

$$\mathcal{L}_\omega := \Pi_S^\perp \big(\omega \cdot \partial_\varphi h + \partial_{xx}(a_1 \partial_x h) + \varepsilon^2 \partial_x (V_0(\theta_0(\varphi), x) h) + \partial_x (a_0 h) + \partial_x \mathcal{R} h \big) \,, \quad \forall h \in H_S^\perp \,, \tag{6.0.14}$$

where

$$\|a_1 - 1\|_{s_0}^{\mathrm{Lip}(\gamma)} \,,\ \|a_0\|_{s_0}^{\mathrm{Lip}(\gamma)} = O(\varepsilon^{\frac{7}{3}}) \,,$$

$\mathcal{R}$ is a finite rank operator as in (5.5.2) satisfying $|\mathcal{R}|_{s_0}^{\mathrm{Lip}(\gamma)} = O(\varepsilon^{\frac{7}{3}})$ and

$$V_0(\theta_0(\varphi), x) := 3 v^2(\theta_0(\varphi), x) \,, \qquad v(\theta, x) := \sum_{j \in S} \sqrt{\xi_j} e^{\mathrm{i}\theta_j} e^{\mathrm{i}jx} \,.$$

We reduce the operator (6.0.14) to constant coefficients by means of the procedure developed in Section 5.6. The smallness condition for the reducibility scheme (see Theorem 5.6.1) is

$$N_0^{C_0} \varepsilon^{\frac{7}{3} - a} \gamma^{-1} \leq \delta_0 \,, \quad \gamma := \varepsilon^{2+a} \,, \quad 0 < a < \frac{1}{6} \,,$$

for some constants $C_0 := C_0(\tau, \nu) > 0$ and $\delta_0 := \delta_0(\tau, \nu) > 0$.

Since

$$\varepsilon^2 [V_0(\theta_0(\varphi), x) - V_0(\varphi, x)] = \varepsilon^2 O\big(\theta_0(\varphi) - \varphi\big) = \varepsilon^2 O(\varepsilon^{\frac{7}{3}} \gamma^{-1}) = O(\varepsilon^{\frac{7}{3} - a}) \,,$$

the only non-perturbative term for the reducibility scheme is $\varepsilon^2\partial_x(V_0(\varphi,x)\cdot)$. This term may be reduced to the constant coefficient term $\varepsilon^2 c(\xi)\partial_x$, where

$$c(\xi) := \int_{\mathbb{T}} V_0(\varphi,x)\,dx = 6\varepsilon^2 \sum_{j\in S^+} \xi_j\,, \tag{6.0.15}$$

by means of a linear BNF-step as the one performed in Section 5.6.5.

Remark 6.0.3. *In the case in which the Hamiltonian density does not depend on x, we can completely eliminate the order ε^2, since*

$$V_0(\varphi,x) := 3\Big(v^2(\varphi,x) - \int_{\mathbb{T}} v^2(\varphi,x)\,dx\Big)$$

has zero average in x.

After the reduction procedure, we get a diagonal operator $\mathcal{L}_\infty$ of the form (5.6.121), where the eigenvalues of the diagonal operator $\mathcal{D}_\infty$ have the asymptotic expansion

$$\mu_j^\infty = \mathrm{i}(-m_3 j^3 + m_1 j + \varepsilon^2 c(\xi) j) + r_j^\infty\,, \tag{6.0.16}$$

where

$$|m_3-1|^{\mathrm{Lip}(\gamma)}\,, |m_1|^{\mathrm{Lip}(\gamma)} = O(\varepsilon^{\frac{7}{3}})\,, \quad \sup_{j\in S^c}|r_j^\infty|^{\mathrm{Lip}(\gamma)} = O(\varepsilon^{\frac{7}{3}-a})\,. \tag{6.0.17}$$

The presence of the term $\varepsilon^2\mathrm{i}c(\xi)j$ in the expansion (6.0.16) makes the measure estimates of the resonant sets $R_{ljk}(i_n)$ in (5.7.39) more delicate. If $j^2+k^2 \geq C_1$ for some constant $C_1 > 0$ large enough, the estimate of Lemma 5.7.4 can be proved with the same argument. To estimate the measure of the resonant sets $R_{ljk}(i_n)$ for $j^2+k^2 \leq C_1$, we need the condition (1.3.11) on the tangential sites S. Let us give the details of such measure estimate, see Lemma 6.0.6. Defining

$$\phi(\omega) := \mathrm{i}\omega\cdot l + \mu_j^\infty(\omega) - \mu_k^\infty(\omega)\,,$$

where $\mu_j^\infty(\omega) := \mu_j^\infty(\omega, i_n(\omega))$ for all $j\in S^c$, the resonant set $R_{ljk}(i_n)$ can be written as

$$R_{ljk}(i_n) := \big\{\omega\in\mathcal{G}_n : |\phi(\omega)| < 2\gamma_n|j^3-k^3|\langle l\rangle^{-\tau}\big\}\,.$$

Using (6.0.16), (6.0.7), the invertibility of the matrix $\mathbb{A}$ and since by (6.0.15)

$$c(\xi) = 6\sum_{j\in S^+}\xi_j = 6\xi\cdot\vec{1}\,, \quad \vec{1} := (1,1,\dots,1)\in\mathbb{R}^\nu\,,$$

we can write

$$\phi(\omega) = a_{jk} + b_{ljk}\cdot\omega + q_{ljk}(\omega)\,,$$

where

$$a_{jk} := -\mathrm{i}(j^3-k^3) - 6\mathrm{i}(j-k)\vec{1}\cdot\mathbb{A}^{-1}[\bar\omega]\,, \quad b_{ljk} := \mathrm{i}l + 6\mathrm{i}(j-k)\mathbb{A}^{-T}[\vec{1}]\,, \tag{6.0.18}$$

and

$$|q_{ljk}|^{\mathrm{Lip}(\gamma)} = O(\varepsilon^{\frac{7}{3}-a})\,. \tag{6.0.19}$$

We prove the following lemma.

Lemma 6.0.5. *Assume* (1.3.11)*. Then for all $j \neq k$, $j^2 + k^2 \leq C_1$, we have $a_{jk} \neq 0$.*

Proof. Using (6.0.9) and since $D_S^{-1}\bar{\omega} = (\bar{j}_1^2, \ldots, \bar{j}_\nu^2)$, one has

$$\mathbb{A}^{-1}[\bar{\omega}] = \frac{1}{3}D_S^{-1}\bar{\omega} - \frac{2}{3(2\nu-1)}UD_S^{-1}\bar{\omega} = \frac{1}{3}D_S^{-1}\bar{\omega} - \frac{2}{3(2\nu-1)}\sum_{j\in S^+} j^2\,\vec{1}\,,$$

$$\vec{1}\cdot\mathbb{A}^{-1}[\bar{\omega}] = -\frac{1}{3(2\nu-1)}\sum_{j\in S^+} j^2\,.$$

Thus by (6.0.18), a_{jk} becomes

$$a_{jk} = -\mathrm{i}\Big(j^3 - k^3 - \frac{2(j-k)}{2\nu-1}\sum_{j'\in S^+} j'^2\Big) = -\mathrm{i}(j-k)\Big\{j^2 + k^2 + jk - \frac{2}{2\nu-1}\sum_{j'\in S^+} j'^2\Big\}\,.$$

Since $j \neq k$, assumption (1.3.11) implies the Lemma. $\square$

The previous lemma implies that

$$\delta := \min\{|a_{jk}| : j^2 + k^2 \leq C_0\} > 0. \tag{6.0.20}$$

Lemma 6.0.6. *If $j^2 + k^2 \leq C_1$, we have $|R_{ljk}(i_n)| \leq C\varepsilon^{2(\nu-1)}\gamma\langle l\rangle^{-\tau}$.*

Proof. For all $j^2 + k^2 \leq C_1$, $\omega \in R_{ljk}(i_n)$, one has

$$|b_{ljk}\cdot\omega| \geq |a_{jk}| - |\phi(\omega)| - |q_{ljk}(\omega)| \overset{(6.0.20),(6.0.19)}{\geq} \delta - 2\gamma_n|j^3 - k^3|\langle l\rangle^{-\tau} - C\varepsilon^{\frac{7}{3}-a} \geq \delta - C\gamma \geq \delta/2$$

for ε small enough. Therefore using that $\omega = \bar{\omega} + O(\varepsilon^2)$

$$|b_{ljk}| \geq \frac{\delta}{2|\omega|} \geq \frac{\delta}{4|\bar{\omega}|} =: \tilde{\delta} > 0\,.$$

By (6.0.19) we have $|q_{ljk}|^{\mathrm{lip}} \leq \gamma^{-1}|q_{ljk}|^{\mathrm{Lip}(\gamma)} \leq C\varepsilon^{\frac{7}{3}-a}\gamma^{-1} = C\varepsilon^{\frac{1}{3}-2a}$. Thus, writing

$$\omega = s\hat{b}_{ljk} + v_{ljk}\,, \quad \hat{b}_{ljk} := b_{ljk}/|b_{ljk}|\,, \quad v_{ljk}\cdot b_{ljk} = 0\,, \quad \psi(s) := \phi(s\hat{b}_{ljk} + v_{ljk})\,,$$

we obtain $|\psi(s_1) - \psi(s_2)| \geq (|b_{ljk}| - |q_{ljk}|^{\mathrm{lip}})|s_1 - s_2| \geq \frac{\tilde{\delta}}{2}|s_1 - s_2|$ for ε small enough (we take $a < 1/6$). Hence

$$\Big|\{s : |\psi(s)| \leq 2\gamma_n|j^3 - k^3|\langle l\rangle^{-\tau}\}\Big| \leq C\gamma\langle l\rangle^{-\tau}$$

and the lemma follows by Fubini's Theorem. $\square$

In conclusion the measure estimates follow as in Section 5.7 and Theorem 6.0.3 is proved (and thus also Theorem 1.3.2).

Remark 6.0.4. *In the case in which the Hamiltonian density $f = f(u, u_x)$ is independent of x, there is no need to prove Lemmata 6.0.5-6.0.6. Indeed, in the expansion of the eigenvalues* (6.0.16)*, the term $\mathrm{i}\varepsilon^2 c(\xi)j$ is zero and the measure estimates follows as in Lemma 5.7.4, using that $\varepsilon^{\frac{7}{3}-a}\gamma^{-1}$ is small enough.*

Chapter 7

Future perspectives

The methods developed in Chapters 4, 5, 6 are general techniques to deal with PDEs with quasi-linear and fully nonlinear perturbations. In this Chapter we mention some natural perspectives.

- **Kirchoff equation**

  We propose to prove existence and stability of quasi-periodic solutions of the forced Kirchoff equation in 1 space dimension (both with Dirichlet and periodic boundary conditions), namely

$$\partial_{tt}u - \Big(1 + \int_{\mathbb{T}} |\partial_x u|^2 \, dx\Big)\partial_{xx}u = \varepsilon f(\omega t, x)\,, \tag{7.0.1}$$

  where f is a differentiable forcing term and $\omega \in \mathbb{R}^\nu$ is a diophantine frequency vector. The equation (7.0.1) is a quasi-linear wave equation. Lax [58], Klainermann and Majda [51] found some classes of quasi-linear wave equations for which all solutions become singular in finite time. Neverthless the existence of periodic solutions of the Kirchoff equation (in any space dimension) has been proved by Baldi in [3]. Such a method does not work for quasi-periodic solutions. We plan to follow the strategy developed in Chapter 4 for the forced equation (1.2.1). The key point concerns the reduction to constant coefficients of the linearized operator. It is thanks to the special structure of the nonlinearity (it depends only on time) that we can reduce the linearized equation to constant coefficients up to a zero order operator. After that, the reducibility scheme will complete the diagonalization.

- **Higher order KdV models**

  The KdV equation arises from fluid dynamics as an approximation of the Euler equation for the water waves, more precisely KdV is a model equation for long surface waves of water in a shallow channel. In [32] Craig derived other approximate models of the Euler equation. One of these models is the *higher order KdV* equation

$$\partial_t u + \partial_{xxx}u + \partial_x(u^2) + \delta\partial_x\Big\{\partial_x^4 u + (\partial_x u)^2 + 2u(\partial_x^2 u)\Big\} = 0\,, \quad x \in \mathbb{T}\,, \tag{7.0.2}$$

  where $\delta > 0$ is a small parameter. We propose to prove existence and stability of quasi-periodic small amplitude solutions for this autonomous equation. Notice that, for each $\delta > 0$ fixed, the equation (7.0.2) is not quasi-linear, indeed the highest order constant coefficients operator is $\delta\partial_x^5$ and the nonlinearity is of order $O(\partial_x^3)$. We propose to study the singular perturbation problem when the parameter $\delta \to 0$.

- **Water waves**

One of the main motivations to study KAM theory for quasi-linear and fully nonlinear PDEs is its possible extension to the water waves equations. The motion of a 2-dimensional perfect, incompressible, irrotational fluid in infinite depth, with periodic boundary conditions and which occupies the free boundary region

$$S_\eta := \big\{(x,y) \in \mathbb{T} \times \mathbb{R} \,:\, y < \eta(t,x)\,, \quad \mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})\big\}\,,$$

is described by the system

$$\begin{cases} \partial_t \Phi + \frac{1}{2}|\nabla\Phi|^2 + g\eta = \kappa \frac{\eta_{xx}}{(1+\eta_x^2)^{3/2}} & \text{at } y = \eta(x) \\ \Delta\Phi = 0 & \text{in } S_\eta \\ \nabla\Phi \to 0 & \text{as } y \to -\infty \\ \partial_t \eta = \partial_y \Phi - \partial_x \eta \cdot \partial_x \Phi & \text{at } \quad y = \eta(x) \end{cases} \tag{7.0.3}$$

where g is the acceleration of gravity, κ is the surface tension coefficient and

$$\frac{\eta_{xx}}{(1+\eta_x^2)^{3/2}} = \partial_x \left(\frac{\eta_x}{\sqrt{1+\eta_x^2}} \right)$$

is the mean curvature of the free surface. The unknowns of the problem are the free surface $y = \eta(x)$ and the velocity potential $\Phi : S_\eta \to \mathbb{R}$, i.e. the irrotational velocity field $v = \nabla_{x,y}\Phi$ of the fluid.

Following Zakharov [74] and Craig-Sulem [35], the evolution problem (7.0.3) may be written as the infinite dimensional Hamiltonian system

$$\begin{cases} \partial_t \eta = G(\eta)\psi, \\ \partial_t \psi + \eta + \dfrac{1}{2}\psi_x^2 - \dfrac{1}{2}\dfrac{\big(G(\eta)\psi + \eta_x\psi_x\big)^2}{1+\eta_x^2} = \kappa \dfrac{\eta_{xx}}{(1+\eta_x^2)^{3/2}} \end{cases} \tag{7.0.4}$$

where $\psi(t,x)$ is the value of the velocity potential on the profile $y = \eta(t,x)$, namely

$$\psi(t,x) = \Phi(t,x,\eta(t,x))$$

and $G(\eta)$ is the so-called Dirichlet-Neumann operator defined by

$$G(\eta)\psi(x) := \sqrt{1+\eta_x^2}\,\partial_n \Phi|_{y=\eta(x)} = (\partial_y\Phi)(x,\eta(x)) - \eta_x(x)\cdot(\partial_x\Phi)(x,\eta(x))\,. \tag{7.0.5}$$

We plan to prove existence of quasi-periodic solutions for the system (7.0.4) for both gravity-capillary water waves (i.e. $\kappa \neq 0$) and gravity water waves (i.e. $\kappa = 0$). The known results concern so far only the existence of periodic solutions. The existence of periodic travelling water waves has been proved by Craig-Nicholls [33] for the capillary case and Iooss-Plotnikov

[47] without capillarity. On the other hand the existence of periodic standing wave solutions has been proved by Iooss-Plotnikov-Toland [45] for the non-capillary case and, recently, by Alazard-Baldi [1] with capillarity.

In order to prove the existence of quasi-periodic solutions we plan to combine the KAM method developed in Chapter 5 with the theory of pseudo-differential and Fourier integral operators, for the analysis of the linearized equations.

Appendix A

Tame and Lipschitz estimates

In this Appendix we present standard tame and Lipschitz estimates for composition of functions and changes of variables which are used in the Thesis.

Let $H^s := H^s(\mathbb{T}^d, \mathbb{C})$ (with norm $\| \, \|_s$) and $W^{s,\infty} := W^{s,\infty}(\mathbb{T}^d, \mathbb{C})$, $d \geq 1$.

Lemma A.0.7. *Let $s_0 > d/2$. Then*
(i) **Embedding.** $\|u\|_{L^\infty} \leq C(s_0)\|u\|_{s_0}$ *for all* $u \in H^{s_0}$.
(ii) **Algebra.** $\|uv\|_{s_0} \leq C(s_0)\|u\|_{s_0}\|v\|_{s_0}$ *for all* $u, v \in H^{s_0}$.
(iii) **Interpolation.** *For* $0 \leq s_1 \leq s \leq s_2$, $s = \lambda s_1 + (1-\lambda)s_2$,

$$\|u\|_s \leq \|u\|_{s_1}^{\lambda} \|u\|_{s_2}^{1-\lambda} , \quad \forall u \in H^{s_2} . \tag{A.0.1}$$

Let $a_0, b_0 \geq 0$ *and* $p, q > 0$. *For all* $u \in H^{a_0+p+q}$, $v \in H^{b_0+p+q}$,

$$\|u\|_{a_0+p}\|v\|_{b_0+q} \leq \|u\|_{a_0+p+q}\|v\|_{b_0} + \|u\|_{a_0}\|v\|_{b_0+p+q} . \tag{A.0.2}$$

Similarly, for the $|u|_{s,\infty} := \sum_{|\beta| \leq s} |D^\beta u|_{L^\infty}$ *norm,*

$$|u|_{s,\infty} \leq C(s_1, s_2)|u|_{s_1,\infty}^{\lambda} |u|_{s_2,\infty}^{1-\lambda} , \quad \forall u \in W^{s_2,\infty} , \tag{A.0.3}$$

and $\forall u \in W^{a_0+p+q,\infty}$, $v \in W^{b_0+p+q,\infty}$,

$$|u|_{a_0+p,\infty}|v|_{b_0+q,\infty} \leq C(a_0, b_0, p, q)\big(|u|_{a_0+p+q,\infty}|v|_{b_0,\infty} + |u|_{a_0,\infty}|v|_{b_0+p+q,\infty}\big) . \tag{A.0.4}$$

(iv) **Asymmetric tame product.** *For* $s \geq s_0$,

$$\|uv\|_s \leq C(s_0)\|u\|_s\|v\|_{s_0} + C(s)\|u\|_{s_0}\|v\|_s , \quad \forall u, v \in H^s . \tag{A.0.5}$$

(v) **Asymmetric tame product in $W^{s,\infty}$.** *For* $s \geq 0$, $s \in \mathbb{N}$,

$$|uv|_{s,\infty} \leq \tfrac{3}{2}\, |u|_{L^\infty}|v|_{s,\infty} + C(s)|u|_{s,\infty}|v|_{L^\infty} , \quad \forall u, v \in W^{s,\infty} . \tag{A.0.6}$$

(vi) **Mixed norms asymmetric tame product.** *For* $s \geq 0$, $s \in \mathbb{N}$,

$$\|uv\|_s \leq \tfrac{3}{2}\, |u|_{L^\infty}\|v\|_s + C(s)|u|_{s,\infty}\|v\|_0 , \quad \forall u \in W^{s,\infty} , \ v \in H^s . \tag{A.0.7}$$

If $u := u(\lambda)$ *and* $v := v(\lambda)$ *depend in a Lipschitz way on* $\lambda \in \Lambda \subset \mathbb{R}^\nu$, *all the previous statements hold if we replace the norms* $\| \cdot \|_s$, $| \cdot |_{s,\infty}$ *with the norms* $\| \cdot \|_s^{\mathrm{Lip}(\gamma)}$, $| \cdot |_{s,\infty}^{\mathrm{Lip}(\gamma)}$.

Proof. The interpolation estimate (A.0.1) for the Sobolev norm (1.2.5) follows by Hölder inequality, see also [62], page 269. Let us prove (A.0.2). Let $a = a_0\lambda + a_1(1-\lambda)$, $b = b_0(1-\lambda) + b_1\lambda$, $\lambda \in [0,1]$. Then (A.0.1) implies

$$\|u\|_a\|v\|_b \le \big(\|u\|_{a_0}\|v\|_{b_1}\big)^{\lambda}\big(\|u\|_{a_1}\|v\|_{b_0}\big)^{1-\lambda} \le \lambda\|u\|_{a_0}\|v\|_{b_1} + (1-\lambda)\|u\|_{a_1}\|v\|_{b_0} \tag{A.0.8}$$

by Young inequality. Applying (A.0.8) with $a = a_0 + p$, $b = b_0 + q$, $a_1 = a_0 + p + q$, $b_1 = b_0 + p + q$, then $\lambda = q/(p+q)$ and we get (A.0.2). Also the interpolation estimates (A.0.3) are classical (see [19]) and (A.0.3) implies (A.0.4) as above.

(iv): see the Appendix of [19]. (v): we write, in the standard multi-index notation,

$$D^\alpha(uv) = \sum_{\beta+\gamma=\alpha} C_{\beta,\gamma}(D^\beta u)(D^\gamma v) = uD^\alpha v + \sum_{\beta+\gamma=\alpha,\beta\neq 0} C_{\beta,\gamma}(D^\beta u)(D^\gamma v)\,. \tag{A.0.9}$$

Using $|(D^\beta u)(D^\gamma v)|_{L^\infty} \le |D^\beta u|_{L^\infty}|D^\gamma v|_{L^\infty} \le |u|_{|\beta|,\infty}|v|_{|\gamma|,\infty}$, and the interpolation inequality (A.0.3) for every $\beta \neq 0$ with $\lambda := |\beta|/|\alpha| \in (0,1]$ (where $|\alpha| \le s$), we get, for any $K > 0$,

$$\begin{aligned} C_{\beta,\gamma}|D^\beta u|_{L^\infty}|D^\gamma v|_{L^\infty} &\le C_{\beta,\gamma}C(s)\big(|v|_{L^\infty}|u|_{s,\infty}\big)^{\lambda}\big(|v|_{s,\infty}|u|_{L^\infty}\big)^{1-\lambda} \\ &= \frac{C(s)}{K}\big[(KC_{\beta,\gamma})^{\frac{1}{\lambda}}|v|_{L^\infty}|u|_{s,\infty}\big]^{\lambda}\big(|v|_{s,\infty}|u|_{L^\infty}\big)^{1-\lambda} \\ &\le \frac{C(s)}{K}\big\{(KC_{\beta,\gamma})^{\frac{|\alpha|}{|\beta|}}|v|_{L^\infty}|u|_{s,\infty} + |v|_{s,\infty}|u|_{L^\infty}\big\}. \end{aligned} \tag{A.0.10}$$

Then (A.0.6) follows by (A.0.9), (A.0.10) taking $K := K(s)$ large enough. (vi): same proof as (v), using the elementary inequality $\|(D^\beta u)(D^\gamma v)\|_0 \le |D^\beta u|_{L^\infty}\|D^\gamma v\|_0$. $\square$

We now recall classical tame estimates for composition of functions, see [62], Section 2, pages 272–275, and [70]-I, Lemma 7 in the Appendix, pages 202–203.

A function $f : \mathbb{T}^d \times B_1 \to \mathbb{C}$, where $B_1 := \{y \in \mathbb{R}^m : |y| < 1\}$, induces the composition operator

$$\tilde f(u)(x) := f(x, u(x), Du(x), \ldots, D^p u(x)) \tag{A.0.11}$$

where $D^k u(x)$ denotes the partial derivatives $\partial_x^\alpha u(x)$ of order $|\alpha| = k$ (the number m of y-variables depends on p, d).

Lemma A.0.8. (Composition of functions) *Assume $f \in C^r(\mathbb{T}^d \times B_1)$. Then*

(i) *For all $u \in H^{r+p}$ such that $|u|_{p,\infty} < 1$, the composition operator* (A.0.11) *is well defined and*

$$\|\tilde f(u)\|_r \le C\|f\|_{C^r}(\|u\|_{r+p} + 1)$$

where the constant C depends on r, d, p. If $f \in C^{r+2}$, then, for all $|u|_{p,\infty}$, $|h|_{p,\infty} < 1/2$,

$$\begin{aligned} \big\|\tilde f(u+h) - \tilde f(u)\big\|_r &\le C\|f\|_{C^{r+1}}\left(\|h\|_{r+p} + |h|_{p,\infty}\|u\|_{r+p}\right), \\ \big\|\tilde f(u+h) - \tilde f(u) - \tilde f'(u)[h]\big\|_r &\le C\|f\|_{C^{r+2}}\,|h|_{p,\infty}(\|h\|_{r+p} + |h|_{p,\infty}\|u\|_{r+p})\,. \end{aligned}$$

(ii) *The previous statement also holds replacing $\|\ \|_r$ with the norms $|\ |_{r,\infty}$.*

Lemma A.0.9. (Lipschitz estimate on parameters) *Let $d \in \mathbb{N}$, $d/2 < s_0 \leq s$, $p \geq 0$, $\gamma > 0$. Let F be a C^1-map satisfying the tame estimates: $\forall \|u\|_{s_0+p} \leq 1$, $h \in H^{s+p}$,*

$$\|F(u)\|_s \leq C(s)(1 + \|u\|_{s+p}) \,, \tag{A.0.12}$$
$$\|\partial_u F(u)[h]\|_s \leq C(s)(\|h\|_{s+p} + \|u\|_{s+p}\|h\|_{s_0+p}) \,. \tag{A.0.13}$$

For $\Lambda \subset \mathbb{R}^\nu$, let $u(\lambda)$ be a Lipschitz family of functions with $\|u\|^{\mathrm{Lip}(\gamma)}_{s_0+p} \leq 1$ (see (3.0.3)). Then

$$\|F(u)\|^{\mathrm{Lip}(\gamma)}_s \leq C(s)\big(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+p}\big).$$

The same statement also holds when all the norms $\| \ \|_s$ are replaced by $| \ |_{s,\infty}$.

Proof. By (A.0.12) we get $\sup_\lambda \|F(u(\lambda))\|_s \leq C(s)(1 + \|u\|^{\mathrm{Lip}(\gamma)}_{s+p})$. Then, denoting $u_1 := u(\lambda_1)$ and $h := u(\lambda_2) - u(\lambda_1)$, we have

$$\begin{aligned}
\|F(u_2) - F(u_1)\|_s &\leq \int_0^1 \|\partial_u F(u_1 + t(u_2 - u_1))[h]\,\|_s \, dt \\
&\overset{(A.0.13)}{\leq_s} \|h\|_{s+p} + \|h\|_{s_0+p} \int_0^1 \big((1-t)\|u(\lambda_1)\|_{s+p} + t\|u(\lambda_2)\|_{s+p}\big)\, dt
\end{aligned}$$

whence

$$\begin{aligned}
\gamma \sup_{\substack{\lambda_1,\lambda_2 \in \Lambda \\ \lambda_1 \neq \lambda_2}} \frac{\|F(u(\lambda_1)) - F(u(\lambda_2))\|_s}{|\lambda_1 - \lambda_2|} &\leq_s \|u\|^{\mathrm{Lip}(\gamma)}_{s+p} + \|u\|^{\mathrm{Lip}(\gamma)}_{s_0+p} \sup_{\lambda_1,\lambda_2} \big(\|u(\lambda_1)\|_{s+p} + \|u(\lambda_2)\|_{s+p}\big) \\
&\leq_s \|u\|^{\mathrm{Lip}(\gamma)}_{s+p} + \|u\|^{\mathrm{Lip}(\gamma)}_{s_0+p} \|u\|^{\mathrm{Lip}(\gamma)}_{s+p} \leq C(s)\|u\|^{\mathrm{Lip}(\gamma)}_{s+p} \,,
\end{aligned}$$

because $\|u\|^{\mathrm{Lip}(\gamma)}_{s_0+p} \leq 1$, and the lemma follows. □

The next lemma is also classical, see for example [45], Appendix G. The present version is proved in [4], except for the part on the Lipschitz dependence on a parameter, which is proved here below.

Lemma A.0.10. (Change of variable) *Let $p : \mathbb{R}^d \to \mathbb{R}^d$ be a 2π-periodic function in $W^{s,\infty}$, $s \geq 1$, with $|p|_{1,\infty} \leq 1/2$. Let $f(x) = x + p(x)$. Then:*

(i) f is invertible, its inverse is $f^{-1}(y) = g(y) = y + q(y)$ where q is 2π-periodic, $q \in W^{s,\infty}(\mathbb{T}^d, \mathbb{R}^d)$, and $|q|_{s,\infty} \leq C|p|_{s,\infty}$. More precisely,

$$|q|_{L^\infty} = |p|_{L^\infty}, \quad |Dq|_{L^\infty} \leq 2|Dp|_{L^\infty}, \quad |Dq|_{s-1,\infty} \leq C|Dp|_{s-1,\infty}. \tag{A.0.14}$$

where the constant C depends on d, s.

Moreover, suppose that $p = p_\lambda$ depends in a Lipschitz way by a parameter $\lambda \in \Lambda \subset \mathbb{R}^\nu$, and suppose, as above, that $|D_x p_\lambda|_{L^\infty} \leq 1/2$ for all λ. Then $q = q_\lambda$ is also Lipschitz in λ, and

$$|q|^{\mathrm{Lip}(\gamma)}_{s,\infty} \leq C\Big(|p|^{\mathrm{Lip}(\gamma)}_{s,\infty} + \big\{\sup_{\lambda \in \Lambda} |p_\lambda|_{s+1,\infty}\big\}\, |p|^{\mathrm{Lip}(\gamma)}_{L^\infty}\Big) \leq C|p|^{\mathrm{Lip}(\gamma)}_{s+1,\infty} \,. \tag{A.0.15}$$

The constant C depends on d, s (and is independent of γ).

(ii) If $u \in H^s(\mathbb{T}^d, \mathbb{C})$, then $u \circ f(x) = u(x + p(x))$ is also in H^s, and, with the same C as in (i),

$$\|u \circ f\|_s \le C(\|u\|_s + |Dp|_{s-1,\infty}\|u\|_1), \tag{A.0.16}$$

$$\|u \circ f - u\|_s \le C\big(|p|_{L^\infty}\|u\|_{s+1} + |p|_{s,\infty}\|u\|_2\big), \tag{A.0.17}$$

$$\|u \circ f\|_s^{\mathrm{Lip}(\gamma)} \le C\left(\|u\|_{s+1}^{\mathrm{Lip}(\gamma)} + |p|_{s,\infty}^{\mathrm{Lip}(\gamma)}\|u\|_2^{\mathrm{Lip}(\gamma)}\right). \tag{A.0.18}$$

(A.0.16), (A.0.17) (A.0.18) *also hold for $u \circ g$.*

(iii) Part (ii) also holds with $\|\cdot\|_k$ replaced by $|\cdot|_{k,\infty}$, and $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$ replaced by $|\cdot|_{s,\infty}^{\mathrm{Lip}(\gamma)}$, namely

$$|u \circ f|_{s,\infty} \le C(|u|_{s,\infty} + |Dp|_{s-1,\infty}|u|_{1,\infty}), \tag{A.0.19}$$

$$|u \circ f|_{s,\infty}^{\mathrm{Lip}(\gamma)} \le C(|u|_{s+1,\infty}^{\mathrm{Lip}(\gamma)} + |Dp|_{s-1,\infty}^{\mathrm{Lip}(\gamma)}|u|_{2,\infty}^{\mathrm{Lip}(\gamma)}). \tag{A.0.20}$$

Proof. The bounds (A.0.14), (A.0.16) and (A.0.19) are proved in [4], Appendix B. Let us prove (A.0.15). Denote $p_\lambda(x) := p(\lambda, x)$, and similarly for $q_\lambda, g_\lambda, f_\lambda$. Since $y = f_\lambda(x) = x + p_\lambda(x)$ if and only if $x = g_\lambda(y) = y + q_\lambda(y)$, one has

$$q_\lambda(y) + p_\lambda(g_\lambda(y)) = 0\,, \quad \forall \lambda \in \Lambda,\ y \in \mathbb{T}^d. \tag{A.0.21}$$

Let $\lambda_1, \lambda_2 \in \Lambda$, and denote, in short, $q_1 = q_{\lambda_1}$, $q_2 = q_{\lambda_2}$, and so on. By (A.0.21),

$$\begin{aligned} q_1 - q_2 &= p_2 \circ g_2 - p_1 \circ g_1 = (p_2 \circ g_2 - p_1 \circ g_2) + (p_1 \circ g_2 - p_1 \circ g_1) \\ &= G_2(p_2 - p_1) + \int_1^2 G_t(D_x p_1)\, dt\, (q_2 - q_1) \end{aligned} \tag{A.0.22}$$

where $G_2 h := h \circ g_2$, $G_t h := h \circ \big(g_1 + (t-1)[g_2 - g_1]\big)$, $t \in [1, 2]$. By (A.0.22), the L^∞ norm of $(q_2 - q_1)$ satisfies

$$|q_2 - q_1|_{L^\infty} \le |G_2(p_2 - p_1)|_{L^\infty} + \int_1^2 |G_t(D_x p_1)|_{L^\infty}\, dt\, |q_2 - q_1|_{L^\infty} \le |p_2 - p_1|_{L^\infty} + |D_x p_1|_{L^\infty}\, |q_2 - q_1|_{L^\infty}$$

whence, using the assumption $|D_x p_1|_{L^\infty} \le 1/2$, we get $|q_2 - q_1|_{L^\infty} \le 2|p_2 - p_1|_{L^\infty}$. By (A.0.22), using (A.0.6), the $W^{s,\infty}$ norm of $(q_2 - q_1)$, for $s \ge 0$, satisfies

$$|q_1 - q_2|_{s,\infty} \le |G_2(p_2 - p_1)|_{s,\infty} + \frac{3}{2} \int_1^2 |G_t(D_x p_1)|_{L^\infty}\, dt\, |q_2 - q_1|_{s,\infty} + C(s) \int_1^2 |G_t(D_x p_1)|_{s,\infty}\, dt\, |q_2 - q_1|_{L^\infty}.$$

Since $|G_t(D_x p_1)|_{L^\infty} = |D_x p_1|_{L^\infty} \le 1/2$,

$$\Big(1 - \frac{3}{4}\Big)|q_2 - q_1|_{s,\infty} \le |G_2(p_2 - p_1)|_{s,\infty} + C(s) \int_1^2 |G_t(D_x p_1)|_{s,\infty}\, dt\, |q_2 - q_1|_{L^\infty}.$$

Using $|q_2 - q_1|_{L^\infty} \le 2|p_2 - p_1|_{L^\infty}$, (A.0.19), (A.0.4) and (A.0.14),

$$|q_2 - q_1|_{s,\infty} \le C(s)\Big(|p_2 - p_1|_{s,\infty} + \big\{\sup_{\lambda \in \Lambda} |p_\lambda|_{s+1,\infty}\big\}|p_2 - p_1|_{L^\infty}\Big)$$

and (A.0.15) follows. The proof of (A.0.17), (A.0.18), (A.0.20) may be obtained similarly. $\square$

Lemma A.0.11. (Composition) *Suppose that for all* $\|u\|_{s_0+\mu_i} \leq 1$ *the operator* $\mathcal{Q}_i(u)$ *satisfies*

$$\|\mathcal{Q}_i h\|_s \leq C(s)\big(\|h\|_{s+\tau_i} + \|u\|_{s+\mu_i}\|h\|_{s_0+\tau_i}\big), \quad i = 1, 2. \tag{A.0.23}$$

Let $\tau := \max\{\tau_1, \tau_2\}$, $\mu := \max\{\mu_1, \mu_2\}$. *Then, for all*

$$\|u\|_{s_0+\tau+\mu} \leq 1\,, \tag{A.0.24}$$

the composition operator $\mathcal{Q} := \mathcal{Q}_1 \circ \mathcal{Q}_2$ *satisfies the tame estimate*

$$\|\mathcal{Q}h\|_s \leq C(s)\big(\|h\|_{s+\tau_1+\tau_2} + \|u\|_{s+\tau+\mu}\|h\|_{s_0+\tau_1+\tau_2}\big). \tag{A.0.25}$$

Moreover, if $\mathcal{Q}_1$, $\mathcal{Q}_2$, u *and* h *depend in a Lipschitz way on a parameter* λ, *then* (A.0.25) *also holds with* $\|\cdot\|_s$ *replaced by* $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$.

Proof. Apply the estimates for (A.0.23) to $\mathcal{Q}_1$ first, then to $\mathcal{Q}_2$, using condition (A.0.24). $\square$

Bibliography

[1] Alazard T., Baldi P., *Gravity capillary standing water waves*, arXiv:1405.1934, preprint (2014).

[2] Arnold V. I., *Small denominators and problems of stability of motion in classical celestial mechanics*, Russian Math. Surveys 18, 85-193 (1963).

[3] Baldi P., *Periodic solutions of forced Kirchhoff equations*, Ann. Scuola Norm. Sup. Pisa, Cl. Sci. (5) 8, 117-141 (2009).

[4] Baldi P., *Periodic solutions of fully nonlinear autonomous equations of Benjamin-Ono type*, Ann. I. H. Poincaré (C) Anal. Non Linéaire 30 no. 1, 33-77 (2013).

[5] Baldi P., Berti M., Montalto R., *A note on KAM theory for quasi-linear and fully nonlinear KdV*, Rend. Lincei Mat. Appl. 24, 437-450 (2013).

[6] Baldi P., Berti M., Montalto R., *KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation*, Math. Annalen 359, 471-536 (2014).

[7] Baldi P., Berti M., Montalto R., *KAM for quasi-linear KdV*, C. R. Acad. Sci. Paris Ser. I 352, 603-607 (2014).

[8] Baldi P., Berti M., Montalto R., *KAM for autonomous quasi-linear perturbations of KdV*, arXiv:1404.3125, preprint (2014).

[9] Baldi P., Berti M., Montalto R., *KAM for autonomous quasi-linear perturbations of modified KdV*, preprint (2014).

[10] Bambusi D., Graffi S., *Time quasi-periodic unbounded perturbations of Schrödinger operators and KAM methods*, Commun. Math. Phys. 219, 465-480 (2001).

[11] Bättig D., Bloch A. M., Guillot J. C., Kappeler T., *On the symplectic structure of the phase space for periodic KdV, Toda, and defocusing NLS*. Duke Math. J. 79, 549-604 (1995).

[12] Bättig D., Kappeler T., Mityagin B., *On the Korteweg-de Vries equation: Convergent Birkhoff normal form*. J. Funct. Anal. 140, 335-358 (1996).

[13] Berti M., *Nonlinear oscillations of Hamiltonian PDEs*, Progress in Nonlin. Diff. Eq. Appl. 74, Birkhäuser Boston, (2008).

[14] Berti M., Biasco L., *Branching of Cantor manifolds of elliptic tori and applications to PDEs*, Comm. Math. Phys. 305 no. 3, 741-796 (2011).

[15] Berti M., Biasco P., Procesi M., *KAM theory for the Hamiltonian DNLW*, Ann. Sci. Éc. Norm. Supér. (4) 46 fascicule 2, 301-373 (2013).

[16] Berti M., Biasco P., Procesi M., *KAM theory for the reversible derivative wave equation*, Arch. Rational Mech. Anal. 212, 905-955 (2014).

[17] Berti M., Bolle P., *Cantor families of periodic solutions for completely resonant nonlinear wave equations*, Duke Math. J. 134, 359-419 (2006).

[18] Berti M., Bolle P., *Cantor families of periodic solutions of wave equations with C^k nonlinearities*, NoDEA 15, 247-276 (2008).

[19] Berti M., Bolle P., Procesi M., *An abstract Nash-Moser theorem with parameters and applications to PDEs*, Ann. I. H. Poincaré 27, 377-399 (2010).

[20] Berti M., Bolle P., *Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential*, Nonlinearity 25, 2579-2613 (2012).

[21] Berti M., Bolle P., *Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential*, Eur. Jour. Math. 15, 229-286 (2013).

[22] Berti M., Bolle P., *A Nash-Moser approach to KAM theory*, preprint (2014).

[23] Berti M., Bolle P., *Quasi-periodic solutions for autonomous NLW on $\mathbb{T}^d$ with a multiplicative potential*, in preparation.

[24] Berti M., Corsi L., Procesi M., *An abstract Nash-Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous spaces*, Comm. Math. Phys., DOI: 10.1007/s00220-014-2128-4, (2014).

[25] Bourgain J., *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices, no. 11 (1994).

[26] Bourgain J., *Gibbs measures and quasi-periodic solutions for nonlinear Hamiltonian partial differential equations*, Gelfand Math. Sem., Birkhäuser, Boston, MA, 23-43 (1996).

[27] Bourgain J., *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Annals of Math. 148, 363-439 (1998).

[28] Bourgain J., *Periodic solutions of nonlinear wave equations*, Harmonic analysis and partial differential equations, Chicago Lectures in Math., Univ. Chicago Press, 69-97 (1999).

[29] Bourgain J., *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton (2005).

[30] Chierchia L., You J., *KAM Tori for 1D Nonlinear Wave Equations with Periodic Boundary Conditions*, Commun. Math. Phys. 211, 497-525 (2000).

[31] Craig W., *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthèses 9, Société Mathématique de France, Paris (2000).

[32] Craig. W., *Transformation theory of Hamiltonian PDEs and the problem of water waves*, proceedings of the Advanced Study Institute on Hamiltonian Dynamical Systems and Applications, NATO Science for Peace and Security Series B: Springer - Verlag, 67-83 (2008).

[33] Craig W., Nicholls D., *Travelling two and three dimensional capillary gravity water waves*, SIAM J. Math. Anal. 32, 323-359 (2000).

[34] Craig W., Wayne C. E., *Newton's method and periodic solutions of nonlinear wave equation*, Comm. Pure Appl. Math. 46, 1409-1498 (1993).

[35] Craig W., Sulem C., *Numerical simulation of gravity waves*, J. Comput. Phys. 108 no. 1, 73-83 (1993).

[36] Eliasson L.H., *Perturbations of stable invariant tori for Hamiltonian systems*, Annali della scuola normale superiore di Pisa, Cl. Sci. (4) 15 no. 1, 115-147 (1988) .

[37] Eliasson H., Kuksin S., *Four lectures on KAM for the non-linear Schrödinger equation*, Hamiltonian dynamical systems and applications (W. Craig editor), NATO advanced studies institute, Montreal, 179-212 (2007).

[38] Eliasson L. H., Kuksin S., *On reducibility of Schrödinger equations with quasiperiodic in time potentials*, Comm. Math. Phys. 286, 125-135 (2009).

[39] Eliasson L.H., Kuksin S., *KAM for non-linear Schrödinger equation*, Annals of Math. 172, 371-435 (2010).

[40] Geng J., Xu X., You J., *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation*, Adv. Math. 226 no. 6, 5361-5402 (2011).

[41] Gentile G., Mastropietro V., Procesi M., *Periodic solutions for completely resonant nonlinear wave equations with Dirichlet boundary conditions*, Comm. Math. Phys. 256 no. 2, 437-490 (2005).

[42] Gentile G., Procesi M., *Periodic solutions for a class of nonlinear partial differential equations in higher dimension*, Comm. Math. Phys. 289 no. 3, 863-906 (2009).

[43] Grebert B., Thomann L., *KAM for the quantum harmonic oscillator*, Comm. Math. Phys. 307 no. 2, 383-427 (2011).

[44] Huang G., Kuksin S., *KdV equation under periodic boundary conditions and its perturbations*, to appear on Nonlinearity, arXiv:1309.1597, preprint (2014)

[45] Iooss G., Plotnikov P.I., Toland J.F., *Standing waves on an infinitely deep perfect fluid under gravity*, Arch. Ration. Mech. Anal. 177 no. 3, 367-478 (2005).

[46] Iooss G., Plotnikov P.I., *Small divisor problem in the theory of three-dimensional water gravity waves*, Mem. Amer. Math. Soc. 200 no. 940 (2009).

[47] Iooss G., Plotnikov P.I., *Asymmetrical three-dimensional travelling gravity waves*, Arch. Ration. Mech. Anal. 200 no. 3, 789-880 (2011).

[48] Kappeler T., Marakov M., *On Birkhoff coordinates for KdV*, Ann. I. H. Poincaré 2, 807-856 (2001).

[49] Kappeler T., Pöschel J., *KAM and KdV*, Springer (2003).

[50] Kappeler T., Schaad B., Topalov P., *mKdV and its Birkhoff coordinates*, Physica D 237, 1655-1662 (2008).

[51] Klainermann S., Majda A., *Formation of singularities for wave equations including the nonlinear vibrating string*, Comm. Pure Appl. Math. 33, 241-263 (1980).

[52] Kolmogorov A.N., *On the persistence of conditionally periodic motions under a small perturbation of the Hamiltonian function*, Dokl. Akad. Nauk. SSSR 98, 527-530 (1954).

[53] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional. Anal. i Prilozhen. 21 no. 3, 22-37, 95 (1987).

[54] Kuksin S., *Perturbation of quasiperiodic solutions of infinite-dimensional Hamiltonian systems*, Mat USSR Izvestiya 32, 39-62 (1989).

[55] Kuksin S., *A KAM theorem for equations of the Korteweg-de Vries type*, Rev. Math. Math Phys. 10 no. 3, 1-64 (1998).

[56] Kuksin S., *Analysis of Hamiltonian PDEs*, Oxford Lecture Series in Mathematics and its Applications 19, Oxford University Press (2000).

[57] Kuksin S., Pöschel J., *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Annals of Math. 143 no. 2, 149-179 (1996).

[58] Lax P., *Development of singularities of solutions of nonlinear hyperbolic partial differential equations*, J. Mathematical Phys. 5, 611-613 (1964).

[59] Liu J., Yuan X., *Spectrum for quantum Duffing oscillator and small-divisor equation with large-variable coefficient*, Comm. Pure Appl. Math. 63 no. 9, 1145-1172 (2010) .

[60] Liu J., Yuan X., *A KAM Theorem for Hamiltonian Partial Differential Equations with Unbounded Perturbations*, Comm. Math. Phys. 307, 629-673 (2011).

[61] Moser J., *A new technique for the construction of solutions of nonlinear differential equations*, Proc. Nat. Acad. Sci. 47, 1824-1831 (1961).

[62] Moser J., *A rapidly convergent iteration method and non-linear partial differential equations - I*, Ann. Sc. Norm. Sup. Pisa, Cl. Sci. (3) 20 no. 2, 265-315 (1966).

[63] Moser J., *Convergent series expansion for quasi-periodic motions*, Math. Ann. 169, 136-176 (1967).

[64] Pöschel J., *On elliptic lower dimensional tori in Hamiltonian systems*, Math Z. 202, 559-608 (1989).

[65] Pöschel J., *A KAM-Theorem for some nonlinear partial differential equations*, Ann. Sc. Norm. Sup. Pisa, Cl. Sci. (4) 23, 119-148 (1996).

[66] Pöschel J., *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv. 71 no. 2, 269-296 (1996).

[67] Procesi M., Procesi C., *A normal form for the Schrödinger equation with analytic non-linearities*, Comm. Math. Phys. 312, 501-557 (2012).

[68] Procesi C., Procesi M., *A KAM algorithm for the completely resonant nonlinear Schrödinger equation*, arXiv: 1211.4242v3, preprint (2013).

[69] Procesi M., Xu X., *Quasi-Töplitz Functions in KAM Theorem*, SIAM J.Math. Anal. 45 no. 4, 2148-2181 (2013).

[70] Rabinowitz P.H., *Periodic solutions of nonlinear hyperbolic partial differential equations*, Part I and II, Comm. Pure Appl. Math. 20, 145-205 (1967), and 22, 15-39 (1969).

[71] Taylor M. E., *Pseudodifferential Operators and Nonlinear PDEs*, Progress in Mathematics, Birkhäuser (1991).

[72] Wang W. M., *Supercritical nonlinear Schrödinger equations I: quasi-periodic solutions*, arXiv:1007.0156, preprint (2010).

[73] Wayne E., *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. 127, 479-528 (1990).

[74] Zakharov V. E., *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, Journal of Applied Mechanics and Technical Physics 9 no. 2, 190-194 (1968).

[75] Zhang J., Gao M., Yuan X., *KAM tori for reversible partial differential equations*, Nonlinearity 24, 1189-1228 (2011).

[76] Zehnder E., *Generalized implicit function theorems with applications to some small divisors problems I-II*, Comm. Pure Appl. Math. 28, 91-140 (1975), and 29, 49-113 (1976).