Available online at www.sciencedirect.com

SciVerse ScienceDirect

Ann. I. H. Poincaré – AN 30 (2013) 33–77

ANNALES DE L'INSTITUT HENRI POINCARÉ
ANALYSE NON LINÉAIRE

www.elsevier.com/locate/anihpc

Periodic solutions of fully nonlinear autonomous equations of Benjamin–Ono type

Pietro Baldi

Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università di Napoli Federico II, Via Cintia, Monte S. Angelo, 80126 Napoli, Italy

Received 24 February 2012; accepted 1 June 2012

Available online 16 June 2012

Abstract

We prove the existence of time-periodic, small amplitude solutions of autonomous quasi-linear or fully nonlinear completely resonant pseudo-PDEs of Benjamin–Ono type in Sobolev class. The result holds for frequencies in a Cantor set that has asymptotically full measure as the amplitude goes to zero.

At the first order of amplitude, the solutions are the superposition of an arbitrarily large number of waves that travel with different velocities (multimodal solutions).

The equation can be considered as a Hamiltonian, reversible system plus a non-Hamiltonian (but still reversible) perturbation that contains derivatives of the highest order.

The main difficulties of the problem are: an infinite-dimensional bifurcation equation, and small divisors in the linearized operator, where also the highest order derivatives have non-constant coefficients.

The main technical step of the proof is the reduction of the linearized operator to constant coefficients up to a regularizing rest, by means of changes of variables and conjugation with simple linear pseudo-differential operators, in the spirit of the method of Iooss, Plotnikov and Toland for standing water waves (ARMA 2005). Other ingredients are a suitable Nash–Moser iteration in Sobolev spaces, and Lyapunov–Schmidt decomposition.

Résumé

Nous démontrons l'existence de solutions de petite amplitude, périodiques en temps, de versions quasi-linéaires ou complètement non linéaires de l'équation de Benjamin–Ono, dans le cas complètement résonnant. Le résultat est démontré dans l'échelle de Sobolev, pour des fréquences dans un ensemble de Cantor de mesure asymptotiquement pleine à l'origine. Nous considérons le cas général où l'équation peut-être vue comme un système Hamiltonien réversible perturbé par une partie réversible, mais qui n'est pas Hamiltonienne, contenant des dérivées d'ordre maximal.

Ces solutions sont, au premier ordre, obtenues par superposition d'un nombre arbitrairement grand d'ondes se propageant à des vitesses différentes (solutions multimodales).

Les principales difficultés du problème sont : la présence d'un noyau de dimension infinie pour l'équation de bifurcation, et l'occurence de petits diviseurs dans la résolution de l'équation linéarisée, où les dérivées de plus haut degré ont des coefficients variables.

Nous montrons que l'opérateur linéarisé est essentiellement conjugué à un opérateur à coefficients constants, modulo un terme régularisant. La démonstration, basée sur des changements de variables et des conjugaisons avec des opérateurs pseudo-différentiels,

E-mail address: pietro.baldi@unina.it.

0294-1449/$ – see front matter
http://dx.doi.org/10.1016/j.anihpc.2012.06.001

est inspirée par la méthode utilisée par Iooss, Plotnikov et Toland (ARMA 2005) pour démontrer l'existence d'ondes de gravité stationnaires. La démonstration utilise également un schéma de Nash–Moser adapté à ce contexte, dans l'échelle des espaces de Sobolev, ainsi qu'une décomposition de Lyapunov–Schmidt.

MSC: 35B10; 37K55; 37K50

Keywords: Benjamin–Ono equation; Fully nonlinear PDEs; Quasi-linear PDEs; Pseudo-PDEs; Periodic solutions; Small divisors; Nash–Moser method; Infinite-dimensional dynamical systems; Reversible dynamical systems

1. The problem and main result

We consider autonomous equations of Benjamin–Ono type

$$u_t + \mathcal{H}u_{xx} + \partial_x\big(u^3\big) + \mathcal{N}_4(u) = 0 \tag{1.1}$$

with periodic boundary conditions $x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$, where the unknown $u(t,x)$ is a real-valued function, $t \in \mathbb{R}$, $\mathcal{H}$ is the periodic Hilbert transform, namely the Fourier multiplier

$$\mathcal{H}e^{ijx} = -i\,\mathrm{sign}(j)e^{ijx}, \quad j \in \mathbb{Z},$$

and $\mathcal{N}_4$ is of type (I) or (II),

$$\text{(I)} \quad \mathcal{N}_4(u) = g_1(x, u, \mathcal{H}u, u_x) + \partial_x\big(g_2(x, u, \mathcal{H}u_x)\big), \tag{1.2}$$

$$\text{(II)} \quad \mathcal{N}_4(u) = g_0(x, u, \mathcal{H}u, u_x, \mathcal{H}u_{xx}). \tag{1.3}$$

(1.1) is a *quasi-linear* problem in case (I) and a *fully nonlinear* problem in case (II).

We assume that the function $g_i(x, y)$ is defined for $y = (y_1, \dots, y_n)$ in the ball $B_1 = \{|y| < 1\}$ of $\mathbb{R}^n$, $n = 2, 3, 4$, g_i is 2π-periodic in the real variable x, and, together with its derivatives in y up to order 4, it is of class C^r in all its arguments (x, y), with

$$\sum_{0 \leqslant |\alpha| \leqslant 4} \big\|\partial_y^\alpha g_i\big\|_{C^r(\mathbb{T}\times B_1)} \leqslant K_{g,r}, \tag{1.4}$$

for some constant $K_{g,r} > 0$. Moreover we assume that at $y = 0$

$$\partial_y^\alpha g_i(x, 0) = 0 \quad \forall \alpha \in \mathbb{N}^n,\ |\alpha| \leqslant 3, \tag{1.5}$$

so that, regarding the amplitude, $\mathcal{N}_4(\varepsilon u) = O(\varepsilon^4)$ as $\varepsilon \to 0$.

We assume that the nonlinearity $\mathcal{N}(u) := \partial_x(u^3) + \mathcal{N}_4(u)$ behaves like the linear part $\partial_t + \mathcal{H}\partial_{xx}$ with respect to the parity of functions $u(t, x)$ in the time–space pair (t, x). This means to assume the *reversibility conditions*

$$g_1(-x, y_1, -y_2, -y_3) = -g_1(x, y_1, y_2, y_3), \qquad g_2(-x, y_1, y_2) = g_2(x, y_1, y_2), \tag{1.6}$$

$$g_0(-x, y_1, -y_2, -y_3, -y_4) = -g_0(x, y_1, y_2, y_3, y_4), \tag{1.7}$$

so that in both cases (I) and (II) $\mathcal{N}(u)$ is odd for all even u, namely

$$u(-t, -x) = u(t, x) \quad \Rightarrow \quad \mathcal{N}(u)(-t, -x) = -\mathcal{N}(u)(t, x). \tag{1.8}$$

Assumptions (1.2), (1.3), (1.6), (1.7) are discussed in the next section.

Remark 1.1. Examples of such nonlinearities are:

$$\text{(I)} \quad \mathcal{N}_4(u) = (\mathcal{H}u_x)^3\mathcal{H}u_{xx} + a(x)u_x^4 + uu_x^3 + b(x)u_x^5, \qquad \text{(II)} \quad \mathcal{N}_4(u) = a(x)(\mathcal{H}u_{xx})^4 + u_x^5,$$

where $a(x)$ is odd and $b(x)$ is even.

We construct small amplitude time-periodic solutions $u(t,x)$ of period $T = 2\pi/\omega$, $\omega > 0$, where the period T is also an unknown of the problem. Rescaling the time $t \to \omega t$, this is equivalent to find 2π-periodic solutions of the equation

$$\omega u_t + \mathcal{H} u_{xx} + \partial_x\big(u^3\big) + \mathcal{N}_4(u) = 0, \tag{1.9}$$

with $u : \mathbb{T}^2 \to \mathbb{R}$, $\omega > 0$.

Regarding the time–space pair (t,x) as a point of the 2-dimensional torus $\mathbb{T}^2$, we consider the L^2-based Sobolev space of real-valued periodic functions

$$H^s = H^s\big(\mathbb{T}^2;\mathbb{R}\big) = \Big\{ u = \sum_{k\in\mathbb{Z}^2} u_k e_k \colon u_{-k} = \bar u_k \in \mathbb{C},\ \|u\|_s^2 := \sum_{k\in\mathbb{Z}^2} |u_k|^2 \langle k\rangle^{2s} < \infty \Big\}, \tag{1.10}$$

where $s \geqslant 0$, $\langle k\rangle := \max\{1, |k|\}$, and $e_k(t,x) := e^{i(k_1 t + k_2 x)}$.

The main result of the paper is the following theorem.

Theorem 1.2. *There exist universal constants $r_0, s_0, c_0 \in \mathbb{N}$ with the following properties.*

Assume hypotheses (1.2)–(1.7) on the nonlinearity $\mathcal{N}$, with $r \geqslant r_0$. Let $m \geqslant 2$ and let $0 < k_1 < k_2 < \cdots < k_m$ be m positive integers that satisfy

$$k_1 + \cdots + k_{m-1} > k_m(m - 3/2), \qquad k_1 + \cdots + k_m \neq (m - 1/2) j \quad \forall j \in \mathbb{N}. \tag{1.11}$$

Then there exist

(i) *a trigonometric polynomial*

$$\bar v_1(t,x) := \sum_{j=1}^{m} a_j \cos\big(k_j x - k_j^2 t\big),$$

even in the pair (t,x), where $a_j \in \mathbb{R}$,

$$a_j^2 = \frac{4}{m - 1/2}\left(\sum_{i=1}^{m} k_i\right) - 4k_j, \qquad j = 1, \ldots, m;$$

(ii) *constants $C, \varepsilon_0^* > 0$ that depend on $k_1, \ldots, k_m, K_{g,r_0}$;*

(iii) *a measurable Cantor-like set $\mathcal{G} \subset (0, \varepsilon_0^*)$ of asymptotically full Lebesgue measure, namely*

$$\frac{|\mathcal{G} \cap (0,\varepsilon_0)|}{\varepsilon_0} \geqslant 1 - \varepsilon_0 C \quad \forall \varepsilon_0 \leqslant \varepsilon_0^*,$$

such that for every $\varepsilon \in \mathcal{G}$ problem (1.9) with frequency

$$\omega = 1 + 3\varepsilon^2$$

has a solution $u_\varepsilon \in H^{s_0}(\mathbb{T}^2, \mathbb{R})$ that satisfies

$$\|u_\varepsilon - \varepsilon \bar v_1\|_{s_0} \leqslant \varepsilon^2 C, \qquad u_\varepsilon(-t,-x) = u_\varepsilon(t,x), \qquad \int_{\mathbb{T}^2} u_\varepsilon(t,x)\, dt\, dx = 0.$$

Moreover $u_\varepsilon \in H^s(\mathbb{T}^2)$ for every s in the interval $s_0 \leqslant s < (r + c_0)/2$.

If g_i, $i = 0,1,2$ in (1.2), (1.3) is of class C^∞, then also $u_\varepsilon \in C^\infty(\mathbb{T}^2)$.

Remark 1.3. (i) The smallest example of $k_1, \ldots, k_m$ satisfying (1.11) is $m = 2$, $k_1 = 2$, $k_2 = 3$. For every $m \geqslant 2$ there exist infinitely many choices of integers $k_1 < \cdots < k_m$ that satisfy (1.11). See also Remark 5.2.

(ii) s_0, r_0 and c_0 can be explicitly calculated: $s_0 = 22$, $c_0 = 28$ (non-sharp calculation); for r_0 see (9.22) and the lines below it.

2. Motivations, questions and comments

The original Benjamin–Ono equation

$$u_t + \mathcal{H} u_{xx} + u u_x = 0 \tag{2.1}$$

models one-dimensional internal waves in deep water [5], and is a completely integrable [1] Hamiltonian partial pseudo-differential equation,

$$\partial_t u = J \nabla H(u), \quad J = -\partial_x, \ H(u) = \int \left(\frac{u \mathcal{H} u_x}{2} + \frac{u^3}{6} \right) dx.$$

The local and global well-posedness in Sobolev class for (2.1) and many generalizations of it (other powers $u^p u_x$, other linear terms $\partial_x |D_x|^\alpha u$, $1 < \alpha < 2$, etc.) have been studied by several authors in the last years: see for example Molinet, Saut and Tzvetkov [29], Colliander, Kenig and Staffilani [12], Tao [35], Kenig and Ionescu [18], Burq and Planchon [11], Molinet [27,28], and the references therein. On the contrary, to the best of our knowledge, there are few works about time-periodic or quasi-periodic solutions of Benjamin–Ono equations. One of them is [2], where 2-mode periodic solutions of (2.1) are studied by numerical methods; another one is [26], which deals with an old very interesting question.

In [26] Liu and Yuan apply a Birkhoff normal form and KAM method to show the existence of quasi-periodic solutions of a Benjamin–Ono equation that is a Hamiltonian analytic perturbation of (2.1), with Hamiltonian of the form

$$H(u) + \varepsilon K(u), \qquad H = \text{Benjamin–Ono}, \qquad \nabla K(u) = \text{bounded operator}.$$

The resulting equation is of the type

$$\partial_t u = -\partial_x \left\{ \mathcal{H} u_x + \frac{1}{2} u^2 + \varepsilon \nabla K(u) \right\} = Au + F(u), \tag{2.2}$$

where the Hamiltonian vector field has a linear part A, which loses $d_A = 2$ derivatives, and a nonlinear part F, which loses $d_F = 1$ derivative and, for this reason, is an *unbounded* operator.

In general, as it was proved in the works of Lax, Klainerman and Majda on the formation of singularities (see for example [23]), the presence of unbounded nonlinear operators can compromise the existence of invariant structure like periodic orbits and KAM tori. In fact, the wide existing literature on KAM and Nash–Moser theory mainly deals with problems where the perturbation is bounded (see Kuksin [25], Craig [13], Berti [6] for a survey. See also Moser [30] where the KAM iteration is applied in problems where the Hamiltonian structure is replaced by reversibility).

For unbounded perturbations, quasi-periodic solutions have been constructed via KAM theory by Kuksin [25] and Kappeler and Pöschel [22] for KdV equations where $d_A = 3$ and the gap between the loss of derivatives of the linear and nonlinear part is $\gamma := (d_A - d_F) = 2$, in analytic class; more recently, in [26] for NLS and (2.2) where $d_A = 2$ and $\gamma = 1$, in C^∞ class; by Zhang, Gao and Yuan [36] for reversible NLS equations with $d_A = 2$ and $\gamma = 1$; and by Berti, Biasco and Procesi [7], where wave equations with a derivative in the nonlinearity become a Hamiltonian system with $d_A = 1$ and $\gamma = 1$, in analytic class. See also Bambusi and Graffi [4] for a related linear result that corresponds to a gap $\gamma > 1$.

Periodic solutions for unbounded perturbations have been obtained for wave equations by Craig [13] for $\gamma > 1$; by Bourgain [10] in the non-Hamiltonian case $u_{tt} - u_{xx} + u + u_t^2 = 0$; by the author in [3] for the quasi-linear equation $u_{tt} - \Delta u (1 + \int |\nabla u|^2 dx) = \varepsilon f(t,x)$, where the integral plays a special role ($\int |\nabla u|^2 dx$ depends only on time). Also the pioneering result of Rabinowitz [34] for fully nonlinear wave equations of the form

$$u_{tt} - u_{xx} + \alpha u_t + \varepsilon F(x, t, u, u_x, u_t, u_{xx}, u_{xt}, u_{tt}) = 0$$

certainly has to be mentioned here; however, the dissipative term $\alpha \neq 0$ destroys any Hamiltonian or reversible structure and completely avoids the resonance phenomenon of the small divisors.

The threshold $\gamma = 1$ in Hamiltonian problems with small divisors has been crossed in the works of Iooss, Plotnikov and Toland [32,21,19,20] about the completely resonant fully nonlinear ($\gamma = 0$) problem of periodic standing water waves on a deep 2D ocean with gravity. So far their very powerful technique, which is a combination of (1) changes of variables and conjugations with pseudo-differential operators to obtain a normal form, and (2) a differentiable Nash–Moser scheme, is essentially the only known method to overcome the small divisors problem in quasi-linear and fully nonlinear PDEs.

Note that recently normal form methods for quasi-linear Hamiltonian PDEs have also been successfully applied to Cauchy problems, see Delort [14].

Thus, some of the general, challenging and open questions that come from the aforementioned works are these:

- Which gap γ is the limit case for the existence of invariant tori for nonlinear Hamiltonian PDEs? How many derivatives can stay in the nonlinearity?
- What is the role of the Hamiltonian structure? Can it be replaced by other structures?

The motivations of the present paper are in these questions. Theorem 1.2 joins the above mentioned results in the aim of approaching an answer, at least in simple cases, and shows that

(i) if the dimension is the lowest for a PDE, $(t,x) \in \mathbb{T}^2$, and
(ii) the derivatives in the nonlinearity have a suitable structure (see (1.2), (1.3), (1.6), (1.7)),

then problem (1.1), where $\gamma = 0$ (the nonlinearity $\mathcal{N}(u)$ loses 2 derivatives like the linear part) admits solutions that bifurcate from the equilibrium $u = 0$. The Hamiltonian structure here is replaced by reversibility: (1.1), in general, is a *non*-Hamiltonian perturbation of the cubic Benjamin–Ono Hamiltonian equation

$$\partial_t u + \mathcal{H}\partial_{xx} u + \partial_x\big(u^3\big) = 0,$$

but $\mathcal{N}(u)$ satisfies the reversibility condition (1.8).

Let us explain the reversible structure in some detail. As a dynamical system, problem (1.1) is

$$\partial_t u(t) = V\big(u(t)\big), \tag{2.3}$$

a first order ordinary differential equation in the infinite-dimensional phase space $L^2(\mathbb{T};\mathbb{R})$, where the vector field $V : H^2(\mathbb{T};\mathbb{R}) \to L^2(\mathbb{T};\mathbb{R})$, $u \mapsto V(u)$ is

$$V(u)(x) = -\mathcal{H}\partial_{xx} u(x) - \partial_x\big(u^3(x)\big) - \mathcal{N}_4(u)(x).$$

The phase space can be split into two subspaces $L^2_e \oplus L^2_o$ of even and odd functions of $x \in \mathbb{T}$ respectively,

$$u = u^e + u^o, \qquad u^e(-x) = u^e(x), \qquad u^o(-x) = -u^o(x), \quad x \in \mathbb{T},\ u \in L^2(\mathbb{T};\mathbb{R}).$$

To decompose $u = u^e + u^o$ means to split the real and imaginary part of each Fourier coefficient of $u \in L^2(\mathbb{T};\mathbb{R})$, namely

$$u(x) = \sum_{j\in\mathbb{Z}} \hat u_j e^{ijx}, \qquad u^e(x) = \sum_{j\in\mathbb{Z}} (\operatorname{Re}\hat u_j) e^{ijx}, \qquad u^o(x) = \sum_{j\in\mathbb{Z}} i(\operatorname{Im}\hat u_j) e^{ijx}.$$

Consider the reflection

$$R : u = u^e + u^o \mapsto Ru = u^e - u^o. \tag{2.4}$$

R is an $\mathbb{R}$-linear bijection of $L^2(\mathbb{T};\mathbb{R})$, and R^2 is the identity map. In terms of Fourier coefficients,

$$R : u(x) = \sum_{j\in\mathbb{Z}} \hat u_j e^{ijx} \mapsto Ru(x) = \sum_{j\in\mathbb{Z}} \overline{\hat u_j} e^{ijx}, \tag{2.5}$$

where $\overline{\hat u_j}$ is the complex conjugate of $\hat u_j$. Note that Ru is real-valued for every real-valued u. (2.3) is a *reversible* system in the sense that

$$V \circ R = -R \circ V. \tag{2.6}$$

It is immediate to check (2.6) for the linear part $\mathcal{H}\partial_{xx}$ of V using (2.5), and for the cubic part $\partial_x(u^3)$ using (2.4). To prove (2.6) for $\mathcal{N}_4(u)$, using (1.6), (1.7) and (2.4) one has

$$\alpha(-x) = -\beta(x), \quad \alpha(x) := \mathcal{N}_4(Ru)(x),\ \beta(x) := \mathcal{N}_4(u)(x).$$

Splitting $\alpha = \alpha^e + \alpha^o$, $\beta = \beta^e + \beta^o$ and projecting the equality $\alpha(-x) = -\beta(x)$ onto L^2_e and L^2_o give $\alpha^e = -\beta^e$ and $\alpha^o = \beta^o$, namely $R\beta = -\alpha$, which is (2.6) for $\mathcal{N}_4$.

(2.6) implies that $V(u) \in L^2_o$ for all $u \in L^2_e \cap H^2$. For, L^2_e is the set of fixed points $u = Ru$, therefore $V(u) = -RV(u)$, whence $(V(u))^e = 0$.

By (2.6), if $u(t)$ solves (2.3), then also $Su(t) := R(u(-t))$ is a solution of (2.3). Thus we look for solutions of (2.3) in the subspace X of the fixed points of S. It is easy to see, using (2.4), (2.5), that X is the space of functions $u(t,x)$ that are even in the time–space pair (t,x), namely $u(-t,-x) = u(t,x)$.

To prove Theorem 1.2 we apply (and slightly modify, under certain technical aspects; see below) the method of Iooss, Plotnikov and Toland. Like in [21], the main difficulties here are: (i) in the bifurcation equation, which is infinite-dimensional (for this reason (1.1) is said to be a *completely resonant* problem); and, especially, (ii) in the inversion of the linearized operator, which has non-constant coefficients also in the highest order derivatives and, therefore, contains small divisors that are not explicitly evident.

The main tool in the inversion proof is the reduction of the linearized operator $\mathcal{L}$ to constant coefficients up to a regularizing rest, by means of changes of variables first (to obtain proportional coefficients in the highest order terms), then by the conjugation with simple linear pseudo-differential operators that imitate the structure of $\mathcal{L}$ (they are the composition of multiplication operators with the Hilbert transform $\mathcal{H}$), to obtain constant coefficients also in terms of lower order, and to lower the degree of the highest non-constant term.

Since we look for periodic solutions, after a finite number of steps this reducibility scheme implies the invertibility of $\mathcal{L}$, by standard Neumann series.

Other, and minor, technical points are the following. Like in [21], the Lyapunov–Schmidt decomposition is not used directly on the nonlinear equation, as it would be made in classical applications (see [6] for the Lyapunov–Schmidt decomposition in completely resonant problems). Instead, it is used a first time at the beginning of the proof, in a formal power series expansion of the nonlinear problem, to look for a suitable starting point of the Nash–Moser iteration. In other words, this means to find a non-degenerate solution of the "unperturbed bifurcation equation". In Theorem 1.2 the existence and the non-degeneracy conditions are the first and the second inequality in (1.11) respectively. Then the Lyapunov–Schmidt decomposition is used a second time in the inversion proof for the linearized operator, in each step of the Nash–Moser scheme.

This method seems to be more complicated than the usual Lyapunov–Schmidt decomposition on the nonlinear problem, at least at a first glance. However, it simplifies the analysis when working with changes of variables (namely compositions with diffeomorphisms of the torus $\mathbb{T}^2$). In fact, changes of variables do not behave very well with respect to the orthogonal projections onto subspaces of L^2, because they are not "close to the identity" in the same way as multiplications operators are (in the language of harmonic analysis, changes of variables are Fourier integral operators, and not pseudo-differential operators. See also Remark 7.3). For this reason, it is simpler to work in the whole function space $H^s(\mathbb{T}^2)$ instead of distinguishing bifurcation and orthogonal subspaces, at least for the first step of reducibility.

Nonetheless, in our setting (4.4) we keep track of the natural "different amount of smallness" between the bifurcation and the orthogonal components of the problem. Thanks to this small change with respect to [21], we avoid factors ε^{-1} in the Nash–Moser scheme and simplify the measure estimate for the small divisors.

Regarding the Nash–Moser scheme, the recent and powerful abstract Nash–Moser theorem for PDEs that is contained in [9] does not apply directly here, as it designed to be used with Galerkin approximations, while in our Nash–Moser scheme, after the reduction to constant coefficients, it is natural to insert the smoothing operators in a different position: see (9.5). Even if our iteration scheme is very close to the usual one, this small difference brings our problem out of the field of applicability of the theorem in [9].

Going back to the "unperturbed bifurcation equation", we point out that the restriction of the functional setting to the subspace X of even functions (a restriction that can be made because of the reversible structure) eliminates a degeneration and makes it possible to prove the non-degeneracy of the solution. Moreover, the solutions we find in Theorem 1.2 are genuinely *multimodal*: for $m = 1$ the second inequality in (1.11) is never satisfied, whereas for every $m \geqslant 2$ there exist suitable integers $k_1, \dots, k_m$ that satisfy (1.11) and produce a non-degenerate solution. This is a nonlinear effect: the solutions of Theorem 1.2 exist as a consequence of the nonlinear interaction of different modes.

Regarding the special structure (1.2),(1.3), the restriction of assuming (I) or (II), instead of considering the more general case

$$\mathcal{N}_4(u) = g(x, u, \mathcal{H}u, u_x, \mathcal{H}u_x, u_{xx}, \mathcal{H}u_{xx}), \tag{2.7}$$

is due to a technical reason: when $\mathcal{N}_4(u)$ is of the type (I) or (II), in the process of reducing the linearized operator $\mathcal{L}$ to constant coefficients we use simple transformations, namely changes of variables, multiplications, the Hilbert transform $\mathcal{H}$ and negative powers of ∂_x (which are Fourier multipliers). On the contrary, in the general case (2.7) these special transformations are not sufficient to conjugate $\mathcal{L}$ to a normal form, and one needs more general transformations: changes of variables should be replaced by general Fourier integral operators. In the intermediate case in which $\mathcal{N}_4$ in (2.7) does not depend on u_{xx} (but it does on $\mathcal{H}u_x$), an additional term of the type $b(t)\partial_x\mathcal{H}$ appears in the transformed linearized operators after the changes of variables. This term could be removed by a simple Fourier integral operator: see Remark 7.1.

Regarding the choice of the leading term $\partial_x(u^3)$ in (1.1) (which is the first natural case to study after the integrable one $\partial_x(u^2)$), we remark that the cubic power has no special reversibility property: $\partial_x(u^p)$ satisfy the reversibility condition (2.6) for every (both even and odd) power $p \in \mathbb{N}$. The proof of this fact is the same as above: if $f(u) = \partial_x(u^p)$, using (2.4) one proves that $\{f(Ru)\}(-x) = -\{f(u)\}(x)$, then $f \circ R = -R \circ f$.

Finally, the coefficient 3 in the frequency–amplitude relation $\omega = 1 + 3\varepsilon^2$ could be replaced by any other positive number: 3 is simply the most convenient choice to do when working with the cubic nonlinearity $\partial_x(u^3)$. On the contrary, what is determined by the nonlinearity in an essential way is *the sign* of that coefficient: for the equation

$$u_t + \mathcal{H}u_{xx} - \partial_x\big(u^3\big) + \mathcal{N}_4(u) = 0,$$

in which the cubic nonlinearity has opposite sign, Theorem 1.2 holds with $\omega = 1 - 3\varepsilon^2$ (the only changes to do are in the bifurcation analysis of Section 5).

The paper is organized as follows. In Section 3 the setting for the problem is introduced. In Section 4 the formal Lyapunov–Schmidt reduction is performed up to order $O(\varepsilon^4)$. In Section 5 non-degenerate solutions $\bar{v}_1$ of the “unperturbed bifurcation equation” are constructed. Here the non-homogeneous dispersion relation of the unperturbed Benjamin–Ono linear part

$$l + j|j| = 0,$$

where l is the Fourier index for the time and j the one for the space, is used in a crucial way. The basic properties of this relation are proved in Appendix A. In Sections 6 and 7 the linearized operator is reduced to constant coefficients. Most of the proofs of the related estimates are in Appendix C and use classical results of Sobolev spaces (tame estimates for changes of variables, compositions and commutators with the Hilbert transform) that are listed in Appendix B. In Section 8 the transformed linearized operator is inverted. In Section 9 the Nash–Moser induction is performed, and the measure of the Cantor set of parameters is estimated.

3. Functional setting

Let

$$\mathcal{F}(u,\omega) := \omega u_t + \mathcal{H}u_{xx} + \mathcal{N}(u), \quad \mathcal{N}(u) := \partial_x\big(u^3\big) + \mathcal{N}_4(u).$$

Let $Z := L^2(\mathbb{T}^2, \mathbb{R})$. Decompose

$$\mathbb{Z}^2 = \mathbb{Z}^2_C + \mathbb{Z}^2_T + \mathbb{Z}^2_E, \quad \mathbb{Z}^2_C = \big\{(0,0)\big\},\ \mathbb{Z}^2_T = \big\{(l,0)\colon l \neq 0\big\},\ \mathbb{Z}^2_E = \big\{(l,j)\colon j \neq 0,\ l \in \mathbb{Z}\big\},$$

let

$$Z_C = \mathbb{R}, \qquad Z_T = \left\{u \in L^2(\mathbb{T})\colon \int_0^{2\pi} u(t)\,dt = 0\right\}, \qquad Z_E = \left\{u \in Z\colon \int_0^{2\pi} u(t,x)\,dx = 0\right\},$$

so that $Z = Z_C \oplus Z_T \oplus Z_E$, namely every $u(t,x) \in Z$ splits into three components

$$u(t,x) = \left(\sum_{\mathbb{Z}^2_C} + \sum_{\mathbb{Z}^2_T} + \sum_{\mathbb{Z}^2_E}\right)\hat{u}_{l,j}e^{i(lt+jx)} = \hat{u}_{0,0} + \sum_{l\neq 0}\hat{u}_{l,0}e^{ilt} + \sum_{j\neq 0}u_j(t)e^{ijx},$$

and denote Π_C, Π_T, Π_E the projections onto Z_C, Z_T, Z_E. Let Z_0 be the space of zero-mean functions, and $\mathbb{P}$ the projection onto Z_0,

$$Z_0 := Z_T \oplus Z_E, \qquad \mathbb{P} := I - \Pi_C = \Pi_T + \Pi_E. \tag{3.1}$$

We define ∂_x^{-1} as the Fourier multiplier

$$\partial_x^{-1} e^{ijx} = \frac{1}{ij} e^{ijx} \quad \forall j \neq 0, \qquad \partial_x^{-1} 1 = 0,$$

and similarly ∂_t^{-1}. Note that $\partial_x^{-1}\partial_x = \Pi_E$, $\mathcal{H}\mathcal{H} = -\Pi_E$.

To eliminate a degeneration that appears in the bifurcation equation, as it was mentioned above where the reversible structure was discussed, we consider the subspaces of even/odd functions with respect to the time–space vector (t,x):

$$X := \{u \in Z\colon u(-t,-x) = u(t,x)\}, \qquad Y := \{u \in Z\colon u(-t,-x) = -u(t,x)\}.$$

In terms of Fourier coefficients, every $u \in Z$ is $u = \sum_{k\in\mathbb{Z}^2} u_k e_k$ with $u_{-k} = \bar{u}_k$ (because u is real-valued), namely $u_k = a_k + ib_k$, with $a_k, b_k \in \mathbb{R}$ and $a_{-k} = a_k$, $b_{-k} = -b_k$, therefore

$$X = \left\{u = \sum_{k\in\mathbb{Z}^2} a_k e_k\colon a_k \in \mathbb{R},\ a_{-k} = a_k\right\}, \qquad Y = \left\{u = \sum_{k\in\mathbb{Z}^2} ib_k e_k\colon b_k \in \mathbb{R},\ b_{-k} = -b_k\right\},$$

and $L^2(\mathbb{T}^2,\mathbb{R}) = Z = X \oplus Y$. The usual rules for even/odd functions hold: $uv \in X$ if both $u, v \in X$ or both $u, v \in Y$, and $uv \in Y$ if $u \in X$, $v \in Y$. Moreover $\mathcal{H}$, ∂_x, ∂_t are all operators that change the parity, namely they map Y into X and vice versa, because they are diagonal operators with respect to the basis $\{e_k\}$ with purely imaginary eigenvalues. Assumption (1.6) implies that the nonlinearity $\mathcal{N}$ maps $X \cap H^2$ into Y, like the linear part $\omega\partial_t + \partial_{xx}\mathcal{H}$ does, therefore $\mathcal{F}(u,\omega) \in Y$ for all $u \in X \cap H^2$.

Also, we denote

$$X_0 := X \cap Z_0,$$

while $Y \cap Z_0 = Y$. We set problem (1.9) in the space X_0 of even functions with zero mean, namely we look for solutions of the equation

$$\mathcal{F}(u,\omega) = 0, \quad u \in X_0. \tag{3.2}$$

Notation. To distinguish L^2- and L^∞-based Sobolev spaces, in the whole paper the following notation is used: two bars for L^2-based Sobolev norms $\|u\|_s$ (1.10), and one bar for L^∞-based Sobolev norms

$$|u|_s = \|u\|_{W^{s,\infty}} = \sum_{0\leqslant|\alpha|\leqslant s} \sup_{(t,x)} \left|\partial^\alpha_{(t,x)} u(t,x)\right|, \quad s \in \mathbb{N}.$$

4. Linearization at zero and formal Lyapunov–Schmidt reduction

Let

$$L := \partial_t + \partial_{xx}\mathcal{H}, \qquad L\left[e^{i(lt+jx)}\right] = i(l + j|j|)e^{i(lt+jx)}.$$

Split $\mathbb{Z}^2 = \mathcal{V} \cup \mathcal{W}$,

$$\mathcal{V} := \{(l,j) \in \mathbb{Z}^2\colon l + j|j| = 0\} = \{(-j|j|, j)\colon j \in \mathbb{Z}\}, \qquad \mathcal{W} := \mathbb{Z}^2 \setminus \mathcal{V}$$

and $Z = V \oplus W$,

$$V := \left\{u = \sum_{k\in\mathcal{V}} u_k e_k \in Z\right\}, \qquad W := \left\{u = \sum_{k\in\mathcal{W}} u_k e_k \in Z\right\}.$$

V is the kernel of L and W is its range. Also, let $V_0 := V \cap Z_0$, so that $Z_0 = V_0 \oplus W$.

We write a finite number of terms of a formal power series expansion to obtain a good starting point for our Nash–Moser scheme. Let

$$\omega = 1 + \sum_{k\geqslant1} \omega_k \varepsilon^k, \qquad u = \sum_{k\geqslant1} u_k \varepsilon^k \in Z_0, \quad u_k = v_k + w_k,\ v_k \in V_0,\ w_k \in W.$$

Then

$$\begin{aligned}\mathcal{F}(u,\omega) &= Lu + (\omega-1)\partial_t u + \partial_x\big(u^3\big) + \mathcal{N}_4(u)\\ &= \varepsilon L u_1 + \varepsilon^2\{Lu_2 + \omega_1\partial_t u_1\} + \varepsilon^3\big\{Lu_3 + \omega_1\partial_t u_2 + \omega_2\partial_t u_1 + \partial_x\big(u_1^3\big)\big\}\\ &\quad + \varepsilon^4\big\{Lu_4 + \omega_1\partial_t u_3 + \omega_2\partial_t u_2 + \omega_3\partial_t u_1 + \partial_x\big(3u_1^2 u_2\big) + \varepsilon^{-4}\mathcal{N}_4(\varepsilon u_1)\big\} + O\big(\varepsilon^5\big)\\ &= \sum_{k\geqslant 1}\varepsilon^k\mathcal{F}_k.\end{aligned}$$

In general, $\mathcal{N}_4(\varepsilon u_1)$ also contains terms of higher order than ε^4; in any case, $\mathcal{N}_4(u) - \mathcal{N}_4(\varepsilon u_1) = O(\varepsilon^5)$.

At order ε, $\mathcal{F}_1 = Lu_1 = 0$ if $w_1 = 0$ and $u_1 = v_1 \in V_0$. Then $\mathcal{F}_2$ becomes

$$\mathcal{F}_2 = Lu_2 + \omega_1\partial_t u_1 = Lw_2 + \omega_1\partial_t v_1.$$

$Lw_2 \in W$ and $\omega_1\partial_t v_1 \in V_0$. Since we look for $v_1 \neq 0$, we have $\mathcal{F}_2 = 0$ if $w_2 = 0$, $\omega_1 = 0$, $u_2 = v_2 \in V_0$.

At order ε^3 the nonlinearity begins to give a contribution: $\mathcal{F}_3 = Lw_3 + \omega_2\partial_t v_1 + \partial_x(v_1^3)$. The "unperturbed bifurcation equation" is the equation $\Pi_V\mathcal{F}_3 = 0$ in the unknown v_1, namely

$$\omega_2\partial_t v_1 + \Pi_V\partial_x\big(v_1^3\big) = 0. \tag{4.1}$$

In the next section (see Proposition 5.3) we construct nontrivial, non-degenerate solutions $\bar v_1$ of (4.1) with $\omega_2 = 3$. A solution v_1 of (4.1) for any other value $\omega_2 > 0$ can be obtained by homogeneity by taking $v_1 = \lambda\bar v_1$, $\lambda = (\omega_2/3)^{1/2}$. Hence there is no loss of generality in fixing $\omega_2 = 3$. At order ε^4,

$$\mathcal{F}_4 = Lu_4 + 3\partial_t v_2 + \omega_3\partial_t v_1 + \partial_x\big(3v_1^2 v_2\big) + \varepsilon^{-4}\mathcal{N}_4(\varepsilon v_1).$$

We fix $\omega_3 = 0$. The "linearized unperturbed bifurcation equation" is the equation $\Pi_V\mathcal{F}_4 = 0$ in the unknown v_2, namely

$$3\partial_t v_2 + \Pi_V\partial_x\big(3v_1^2 v_2\big) = -\varepsilon^{-4}\Pi_V\mathcal{N}_4(\varepsilon v_1), \tag{4.2}$$

which has a unique solution $\bar v_2(\varepsilon)$ because $\bar v_1$ is a non-degenerate solutions of (4.1). Thus, at $u = \varepsilon\bar v_1 + \varepsilon^2\bar v_2(\varepsilon)$ and $\omega = 1 + 3\varepsilon^2$,

$$\begin{aligned}\mathcal{F}\big(\varepsilon\bar v_1 + \varepsilon^2\bar v_2, 1+3\varepsilon^2\big) &= \varepsilon^3\Pi_W\partial_x\big(\bar v_1^3\big) + \varepsilon^4\Pi_W\partial_x\big(3\bar v_1^2\bar v_2\big) + \mathcal{N}_4\big(\varepsilon\bar v_1 + \varepsilon^2\bar v_2\big) - \mathcal{N}_4(\varepsilon\bar v_1)\\ &\quad + \Pi_W\mathcal{N}_4(\varepsilon\bar v_1) + \varepsilon^5\partial_x\big(3\bar v_1\bar v_2^2\big) + \varepsilon^6\partial_x\big(\bar v_2^3\big).\end{aligned} \tag{4.3}$$

With these power of ε, the sufficient accuracy is achieved to start the quadratic Nash–Moser scheme (see Section 9). Hence, for $\varepsilon > 0$, let

$$\begin{aligned}F(u,\varepsilon) &:= \big(\varepsilon^{-4}\Pi_V + \varepsilon^{-2}\Pi_W\big)\mathcal{F}\big(\varepsilon\bar v_1 + \varepsilon^2 u, \omega\big)\\ &= \varepsilon^{-2}P_\varepsilon^{-1}\mathcal{F}\big(\varepsilon\bar v_1 + \varepsilon^2 u, 1+3\varepsilon^2\big) \qquad (4.4)\\ &= \Pi_V\big\{3\partial_t u + \partial_x\big(3\bar v_1^2 u + \varepsilon 3\bar v_1 u^2 + \varepsilon^2 u^3\big) + \varepsilon^{-4}\mathcal{N}_4\big(\varepsilon\bar v_1 + \varepsilon^2 u\big)\big\}\\ &\quad + \Pi_W\big\{Lu + \varepsilon^2 3\partial_t u + \varepsilon\partial_x\big[(v_1+\varepsilon u)^3\big] + \varepsilon^{-2}\mathcal{N}_4\big(\varepsilon\bar v_1 + \varepsilon^2 u\big)\big\}, \qquad (4.5)\end{aligned}$$

$$\omega := 1 + 3\varepsilon^2, \qquad P_\varepsilon := \varepsilon^2\Pi_V + \Pi_W, \qquad P_\varepsilon^{-1} = \varepsilon^{-2}\Pi_V + \Pi_W.$$

By (4.3), $F(\bar v_2, \varepsilon) = O(\varepsilon)$ (see Lemma 8.5 for precise estimates). For $\varepsilon > 0$, problem (3.2) becomes

$$F(u,\varepsilon) = 0, \quad u \in X_0. \tag{4.6}$$

Like $\mathcal{F}$ does, F also maps X_0 into Y.

5. Bifurcation

In this section we construct a solution $v \in V_0$ of (4.1) and prove its non-degeneracy. Recall that in $\mathcal{V}$ it is $l + j|j| = 0$. Let

$$q_j(t,x) := e^{i(-j|j|t + jx)}, \quad j \in \mathbb{Z}. \tag{5.1}$$

Note that $q_{j_1}q_{j_2} = 1 = q_0$ if $j_1 + j_2 = 0$.

Lemma 5.1. 1) (Product of two terms). *Let $j_1, j_2 \in \mathbb{Z}$ be both nonzero integers. Then $\Pi_V(q_{j_1} q_{j_2}) = 0$ except the case when $j_1 + j_2 = 0$.*

2) (Product of three terms). *Let $j_1, j_2, j_3 \in \mathbb{Z}$ be all nonzero integers. Then $\Pi_V(q_{j_1} q_{j_2} q_{j_3}) = 0$ except the case when $j_1 + j_2 = 0$ or $j_1 + j_3 = 0$ or $j_2 + j_3 = 0$.*

Proof. See Appendix A. □

Consider m positive distinct integers $0 < k_1 < k_2 < \cdots < k_m$, $m \geqslant 1$, and let

$$\mathcal{K} := \{k_1, k_2, \dots, k_m, -k_1, -k_2, \dots, -k_m\}.$$

Consider three elements $v, v', v'' \in V_0 \cap X$ with only Fourier modes in $\mathcal{K}$,

$$v = \sum_{j \in \mathcal{K}} a_j q_j, \qquad v' = \sum_{j \in \mathcal{K}} b_j q_j, \qquad v'' = \sum_{j \in \mathcal{K}} c_j q_j,$$

with $a_{-j} = a_j \in \mathbb{R}$, and similar for b_j, c_j. Then

$$vv'v'' = \sum_{j_1, j_2, j_3 \in \mathcal{K}} a_{j_1} b_{j_2} c_{j_3} q_{j_1} q_{j_2} q_{j_3}, \qquad \Pi_V\big(vv'v''\big) = \sum_{j_1, j_2, j_3 \in \mathcal{K}} a_{j_1} b_{j_2} c_{j_3} \Pi_V(q_{j_1} q_{j_2} q_{j_3}).$$

Develop the sum with respect to j_1. Let $k \in \mathcal{K}$. For $j_1 = k$, $\Pi_V(q_{j_1} q_{j_2} q_{j_3})$ is nonzero only if:

$$\begin{pmatrix} j_1 = k \\ j_2 = k \\ j_3 = -k \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} j_1 = k \\ j_2 = -k \\ j_3 \in \mathcal{K} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} j_1 = k \\ j_2 \neq \pm k \\ j_3 = -k \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} j_1 = k \\ j_2 \neq \pm k \\ j_3 = -j_2 \end{pmatrix}. \tag{5.2}$$

Hence in the sum only these four cases give a nonzero contribution:

$$\Pi_V\big(vv'v''\big) = \sum_{k \in \mathcal{K}} a_k b_k c_k q_k + \sum_{k, j \in \mathcal{K}} a_k b_k c_j q_j + \sum_{k \in \mathcal{K}, j \neq \pm k} a_k b_j c_k q_j + \sum_{k \in \mathcal{K}, j \neq \pm k} a_k b_j c_j q_k. \tag{5.3}$$

Since $\sum_{k \in \mathcal{K}, j \neq \pm k} = \sum_{k, j \in \mathcal{K}} - \sum_{k \in \mathcal{K}, j = k} - \sum_{k \in \mathcal{K}, j = -k}$, the third sum in (5.3) is

$$\begin{aligned} \sum_{k \in \mathcal{K}, j \neq \pm k} a_k b_j c_k q_j &= \sum_{k, j \in \mathcal{K}} a_k b_j c_k q_j - \sum_{k \in \mathcal{K}} a_k b_k c_k q_k - \sum_{k \in \mathcal{K}} a_k b_k c_k q_{-k} \\ &= \sum_{k, j \in \mathcal{K}} a_k b_j c_k q_j - \sum_{k \in \mathcal{K}} a_k b_k c_k q_k - \sum_{k \in \mathcal{K}} a_k b_k c_k q_k \end{aligned}$$

(in the last equality we have made the change of summation variable $k = -k'$). Analogously, the fourth sum in (5.3) is

$$\sum_{k \in \mathcal{K}, j \neq \pm k} a_k b_j c_j q_k = \sum_{k, j \in \mathcal{K}} a_k b_j c_j q_k - \sum_{k \in \mathcal{K}} a_k b_k c_k q_k - \sum_{k \in \mathcal{K}} a_k b_k c_k q_k.$$

Thus

$$\Pi_V\big(vv'v''\big) = \sum_{k \in \mathcal{K}} \left\{ -3 a_k b_k c_k + a_k \left(\sum_{j \in \mathcal{K}} b_j c_j \right) + b_k \left(\sum_{j \in \mathcal{K}} a_j c_j \right) + c_k \left(\sum_{j \in \mathcal{K}} a_j b_j \right) \right\} q_k. \tag{5.4}$$

The formula for $\Pi_V[\partial_x(vv'v'')] = \partial_x \Pi_V(vv'v'')$ simply has ikq_k instead of q_k. For $v = v' = v''$, (5.4) gives

$$\Pi_V\big(v^3\big) = 3 \sum_{k \in \mathcal{K}} \left(-a_k^2 + \sum_{j \in \mathcal{K}} a_j^2 \right) a_k q_k.$$

Then

$$3\partial_t v + \Pi_V\big[\partial_x\big(v^3\big)\big] = 3 \sum_{k \in \mathcal{K}} \left(-|k| - a_k^2 + \sum_{j \in \mathcal{K}} a_j^2 \right) a_k i k q_k.$$

This is zero if

$$\left(\sum_{j\in\mathcal{K}} a_j^2\right) - a_k^2 = |k| \quad \forall k \in \mathcal{K}. \tag{5.5}$$

Since $\sum_{j\in\mathcal{K}} a_j^2 = 2(a_{k_1}^2 + \cdots + a_{k_m}^2)$, (5.5) is equivalent to

$$\begin{cases} a_{k_1}^2 + 2a_{k_2}^2 + 2a_{k_3}^2 + \cdots + 2a_{k_m}^2 = k_1, \\ 2a_{k_1}^2 + a_{k_2}^2 + 2a_{k_3}^2 + \cdots + 2a_{k_m}^2 = k_2, \\ \cdots \\ 2a_{k_1}^2 + 2a_{k_2}^2 + 2a_{k_3}^2 + \cdots + a_{k_m}^2 = k_m, \end{cases} \tag{5.6}$$

which is a system of m equations in the m unknowns $a_{k_1}^2, \ldots, a_{k_m}^2$. Let M the $m \times m$ matrix that has 1 on the principal diagonal and 2 everywhere else. M is invertible, and its inverse M^{-1} is the $m \times m$ matrix that has α on the principal diagonal and β everywhere else, with

$$\alpha = -\frac{m-3/2}{m-1/2}, \qquad \beta = \frac{1}{m-1/2}.$$

Hence (5.6) is equivalent to

$$a_{k_1}^2 = \rho_1, \qquad a_{k_2}^2 = \rho_2, \quad \ldots, \quad a_{k_m}^2 = \rho_m, \tag{5.7}$$

where $(\rho_1, \ldots, \rho_m) := M^{-1}(k_1, \ldots, k_m)$, namely

$$\rho_i := \alpha k_i + \beta \sum_{j\neq i} k_j = \frac{1}{m-1/2}\left(\sum_{j=1}^m k_j\right) - k_i, \quad i = 1, \ldots, m. \tag{5.8}$$

(5.7) has solutions with all $a_j \neq 0$ if all ρ_j are positive. Note that $\rho_j > \rho_{j+1}$, because $\beta - \alpha = 1$ and

$$\rho_j - \rho_{j+1} = \alpha k_j + \beta k_{j+1} - \beta k_j - \alpha k_{j+1} = k_{j+1} - k_j > 0.$$

Hence all $\rho_j > 0$ if $\rho_m > 0$, namely if

$$k_1 + \cdots + k_{m-1} > k_m(m - 3/2). \tag{5.9}$$

When a_j satisfy (5.7),

$$\sum_{j\in\mathcal{K}} a_j^2 = 2\left(a_{k_1}^2 + \cdots + a_{k_m}^2\right) = \frac{1}{m-1/2}\sum_{i=1}^m k_i. \tag{5.10}$$

Remark 5.2. $k_1, \ldots, k_m$ satisfy (5.9) if they are sufficiently close, as if they form a "packet" of integers. Note also that if the smallest and the biggest integers satisfy the stronger condition

$$\frac{k_m}{k_1} < \frac{m-1}{m-3/2}, \tag{5.11}$$

then $k_1, k_2, \ldots, k_m$ satisfy (5.9) for every choice of the intermediate integers $k_2, \ldots, k_{m-1}$, because

$$k_1 + k_2 + \cdots + k_{m-1} > (m-1)k_1 > (m-3/2)k_m.$$

(5.11) is meaningful because $(m-1)/(m-3/2) > 1$.

Now we prove that for every $f \in V_0 \cap Y$ there is a unique $h \in V_0 \cap X$ such that

$$3\partial_t h + \Pi_V \partial_x\left(3v^2 h\right) = f. \tag{5.12}$$

Let $f \in V \cap Y$ and $h \in V \cap X$,

$$f = \sum_{j\neq 0} i y_j q_j \in V \cap Y, \quad y_{-j} = -y_j \in \mathbb{R}, \qquad h = \sum_{j\neq 0} h_j q_j \in V \cap X, \quad h_{-j} = h_j \in \mathbb{R}.$$

Split

$$f = \Pi_{\mathcal{K}} f + \Pi_{\mathcal{K}}^{\perp} f, \quad \Pi_{\mathcal{K}} f := \sum_{j \in \mathcal{K}} i y_j q_j, \ \Pi_{\mathcal{K}}^{\perp} f := \sum_{j \notin \mathcal{K}} i y_j q_j,$$

and similarly $h = \Pi_{\mathcal{K}} h + \Pi_{\mathcal{K}}^{\perp} h$. The formula for $\Pi_V \partial_x (v^2 \Pi_{\mathcal{K}} h)$ is obtained from (5.4) with $b_j = a_j$ and $c_j = h_j$, namely

$$\Pi_V \partial_x \big(v^2 (\Pi_{\mathcal{K}} h)\big) = \sum_{k \in \mathcal{K}} \left\{ -3a_k^2 h_k + 2a_k \left(\sum_{j \in \mathcal{K}} a_j h_j \right) + h_k \left(\sum_{j \in \mathcal{K}} a_j^2 \right) \right\} i k q_k.$$

Hence

$$3\partial_t (\Pi_{\mathcal{K}} h) + \Pi_V \partial_x \big(3v^2 \Pi_{\mathcal{K}} h\big) = 3 \sum_{k \in \mathcal{K}} \left\{ -|k| h_k - 3a_k^2 h_k + 2a_k \left(\sum_{j \in \mathcal{K}} a_j h_j \right) + h_k \left(\sum_{j \in \mathcal{K}} a_j^2 \right) \right\} i k q_k$$

which is, replacing $|k|$ by (5.5),

$$= 3 \sum_{k \in \mathcal{K}} \left\{ -2a_k^2 h_k + 2a_k \left(\sum_{j \in \mathcal{K}} a_j h_j \right) \right\} i k q_k = 6 \sum_{k \in \mathcal{K}} \left\{ -a_k h_k + \sum_{j \in \mathcal{K}} a_j h_j \right\} a_k i k q_k.$$

Note that this sum has only Fourier modes in $\mathcal{K}$; in other words, the space of functions in V that are Fourier-supported on $\mathcal{K}$ is an invariant subspace for the operator $3\partial_t + \Pi_V \partial_x (3v^2 \cdot)$ (with, of course, the change of parity $X \to Y$).

Thus, the equation $3\partial_t (\Pi_{\mathcal{K}} h) + \Pi_V \partial_x (3v^2 (\Pi_{\mathcal{K}} h)) = \Pi_{\mathcal{K}} f$ is equivalent to

$$-a_k h_k + \sum_{j \in \mathcal{K}} a_j h_j = \frac{y_k}{6 k a_k} =: y_k' \quad \forall k \in \mathcal{K},$$

namely to the system

$$M \begin{pmatrix} a_{k_1} h_{k_1} \\ \vdots \\ a_{k_m} h_{k_m} \end{pmatrix} = \begin{pmatrix} y_{k_1}' \\ \vdots \\ y_{k_m}' \end{pmatrix} \tag{5.13}$$

because $y_{-k}' = y_k'$ for all $k \in \mathcal{K}$, where M is the $m \times m$ matrix defined above (1 on the principal diagonal and 2 everywhere else). Therefore there exists a unique solution of (5.13),

$$h_{k_i} = \frac{1}{a_{k_i}} \left(\alpha y_{k_i}' + \beta \sum_{j \neq i} y_{k_j}' \right).$$

Since a_j solve (5.7),

$$\sum_{j \in \mathcal{K}} h_j^2 \leqslant C \sum_{j \in \mathcal{K}} y_j^2,$$

where $C > 0$ depends only on $k_1, \dots, k_m$ and m.

Now consider $\Pi_{\mathcal{K}}^{\perp} h$, $\Pi_{\mathcal{K}}^{\perp} f$. In the product

$$v^2 \big(\Pi_{\mathcal{K}}^{\perp} h\big) = \sum_{j_1, j_2 \in \mathcal{K}, j_3 \notin \mathcal{K}} a_{j_1} a_{j_2} h_{j_3} q_{j_1} q_{j_2} q_{j_3}$$

only the second case of (5.2) occurs, namely $j_1 = k = -j_2 \in \mathcal{K}$, $j_3 \notin \mathcal{K}$. Hence

$$\Pi_V \partial_x \big(v^2 \big(\Pi_{\mathcal{K}}^{\perp} h\big)\big) = \sum_{k \in \mathcal{K}, j \notin \mathcal{K}} a_k^2 h_j i j q_j = \left(\sum_{k \in \mathcal{K}} a_k^2 \right) \sum_{j \notin \mathcal{K}} i j h_j q_j = \frac{k_1 + \cdots + k_m}{m - 1/2} \partial_x \big(\Pi_{\mathcal{K}}^{\perp} h\big)$$

by (5.10). Therefore

$$3\partial_t \big(\Pi_{\mathcal{K}}^{\perp} h\big) + \Pi_V \partial_x \big(3v^2 \big(\Pi_{\mathcal{K}}^{\perp} h\big)\big) = 3 \sum_{j \notin \mathcal{K}} \left(-|j| + \frac{k_1 + \cdots + k_m}{m - 1/2} \right) i j h_j q_j.$$

Analogously as above, note that this sum has only Fourier modes out of $\mathcal{K}$; in other words, the space of functions in V that are Fourier-supported on the complementary of $\mathcal{K}$ is invariant for the operator $3\partial_t + \Pi_V \partial_x(3v^2\cdot)$ (with the change of parity $X \to Y$). The condition for the invertibility is

$$\frac{k_1 + \cdots + k_m}{m - 1/2} \neq |j| \quad \forall j \notin \mathcal{K}. \tag{5.14}$$

When (5.9) holds, $k_1 + \cdots + k_m > k_m(m-1/2)$, therefore $(k_1 + \cdots + k_m)/(m-1/2)$ is automatically out of $\mathcal{K}$. Hence (5.14) can be more easily written in this equivalent form:

$$\frac{k_1 + \cdots + k_m}{m - 1/2} \notin \mathbb{N}. \tag{5.15}$$

(5.15) implies that

$$\left| -|j| + \frac{k_1 + \cdots + k_m}{m - 1/2} \right| \geqslant \delta |j| \quad \forall j \neq 0, \tag{5.16}$$

where $\delta > 0$ depends only on $k_1, \ldots, k_m$ and m. Therefore the equation $3\partial_t(\Pi_{\mathcal{K}}^{\perp} h) + \Pi_V \partial_x(3v^2(\Pi_{\mathcal{K}}^{\perp} h)) = \Pi_{\mathcal{K}}^{\perp} g$ has a unique solution $\Pi_{\mathcal{K}}^{\perp} h$, with

$$|h_j| \leqslant \frac{C}{|j|^2} |y_j| \quad \forall j \neq 0, \ j \notin \mathcal{K}.$$

Also, by (5.10) and Lemma 5.1, $(k_1 + \cdots + k_m)/(m - 1/2) = \Pi_C(v^2)$, therefore (5.16) can be written as $|\Pi_C(v^2) - |j|| \geqslant \delta |j|$ for all $j \neq 0$.

We have proved the following result:

Proposition 5.3 *(Bifurcation for cubic nonlinearities). Let $m \geqslant 2$. Let $0 < k_1 < k_2 < \cdots < k_m$ be m positive integers that satisfy (5.9) and (5.15). Then there exist m positive numbers $\rho_1, \ldots, \rho_m > 0$, given by (5.8), and constants $C, \delta > 0$ that depend only on $k_1, \ldots, k_m$ and have the following property.*

Let $\mathcal{K} := \{k_1, \ldots, k_m, -k_1, \ldots, -k_m\}$. Every function $v = \sum_{j \in \mathcal{K}} a_j q_j \in V_0 \cap X$ which is Fourier-supported on $\mathcal{K}$ with

$$a_{k_1}^2 = \rho_1, \quad \ldots, \quad a_{k_m}^2 = \rho_m$$

is a solution of the unperturbed bifurcation equation $3\partial_t v + \Pi_V \partial_x(v^3) = 0$.

For every $f \in V_0 \cap Y$ there exists a unique $h \in V_0 \cap X$ such that $3\partial_t h + \Pi_V \partial_x(3v^2 h) = f$.

If $f \in H^s$, $s \geqslant 0$, then $h \in H^{s+1}$, with $\|h\|_{s+1} \leqslant C\|f\|_s$. Moreover

$$\left| \Pi_C(v^2) - |j| \right| \geqslant \delta |j| \quad \forall j \in \mathbb{Z}, \ j \neq 0.$$

6. The linearized equation

Remember that

$$F(u, \varepsilon) = \varepsilon^{-2} P_\varepsilon^{-1} \mathcal{F}(\varepsilon \bar{v} + \varepsilon^2 u, \omega), \quad \omega = 1 + 3\varepsilon^2, \ P_\varepsilon^{-1} = \varepsilon^{-2} \Pi_V + \Pi_W,$$

where $\bar{v} := \bar{v}_1$ is a solution of the unperturbed bifurcation equation (4.1) as in Proposition 5.3. The linearized operator $F'(u, \varepsilon)$ applied to h, namely the Fréchet derivative $\partial_u F(u, \varepsilon)[h]$ of F with respect to u in the direction h, is then

$$F'(u, \varepsilon)h = \varepsilon^{-2} P_\varepsilon^{-1} \mathcal{F}'(\varepsilon \bar{v} + \varepsilon^2 u, \omega)\left[\varepsilon^2 h\right] = P_\varepsilon^{-1} \mathcal{L}(u, \varepsilon)h,$$

$$\mathcal{L}(u, \varepsilon)h := \mathcal{F}'(\varepsilon \bar{v} + \varepsilon^2 u, \omega)[h] = \omega \partial_t h + (1 + a_1)\mathcal{H}\partial_{xx} h + a_2 \mathcal{H}\partial_x h + a_3 \partial_x h + a_4 \mathcal{H} h + a_5 h$$

where the coefficients $a_i = a_i(t, x) = a_i(u, \varepsilon)(t, x)$ are periodic functions of (t, x), depending on u, ε, and are obtained from $\partial_x(U^3)$ and the partial derivatives of g_1, g_2 or g_0 evaluated at $(x, U(t, x), \mathcal{H}U(t, x), \ldots)$, $U := \varepsilon \bar{v} + \varepsilon^2 u$. For example, in case (I)

$$a_1(t, x) = (\partial_{y_2} g_2)\big(x, U(t, x), \mathcal{H}U_x(t, x)\big), \qquad a_2(t, x) = \partial_x a_1(t, x), \tag{6.1}$$

and in case (II)

$$a_1(t,x) = (\partial_{y_4} g_0)\big(x, U(t,x), \mathcal{H}U(t,x), U_x(t,x), \mathcal{H}U_{xx}(t,x)\big), \qquad a_2(t,x) = 0. \tag{6.2}$$

$\mathcal{N}(U) = \partial_x(U^3) + O(U^4)$, and $U = \varepsilon\bar{v} + \varepsilon^2 u = O(\varepsilon)$, therefore $a_1, a_2, a_4 = O(\varepsilon^3)$, $a_3, a_5 = O(\varepsilon^2)$. More precisely: let $\delta_0 \in (0,1)$ be a universal constant such that

$$\big\|(U, \mathcal{H}U, U_x, \mathcal{H}U_x, \mathcal{H}U_{xx})\big\|_{L^\infty} < 1 \quad \forall U \in H^4(\mathbb{T}^2),\ \|U\|_4 < \delta_0. \tag{6.3}$$

Proposition 6.1. *Let $K > 0$. There exists $\varepsilon_0 \in (0,1)$, depending on K, with the following property: if $\varepsilon \in (0, \varepsilon_0)$, $\|u\|_4 \leqslant K$, and*

$$\big\|\varepsilon\bar{v}_1 + \varepsilon^2 u\big\|_4 \leqslant \varepsilon_0\|\bar{v}_1\|_4 + \varepsilon_0^2\|u\|_4 < \delta_0, \tag{6.4}$$

then the coefficients $a_i(u,\varepsilon)(t,x)$, $i = 1,\dots,5$ satisfy

$$|a_1|_s + |a_2|_s + \big|a_3 - \varepsilon^2 3\bar{v}^2\big|_s + |a_4|_s + \big|a_5 - \varepsilon^2\big(3\bar{v}^2\big)_x\big|_s \leqslant \varepsilon^3 C(s,K)\big(1 + \|u\|_{s+4}\big), \quad 0 \leqslant s \leqslant r. \tag{6.5}$$

a_i is of class C^1 as a function of (u,ε), with

$$\sum_{i=1,2,4} \big|\partial_u a_i(u,\varepsilon)[h]\big|_s + \big|\partial_u a_3(u,\varepsilon)[h] - \varepsilon^3 6\bar{v}h\big|_s + \big|\partial_u a_5(u,\varepsilon)[h] - \varepsilon^3(6\bar{v}h)_x\big|_s$$
$$\leqslant \varepsilon^4 C(s,K)\big(\|h\|_{s+4} + \|u\|_{s+4}\|h\|_4\big), \tag{6.6}$$
$$\sum_{i=1,2,4} \big|\partial_\varepsilon a_i(u,\varepsilon)\big|_s + \big|\partial_\varepsilon a_3(u,\varepsilon) - \varepsilon 6\bar{v}^2\big|_s + \big|\partial_\varepsilon a_5(u,\varepsilon) - \varepsilon\big(6\bar{v}^2\big)_x\big|_s \leqslant \varepsilon^2 C(s,K)\big(1 + \|u\|_{s+4}\big), \tag{6.7}$$

for $0 \leqslant s \leqslant r$. The constant $C(s,K) > 0$ depend on s, K, and $K_{g,r}$ of (1.4). In these estimates the norm $\|\bar{v}_1\|_{s+4}$ appears like a constant $C(s)$ depending on s.

Proof. In Appendix C. □

Remark 6.2. In general, the inequality $\|\mathcal{H}u\|_{L^\infty} \leqslant C\|u\|_{L^\infty}$ is false (see, for example, [24]), while it is trivially true that $\|\mathcal{H}u\|_s \leqslant \|u\|_s$ for all s. Therefore to obtain the estimate $\|\mathcal{H}u_{xx}\|_{L^\infty} \leqslant C\|u\|_4$ (which is used to prove (6.3)) the right chain of inequalities is $\|\mathcal{H}u_{xx}\|_{L^\infty} \leqslant C\|\mathcal{H}u_{xx}\|_2 \leqslant C\|u_{xx}\|_2 \leqslant C\|u\|_4$.

Since $\bar{v}, u \in X$,

$$a_1, a_3, a_4 \in X, \qquad a_2, a_5 \in Y,$$

and $\mathcal{L}(u,\varepsilon)$ maps $X \cap H^2 \to Y$.

As a pseudo-differential operator, we write

$$\mathcal{L} := \mathcal{L}(u,\varepsilon) = \omega\partial_t + \big(1 + a_1(t,x)\big)\mathcal{H}\partial_{xx} + a_2(t,x)\mathcal{H}\partial_x + a_3(t,x)\partial_x + a_4(t,x)\mathcal{H} + a_5(t,x).$$

In this operator notation a function $p(t,x)$ is identified with the multiplication operator $h \mapsto p(t,x)h$, and the composition is understood: for example, $\partial_x p$ is the operator $p\partial_x + p_x$, because $\partial_x(ph) = p\partial_x h + p_x h$.

To emphasize that we are in the space of zero mean functions, write

$$\tilde{\mathcal{L}} := \mathbb{P}\mathcal{L}\mathbb{P},$$

where $\mathbb{P} = I - \Pi_C$ is defined in (3.1). Since F maps $X_0 \to Y$, also $F'(u,\varepsilon)$ maps $X_0 \to Y$, therefore

$$\tilde{\mathcal{L}}h = \mathcal{L}h \quad \forall h \in X_0$$

because $\mathbb{P}h = h$ and $\mathbb{P}f = f$ for all $h \in X_0$, $f \in Y$.

7. Reduction to constant coefficients

In this section the linearized operator is conjugated to a linear operator with constant coefficients plus a regularizing rest. The transformation is performed in several steps.

7.1. Change of variables

As a first step in the reduction proof, we construct a change of variables that transforms $\mathcal{L}$ into a new operator with constant coefficients in the highest order derivatives ∂_t and $\mathcal{H}\partial_{xx}$. Since $\mathcal{L}$ maps X_0 into Y, we want that our transformation maps $X_0 \to X_0$ and $Y \to Y$.

We consider diffeomorphisms of the torus $(t,x) \in \mathbb{T}^2$ which are the composition of (i) a time-dependent change of the space variable $x \to x + \beta(t,x)$, and (ii) a change of the time variable $t \to t + \alpha(t)$ that does not depend on space. Diffeomorphisms of this type preserve the special role of the time variable as "a parameter" with respect to pseudo-differential operators of the space variable like $\mathcal{H}$.

Let

$$\psi : \mathbb{T}^2 \to \mathbb{T}^2, \qquad \psi(t,x) := \big(t + \alpha(t), x + \beta(t,x)\big) = (\tau, y)$$

and let Ψ be the transformation $\Psi : u \mapsto \Psi u$,

$$(\Psi u)(t,x) := u\big(\psi(t,x)\big) = u\big(t+\alpha(t), x + \beta(t,x)\big) = u(\tau, y).$$

$\alpha(t)$ and $\beta(t,x)$ are periodic functions in Y to be determined.

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

$$\Psi^{-1}\partial_t\Psi = \big[1 + \big(\Psi^{-1}\alpha'\big)(\tau)\big]\partial_\tau + \big(\Psi^{-1}\beta_t\big)(\tau,y)\partial_y, \qquad \Psi^{-1}\partial_x\Psi = \big[1 + \big(\Psi^{-1}\beta_x\big)(\tau,y)\big]\partial_y,$$
$$\Psi^{-1}\partial_{xx}\Psi = \big[1 + \big(\Psi^{-1}\beta_x\big)(\tau,y)\big]^2\partial_{yy} + \big(\Psi^{-1}\beta_{xx}\big)(\tau,y)\partial_y, \qquad \Psi^{-1}\mathcal{H}\Psi = \mathcal{H} + \mathcal{R}_{\mathcal{H}},$$

where $\mathcal{R}_{\mathcal{H}}$ is defined by the last equality, and it is regularizing in space, bounded in time, see Lemma B.5(iii).

Since $\alpha, \beta \in Y$, Ψ maps $X \to X$ and $Y \to Y$. However, in general, Ψ does not map X_0 into X_0.[1] To obtain a transformation of X_0 onto itself, consider the projection onto Z_0,

$$\tilde{\Psi} := \mathbb{P}\Psi\mathbb{P}.$$

Since $\Psi^{-1}\Pi_C = \Pi_C$, one has $\mathbb{P}\Psi^{-1}\Pi_C = \mathbb{P}\Pi_C = 0$, and

$$\mathbb{P}\Psi^{-1}\mathbb{P} = \mathbb{P}\Psi^{-1}(I - \Pi_C) = \mathbb{P}\Psi^{-1}. \tag{7.1}$$

As a consequence,

$$\big(\mathbb{P}\Psi^{-1}\mathbb{P}\big)(\mathbb{P}\Psi\mathbb{P}) = \mathbb{P}\Psi^{-1}\mathbb{P}\Psi\mathbb{P} = \mathbb{P}\Psi^{-1}\Psi\mathbb{P} = \mathbb{P},$$

therefore $\tilde{\Psi} : Z_0 \to Z_0$ is invertible, with inverse

$$(\tilde{\Psi})^{-1} = (\mathbb{P}\Psi\mathbb{P})^{-1} = \mathbb{P}\Psi^{-1}\mathbb{P}.$$

Thus $\tilde{\Psi}$ is a linear bijective operator of $X_0 \to X_0$ and $Y \to Y$. Also,

$$[\Psi, \mathbb{P}]h = [\Pi_C, \Psi]h = \Pi_C\big(\tilde{\alpha}' + \tilde{\beta}_y + \tilde{\alpha}'\tilde{\beta}_y\big)h = \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2} h\big(\tilde{\alpha}' + \tilde{\beta}_y + \tilde{\alpha}'\tilde{\beta}_y\big)\,d\tau\,dy, \tag{7.2}$$

where $(\tau,y) \mapsto (\tau + \tilde{\alpha}(\tau), y + \tilde{\beta}(\tau,y)) = \psi^{-1}(\tau,y)$ is the inverse of ψ, and similarly

$$\big[\Psi^{-1}, \mathbb{P}\big] = \big[\Pi_C, \Psi^{-1}\big] = \Pi_C\big(\alpha' + \beta_x + \alpha'\beta_x\big).$$

These commutators are regularizing operators, both in space and time (by integrations by parts, any derivative applied to the argument h moves to α, β or $\tilde{\alpha}, \tilde{\beta}$).

By (7.1),

$$\tilde{\mathcal{L}}_1 := \tilde{\Psi}^{-1}\tilde{\mathcal{L}}\tilde{\Psi} = \mathbb{P}\Psi^{-1}\mathbb{P}\mathcal{L}\mathbb{P}\Psi\mathbb{P} = \mathbb{P}\Psi^{-1}\mathcal{L}\mathbb{P}\Psi\mathbb{P} = \mathbb{P}\mathcal{L}_1\mathbb{P},$$

[1] For example: let $u(t,x) = \cos t \in X_0$, $\beta = 0$ and α such that the inverse of $t \mapsto t + \alpha(t)$ is $\tau \mapsto \tau + (1/2)\sin\tau$. Changing variable in the integral, $\int_{\mathbb{T}^2}(\Psi u)\,dt\,dx = (1/2)\int_{\mathbb{T}^2}\cos^2\tau\,d\tau\,dy > 0$, therefore $\Psi u \notin X_0$.

where

$$
\begin{aligned}
\mathcal{L}_1 &= \omega\big[1+\big(\Psi^{-1}\alpha'\big)(\tau)\big]\partial_\tau + \big[1+\big(\Psi^{-1}a_1\big)(\tau,y)\big]\big[1+\big(\Psi^{-1}\beta_x\big)(\tau,y)\big]^2\partial_{yy}\mathcal{H} \\
&\quad + \big\{\big[1+\big(\Psi^{-1}a_1\big)(\tau,y)\big]\big(\Psi^{-1}\beta_{xx}\big)(\tau,y) + \big(\Psi^{-1}a_2\big)(\tau,y)\big[1+\big(\Psi^{-1}\beta_x\big)(\tau,y)\big]\big\}\partial_y\mathcal{H} \\
&\quad + \big\{\omega\big(\Psi^{-1}\beta_t\big)(\tau,y) + \big(\Psi^{-1}a_3\big)(\tau,y)\big[1+\big(\Psi^{-1}\beta_x\big)(\tau,y)\big]\big\}\partial_y \\
&\quad + \big(\Psi^{-1}a_4\big)(\tau,y)\mathcal{H} + \big(\Psi^{-1}a_5\big)(\tau,y) + \mathcal{R}_1, \\
\mathcal{R}_1 &= \big[1+\big(\Psi^{-1}a_1\big)(\tau,y)\big]\big[1+\big(\Psi^{-1}\beta_x\big)(\tau,y)\big]^2\partial_{yy}\mathcal{R}_{\mathcal{H}} \\
&\quad + \big\{\big[1+\big(\Psi^{-1}a_1\big)(\tau,y)\big]\big(\Psi^{-1}\beta_{xx}\big)(\tau,y) + \big(\Psi^{-1}a_2\big)(\tau,y)\big[1+\big(\Psi^{-1}\beta_x\big)(\tau,y)\big]\big\}\partial_y\mathcal{R}_{\mathcal{H}} \\
&\quad + \big(\Psi^{-1}a_4\big)(\tau,y)\mathcal{R}_{\mathcal{H}} - \mathbb{P}\big(\Psi^{-1}a_5\big)(\tau,y)[\Pi_C,\Psi]
\end{aligned}
\tag{7.3}
$$

because $\mathcal{L}\Pi_C = a_5\Pi_C$. We look for α,β such that the coefficients of ∂_τ and $\partial_{yy}\mathcal{H}$ are proportional, namely

$$\big[1+\big(\Psi^{-1}a_1\big)(\tau,y)\big]\big[1+\big(\Psi^{-1}\beta_x\big)(\tau,y)\big]^2 = \mu_2\big[1+\big(\Psi^{-1}\alpha'\big)(\tau)\big] \tag{7.4}$$

for some $\mu_2 \in \mathbb{R}$. (7.4) is equivalent to

$$\big(1+a_1(t,x)\big)\big(1+\beta_x(t,x)\big)^2 = \mu_2\big(1+\alpha'(t)\big). \tag{7.5}$$

Take the square root of (7.5),

$$1+\beta_x(t,x) = \mu_2^{1/2}\big(1+\alpha'(t)\big)^{1/2}\big(1+a_1(t,x)\big)^{-1/2}, \tag{7.6}$$

and integrate in dx,

$$1 = \mu_2^{1/2}\big(1+\alpha'(t)\big)^{1/2}\frac{1}{2\pi}\int_0^{2\pi}(1+a_1)^{-1/2}\,dx.$$

Take the square,

$$\mu_2\big(1+\alpha'(t)\big) = \left(\frac{1}{2\pi}\int_0^{2\pi}(1+a_1)^{-1/2}\,dx\right)^{-2} =: \rho(t). \tag{7.7}$$

Integrating in dt determines $\mu_2 \in \mathbb{R}$,

$$\mu_2 = \Pi_C(\rho) = \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi}(1+a_1)^{-1/2}\,dx\right)^{-2}dt,$$

then $\alpha(t) \in Y$ is also determined,

$$\alpha(t) = \frac{1}{\mu_2}\partial_t^{-1}(\Pi_T\rho)(t).$$

Since $a_1 \in X$, also $\rho \in X$, therefore $\alpha \in Y$, as it was required. (7.6) gives

$$\beta_x = \rho^{1/2}(1+a_1)^{-1/2} - 1 = \frac{p}{\Pi_{T+C}(p)} - 1 = \frac{\Pi_E(p)}{\Pi_{T+C}(p)}, \quad p := (1+a_1)^{-1/2}, \tag{7.8}$$

therefore the Z_E-component of β is determined,

$$(\Pi_E\beta)(t,x) = \frac{1}{(\Pi_T p)(t) + \Pi_C(p)}\big(\partial_x^{-1}\Pi_E p\big)(t,x).$$

Since $a_1 \in X$, also $p \in X$, and $\Pi_E\beta \in Y$, as it was required. The Z_T-component of β will be determined later. With this choice of α,β, (7.4) is satisfied. By (7.4),

$$\mathcal{L}_1 = \mathcal{M}\mathcal{L}_2,$$

where $\mathcal{M}$ is the multiplication operator of factor $[1+(\Psi^{-1}\alpha')(\tau)]$,

$$\mathcal{L}_2 = \omega\partial_\tau + \mu_2\partial_{yy}\mathcal{H} + a_6(\tau,y)\partial_y\mathcal{H} + a_7(\tau,y)\partial_y + a_8(\tau,y)\mathcal{H} + a_9(\tau,y) + \mathcal{R}_2,$$
$$a_6(\tau,y) := \Psi^{-1}\left(\frac{(1+a_1)\beta_{xx} + a_2(1+\beta_x)}{1+\alpha'}\right)(\tau,y), \qquad a_8(\tau,y) := \Psi^{-1}\left(\frac{a_4}{1+\alpha'}\right)(\tau,y),$$
$$a_7(\tau,y) := \Psi^{-1}\left(\frac{\omega\beta_t + a_3(1+\beta_x)}{1+\alpha'}\right)(\tau,y), \qquad a_9(\tau,y) := \Psi^{-1}\left(\frac{a_5}{1+\alpha'}\right)(\tau,y),$$
$$\mathcal{R}_2 := \frac{1}{1+(\Psi^{-1}\alpha')(\tau)}\mathcal{R}_1. \tag{7.9}$$

We show that

$$a_6(\tau,y) \in Z_E. \tag{7.10}$$

For each fixed $\tau = t+\alpha(t)$, changing variable $y = x+\beta(t,x)$, $dy = (1+\beta_x(t,x))\,dx$ in the integral,

$$\int_0^{2\pi} a_6(\tau,y)\,dy = \int_0^{2\pi} \frac{(1+a_1(t,x))\beta_{xx}(t,x) + a_2(t,x)(1+\beta_x(t,x))}{1+\alpha'(t)}\big(1+\beta_x(t,x)\big)\,dx.$$

By (7.5),

$$\frac{(1+a_1)\beta_{xx} + a_2(1+\beta_x)}{1+\alpha'}(1+\beta_x) = \mu_2\frac{(1+a_1)\beta_{xx} + a_2(1+\beta_x)}{(1+a_1)(1+\beta_x)}.$$

In case (I) $a_2 = (a_1)_x$ (see (6.1)), therefore

$$\frac{(1+a_1)\beta_{xx} + a_2(1+\beta_x)}{(1+a_1)(1+\beta_x)} = \frac{[(1+a_1)(1+\beta_x)]_x}{(1+a_1)(1+\beta_x)} = \partial_x\big\{\log\big[(1+a_1)(1+\beta_x)\big]\big\};$$

in case (II) $a_2 = 0$ (see (6.2)), therefore

$$\frac{(1+a_1)\beta_{xx} + a_2(1+\beta_x)}{(1+a_1)(1+\beta_x)} = \frac{\beta_{xx}}{1+\beta_x} = \partial_x\big\{\log(1+\beta_x)\big\}.$$

Hence in both cases (I) and (II), by periodicity, $\int_0^{2\pi} a_6\,dy = 0$, which is (7.10).

Remark 7.1. The assumptions (I), (II) on the nonlinearity $\mathcal{N}_4(u)$ have been used to prove (7.10). In more general situations, when (I), (II) are not satisfied, a term $b(\tau)\mathcal{H}\partial_y$ also appears, where $b(\tau) \in Z_T$ is the Z_T-component of the coefficient a_6 (which here is zero by (7.10)). This term can be removed by using the Fourier integral operator

$$u(\tau,y) = \sum_{j\in\mathbb{Z}} u_j(\tau)e^{ijy} \ \mapsto\ Au(\tau,y) = \sum_{j\in\mathbb{Z}} u_j(\tau)e^{ijy+|j|p(\tau)},$$

where $p(\tau) = \partial_\tau^{-1}b(\tau)$.

Now we choose the Z_T-component of β so that $\Pi_T a_7 = 0$. Denote $\gamma(t) := (\Pi_T\beta)(t)$. As above,

$$\frac{1}{2\pi}\int_{\mathbb{T}} a_7(\tau,y)\,dy = \frac{1}{2\pi}\int_0^{2\pi} \frac{\omega\beta_t(t,x) + a_3(t,x)(1+\beta_x(t,x))}{1+\alpha'(t)}\big(1+\beta_x(t,x)\big)\,dx.$$

This integral is equal to some constant $\mu_1 \in \mathbb{R}$ if and only if

$$\omega\gamma'(t) + \sigma(t) = \mu_1\big(1+\alpha'(t)\big), \quad \sigma(t) := \frac{1}{2\pi}\int_0^{2\pi}\Big(\omega\beta_t^E\big(1+\beta_x^E\big) + a_3\big(1+\beta_x^E\big)^2\Big)\,dx, \ \beta^E := \Pi_E\beta. \tag{7.11}$$

Hence an integration in dt on $\mathbb{T}$ determines $\mu_1 \in \mathbb{R}$ and $\gamma \in Z_T$,

$$\mu_1 = \Pi_C(\sigma), \qquad \gamma(t) = \frac{\mu_1\alpha(t) - (\partial_t^{-1}\Pi_T\sigma)(t)}{\omega} \in Z_T. \tag{7.12}$$

Thus

$$\Pi_C(a_7) = \mu_1, \quad a_7 - \mu_1 \in Z_E. \tag{7.13}$$

$\sigma \in X$ because $a_3 \in X$, therefore $\gamma \in Y$ as it was required. Hence $\beta = \gamma + (\Pi_E \beta) \in Y$. As a consequence,

$$a_6, a_9 \in Y, \qquad a_7, a_8 \in X. \tag{7.14}$$

Since $I = \mathbb{P} + \Pi_C$,

$$\tilde{\mathcal{L}}_1 = \mathbb{P}\mathcal{L}_1\mathbb{P} = \mathbb{P}\mathcal{M}\mathcal{L}_2\mathbb{P} = (\mathbb{P}\mathcal{M}\mathbb{P})(\mathbb{P}\mathcal{L}_2\mathbb{P}) - \mathbb{P}\mathcal{M}\Pi_C\mathcal{L}_2\mathbb{P} = \tilde{\mathcal{M}}\tilde{\mathcal{L}}_3,$$

where

$$\tilde{\mathcal{M}} := \mathbb{P}\mathcal{M}\mathbb{P}, \qquad \tilde{\mathcal{L}}_3 := \mathbb{P}\mathcal{L}_3\mathbb{P}, \quad \mathcal{L}_3 = \mathcal{L}_2 - \tilde{\mathcal{M}}^{-1}\mathcal{M}\Pi_C\mathcal{L}_2.$$

Thus

$$\mathcal{L}_3 = \omega\partial_\tau + \mu_2\partial_{yy}\mathcal{H} + a_6(\tau, y)\partial_y\mathcal{H} + a_7(\tau, y)\partial_y + a_8(\tau, y)\mathcal{H} + a_9(\tau, y) + \mathcal{R}_3,$$
$$\mathcal{R}_3 := \mathcal{R}_2 - \tilde{\mathcal{M}}^{-1}\mathcal{M}\Pi_C\mathcal{L}_2.$$

$\tilde{\mathcal{M}}$ is invertible, its inverse $\tilde{\mathcal{M}}^{-1}$ maps $X_0 \to X_0$ and $Y \to Y$, and

$$\tilde{\mathcal{M}}^{-1}h = mh - \frac{m}{\Pi_C(m)}\Pi_C(mh), \quad m(\tau) := \frac{1}{1 + (\Psi^{-1}\alpha')(\tau)}, \tag{7.15}$$

whence

$$\tilde{\mathcal{M}}^{-1}\mathcal{M}\Pi_C = -\left(\frac{(\mathbb{P}m)}{\Pi_C(m)}\right)\Pi_C.$$

Formula (7.15) can be proved by a direct calculation: $\tilde{\mathcal{M}}\tilde{\mathcal{M}}^{-1}h = \tilde{\mathcal{M}}^{-1}\tilde{\mathcal{M}}h = h$ for all $h \in Z_0$.

From Proposition 6.1 and the explicit formulae above, $\mu_2, \mu_1, \rho, \alpha, \beta, \gamma$ all depend on (u, ε) in a C^1 way, and the following estimates hold.

Proposition 7.2. *Let $K > 0$. There exists $\varepsilon_0 \in (0, 1)$, depending on K, such that, if $\varepsilon \in (0, \varepsilon_0)$, $\|u\|_8 \leqslant K$, and $\|u\|_4, \varepsilon_0$ satisfy (6.4), then all the following inequalities hold.*

$\mu_2(u, \varepsilon)$ and $\mu_1(u, \varepsilon)$ satisfy

$$|\mu_2 - 1| \leqslant \varepsilon^3 C(K), \qquad \left|\partial_u\mu_2[h]\right| \leqslant \varepsilon^4 C(K)\|h\|_4, \qquad |\partial_\varepsilon\mu_2| \leqslant \varepsilon^2 C(K), \tag{7.16}$$

$$\left|\mu_1 - \varepsilon^2\Pi_C\left(3\bar{v}^2\right)\right| \leqslant \varepsilon^3 C(K), \qquad \left|\partial_u\mu_1[h]\right| \leqslant \varepsilon^4 C(K)\|h\|_5, \qquad \left|\partial_\varepsilon\mu_1 - \varepsilon\Pi_C\left(6\bar{v}^2\right)\right| \leqslant \varepsilon^2 C(K). \tag{7.17}$$

$\psi(t, x) = (t + \alpha(t), x + \beta(t, x))$ and its inverse $\psi^{-1}(\tau, y) = (\tau + \tilde{\alpha}(\tau), y + \tilde{\beta}(\tau, y))$ are diffeomorphisms of $\mathbb{T}^2$, with

$$|\alpha|_1 + |\beta|_1 + |\tilde{\alpha}|_1 + |\tilde{\beta}|_1 < \varepsilon^3 C(K) < 1/2, \qquad |\alpha|_s + |\beta|_s + |\tilde{\alpha}|_s + |\tilde{\beta}|_s \leqslant \varepsilon^3 C(s, K)\left(1 + \|u\|_{s+4}\right), \tag{7.18}$$

for all $1 \leqslant s \leqslant r$. $\alpha, \beta, \tilde{\alpha}, \tilde{\beta}$ are C^1 functions of (u, ε). For $1 \leqslant s \leqslant r - 1$, their derivatives satisfy

$$\left|\partial_u\alpha[h]\right|_s + \left|\partial_u\beta[h]\right|_s + \left|\partial_u\tilde{\alpha}[h]\right|_s + \left|\partial_u\tilde{\beta}[h]\right|_s \leqslant \varepsilon^4 C(s, K)\left(\|h\|_{s+4} + \|u\|_{s+5}\|h\|_5\right), \tag{7.19}$$

$$|\partial_\varepsilon\alpha|_s + |\partial_\varepsilon\beta|_s + |\partial_\varepsilon\tilde{\alpha}|_s + |\partial_\varepsilon\tilde{\beta}|_s \leqslant \varepsilon^2 C(s, K)\left(1 + \|u\|_{s+5}\right). \tag{7.20}$$

The operators Ψ, Ψ^{-1} satisfy

$$\|\Psi f\|_s + \left\|\Psi^{-1} f\right\|_s \leqslant C(s, K)\left(\|f\|_s + \|u\|_{s+4}\|f\|_1\right), \qquad \|\Psi f\|_0 + \left\|\Psi^{-1} f\right\|_0 \leqslant 2\|f\|_0, \tag{7.21}$$

$$\left\|(\Psi - I) f\right\|_s + \left\|\left(\Psi^{-1} - I\right) f\right\|_s \leqslant \varepsilon^3 C(s, K)\left(\|f\|_{s+1} + \|u\|_{s+5}\|f\|_1\right), \tag{7.22}$$

for all $1 \leqslant s \leqslant r$. (7.21), (7.22) also hold for $\tilde{\Psi}, \tilde{\Psi}^{-1}$. Moreover, for $1 \leqslant s \leqslant r$,

$$|\Psi f|_s + \left|\Psi^{-1} f\right|_s \leqslant C(s, K)\left(|f|_s + \|u\|_{s+4}|f|_1\right), \qquad |\Psi f|_0 = \left|\Psi^{-1} f\right|_0 = |f|_0, \tag{7.23}$$

$$\left|(\Psi - I) f\right|_s + \left|\left(\Psi^{-1} - I\right) f\right|_s \leqslant \varepsilon^3 C(s, K)\left(|f|_{s+1} + \|u\|_{s+5}|f|_1\right). \tag{7.24}$$

The operators Ψ, Ψ^{-1} depend on (u, ε) via α, β. The derivatives of Ψf, $\Psi^{-1} f$ with respect to u in the direction h and with respect to ε satisfy

$$\left\|\partial_u(\Psi f)[h]\right\|_s + \left\|\partial_u\big(\Psi^{-1} f\big)[h]\right\|_s \leqslant \varepsilon^4 C(s, K)\big(\|f\|_{s+1}\|h\|_5 + \|f\|_1\|h\|_{s+4} + \|u\|_{s+5}\|f\|_1\|h\|_5\big), \tag{7.25}$$

$$\|\partial_\varepsilon \Psi f\|_s + \left\|\partial_\varepsilon \Psi^{-1} f\right\|_s \leqslant \varepsilon^2 C(s, K)\big(\|f\|_{s+1} + \|u\|_{s+5}\|f\|_1\big), \tag{7.26}$$

for all $1 \leqslant s \leqslant r-1$. (7.25) and (7.26) also hold with $|\,|_s$ instead of $\|\,\|_s$ on the left-hand side and on f. (7.25) and (7.26) also hold for $\tilde\Psi, \tilde\Psi^{-1}$.

For $2 \leqslant s \leqslant r$,

$$\left\|(\tilde{\mathcal M} - I) f\right\|_s + \left\|\big(\tilde{\mathcal M}^{-1} - I\big) f\right\|_s \leqslant \varepsilon^3 C(s, K)\big(\|f\|_s + \|u\|_{s+4}\|f\|_2\big). \tag{7.27}$$

The derivatives of $\tilde{\mathcal M} f$, $\tilde{\mathcal M}^{-1} f$ with respect to u in the direction h and with respect to ε satisfy

$$\left\|\partial_u(\tilde{\mathcal M} f)[h]\right\|_s + \left\|\partial_u\big(\tilde{\mathcal M}^{-1} f\big)[h]\right\|_s \leqslant \varepsilon^4 C(s, K)\big(\|f\|_s\|h\|_6 + \|f\|_2\|h\|_{s+5} + \|u\|_{s+6}\|f\|_2\|h\|_5\big), \tag{7.28}$$

$$\|\partial_\varepsilon \tilde{\mathcal M} f\|_s + \left\|\partial_\varepsilon \tilde{\mathcal M}^{-1} f\right\|_s \leqslant \varepsilon^2 C(s, K)\big(\|f\|_s + \|u\|_{s+6}\|f\|_2\big), \tag{7.29}$$

for $2 \leqslant s \leqslant r-2$.

The coefficients of $\mathcal L_3$ satisfy

$$|a_6|_s + \left|a_7 - \varepsilon^2 3\bar v^2\right|_s + |a_8|_s + \left|a_9 - \varepsilon^2\big(3\bar v^2\big)_x\right|_s \leqslant \varepsilon^3 C(s, K)\big(1 + \|u\|_{s+6}\big), \tag{7.30}$$

$$\left|\partial_u a_6[h]\right|_s + \left|\partial_u a_7[h]\right|_s + \left|\partial_u a_8[h]\right|_s + \left|\partial_u a_9[h]\right|_s \leqslant \varepsilon^4 C(s, K)\big(\|h\|_{s+4} + \|u\|_{s+6}\|h\|_5\big), \tag{7.31}$$

$$|\partial_\varepsilon a_6|_s + \left|\partial_\varepsilon a_7 - \varepsilon 6\bar v^2\right|_s + |\partial_\varepsilon a_8|_s + \left|\partial_\varepsilon a_9 - \varepsilon\big(6\bar v^2\big)_x\right|_s \leqslant \varepsilon^2 C(s, K)\big(1 + \|u\|_{s+6}\big). \tag{7.32}$$

For $s, m_1, m_2 \geqslant 0$, $m = m_1 + m_2$, $m + s + 1 \leqslant r$,

$$\left\|\partial_x^{m_1} \mathcal R_{\mathcal H} \partial_x^{m_2} f\right\|_s \leqslant \varepsilon^3 C(s, m, K)\big(\|f\|_s\big(1 + \|u\|_{m+5}\big) + \|u\|_{s+m+5}\|f\|_0\big). \tag{7.33}$$

For $m, s \geqslant 0$, $m + s + 3 \leqslant r$,

$$\left\|\mathcal R_i \partial_y^m f\right\|_s \leqslant \varepsilon^3 C(s, m, K)\big(\|f\|_s\big(1 + \|u\|_{m+7}\big) + \|f\|_0\|u\|_{s+m+7}\big), \quad i = 1, 2, 3. \tag{7.34}$$

Proof. In Appendix C. □

Remark 7.3. The loss of one derivative for the difference $\Psi - I$ in (7.22), (7.24) is typical of any change of variables: in general, if we want to estimate a difference $h(x + p(x)) - h(x)$ with a factor of size p, we can do nothing but making a derivative, $h(x + p(x)) - h(x) \simeq h'(x)p(x)$.

7.2. Descent method: conjugation with pseudo-differential operators

We construct an invertible linear operator $\tilde\Phi = \mathbb P \Phi \mathbb P$ that maps $X_0 \to X_0$ and $Y \to Y$ and conjugates $\tilde{\mathcal L}_3$ to a new operator

$$\tilde{\mathcal L}_4 := \tilde\Phi^{-1} \tilde{\mathcal L}_3 \tilde\Phi = \mathbb P \mathcal L_4 \mathbb P, \quad \mathcal L_4 = \mathcal D + \mathcal R, \tag{7.35}$$

where $\mathcal D$ has constant coefficients and the remainder $\mathcal R$ is regularizing in space, bounded in time. We look for $\mathcal D$ of the form

$$\mathcal D = \omega\partial_\tau + \mu_2\partial_{yy}\mathcal H + \mu_1\partial_y + \nu_0' + \nu_0\mathcal H + \big(\nu_{-1}' + \nu_{-1}\mathcal H\big)\partial_y^{-1} + \big(\nu_{-2}' + \nu_{-2}\mathcal H\big)\partial_y^{-2},$$

where μ_2, μ_1 are the constants calculated in the previous section, ν_k, ν_k', $k = 0, -1, -2$ are constants to be determined. We look for Φ such that $(\mathbb P \mathcal L_3 \mathbb P)(\mathbb P \Phi \mathbb P) - (\mathbb P \Phi \mathbb P)(\mathbb P \mathcal D \mathbb P)$ is an operator of order $\leqslant -3$ in y. Write Φ as

$$\Phi = \Phi_0 + \Phi_1 + \Phi_2 + \Phi_3, \quad \Phi_k = \big(\alpha^{(k)} + \mathcal H\beta^{(k)}\big)\partial_y^{-k}, \ k = 0, 1, 2, 3,$$

namely $\Phi_k h = \alpha^{(k)}\partial_y^{-k} h + \mathcal H(\beta^{(k)}\partial_y^{-k} h)$, where $\alpha^{(k)}(\tau, y)$, $\beta^{(k)}(\tau, y)$ are functions to be determined. Φ is close to the identity if $\alpha^{(0)}$ is close to 1 and all the other $\alpha^{(k)}, \beta^{(k)}$ are small.

Calculate and write the terms of order $1, 0, -1, -2$ in y, and move all the '$\mathcal{H}$' on the left-hand side, introducing the corresponding commutators (for example, write $\alpha\mathcal{H}$ as $\mathcal{H}\alpha + [\alpha, \mathcal{H}]$). Note that

$$\mathcal{H}^2 = \mathcal{H}\mathcal{H} = -\Pi_E = -I + \Pi_E^{\perp}, \quad \Pi_E^{\perp} := I - \Pi_E = \Pi_T + \Pi_C.$$

$\Pi_E^{\perp}$ is regularizing in y because it is the operator that takes the mean of a function with respect to y. Therefore, up to a regularizing rest, sums and products of terms of the type $(\alpha + \mathcal{H}\beta)$ follow the same algebraic rules as those of complex numbers, where the role of i is played by $\mathcal{H}$. As a consequence, to perform the calculations up to terms containing $\Pi_E^{\perp}$ or commutators with $\mathcal{H}$ it is comfortable to introduce the complex notation:

$$\begin{cases} f^{(k)} := \alpha^{(k)} + i\beta^{(k)}, \qquad \mathcal{L}_3 = \omega\partial_\tau + \mu_2 i\partial_{yy} + a_{76}\partial_y + a_{98} + \mathcal{R}_3, \quad a_{76} := a_7 + ia_6,\ a_{98} := a_9 + ia_8, \\ \mathcal{D} = \omega\partial_\tau + \mu_2 i\partial_{yy} + \mu_1\partial_y + c_0 + c_{-1}\partial_y^{-1} + c_{-2}\partial_y^{-2}, \quad c_{-k} := \nu'_{-k} + i\nu_{-k}, \\ \text{where } i \text{ means } \mathcal{H}. \end{cases}$$

We stress that this is only a notation, as $\mathcal{H}$ maps real-valued functions into real-valued functions, and therefore $\alpha + \mathcal{H}\beta$ is real when α, β are real. Straightforward calculations (use $\mathbb{P} = I - \Pi_C$ for a_9) give

$$\tilde{\mathcal{L}}_3\tilde{\Phi} - \tilde{\Phi}\tilde{\mathcal{D}} = \mathbb{P}\big(T_1\partial_y + T_0 + T_{-1}\partial_y^{-1} + +T_{-2}\partial_y^{-2} + \mathcal{R}_4\big)\mathbb{P}, \tag{7.36}$$

where the coefficients T_k are

$$\begin{aligned} &T_1 = Qf^{(0)}, \qquad T_{-1} = Qf^{(2)} + Sf^{(1)} - c_{-1}f^{(0)}, \\ &T_0 = Qf^{(1)} + Sf^{(0)}, \qquad T_{-2} = Qf^{(3)} + Sf^{(2)} - c_{-1}f^{(1)} - c_{-2}f^{(0)}, \end{aligned} \tag{7.37}$$

Q, S mean

$$Qf := 2i\mu_2 f_y + (a_{76} - \nu)f, \qquad Sf := (\mathcal{L}_3 - \mathcal{R}_3 - c_0)f = \omega f_\tau + i\mu_2 f_{yy} + a_{76}f_y + (a_{98} - c_0)f,$$

and the rest $\mathcal{R}_4$ is the sum $\mathcal{R}_3\mathbb{P}\Phi - a_9\Pi_C\Phi$ + terms of order ∂_y^{-3} + other regularizing terms that

(a) contain a commutator $[g, \mathcal{H}]$, where $g \in \{a_j, \alpha^{(k)}, \beta^{(k)}\colon j = 6, 7, 8, 9,\ k = 0, 1, 2, 3\}$; or
(b) contain $\Pi_E^{\perp}$.

The complete formula for $\mathcal{R}_4$ is in Appendix C. For example, typical terms are

$$\Pi_E^{\perp}\beta^{(0)}\partial_y^2, \qquad a_6\Pi_E^{\perp}\beta_y^{(1)}\partial_y^{-1}, \qquad [a_6, \mathcal{H}]\alpha_y^{(0)}, \qquad \big[\beta^{(1)}, \mathcal{H}\big]\partial_y.$$

Now we choose $\nu_i, \alpha^{(k)}, \beta^{(k)}$ such that all T_n, $n = 1, 0, -1, -2$, vanish. Every T_n is an operator of the form $T_n h = p_n h + \mathcal{H}(q_n h)$ for some functions $p_n(\tau, y), q_n(\tau, y)$. Thus $T_n = 0$ if

$$p_n = 0, \qquad q_n = 0. \tag{7.38}$$

To solve (7.38), which is a system of two equations in the real-valued unknowns $\alpha^{(k)}, \beta^{(k)}$, we use complex notation again. Consider the complex-valued unknown $f^{(k)} = \alpha^{(k)} + i\beta^{(k)}$, where now i is the standard imaginary unit of $\mathbb{C}$. Then the real system (7.38) is equivalent to the complex ODE $Qf^{(0)} = 0$ for $n = 1$, and similar complex equations for $n = 0, -1, -2$, according to (7.37). Hence we look for complex-valued solutions $f^{(k)}$ of the four complex equations $T_n = 0$, $n = 1, 0, -1, -2$.

Reduction of T_1. Let

$$a_{76}^E(\tau, y) := a_{76}(\tau, y) - \mu_1 = a_7(\tau, y) - \mu_1 + ia_6(\tau, y).$$

Remember that $a_7 - \nu, a_6 \in Z_E$ (see (7.10), (7.13)). $T_1 = 0$ if

$$Qf^{(0)} = 2i\mu_2 f_y^{(0)} + a_{76}^E(\tau, y)f^{(0)} = 0. \tag{7.39}$$

The solutions of (7.39) are the exponentials $f^{(0)} = \exp(\varphi)$, where $\varphi(\tau, y)$ satisfies

$$2i\mu_2\varphi_y + a_{76}^E(\tau, y) = 0. \tag{7.40}$$

(7.40) determines the Z_E-component of φ,

$$(\Pi_E\varphi)(\tau, y) = \frac{i}{2\mu_2}\big(\partial_y^{-1}a_{76}^E\big)(\tau, y) = -\frac{1}{2\mu_2}\big(\partial_y^{-1}a_6\big)(\tau, y) + i\frac{1}{2\mu_2}\big(\partial_y^{-1}\Pi_E a_7\big)(\tau, y).$$

Reduction of T_0. Since $f^{(0)} = \exp(\varphi)$,

$$Sf^{(0)} = f^{(0)} g^{(0)}, \quad g^{(0)} := \omega\varphi_\tau + i\mu_2(\varphi_y^2 + \varphi_{yy}) + a_{76}\varphi_y + (a_{98} - c_0). \tag{7.41}$$

Moreover

$$i\mu_2\varphi_y^2 + a_{76}\varphi_y = \frac{i}{4\mu_2}(a_{76}^E)^2 + \frac{i}{2\mu_2}\nu a_{76}^E$$

by (7.40) and because $a_{76} = a_{76}^E + \nu$. Since $Qf^{(0)} = 0$, we solve the equation $T_0 = 0$ by variation of constants: $f^{(1)} = \eta^{(1)} f^{(0)}$ is a solution of $T_0 = Qf^{(1)} + Sf^{(0)} = 0$ if $\eta^{(1)}$ solves

$$2i\mu_2\eta_y^{(1)} + g^{(0)} = 0. \tag{7.42}$$

(7.42) has a periodic solution $\eta^{(1)}$ if $g^{(0)} \in Z_E$. The condition

$$\Pi_C(g^{(0)}) = \frac{i}{4\mu_2}\Pi_C((a_{76}^E)^2) + \Pi_C(a_{98}) - c_0 = 0$$

determines the constant c_0,

$$c_0 = \frac{i}{4\mu_2}\Pi_C((a_{76}^E)^2) + \Pi_C(a_{98}) \in \mathbb{C}.$$

The condition

$$\Pi_T(g^{(0)}) = \omega(\Pi_T\varphi)_\tau + \frac{i}{4\mu_2}\Pi_T((a_{76}^E)^2) + \Pi_T(a_{98}) = 0$$

determines the Z_T-component of φ,

$$(\Pi_T\varphi)(\tau) = -\frac{i}{4\mu_2\omega}(\partial_\tau^{-1}\Pi_T(a_{76}^E)^2)(\tau) - \frac{1}{\omega}(\partial_\tau^{-1}\Pi_T a_{98})(\tau) \in Z_T.$$

So $g^{(0)} \in Z_E$, (7.42) can be solved, and the Z_E-component of $\eta^{(1)}$ is determined,

$$(\Pi_E\eta^{(1)})(\tau, y) = \frac{i}{2\mu_2}(\partial_y^{-1}g^{(0)})(\tau, y) \in Z_E. \tag{7.43}$$

Reduction of T_{-1}. Since $f^{(1)} = \eta^{(1)} f^{(0)}$, $Sf^{(0)} = f^{(0)} g^{(0)}$, by (7.40) and the definition of S,

$$Sf^{(1)} - c_{-1}f^{(0)} = \eta^{(1)}Sf^{(0)} + \eta_y^{(1)}[2i\mu_2 f_y^{(0)} + a_{76}f^{(0)}] + f^{(0)}[\omega\eta_\tau^{(1)} + i\mu_2\eta_{yy}^{(1)} - c_{-1}] = f^{(0)}g^{(1)},$$

where

$$g^{(1)} := \eta^{(1)}g^{(0)} + \omega\eta_\tau^{(1)} + i\mu_2\eta_{yy}^{(1)} + \mu_1\eta_y^{(1)} - c_{-1}. \tag{7.44}$$

By variation of constants, $f^{(2)} = \eta^{(2)} f^{(0)}$ is a solution of $T_{-1} = Qf^{(2)} + Sf^{(1)} - c_{-1}f^{(0)} = 0$ if $\eta^{(2)}$ solves

$$2i\mu_2\eta_y^{(2)} + g^{(1)} = 0. \tag{7.45}$$

(7.45) has a periodic solution $\eta^{(2)}$ if $g^{(1)} \in Z_E$. By (7.42), $g^{(0)} = -2i\mu_2\eta_y^{(1)}$, therefore

$$\eta^{(1)}g^{(0)} = -2i\mu_2\eta^{(1)}\eta_y^{(1)} = -i\mu_2\partial_y\{(\eta^{(1)})^2\} \in Z_E.$$

As a consequence, the condition $g^{(1)} \in Z_E$ determines

$$\Pi_T(\eta^{(1)}) = 0, \qquad c_{-1} = 0. \tag{7.46}$$

Thus (7.45) can be solved, and the Z_E-component of $\eta^{(2)}$ is determined,

$$(\Pi_E\eta^{(2)})(\tau, y) = \frac{i}{2\mu_2}(\partial_y^{-1}g^{(1)})(\tau, y). \tag{7.47}$$

Reduction of T_{-2}. Since $c_{-1} = 0$, $T_{-2} = Qf^{(3)} + Sf^{(2)} - c_{-2}f^{(0)}$. By the same calculations as above,

$$Sf^{(2)} - c_{-2}f^{(0)} = \eta^{(2)}Sf^{(0)} + \eta_y^{(2)}[2i\mu_2 f_y^{(0)} + a_{76}f^{(0)}] + f^{(0)}[\omega\eta_\tau^{(2)} + i\mu_2\eta_{yy}^{(2)} - c_{-2}] = f^{(0)}g^{(2)},$$

where

$$g^{(2)} := \eta^{(2)} g^{(0)} + \omega \eta^{(2)}_\tau + i\mu_2 \eta^{(2)}_{yy} + \mu_1 \eta^{(2)}_y - c_{-2}. \tag{7.48}$$

By variation of constants, $f^{(3)} = \eta^{(3)} f^{(0)}$ is a solution of $T_{-2} = Qf^{(3)} + Sf^{(2)} - c_{-2} f^{(0)} = 0$ if $\eta^{(3)}$ solves

$$2i\mu_2 \eta^{(3)}_y + g^{(2)} = 0. \tag{7.49}$$

(7.49) has a periodic solution $\eta^{(3)}$ if $g^{(2)} \in Z_E$. Both $(\Pi_T \eta^{(2)}) g^{(0)}$ and $(\Pi_C \eta^{(2)}) g^{(0)}$ belongs to Z_E because $g^{(0)} \in Z_E$. Hence

$$\Pi_T\big(\eta^{(2)} g^{(0)}\big) = \Pi_T\big[\big(\Pi_C \eta^{(2)}\big) g^{(0)} + \big(\Pi_T \eta^{(2)}\big) g^{(0)} + \big(\Pi_E \eta^{(2)}\big) g^{(0)}\big] = \Pi_T\big[\big(\Pi_E \eta^{(2)}\big) g^{(0)}\big],$$

and the same for $\Pi_C(\eta^{(2)} g^{(0)})$. $\Pi_E \eta^{(2)}$ is given by (7.47). The condition $\Pi_T g^{(2)} = 0$ determines

$$\Pi_T \eta^{(2)} = -\frac{1}{\omega} \partial_\tau^{-1} \Pi_T\big[\big(\Pi_E \eta^{(2)}\big) g^{(0)}\big], \tag{7.50}$$

the condition $\Pi_C g^{(2)} = 0$ determines

$$c_{-2} = \Pi_C\big[\big(\Pi_E \eta^{(2)}\big) g^{(0)}\big].$$

Thus $g^{(2)} \in Z_E$, (7.49) can be solved, and the Z_E-component of $\eta^{(3)}$ is determined,

$$\big(\Pi_E \eta^{(3)}\big)(\tau, y) = \frac{i}{2\mu_2}\big(\partial_y^{-1} g^{(2)}\big)(\tau, y). \tag{7.51}$$

The only terms that have not been determined by the four equations $T_1 = 0, \ldots, T_{-2} = 0$ are $\Pi_C(\varphi)$, $\Pi_C(\eta^{(1)})$, $\Pi_C(\eta^{(2)})$, $\Pi_C(\eta^{(3)})$, and $\Pi_T(\eta^{(3)})$. Fix all of them to be 0. Split real and imaginary part,

$$\mathrm{Re}(\varphi) = \frac{1}{2\mu_2\omega} \partial_\tau^{-1} \Pi_T\big[(\Pi_E a_7) a_6\big] - \frac{1}{\omega} \partial_\tau^{-1} \Pi_T(a_9) - \frac{1}{2\mu_2}\big(\partial_y^{-1} a_6\big), \tag{7.52}$$

$$\mathrm{Im}(\varphi) = -\frac{1}{4\mu_2\omega} \partial_\tau^{-1} \Pi_T\big[(\Pi_E a_7)^2 - (a_6)^2\big] - \frac{1}{\omega} \partial_\tau^{-1} \Pi_T(a_8) + \frac{1}{2\mu_2}\big(\partial_y^{-1} \Pi_E a_7\big), \tag{7.53}$$

$$\alpha^{(0)} = e^{\mathrm{Re}(\varphi)} \cos\big(\mathrm{Im}(\varphi)\big), \qquad \beta^{(0)} = e^{\mathrm{Re}(\varphi)} \sin\big(\mathrm{Im}(\varphi)\big). \tag{7.54}$$

By (7.14),

$$\mathrm{Re}(\varphi) \in X, \qquad \mathrm{Im}(\varphi) \in Y, \qquad \alpha^{(0)} \in X, \qquad \beta^{(0)} \in Y.$$

As a consequence, $g^{(0)}, \eta^{(1)}, g^{(2)}, \eta^{(3)} \in Y + iX$, $g^{(1)}, \eta^{(2)} \in X + iY$, and

$$\alpha^{(1)} \in Y, \qquad \beta^{(1)} \in X, \qquad \alpha^{(2)} \in X, \qquad \beta^{(2)} \in Y, \qquad \alpha^{(3)} \in Y, \qquad \beta^{(3)} \in X.$$

Hence Φ preserves the parity, namely Φ maps $X \to X$ and $Y \to Y$.

By (7.14), $(\Pi_E a_7) a_6 \in Y$, $a_9 \in Y$, therefore

$$\nu_0' = \mathrm{Re}(c_0) = 0, \qquad \nu_0 = \mathrm{Im}(c_0) = \frac{1}{4\mu_2} \Pi_C\big[(\Pi_E a_7)^2 - a_6^2\big] + \Pi_C(a_8). \tag{7.55}$$

$\nu_{-1} = \nu_{-1}' = 0$, and

$$\nu_{-2}' = \mathrm{Re}(c_{-2}) = 0, \qquad \nu_{-2} = \mathrm{Im}(c_{-2}) = \mathrm{Im}\big\{\Pi_C\big[\big(\Pi_E \eta^{(2)}\big) g^{(0)}\big]\big\}. \tag{7.56}$$

Put

$$\mu_0 := \nu_0, \qquad \mu_{-2} := \nu_{-2}.$$

Since T_1, T_0, T_{-1}, T_{-2} vanish, (7.36) becomes $\tilde{\mathcal{L}}_3 \tilde{\Phi} - \tilde{\Phi} \tilde{\mathcal{D}} = \mathbb{P}\mathcal{R}_4\mathbb{P}$, and (7.35) holds with

$$\mathcal{L}_4 = \mathcal{D} + \mathcal{R}, \quad \mathcal{D} = \omega\partial_\tau + \mu_2 \mathcal{H} \partial_{yy} + \mu_1 \partial_y + \mu_0 \mathcal{H} + \mu_{-2} \mathcal{H} \partial_y^{-2}, \ \mathcal{R} := \tilde{\Phi}^{-1} \mathbb{P} \mathcal{R}_4. \tag{7.57}$$

If $\tilde{\Phi}$ is invertible, we have transformed $\tilde{\mathcal{L}}$ into $\tilde{\mathcal{L}}_4$, namely

$$\tilde{\mathcal{L}} = \tilde{\Psi} \tilde{\mathcal{M}} \tilde{\Phi} \tilde{\mathcal{L}}_4 \tilde{\Phi}^{-1} \tilde{\Psi}^{-1}, \quad \tilde{\mathcal{L}}_4 = \tilde{\Phi}^{-1} \tilde{\mathcal{M}}^{-1} \tilde{\Psi}^{-1} \tilde{\mathcal{L}} \tilde{\Psi} \tilde{\Phi}. \tag{7.58}$$

From the formulae above, $\mu_0, \mu_{-2}, \alpha^{(k)}, \beta^{(k)}$ are C^1 functions of (u, ε), and the following estimates hold.

Proposition 7.4. *Let $K > 0$. There exists $\varepsilon_0 \in (0,1)$, depending on K, such that, if $\varepsilon \in (0,\varepsilon_0)$, $\|u\|_{19} \leqslant K$, and $\|u\|_4, \varepsilon_0$ satisfy (6.4), then all the following inequalities hold.*

$$|\mu_0| \leqslant \varepsilon^3 C(K), \qquad \left|\partial_u \mu_0[h]\right| \leqslant \varepsilon^4 C(K)\|h\|_5, \qquad |\partial_\varepsilon \mu_0| \leqslant \varepsilon^2 C(K), \tag{7.59}$$

$$|\mu_{-2}| \leqslant \varepsilon^4 C(K), \qquad \left|\partial_u \mu_{-2}[h]\right| \leqslant \varepsilon^6 C(K)\|h\|_{12}, \qquad |\partial_\varepsilon \mu_{-2}| \leqslant \varepsilon^3 C(K). \tag{7.60}$$

The operator $\tilde\Phi : Z_0 \to Z_0$ is invertible, and maps $X_0 \to X_0$ and $Y \to Y$. $\tilde\Phi, \tilde\Phi^{-1}$ satisfy

$$\left\|(\tilde\Phi - I)f\right\|_s + \left\|\left(\tilde\Phi^{-1} - I\right)f\right\|_s \leqslant \varepsilon^2 C(s,K)\big(\|f\|_s + \|u\|_{s+12}\|f\|_2\big) \quad \forall f \in Z_0, \tag{7.61}$$

for all $2 \leqslant s \leqslant r-7$. The derivatives of $\tilde\Phi f, \tilde\Phi^{-1} f$ with respect to u in the direction h and with respect to ε satisfy

$$\left\|\partial_u(\tilde\Phi f)[h]\right\|_s + \left\|\partial_u\big(\tilde\Phi^{-1} f\big)[h]\right\|_s \leqslant \varepsilon^4 C(s,K)\big(\|f\|_s\|h\|_{14} + \|f\|_2\|h\|_{s+12} + \|u\|_{s+12}\|f\|_2\|h\|_{14}\big), \tag{7.62}$$

$$\|\partial_\varepsilon \tilde\Phi f\|_s + \left\|\partial_\varepsilon \tilde\Phi^{-1} f\right\|_s \leqslant \varepsilon C(s,K)\big(\|f\|_s + \|u\|_{s+12}\|f\|_2\big). \tag{7.63}$$

Moreover

$$\left\|\partial_\tau(\tilde\Phi - I)f\right\|_s \leqslant \varepsilon^2 C(s,K)\big(\|\partial_\tau f\|_s + \|f\|_s + \|u\|_{s+13}\big(\|\partial_\tau f\|_2 + \|f\|_2\big)\big), \tag{7.64}$$

$$\left\|\partial_y^k(\tilde\Phi - I)f\right\|_s \leqslant \varepsilon^2 C(s,K)\big(\left\|\partial_y^k f\right\|_s + \|f\|_s + \|u\|_{s+14}\big(\left\|\partial_y^k f\right\|_2 + \|f\|_2\big)\big), \quad k = 1,2, \tag{7.65}$$

for $2 \leqslant s \leqslant r-9$, for all $f \in Z_0$.

The operators $\tilde\Psi\tilde\Phi$, $\tilde\Psi\tilde{\mathcal M}\tilde\Phi$, $\tilde\Phi^{-1}\tilde\Psi^{-1}$, $\tilde\Phi^{-1}\tilde{\mathcal M}^{-1}\tilde\Psi^{-1}$ are all of the type $I+S$, where S satisfies

$$\|Sf\|_s \leqslant \varepsilon^2 C(s,K)\big(\|f\|_{s+1} + \|u\|_{s+12}\|f\|_2\big), \quad 2 \leqslant s \leqslant r-7. \tag{7.66}$$

The rest $\mathcal R$ satisfies

$$\left\|\mathcal R \partial_y^m f\right\|_s \leqslant \varepsilon^2 C(s,K)\big(\|f\|_s + \|u\|_{s+17}\|f\|_2\big), \quad 0 \leqslant m \leqslant 3,\ 2 \leqslant s \leqslant r-12. \tag{7.67}$$

Proof. The proof is in Appendix C. □

8. Inversion of the transformed linearized operator

In view of the Nash–Moser iteration, we invert $\tilde{\mathcal L}_4 = \tilde{\mathcal D} + \tilde{\mathcal R}$ on a subspace of Fourier-truncated functions. Let

$$Z_N := \left\{ u = \sum_{|k| \leqslant N} u_k e_k \right\} \subset Z, \quad k = (l,j) \in \mathbb Z^2,\ |k| = |l| + |j|,\ Z_{0N} := Z_0 \cap Z_N,$$

with $N > 0$ sufficiently large to have $\bar v \in Z_N$, namely $\mathcal K \subseteq [-N,N]$, where $\mathcal K$ is defined in Section 5 (see Proposition 5.3). Let $\Pi_N, \Pi_N^\perp$ denote the orthogonal projections onto Z_N and $Z_N^\perp$ respectively. Let

$$X_{0N} := X_0 \cap Z_N, \qquad Y_N := Y \cap Z_N, \qquad V_{0N} := V_0 \cap Z_N, \qquad W_N := W \cap Z_N.$$

$\Pi_N \tilde{\mathcal L}_4 \Pi_N$ maps $X_{0N} \to Y_N$ because $\tilde{\mathcal L}_4 : X_0 \to Y$. Since $Z_{0N} = V_{0N} \oplus W_N$, to prove that $\Pi_N \tilde{\mathcal L}_4 \Pi_N : X_{0N} \to Y_N$ is invertible, we project on the subspaces V_{0N} and W_N (Lyapunov–Schmidt decomposition, like in Section 4): given $f \in Y_N$,

$$\Pi_N \tilde{\mathcal L}_4 \Pi_N h = f \quad \Longleftrightarrow \quad \begin{cases} \Pi_{V_{0N}} \tilde{\mathcal L}_4 \Pi_{V_{0N}} h + \Pi_{V_{0N}} \tilde{\mathcal L}_4 \Pi_{W_N} h = \Pi_{V_{0N}} f, \\ \Pi_{W_N} \tilde{\mathcal L}_4 \Pi_{V_{0N}} h + \Pi_{W_N} \tilde{\mathcal L}_4 \Pi_{W_N} h = \Pi_{W_N} f. \end{cases} \tag{8.1}$$

Since $\mathcal D$ is diagonal, $\mathcal D$ maps $V \to V$ and $W \to W$, therefore

$$\Pi_V \tilde{\mathcal L}_4 \Pi_W = \Pi_V \tilde{\mathcal R} \Pi_W, \qquad \Pi_W \tilde{\mathcal L}_4 \Pi_V = \Pi_W \tilde{\mathcal R} \Pi_V. \tag{8.2}$$

Lemma 8.1 *(Inversion on V_{0N}). Let $K > 0$. There exists $\varepsilon_0 \in (0,1)$, depending on K, such that, if $\varepsilon \in (0,\varepsilon_0)$, $\|u\|_{19} \leqslant K$, and $\|u\|_4, \varepsilon_0$ satisfy (6.4), then*

$$\Pi_{V_{0N}} \tilde{\mathcal L}_4 \Pi_{V_{0N}} : V_{0N} \cap X_0 \to V_{0N} \cap Y$$

is invertible, with

$$\left\|(\Pi_{V_{0N}} \tilde{\mathcal L}_4 \Pi_{V_{0N}})^{-1} h\right\|_s \leqslant \frac{C(s,K)}{\varepsilon^2}\big(\|h\|_{s-1} + \|u\|_{s+13}\|h\|_2\big), \quad 3 \leqslant s \leqslant r-8. \tag{8.3}$$

Proof. $\tilde{\mathcal{L}}_4 = \tilde{\Phi}^{-1}\tilde{\mathcal{L}}_3\tilde{\Phi}$ (see (7.35)). Split $\mathcal{L}_3 = L + \varepsilon^2 A + \varepsilon^3 B$, where

$$L = \partial_\tau + \partial_{yy}\mathcal{H}, \qquad Ah = 3\partial_\tau h + \partial_y\big(3\bar{v}^2 h\big),$$
$$B = \varepsilon^{-3}\big\{(\mu_2 - 1)\partial_{yy}\mathcal{H} + a_6\partial_y\mathcal{H} + \big(a_7 - \varepsilon^2 3\bar{v}^2\big)\partial_y + a_8\mathcal{H} + \big(a_9 - \varepsilon^2\big(3\bar{v}^2\big)_y\big) + \mathcal{R}_3\big\}.$$

By (7.16),(7.30),(7.34),

$$\|Bh\|_s \leqslant C(s,K)\big(\|h_{yy}\|_s + \|h_y\|_s + \|h\|_s + \|u\|_{s+7}\big(\|h_y\|_0 + \|h\|_0\big)\big), \quad 2 \leqslant s \leqslant r-3. \tag{8.4}$$

Let $S_i : Z_0 \to Z_0$, $S_1 := \varepsilon^{-2}(\tilde{\Phi} - I)$, $S_2 := \varepsilon^{-2}(\tilde{\Phi}^{-1} - I)$ (recall that $\mathbb{P} = I$ on Z_0). Since $\Pi_V L = L\Pi_V = 0$,

$$\Pi_{V_{0N}}\tilde{\mathcal{L}}_4\Pi_{V_{0N}} = \Pi_{V_{0N}}\tilde{\Phi}^{-1}\tilde{\mathcal{L}}_3\tilde{\Phi}\Pi_{V_{0N}} = \Pi_{V_{0N}}\big(I + \varepsilon^2 S_2\big)\mathbb{P}\big(L + \varepsilon^2 A + \varepsilon^3 B\big)\mathbb{P}\big(I + \varepsilon^2 S_1\big)\Pi_{V_{0N}}$$
$$= \varepsilon^2\Pi_{V_{0N}}(A + \varepsilon B_1)\Pi_{V_{0N}}, \tag{8.5}$$

where

$$B_1 = \varepsilon S_2\mathbb{P}L\mathbb{P}S_1 + \varepsilon S_2\mathbb{P}A + \varepsilon A\mathbb{P}S_1 + \varepsilon^3 S_2\mathbb{P}A\mathbb{P}S_1 + \tilde{\Phi}^{-1}\mathbb{P}B\mathbb{P}\tilde{\Phi}.$$

By Proposition 5.3, $\Pi_{V_{0N}}A\Pi_{V_{0N}} : V_{0N}\cap X_0 \to V_{0N}\cap Y$ is invertible, with

$$\big\|(\Pi_{V_{0N}}A\Pi_{V_{0N}})^{-1}h\big\|_s \leqslant C\|h\|_{s-1} \quad \forall h \in V_{0N}\cap Y,\ \forall s \geqslant 0, \tag{8.6}$$

where $C > 0$ depends only on the set $\mathcal{K}$. By (7.61),(7.64),(7.65), for $2 \leqslant s \leqslant r-9$,

$$\|S_1 h\|_s + \|S_2 h\|_s \leqslant C(s,K)\big(\|h\|_s + \|u\|_{s+12}\|h\|_2\big),$$
$$\|\partial_\cdot S_1 h\|_s \leqslant C(s,K)\big(\|\partial_\cdot h\|_s + \|h\|_s + \|u\|_{s+14}\big(\|\partial_\cdot h\|_2 + \|h\|_2\big)\big), \quad \partial_\cdot = \partial_\tau, \partial_y, \partial_{yy},$$

for all $h \in Z_0$. Then, since $L = \partial_\tau + \mathcal{H}\partial_y^2$, $Ah = 3\partial_\tau h + 3\bar{v}^2\partial_y h + (3\bar{v}^2)_y h$, and by (8.4),

$$\|\Pi_{V_{0N}}B_1\Pi_{V_{0N}}h\|_s \leqslant C(s,K)\big(\|h\|_{s+1} + \|u\|_{s+14}\|h\|_3\big), \quad 2 \leqslant s \leqslant r-9, \tag{8.7}$$

because $\|\partial_y^2 h\|_s = \|\mathcal{H}\partial_y^2 h\|_s = \|\partial_\tau h\|_s \leqslant \|h\|_{s+1}$ for all $h \in V$. Thus, by (8.6), (8.7),

$$\big\|(\Pi_{V_{0N}}B_1\Pi_{V_{0N}})(\Pi_{V_{0N}}A\Pi_{V_{0N}})^{-1}h\big\|_s \leqslant C(s,K)\big(\|h\|_s + \|u\|_{s+14}\|h\|_2\big), \quad 2 \leqslant s \leqslant r-9,$$

for all $h \in V_{0N}\cap Y$. Since B_1 maps X into Y, $B_2 := (\Pi_{V_{0N}}B_1\Pi_{V_{0N}})(\Pi_{V_{0N}}A\Pi_{V_{0N}})^{-1}$ maps Y into Y. By standard Neumann series with tame estimates (see Lemma B.2), $I + \varepsilon B_2$ is invertible as an operator of $V_{0N}\cap Y$ onto itself, with

$$\big\|(I + \varepsilon B_2)^{-1}h\big\|_s \leqslant C(s,K)\big(\|h\|_s + \|u\|_{s+14}\|h\|_2\big), \quad 2 \leqslant s \leqslant r-9, \tag{8.8}$$

provided that $\varepsilon C(K) < 1/2$, for some $C(K) > 0$ depending on $K, K_{g,r}, \|\bar{v}\|_{19}$. By (8.6) and (8.8), $\Pi_{V_{0N}}(A + \varepsilon B_1)\Pi_{V_{0N}} = (I + \varepsilon B_2)^{-1}(\Pi_{V_{0N}}A\Pi_{V_{0N}}) : X_0\cap V_{0N} \to Y\cap V_{0N}$ is invertible, with

$$\big\|\big\{\Pi_{V_{0N}}(A + \varepsilon B_1)\Pi_{V_{0N}}\big\}^{-1}h\big\|_s \leqslant C(s,K)\big(\|h\|_{s-1} + \|u\|_{s+13}\|h\|_2\big), \quad 3 \leqslant s \leqslant r-8.$$

By (8.5) the thesis is proved. □

By Lemma 8.1, the V_{0N}-equation of system (8.1) can be solved for $\Pi_{V_{0N}}h$,

$$\Pi_{V_{0N}}h = (\Pi_{V_{0N}}\tilde{\mathcal{L}}_4\Pi_{V_{0N}})^{-1}[\Pi_{V_{0N}}f - \Pi_{V_{0N}}\tilde{\mathcal{L}}_4\Pi_{W_N}h]. \tag{8.9}$$

Substituting $\Pi_{V_{0N}}h$, and using (8.2), the W_N-equation of system (8.1) becomes

$$\mathcal{A}(\Pi_{W_N}h) = f_1, \tag{8.10}$$

where

$$\mathcal{A} := \Pi_{W_N}\tilde{\mathcal{L}}_4\Pi_{W_N} - (\Pi_{W_N}\tilde{\mathcal{R}}\Pi_{V_{0N}})(\Pi_{V_{0N}}\tilde{\mathcal{L}}_4\Pi_{V_{0N}})^{-1}(\Pi_{V_{0N}}\tilde{\mathcal{R}}\Pi_{W_N}), \tag{8.11}$$
$$f_1 := \Pi_{W_N}f - (\Pi_{W_N}\tilde{\mathcal{R}}\Pi_{V_{0N}})(\Pi_{V_{0N}}\tilde{\mathcal{L}}_4\Pi_{V_{0N}})^{-1}\Pi_{V_{0N}}f. \tag{8.12}$$

$\tilde{\mathcal{L}}_4 = \mathcal{D} + \tilde{\mathcal{R}}$, where $\mathcal{D} = \omega\partial_\tau + \mu_2\mathcal{H}\partial_{yy} + \mu_1\partial_y + \mu_0\mathcal{H} + \mu_{-2}\mathcal{H}\partial_y^{-2}$, which is (7.57). In the basis $\{e^{i(l\tau+jy)}\}_{l,j}$, $\mathcal{D}$ is diagonal with eigenvalues

$$\lambda_{l,j} = \lambda_{l,j}(u,\varepsilon) = i\big(\omega l + \mu_2 j|j| + \mu_1 j - \mu_0 \operatorname{sign}(j) - \mu_{-2}\operatorname{sign}(j)(ij)^{-2}\big), \tag{8.13}$$

where $\omega = 1 + 3\varepsilon^2$ and $\mu_i(u,\varepsilon)$ are C^1 functions of (u,ε). By (7.16), (7.17), (7.59), (7.60),

$$|\omega - 1| + |\mu_2 - 1| + |\mu_1| + |\mu_0| + |\mu_{-2}| < 1/2 \tag{8.14}$$

for $\varepsilon < \varepsilon_0$ sufficiently small. Remember the notation $\langle j\rangle = \max\{1, |j|\}$.

Lemma 8.2 *(Inversion on W_N). Let $K > 0$. There exists $\varepsilon_0 \in (0,1)$, depending on K, with the following property. Let $\varepsilon \in (0,\varepsilon_0)$, $\|u\|_{19} \leqslant K$, and assume that $\|u\|_4$, ε_0 satisfy (6.4). Let*

$$\big|\lambda_{l,j}(u,\varepsilon)\big| > \frac{1}{2\langle j\rangle^3} \quad \forall (l,j) \in \mathcal{W}_N, \tag{8.15}$$

where

$$\mathcal{W}_N := \big\{(l,j) \in \mathcal{W}\colon |j| \leqslant N\big\} = \big\{(l,j) \in \mathbb{Z}^2\colon l + j|j| \neq 0,\ |j| \leqslant N\big\}.$$

Then $\mathcal{A}: X_0 \cap W_N \to Y \cap W_N$ is invertible, with

$$\big\|\mathcal{A}^{-1}h\big\|_s \leqslant C(s,K)\big(\|h\|_{s+3/2} + \|u\|_{s+16+3/2}\|h\|_2\big), \quad 3/2 \leqslant s \leqslant r - 12 - 3/2. \tag{8.16}$$

Proof. Since $\tilde{\mathcal{L}}_4 = \tilde{\mathcal{D}} + \tilde{\mathcal{R}}$, we have $\mathcal{A} = \mathcal{D}_{W_N} + \mathcal{R}_{W_N}$, where

$$\mathcal{D}_{W_N} := \Pi_{W_N}\mathcal{D}\Pi_{W_N}, \qquad \mathcal{R}_{W_N} := \Pi_{W_N}\tilde{\mathcal{R}}\Pi_{W_N} - (\Pi_{W_N}\tilde{\mathcal{R}}\Pi_{V_{0N}})(\Pi_{V_{0N}}\tilde{\mathcal{L}}_4\Pi_{V_{0N}})^{-1}(\Pi_{V_{0N}}\tilde{\mathcal{R}}\Pi_{W_N}).$$

Like $\mathcal{A}$, also $\mathcal{D}_{W_N}$ and $\mathcal{R}_{W_N}$ map X into Y. $\mathcal{D}_{W_N}: W_N \to W_N$ is invertible because $\lambda_{l,j} \neq 0$ for all $(l,j) \in \mathcal{W}_N$. Let

$$\mathcal{U} := \partial_y^3 + \Pi_T + \Pi_C, \qquad \mathcal{U}e^{i(l\tau+jy)} = \mathcal{U}_j e^{i(l\tau+jy)}, \quad \mathcal{U}_j = (ij)^3\ \forall j \neq 0,\ \mathcal{U}_0 = 1.$$

$|\lambda_{l,j}||\mathcal{U}_j| > 1/2$ for every $(l,j) \in \mathcal{W}_N$ because $|\mathcal{U}_j| = \langle j\rangle^3$. As a consequence,

$$\big\|\mathcal{U}^{-1}\mathcal{D}_{W_N}^{-1}h\big\|_s \leqslant 2\|h\|_s \quad \forall h \in W_N,\ \forall s \geqslant 0.$$

By (7.67) and (8.3),

$$\|\mathcal{R}_{W_N}\mathcal{U}h\|_s \leqslant \big\|\mathcal{R}_{W_N}\partial_y^3 h\big\|_s + \big\|\mathcal{R}_{W_N}(\Pi_T + \Pi_C)h\big\|_s \leqslant \varepsilon^2 C(s,K)\big(\|h\|_s + \|u\|_{s+16}\|h\|_2\big)$$

for $3 \leqslant s \leqslant r - 12$, whence

$$\big\|\mathcal{R}_{W_N}\mathcal{D}_{W_N}^{-1}h\big\|_s = \big\|(\mathcal{R}_{W_N}\mathcal{U})\big(\mathcal{U}^{-1}\mathcal{D}_{W_N}^{-1}\big)h\big\|_s \leqslant \varepsilon^2 C(s,K)\big(\|h\|_s + \|u\|_{s+16}\|h\|_2\big), \quad 3 \leqslant s \leqslant r - 12.$$

For $s = 3$, $\|\mathcal{R}_{W_N}\mathcal{D}_{W_N}^{-1}h\|_3 \leqslant \varepsilon^2 C(K)\|h\|_3$. By Lemma B.2, $I + \mathcal{R}_{W_N}\mathcal{D}_{W_N}^{-1}$ is invertible on W_N, with

$$\big\|\big(I + \mathcal{R}_{W_N}\mathcal{D}_{W_N}^{-1}\big)^{-1}h\big\|_s \leqslant C(s,K)\big(\|h\|_s + \|u\|_{s+16}\|h\|_2\big), \quad 3 \leqslant s \leqslant r - 12,$$

if $\varepsilon^2 C(K) < 1/2$. Therefore $\mathcal{A} = (I + \mathcal{R}_{W_N}\mathcal{D}_{W_N}^{-1})\mathcal{D}_{W_N}$ is also invertible. Now $\|\mathcal{D}_{W_N}^{-1}h\|_s \leqslant C\|h\|_{s+3/2}$ because, for indices $(l,j) \in \mathcal{W}$ such that $|\lambda_{l,j}| < 1$, one has $|j|^2 \leqslant C|l|$ by the triangular inequality and (8.14), so that $1/|\lambda_{l,j}| \leqslant 2\langle j\rangle^3 \leqslant C\langle l\rangle^{3/2}$. Hence (8.16) follows. □

Remember the definition $P_\varepsilon := \varepsilon^2\Pi_V + \Pi_W$.

Lemma 8.3 *(Inversion of $\Pi_N\tilde{\mathcal{L}}_4\Pi_N$). Assume the hypotheses of Lemmata 8.1 and 8.2. Then for every $f \in Y_N$ there exists a unique $h \in X_{0N}$ such that $\Pi_N\tilde{\mathcal{L}}_4\Pi_N h = f$. The inverse operator $(\Pi_N\tilde{\mathcal{L}}_4\Pi_N)^{-1}$ maps $Y_N \to X_{0N}$, with*

$$\big\|(\Pi_N\tilde{\mathcal{L}}_4\Pi_N)^{-1}f\big\|_s \leqslant \varepsilon^{-2}C(s,K)\big(\|f\|_{s+3/2} + \|u\|_{s+17+3/2}\|f\|_2\big), \tag{8.17}$$

$$\big\|(\Pi_N\tilde{\mathcal{L}}_4\Pi_N)^{-1}P_\varepsilon f\big\|_s + \big\|P_\varepsilon(\Pi_N\tilde{\mathcal{L}}_4\Pi_N)^{-1}f\big\|_s \leqslant C(s,K)\big(\|f\|_{s+3/2} + \|u\|_{s+17+3/2}\|f\|_2\big), \tag{8.18}$$

$3/2 \leqslant s \leqslant r - 12 - 3/2.$

Proof. Use (8.1), (8.9), (8.10), (8.11), (8.12), (8.3) and (8.16). □

Lemma 8.4 *(Derivatives of $(\Pi_N \tilde{\mathcal{L}}_4 \Pi_N)^{-1}$). Let $K > 0$. There exists $\varepsilon_0 \in (0,1)$, depending on $4K$, with the following property.*

Let $\varepsilon \in (0, \varepsilon_0)$, $\|u\|_{22} \leqslant K$, assume that $\|u\|_4, \varepsilon_0$ satisfy (6.4), and that (8.15) holds. Then, for $2 \leqslant s \leqslant r - 18$,

$$\begin{aligned}
&\big\|\partial_u (\Pi_N \tilde{\mathcal{L}}_4 \Pi_N)^{-1}[h] f\big\|_s \leqslant \varepsilon^{-1} C(s,K)\big(\|f\|_{s+6}\|h\|_{14} + \|f\|_8\big(\|h\|_{s+16} + \|u\|_{s+23}\|h\|_{14}\big)\big),\\
&\big\|\partial_\varepsilon (\Pi_N \tilde{\mathcal{L}}_4 \Pi_N)^{-1} f\big\|_s \leqslant \varepsilon^{-3} C(s,K)\big(\|f\|_{s+6} + \|u\|_{s+23}\|f\|_8\big),\\
&\big\|\partial_u (\Pi_N \tilde{\mathcal{L}}_4 \Pi_N)^{-1}[h] P_\varepsilon f\big\|_s + \big\|P_\varepsilon \partial_u (\Pi_N \tilde{\mathcal{L}}_4 \Pi_N)^{-1}[h] f\big\|_s\\
&\quad \leqslant \varepsilon C(s,K)\big(\|f\|_{s+6}\|h\|_{14} + \|f\|_8\big(\|h\|_{s+16} + \|u\|_{s+23}\|h\|_{14}\big)\big),\\
&\big\|\big\{\partial_\varepsilon (\Pi_N \tilde{\mathcal{L}}_4 \Pi_N)^{-1}\big\} P_\varepsilon f\big\|_s + \big\|P_\varepsilon \big\{\partial_\varepsilon (\Pi_N \tilde{\mathcal{L}}_4 \Pi_N)^{-1}\big\} f\big\|_s \leqslant \varepsilon^{-1} C(s,K)\big(\|f\|_{s+6} + \|u\|_{s+23}\|f\|_8\big).
\end{aligned}$$

Proof. By Proposition 6.1, for all $0 \leqslant s \leqslant r$,

$$\begin{aligned}
&\|\tilde{\mathcal{L}} f\|_s \leqslant C(s,K)\big(\|f\|_{s+2} + \|u\|_{s+4}\|f\|_2\big),\\
&\big\|\partial_u \tilde{\mathcal{L}}[h] f\big\|_s \leqslant \varepsilon^3 C(s,K)\big(\|f\|_{s+2}\|h\|_4 + \|f\|_2\big(\|h\|_{s+4} + \|u\|_{s+4}\|h\|_4\big)\big),\\
&\|\partial_\varepsilon \tilde{\mathcal{L}} f\|_s \leqslant \varepsilon C(s,K)\big(\|f\|_{s+2} + \|u\|_{s+4}\|f\|_2\big).
\end{aligned}$$

Hence, from formula (7.58), using the estimates (7.25), (7.26), (7.28), (7.29), (7.62), (7.63) for $\tilde{\Phi}, \tilde{\Psi}, \tilde{\mathcal{M}}$ and their inverse,

$$\begin{aligned}
&\|\tilde{\mathcal{L}}_4 f\|_s \leqslant C(s,K)\big(\|f\|_{s+2} + \|u\|_{s+14}\|f\|_2\big),\\
&\big\|\partial_u \tilde{\mathcal{L}}_4[h] f\big\|_s \leqslant \varepsilon^3 C(s,K)\big(\|f\|_{s+3}\|h\|_{14} + \|f\|_5\big(\|h\|_{s+14} + \|u\|_{s+15}\|h\|_{14}\big)\big),\\
&\|\partial_\varepsilon \tilde{\mathcal{L}}_4 f\|_s \leqslant \varepsilon C(s,K)\big(\|f\|_{s+3} + \|u\|_{s+15}\|f\|_5\big),
\end{aligned}$$

for $2 \leqslant s \leqslant r - 10$. The Lemma follows from formula (B.9) and Lemma 8.3. □

8.1. Further estimates

In this section we collect some tame estimates that will be used in the Nash–Moser iteration.

Lemma 8.5 *(Tame estimates for F).* (i) *There exists $\varepsilon_0 \in (0,1)$, depending only on $\|\bar{v}_1\|_5$, such that*

$$\varepsilon\|\bar{v}_1\|_4 + \varepsilon^2\|\bar{v}_2\|_4 < \delta_0, \qquad \big\|\bar{v}_2(\varepsilon)\big\|_s \leqslant C(s), \qquad \big\|\partial_\varepsilon \bar{v}_2(\varepsilon)\big\|_s \leqslant \varepsilon^{-1} C(s), \tag{8.19}$$

$$\big\|F\big(\bar{v}_2(\varepsilon), \varepsilon\big)\big\|_s \leqslant \varepsilon C(s), \qquad \big\|\partial_\varepsilon\big\{F\big(\bar{v}_2(\varepsilon), \varepsilon\big)\big\}\big\|_s \leqslant C(s), \tag{8.20}$$

for every $\varepsilon \in (0, \varepsilon_0)$, $2 \leqslant s \leqslant r$.

(ii) *Assume that ε_0, u, h satisfy $\varepsilon_0\|\bar{v}_1\|_4 + \varepsilon_0^2(\|u\|_4 + \|h\|_4) < \delta_0$ (δ_0 is the universal constant of (6.4)), and $\|u\|_4 + \|h\|_4 \leqslant K$. Let*

$$Q(u,h,\varepsilon) := F(u+h,\varepsilon) - F(u,\varepsilon) - \partial_u F(u,\varepsilon)[h]. \tag{8.21}$$

Then, for $2 \leqslant s \leqslant r$, $\varepsilon \in (0, \varepsilon_0)$,

$$\big\|Q(u,h,\varepsilon)\big\|_s \leqslant C(s,K)\|h\|_4\big(\|h\|_{s+2} + \|u\|_{s+2}\|h\|_4\big). \tag{8.22}$$

(iii) *Assume that $\varepsilon_0\|\bar{v}_1\|_4 + \varepsilon_0^2\|u\|_4 < \delta_0$, namely (6.4), and $\|u\|_4 \leqslant K$. Then*

$$\big\|F(u,\varepsilon)\big\|_s \leqslant C(s,K)\big(1 + \|u\|_{s+2}\big), \tag{8.23}$$

$$\big\|\partial_u F(u,\varepsilon)[h]\big\|_s \leqslant C(s,K)\big(\|h\|_{s+2} + \|u\|_{s+2}\|h\|_4\big), \tag{8.24}$$

$$\big\|\partial_\varepsilon F(u,\varepsilon)[h]\big\|_s \leqslant \varepsilon^{-1} C(s,K)\big(1 + \|u\|_{s+2}\big), \tag{8.25}$$

for all $2 \leqslant s \leqslant r$, $\varepsilon \in (0, \varepsilon_0)$.

Proof. In Appendix C. □

Remark 8.6. Estimate (8.22) actually holds with an additional factor ε on the right-hand side. However, this makes no essential difference in our iteration proof below.

Lemma 8.7. *Assume the hypotheses of Lemma 8.4. Then*

$$\big\|\tilde\Psi\tilde\Phi(\Pi_N\tilde{\mathcal L}_4\Pi_N)^{-1}\Pi_N\tilde\Phi^{-1}\tilde{\mathcal M}^{-1}\tilde\Psi^{-1}P_\varepsilon f\big\|_s \leqslant C(s,K)\big(\|f\|_{s+5/2}+\|u\|_{s+17+5/2}\|f\|_2\big) \tag{8.26}$$

for $2\leqslant s\leqslant r-12-3/2$.

Proof. By (7.21) and (7.61), the term on the left-hand side in (8.26) is

$$\leqslant C(s,K)\big(\big\|(\Pi_N\tilde{\mathcal L}_4\Pi_N)^{-1}\Pi_N\tilde\Phi^{-1}\tilde{\mathcal M}^{-1}\tilde\Psi^{-1}P_\varepsilon f\big\|_s+\|u\|_{s+12}\big\|(\Pi_N\tilde{\mathcal L}_4\Pi_N)^{-1}\Pi_N\tilde\Phi^{-1}\tilde{\mathcal M}^{-1}\tilde\Psi^{-1}P_\varepsilon f\big\|_2\big)$$

for $2\leqslant s\leqslant r-7$. Write $\tilde\Phi^{-1}\tilde{\mathcal M}^{-1}\tilde\Psi^{-1}$ as $I+S$, where S satisfies (7.66). Since $\Pi_NP_\varepsilon=P_\varepsilon\Pi_N$,

$$(\Pi_N\tilde{\mathcal L}_4\Pi_N)^{-1}\Pi_N\tilde\Phi^{-1}\tilde{\mathcal M}^{-1}\tilde\Psi^{-1}P_\varepsilon f=(\Pi_N\tilde{\mathcal L}_4\Pi_N)^{-1}P_\varepsilon\Pi_Nf+(\Pi_N\tilde{\mathcal L}_4\Pi_N)^{-1}\Pi_NSP_\varepsilon f,$$

then use (8.18) for $(\Pi_N\tilde{\mathcal L}_4\Pi_N)^{-1}P_\varepsilon\Pi_Nf$, and use (8.17), (7.66) for $(\Pi_N\tilde{\mathcal L}_4\Pi_N)^{-1}\Pi_NSP_\varepsilon f$. □

9. Nash–Moser iteration and Cantor set of parameters

Let

$$\chi:=3/2,\qquad \bar a>0,\qquad N_n:=\exp\big(\bar a\chi^n\big),\quad n\in\mathbb N, \tag{9.1}$$

with $N_0=\exp(\bar a)$ sufficiently large to have $\mathcal K\subseteq[-N_0,N_0]$ ($\mathcal K$ is defined in Section 5). Consider the corresponding increasing sequence of finite-dimensional subspaces $Z_n:=Z_{N_n}$, with respective projections $\Pi_n:=\Pi_{N_n}$. For all $s,\alpha\geqslant 0$, Π_n enjoys the smoothing properties

$$\|\Pi_nu\|_{s+\alpha}\leqslant N_n^\alpha\|u\|_s\quad \forall u\in H^s, \tag{9.2}$$

$$\big\|\Pi_n^\perp u\big\|_s\leqslant N_n^{-\alpha}\|u\|_{s+\alpha}\quad \forall u\in H^{s+\alpha}, \tag{9.3}$$

where $\Pi_n^\perp=I-\Pi_n$. Note that (9.2), (9.3) hold even if $N_n>0$ is not an integer number.

In the previous sections we have proved the transformation

$$F'(u,\varepsilon)=P_\varepsilon^{-1}\mathcal L(u,\varepsilon)=P_\varepsilon^{-1}\tilde{\mathcal L}(u,\varepsilon)=P_\varepsilon^{-1}\tilde\Psi\tilde{\mathcal M}\tilde\Phi\tilde{\mathcal L}_4\tilde\Phi^{-1}\tilde\Psi^{-1} \tag{9.4}$$

where $\tilde\Psi,\tilde{\mathcal M},\tilde\Phi,\tilde{\mathcal L}_4$ all depend on (u,ε). Following a suitable Nash–Moser scheme, we construct a sequence $(u_n)\subset C^\infty(\mathbb T^2)$ of ε-dependent trigonometric polynomials by setting $u_0:=\bar v_2$ as defined in Section 5, $h_0:=0$, and

$$u_{n+1}:=u_n+h_{n+1},\quad h_{n+1}:=-\Pi_{n+1}\tilde\Psi_n\tilde\Phi_n(\Pi_{n+1}\tilde{\mathcal L}_{4,n}\Pi_{n+1})^{-1}\Pi_{n+1}\tilde\Phi_n^{-1}\tilde{\mathcal M}_n^{-1}\tilde\Psi_n^{-1}P_\varepsilon F(u_n), \tag{9.5}$$

provided that the inverse operator $\mathcal I_n:=(\Pi_{n+1}\tilde{\mathcal L}_4(u_n)\Pi_{n+1})^{-1}$ is well-defined on Z_{n+1}. The notation in (9.5) means

$$\tilde{\mathcal L}_{4,n}:=\tilde{\mathcal L}_4(u_n)=\tilde{\mathcal L}_4\big(u_n(\varepsilon),\varepsilon\big),\qquad \Psi_n:=\Psi(u_n)=\Psi\big(u_n(\varepsilon),\varepsilon\big),$$

and similarly for $\tilde{\mathcal M},\tilde\Phi$. Also, $\mathcal L_{4,n}=\mathcal D_n+\mathcal R_n$. We omit to write explicitly the dependence on ε only to shorten the notation. At a first glance, (9.5) could seem an unusual and excessively complicated Nash–Moser scheme. However, in some sense it is “the most natural” for the present problem, as the “normal form” for the linearized operator is given by $\mathcal L_{4,n}=\mathcal D_n+\mathcal R_n$, therefore it is natural to impose Diophantine conditions on the eigenvalues of $\mathcal D_n$ and to insert smoothing operators Π_n before and after it.

With h_{n+1} defined by (9.5), one has $h_{n+1}=-\Pi_{n+1}\tilde\Psi_n\tilde\Phi_n\mathcal I_n\Pi_{n+1}c_n$,

$$F(u_n)+F'(u_n)h_{n+1}=r_n:=P_\varepsilon^{-1}\tilde\Psi_n\tilde{\mathcal M}_n\tilde\Phi_n\big\{\Pi_{n+1}^\perp c_n-\Pi_{n+1}^\perp\tilde{\mathcal R}_n\Pi_{n+1}\mathcal I_n\Pi_{n+1}c_n+\tilde{\mathcal L}_{4,n}b_n\big\} \tag{9.6}$$

where

$$c_n:=\tilde\Phi_n^{-1}\tilde{\mathcal M}_n^{-1}\tilde\Psi_n^{-1}P_\varepsilon F(u_n),\qquad b_n:=\tilde\Phi_n^{-1}\tilde\Psi_n^{-1}\Pi_{n+1}^\perp\tilde\Psi_n\tilde\Phi_n\mathcal I_n\Pi_{n+1}c_n.$$

(9.6) follows directly from (9.5), and is proved in Appendix C. Hence

$$F(u_{n+1}) = r_n + Q(u_n, h_{n+1}), \tag{9.7}$$

where Q is defined in (8.21).

By Lemma 8.3, $\Pi_{n+1}\tilde{\mathcal{L}}_4(u_n)\Pi_{n+1}$ is invertible if the eigenvalues $\lambda_{l,j}(u_n,\varepsilon)$ of $\mathcal{D}_n$ satisfy the Diophantine condition (8.15) for $u = u_n$ and $N = N_{n+1}$. Let $\mathcal{W}_n := \mathcal{W}_{N_n}$. Define recursively the set of the "good" parameters ε, those for which (8.15) holds: let $\mathcal{G}_0 := (0, \varepsilon_0)$, and define

$$\mathcal{G}_{n+1} := \left\{\varepsilon \in \mathcal{G}_n\colon \left|\lambda_{l,j}(u_n,\varepsilon)\right| > \frac{1}{2\langle j\rangle^3}\ \forall (l,j) \in \mathcal{W}_{n+1}\right\}, \quad n \geqslant 0. \tag{9.8}$$

$\mathcal{G}_n$ is the set of the parameters ε for which $(u_k, h_k, A_k, \mathcal{G}_k)$ can be defined recursively for $k = 0, \dots, n$. On the contrary, after constructing $(u_k, h_k, A_k, \mathcal{G}_k)$ for $k \leqslant n$,

$$\mathcal{B}_{n+1} := \mathcal{G}_n \setminus \mathcal{G}_{n+1}$$

is the set of the "bad" parameters ε for which the Diophantine condition (8.15) on the eigenvalues $\lambda_{l,j}(u_n,\varepsilon)$ is violated on $|l| + |j| \leqslant N_{n+1}$, the inverse of $(\Pi_{n+1}\mathcal{L}_4(u_n)\Pi_{n+1})$ is not well-defined, h_{n+1} cannot be defined by (9.5), and the recursive construction stops. Therefore at the n-th step we eliminate the bad set $\mathcal{B}_{n+1}$, and restrict the parameter set to the subset $\mathcal{G}_{n+1} \subseteq \mathcal{G}_n$. For convenience, put $\mathcal{B}_0 := \emptyset$.

Proposition 9.1 *(Nash–Moser induction and measure estimate for the parameter set). There exist universal constants $r_0, s_0 > 0$ and constants $C, C', c_0, \bar{a}, \bar{b}, \varepsilon_0^* > 0$ depending only on $\bar{v}_1, K_{g,r_0}$ such that if $\mathcal{G}_0 = (0, \varepsilon_0)$, $\varepsilon_0 \leqslant \varepsilon_0^*$, $r \geqslant r_0$, and $\bar{a}$ defines N_n in* (9.1)*, then the following induction hold.*

Let $(P_n) = \{(P_n)(\mathrm{i}), (P_n)(\mathrm{ii})\}$, $n \geqslant 1$, be the following set of statements.

- (P_n)(i). *$\mathcal{G}_n$ is an open set. The Lebesgue measure of $\mathcal{B}_n$ satisfies $|\mathcal{B}_n| \leqslant \varepsilon_0^2 C b_n$, where the sequence (b_n) satisfies $\sum_{n=0}^{\infty} b_n = C' < \infty$.*
- (P_n)(ii). *For every $\varepsilon \in \mathcal{G}_n$, $h_n(\varepsilon) \in Z_n$ is well-defined. $h_n : \mathcal{G}_n \to Z_n$, $\varepsilon \mapsto h_n(\varepsilon)$ is of class C^1 as a function of ε, with*

$$\left\|h_n(\varepsilon)\right\|_{s_0} < \exp\left(-\bar{b}\chi^n\right), \qquad \left\|\partial_\varepsilon h_n(\varepsilon)\right\|_{s_0} \leqslant \varepsilon^{-1}\exp\left(-\bar{b}\chi^n\right). \tag{9.9}$$

(P_1) holds. If (P_n) holds, then, using (9.5), (9.8) *to define h_{n+1} and $\mathcal{G}_{n+1}$, (P_{n+1}) also holds.*

As a consequence, the Cantor set $\mathcal{G}_\infty := \bigcap_{n\geqslant 0}\mathcal{G}_n \subset (0,\varepsilon_0)$ has Lebesgue measure

$$|\mathcal{G}_\infty| \geqslant \varepsilon_0(1 - \varepsilon_0 C).$$

For every $\varepsilon \in \mathcal{G}_\infty$, the sequence $(u_n(\varepsilon))$ converges in $H^{s_0}(\mathbb{T}^2)$ to a limit $u_\infty(\varepsilon)$, which solves

$$F\left(u_\infty(\varepsilon), \varepsilon\right) = 0.$$

Moreover, $u_\infty(\varepsilon) \in H^s(\mathbb{T}^2)$ for every s in the interval $s_0 \leqslant s < (r + c_0)/2$.

If g_i, $i = 0, 1, 2$ in (1.2), (1.3) *is of class C^∞, then also $u_\infty(\varepsilon) \in C^\infty(\mathbb{T}^2)$.*

s_0, r_0 and c_0 can be explicitly calculated: $s_0 = 22$, $c_0 = 28$; for r_0 see (9.22) *and below.*

We split the proof of Proposition 9.1 into two parts: the Nash–Moser sequence (P_n)(ii) with its regularity in Section 9.1, then the measure estimate (P_n)(i) for the parameter set in Section 9.2

9.1. *Proof of the Nash–Moser iteration*

First step. Let us prove (P_1)(ii). For $\varepsilon \in \mathcal{G}_1$, (9.5) defines $h_1 = h_1(\varepsilon)$. By (8.19), the condition (6.4) holds. By (8.19), if $22 \leqslant r$, then $\|\bar{v}_2(\varepsilon)\|_{22} \leqslant C$ for all $\varepsilon \in (0, \varepsilon_0)$, for some constant C. Take this constant C as the "K" in all the lemmata of the previous sections, so that the assumption $K \geqslant \|u\|_{22}$ is satisfied for $u = u_0 = \bar{v}_2(\varepsilon)$, for all $\varepsilon \in (0, \varepsilon_0)$. In this way, to indicate the dependence on K in all the constants $C(s, K)$ is redundant, and we simply write $C(s, K) = C(s)$. By (9.5), (8.26), (8.19) and (8.20),

$$\|h_1\|_s = \|\tilde{\Psi}_0\tilde{\Phi}_0\mathcal{I}_0\Pi_1 c_0\|_s \leqslant C(s)\left(\left\|F(u_0)\right\|_{s+5/2} + \|u_0\|_{s+17+5/2}\left\|F(u_0)\right\|_2\right) \leqslant \varepsilon C(s)$$

if $s+17+5/2 \leqslant r$. Hence the first inequality in (P_1)(ii) holds if

$$\varepsilon_0 C(s) \leqslant \exp(-\bar{b}\chi). \tag{9.10}$$

$\partial_\varepsilon h_1$ is obtained by differentiating every term in formula (9.5) with respect to ε and applying the estimates for $\partial_\varepsilon \tilde{\Psi}$, $\partial_\varepsilon \tilde{\Phi}$, $\partial_\varepsilon\{(\Pi_1 \tilde{\mathcal{L}}_4(u_0(\varepsilon),\varepsilon)\Pi_1)^{-1}\}$, etc.; using (8.19) for $\partial_\varepsilon \bar{v}_2$, and (8.20) for $\partial_\varepsilon\{F(\bar{v}_2(\varepsilon),\varepsilon)\}$, we get

$$\|\partial_\varepsilon h_1(\varepsilon)\|_s \leqslant C(s)$$

for $\varepsilon \in (0,\varepsilon_0)$, $s+17+5/2 \leqslant r$. Therefore the second inequality in (P_1)(ii) holds if (9.10) holds (with a possibly different constant $C(s)$, as usual).

Inductive step. Now assume that (P_n) holds, $n \geqslant 1$, and prove (P_{n+1})(ii). By (9.9),

$$\|u_n\|_s \leqslant \|u_0\|_s + \sum_{k=1}^{n} \|h_k\|_s \leqslant \|\bar{v}_2\|_s + C(\bar{b}), \quad C(\bar{b}) := \sum_{k=1}^{\infty} \exp(-\bar{b}\chi^k). \tag{9.11}$$

Note that $C(\bar{b})$ is independent on n, it is decreasing as a function of $\bar{b}$, and $C(\bar{b}) \to 0$ as $\bar{b} \to +\infty$. Hence, for $s \geqslant 22$, $\|u_n\|_{22} \leqslant \|\bar{v}_2\|_{22} + C(\bar{b}) \leqslant 2\|\bar{v}_2\|_{22} = C$ for all $\varepsilon \in (0,\varepsilon_0)$ if

$$\bar{b} \geqslant C, \tag{9.12}$$

for some $C>0$. As in the previous step, take this constant C as the "K", and replace $C(s,K)$ with $C(s)$ in all the lemmata of the previous sections. Moreover, (6.4) is satisfied for $u=u_n$ if ε_0 is sufficiently small, independently on the parameters. Also, $\|u_n\|_s \leqslant C(s)$.

By (9.5), (9.2) and (8.26), for $\alpha \geqslant 0$, $2 \leqslant s-\alpha \leqslant r-12-3/2$,

$$\begin{aligned} \|h_{n+1}\|_s &\leqslant N_{n+1}^{\alpha} \|\tilde{\Psi}_n \tilde{\Phi}_n \mathcal{I}_n \Pi_{n+1} c_n\|_{s-\alpha} \\ &\leqslant N_{n+1}^{\alpha} C(s-\alpha)\big(\|F(u_n)\|_{s-\alpha+5/2} + \|u_n\|_{s-\alpha+17+5/2}\|F(u_n)\|_2\big). \end{aligned} \tag{9.13}$$

Take $\alpha := 17+5/2$, and denote $s' := s-17$. Since $s' \geqslant 2$,

$$\|h_{n+1}\|_s \leqslant (9.13) \leqslant N_{n+1}^{\alpha} C(s)\big(\|F(u_n)\|_{s'} + \|u_n\|_s \|F(u_n)\|_2\big) \leqslant N_{n+1}^{\alpha} C(s)\|F(u_n)\|_{s'}$$

because $\|u_n\|_s \leqslant C(s)$ by (9.11). By (9.7), $F(u_n) = r_{n-1} + Q(u_{n-1},h_n)$. Therefore

$$\|h_{n+1}\|_s \leqslant A_r + A_Q, \quad A_r := N_{n+1}^{\alpha} C(s)\|r_{n-1}\|_{s'}, \ A_Q := N_{n+1}^{\alpha} C(s)\|Q(u_{n-1},h_n)\|_{s'}. \tag{9.14}$$

By (9.6), r_{n-1} is the sum of 3 terms, say (I) + (II) + (III). The first one is

$$(\mathrm{I}) = P_\varepsilon^{-1} \tilde{\Psi}_{n-1} \tilde{\mathcal{M}}_{n-1} \tilde{\Phi}_{n-1} \Pi_n^{\perp} \tilde{\Phi}_{n-1}^{-1} \tilde{\mathcal{M}}_{n-1}^{-1} \tilde{\Psi}_{n-1}^{-1} P_\varepsilon F(u_{n-1}).$$

Using (7.66), like in the proof of Lemma 8.7, no negative power of ε appears in the estimate of (I). Using (9.3) to deal with $\Pi_n^{\perp}$, for $\beta \geqslant 0$, $2 \leqslant s'+\beta \leqslant r-8$, one has

$$\|(\mathrm{I})\|_{s'} \leqslant C(s+\beta) N_n^{-\beta}\big(\|F(u_{n-1})\|_{s'+\beta+2} + \|u_{n-1}\|_{s'+\beta+13}\|F(u_{n-1})\|_2\big).$$

The same argument applies to (II) and (III), whence

$$\|r_{n-1}\|_{s'} \leqslant C(s'+\beta) N_n^{-\beta}\big(\|F(u_{n-1})\|_{s'+\beta+8} + \|u_{n-1}\|_{s'+\beta+19}\|F(u_{n-1})\|_2\big),$$

$2 \leqslant s'+\beta \leqslant r-16$. Applying (8.23),

$$\|r_{n-1}\|_{s'} \leqslant C(s'+\beta) N_n^{-\beta}\big(1 + \|u_{n-1}\|_{s'+\beta+19}\big) = C(s+\beta) N_n^{-\beta}\big(1 + \|u_{n-1}\|_{s+\beta+2}\big). \tag{9.15}$$

Now estimate the "high norm" $B_k := \|h_k\|_{s+\beta+2}$. To each $k = 0,\dots,n$, apply (9.13) with $s+\beta+2$ instead of s, and use (8.23): for $2 \leqslant (s+\beta+2)-\alpha \leqslant r-12-3/2$,

$$\begin{aligned} \|h_{k+1}\|_{s+\beta+2} &\leqslant N_{k+1}^{\alpha} C(s+\beta+2-\alpha)\big(\|F(u_k)\|_{s+\beta+2-\alpha+5/2} + \|u_k\|_{s+\beta+2-\alpha+17+5/2}\|F(u_k)\|_2\big) \\ &\leqslant N_{k+1}^{\alpha} C(s+\beta)\big(1 + \|u_k\|_{s+\beta+2}\big) \end{aligned} \tag{9.16}$$

where, as above, $\alpha := 17 + 5/2$. For (8.19), $\|u_0\|_{s+\beta+2} \leqslant C(s+\beta)$ if $s+\beta+2 \leqslant r$. Then, by (9.16), $B_1 = \|h_1\|_{s+\beta+2} \leqslant N_1^\alpha C(s+\beta)$, and

$$B_{k+1} \leqslant N_{k+1}^\alpha C(s+\beta)\left(1 + \|u_0\|_{s+\beta+2} + \sum_{j=1}^k \|h_j\|_{s+\beta+2}\right) \leqslant N_{k+1}^\alpha C(s+\beta)\left(1 + \sum_{j=1}^k B_j\right) \tag{9.17}$$

for $1 \leqslant k \leqslant n$. By (9.1), this implies that

$$\|h_k\|_{s+\beta+2} = B_k \leqslant \exp\bigl(\bar b \chi^k\bigr), \tag{9.18}$$

$k = 1, \dots, n+1$. For, by induction: (9.18) holds for $k=1$ if $C(s+\beta)\exp[(\bar a\alpha - \bar b)\chi] \leqslant 1$, namely if $(\bar b - \bar a\alpha)$ is larger than some constant depending on $(s+\beta)$. Suppose that (9.18) holds for all $j \in [1,k]$, $k \geqslant 1$. For $\bar b \geqslant 1$,

$$1 + \sum_{j=1}^k \exp\bigl(\bar b\chi^j\bigr) \leqslant C\exp\bigl(\bar b\chi^k\bigr), \quad \forall k \in \mathbb{N},$$

for some universal constant C. Then, by (9.17), (9.18) also holds for $k+1$ if $C(s+\beta)\exp[\chi^k(\bar a\alpha\chi - \bar b\chi + \bar b)] \leqslant 1$, namely if

$$\bar b - 3\bar a\alpha \geqslant C(s+\beta) \tag{9.19}$$

for some $C(s+\beta) > 0$, and (9.18) is proved. Thus $\|u_{n-1}\|_{s+\beta+2} \leqslant C(s+\beta)\exp(\bar b\chi^{n-1})$, and, by (9.15),

$$\|r_{n-1}\|_{s'} \leqslant C(s+\beta)\exp\bigl[\chi^{n-1}(\bar b - \beta\bar a\chi)\bigr], \qquad A_r \leqslant C(s+\beta)\exp\bigl[\chi^{n-1}\bigl(\bar b + \alpha\bar a\chi^2 - \beta\bar a\chi\bigr)\bigr].$$

As a consequence, $A_r \leqslant \frac{1}{2}\exp(-\bar b\chi^{n+1})$ if

$$\bar a\bigl(\beta\chi - \alpha\chi^2\bigr) - \bar b\bigl(1+\chi^2\bigr) \geqslant C(s+\beta) \tag{9.20}$$

for some $C(s+\beta) > 0$.

Estimate A_Q. Since $\|u_{n-1}\|_{s'+2} = \|u_{n-1}\|_{s-15} \leqslant C(s)$, by (8.22) we have $A_Q \leqslant N_{n+1}^\alpha C(s)\|h_n\|_s^2$. This is $\leqslant \frac{1}{2}\exp(-\bar b\chi^{n+1})$ if

$$\bar b - 3\alpha\bar a \geqslant C(s) \tag{9.21}$$

for some $C(s) > 0$. Now fix

$$\bar b := (3\alpha+1)\bar a, \qquad \beta := \bigl[\alpha\chi^2 + \bigl(1+\chi^2\bigr)(3\alpha+1)\bigr]\chi^{-1}. \tag{9.22}$$

Since $\chi = 3/2$ and $\alpha = 17 + 5/2$, β is a universal constant, and the constants $C(s+\beta)$ can be written as $C(s)$. Fix $\bar a \geqslant C(s)$ sufficiently large to satisfy (9.19), (9.20), (9.21) and (9.12). Then fix $\varepsilon_0 \leqslant C(s)$ sufficiently small to satisfy (9.10). All the above conditions on s hold if

$$22 \leqslant s \leqslant r - 2 - \beta.$$

Hence the minimal value for r is $r_0 := 24 + \beta$. Put $s_0 := 22$. For $s = s_0 = 22$ and $r = r_0$, all the above constants that depend on s and $K_{g,r}$ become constants depending only on K_{g,r_0}. With this choice of parameters, the first estimate of (P_{n+1})(ii) is proved.

The second estimate of (P_{n+1})(ii) can be proved by the same arguments. Observe that in every estimate for ∂_ε there is an additional factor $1/\varepsilon$: indeed, terms like ε^p or P_ε, after being differentiated, have one degree less as powers of ε. Terms like $F(u_n,\varepsilon)$, $\tilde\Psi(u_n,\varepsilon)$, …, after being differentiated with respect to ε, contain also terms like $\partial_u F(u_n,\varepsilon)[\partial_\varepsilon u_n]$, $\partial_u\tilde\Psi(u_n,\varepsilon)[\partial_\varepsilon u_n]$, …, and the loss of one degree as a power of ε comes from (9.9). The estimates for ∂_u and ∂_ε of all the terms are given in the previous sections (and remind formula (4.5) for $F(u,\varepsilon)$).

For each ε for which the sequence $(u_n(\varepsilon))$ can be constructed, by (9.9) $u_n = u_0 + \sum_{k=1}^n h_k$ is a Cauchy sequence in $H^{s_0}(\mathbb{T}^2)$, therefore $u_n(\varepsilon)$ converges in H^{s_0} to some limit $u_\infty(\varepsilon) \in H^{s_0}$ as $n \to \infty$. Since the map $H^{s_0} \to H^{s_0-2}$, $u \mapsto F(u,\varepsilon)$ is continuous, $\|F(u_n,\varepsilon) - F(u_\infty,\varepsilon)\|_{s_0-2} \to 0$. On the other hand, we have proved that

$$\bigl\|F(u_n,\varepsilon)\bigr\|_{s'} \leqslant \|r_{n-1}\|_{s'} + \bigl\|Q(u_{n-1},h_n)\bigr\|_{s'} = C(s_0)N_{n+1}^{-\alpha}(A_r + A_Q) \leqslant C(s_0)N_{n+1}^{-\alpha}\exp\bigl(-\bar b\chi^{n+1}\bigr) \to 0$$

as $n \to \infty$, where $s' = s_0 - 17 = 5$. Thus $F(u_\infty,\varepsilon) = 0$.

Now let $22 = s_0 < s_1 < s_2$, with $s_1 = \lambda s_0 + (1-\lambda)s_2$, and $\lambda \in (1/2, 1)$. Apply (9.16) with s_2 instead of $s+\beta+2$: for $s_2 - \alpha \leqslant r - 12 - 3/2$ we get

$$\|h_{k+1}\|_{s_2} \leqslant N_{k+1}^{\alpha} C(s_2)\big(1 + \|u_k\|_{s_2}\big) \quad \forall k \geqslant 0,$$

for some constant $C(s_2)$ depending on s_2. For (8.19), $\|u_0\|_{s_2} \leqslant C(s_2)$ if $s_2 \leqslant r$. Then the "very high norms" $B'_k := \|h_k\|_{s_2}$ satisfy $B'_1 = \|h_1\|_{s_2} \leqslant N_1^{\alpha} C(s_2)$, and

$$B'_{k+1} \leqslant N_{k+1}^{\alpha} C(s_2)\left(1 + \sum_{j=1}^{k} B'_j\right), \quad k \geqslant 1.$$

Therefore there is a constant $K(s_2)$ such that

$$\|h_k\|_{s_2} = B'_k \leqslant K(s_2)\exp\big(\bar b \chi^k\big), \quad k \geqslant 1. \tag{9.23}$$

Let us prove (9.23). Since $\bar b - 3\alpha\bar a > 0$, where $\bar a, \bar b$ have been fixed above, the inductive step ($k \Rightarrow k+1$) holds for all $k \geqslant k_0(s_2)$, for some $k_0(s_2)$ depending on s_2 which is sufficiently large. Note that the constant $K(s_2)$ have no role in the inductive step. Then choose $K(s_2) := \max\{\|h_k\|_{s_2}\exp(-\bar b\chi^k)\colon 1 \leqslant k \leqslant k_0(s_2)\}$, so that (9.23) holds for all $k \geqslant 1$. Now, by (B.1), (9.23) and (9.9),

$$\|h_k\|_{s_1} \leqslant 2\|h_k\|_{s_0}^{\lambda}\|h_k\|_{s_2}^{1-\lambda} \leqslant 2K(s_2)^{1-\lambda}\exp\big(-\lambda\bar b\chi^k\big)\exp\big((1-\lambda)\bar b\chi^k\big) = C(s_2,\lambda)\exp\big((1-2\lambda)\bar b\chi^k\big),$$

and the series $\sum_{k\geqslant 1}\exp((1-2\lambda)\bar b\chi^k)$ converges because $(1-2\lambda) < 0$. This implies that $\|u_\infty\|_{s_1} \leqslant \|u_0\|_{s_1} + \sum_{k\geqslant 1}\|h_k\|_{s_1} < \infty$. Since $s_1 < (s_0+s_2)/2$ and $s_2 < r - 12 - 3/2 + \alpha$, $\alpha = 17 + 5/2$, this argument holds if

$$s_1 < \frac{r+28}{2}.$$

If g_i, $i = 0,1,2$ that defines the nonlinearity $\mathcal N$ is of class C^∞, then there is no upper bound for s_1, and the argument applies for every $s_1 \geqslant s_0$, whence $u_\infty \in C^\infty$.

9.2. Proof of the measure estimate

$\mathcal G_0 = (0, \varepsilon_0)$, $\mathcal B_0 = \emptyset$. Let us estimate $\mathcal G_{n+1}, \mathcal B_{n+1}$, $n \geqslant 0$.

The set $\mathcal G_{n+1}$ is defined by (9.8). $u_n(\varepsilon)$ is a C^1 function of ε, and $\mu_k(u,\varepsilon)$, $k = 2,1,0,-2$ is a C^1 function of (u,ε). Therefore each eigenvalue $\lambda_{l,j}(u_n(\varepsilon),\varepsilon)$ is C^1 in ε. $\mathcal B_{n+1}$ is the union

$$\mathcal B_{n+1} = \bigcup_{(l,j)\in\mathcal W_{n+1}} \Omega^n_{l,j}, \quad \Omega^n_{l,j} := \left\{\varepsilon \in \mathcal G_n\colon \big|\lambda_{l,j}(u_n,\varepsilon)\big| \leqslant \frac{1}{2\langle j\rangle^3}\right\}. \tag{9.24}$$

Write the eigenvalues $\lambda_{l,j}(u_n(\varepsilon),\varepsilon)$ as

$$\lambda_{l,j}\big(u_n(\varepsilon),\varepsilon\big) = i\omega\big(l + p^n_j(\varepsilon)\big),$$
$$p^n_j(\varepsilon) := \frac{\mu_2(u_n(\varepsilon),\varepsilon)}{1+3\varepsilon^2} j|j| + \frac{\mu_1(u_n(\varepsilon),\varepsilon)}{1+3\varepsilon^2} j + \frac{-\mu_0(u_n(\varepsilon),\varepsilon)}{1+3\varepsilon^2}\operatorname{sign}(j) + \frac{\mu_{-2}(u_n(\varepsilon),\varepsilon)}{1+3\varepsilon^2}\frac{\operatorname{sign}(j)}{j^2}$$

(where we mean $\operatorname{sign}(j)j^{-2} = 0$ for $j = 0$). Since $\omega = 1+3\varepsilon^2 > 1$, $|\lambda_{l,j}(u_n(\varepsilon),\varepsilon)| \geqslant |l + p^n_j(\varepsilon)|$, and

$$\Omega^n_{l,j} \subseteq \tilde\Omega^n_{l,j} := \left\{\varepsilon \in \mathcal G_n\colon \big|l + p^n_j(\varepsilon)\big| \leqslant \frac{1}{2\langle j\rangle^3}\right\} \quad \forall (l,j) \in \mathcal W_{n+1}. \tag{9.25}$$

For $j = 0$, $p^n_j(\varepsilon) = p^n_0(\varepsilon) = 0$, therefore $\tilde\Omega^n_{l,0} = \emptyset$ for all $l \neq 0$. The pair $(l,j) = (0,0)$ does not belong to $\mathcal W_{n+1}$, hence the case $j = 0$ gives no contribution to the union (9.24). So let $j \neq 0$.

$$\frac{\mu_2(u_n(\varepsilon),\varepsilon)}{1+3\varepsilon^2} = 1 - 3\varepsilon^2 + O\big(\varepsilon^3\big), \qquad \frac{\mu_1(u_n(\varepsilon),\varepsilon)}{1+3\varepsilon^2} = 3b\varepsilon^2 + O\big(\varepsilon^3\big),$$
$$\frac{\mu_k(u_n(\varepsilon),\varepsilon)}{1+3\varepsilon^2} = O\big(\varepsilon^3\big), \quad k = 0,-2,$$

where $b := \Pi_C(\bar{v}_1^2)$, and the precise meaning of $O(\varepsilon^3)$ is given by (7.16), (7.17), (7.59), (7.60). Therefore

$$p_j^n(\varepsilon) = j|j|\big(1+\varepsilon^2 r_j^n(\varepsilon)\big), \quad r_j^n(\varepsilon) := \frac{1}{\varepsilon^2}\bigg(\frac{p_j^n(\varepsilon)}{j|j|} - 1\bigg) = -3 + \frac{3b}{|j|} + O(\varepsilon).$$

$|r_j^n(\varepsilon)| \leqslant C$ for some $C > 0$ independent of j, n, ε. Also, by Proposition 5.3,

$$\big|b - |j|\big| \geqslant \delta|j|, \qquad \left|-3 + \frac{3b}{|j|}\right| \geqslant 3\delta \quad \forall j \in \mathbb{N},\ j \neq 0.$$

As a consequence,

$$2\delta \leqslant \big|r_j^n(\varepsilon)\big| \leqslant C$$

for $\varepsilon < \varepsilon_0$ sufficiently small to have $|r_j^n(\varepsilon) + 3 - 3b/|j|| \leqslant \delta$. Suppose that $\varepsilon \in \tilde{\Omega}_{l,j}^n \neq \emptyset$. Then, by the triangular inequality,

$$\big|l + j|j|\big| \leqslant \big|l + p_j^n(\varepsilon)\big| + \big|-p_j^n(\varepsilon) + j|j|\big| \leqslant \frac{1}{2\langle j\rangle^3} + \varepsilon^2|j|^2\big|r_j^n(\varepsilon)\big| \leqslant \frac{1}{2} + C\varepsilon^2|j|^2. \tag{9.26}$$

$|l + j|j|| \geqslant 1$ because $l + j|j|$ is a nonzero integer. Thus we have a “cut-off”: if $\tilde{\Omega}_{l,j}^n \neq \emptyset$, then $1 \leqslant 1/2 + C\varepsilon^2|j|^2$, and

$$C \leqslant \varepsilon|j| \leqslant \varepsilon_0|j|, \tag{9.27}$$

for some $C > 0$. Moreover, by (9.26), l belongs to the interval

$$-j|j| - 1/2 - C\varepsilon_0^2|j|^2 \leqslant l \leqslant -j|j| + 1/2 + C\varepsilon_0^2|j|^2. \tag{9.28}$$

As a consequence, for any fixed j with $|j| \geqslant C/\varepsilon_0$, the number of integers l such that $\tilde{\Omega}_{l,j}^n \neq \emptyset$ does not exceed the number of integers l in the interval (9.28), namely

$$\sharp\big\{l\colon \tilde{\Omega}_{l,j}^n \neq \emptyset\big\} \leqslant 2\big(1/2 + C\varepsilon_0^2|j|^2\big) + 1 \leqslant C'\varepsilon_0^2|j|^2 \tag{9.29}$$

because $2 \leqslant C\varepsilon_0^2|j|^2$ by (9.27) (and the number of integers in an interval $[a,b]$ is at most $(b-a+1)$). By (9.25), (9.29) implies that $\mathcal{B}_{n+1}$ is the union of a finite number of closed sets, hence $\mathcal{G}_{n+1}$ is open.

From the chain rule, (7.16), (7.17), (7.59), (7.60), and $\|\partial_\varepsilon u_n(\varepsilon)\|_{12} \leqslant \varepsilon^{-1}C$ (which follows from (9.9)),

$$\partial_\varepsilon p_j^n(\varepsilon) = j|j|\varepsilon\bigg(-6 + \frac{6b}{|j|} + O(\varepsilon)\bigg).$$

Hence, for any fixed j, the sign of $\partial_\varepsilon p_j^n(\varepsilon)$ is the sign of $j(-1 + b/|j|)$, which is constant with respect to ε. By (9.27),

$$\big|\partial_\varepsilon p_j^n(\varepsilon)\big| = |j|^2\varepsilon\left|-6 + \frac{6b}{|j|} + O(\varepsilon)\right| \geqslant |j|^2\varepsilon\delta \geqslant C|j|$$

if ε_0 is sufficiently small. So p_j^n is strictly monotone as a function of ε, and, as a consequence, $\tilde{\Omega}_{l,j}^n$ is an interval, say $[\varepsilon_1, \varepsilon_2]$. If p_j^n is increasing, then

$$\frac{1}{|j|^3} \geqslant p_j^n(\varepsilon_2) - p_j^n(\varepsilon_1) = \int_{\varepsilon_1}^{\varepsilon_2} \partial_\varepsilon p_j^n(\varepsilon)\,d\varepsilon \geqslant C|j|(\varepsilon_2 - \varepsilon_1) = C|j|\big|\tilde{\Omega}_{l,j}^n\big|,$$

and analogous calculation if p_j^n is decreasing. Thus

$$\big|\tilde{\Omega}_{l,j}^n\big| \leqslant \frac{C}{|j|^4}. \tag{9.30}$$

Also, $|\Omega_{l,j}^n| \leqslant |\tilde{\Omega}_{l,j}^n|$ because $\Omega_{l,j}^n \subseteq \tilde{\Omega}_{l,j}^n$.

Now split the union (9.24) into two parts, the union over the "old" indices $(l, j) \in \mathcal{W}_{n+1} \cap \mathcal{W}_n = \mathcal{W}_n$ and the one over the "new" indices $(l, j) \in \mathcal{W}_{n+1} \setminus \mathcal{W}_n$. By (9.29) and (9.30), the Lebesgue measure of the union over the new indices is

$$\Big|\bigcup_{\text{new}} \Omega^n_{l,j}\Big| \leqslant \sum_{\text{new}} |\Omega^n_{l,j}| \leqslant \sum_{N_n<|j|\leqslant N_{n+1}} \frac{C}{|j|^4}\varepsilon_0^2|j|^2 = C\varepsilon_0^2 \sum_{N_n<|j|\leqslant N_{n+1}} \frac{1}{|j|^2} = C\varepsilon_0^2 c_{n+1},$$

where

$$c_0 := \sum_{1\leqslant|j|\leqslant N_0} \frac{1}{|j|^2}, \qquad c_{n+1} := \sum_{N_n<|j|\leqslant N_{n+1}} \frac{1}{|j|^2}, \qquad \text{and} \quad \sum_{n=0}^{\infty} c_n = \sum_{|j|=1}^{\infty} \frac{1}{|j|^2} = C < \infty.$$

For old indices, let $\varepsilon \in \tilde{\Omega}^n_{l,j}$, with $(l, j) \in \mathcal{W}_n$. By the triangular inequality, $u_n = u_{n-1} + h_n$, and estimates (7.16), (7.17), (7.59), (7.60) for $\partial_u \mu_k(u, \varepsilon)$,

$$\big|l + p_j^{n-1}(\varepsilon)\big| \leqslant \big|l + p_j^n(\varepsilon)\big| + \big|p_j^n(\varepsilon) - p_j^{n-1}(\varepsilon)\big| \leqslant \frac{1}{2|j|^3} + C\varepsilon^4|j|^2\big\|h_n(\varepsilon)\big\|_{12}.$$

Since $\tilde{\Omega}^n_{l,j} \subseteq \mathcal{G}_n$, and $(l, j) \in \mathcal{W}_n$,

$$\tilde{\Omega}^n_{l,j} \subseteq \left\{\varepsilon \in \mathcal{G}_n\colon \frac{1}{2|j|^3} < \big|l + p_j^{n-1}(\varepsilon)\big| \leqslant \frac{1}{2|j|^3} + C\varepsilon^4|j|^2\big\|h_n(\varepsilon)\big\|_{12}\right\}.$$

As above, p_j^{n-1} is strictly monotone as a function of ε, $|\partial_\varepsilon p_j^{n-1}(\varepsilon)| \geqslant C|j|$, and $\|h_n(\varepsilon)\|_{12} \leqslant \exp(-\bar{b}\chi^n)$ by (9.9). Hence

$$\big|\tilde{\Omega}^n_{l,j}\big| \leqslant C\varepsilon_0^4|j|^2 \exp\big(-\bar{b}\chi^n\big)\frac{1}{|j|} \leqslant C\varepsilon_0^4 N_n \exp\big(-\bar{b}\chi^n\big)$$

because $|j| \leqslant N_n$. By (9.29) and (9.1), the Lebesgue measure of the union over the old indices is then

$$\Big|\bigcup_{\text{old}} \Omega^n_{l,j}\Big| \leqslant \sum_{\text{old}} |\Omega^n_{l,j}| \leqslant C\varepsilon_0^4 \sum_{|j|\leqslant N_n} N_n^3 \exp\big(-\bar{b}\chi^n\big) \leqslant C\varepsilon_0^4 N_n^4 \exp\big(-\bar{b}\chi^n\big) = C\varepsilon_0^4 \exp\big[\chi^n(-\bar{b} + 4\bar{a})\big].$$

Since $\bar{b} - 4\bar{a} > \bar{a} \geqslant 1$ by (9.22), $\sum_{n=0}^{\infty} \exp[\chi^n(-\bar{b} + 4\bar{a})] = C < \infty$. We have proved that

$$|\mathcal{B}_{n+1}| \leqslant C\varepsilon_0^2 b_{n+1}, \qquad \sum_{n=0}^{\infty} b_n = C < \infty.$$

Therefore $|\bigcup_{n\geqslant 1} \mathcal{B}_n| \leqslant \varepsilon_0^2 C$, whence $|\mathcal{G}_\infty| \geqslant \varepsilon_0(1 - \varepsilon_0 C)$.

Acknowledgements

I express my gratitude to Massimiliano Berti for many fruitful discussions and suggestions, Pavel Plotnikov, Gérard Iooss and Thomas Alazard for useful conversations, and John Toland for introducing me to the problem.

This work is partially supported by the Italian PRIN2009 grant *Critical Point Theory and Perturbative Methods for Nonlinear Differential Equations*, and by the European Research Council, FP7, project *New connections between Dynamical Systems and Hamiltonian PDEs with Small Divisors Phenomena*.

Appendix A. Kernel properties

Proof of Lemma 5.1. 1) Let j_1, j_2 be nonzero. $q_{j_1}q_{j_2} = q_{j_3} \in V$ for some $j_3 \in \mathbb{Z}$ if and only if

$$j_1 + j_2 = j_3, \qquad -j_1|j_1| - j_2|j_2| = -j_3|j_3|.$$

Let $n_k := |j_k|$ and $j_k = \sigma_k n_k$, $\sigma_k \in \{1, -1\}$, $k = 1, 2$. If $\sigma_1 = \sigma_2$, then

$$j_3 = j_1 + j_2 = \sigma_1(n_1 + n_2), \qquad j_3|j_3| = j_1|j_1| + j_2|j_2| = \sigma_1\big(n_1^2 + n_2^2\big),$$

therefore $|j_3|^2 = (n_1 + n_2)^2 = (n_1^2 + n_2^2)$, and this is impossible because $n_1 n_2 > 0$. If $\sigma_1 = -\sigma_2$, then

$$j_3 = j_1 + j_2 = \sigma_1(n_1 - n_2), \qquad j_3|j_3| = j_1|j_1| + j_2|j_2| = \sigma_1\big(n_1^2 - n_2^2\big),$$

whence $|n_2 - n_1|(n_1 + n_2 - |n_2 - n_1|) = 0$. This holds only for $n_2 = n_1$.

2) Let j_1, j_2, j_3 all nonzero. $q_{j_1} q_{j_2} q_{j_3} = q_{j_4} \in V$ for some $j_4 \in \mathbb{Z}$ if and only if

$$j_1 + j_2 + j_3 = j_4, \qquad -j_1|j_1| - j_2|j_2| - j_3|j_3| = -j_4|j_4|.$$

Let $n_k := |j_k|$, $j_k = \sigma_k n_k$, $k = 1, 2, 3, 4$, with $\sigma_1, \sigma_2, \sigma_3 \in \{1, -1\}$ and $\sigma_4 \in \{1, 0, -1\}$. If $\sigma_1 = \sigma_2 = \sigma_3$, then

$$-n_1^2 - n_2^2 - n_3^2 + (n_1 + n_2 + n_3)^2 = 0,$$

which is impossible because $n_1, n_2, n_3 > 0$. If $\sigma_1, \sigma_2, \sigma_3$ are not all equal, say $\sigma_1 = \sigma_2 = -\sigma_3$, then

$$\sigma_4 n_4 = j_4 = j_1 + j_2 + j_3 = \sigma_1(n_1 + n_2 - n_3),$$
$$\sigma_4 n_4^2 = j_4|j_4| = j_1|j_1| + j_2|j_2| + j_3|j_3| = \sigma_1\big(n_1^2 + n_2^2 - n_3^2\big).$$

If $j_4 = 0$, then

$$n_1 + n_2 = n_3, \qquad n_1^2 + n_2^2 = n_3^2,$$

which is impossible because $n_1 n_2 > 0$. Thus $j_4 \neq 0$, $\sigma_4 \neq 0$. As a consequence,

$$n_1 + n_2 - n_3 = \sigma n_4, \qquad n_1^2 + n_2^2 - n_3^2 = \sigma n_4^2, \qquad \sigma := \sigma_1 \sigma_4 \in \{1, -1\}.$$

If $\sigma = -1$, then

$$n_1 + n_2 + n_4 = n_3, \qquad n_1^2 + n_2^2 + n_4^2 = n_3^2,$$

which is impossible, as already observed. Thus $\sigma = 1$ and

$$n_1 - n_3 = n_4 - n_2, \qquad (n_1 - n_3)(n_1 + n_3) = (n_4 - n_2)(n_4 + n_2).$$

If $n_1 \neq n_3$, then the second equality implies $n_1 + n_3 = n_4 + n_2$. Therefore the sum of the two equalities gives

$$n_1 = n_4, \qquad n_3 = n_2,$$

hence $j_2 + j_3 = 0$ because $\sigma_2 = -\sigma_3$. If, instead, $n_1 = n_3$, then also $n_2 = n_4$, and $j_1 + j_3 = 0$ because $\sigma_1 = -\sigma_3$. □

Appendix B. Tame estimates

In this appendix we remind classical tame estimates for changes of variables, composition of functions and the Hilbert transform, in Sobolev class on the torus, which are used in the paper. For these classical estimates see also, for example: [21, Appendix G]; [16, Appendix]; [8, Section 2]; [17]. Before that, remind standard Sobolev norms properties (Lemma B.1) and tame estimates for operators (Lemma B.2).

Lemma B.1. *Let $d \in \mathbb{N}$, $d \geqslant 1$, and $s_0 > d/2$. There exists an increasing function $C(s) > 0$, $s \geqslant s_0$, with the following properties.*

(i) *Embedding.* $\|u\|_{L^\infty} \leqslant C(s_0)\|u\|_{s_0}$ *for all* $u \in H^{s_0}(\mathbb{T}^d, \mathbb{C})$.
(ii) *Algebra.* $\|uv\|_{s_0} \leqslant C(s_0)\|u\|_{s_0}\|v\|_{s_0}$ *for all* $u, v \in H^{s_0}(\mathbb{T}^d, \mathbb{C})$.
(iii) *Interpolation. For* $0 \leqslant s_1 \leqslant s \leqslant s_2$, $s = \lambda s_1 + (1 - \lambda)s_2$,

$$\|u\|_s \leqslant 2\|u\|_{s_1}^{\lambda}\|u\|_{s_2}^{1-\lambda} \quad \forall u \in H^{s_2}\big(\mathbb{T}^d, \mathbb{C}\big). \tag{B.1}$$

For $0 \leqslant s_1 \leqslant \sigma_1 \leqslant \sigma_2 \leqslant s_2$,

$$\|u\|_{\sigma_1}\|u\|_{\sigma_2} \leqslant 4\|u\|_{s_1}\|u\|_{s_2} \quad \forall u \in H^{s_2}\big(\mathbb{T}^d, \mathbb{C}\big). \tag{B.2}$$

(B.1), (B.2) *also hold with all* $\|u\|_s$ *replaced by* $|u|_s$, $u \in W^{s,\infty}(\mathbb{T}^d)$, $s \in \mathbb{N}$.

(iv) *Asymmetric tame product. For $s \geqslant s_0$,*

$$\|uv\|_s \leqslant C(s)\|u\|_s\|v\|_{s_0} + C(s_0)\|u\|_{s_0}\|v\|_s \quad \forall u, v \in H^s(\mathbb{T}^d). \tag{B.3}$$

(v) *Mixed norms tame product. For $s \geqslant 0$, $s \in \mathbb{N}$,*

$$\|uv\|_s \leqslant C(s)\big(\|u\|_s|v|_0 + \|u\|_0|v|_s\big) \quad \forall u \in H^s(\mathbb{T}^d),\ v \in W^{s,\infty}(\mathbb{T}^d). \tag{B.4}$$

Proof. (iii): see [31, p. 269]. (iv): see the appendix of [9]. (v): write $D^\alpha(uv) = \sum_{\beta+\gamma=\alpha}(D^\beta u)(D^\gamma v)$, use the elementary inequality $\|(D^\beta u)(D^\gamma v)\|_0 \leqslant \|D^\beta u\|_0|D^\gamma v|_0$, then the interpolation (iii). □

Lemma B.2. *Let $0 \leqslant s_0 \leqslant s$, and $c_0, c_s > 0$. Let S be a closed linear subspace of Z (for example, $S = Z_0$ or $S = Z_{0N} \cap Y$). Let $T : S \cap H^{s_0} \to S \cap H^{s_0}$ be a linear operator.*

(i) *Tame Neumann series. Let $c_0 \leqslant 1/2$. Assume that*

$$\big\|(T-I)f\big\|_s \leqslant c_0\|f\|_s + c_s\|f\|_{s_0}, \qquad \big\|(T-I)f\big\|_{s_0} \leqslant c_0\|f\|_{s_0} \tag{B.5}$$

for all $f \in S \cap H^{s_0}$. Then $T : S \cap H^{s_0} \to S \cap H^{s_0}$ is invertible, with

$$\big\|\big(T^{-1}-I\big)f\big\|_s \leqslant 2c_0\|f\|_s + 4c_s\|f\|_{s_0}, \qquad \big\|\big(T^{-1}-I\big)f\big\|_{s_0} \leqslant 2c_0\|f\|_{s_0}. \tag{B.6}$$

(ii) *Tame derivative of the inverse with respect to a parameter. Let*

$$\big\|T^{-1}f\big\|_s \leqslant c_0\|f\|_s + c_s\|f\|_{s_0}, \qquad \big\|T^{-1}f\big\|_{s_0} \leqslant c_0\|f\|_{s_0} \tag{B.7}$$

for all $f \in S \cap H^{s_0}$. Assume that T depends in a C^1 way on a parameter λ in a Banach space, and the derivative $(\partial_\lambda T)[\hat\lambda]f$ of Tf with respect to λ in the direction $\hat\lambda$ satisfies

$$\big\|(\partial_\lambda T)[\hat\lambda]f\big\|_s \leqslant b_0\|f\|_s + b_s\|f\|_{s_0}, \qquad \big\|(\partial_\lambda T)[\hat\lambda]f\big\|_{s_0} \leqslant b_0\|f\|_{s_0} \tag{B.8}$$

for all $f \in S \cap H^{s_0}$, for some constants $b_0, b_s > 0$. Then T^{-1} is also a C^1 function of λ,

$$\partial_\lambda T^{-1}[\hat\lambda] = -T^{-1}\big(\partial_\lambda T[\hat\lambda]\big)T^{-1}, \tag{B.9}$$

$$\big\|\partial_\lambda T^{-1}[\hat\lambda]f\big\|_s \leqslant \big(4c_0^2 b_0\big)\|f\|_s + \big(16c_0b_0c_s + 4c_0^2 b_s\big)\|f\|_{s_0}, \qquad \big\|\partial_\lambda T^{-1}[\hat\lambda]f\big\|_{s_0} \leqslant c_0^2 b_0\|f\|_{s_0}. \tag{B.10}$$

Proof. (i). Let $A := I - T$. By induction,

$$\big\|A^n f\big\|_s \leqslant c_0^n\|f\|_s + c_s n c_0^{n-1}\|f\|_{s_0}, \qquad \big\|A^n f\big\|_{s_0} \leqslant c_0^n\|f\|_{s_0}, \quad n \geqslant 1,$$

where $A^2 f$ means $A(Af)$ and so on. Since $c_0 \leqslant 1/2$,

$$\sum_{n=1}^\infty \big\|A^n f\big\|_s \leqslant c_0\left(\sum_{n=0}^\infty c_0^n\right)\|f\|_s + c_s\left(\sum_{n=1}^\infty n c_0^{n-1}\right)\|f\|_{s_0} \leqslant 2c_0\|f\|_s + 4c_s\|f\|_{s_0}.$$

Hence, by Neumann series, T is invertible, and $T^{-1} - I = \sum_{n=1}^\infty A^n$ satisfies (B.6).

(ii) Formula (B.9) follows from differentiating the equality $TT^{-1}f = f$ with respect to the parameter λ. (B.7), (B.8), (B.9) give (B.10). □

Lemma B.3 *(Composition of functions).* (i) *Let $f(x,y)$ be defined for $y = (y_1,\ldots,y_m)$ in the ball $B_1 = \{y \in \mathbb{R}^m: |y|^2 = \sum_{i=1}^m |y_i|^2 < 1\}$ and all $x = (x_1,\ldots,x_d) \in \mathbb{R}^d$, and let f be 2π periodic in $x_1,\ldots,x_d$. Assume that f has continuous derivatives up to order $r \geqslant 0$ which are bounded by $\|f\|_{C^r} < \infty$. Let $u \in H^r(\mathbb{T}^d, \mathbb{R}^m)$, with $u(x) \in B_1$ for all x. Let $\tilde f(u)(x) = f(x, u(x))$. Then*

$$\big\|\tilde f(u)\big\|_r \leqslant C\|f\|_{C^r}\big(\|u\|_r + 1\big).$$

The constant C depends on r, d, m.

(ii) *Let $f, \tilde f$ be like in* (i)*, and assume that $\|\partial_y^\alpha f\|_{C^r} \leqslant K_r$ for all $|\alpha| \leqslant N+1$. Let $\tilde f^{(n)}(u)[h]^n$ denote the n-th Fréchet derivative of $\tilde f$ at u in the direction $[h]^n = [h,\ldots,h]$. ($\tilde f^{(n)}(u)(x)$ is simply the n-th Fréchet derivative* *of* $f(x,y)$ *with respect to the variable* y*, evaluated at the point* $(x,y)=(x,u(x))$*.) If* $u,h\in H^r(\mathbb{T}^d,\mathbb{R}^m)$*, with* $u(x), u(x)+h(x)\in B_1$ *for all* x*, then*

$$\left\| \tilde f(u+h) - \sum_{n=0}^{N} \frac{1}{n!}\tilde f^{(n)}(u)[h]^n \right\|_r \leqslant CK_r \|h\|_{L^\infty}^N \big(\|h\|_r + \|h\|_{L^\infty}\|u\|_r\big).$$

C depends on r,d,m,N.

(iii) *Let* $u\in H^{r+p}(\mathbb{T}^d,\mathbb{R})$*. Let* $D^k u(x)$ *be the list of all partial derivatives* $\partial_x^\alpha u(x)$ *of order* $|\alpha|=k$*. Let* $\tilde f(u)(x) = f(x,u(x),Du(x),\dots,D^p u(x))$*, where* f *is like in (i) for a suitable* m*. Then*

$$\|\tilde f(u)\|_r \leqslant C\|f\|_{C^r}\big(\|u\|_{r+p}+1\big)$$

provided $(u(x),Du(x),\dots,D^p u(x))\in B_1$ *for all* x*. C depends on* r,d,p.

If, in addition, $\|\partial_y^\alpha f\|_{C^r}\leqslant K_r$ *for all* $|\alpha|\leqslant N+1$*, then*

$$\left\| \tilde f(u+h) - \sum_{n=0}^{N} \frac{1}{n!}\tilde f^{(n)}(u)[h]^n \right\|_r \leqslant CK_r \|h\|_{W^{p,\infty}}^N \big(\|h\|_{r+p} + \|h\|_{W^{p,\infty}}\|u\|_{r+p}\big). \tag{B.11}$$

C depends on r,d,p,N.

(iv) *The previous statements also hold when all the* L^2*-based Sobolev norms* $\|u\|_r$ *are replaced by the* L^∞*-based Sobolev norms* $|u|_r = \|u\|_{W^{r,\infty}} = \sum_{k\leqslant r}\|D^k u\|_{L^\infty}$.

Proof. (i) See [31, Section 2, pp. 272–275]. (ii) Use Taylor's formula with integral rest and the inequality $\|\int_0^1 u(\lambda,\cdot)\,d\lambda\|_r^2 \leqslant \int_0^1 \|u(\lambda,\cdot)\|_r^2\,d\lambda$, which holds for $u(\lambda,x)\in H^r(\mathbb{T}_x^d)$, depending on the parameter λ, by Hölder's inequality. As an alternative, see [33, Lemma 7 in Appendix, pp. 202–203]. (iii) Consider $\tilde u = (u,Du,\dots,D^p u)$ and apply (i), (ii). See also [31, p. 275]. (iv) See [15, Lemma 2.3.4, p. 147], for (i) in the $W^{r,\infty}$ case. (ii), (iii) can be adapted with no difficulty (the $W^{r,\infty}$ norms satisfy the algebra and interpolation properties, which are the core of the proofs). □

(iii) of Lemma B.3 is used for the nonlinearity $\mathcal{N}(u)$. (ii) is also used for $N=0$, $u=0$, mainly for $f(y)=e^y$, $f(y)=\cos(y)$, $f(y)=(1+y)^p$, $p\in\mathbb{R}$:

$$\big|f(h)-f(0)\big|_s \leqslant C|h|_s \quad \forall h\in W^{s,\infty}\big(\mathbb{T}^2,\mathbb{R}\big),\ |h|_0<1, \tag{B.12}$$

where C depends on f and s.

The next lemma is also classical, see for example [16, Appendix], and [21, Appendix G]. However, in those papers it is stated slightly differently than in Lemma B.4, especially part (i), therefore we prove it, adapting Lemma 2.3.6 on p. 149 of [15].

Lemma B.4 *(Change of variable). Let* $p:\mathbb{R}^d\to\mathbb{R}^d$ *be a* 2π*-periodic function in* $W^{m,\infty}$*,* $m\geqslant 1$*, with* $|Dp|_0\leqslant 1/2$*. 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 periodic,* $q\in W^{m,\infty}(\mathbb{T}^d,\mathbb{R}^d)$*, and* $|q|_m\leqslant C|p|_m$*. More precisely,*

$$|q|_0=|p|_0, \qquad |Dq|_0\leqslant 2|Dp|_0\leqslant 1, \qquad |Dq|_{m-1}\leqslant C|Dp|_{m-1}.$$

The constant C depends on d,m.

(ii) *If* $u\in H^m(\mathbb{T}^d,\mathbb{C})$*, then* $u\circ f(x)=u(x+p(x))$ *is also in* H^m*, and, with the same C as in* (i),

$$\|u\circ f\|_m \leqslant C\big(\|u\|_m + |Dp|_{m-1}\|u\|_1\big).$$

(iii) *Part* (ii) *also holds with* $\|\ \|_k$ *replaced by* $|\ |_k$*, namely* $|u\circ f|_m\leqslant C(|u|_m+|Dp|_{m-1}|u|_1)$.

Proof. (i) For every $y\in\mathbb{R}^d$, the map $G_y:\mathbb{R}^d\to\mathbb{R}^d$, $G_y(x)=y-p(x)$ is a contraction because $|Dp|_0\leqslant 1/2$, therefore G_y has a unique fixed point $x=G_y(x)$ in $\mathbb{R}^d$, and the inverse function $g=f^{-1}:\mathbb{R}^d\to\mathbb{R}^d$ is globally defined. Let $q(y):=g(y)-y$.

Since p is periodic, $f(x+2\pi m)=f(x)+2\pi m$ for all $m\in\mathbb{Z}^d$. Applying g to this equality gives $x+2\pi m=g(f(x)+2\pi m)$, namely $g(y)+2\pi m=g(y+2\pi m)$ where $y=f(x)$, and this means that q is periodic. Hence g, like f, is also a bijection of $\mathbb{T}^d$ onto itself.

The identity $f(g(y))=y$ gives

$$q(y)+p\big(y+q(y)\big)=0, \qquad q\big(x+p(x)\big)+p(x)=0 \quad \forall x,y\in\mathbb{R}^d. \tag{B.13}$$

(B.13) implies that $|q|_0=|p|_0$. By Neumann series, the matrix $Df(x)=I+Dp(x)$ is invertible for a.e. x, $(Df(x))^{-1}=\sum_{n=0}^{\infty}(-Dp(x))^n$, and $|(Df)^{-1}|_0\leqslant 2$. Differentiating (B.13),

$$Dq(y)=-\big[Df\big(y+q(y)\big)\big]^{-1}Dp\big(y+q(y)\big)=\sum_{n=1}^{\infty}\big[-Dp\big(g(y)\big)\big]^n, \tag{B.14}$$

whence $|Dq|_0\leqslant 2|Dp|_0\leqslant 1$. Differentiating (B.14),

$$\big(D^2q\big)(y)=-\big[(Df)\big(g(y)\big)\big]^{-1}\big(D^2p\big)\big(g(y)\big)Dg(y)Dg(y),$$

and $|D^2q|_0\leqslant 8|D^2p|_0$. (i) is proved for $m=1$ and $m=2$.

In general, by the "chain rule", the m-th Fréchet derivative of the composition of functions $u\circ v$ is

$$D^m(u\circ v)(x)=\sum_{k=1}^{m}\sum_{j_1+\ldots+j_k=m}C_{kj}\big(D^ku\big)\big(v(x)\big)\big[D^{j_1}v(x),\ldots,D^{j_k}v(x)\big], \tag{B.15}$$

where $j_1,\ldots,j_k\geqslant 1$, and C_{kj} are constants depending on $k,j_1,\ldots,j_k$ [15, p. 147]. Apply (B.15) to $f\circ g$: since $f(g(y))=y$, $D^m(f\circ g)=0$ for all $m\geqslant 2$. Separate $k=1$ from $k\geqslant 2$ in the sum (B.15) and solve for D^mg,

$$D^mg(y)=-Dg(y)\sum_{k=2}^{m}\sum_{j_1+\ldots+j_k=m}C_{kj}\big(D^kf\big)\big(g(x)\big)\big[D^{j_1}g(y),\ldots,D^{j_k}g(y)\big].$$

$D^mg=D^mq$ and $D^kf=D^kp$ because $k,m\geqslant 2$. Since $k\geqslant 2$, it is $1\leqslant j_i\leqslant m-1$ for all $i=1,\ldots,k$, because there are at least two j_1,j_2, each of them $\geqslant 1$, and $\sum j_i=m$. For $k=m$ one has $j_i=1$ for all $i=1,\ldots,m$, and the corresponding term in the sum is estimated

$$\big|\big(D^mp\big)\circ g[Dg,\ldots,Dg]\big|_0\leqslant\big|D^mp\big|_0|Dg|_0^m\leqslant C|Dp|_{m-1},$$

because $|Dg|_0=|I+Dq|_0\leqslant 2$. For $2\leqslant k\leqslant m-1$, at least one among $j_1,\ldots,j_k$ is $\geqslant 2$ (otherwise $k=m$). Let ℓ be the number of indices j_i that are $\geqslant 2$, so that $1\leqslant\ell\leqslant k$. It remains to estimate

$$\sum_{k=2}^{m-1}\sum_{\ell=1}^{k}\sum_{\sigma_1+\cdots+\sigma_\ell=m-k+\ell}C_{k\ell\sigma}\big(D^kp\big)\big(g(y)\big)\big[Dg(y)\big]^{k-\ell}\big[D^{\sigma_1}q(y),\ldots,D^{\sigma_\ell}q(y)\big], \tag{B.16}$$

where indices $j_i\geqslant 2$ have been renamed $\sigma_1,\ldots\sigma_\ell$, the number of indices $j_i=1$ is $k-\ell$, and $D^{\sigma_i}g=D^{\sigma_i}q$ because $\sigma_i\geqslant 2$. Every factor Dg in (B.16) is estimated by $|Dg|_0\leqslant 2$. For the remaining factors use the interpolation between 0 and $m-2$, which is possible because $1\leqslant\sigma_i-1\leqslant m-2$, and use the formula $\sigma_1+\cdots+\sigma_\ell=m-k+\ell$,

$$\begin{aligned}
\big|\big(D^kp\big)\circ g\big(D^{\sigma_1}q\big)\cdots\big(D^{\sigma_\ell}q\big)\big|_0 &\leqslant\big|D^{k-2}D^2p\big|_0\big|D^{\sigma_1-1}Dq\big|_0\cdots\big|D^{\sigma_\ell-1}Dq\big|_0\\
&\leqslant C\big|D^2p\big|_0^{\frac{m-2-(k-2)}{m-2}}\big|D^2p\big|_{m-2}^{\frac{k-2}{m-2}}\prod_{i=1}^{\ell}|Dq|_0^{\frac{m-2-(\sigma_i-1)}{m-2}}|Dq|_{m-2}^{\frac{\sigma_i-1}{m-2}}\\
&=C|Dq|_0^{\ell-1}\big(\big|D^2p\big|_0|Dq|_{m-2}\big)^{1-\frac{k-2}{m-2}}\big(\big|D^2p\big|_{m-2}|Dq|_0\big)^{\frac{k-2}{m-2}}\\
&\leqslant C|Dq|_0^{\ell-1}\big(\big|D^2p\big|_0|Dq|_{m-2}+\big|D^2p\big|_{m-2}|Dq|_0\big)\\
&\leqslant C\big(|Dq|_{m-2}+|Dp|_{m-1}\big).
\end{aligned}$$

Collecting all the terms in the sum, we have proved that

$$\big|D^mq\big|_0\leqslant C\big(|Dp|_{m-1}+|Dq|_{m-2}\big). \tag{B.17}$$

Now use the induction on m. We have already proved (P_m) $|Dq|_{m-1} \leqslant C|Dp|_{m-1}$ for $m=2$. Assume that (P_{m-1}) holds. Then (P_m) follows from (B.17).

(iii) follows a similar argument, using formula (B.15) and interpolation for $W^{k,\infty}$ norms; see [15, Lemma 2.3.4, p. 147].

(ii) $\|u \circ f\|_0 \leqslant C\|u\|_0$, because, changing variable $x = g(y)$ in the integral,

$$\|u \circ f\|_0^2 = \int_{\mathbb{T}^d} \big|u\big(f(x)\big)\big|^2\, dx = \int_{\mathbb{T}^d} \big|u(y)\big|^2 \big|\det Dg(y)\big|\, dy \leqslant \|\det Dg\|_{L^\infty} \int_{\mathbb{T}^d} \big|u(y)\big|^2 dy \leqslant C\|u\|_0^2. \tag{B.18}$$

The m-th derivative of $u \circ f$, $m \geqslant 1$, is given by formula (B.15). The L^2 norm of a typical term of the sum is estimated by

$$\big\|D^k u\big(f(x)\big)\big[D^{j_1} f(x), \dots, D^{j_k} f(x)\big]\big\|_0 \leqslant \big\|\big(D^k u\big) \circ f\big\|_0 \big\|D^{j_1} f\big\|_{L^\infty} \cdots \big\|D^{j_k} f\big\|_{L^\infty}.$$

$\|(D^k u) \circ f\|_0 \leqslant C\|D^k u\|_0 \leqslant C\|Du\|_{k-1}$ by (B.18). Use interpolation (B.1) for $\|Du\|_{k-1}$ and interpolation with $W^{k,\infty}$ norms for all $D^{j_i-1}Df$ between 0 and $m-1$, which is possible because $k-1$, j_i-1 are all in the interval $[0, m-1]$. (Remember that Df is periodic, while f is not.) We get

$$\big\|D^k u\big\|_0 \big\|D^{j_1} f\big\|_{L^\infty} \cdots \big\|D^{j_k} f\big\|_{L^\infty} \leqslant C\|Df\|_{L^\infty}^{k-1} \big(\|Du\|_{m-1}\|Df\|_{L^\infty} + \|Du\|_0 \|Df\|_{W^{m-1,\infty}}\big).$$

Now $\|Df\|_{L^\infty} \leqslant 2$, and $\|Df\|_{W^{m-1,\infty}} \leqslant C(1 + \|Dp\|_{W^{m-1,\infty}})$. The sum gives the thesis. □

The next lemma estimates the commutator of $\mathcal{H}$ with multiplication operators and changes of variables that are used in the paper. See also [21, Appendices H and I].

Lemma B.5 *(Commutators of $\mathcal{H}$). 1) Let $s, m_1, m_2 \in \mathbb{N}$, with $s \geqslant 2$, $m_1, m_2 \geqslant 0$, $m = m_1 + m_2$. Let $f(t,x) \in H^{s+m}(\mathbb{T}^2, \mathbb{C})$. Then $[f, \mathcal{H}]u = f\mathcal{H}u - \mathcal{H}(fu)$ satisfies*

$$\big\|\partial_x^{m_1}[f, \mathcal{H}]\partial_x^{m_2} u\big\|_s \leqslant C(s)\big(\|u\|_s \|f\|_{m+2} + \|u\|_2 \|f\|_{m+s}\big).$$

2) Let $a : \mathbb{T} \to \mathbb{T}$ a function, and $Au(t,x) = u(a(t), x)$. Then $[A, \mathcal{H}] = 0$.

3) There exists a universal constant $\delta \in (0,1)$ with the following property. Let $s, m_1, m_2 \in \mathbb{N}$, $m = m_1 + m_2$, $\beta(t,x) \in W^{s+m+1,\infty}(\mathbb{T}^2, \mathbb{R})$, with $|\beta|_1 \leqslant \delta$. Let $Bh(t,x) = h(t, x + \beta(t,x))$, $h \in H^s(\mathbb{T}^2, \mathbb{C})$. Then

$$\big\|\partial_x^{m_1}\big(B^{-1}\mathcal{H}B - \mathcal{H}\big)\partial_x^{m_2} h\big\|_s \leqslant C(s,m)\big(|\beta|_{m+1}\|h\|_s + |\beta|_{s+m+1}\|h\|_0\big).$$

Proof. 1) Let $u(t,x) = \sum_{k\in\mathbb{Z}} u_k(t)e^{ikx}$, $f(t,x) = \sum_{k\in\mathbb{Z}} f_k(t)e^{ikx}$, and

$$S = \big\{(k,j) \in \mathbb{Z}^2\colon\ \mathrm{sign}(k) - \mathrm{sign}(j) \neq 0\big\}, \qquad S(k) = \big\{j \in \mathbb{Z}\colon (k,j) \in S\big\}.$$

Since $\mathcal{H}(e^{ikx}) = -i\,\mathrm{sign}(k)e^{ikx}$,

$$\partial_x^{m_1}[f, \mathcal{H}]\partial_x^{m_2} u = \sum_{k,j\in\mathbb{Z}} f_{j-k}(t)u_k(t)\delta(k,j)(ij)^{m_1}(ik)^{m_2}e^{ijx} = \sum_{(k,j)\in S} (\text{the same}),$$

where $\delta(k,j) := -i(\mathrm{sign}(k) - \mathrm{sign}(j))$. If $(k,j) \in S$, then

$$|k-j| = |k| + |j|, \quad |j| \leqslant |j-k|, \ |k| \leqslant |j-k|.$$

Therefore $|j^{m_1}k^{m_2}| \leqslant |k-j|^m$. If j, k are Fourier indices for the space and n, l for the time,

$$\big\|\partial_x^{m_1}[f, \mathcal{H}]\partial_x^{m_2} u\big\|_s^2 \leqslant \sum_{n,j}\bigg(\sum_{l,k} |f_{(n-l,j-k)}||j-k|^m |u_{(l,k)}|\bigg)^2 \big\langle (n,j)\big\rangle^{2s} \leqslant \sum_{a\in\mathbb{Z}^2}\bigg(\sum_{b\in\mathbb{Z}^2} \big|\big(\partial_x^m f\big)_{a-b}\big||u_b|\bigg)^2 \langle a\rangle^{2s}$$

and this gives the usual tame estimate for the product $(\partial_x^m f)u$. The estimate holds with $\|\ \|_{s_0}$ with $s_0 > d/2 = 2/2 = 1$, so we fix $s_0 = 2$.

2) Trivially $A\mathcal{H}u(t,x) = \sum_k u_k(a(t))(-i\,\mathrm{sign}\,k)e^{ikx} = \mathcal{H}Au(t,x)$.

3) Following [21, Appendix I], it is convenient to use the representation of $\mathcal{H}$ as a principal value integral,

$$\mathcal{H}u(t,x) = \frac{-1}{2\pi}\,\text{p.v.}\int_{\mathbb{T}} \frac{u(t,x')}{\tan\frac{1}{2}(x-x')}\,dx' = \frac{-1}{2\pi}\lim_{\varepsilon\to 0^+}\left\{\int_{x-\pi}^{x-\varepsilon} + \int_{x+\varepsilon}^{x+\pi}\right\}\frac{u(t,x')}{\tan\frac{1}{2}(x-x')}\,dx'. \tag{B.19}$$

Let $I+\tilde\beta$ be the inverse of $I+\beta$, namely $x+\beta(t,x)=y$ if and only if $x=y+\tilde\beta(t,y)$. Changing variable $x'+\beta(t,x')=y'$, $dx'=(1+\tilde\beta_{y'}(t,y'))\,dy'$ in the integral,

$$B^{-1}\mathcal{H}Bu(t,y) = \frac{1}{\pi}\,\text{p.v.}\int_{-\pi}^{\pi} u(t,y')\partial_{y'}\left\{\log\sin\left(\frac{1}{2}\left[y+\tilde\beta(t,y)-y'-\tilde\beta(t,y')\right]\right)\right\}dy',$$

therefore

$$\left(B^{-1}\mathcal{H}B-\mathcal{H}\right)u(t,y) = \int_{\mathbb{T}} u(t,y')K(t,y,y')\,dy', \tag{B.20}$$

where the kernel K is

$$K(t,y,y') = \frac{1}{\pi}\partial_{y'}\log\left(\frac{\sin\frac{1}{2}[y+\tilde\beta(t,y)-y'-\tilde\beta(t,y')]}{\sin\frac{1}{2}(y-y')}\right).$$

If β is sufficiently regular, then K is bounded, and the integral in (B.20) is no longer a singular one. Denote $\mathcal{R} = B^{-1}\mathcal{H}B-\mathcal{H}$. Then

$$\partial_y^{m_1}\mathcal{R}\partial_y^{m_2}u(t,y) = \int_{\mathbb{T}}\left(\partial_{y'}^{m_2}u\right)(t,y')\partial_y^{m_1}K(t,y,y')\,dy' = \int_{\mathbb{T}} u(t,y')(-1)^{m_2}\partial_{y'}^{m_2}\partial_y^{m_1}K(t,y,y')\,dy',$$

every space derivative goes on K and does not affect u. Hence

$$\|\mathcal{R}u\|_0^2 = \int_{\mathbb{T}^2}\left|\int_{\mathbb{T}} u(t,y')K(t,y,y')\,dy'\right|^2 dy\,dt \leqslant C\int_{\mathbb{T}^3}\left|u(t,y')\right|^2\left|K(t,y,y')\right|^2 dy'\,dy\,dt \leqslant C|K|_0^2\|u\|_0^2,$$

for $\|\partial_y^s(\partial_y^{m_1}\mathcal{R}\partial_y^{m_2}u)\|_0$ replace K with $\partial_y^{s+m_1}\partial_{y'}^{m_2}K$ and for $\|\partial_t^s(\partial_y^{m_1}\mathcal{R}\partial_y^{m_2}u)\|_0$ calculate the usual derivatives of a product. Thus

$$\left\|\partial_y^{m_1}\mathcal{R}\partial_y^{m_2}u\right\|_s \leqslant C\left(\|u\|_s|K|_m + \|u\|_0|K|_{s+m}\right).$$

Now write $K=(1/\pi)\partial_{y'}\log(1+f)$, where

$$f(t,y,y') = \frac{\sin\frac{1}{2}[y+\tilde\beta(t,y)-y'-\tilde\beta(t,y')] - \sin\frac{1}{2}(y-y')}{\sin\frac{1}{2}(y-y')},$$

and decompose $f=abc$,

$$a(y,y') = \frac{\frac{1}{2}(y-y')}{\sin\frac{1}{2}(y-y')}, \qquad b(t,y,y') = \frac{\tilde\beta(t,y)-\tilde\beta(t,y')}{y-y'} = \int_0^1 \tilde\beta_y\left(t,\lambda y+(1-\lambda)y'\right)d\lambda,$$

$$c(t,y,y') = \int_0^1\cos\left(\frac{y-y'+\lambda[\tilde\beta(t,y)-\tilde\beta(t,y')]}{2}\right)d\lambda.$$

$a\in C^\infty$ for $|y'-y|\leqslant\pi$ (by periodicity, take $\mathbb{T}=[y-\pi,y+\pi]$ when integrating in dy'). $|b|_s\leqslant C|\tilde\beta|_{s+1}\leqslant C|\beta|_{s+1}$ by Lemma B.4(i). All the derivatives of c of order $\leqslant s$ are bounded if $\tilde\beta\in W^{s,\infty}$, with tame estimate

$$|c|_s\leqslant C\left(s,|\tilde\beta|_0\right)\left(1+|\tilde\beta|_s\right)\leqslant C\left(s,|\beta|_0\right)\left(1+|\beta|_s\right).$$

As a consequence $|f|_0\leqslant 1/2$ if $|\beta|_1\leqslant\delta$ for some universal $\delta\in(0,1)$, and $|K|_s\leqslant C(s)|\beta|_{s+1}$. □

Remark B.6. Inequality 1) of Lemma B.5 can also be proved in a simple way using (B.19), see [21, Appendix H].

Appendix C. Proofs

Proof of Proposition 6.1. Apply Lemma B.3(iv): let $f(x,y) = \partial_y^\alpha g_i(x,y)$, $|\alpha| = 1$. By (1.5), $\partial_y^\beta f(x,0) = 0$ for all $|\beta| \leqslant 2$, and, by Taylor's formula (B.11) for $N = 2$ (with $\tilde f$ defined as in Lemma B.3),

$$\big|\tilde f(U)\big|_s = \left|\tilde f(U) - \sum_{n=0}^{2} \frac{1}{n!} \tilde f^{(n)}(0)[U]^n\right|_s \leqslant C(s)|U|_2^2 |U|_{s+2} \leqslant C(s)\|U\|_4^2 \|U\|_{s+4}. \tag{C.1}$$

Suppose that $a_1 = (\partial_y^\alpha g_i)(x, U, \mathcal{H}U, \ldots) = \tilde f(U)$, where $U = \varepsilon \bar v + \varepsilon^2 u$. Then (C.1) gives

$$\begin{aligned}|a_1|_s &\leqslant C(s)\big\|\varepsilon\bar v + \varepsilon^2 u\big\|_4^2 \big\|\varepsilon \bar v + \varepsilon^2 u\big\|_{s+4} \leqslant \varepsilon^3 C(s)\big(\|\bar v\|_4 + \varepsilon K\big)^2 \big(\|\bar v\|_{s+4} + \varepsilon\|u\|_{s+4}\big)\\ &\leqslant \varepsilon^3 C(s,K)\big(1 + \|u\|_{s+4}\big)\end{aligned}$$

because $\|u\|_4 \leqslant K$ and $\|\bar v\|_{s+4}$ is a certain constant $C(s)$ depending on s. Also $a_2, a_4, a_3 - 3U^2$ and $a_5 - 3(U^2)_x$ are of the type $(\partial_y^\alpha g_i)(x, U, \mathcal{H}U, \ldots)$, therefore they satisfy the same estimate as a_1. The additional part in a_3 and a_5 comes from the cubic term $\partial_x(U^3)$ of the nonlinearity $\mathcal{N}(U)$. One has

$$\big|U^2 - \varepsilon^2 \bar v^2\big|_s = \varepsilon^3 \big|2\bar v u + \varepsilon u^2\big|_s \leqslant \varepsilon^3 C(s,K)|u|_s \leqslant \varepsilon^3 C(s,K)\|u\|_{s+2}$$

because $U = \varepsilon\bar v + \varepsilon^2 u$, and the estimate for $a_3 - \varepsilon^2 3\bar v^2$ follows. Similarly for a_5.

The derivatives $\partial_u a_1$ and $\partial_\varepsilon a_1$ are obtained differentiating the equality $a_1 = (\partial_y^\alpha g_i)(x, U, \mathcal{H}U, \ldots)$, therefore they involve $\partial_y^\beta g_i$ with $|\beta| = 2$. Then apply Taylor's formula (B.11) with $N = 1$ and evaluate at U, as above. □

Remark C.1. In the estimate for $\partial_u a_i$ there is a factor ε^2 more than in the one for $\partial_\varepsilon a_i$ because $\partial_u U[h] = \varepsilon^2 h = O(\varepsilon^2)$, while $\partial_\varepsilon U = \bar v + 2\varepsilon u = O(1)$. The point becomes very evident in the simplest case $g(x, U, \ldots) = U^4$.

Proof of Proposition 7.2. By Proposition 6.1, for $s = 0$ and $\varepsilon < \varepsilon_0$, $|a_1|_0 \leqslant \varepsilon^3 C(K) \leqslant \varepsilon_0^3 C(K) \leqslant 1/2$ if ε_0 is small enough. $|\int a\,dx|_s \leqslant 2\pi|a|_s$ for all $a(t,x)$. Applying (B.12) with $f(y) = (1+y)^p$, $p = -1/2, -2$ gives

$$|\rho - 1|_s \leqslant C(s,K)|a_1|_s \leqslant \varepsilon^3 C(s,K)\big(1 + \|u\|_{s+4}\big), \quad 0 \leqslant s \leqslant r. \tag{C.2}$$

Differentiating the formula for $\rho(u,\varepsilon)$, and using estimates on a_1,

$$\big|\partial_u \rho(u,\varepsilon)[h]\big|_s \leqslant C(s,K)\big(\big|\partial_u a_1[h]\big|_s + |a_1|_s \big|\partial_u a_1[h]\big|_0\big) \leqslant \varepsilon^4 C(s,K)\big(\|h\|_{s+4} + \|u\|_{s+4}\|h\|_4\big), \tag{C.3}$$

and similarly $|\partial_\varepsilon \rho(u,\varepsilon)|_s \leqslant \varepsilon^2 C(s,K)(1 + \|u\|_{s+4})$, for all $0 \leqslant s \leqslant r$.

$\mu_2 = \Pi_C(\rho)$, therefore, using (C.2) with $s = 0$, $|\mu_2 - 1| = |\Pi_C(\rho - 1)| \leqslant |\rho - 1|_0 \leqslant \varepsilon^3 C(0,K)\|u\|_4 = \varepsilon^3 C(K) \leqslant 1/2$. Also, $|\partial_u \mu_2(u,\varepsilon)[h]| = |\Pi_C(\partial_u\rho(u,\varepsilon)[h])| \leqslant |\partial_u \rho(u,\varepsilon)[h]|_0$, then use (C.3) with $s = 0$. Similarly for $\partial_\varepsilon \mu_2$.

α satisfies (7.7), namely $\mu_2(1+\alpha') = \rho$. Thus $\alpha' = \mu_2^{-1}[(\rho - 1) + (1 - \mu_2)]$, whence $|\alpha'|_s \leqslant 2(|\rho - 1|_s + |\mu_2 - 1|)$. Moreover $|\alpha|_{s+1} \leqslant C|\alpha'|_s$ because $\alpha \in Y$, $\alpha(0) = 0$, and $|\alpha(t)| = |\alpha(t) - \alpha(0)| \leqslant \pi|\alpha'|_0$ for all $|t| \leqslant \pi$ (Poincaré inequality for odd functions). The derivatives of α are obtained differentiating the equality $\mu_2(1+\alpha') = \rho$. Similar argument for $\Pi_E\beta$ using (B.12), because $\Pi_E\beta_x = \rho^{1/2}(1+a_1)^{-1/2} - 1$ by (7.8). Thus $\alpha(u,\varepsilon)$ and $\Pi_E\beta(u,\varepsilon)$ satisfy

$$|\alpha|_{s+1} + |\Pi_E\beta|_s + |\Pi_E\beta_x|_s \leqslant \varepsilon^3 C(s,K)\big(1 + \|u\|_{s+4}\big), \tag{C.4}$$

$$\big|\partial_u \alpha[h]\big|_{s+1} + \big|\partial_u(\Pi_E\beta)[h]\big|_s \leqslant \varepsilon^4 C(s,K)\big(\|h\|_{s+4} + \|u\|_{s+4}\|h\|_4\big), \quad 0 \leqslant s \leqslant r, \tag{C.5}$$

$$|\partial_\varepsilon \alpha|_{s+1} + |\partial_\varepsilon \Pi_E\beta|_s \leqslant \varepsilon^2 C(s,K)\big(1 + \|u\|_{s+4}\big). \tag{C.6}$$

σ is defined in (7.11), namely $\sigma = \Pi_{T+C}\{\omega(\Pi_E\beta)_t(1 + \Pi_E\beta_x) + a_3(1 + \Pi_E\beta_x)^2\}$. Since $\Pi_E\beta = O(\varepsilon^3)$, the only term of order ε^2 in σ comes from a_3 and it is $\varepsilon^2\Pi_{T+C}(3\bar v^2)$. $\bar v$ is a finite sum of q_j (5.1), therefore $\Pi_T(\bar v^2) = 0$. As a consequence,

$$\sigma - \varepsilon^2\Pi_C\big(3\bar v^2\big) = \Pi_{T+C}\big\{\omega(\Pi_E\beta_t)(1 + \Pi_E\beta_x) + a_3(\Pi_E\beta_x)(2 + \Pi_E\beta_x) + \big(a_3 - \varepsilon^2 3\bar v^2\big)\big\}.$$

Then, using the estimates for $\Pi_E\beta$, $(a_3 - \varepsilon^2 3\bar v^2)$ and their derivatives,

$$\left|\sigma - \varepsilon^2 \Pi_C\left(3\bar{v}^2\right)\right|_{s-1} \leqslant \varepsilon^3 C(s,K)\left(1 + \|u\|_{s+4}\right), \tag{C.7}$$

$$\left|\partial_u \sigma(u,\varepsilon)[h]\right|_{s-1} \leqslant \varepsilon^4 C(s,K)\left(\|h\|_{s+4} + \|u\|_{s+4}\|h\|_4\right), \quad 1 \leqslant s \leqslant r, \tag{C.8}$$

$$\left|\partial_\varepsilon \sigma(u,\varepsilon) - \varepsilon \Pi_C\left(6\bar{v}^2\right)\right|_{s-1} \leqslant \varepsilon^2 C(s,K)\left(1 + \|u\|_{s+4}\right) \tag{C.9}$$

($s-1$ because $|\Pi_E \beta_t|_{s-1} \leqslant |\Pi_E \beta|_s$.)

By (7.12), $\mu_1 = \Pi_C(\sigma)$, and the estimates for μ_1 follow from (C.7), (C.8), (C.9) with $s = 1$.

Since $\sigma - \mu_1 = \sigma - \Pi_C(\sigma) = \Pi_T(\sigma)$, by (7.11) $\omega\gamma' = \mu_1(1+\alpha') - \sigma = \mu_1\alpha' - \Pi_T(\sigma)$. By Poincaré inequality, $|\gamma|_s \leqslant C|\gamma'|_{s-1}$ because $\gamma \in Y$. The estimates for $\gamma = \Pi_T \beta$ follow from those for σ, α, μ_1 and their derivatives, using the fact that $\omega = 1 + 3\varepsilon^2$. Hence (C.4), (C.5), (C.6) hold not only for $\Pi_E \beta$, but also for $\gamma = \Pi_T \beta$, and, as a consequence, for β too, for $1 \leqslant s \leqslant r$.

By Lemma B.4(i), $|\tilde{\alpha}|_s + |\tilde{\beta}|_s \leqslant C(s)(|\alpha|_s + |\beta_s)$. Choose a smaller ε_0, if necessary, to have $\varepsilon_0^3 C(K) < 1/2$ in (7.18). (7.21), (7.23) hold by Lemma B.4. Since

$$\alpha(t) + \tilde{\alpha}\big(t + \alpha(t)\big) = 0, \qquad \beta(t,x) + \tilde{\beta}\big(t + \alpha(t), x + \beta(t,x)\big) = 0 \quad \forall (t,x) \in \mathbb{T}^2, \tag{C.10}$$

the derivatives of $\tilde{\alpha}, \tilde{\beta}$ with respect to the parameters (u,ε) are obtained by differentiating (C.10) with respect to u or ε, whence

$$\partial_u \tilde{\alpha}[h] = -(1 + \tilde{\alpha}_\tau)\Psi^{-1}\{\partial_u \alpha[h]\}, \qquad \partial_u \tilde{\beta}[h] = -(1 + \tilde{\beta}_y)\Psi^{-1}\{\partial_u \beta[h]\} - \tilde{\beta}_\tau \Psi^{-1}\{\partial_u \alpha[h]\},$$

and similarly for $\partial_\varepsilon \tilde{\alpha}$, $\partial_\varepsilon \tilde{\beta}$. (Given a diffeomorphism depending on a parameter, this is nothing but the formula for the derivative of the inverse diffeomorphism with respect to the parameter.) Using (C.5), (C.6) and (7.23), for $s + 1 \leqslant r$ we get

$$\left|\partial_u \tilde{\beta}[h]\right|_s \leqslant \varepsilon^4 C(s,K)\left(\|h\|_{s+4} + \|u\|_{s+5}\|h\|_5\right), \qquad |\partial_\varepsilon \tilde{\beta}|_s \leqslant \varepsilon^2 C(s,K)\left(1 + \|u\|_{s+5}\right),$$

and the same for $\tilde{\alpha}$. These inequalities also hold for α, β (actually, α, β satisfy (C.5), (C.6), which are stronger).

To prove (7.22), consider the one-parameter family of changes of variables

$$(\Psi_\lambda f)(t,x) = f\big(\psi_\lambda(t,x)\big), \quad \psi_\lambda(t,x) = \big(t + \lambda\alpha(t), x + \lambda\beta(t,x)\big), \ 0 \leqslant \lambda \leqslant 1.$$

One has

$$(\Psi - I)f(t,x) = f\big(\psi_1(t,x)\big) - f\big(\psi_0(t,x)\big) = \int_0^1 (\nabla f)\big(\psi_\lambda(t,x)\big) \cdot \big(\alpha(t), \beta(t,x)\big)\, d\lambda.$$

Use Lemma B.4 to estimate $\|\Psi_\lambda f_t\|_s$ and $\|\Psi_\lambda f_x\|_s$, then use (B.4). The same holds for Ψ^{-1}. The estimate for $\tilde{\Psi}, \tilde{\Psi}^{-1}$ hold because $\|\mathbb{P}h\|_s \leqslant \|h\|_s$ for all s. Repeat the same argument with norms $|\,|_s$ to prove (7.24). By the chain rule, the derivative of Ψf with respect to u in the direction h is

$$\partial_u(\Psi f)[h] = \partial_u\big\{f\big(t + \alpha(t), x + \beta(t,x)\big)\big\}[h] = (\Psi f_t)\partial_u\alpha[h] + (\Psi f_x)\partial_u\beta[h],$$

therefore (7.25) follows using the interpolation (B.4) for products. Similarly for (7.26).

Since

$$\big[1 + \big(\psi^{-1}\alpha'\big)(\tau)\big]\big(1 + \tilde{\alpha}'(\tau)\big) = 1,$$

$(\mathcal{M} - I)$ is the multiplication by the factor $(\Psi^{-1}\alpha') = -\tilde{\alpha}'/(1 + \tilde{\alpha}') =: p$. Hence $(\tilde{\mathcal{M}} - I)f = \mathbb{P}(\mathcal{M} - I)f = \mathbb{P}(pf)$ for all $f \in Z_0$, because $\mathbb{P} = I$ on Z_0. By Lemma B.3, p satisfies the same estimate as $\tilde{\alpha}'$, and $|\tilde{\alpha}'|_s \leqslant C(s)|\alpha'|_s \leqslant C(s)|\alpha|_{s+1}$, then use (C.4) and apply (B.3) to get

$$\|pf\|_s \leqslant \varepsilon^3 C(K)\|f\|_s + \varepsilon^3 C(s,K)\big(1 + \|u\|_{s+4}\big)\|f\|_2, \quad 2 \leqslant s \leqslant r.$$

For the derivatives $\partial_u \mathcal{M}[h]$, $\partial_\varepsilon \mathcal{M}$ use (7.19), (7.20). Apply Lemma B.2 to obtain the estimates for $(\tilde{\mathcal{M}}^{-1} - I)$ and its derivatives.

The estimates for a_i, $i = 6, \dots, 9$ follow from formulae (7.9) and the estimates for Ψ^{-1}. In a_7 put the term $\varepsilon^2 3\bar{v}^2$ in evidence, namely write

$$\frac{\omega\beta_t + a_3(1 + \beta_x)}{1 + \alpha'} = b + q, \quad b := \varepsilon^2 3\bar{v}^2, \ q := \frac{\omega\beta_t + (a_3 - b)(1 + \beta_x) + b(\beta_x - \alpha')}{1 + \alpha'},$$

estimate $\Psi^{-1}(q)$ using (7.23), the inequalities for $\alpha, \beta, (a_3 - b)$, and $|b|_s = C(s)$. For $\Psi^{-1}(b) = b + (\Psi^{-1} - I)b$, use (7.22). Similarly for a_9. Similar calculations for the derivatives $\partial_u a_i[h]$, $\partial_\varepsilon a_i$.

To prove (7.33), write Ψ as the composition of the two changes of variables A, B,

$$\Psi = AB, \qquad Ah(t,x) = h\big(t + \alpha(t), x\big), \qquad Bh(t,x) = h\big(t, x + \beta_1(t,x)\big),$$

where $\beta_1 := A^{-1}(\beta)$, namely $\beta_1(t + \alpha(t), x) = \beta(t,x)$. By Lemma B.5(ii), $\Psi^{-1}\mathcal{H}\Psi = B^{-1}A^{-1}\mathcal{H}AB = B^{-1}\mathcal{H}B$. By the inequality (7.23) for the change of variable A, $|\beta_1|_s \leqslant \varepsilon^3 C(s,K)(1 + \|u\|_{s+4})$. Then apply Lemma B.5(iii).

In $\mathcal{R}_1$ (see (7.3)) the coefficients of $\partial_y^k \mathcal{R}_{\mathcal{H}}$, $k = 0,1,2$, are functions f_k that satisfy $|f_k|_s \leqslant C(s,K)(1 + \|u\|_{s+5})$ for $s + 1 \leqslant r$ (two of them are a_6, a_8 without the denominator $(1 + \alpha')$, the other one is (7.4)). By (B.4), (B.2), and (7.33),

$$\big\| f_k \partial_y^k \mathcal{R}_{\mathcal{H}} \partial_y^m h\big\|_s \leqslant \varepsilon^3 C(s,m,K)\big(\|h\|_s\big(1 + \|u\|_{m+7}\big) + \|h\|_0 \|u\|_{s+m+7}\big), \quad k = 0,1,2,$$

for $m \geqslant 0$, $s + m + 3 \leqslant r$. For the last term in $\mathcal{R}_1$ use (7.2), the estimate for $\Psi^{-1}a_5$, integration by parts $|\Pi_C(f\partial_y^m h)| = |\Pi_C[(\partial_y^m f)h]|$, the inequality $|\Pi_C(fh)| \leqslant C|f|_0\|h\|_0$, Lemma B.4(i) to pass from $\tilde\alpha, \tilde\beta$ to α, β, and (B.2):

$$\big\|\mathbb{P}\big(\Psi^{-1}a_5\big)[\Pi_C, \Psi]\partial_y^m h\big\|_s = \big\|\Psi^{-1}a_5\big\|_s \big|[\Pi_C, \Psi]\partial_y^m h\big| \leqslant \varepsilon^5 C(s,m)\big(1 + \|u\|_{s+m+4}\big)\|h\|_0. \tag{C.11}$$

The estimate for $\mathcal{R}_1$ follows. $\mathcal{R}_2$ satisfies the same estimate as $\mathcal{R}_1$ because $\mathcal{R}_2 = \mathcal{M}^{-1}\mathcal{R}_1$. For $\mathcal{R}_3$, note that $\Pi_C \mathcal{L}_2 = \Pi_C(a_9 + \mathcal{R}_2)$. Use (7.27) for $\mathcal{M}^{-1}$, then the same arguments as for (C.11). $\square$

Formula for $\mathcal{R}_4$.

$$\begin{aligned}
\mathcal{R}_4 = {} & \mathcal{R}_3 \mathbb{P}\Phi - a_9 \Pi_C \Phi \\
& + \sum_{k=0}^{3} \big\{ \Pi_E^\perp \mu_2 \big(\beta_{yy}^{(k)} \partial_y^{-k} + 2\beta_y^{(k)} \partial_y^{-k+1} + \beta^{(k)} \partial_y^{-k+2}\big) + a_6 \Pi_E^\perp \big(\beta_y^{(k)} \partial_y^{-k} + \beta^{(k)} \partial_y^{-k+1}\big) \\
& + a_8 \Pi_E^\perp \beta^{(k)} \partial_y^{-k} - \big(\mu_2 \beta^{(k)} \partial_y^{-k+2} + \mu_0 \beta^{(k)} \partial_y^{-k} + \mu_{-2} \beta^{(k)} \partial_y^{-k-2}\big) \Pi_E^\perp \big\} \\
& + \big(-\mathcal{H}\big(2\mu_2 \alpha_y^{(1)} + a_6 \alpha^{(1)}\big) - (a_7 - \mu_1)\alpha^{(1)}\big)\Pi_E^\perp + \big(\big(2\mu_2 \beta_y^{(1)} + a_6 \beta^{(1)}\big) - \mathcal{H}(a_7 - \mu_1)\beta^{(1)}\big)\Pi_E^\perp \\
& + \sum_{k=0}^{3} \big\{ [a_6, \mathcal{H}]\big(\alpha_y^{(k)} \partial_y^{-k} + \alpha^{(k)} \partial_y^{-k+1}\big) + [a_7, \mathcal{H}]\big(\beta_y^{(k)} \partial_y^{-k} + \beta^{(k)} \partial_y^{-k+1}\big) + [a_8, \mathcal{H}]\alpha^{(k)} \partial_y^{-k} \\
& + [a_9, \mathcal{H}]\beta^{(k)} \partial_y^{-k} \big\} + \sum_{k=0}^{3} \big[\beta^{(k)} - \alpha^{(k)}, \mathcal{H}\big]\big(\mu_2 \partial_y^{-k+2} + \mu_0 \partial_y^{-k} + \mu_{-2} \partial_y^{-k-2}\big) \\
& + \left(\omega \alpha_\tau^{(3)} - \mu_2 \beta_{yy}^{(3)} - a_6 \beta_y^{(3)} + a_7 \alpha_y^{(3)} - (a_8 - \mu_0)\beta^{(3)} + a_9 \alpha^{(3)} + \mu_{-2} \sum_{k=1}^{3} \beta^{(k)} \partial_y^{-k-2}\right) \partial_y^{-3} \\
& + \mathcal{H}\left(\omega \beta_\tau^{(3)} + \mu_2 \alpha_{yy}^{(3)} + a_6 \alpha_y^{(3)} + a_7 \beta_y^{(3)} + (a_8 - \mu_0)\alpha^{(3)} + a_9 \beta^{(3)} - \mu_{-2} \sum_{k=1}^{3} \alpha^{(k)} \partial_y^{-k-2}\right) \partial_y^{-3}.
\end{aligned}$$

Proof of Proposition 7.4. From the estimates for $\mu_2, \mu_1, a_6, a_7, a_8, a_9$ of Proposition 7.2 and formulae (7.52), (7.53) for φ it follows that

$$\big\|\mathrm{Re}(\varphi)\big\|_s + \big\|\mathrm{Im}(\varphi)\big\|_s \leqslant \varepsilon^2 C(s,K)\big(1 + \|u\|_{s+c}\big), \tag{C.12}$$

$$\big\|\partial_u \mathrm{Re}(\varphi)[h]\big\|_s + \big\|\partial_u \mathrm{Im}(\varphi)[h]\big\|_s \leqslant \varepsilon^4 C(s,K)\big(\|h\|_{s+c} + \|u\|_{s+c}\|h\|_4\big), \tag{C.13}$$

$$\big\|\partial_\varepsilon \mathrm{Re}(\varphi)\big\|_s + \big\|\partial_\varepsilon \mathrm{Im}(\varphi)\big\|_s \leqslant \varepsilon C(s,K)\big(1 + \|u\|_{s+c}\big), \tag{C.14}$$

for $2 \leqslant s \leqslant r - 1$, where $c = 6$ (in this proof we use (B.3) to estimate any product). As a consequence, by Lemma B.3 and (7.54), $\alpha^{(0)} - 1$ and $\beta^{(0)}$ and their derivatives satisfy the same estimates (C.12), (C.13), (C.14), with $c = 6$.

$g^{(0)}$ is given by (7.41), therefore its real and imaginary part satisfy (C.12), (C.13), (C.14), with $c = 8$, for $2 \leqslant s \leqslant r - 3$. The same for $\eta^{(1)}$ because of (7.43), (7.46). By formulae (7.44), (7.47), (7.50), (7.48), (7.51), the same holds for $g^{(1)}, \eta^{(2)}$, with $c = 10$, $2 \leqslant s \leqslant r - 5$, and for $g^{(2)}, \eta^{(3)}$, with $c = 12$, $2 \leqslant s \leqslant r - 7$. Since $f^{(k)} = \eta^{(k)} f^{(0)}$, $k = 1, 2, 3$, all coefficients $\alpha^{(k)}, \beta^{(k)}$, $k = 1, 2, 3$ and their derivatives satisfy (C.12), (C.13), (C.14), with $c = 12$, for all $2 \leqslant s \leqslant r - 7$. By (B.3),

$$\left\|(\Phi - I)f\right\|_s \leqslant C\|\mathrm{coeff}\|_2\|f\|_s + C(s)\|\mathrm{coeff}\|_s\|f\|_2,$$

where 'coeff' are $(\alpha^{(0)} - 1), \beta^{(0)}, \alpha^{(k)}, \beta^{(k)}$, $k = 1, 2, 3$, and C does not depend on s. Therefore

$$\left\|(\Phi - I)f\right\|_s \leqslant \varepsilon^2 C(K)\|f\|_s + \varepsilon^2 C(s, K)\big(1 + \|u\|_{s+12}\big)\|f\|_2.$$

The estimates for $\partial_u \Phi[h]$ and $\partial_\varepsilon \Phi$ are obtained in the same way, using the estimates for the derivatives of the coefficients. Similarly, (7.64), (7.65) follow because $\partial_\tau(\Phi - I)f = (\Phi - I)\partial_\tau f + \Phi_\tau f$, where Φ_τ is the operator of the same type as Φ that has coefficients $\alpha_\tau^{(k)}, \beta_\tau^{(k)}$ instead of $\alpha^{(k)}, \beta^{(k)}$, $k = 0, \dots, 3$. Since $\|\mathbb{P}f\|_s \leqslant \|f\|_s$, all the estimate for $\Phi - I$ also hold for $\tilde\Phi - \mathbb{P} = \mathbb{P}(\Phi - I)\mathbb{P}$. (7.61), (7.62) and (7.63) also hold for $\tilde\Phi^{-1}$ by Lemma B.2.

To prove (7.66) for $\tilde\Phi^{-1}\tilde{\mathcal{M}}^{-1}\tilde\Psi^{-1}$, write

$$\tilde\Phi^{-1}\tilde{\mathcal{M}}^{-1}\tilde\Psi^{-1} = I + S, \quad S := \big(\tilde\Psi^{-1} - I\big) + \big(\tilde{\mathcal{M}}^{-1} - I\big)\tilde\Psi^{-1} + \big(\tilde\Phi^{-1} - I\big)\tilde{\mathcal{M}}^{-1}\tilde\Psi^{-1},$$

then apply (7.22), (7.21), (7.27) and (7.61). Similarly for the other operators.

The estimates for μ_0, μ_{-2} and their derivatives follow from formulae (7.55), (7.56) and the estimates for $\mu_2, a_6, a_7, a_8, a_9, \eta^{(2)}, g^{(0)}$.

Now study the rest $\mathcal{R}$. By (7.34), for $2 \leqslant s \leqslant r - 6$,

$$\left\|\mathcal{R}_3\partial_y^m f\right\|_s \leqslant \varepsilon^3 C(s, K)\big(\|f\|_s + \|f\|_0\|u\|_{s+10}\big), \quad 0 \leqslant m \leqslant 3. \tag{C.15}$$

By definition, Φ is a combination of multiplications and $\mathcal{H}, \partial_y^{-1}$. Every ∂_y can be moved from the right to the left of any multiplication operator with elementary calculus: $[a, \partial_y] = -a_y$, namely, for every a, f,

$$a\partial_y f = \partial_y(af) - a_y f, \qquad a\partial_y^2 f = \partial_y^2(af) - 2\partial_y(a_y f) + a_{yy} f,$$
$$a\partial_y^3 f = \partial_y^3(af) - 3\partial_y^2(a_y h) + 3\partial_y(a_{yy} f) - a_{yyy} f.$$

Recall that the coefficients $\alpha^{(k)}, \beta^{(k)}$ satisfy (C.12), (C.13), (C.14), with $c = 12$, $2 \leqslant s \leqslant r - 7$. Moving ∂_y^m to the left of Φ, $m = 0, 1, 2, 3$, the coefficients $\alpha^{(k)}, \beta^{(k)}$ are subject to up to 3 derivatives in y. So applying (C.15) gives

$$\left\|\mathcal{R}_3\mathbb{P}\Phi\partial_y^m f\right\|_s \leqslant \varepsilon^5 C(s, K)\big(\|f\|_s + \|u\|_{s+10}\|f\|_2\big), \quad 0 \leqslant m \leqslant 3,\ 2 \leqslant s \leqslant r - 10.$$

Each term $R_{(a)}$ of type (a) containing $[b, \mathcal{H}]$ can be estimated by Lemma B.5(i), whence

$$\left\|R_{(a)}\partial_y^m f\right\|_s \leqslant \varepsilon^2 C(s, K)\big(\|f\|_s + \|u\|_{s+17}\|f\|_2\big), \quad 0 \leqslant m \leqslant 3,\ 2 \leqslant s \leqslant r - 12,$$

and the same inequality also holds for each term $R_{(b)}$ of type (b) that contains $\Pi_E^\perp$. Thus it holds for $\|\mathcal{R}_4\partial_y^m f\|_s$. Since $\mathcal{R} := \tilde\Phi^{-1}\mathbb{P}\mathcal{R}_4$ by (7.57), the estimate for $\mathcal{R}\partial_y^m$ follows from (7.61). □

Proof of (9.6). (The meaning of A, B, a, b, c in the following proof is independent on the rest of the paper.) By (9.4),

$$\begin{aligned} F(u_n) + F'(u_n)h_{n+1} &= F(u_n) + P_\varepsilon^{-1}\tilde\Psi_n\tilde{\mathcal{M}}_n\tilde\Phi_n\tilde{\mathcal{L}}_4(u_n)\tilde\Phi_n^{-1}\tilde\Psi_n^{-1}h_{n+1} \\ &= P_\varepsilon^{-1}\tilde\Psi_n\tilde{\mathcal{M}}_n\tilde\Phi_n\big\{\tilde\Phi_n^{-1}\tilde{\mathcal{M}}_n^{-1}\tilde\Psi_n^{-1}P_\varepsilon F(u_n) + \tilde{\mathcal{L}}_4(u_n)\tilde\Phi_n^{-1}\tilde\Psi_n^{-1}h_{n+1}\big\}. \end{aligned} \tag{C.16}$$

Let $p = \{\dots\}$ be the quantity in parentheses in (C.16). Let

$$c := \tilde\Phi_n^{-1}\tilde{\mathcal{M}}_n^{-1}\tilde\Psi_n^{-1}P_\varepsilon F(u_n) = \Pi_{n+1}c + \Pi_{n+1}^\perp c,$$
$$\tilde{\mathcal{L}}_4(u_n) = A + B, \quad A := \Pi_{n+1}\tilde{\mathcal{L}}_4(u_n)\Pi_{n+1},\ B := \Pi_{n+1}^\perp\tilde{\mathcal{L}}_4(u_n)\Pi_{n+1} + \tilde{\mathcal{L}}_4(u_n)\Pi_{n+1}^\perp.$$

With these abbreviations, by the definition (9.5) $h_{n+1} = -\Pi_{n+1}\tilde\Psi_n\tilde\Phi_n A^{-1}\Pi_{n+1}c$, whence

$$\tilde\Phi_n^{-1}\tilde\Psi_n^{-1}h_{n+1} = a + b, \quad a := -A^{-1}\Pi_{n+1}c,\ b := \tilde\Phi_n^{-1}\tilde\Psi_n^{-1}\Pi_{n+1}^\perp\tilde\Psi_n\tilde\Phi_n A^{-1}\Pi_{n+1}c.$$

Now $p = c + (A + B)(a + b)$, and $Aa + \Pi_{n+1}c = 0$. Therefore

$$p = \Pi_{n+1}^\perp c + Ba + (A + B)b.$$

$\Pi_{n+1}^{\perp}\tilde{\mathcal{L}}_4(u_n)\Pi_{n+1} = \Pi_{n+1}^{\perp}\tilde{\mathcal{R}}\Pi_{n+1}$ because $\tilde{\mathcal{L}}_4(u_n) = \tilde{\mathcal{D}} + \tilde{\mathcal{R}}$ and $\tilde{\mathcal{D}}$ is diagonal. Moreover $\tilde{\mathcal{L}}_4(u_n)\Pi_{n+1}^{\perp}a = 0$ because $a \in Z_n$. Thus (9.6) follows. □

Proof of Lemma 8.5. (i) Lemma (8.5) simply follows from Lemma B.3. In particular, $\bar{v}_2(\varepsilon)$ satisfies (4.2). By Proposition 5.3, $(\Pi_V A \Pi_V) : V \cap X \to V \cap Y$, $h \mapsto 3\partial_t h + \Pi_V \partial_x(3\bar{v}_1^2 h)$ is invertible, with

$$\left\| (\Pi_V A \Pi_V)^{-1} h \right\|_s \leqslant C \|h\|_{s-1} \quad \forall h \in V \cap Y,\ s \geqslant 1, \tag{C.17}$$

where C depends only on the set $\mathcal{K}$, like in (8.6). By (1.5) and (B.11), $\|\mathcal{N}_4(h)\|_s \leqslant C(s)\|h\|_4^3 \|h\|_{s+2}$ for $0 \leqslant s \leqslant r$. Hence

$$\left\| \bar{v}_2(\varepsilon) \right\|_s \leqslant C\varepsilon^{-4} \left\| \mathcal{N}_4(\varepsilon \bar{v}_1) \right\|_{s-1} \leqslant C(s) \|\bar{v}_1\|_4^3 \|\bar{v}_1\|_{s+1} = C'(s) \tag{C.18}$$

where $C'(s)$ depends on s and $\|\bar{v}_1\|_{s+1}$. (C.18) for $s = 4$ implies that $\varepsilon\|\bar{v}_1\|_4 + \varepsilon^2 \|\bar{v}_2\|_4 < \delta_0$ for all $\varepsilon < \varepsilon_0$, for some ε_0 depending on $\|\bar{v}_1\|_5$.

To complete the proof of (8.19), differentiate (4.2) with respect to ε, then use (C.17),

$$\left\| \partial_\varepsilon \bar{v}_2(\varepsilon) \right\|_s \leqslant C\left(4\varepsilon^{-5} \left\| \Pi_V \mathcal{N}_4(\varepsilon \bar{v}_1) \right\|_{s-1} + \varepsilon^{-4} \left\| \Pi_V \mathcal{N}_4'(\varepsilon \bar{v}_1)[\bar{v}_1] \right\|_{s-1}\right) \leqslant \varepsilon^{-1} C(s).$$

(8.20) follows from formula (4.3) and estimates (8.19). To prove (ii), observe that

$$\begin{aligned} Q(u,h,\varepsilon) = \varepsilon^{-2} P_\varepsilon^{-1}\Big(&\partial_x\big\{3(\varepsilon\bar{v}_1 + \varepsilon^2 u)(\varepsilon^2 h)^2 + (\varepsilon^2 h)^3\big\} + \mathcal{N}_4(\varepsilon\bar{v}_1 + \varepsilon^2 u + \varepsilon^2 h) - \mathcal{N}_4(\varepsilon\bar{v}_1 + \varepsilon^2 u) \\ &- \mathcal{N}_4'(\varepsilon\bar{v}_1 + \varepsilon^2 u)[\varepsilon^2 h]\Big), \end{aligned}$$

then apply (B.11) to $\mathcal{N}_4$.

(iii) follows from (4.5) by the usual tame estimates. □

References

[1] M. Ablowitz, A. Fokas, The inverse scattering transform for the Benjamin–Ono equation, a pivot for multidimensional problems, Stud. Appl. Math. 68 (1983) 1–10.

[2] D.M. Ambrose, J. Wilkening, Computation of time-periodic solutions of the Benjamin–Ono equation, J. Nonlinear Sci. 20 (3) (2010) 277–308.

[3] P. Baldi, Periodic solutions of forced Kirchhoff equations, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 8 (2009) 117–141.

[4] D. Bambusi, S. Graffi, Time quasi-periodic unbounded perturbations of Schrödinger operators and KAM methods, Commun. Math. Phys. 219 (2001) 465–480.

[5] T. Benjamin, Internal waves of permanent form in fluids of great depth, J. Fluid Mech. 29 (1967) 559–562.

[6] M. Berti, Nonlinear Oscillations of Hamiltonian PDEs, Progress in Nonlin. Diff. Eq. Appl., vol. 74, Birkhäuser, Boston, 2008, pp. 1–181.

[7] M. Berti, L. Biasco, M. Procesi, KAM theory for the Hamiltonian derivative wave equation, preprint, arXiv:1111.3905, 2011.

[8] M. Berti, Ph. Bolle, Cantor families of periodic solutions of wave equations with C^k nonlinearities, NoDEA 15 (2008) 247–276.

[9] M. Berti, Ph. Bolle, M. Procesi, An abstract Nash–Moser theorem with parameters and applications to PDEs, Ann. Institute H. Poincaré (C) Anal. Non Linéaire 27 (2010) 377–399.

[10] J. Bourgain, Periodic solutions of nonlinear wave equations, in: Harmonic Analysis and Partial Differential Equations, in: Chicago Lectures in Math., Univ. Chicago Press, Chicago, IL, 1999, pp. 69–97.

[11] N. Burq, F. Planchon, On well-posedness for the Benjamin–Ono equation, Math. Ann. 340 (3) (2005) 497–542.

[12] J. Colliander, C. Kenig, G. Staffilani, Local well-posedness for dispersion generalized Benjamin–Ono equations, Diff. Integral Eq. Internat. J. Theory Appl. 16 (12) (2003) 1441–1472.

[13] W. Craig, Problèmes de petits diviseurs dans les équations aux dérivées partielles, Panoramas et Synthèses, vol. 9, Société Mathématique de France, Paris, 2000.

[14] J.-M. Delort, A quasi-linear Birkhoff normal forms method. Application to the quasi-linear Klein–Gordon equation on $\mathbb{S}^1$, Astérisque 341 (2012).

[15] R.S. Hamilton, The inverse function theorem of Nash and Moser, Bull. Amer. Math. Soc. 7 (1) (1982) 65–222.

[16] L. Hörmander, The boundary problems of physical geodesy, Arch. Ration. Mech. Anal. 62 (1) (1976) 1–52.

[17] H. Inci, T. Kappeler, P. Topalov, On the regularity of the composition of diffeomorphisms, Mem. Amer. Math. Soc., in press.

[18] A.D. Ionescu, C.E. Kenig, Global well-posedness of the Benjamin–Ono equation in low-regularity spaces, J. Amer. Math. Soc. 20 (2007) 753–798.

[19] G. Iooss, P.I. Plotnikov, Multimodal standing gravity waves: a completely resonant system, J. Math. Fluid Mech. 7 (2005) 110–126.

[20] G. Iooss, P.I. Plotnikov, Existence of multimodal standing gravity waves, J. Math. Fluid Mech. 7 (2005) 349–364.

[21] G. Iooss, P.I. Plotnikov, J.F. Toland, Standing waves on an infinitely deep perfect fluid under gravity, Arch. Ration. Mech. Anal. 177 (3) (2005) 367–478.

[22] T. Kappeler, J. Pöschel, KdV & KAM, Springer-Verlag, 2003.

[23] S. Klainerman, A. Majda, Formation of singularities for wave equations including the nonlinear vibrating string, Comm. Pure Appl. Math. 33 (1980) 241–263.
[24] P. Koosis, Introduction to H^p Spaces, Cambridge University Press, 1998.
[25] S. Kuksin, Analysis of Hamiltonian PDEs, Oxford Lectures Ser. Math. Appl., vol. 19, Oxford University Press, 2000.
[26] J. Liu, X. Yuan, A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations, Comm. Math. Phys. 307 (3) (2011) 629–673.
[27] L. Molinet, Global well-posedness in the energy space for the Benjamin–Ono equation on the circle, Math. Ann. 337 (2) (2007) 353–383.
[28] L. Molinet, Global well-posedness in L^2 for the periodic Benjamin–Ono equation, Amer. J. Math. 130 (3) (2008) 635–683.
[29] L. Molinet, J.C. Saut, N. Tzvetkov, Ill-posedness issues for the Benjamin–Ono and related equations, SIAM J. Math. Anal. 33 (2001) 982–988.
[30] J. Moser, Convergent series expansions for quasi-periodic motions, Math. Ann. 169 (1967) 136–176.
[31] J. Moser, A rapidly convergent iteration method and non-linear partial differential equations – I, Ann. Sc. Norm. Super. Pisa Cl. Sci. III Ser. 20 (2) (1966) 265–315.
[32] P.I. Plotnikov, J.F. Toland, Nash–Moser theory for standing water waves, Arch. Ration. Mech. Anal. 159 (2001) 1–83.
[33] P.H. Rabinowitz, Periodic solutions of nonlinear hyperbolic partial differential equations, Comm. Pure Appl. Math. 20 (1967) 145–205.
[34] P.H. Rabinowitz, Periodic solutions of nonlinear hyperbolic partial differential equations. II, Comm. Pure Appl. Math. 22 (1969) 15–39.
[35] T. Tao, Global well-posedness of the Benjamin–Ono equation in $H^1(\mathbb{R})$, J. Hyperbolic Diff. Eq. 1 (2004) 27–49.
[36] J. Zhang, M. Gao, X. Yuan, KAM tori for reversible partial differential equations, Nonlinearity 24 (4) (2011) 1189–1229.