ANALYSIS & PDE

Volume 16 No. 9 2023

ÁNGEL CASTRO, DIEGO CÓRDOBA AND FAN ZHENG

STABILITY OF TRAVELING WAVES FOR THE BURGERS–HILBERT EQUATION

msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 9, pp. 2109–2145
DOI: 10.2140/apde.2023.16.2109

STABILITY OF TRAVELING WAVES FOR THE BURGERS–HILBERT EQUATION

ÁNGEL CASTRO, DIEGO CÓRDOBA AND FAN ZHENG

We consider smooth solutions of the Burgers–Hilbert equation that are a small perturbation δ from a global periodic traveling wave with small amplitude ϵ. We use a modified energy method to prove the existence time of smooth solutions on a time scale of $1/(\epsilon\delta)$, with $0<\delta\ll\epsilon\ll 1$, and on a time scale of ϵ/δ^2, with $0<\delta\ll\epsilon^2\ll 1$. Moreover, we show that the traveling wave exists for an amplitude ϵ in the range $(0,\epsilon^*)$, with $\epsilon^*\sim 0.23$, and fails to exist for $\epsilon>2/e$.

1. Introduction

1A. ***The Burgers–Hilbert equation (BH).*** We study the size and stability of traveling waves of the Burgers–Hilbert equation (BH),

$$f_t = Hf + ff_x \quad \text{for } (x,t)\in\Omega\times\mathbb{R}, \tag{1-1}$$
$$f(x,0)=f_0(x), \tag{1-2}$$

where Ω is the real line $\mathbb{R}$ or the torus $\mathbb{T}=\mathbb{R}/2\pi\mathbb{Z}$ and Hf is the Hilbert transform which is defined for $f:\mathbb{R}$ (resp. $\mathbb{T}$) $\to\mathbb{R}$ by

$$Hf(x)=\frac{1}{\pi}\,\text{P. V.}\int_{\mathbb{R}}\frac{f(y)}{x-y}\,dy \quad \text{resp. } Hf(x)=\frac{1}{2\pi}\,\text{P. V.}\int_0^{2\pi} f(y)\cot\frac{x-y}{2}\,dy.$$

Its action in the frequency space is $\widehat{Hf}(k)=-i\,\operatorname{sgn}k\,\hat{f}(k)$ for $k\neq 0$, and $\widehat{Hf}(0)=0$.

This equation arose in [Marsden and Weinstein 1983] as a quadratic approximation for the evolution of the boundary of a simply connected vorticity patch in two dimensions. Later, Biello and Hunter [2010] proposed the model as an approximation for describing the dynamics of small slope vorticity fronts in the 2-dimensional incompressible Euler equations. Recently, the validity of this approximation was proved in [Hunter et al. 2022].

MSC2020: 35F25, 76B47.
Keywords: Burgers–Hilbert, normal forms, traveling waves.

By standard energy estimates the initial value problem for (BH) is locally well-posed in H^s for $s > \frac{3}{2}$. Bressan and Nguyen [2014] established in global existence of weak solutions for initial data $f_0 \in L^2(\mathbb{R})$, with $f(x,t) \in L^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$ for all $t > 0$. Bressan and Zhang [2017] constructed locally in time piecewise continuous solutions to the BH equation with a single discontinuity where the Hilbert transform generates a logarithmic singularity. Uniqueness for general global weak solutions of [Bressan and Nguyen 2014] is open. But piecewise continuous solutions are shown to be unique in [Krupa and Vasseur 2020].

The Burgers–Hilbert equation can indeed form shocks in finite time. Various numerical simulations have been performed in [Biello and Hunter 2010; Hunter 2018; Klein and Saut 2015]. Finite time singularities, in the $C^{1,\delta}$ norm, with $0 < \delta < 1$, were shown to exist in [Castro et al. 2010] for initial data f_0 in $L^2(\mathbb{R}) \cap C^{1,\delta}(\mathbb{R})$ that has a point $x_0 \in \mathbb{R}$ such that $H(f_0)(x_0) > 0$ and $f_0(x_0) \geq (32\pi \|f_0\|_{L^2})^{1/3}$. Recently, with a different approach, Saut and Wang [2022] proved shock formation in finite time for (BH) and Yang [2021] constructed solutions that develop an asymptotic self-similar shock at one single point with an explicitly computable blowup profile for (BH).

In this paper we are concerned with the dynamics in the small amplitude regime where (BH) can be viewed as a perturbation of the linearized (BH) equation $f_t = H[f]$. Since the nonlinear term in (1-1) is quadratic and the Hilbert transform is orthogonal in L^2, standard energy estimates yield a time of existence of smooth solutions $T \sim 1/\|f_0\|$. Thanks to the effect of the Hilbert transform and using the normal form method, Hunter, Ifrim, Tataru and Wong (see [Hunter and Ifrim 2012; Hunter et al. 2015]) were able to improve this time of existence. More precisely, if ϵ is the size of the initial data, they prove a lifespan $T \sim 1/\epsilon^2$ for small enough ϵ (see also [Ehrnström and Wang 2019] for a similar approach with a modified version of the (BH) equation). The proofs are based on the normal form method and on the modified energy method. Furthermore, Hunter [2018] showed for $0 < \epsilon \ll 1$ the existence of C^∞-traveling wave solutions of the form

$$f_\epsilon(x,t) = u_\epsilon(x + v_\epsilon t),$$

with

$$u_\epsilon(x) = \epsilon \cos(x) + O(\epsilon^2), \tag{1-3}$$

$$v_\epsilon = -1 + O(\epsilon^2). \tag{1-4}$$

Notice that, $(u_\epsilon(nx)/n, v_\epsilon/n)$ is also a C^∞-traveling wave solution.

Throughout the paper we will assume that the initial data f_0 has zero mean. Since (1-1) preserves the mean,

$$\int_0^{2\pi} f(x,t)\,dx = 0 \quad \text{for all } t.$$

Since in the construction above u_ϵ also has zero mean,

$$\int_0^{2\pi} f(x,t)\,dx = 0 \quad \text{for all } t.$$

1B. ***The main theorem.*** In the present work we extend the results in the small amplitude regime in the following way:

(1) Size of the traveling waves: We show that the traveling waves exist for an amplitude ϵ in the range $(0, \epsilon^*)$, with $\epsilon^* \sim 0.23$, and fail to exist for $\epsilon > 2/e$.

(2) Extended lifespan from a traveling wave: We prove that a δ-perturbation of u_ϵ lives, at least, for a time $T \sim 1/(\delta\epsilon)$ for $0 < \delta \ll \epsilon \ll 1$, and for a time $T \sim \epsilon/\delta^2$ for $0 < \delta \ll \epsilon^2 \ll 1$. This is an improvement compared with the time $T \sim 1/\epsilon^2$ provided by the results in [Hunter and Ifrim 2012; Hunter et al. 2015]. Indeed, our main theorem reads:

Theorem 1.1. *For $0 < |\epsilon|,\ \delta \ll 1$ let $(u_\epsilon, v_\epsilon) \in C^\infty(\mathbb{T}) \times \mathbb{R}$ be a traveling wave solution of (1-1) as in (1-3) and (1-4) and*

$$\|f_0 - u_\epsilon\|_{H^4(\mathbb{T})} < \delta.$$

Then there exist $0 < \epsilon_0 \ll 1,\ T(\epsilon, \delta) > 0$ and a solution of (1-1)

$$f(x, t) \in C([0, T(\epsilon, \delta)); H^4(\mathbb{T}))$$

such that

(1) *if $\delta \ll |\epsilon|$ and $|\epsilon| \le \epsilon_0$, then $T(\epsilon, \delta) \sim 1/(\epsilon\delta)$,*

(2) *if $\delta \ll \epsilon^2$ and $|\epsilon| \le \epsilon_0$, then $T(\epsilon, \delta) \sim \epsilon/\delta^2$.*

Moreover, there are two differentiable functions $\epsilon(t)$ and $a(t)$ such that

$$\|f(x, t) - u_{\epsilon(t)}(x + a(t))\|_{H^4} \lesssim \delta.$$

1C. ***Sketch of the proof of Theorem 1.1.*** Now we briefly describe the proof of Theorem 1.1. Assume that the solution

$$f(x, t) = u_\epsilon(x + v_\epsilon t) + g(x + v_\epsilon t, t)$$

is a small perturbation around the traveling wave $u_\epsilon(x+v_\epsilon t)$. Then the linearization of the Burgers–Hilbert equation (1-1) is

$$L_\epsilon g := -v_\epsilon g_x + Hg + (u_\epsilon(x)g)_x = 0$$

so to the first order, the perturbation g solves the equation $g_t = L_\epsilon g$, with solution

$$g(x, t) = e^{tL_\epsilon} g(x, 0).$$

Therefore the linear evolution of g is determined by the eigenvalues of L_ϵ.

The full nonlinear evolution of g is

$$g_t = L_\epsilon g + N(g, g),$$

where $N(g, g)$ is a nonlinearity that is (at least) quadratic in g. We plug in the linear solution to get

$$g_t = e^{tL_\epsilon} L_\epsilon g(x, 0) + N(e^{tL_\epsilon} g(x, 0), e^{tL_\epsilon} g(x, 0))$$

to second order, which integrates to

$$g(x, t) = e^{tL_\epsilon} g(x, 0) + e^{tL_\epsilon} \int_0^t e^{-sL_\epsilon} N(e^{sL_\epsilon} g(x, 0), e^{sL_\epsilon} g(x, 0))\, ds.$$

Expand (at least formally) the initial data and the nonlinearity in terms of the eigenvectors of L_ϵ as

$$g(x,0)=\sum_n c_n\varphi_n(x),\qquad N(\varphi_l,\varphi_m)=\sum_n c_{lmn}\varphi_n,$$

where the eigenvalue of φ_n is λ_n. Then

$$g(x,t)\approx\sum_n c_n e^{\lambda_n t}\varphi_n(x)+\sum_{l,m,n}\frac{e^{(\lambda_l+\lambda_m)t}-e^{\lambda_n t}}{\lambda_l+\lambda_m-\lambda_n}c_{lmn}\varphi_n(x)\tag{1-5}$$

to second order, provided that the denominator $\lambda_l+\lambda_m-\lambda_n$ is not equal to 0, i.e., that the eigenvalues are "nonresonant". Then we can integrate (1-1) up to a cubic error term, yielding the "cubic lifespan", i.e., initial data of size ϵ leads to a solution that exists for a time at least comparable to ϵ^{-2}. This is the "normal form transformation", first proposed by Poincaré in the setting of ordinary differential equations (see [Arnold 1983] for a book reference). Its application to partial differential equations was initiated by Shatah [1985] in the study of the nonlinear Klein–Gordon equation, and then extended to the water wave problem by Germain, Masmoudi and Shatah [Germain et al. 2012; 2015] and Ionescu and Pusateri [2015; 2018], the Burgers–Hilbert equation by Hunter, Ifrim, Tataru and Wang [Hunter et al. 2015], and more recently, the Einstein–Klein–Gordon equation by Ionescu and Pausader [2022].

Unfortunately, nonresonance fails for L_ϵ because 0 is an eigenvalue, and $0+\lambda_n-\lambda_n=0$. The eigenvalue 0 arises from the symmetry of (1-1). Indeed, the initial data $u_\epsilon(x+\delta)\approx u_\epsilon(x)+\delta u_\epsilon'(x)$ produces the solution

$$f(x,t)=u_\epsilon(x+v_\epsilon t+\delta)\approx u_\epsilon(x+v_\epsilon t)+\delta u_\epsilon'(x+v_\epsilon t).$$

In this case $g(x,t)=\delta u_\epsilon'(x)$, with $g_t=0$, so $u_\epsilon'\in\ker L_\epsilon$. Also, the initial data $u_{\epsilon+\delta}(x)\approx u_\epsilon(x)+\delta\partial_\epsilon u_\epsilon(x)$ produces the solution

$$f(x,t)=u_{\epsilon+\delta}(x+v_{\epsilon+\delta}t)\approx u_\epsilon(x+v_\epsilon t)+\delta\partial_\epsilon u_\epsilon(x+v_\epsilon t)+\delta v_\epsilon' t u_\epsilon'(x+v_\epsilon t).$$

In this case $g(x,t)=\delta\partial_\epsilon u_\epsilon(x)+\delta v_\epsilon' t u_\epsilon'(x)$, so

$$L_\epsilon g=\delta L_\epsilon\partial_\epsilon u_\epsilon=g_t=\delta v_\epsilon' u_\epsilon'\in\ker L_\epsilon,$$

and thus $\partial_\epsilon u_\epsilon$ is in the generalized eigenspace corresponding to the eigenvalue 0.

These perturbations generate translations and variations along the bifurcation curve. We treat them separately using a more sophisticated ansatz

$$f(x,t)=u_{\epsilon(t)}(x+a(t))+g(x+a(t),t).$$

We will show in Proposition 4.1 that if $|\epsilon_0|$ and $\|f-u_{\epsilon_0}\|_{H^2}/|\epsilon_0|$ are sufficiently small, then f can always be put in the form above, with $|\epsilon-\epsilon_0|/|\epsilon_0|$ also small and the expansion of g not involving any eigenvector with eigenvalue 0. This way we remove the resonance caused by the eigenvalue 0 from the evolution of g.

We also need to analyze the other eigenvalues of L_ϵ, a first-order differential operator with variable coefficients, and a quasilinear perturbation from $L_0=\partial_x+H$, whose eigenvectors are the Fourier modes e^{inx}. Just like the Schrödinger operator with potential $-\Delta+V$, with a basis of eigenvectors known as the "Jost functions”, giving rise to the “distorted Fourier transform” (see [Agmon 1975]), L_ϵ can also be diagonalized using a combination of conjugation and perturbative analysis. More precisely, let $g = h_x$. Then

$$L_\epsilon g = ((u_\epsilon(x) - v_\epsilon)g)_x + Hg = ((u_\epsilon(x) - v_\epsilon)h_x + Hh)_x,$$

so L_ϵ is conjugate to the operator $h \mapsto (u_\epsilon(x) - v_\epsilon)h_x + Hh$. Let $h = \tilde{h} \circ \phi_\epsilon$, where $\phi_\epsilon'(x)$ is proportional to $(u_\epsilon(x) - v_\epsilon)^{-1}$. Then

$$L_\epsilon g = ((c_\epsilon \partial_x + H + R_\epsilon)\tilde{h} \circ \phi_\epsilon)_x,$$

where $c_\epsilon \to 1$ as $\epsilon \to 0$, and R_ϵ is a small smoothing remainder (i.e., it gains derivatives of arbitrarily high orders). Thus L_ϵ is conjugate to $c_\epsilon \partial_x + H + R_\epsilon$, whose eigenvalues can be approximated by those of $c_\epsilon \partial_x + H$, which are $\pm(nc_\epsilon i - i)$, $n = 1, 2, \ldots$. The general theory of unbounded analytic operators developed in [Kato 1976] allows us to justify this approximation up to $O(\epsilon^6)$ (see Corollary 3.10), and to relate the eigenvectors of L_ϵ to the Fourier modes (see Lemma 3.7), in the sense that another linear map $\tilde{h} \mapsto \mathfrak{h}$ conjugates L_ϵ into a Fourier multiplier whose action on $e^{i(n+\operatorname{sgn} n)x}$ is multiplication by λ_n ($n \neq 0$).

At the end of the day we have the following estimate for small ϵ:

$$|\lambda_l + \lambda_m - \lambda_n| > \begin{cases} \frac{1}{2}, & l + m \neq n, \\ \frac{1}{5}\epsilon^2, & l + m = n; \end{cases}$$

see Proposition 3.11. Because this value appears in the denominator in (1-5), if g has size δ, a direct application of the normal form transformation yields a lifespan comparable to ϵ^2/δ^2. To improve on this, we will make use of the structure of the nonlinearity:

$$N(\mathfrak{h}, \mathfrak{h}) = \tfrac{1}{2}\mathfrak{h}_x^2 + O(|\epsilon|).$$

The first term is the usual product-style nonlinearity, which imposes the restriction $l + \operatorname{sgn} l + m + \operatorname{sgn} m = n + \operatorname{sgn} n$, and implies $l + m - n = \pm 1 \neq 0$, so the normal form transformation can be carried out as before. The second term is of size $|\epsilon|$ and gains a factor of $1/|\epsilon|$ in the lifespan. Thus the usual energy estimate can show a lifespan comparable to $1/|\epsilon\delta|$, and the normal form transformation can show a lifespan comparable to $|\epsilon|/\delta^2$. This decomposition of the nonlinearity into one part satisfying classical additive frequency restrictions and another part enjoying better estimates analytically was first used by Germain, Pusateri and Rousset [Germain et al. 2018] to show global well-posedness of the 1-dimensional Schrödinger equation with potential (see also [Chen and Pusateri 2022]). Our result shows that this approach can be adapted to quasilinear equations and to the case of discrete spectrum.

1D. ***Outline of the paper.*** In Section 2 we study the traveling waves solutions for (1-1). For sake of completeness we sketch the proof of existence which follows from bifurcation theory. In addition we analyze the size of the traveling waves. In Section 3 we study the linearization of (1-1) around the traveling waves. In Section 4, we introduce a new frame of reference which will help us to avoid the resonances found in Section 3. Finally, in Section 5 we prove Theorem 1.1.

2. Traveling waves

The existence of traveling waves for (1-1) was shown in [Hunter 2018]. Here we will study their size after we give some details about the existence proof. We look for solutions of (1-1) of the form

$$f_\epsilon(x,t) = u_\epsilon(x + v_\epsilon t);$$

thus we have to find $(u_\epsilon,\ v_\epsilon)$ solving

$$Hu_\epsilon - v_\epsilon u_\epsilon' + u_\epsilon u_\epsilon' = 0. \tag{2-1}$$

If (u_ϵ, v_ϵ) is a solution, so is $(u_\epsilon^n(x), v_\epsilon^n) = (u_\epsilon(nx)/n, v_\epsilon/n)$. Thus from one solution we can get n-fold symmetric solutions for all $n \geq 1$.

To solve (2-1) we can apply the Crandall–Rabinowitz theorem [1971] to

$$\begin{aligned} F : H_r^{k,+}(\mathbb{T}) \times \mathbb{C} &\to H_r^{k-1,-}(\mathbb{T}), \\ (u, \mu) &\mapsto Hu + uu' - (-1+\mu)u', \end{aligned}$$

where

$$H_r^{k,+}(\mathbb{T}) = \{2\pi\text{-periodic, mean zero, even functions analytic in the strip } \{|\mathrm{Im}(z)| < r\}\},$$

endowed with the norm

$$\|f\|_{H_r^{k,+}(\mathbb{T})} = \sum_{\pm} \|f(\cdot \pm ir)\|_{H^k(\mathbb{T})},$$

and

$$H_r^{k,-}(\mathbb{T}) = \{2\pi\text{-periodic, odd functions analytic in the strip } \{|\mathrm{Im}(z)| < r\}\},$$

endowed with the norm

$$\|f\|_{H_r^{k,-}(\mathbb{T})} = \sum_{\pm} \|f(\cdot \pm ir)\|_{H^k(\mathbb{T})}.$$

Here $\|\cdot\|_{H^k(\mathbb{T})}$ is the usual Sobolev norm, and it is enough to take $k \geq 1$ and $r = 1$.

We notice that $F(0,\mu) = 0$ and the derivative of F at $u = 0,\ \mu = 0$,

$$D_u F(0,0)h = Hh + h'$$

has a nontrivial element in its kernel belonging to $H_r^{k,+}(\mathbb{T})$, namely, $h = \cos(x)$.

Thus, the application of the Crandall–Rabinowitz theorem allows to show the existence of a branch of solutions $(u_\epsilon, v_\epsilon) \in (H_1^{1,+}, \mathbb{R})$, bifurcating from $(0,-1)$ for (2-1) with the leading-order term

$$u_\epsilon(x) = \epsilon\cos(x) + O(\epsilon^2), \quad v_\epsilon = -1 + O(\epsilon).$$

We remark that we obtain a bifurcation curve

$$\begin{aligned} \epsilon &\to (u_\epsilon, v_\epsilon), \\ B_\delta = \{z \in \mathbb{C} : |z| < \delta\} &\to (H_r^{k-1,-}, \mathbb{R}), \end{aligned} \tag{2-2}$$

which is differentiable and hence analytic on B_δ for δ small enough.

The rest of this section is devoted to proving further properties of these solutions.

Introducing the asymptotic expansion

$$u_\epsilon(x)=\sum_{n=1}^{\infty}u_n(x)\epsilon^n,\qquad v_\epsilon=\sum_{n=0}^{\infty}v_n\epsilon^n, \tag{2-3}$$

taking $u_1=\cos(x)$, $\lambda_0=-1$ and comparing the coefficient in ϵ^n we obtain that

$$u_n'+Hu_n=-v_{n-1}\sin(x)+\sum_{m=1}^{n-2}v_m u_{n-m}'-\frac{1}{2}\partial_x\sum_{m=1}^{n-1}u_{n-m}u_m=-v_{n-1}\sin(x)+f_n$$

for $n=2,3,\dots$.

We notice that in order to solve the equation $Hu+u'=f$ we need $(f,\sin(x))=0$. Therefore we have to choose $v_{n-1}=\frac{1}{\pi}(\sin(x),f_n)$. This gives us a recurrence for (u_n,v_{n-1}), $n\geq 2$, in terms of $\{(u_m,v_{m-1})\}_{m=1}^{n-1}$. In order to study this recurrence we will introduce the ansatz

$$u_n=\sum_{k=2}^{n}u_{n,k}\cos(kx). \tag{2-4}$$

By induction, one can check that the rest of coefficients in the expansion on cosines of u_n must be zero. In addition, if $u_\epsilon(x)$ solves (2-1), $u_{-\epsilon}(x+\pi)$ is also a bifurcation curve in the direction of $\cos(x)$, and then by uniqueness, $u_\epsilon(x)=u_{-\epsilon}(x+\pi)$, which yields $u_{n,k}=0$ if $n-k=1 \pmod 2$.

Comparing the coefficient of $\sin(kx)$, with $k=n \pmod 2$, and $2\leq k\leq n$, we have

$$(1-k)u_{n,k}+k\sum_{m=1}^{n-k}v_m u_{n-m,k}-\frac{k}{4}\sum_{m=1}^{n-1}\sum_{l=\max(1,k-n+m)}^{\min(m,k-1)}u_{m,l}u_{n-m,k-l}-\frac{k}{2}\sum_{m=1}^{n-1}\sum_{l=1}^{\min(m,n-m-k)}u_{m,l}u_{n-m,k+l}=0. \tag{2-5}$$

And comparing with $\sin(x)$ we have

$$v_{n-1}=\frac{1}{2}\sum_{m=1}^{n-1}\sum_{l=1}^{\min(m,n-m-1)}u_{m,l}u_{n-m,1+l}. \tag{2-6}$$

Up to order $O(\epsilon^4)$ we find

$$\begin{aligned}u_\epsilon(x)&=\epsilon\cos x-\tfrac{1}{2}\epsilon^2\cos 2x+\tfrac{3}{8}\epsilon^3\cos 3x+O(\epsilon^4),\\ v_\epsilon&=-1-\tfrac{1}{4}\epsilon^2+O(\epsilon^4).\end{aligned} \tag{2-7}$$

The recurrence (2-5)–(2-6) allows us to prove the following result.

Theorem 2.1. *The radius of convergence of the series* (2-3), *with the coefficients given by* (2-4)–(2-6), *is not bigger than* $2/e$.

Proof. From (2-5) and (2-6) we have

$$(1-n)u_{n,n}=\frac{1}{2}\sum_{k=1}^{n-1}(n-k)u_{k,k}u_{n-k,n-k}.$$

Let

$$y=y(x)=x+\sum_{n=2}^{\infty}u_{n,n}x^n.$$

Then $y - xy' = xyy'/2$, which, together with $y \sim x$ for small x, yields $y = 2W(x/2)$, where W is the Lambert W-function. Since the radius of convergence of W at 0 is $1/e$, the radius of convergence of y at 0 is $2/e$, so the radius of convergence of (2-5) and (2-6) is at most $2/e$. $\square$

In addition we can get a bound for how large the traveling wave can be.

Theorem 2.2. *The series* (2-3), *with the coefficients given by* (2-4)–(2-6), *converges for any* $\epsilon < x^* \sim 0.23$.

Proof. This proof is based on the implicit function theorem.

First we introduce the spaces

$$L^{2,-} = \{\text{odd functions } f \in L^2(\mathbb{T})\},$$
$$H^{1,+} = \{\text{even functions } f \in H^1(\mathbb{T})\}.$$

The space X is the orthogonal complement of the span of $\cos(x)$ in $H^{1,+}$. We will equip $L^{2,-}$ with the norm

$$\|u\|^2_{L^{2,-}} = \frac{1}{\pi}\int_{-\pi}^{\pi} |u(x)|^2\,dx \tag{2-8}$$

in such a way that $\|\sin(nx)\|_{L^{2,-}} = 1$ for $n \geq 1$. We also define

$$\|u\|^2_X = \frac{1}{\pi}\int_{-\pi}^{\pi} (|u'(x)|^2 + |u(x)|^2 - 2u(x)\Lambda u(x))\,dx. \tag{2-9}$$

Thus $\|\cos(nx)\|_X = n - 1$ for $n \geq 2$. The reason why we take these norms is technical and it will arise below. Finally we define

$$\mathcal{X} = X \times \mathbb{R}$$

equipped with the norm

$$\|(\tilde{u}, \nu)\|_{\mathcal{X}} = \sqrt{\|\tilde{u}\|^2_X + |\nu|^2}.$$

Since $u_\epsilon = \epsilon\cos x - \frac{1}{2}\epsilon^2\cos 2x + O(\epsilon^3)$ and $v_\epsilon = -1 + O(\epsilon^2)$, we can let

$$\begin{aligned}
&G(\epsilon, \tilde{u}, \mu)\\
&= \frac{1}{\epsilon^2} F\big(\epsilon\cos x - \tfrac{1}{2}\epsilon^2\cos 2x + \epsilon^2\tilde{u}, \epsilon\mu\big)\\
&= H\tilde{u} + \epsilon\big(\cos x(\sin 2x + \tilde{u}') + \big(\tfrac{1}{2}\cos 2x - \tilde{u}\big)(\sin x - \epsilon\sin 2x - \epsilon\tilde{u}')\big) + \tilde{u}' - \mu(\sin x - \epsilon\sin 2x - \epsilon\tilde{u}')
\end{aligned}$$

map $\mathbb{R} \times \mathcal{X}$ to $L^{2,-}$.

Because of the existence of traveling waves, we already know that there exists ϵ^* such that, for every $\epsilon \in [0, \epsilon^*)$, there exist $\tilde{u}_\epsilon$ and μ_ϵ satisfying

$$G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon) = 0.$$

In addition we have

$$\begin{aligned}
\left.\frac{dG(\epsilon, \tilde{u}_\epsilon + s\tilde{v}, \mu + s\nu)}{ds}\right|_{s=0} &\equiv dG_{\epsilon,\tilde{u}_\epsilon,\mu}(\tilde{v}, \nu)\\
&= H\tilde{v} + \epsilon\big(\tilde{v}'\cos x - \tilde{v}(\sin x - \epsilon\sin 2x - \epsilon\tilde{u}') - \epsilon\big(\tfrac{1}{2}\cos 2x - \tilde{u}\big)\tilde{v}'\big)\\
&\qquad + \tilde{v}' - \nu(\sin x - \epsilon\sin 2x - \epsilon\tilde{u}') + \epsilon\mu\tilde{v}'
\end{aligned}$$

maps $(\tilde{v}, \nu) \in \mathcal{X}$ linearly to $L^{2,-}$.

Thus as far as $dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon}(\tilde{u},\mu)$ is invertible from $\mathcal{X}$ to $L^{2,-}$ for $\epsilon \in [0, x^*)$ we will be able to extend the solution $(u_\epsilon, \mu_\epsilon)$ from $[0, \epsilon^*)$ to $[0, x^*)$ by the implicit function theorem.

Note that

$$dG_{0,0,0}(\tilde{v}, \nu) = H\tilde{v} + \tilde{v}' - \nu \sin x$$

is an isometry from $\mathcal{X}$ to $L^{2,-}$ under the norms given by (2-8) and (2-9). Therefore one can compute

$$dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon} = dG^{-1}_{0,0,0}(\mathbb{I} + dG^{-1}_{0,0,0}(dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon} - dG_{0,0,0})).$$

By the Neumann series and the fact that $dG_{0,0,0}$ is an isometry, $dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon}$ will be invertible, as long as $\|dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon} - dG_{0,0,0}\|_{\mathcal{X}\to L^{2,-}} < 1$. In order to show this last inequality we will bound

$$A_\epsilon := \|dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon} - dG_{0,0,0}\|_{\mathcal{X}\to L^{2,-}}$$

in terms of $\|\tilde{u}_\epsilon\|_X$ and μ_ϵ. After that we will bound $\|\tilde{u}_\epsilon\|_X$ and μ_ϵ. To do it we will use the information we have about $\partial_\epsilon \tilde{u}_\epsilon$ and $\partial_\epsilon \mu_\epsilon$.

Along the bifurcation curve,

$$\begin{aligned} &dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon}(\partial_\epsilon \tilde{u}_\epsilon, \mu'_\epsilon) \\ &\qquad = -\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon) \\ &\qquad = \cos x(\sin 2x + \tilde{u}'_\epsilon) + \tfrac{1}{2}\sin x(\cos 2x - 2\tilde{u}_\epsilon) - \epsilon(\cos 2x - 2\tilde{u}_\epsilon)(\sin 2x + \tilde{u}'_\epsilon) + \mu_\epsilon(\sin 2x + \tilde{u}'_\epsilon). \end{aligned} \tag{2-10}$$

Thus

$$(\partial_\epsilon \tilde{u}_\epsilon, \mu'_\epsilon) = dG^{-1}_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon}(-\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)).$$

Therefore

$$\sqrt{\|\partial_\epsilon \tilde{u}_\epsilon\|^2_X + |\mu'_\epsilon|^2} \le \frac{1}{1 - A_\epsilon}\|\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)\|_{L^{2,-}}. \tag{2-11}$$

In addition we have, for $r_\epsilon = \sqrt{\|\tilde{u}_\epsilon\|^2_X + |\mu_\epsilon|^2}$,

$$\partial_\epsilon r_\epsilon \le \sqrt{\|\partial_\epsilon \tilde{u}_\epsilon\|^2_X + |\mu'_\epsilon|^2} \le \frac{1}{1 - A_\epsilon}\|\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)\|_{L^{2,-}}.$$

Thus, explicit estimates for A_ϵ and $\|\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)\|_{L^{2,-}}$ in terms of r_ϵ and ϵ give a differential inequality for r_ϵ which can be used to bound A_ϵ.

We will need the following lemmas to bound A_ϵ and the norm $\|\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)\|_{L^{2,-}}$, where $\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)$ is given by the right-hand side of (2-10).

Lemma 2.3. *If $f \in X$ then $\|f \sin x - f' \cos x\|_{L^2} \le \sqrt{3}\|f\|_X$.*

Lemma 2.4. *If $f \in X$ then $\|2f \sin 2x - f' \cos 2x\|_{L^2} \le \frac{1}{2}\sqrt{17}\|f\|_X$.*

Proof. We only show Lemma 2.3. The proof of Lemma 2.4 is similar.

Let $f = \sum_{n=2}^{\infty} f_n \cos nx$. Then

$$2(f \sin x - f' \cos x) = f_2 \sin x + 2 f_3 \sin 2x + \sum_{n=3}^{\infty} n(f_{n-1} + f_{n+1}) \sin nx,$$

and

$$
\begin{aligned}
4\|f\sin(x)-f'\cos(x)\|_{L^2}^2 &= f_2^2+4f_3^2+\sum_{n=3}^{\infty} n^2(f_{n-1}+f_{n+1})^2\\
&\le 10f_2^2+20f_3^2+59f_4^2+88f_5^2+18f_2f_4+32f_3f_5+4\sum_{n=6}^{\infty}(n^2+1)f_n^2.
\end{aligned}
$$

The infinite sum is bounded by $1.48\sum_{n=6}^{\infty}(n-1)^2 f_n^2$, and it remains a finite-dimensional problem to show that the remaining terms are bounded by $12\sum_{n=2}^{5}(n-1)^2 f_n^2$. $\square$

Lemma 2.5. *If $f, g \in X$ then $\|(fg)'\|_{L^2} \le B\|f\|_X\|g\|_X$, where*

$$
B=\sqrt{\frac{\pi^2}{3}+\frac{869}{144}}\approx 3.05.
$$

Proof. Let $f=\sum_{n=2}^{\infty} f_n\cos nx,\ g=\sum_{n=2}^{\infty} g_n\cos nx \in X$. Then

$$
(fg)'=-\frac{1}{2}\sum_{n\ge 1} n \sum_{|m|\ge 2, |n-m|\ge 2} f_{|m|}g_{|n-m|}\sin nx,
$$

so by Cauchy–Schwarz,

$$
\|(fg)'\|_{L^2}^2=\frac{1}{8}\sum_{|n|\ge 1} n^2\biggl(\sum_{|m|\ge 2, |n-m|\ge 2} f_{|m|}g_{|n-m|}\biggr)^2 \le C\|f\|_X^2\|g\|_X^2,
$$

where

$$
C=\frac{1}{2}\sup_{n=1}^{\infty}\sum_{|m|\ge 2, |n-m|\ge 2}\frac{n^2}{(|m|-1)^2(|n-m|-1)^2}=\frac{\pi^2}{3}+\frac{869}{144}. \qquad \square
$$

Now, with Lemmas 2.3, 2.5 and 2.4 we are ready to bound the right-hand side of (2-10). Indeed,

$$
\|\text{right-hand side of (2-10)}\|_{L^2} \le \frac{\sqrt{10+4\epsilon^2}}{4}+2r_\epsilon+\frac{\sqrt{17}}{2}\epsilon\|\tilde{u}_\epsilon\|_X+B\epsilon\|\tilde{u}_\epsilon\|_X^2+\|\tilde{u}_\epsilon\|_X^2+|\mu_\epsilon|^2.
$$

Turning to the other side, we have

$$
\begin{aligned}
&(dG_{(\epsilon,\tilde{u}_\epsilon,\mu_\epsilon)}-dG_{(0,0,0)})(\tilde{v},\nu)\\
&\quad=\epsilon\bigl(\tilde{v}'\cos x-\tilde{v}(\sin x-\epsilon\sin 2x-\epsilon\tilde{u}_\epsilon')-\epsilon\bigl(\tfrac{1}{2}\cos 2x-\tilde{u}_\epsilon\bigr)\tilde{v}'\bigr)+\epsilon\nu(\sin 2x+\tilde{u}_\epsilon')+\epsilon\mu_\epsilon\tilde{v}'. \qquad (2\text{-}12)
\end{aligned}
$$

Again by Lemmas 2.3, 2.5 and 2.4 we find

$$
\begin{aligned}
\|\text{left-hand side of (2-12)}\|_{L^2} &\le \Bigl(\sqrt{3}\epsilon+\frac{\sqrt{17}}{4}\epsilon^2+B\epsilon^2\|\tilde{u}_\epsilon\|_X+2\epsilon|\mu_\epsilon|\Bigr)\|\tilde{v}\|_X+\epsilon(1+2\|\tilde{u}_\epsilon\|_X)|\nu|\\
&=\epsilon(\sqrt{3}+2|\mu_\epsilon|, 1+2\|\tilde{u}_\epsilon\|_X)\cdot(\|\tilde{v}\|_X,|\nu|)+\Bigl(\frac{\sqrt{17}}{4}\epsilon^2+B\epsilon^2\|\tilde{u}_\epsilon\|_X\Bigr)\|\tilde{v}\|_X,
\end{aligned}
$$

so

$$
A_\epsilon \le 2\epsilon+2\epsilon r_\epsilon+\frac{\sqrt{17}}{4}\epsilon^2+B\epsilon^2 r_\epsilon.
$$

Since $dG_{(0,0,0)}$ is an isometry, the Neumann series $(1-T)^{-1}=\sum_{n=0}^{\infty}T^n$ shows that if $A_\epsilon<1$, then $dG_{(\epsilon,\tilde{u}_\epsilon,\mu_\epsilon)}$ is invertible, and $\|dG^{-1}_{(\epsilon,\tilde{u}_\epsilon,\mu_\epsilon)}\|\le(1-A_\epsilon)^{-1}$, so

$$
\sqrt{\|\partial_\epsilon\tilde{u}_\epsilon\|_X^2+|\mu_\epsilon'|^2}\le\frac{1}{1-A_\epsilon}\Bigl(\frac{\sqrt{10+4\epsilon^2}}{4}+2r_\epsilon+\frac{\sqrt{17}}{2}\epsilon\|\tilde{u}_\epsilon\|_X+B\epsilon r_\epsilon^2+r_\epsilon^2\Bigr).
$$

Then $r_0 = 0$ and

$$r'_\epsilon \le \frac{\frac{1}{4}\sqrt{10+4\epsilon^2} + \big(2 + \frac{\sqrt{17}}{2}\epsilon\big) r_\epsilon + B\epsilon r_\epsilon^2 + r_\epsilon^2}{1 - 2\epsilon - 2\epsilon r_\epsilon - \frac{\sqrt{17}}{4}\epsilon^2 - B\epsilon^2 r_\epsilon}.$$

By the comparison principle, r_ϵ is bounded from above by the solution to

$$\frac{dy}{dx} = y' = \frac{\sqrt{10+4x^2} + (8 + 2\sqrt{17}x)y + 4Bxy^2 + 4y^2}{4 - 8x - 8xy - \sqrt{17}x^2 - 4Bx^2 y}, \tag{2-13}$$

with $y(0) = 0$, which is

$$(2Bx^2 + 4x)y^2 + (8x + \sqrt{17}x^2 - 4)y + x\sqrt{x^2 + 2.5} + 2.5\sinh^{-1}(\sqrt{0.4}x) = 0.$$

When $x > 0$, the quadratic coefficient and the constant are positive, so this equation has a nonnegative root if and only if

$$8x + \sqrt{17}x^2 - 4 \le -2\sqrt{(2Bx^2 + 4x)(x\sqrt{x^2 + 2.5} + 2.5\sinh^{-1}(\sqrt{0.4}x))},$$

whose solution is $x \le x^* \approx 0.23$ numerically. Hence the solution can be extended to $\epsilon = x^* \approx 0.23$. In order to achieve this last conclusion we notice that the solution to (2-13), with $y(0) = 0$ can be extended only if $A_\epsilon < 1$, since $1 - A_\epsilon$ arises in the denominator.

The above argument shows that for $\epsilon \in (-x^*, x^*)$, the bifurcation curve produces a traveling wave $u_\epsilon = \epsilon\cos x - \frac{1}{2}\epsilon^2\cos 2x + \epsilon^2\tilde{u}_\epsilon$, which travels at speed $v_\epsilon = -1 - \epsilon\mu_\epsilon$. Since all the operators involved are analytic in all its arguments, the bifurcation curve is analytic in ϵ on $(-x^*, x^*)$. It may be the case, however, that the power series for u_ϵ and v_ϵ around $\epsilon = 0$ has a smaller radius of convergence than x^* (for example, the function $f(x) = (x^2+1)^{-1}$ is analytic on the whole real line, but the radius of convergence of its power series around 0 is only 1.) We now show that the radius of convergence of the power series for u_ϵ and v_ϵ are indeed at least x^*.

We note that the above argument also works if ϵ is replaced with ϵe^{ia} ($a \in \mathbb{R}$), so the bifurcation curve (u_ϵ, v_ϵ) is also analytic in a neighborhood of $\{\epsilon e^{ia} : \epsilon \in (-x^*, x^*)\}$. Hence the curve is analytic in the disk of radius x^* centered at 0, so the radius of convergence of its power series around 0 is at least x^*. □

3. Linearization around traveling waves

In this section we will analyze the spectrum of the operator

$$L_\epsilon g = -v_\epsilon g_x + Hg + (u_\epsilon(x)g)_x$$

corresponding to the linearization of (1-1) around the traveling wave (u_ϵ, v_ϵ) bifurcating from zero in the direction of the cosine studied in the previous section.

Actually, let

$$f(x,t) = f_\epsilon(x,t) + g(x + v_\epsilon t, t),$$

with $f_\epsilon(x,t) = u_\epsilon(x + v_\epsilon t)$. Then

$$f_t(x,t) = \partial_t f_\epsilon(x,t) + (v_\epsilon g_x + g_t)(x + v_\epsilon t, t)$$

and

$$(Hf + ff_x)(x,t) = (Hf_\epsilon + f_\epsilon \partial_x f_\epsilon)(x,t) + Hg(x+v_\epsilon t, t) + \partial_x(f_\epsilon(x,t)g(x+v_\epsilon t,t)) + g(x+v_\epsilon t,t)\partial_x g(x+v_\epsilon t,t).$$

Putting these in (1-1), we get the equation for $g(x,t)$

$$\partial_t g(x,t) = -v_\epsilon g(x,t)_x + Hg(x,t) + (u_\epsilon(x)g(x,t))_x + g(x,t)g(x,t)_x.$$

The linearization around $g=0$ is

$$\partial_t g = L_\epsilon g,$$

where

$$L_\epsilon g = -v_\epsilon g_x + Hg + (u_\epsilon g)_x = \underbrace{Hg + g_x}_{Lg} + \sum_{n=1}^\infty \epsilon^n \underbrace{((u^{(n)} - v^{(n)})g)_x}_{L^{(n)}g}. \tag{3-1}$$

3A. ***The eigenvalue 0.*** The action of L on the Fourier modes is

$$\mathcal{F}(Lg)(m) = i(m - \operatorname{sgn} m)\hat{g}(m),$$

with eigenvalues 0 (double), $\pm i$, $\pm 2i$, ... (on $L^2(\mathbb{T})$ with zero mean). We first study the perturbation of the eigenspace corresponding to the double eigenvalue of 0. By translational symmetry, for any $\delta \in \mathbb{R}$, $u_\epsilon(x+\delta)$ is also a solution to

$$Hu - v_\epsilon u + uu' = 0.$$

Differentiation with respect to δ then shows that

$$L_\epsilon u_\epsilon' = Hu_\epsilon' - v_\epsilon u_\epsilon' + (u_\epsilon u_\epsilon')' = 0.$$

Also, since u_ϵ lies on a bifurcation curve, we can differentiate

$$Hu_\epsilon - v_\epsilon u_\epsilon' + u_\epsilon u_\epsilon' = 0,$$

with respect to ϵ to get

$$L_\epsilon \partial_\epsilon u_\epsilon = H\partial_\epsilon u_\epsilon - (\partial_\epsilon v_\epsilon)u_\epsilon' + u_\epsilon \partial_\epsilon u_\epsilon' + u_\epsilon' \partial_\epsilon u_\epsilon = (\partial_\epsilon v_\epsilon)u_\epsilon',$$

so on the span V_ϵ of u_ϵ' and $\partial_\epsilon u_\epsilon$, L_ϵ acts nilpotently by the matrix

$$\begin{pmatrix} 0 & \partial_\epsilon v_\epsilon \\ 0 & 0 \end{pmatrix}.$$

3B. ***Simplifying the linearized operator.*** We want to solve the eigenvalue problem

$$L_\epsilon g = ((u_\epsilon - v_\epsilon)g)' + Hg = \lambda(\epsilon)g.$$

Let $g = h'$. Then the antiderivative of the above is

$$(u_\epsilon - v_\epsilon)h' + Hh = \lambda(\epsilon)h \pmod 1. \tag{3-2}$$

Let $h = \tilde{h} \circ \phi_\epsilon$, where ϕ_ϵ satisfies

$$\phi_\epsilon' = \frac{2\pi}{u_\epsilon - v_\epsilon}\left(\int_0^{2\pi} \frac{dy}{u_\epsilon(y) - v_\epsilon}\right)^{-1}. \tag{3-3}$$

Then

$$(u_\epsilon - v_\epsilon)\phi_\epsilon'(\tilde h' \circ \phi_\epsilon) + H(\tilde h \circ \phi_\epsilon) = \lambda(\epsilon)\tilde h \circ \phi_\epsilon \pmod 1.$$

When ϵ is small enough, ϕ_ϵ is a diffeomorphism of $\mathbb{R}/2\pi\mathbb{Z}$, so

$$2\pi\left(\int_0^{2\pi} \frac{dy}{u_\epsilon(y) - v_\epsilon}\right)^{-1} \tilde h' + H(\tilde h \circ \phi_\epsilon) \circ \phi_\epsilon^{-1} = \lambda(\epsilon)\tilde h \pmod 1.$$

By the change of variable $z = \phi_\epsilon(y)$,

$$\begin{aligned} H(\tilde h \circ \phi_\epsilon) \circ \phi_\epsilon^{-1}(x) &= \frac{1}{2\pi}\int_0^{2\pi} \tilde h(\phi_\epsilon(y)) \cot\left(\frac{\phi_\epsilon^{-1}(x) - y}{2}\right) dy \\ &= \frac{1}{2\pi}\int_0^{2\pi} \tilde h(z) \cot\left(\frac{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)}{2}\right)(\phi_\epsilon^{-1})'(z)\, dz. \end{aligned}$$

The convolution kernel of the operator

$$R_\epsilon \tilde h = H(\tilde h \circ \phi_\epsilon) \circ \phi_\epsilon^{-1} - H\tilde h$$

is

$$K_\epsilon(x,z) = \cot\left(\frac{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)}{2}\right)(\phi_\epsilon^{-1})'(z) - \cot\left(\frac{x-z}{2}\right) \tag{3-4}$$

and the ϵ-derivative of the kernel is

$$\begin{aligned} \partial_\epsilon K_\epsilon(x,z) = -\csc^2\left(\frac{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)}{2}\right) \frac{\partial_\epsilon\phi_\epsilon^{-1}(x) - \partial_\epsilon\phi_\epsilon^{-1}(z)}{2}(\phi_\epsilon^{-1})'(z) \\ + \cot\left(\frac{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)}{2}\right)\partial_\epsilon(\phi_\epsilon^{-1})'(z). \end{aligned}$$

Near $x = 0$, $\csc x - 1/x^2$ and $\cot x - 1/x$ are smooth, and $(\phi_\epsilon^{-1})'$ is smooth everywhere, so when $x - z$ is small enough, up to a smooth function in (x, z),

$$\begin{aligned} \frac{\partial_\epsilon K_\epsilon(x,z)}{2} &= -\frac{(\partial_\epsilon\phi_\epsilon^{-1}(x) - \partial_\epsilon\phi_\epsilon^{-1}(z))(\phi_\epsilon^{-1})'(z)}{(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z))^2} + \frac{\partial_\epsilon(\phi_\epsilon^{-1})'(z)}{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)} \\ &= \frac{\partial_\epsilon(\phi_\epsilon^{-1})'(z)(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)) - (\partial_\epsilon\phi_\epsilon^{-1}(x) - \partial_\epsilon\phi_\epsilon^{-1}(z))(\phi_\epsilon^{-1})'(z)}{(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z))^2} \\ &= \frac{\partial_\epsilon(\phi_\epsilon^{-1})'(z)(x-z)^2\int_0^1 (1-t)(\phi_\epsilon^{-1})''((1-t)z+tx)\,dt}{(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z))^2} \\ &\qquad - \frac{(\phi_\epsilon^{-1})'(z)(x-z)^2\int_0^1 (1-t)\partial_\epsilon(\phi_\epsilon^{-1})''((1-t)z+tx)\,dt}{(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z))^2}, \end{aligned}$$

which is itself a smooth function of (x, z) when $x - z$ is small enough (because ϕ_ϵ^{-1} is smooth). Then

$$\|\partial_\epsilon R_\epsilon \tilde h^{(m)}\|_{\dot H^k} \lesssim_{k,m} \|\tilde h\|_{L^2/(1)}, \quad k, m = 0, 1, \ldots,$$

where the constant does not depend on ϵ, for all $\tilde{h} \in H^m/(1)$, or, equivalently,

$$\|\partial_\epsilon R_\epsilon \tilde{h}\|_{\dot{H}^k} \lesssim_{k,m} \|\tilde{h}\|_{\dot{H}^{-m}}, \quad k, m = 0, 1, \ldots, \tag{3-5}$$

where the dot over H means that the norm does not measure frequency zero.

Definition 3.1. We say an operator is of class $\mathcal{S}$ if it satisfies (3-5). We say a family of operators is of class $\mathcal{S}$ uniformly if for each k and m there is an implicit constant that makes (3-5) true for all operators in the family.

Thus $\partial_\epsilon R_\epsilon$ is of class $\mathcal{S}$ uniformly in ϵ. Since $R_0 = 0$, R_ϵ/ϵ is also of class $\mathcal{S}$ uniformly in ϵ.

Now the eigenvalue problem for $\tilde{h}$ is of the form

$$(c_\epsilon \partial_x + H + R_\epsilon)\tilde{h} = \lambda(\epsilon)\tilde{h} \pmod 1$$

or, equivalently,

$$(\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon)\tilde{h} = c_\epsilon^{-1}\lambda(\epsilon)\tilde{h} \pmod 1, \tag{3-6}$$

where

$$c_\epsilon = 2\pi\left(\int_0^{2\pi} \frac{dy}{u_\epsilon(y) - v_\epsilon}\right)^{-1} \tag{3-7}$$

and R_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ. Note that since u_ϵ and v_ϵ are analytic functions of ϵ on a neighborhood of 0, with $u_0 = 0$ and $v_0 = -1$, so are ϕ_ϵ, R_ϵ and c_ϵ with $\phi_0 = I$, $R_0 = 0$ and $c_0 = 1$.

3C. *Spectral analysis of the linearization.* The eigenvalue problem (3-6) is a perturbation of the eigenvalue problem

$$\tilde{h}' + H\tilde{h} = \lambda\tilde{h} \pmod 1,$$

with explicit eigenvalues

$$0 \text{ (double)}, \; ni, \quad n = \pm1, \pm2, \ldots,$$

and eigenfunctions

$$e^{\pm ix}, \; e^{i(n+\operatorname{sgn} n)x}, \quad n = \pm1, \pm2, \ldots.$$

They form an orthogonal basis of $H^k/(1)$ for any nonnegative integer k.

Definition 3.2. Let $T : \dot{H}^k(\mathbb{T}) \to \dot{H}^k(\mathbb{T})$ for $k \in \mathbb{N}$ be a linear operator. We will define

$$\|T\| := \|T\|_{\dot{H}^k(\mathbb{T}) \to \dot{H}^k(\mathbb{T})}.$$

The resolvent $(\partial_x + H - z)^{-1}$ is also a Fourier multiplier whose action on Fourier modes is

$$(\partial_x + H - z)^{-1} e^{\pm i(n+1)x} = (\pm ni - z)^{-1} e^{\pm i(n+1)x}, \quad n = 0, 1, \ldots. \tag{3-8}$$

The circle

$$\Gamma_n = \left\{z : |z - ni| = \tfrac{1}{2}\right\}, \quad n = \pm1, \pm2, \ldots,$$

encloses a single eigenvalue $\pm ni$, and the circle

$$\Gamma_0 = \left\{z : |z| = \tfrac{1}{2}\right\}$$

encloses the double eigenvalue 0. On Γ_n and Γ_0 we have

$$|z - mi| \geq \tfrac{1}{2}, \quad m \in \mathbb{Z}, \tag{3-9}$$

so by (3-8),

$$\|(\partial_x + H - z)^{-1}\| \leq 2, \quad z \in \Gamma_n,\ n \in \mathbb{Z}. \tag{3-10}$$

Moreover the projection

$$P_n = -\frac{1}{2\pi i}\int_{\Gamma_n} (\partial_x + H - z)^{-1}\, dz, \quad n = \pm 1, \pm 2, \ldots,$$

is the projection on the span of $e^{i(n+\operatorname{sgn} n)x}$ and the projection

$$P_0 = -\frac{1}{2\pi i}\int_{\Gamma_0} (\partial_x + H - z)^{-1}\, dz$$

is the projection on the span of e^{ix} and e^{-ix}.

Now when ϵ is small enough and $z \in \Gamma_n$, we have

$$\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z = (\partial_x + H - z)(1 + (\partial_x + H - z)^{-1} R'_\epsilon),$$

where

$$R'_\epsilon = (\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon) - (\partial_x + H) = (c_\epsilon^{-1} - 1)H + c_\epsilon^{-1} R_\epsilon \tag{3-11}$$

is analytic in ϵ near 0, with $R'_0 = 0$, thanks to the analyticity of c_ϵ. Taking the inverse gives

$$(\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z)^{-1} = (1 + (\partial_x + H - z)^{-1} R'_\epsilon)^{-1} (\partial_x + H - z)^{-1}$$

and the Neumann series

$$(1 + (\partial_x + H - z)^{-1} R'_\epsilon)^{-1} = \sum_{n=0}^{\infty} ((\partial_x + H - z)^{-1} R'_\epsilon)^n \tag{3-12}$$

converges because

$$\|(\partial_x + H - z)^{-1} R'_\epsilon\| \leq 2\|R'_\epsilon\| \lesssim_k \epsilon < 1$$

when ϵ is small enough (depending on k). Moreover,

$$\|(1 + (\partial_x + H - z)^{-1} R'_\epsilon)^{-1} - I\| \lesssim_k \epsilon$$

and so

$$\|(\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z)^{-1} - (\partial_x + H - z)^{-1}\| \lesssim \epsilon$$

uniformly for $z \in \Gamma_n$. Hence the projections

$$Q_n(\epsilon) = -\frac{1}{2\pi i}\int_{\Gamma_n} (\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z)^{-1}\, dz, \quad n \in \mathbb{Z}, \tag{3-13}$$

exist and satisfy

$$\|Q_n(\epsilon) - P_n\| \lesssim_k \epsilon, \quad n \in \mathbb{Z}, \tag{3-14}$$

uniformly in n. Then by [Kato 1976, Chapter I, Section 4.6], when ϵ is small enough, $Q_n(\epsilon)$ is conjugate to P_n. Thus $\dim \operatorname{ran} Q_n(\epsilon) = 1$ for $n \neq 0$ and $\dim \operatorname{ran} Q_0(\epsilon) = 2$. So $\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon$ has a single eigenvalue enclosed by Γ_n for $n \neq 0$. In Section 3A we showed that the action on the range of $Q_0(\epsilon)$ is given by a nonzero nilpotent 2-by-2 matrix. If z is outside all these circles, then (3-10) still holds and the Neumann series (3-12) still converges to show that $\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z$ is invertible, so it has no other eigenvalues.

3D. ***Analyticity of eigenvalues and eigenvectors.*** By (3-8) and (3-9), $(\partial_x + H - z)^{-1}$ is analytic in (z, ϵ) for z in a neighborhood U of $\bigcup_{n\in\mathbb{Z}} \Gamma_n$, and ϵ near 0. By (3-11), R'_ϵ is analytic in ϵ near 0, so the series (3-12) shows that $(\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z)^{-1}$ is analytic in (z, ϵ) for $z \in U$ and ϵ near 0, and the integral (3-13) shows that all the projections $Q_n(\epsilon)$ $(n \in \mathbb{Z})$ are analytic in a neighborhood of 0 independent of n.

Let $\psi_n(\epsilon)$ be the corresponding eigenvectors to $Q_n(\epsilon)$ for $n \neq 0$. Thanks to (3-14), a good choice is $\psi_n(\epsilon) = Q_n(\epsilon) e^{i(n+\operatorname{sgn} n)x}$, which is nonzero and analytic in a neighborhood of 0 independent of n. Then by (3-6),

$$Q_n(\epsilon)(\partial_x + c_\epsilon^{-1}(H + R_\epsilon)) e^{i(n+\operatorname{sgn} n)x} = (\partial_x + c_\epsilon^{-1}(H + R_\epsilon))\psi_n(\epsilon) = c_\epsilon^{-1} \lambda_n(\epsilon) \psi_n(\epsilon).$$

On the other hand, the left-hand side equals

$$(n + \operatorname{sgn} n) i\, Q_n(\epsilon) e^{i(n+\operatorname{sgn} n)x} + c_\epsilon^{-1} Q_n(\epsilon)(H + R_\epsilon) e^{\pm i(n+\operatorname{sgn} n)x},$$

which is another vector analytic in ϵ near 0. Then by the next lemma, all the eigenvalues $c_\epsilon^{-1}\lambda_n(\epsilon)$, and hence $\lambda_n(\epsilon)$, are analytic in a neighborhood of 0 independent of n.

Lemma 3.3. *Let $u(\epsilon)$ and $v(\epsilon)$ be two vectors analytic in $\epsilon \in U$ satisfying*

$$u(\epsilon) \neq 0 \quad \text{and} \quad v(\epsilon) = \lambda(\epsilon) u(\epsilon), \qquad \epsilon \in U.$$

Then $\lambda(\epsilon)$ is analytic in $\epsilon \in U$.

Proof. Without loss of generality assume that $0 \in U$. Since the result is local in ϵ, it suffices to show that $\lambda(\epsilon)$ is analytic in a smaller neighborhood of 0.

Since $u(0) \neq 0$, we can find a linear functional f such that $f(u(0)) \neq 0$. Then $f(u(\epsilon)) \neq 0$ in a neighborhood of 0, and so

$$\lambda(\epsilon) = \frac{f(v(\epsilon))}{f(u(\epsilon))}$$

is analytic in a neighborhood of 0. □

Regarding the double eigenvalue 0, in Section 3A we showed that u'_ϵ and $\partial_\epsilon u_\epsilon$ are two generalized eigenvectors of the operator L_ϵ. Using the relation given in Section 3B, they correspond to two generalized eigenvectors $\psi_0^-(\epsilon)$ and $\psi_0^+(\epsilon)$ of the operator $\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon$, via the relation $(\psi_0^-(\epsilon) \circ \phi_\epsilon)' = u'_\epsilon$ and $(\psi_0^+(\epsilon) \circ \phi_\epsilon)' = \partial_\epsilon u_\epsilon$. Then clearly $\psi_0^\pm(\epsilon)$ are both analytic in ϵ.

From the analyticity of the eigenvalues $c_\epsilon^{-1}\lambda_n(\epsilon)$, it is easy to derive bounds on their Taylor coefficients.

Proposition 3.4. *For $k \geq 1$ and $n \neq 0$, the coefficient of ϵ^k in $c_\epsilon^{-1}\lambda_n(\epsilon)$ is bounded in absolute value by C^k for a constant $C > 0$ independent of n,* *Proof.* At the end of Section 3C we showed that when ϵ is in a neighborhood of 0 independent of n, the eigenvalues $c_\epsilon^{-1}\lambda_n(\epsilon)$ are enclosed in the circle Γ_n. Then

$$|c_\epsilon^{-1}\lambda_n(\epsilon)-ni|<\tfrac{1}{2},\qquad n=\pm 1,\pm 2,\ldots.$$

The result follows from Cauchy's integral formula for Taylor coefficients. □

Corollary 3.5. *For $k\geq 0$ and $n\neq 0$, the coefficient of ϵ^k in $\lambda_n(\epsilon)$ is bounded in absolute value by $|n|C^k$ for a constant $C>0$ independent of n.*

Proof. Since c_ϵ is analytic in ϵ near 0 with $c_0=1$, and $\lambda_n(0)=ni$, the result follows from Leibniz's rule. □

3E. ***Conjugation to a Fourier multiplier.*** We have conjugated the eigenspaces of $T=\partial_x+c_\epsilon^{-1}H+c_\epsilon^{-1}R_\epsilon$ (and also of $c_\epsilon\partial_x+H+R_\epsilon$) to Fourier modes via the operator

$$1+W_\epsilon=\sum_{n\in\mathbb{Z}}P_nQ_n(\epsilon),$$

where P_0 is the projection onto the span of $e^{\pm ix}$, $Q_0(\epsilon)$ is the projection onto the span of $\psi_0^{\pm}(\epsilon)$, P_n is the projection onto the span of $e^{i(n+\operatorname{sgn} n)x}$, and $Q_n(\epsilon)$ is the projection onto the span of $\psi_n(\epsilon)$, $n=\pm 1,\pm 2,\ldots$.

We will view T as a perturbation of $\partial_x+c_\epsilon^{-1}H$ and follow the proof of [Kato 1976, Chapter V, Theorem 4.15a]. In the process we will extract more information from the fact that R_ϵ is of class $\mathcal{S}$. Since

$$P_n^2=P_n,\qquad \sum_{n\in\mathbb{Z}}P_n=1, \tag{3-15}$$

we have

$$W_\epsilon=\sum_{n\in\mathbb{Z}}P_n(Q_n(\epsilon)-P_n) \tag{3-16}$$

and $W_0=0$.

Proposition 3.6. *W_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ.*

Proof. We bound each term on the right-hand side separately. By [Kato 1976, Chapter V, (4.38)],

$$Q_n(\epsilon)-P_n=-c_\epsilon^{-1}Q_n(\epsilon)R_\epsilon Z_n(\epsilon)-c_\epsilon^{-1}Z_n'(\epsilon)R_\epsilon P_n,$$

where

$$Z_n(\epsilon)=\frac{1}{2\pi i}\int_{\Gamma_n}(z-(n+(1-c_\epsilon^{-1})\operatorname{sgn} n)i)^{-1}(\partial_x+c_\epsilon^{-1}H-z)^{-1}\,dz,$$
$$Z_n'(\epsilon)=\frac{1}{2\pi i}\int_{\Gamma_n}(z-c_\epsilon^{-1}\lambda_n(\epsilon))^{-1}(T-z)^{-1}\,dz.$$

We now bound the operator norms of the right-hand side, with uniformity in ϵ and decay in n, in order to show that the sum in n converges.

First note that it is clear from the frequency side that when ϵ is in a neighborhood of 0 independent of n and $z\in\bigcup_{n\in\mathbb{Z}}\Gamma_n$ for all $m\geq 0$, the operator $(\partial_x+c_\epsilon^{-1}H-z)^{-1}$ is bounded from H^m to H^m, uniformly in ϵ and z. Since R_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ (see (3-5) and notice that $R_0=0$), it follows from the Neumann series that $\|(T-z)^{-1}\|_{\dot H^m\to\dot H^m}$ is finite and only depends on m. Since $|z-(n+(1-c_\epsilon^{-1})\operatorname{sgn} n)i|$ and $|z-c_\epsilon^{-1}\lambda_n(\epsilon)|$ are uniformly bounded from below, both $Z_n(\epsilon)$ and $Z_n'(\epsilon)$ are bounded from $\dot H^m$

to $\dot H^m$, uniformly in ϵ and n. Since $Q_n(\epsilon)$ is given by a similar integral (3-13), it also has this property, which is also trivially true for P_n. Now, for all $n, m, k \in \mathbb{Z}$, $m, k \geq 0$ and $\tilde h \in L^2$,

$$\begin{aligned}\|Z_n'(\epsilon)R_\epsilon P_n\tilde h\|_{\dot H^k} \lesssim_k \|R_\epsilon P_n\tilde h\|_{\dot H^k} &\lesssim_{m,k} |\epsilon|\,\|P_n\tilde h\|_{\dot H^{-m-2}}\\ &\lesssim_{m,k} |\epsilon|(1+|n|)^{-2}\|\tilde h\|_{\dot H^{-m}}\end{aligned} \tag{3-17}$$

because P_n is the projection onto very specific Fourier modes. For the first term we have

$$\|R_\epsilon Z_n(\epsilon)\tilde h\|_{\dot H^k} \lesssim_{m,k} |\epsilon|\,\|Z_n(\epsilon)\tilde h\|_{\dot H^{-m}} \lesssim_{m,k} |\epsilon|\,\|\tilde h\|_{\dot H^{-m}}.$$

To introduce the action of $Q_n(\epsilon)$, note that the image of $Q_n(\epsilon)$ lies in the eigenspace of the operator $c_\epsilon\partial_x + H + R_\epsilon$, with eigenvalue $\lambda_n(\epsilon)$, so for $n \neq 0$ and $u \in \operatorname{Im} Q_n(\epsilon)$ we have

$$u = \lambda_n(\epsilon)^{-1}(c_\epsilon u' + Hu + R_\epsilon u),$$

so $\|u\|_{\dot H^k} \lesssim_k |\lambda_n(\epsilon)|^{-1}\|u\|_{\dot H^{k+1}} \lesssim |n|^{-1}\|u\|_{\dot H^{k+1}}$. Hence

$$\|Q_n(\epsilon)R_\epsilon Z_n(\epsilon)\|_{\dot H^k} \lesssim_k n^{-2}\|R_\epsilon Z_n(\epsilon)\|_{\dot H^{k+2}} \lesssim_{m,k} |\epsilon|(1+|n|)^{-2}\|\tilde h\|_{\dot H^{-m}}. \tag{3-18}$$

This also holds for $n = 0$ because R_ϵ/ϵ is of class $\mathcal{S}$ uniformly, so W_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ thanks to the convergence of $\sum_{n\in\mathbb{Z}}(1+|n|)^{-2}$. □

Now for $k = 0, 1, \dots,$ there is a neighborhood of 0 such that when ϵ is in this neighborhood, $\|W_\epsilon\|_{\dot H^k\to\dot H^k} < 1$, so $1 + W_\epsilon : \dot H^k \to \dot H^k$ is invertible. By (3-15) and (3-16) it follows easily that

$$(1+W_\epsilon)Q_n(\epsilon) = P_n(1+W_\epsilon), \tag{3-19}$$

so the eigenspace of T is conjugated to the (span of) Fourier modes, and hence T is conjugated to a Fourier multiplier.

We have proven the following lemma:

Lemma 3.7. *For ϵ small enough, there exists an operator W_ϵ such that W_ϵ/ϵ is of class $\mathcal{S}$, uniformly in ϵ. Moreover*:

(1) $1 + W_\epsilon : \dot H^k \to \dot H^k$ *is invertible.*

(2) $(1+W_\epsilon)Q_n(\epsilon) = P_n(1+W_\epsilon),\ n \in \mathbb{Z}$.

(3) *If ψ is in the closed linear span of the eigenvectors $\psi_n(\epsilon)$ $(n \neq 0)$ of $c_\epsilon\partial_x + H + R_\epsilon$, then*

$$(1+W_\epsilon)(c_\epsilon\partial_x + H + R_\epsilon)\psi = \Lambda_\epsilon(1+W_\epsilon)\psi,$$

where Λ_ϵ is a multiplier such that

$$\Lambda_\epsilon e^{i(n+\operatorname{sgn} n)x} = \lambda_n(\epsilon)e^{i(n+\operatorname{sgn} n)x}, \quad n = \pm1, \pm2, \dots.$$

3F. ***Taylor expansion of eigenvalues.*** Now we Taylor expand the eigenvalues $\lambda_n(\epsilon)$ for $n \neq 0$. To do so it is more convenient to study the eigenvalue problem (3-2) for h:

$$L_\epsilon g := ((u_\epsilon - v_\epsilon)g)' + Hg = \lambda(\epsilon)g.$$

Recall the operator $L = L_0 = \partial_x + H$ whose action on the Fourier modes is

$$\mathcal{F}(Lg)(m) = i(m - \operatorname{sgn} m)\hat{g}(m),$$

with eigenvalues 0 (double), $\pm i$, $\pm 2i$, ... (g mean zero).

Since (u_ϵ, v_ϵ) is analytic in ϵ on a neighborhood of 0, and

$$\|h'\|_{L^2} \le \|h' + Hh\|_{L^2} + \|Hh\|_{L^2} = \|Lh\|_{L^2} + \|h\|_{L^2},$$

by [Kato 1976, Chapter VII, Theorem 2.6], L_ϵ is a holomorphic family of operators of type (A), so by Chapter VII, Section 2.3, all the results in Chapter II, Sections 1 and 2 apply, and we can compute the Taylor coefficients of $\lambda(\epsilon)$ as if L_ϵ acted on a finite-dimensional vector space.

We start with computing the resolvent of L,

$$R(z) = (L - z)^{-1}$$

whose action on the Fourier modes is

$$\mathcal{F}(R(z)g)(m) = (i(m - \operatorname{sgn} m) - z)^{-1}\hat{g}(m).$$

Around the eigenvalue ni ($n = \pm 1, \pm 2, \ldots$) we have the expansion

$$R(z) = (ni - z)^{-1} P_n + \sum_{k=0}^{\infty} (z - ni)^k S_n^{k+1},$$

where P_n is the projection on the span of $e^{i(n+\operatorname{sgn} n)x}$ and

$$\mathcal{F}(S_n g)(m) = \frac{\hat{g}(m)}{i(m - \operatorname{sgn} m - n)}, \quad m \ne n + \operatorname{sgn} n. \tag{3-20}$$

By [Kato 1976, (II.2.33)],

$$\lambda_n(\epsilon) = ni + \sum_{k=1}^{\infty} \epsilon^k \lambda_n^{(k)}, \quad n = \pm 1, \pm 2, \ldots,$$

where

$$\lambda_n^{(k)} = \sum_{p=1}^{k} \frac{(-1)^p}{p} \sum_{\substack{v_1+\cdots+v_p=n,\ v_j \ge 1 \\ h_1+\cdots+h_p=p-1}} \operatorname{Tr} L^{(v_p)} S_n^{(h_p)} \cdots L^{(v_1)} S_n^{(h_1)},$$

where $S_n^{(0)} = -P_n$ and, for $h \ge 1$, $S_n^{(h)} = S_n^h$, with S_n defined in (3-20), and $L^{(v)}$ is the coefficient of ϵ^v in the Taylor expansion of L_ϵ. Note that the constraints in the summation imply that there is some $j \in \{1, \ldots, p\}$ such that $h_j = 0$ and so $S_n^{(h_j)} = -P_n$, so every summand is a finite-rank operator whose trace is thus well-defined.

Lemma 3.8. *If A is a finite-rank operator, then* $\operatorname{Tr} AB = \operatorname{Tr} BA$.

Proof. By linearity we can assume A has the form $A(\cdot) = f(\cdot)v$ for some (not necessarily continuous) linear functional f. Then $\operatorname{Tr} A = f(v)$. Since $AB(\cdot) = f(B\cdot)v$ and $BA(\cdot) = f(\cdot)Bv$, it follows that $\operatorname{Tr} AB = f(Bv) = \operatorname{Tr} BA$. □

Using the lemma above, we can simplify the sum in $\lambda_n^{(k)}$ a little. Indeed, there are p circular rotations of the tuple $(h_1, \dots, h_p)$. Since $\left(\sum_j h_j, p\right) = 1$, the p circular rotations are all distinct, so we can choose the lexicographically smallest one as a representative. For such a representative, $h_1 = \min_j h_j = 0$, so $S_n^{(h_1)} = -P_n$, and thus we only need to act $L^{(v_p)} S_n^{(h_p)} \cdots L^{(v_1)}$ on $e^{i(n+\operatorname{sgn} n)x}$ and take the $(n+\operatorname{sgn} n)$-th mode to compute the trace. Thus

$$\lambda_n^{(k)} = \sum_{p=1}^{k} (-1)^{p-1} \sum_{\substack{v_1+\cdots+v_p=k,\ v_j\ge 1 \\ h_1+\cdots+h_p=p-1 \\ (h_1,\dots,h_p) \text{ is a representative}}} \mathcal{F}[L^{(v_p)} S_n^{(h_p)} \cdots L^{(v_1)} e^{i(n+\operatorname{sgn} n)x}](n+\operatorname{sgn} n). \tag{3-21}$$

Let us compute some terms $\lambda_n^{(k)}$ by using the formula (3-21). We have

$$\lambda_n^{(1)} = \operatorname{Tr} L^{(1)} P_n = 0$$

because L_1 shifts the mode by 1, and

$$\lambda_n^{(2)} = \operatorname{Tr}(L^{(2)} P_n - L^{(1)} S_n L^{(1)} P_n).$$

Put $s = \operatorname{sgn} n$. We extract the $(n+s)$-th mode of each term:

$$\begin{aligned} \operatorname{Tr} L^{(2)} P_n &= \mathcal{F}[L^{(2)} 2 e^{i(n+s)x}](n+s) = \frac{i(n+s)}{4}, \\ L^{(1)} S_n L^{(1)} e^{i(n+s)x} &= \frac{i L^{(1)} S_n}{2}((n+s+1)e^{i(n+s+1)x} + (n+s-1)e^{i(n+s-1)x}) \\ &= \frac{L^{(1)}}{2}((n+s+1)e^{i(n+s+1)x} - (n+s-1)e^{i(n+s-1)x}), \\ \operatorname{Tr} L^{(1)} S_n L^{(1)} P_n &= \frac{i(n+s+1)(n+s) - i(n+s-1)(n+s)}{4} = \frac{i(n+s)}{2}, \end{aligned}$$

so

$$\lambda_n^{(2)} = \frac{i(n+s)}{4} - \frac{2i(n+s)}{4} = -\frac{i(n+s)}{4}.$$

We can further compute that

$$\lambda_n(\epsilon) = in - \frac{\epsilon^2 i(n+s)}{4} - \frac{11\epsilon^4 i(n+s)}{32} - \frac{527 i \epsilon^6 (n+s)}{768} + O_n(\epsilon^7)$$

for $n = \pm 1, \pm 2, \pm 3, \dots$.

Proposition 3.9. *For $n = \pm 1, \pm 2, \dots$,*

$$\lambda_n^{(k)} = \begin{cases} 0, & 2 \nmid k, \\ i c^{(k)} (n + \operatorname{sgn} n), & k \le 2|n| + 2, \end{cases}$$

where $c^{(k)}$ is the k-th Taylor coefficient of c_ϵ as defined in (3-7).

When $k \ge 2|n| + 4$, $\lambda_n^{(k)}$ is still purely imaginary but the formula $\lambda_n^{(k)} = i c^{(k)}(n + \operatorname{sgn} n)$ does not hold in general.

Proof. Firstly we notice that, for $n = \pm 1$, the coefficient of ϵ^6 in $\lambda_{\pm 1}(\epsilon)$ is

$$\lambda_1(\epsilon) = i - \frac{\epsilon^2 i}{2} - \frac{11\epsilon^4 i}{16} - \frac{529\epsilon^6 i}{384} + O(\epsilon^7),$$

which does not hold for $\lambda_{\pm 1}^{(6)} = \pm 2ic^{(6)}$.

Next, we prove the fist part of the lemma. In each summand of (3-21), all the coefficients are real, except that each operator L brings a factor of i to the Fourier coefficients (via the operator ∂_x), and each operator S_n removes a factor of i (see (3-20)). Hence each summand is purely imaginary, and so is $\lambda_n^{(k)}$.

In each summand of (3-21), the operator $S_n^{(h_j)}$ is a Fourier multiplier that does not shift the modes, while the operator $L^{(m)}g = ((u^{(m)} - v^{(m)})g)'$ shifts the modes by at most m because $u^{(m)}$ only contains modes up to $e^{\pm imx}$. Also the amount of shift is equal to $m \pmod 2$. Thus when acting the sequence $L^{(v_p)} S_n^{(h_p)} \cdots L^{(v_1)}$ on $e^{i(n+s)x}$, the mode is consecutively shifted by at most $v_1, v_2, \dots, v_p$, and the total amount of shift is equal to $\sum_j v_j = k \pmod 2$. Since in the end we are taking the $(n+s)$-th mode, the total amount of shift must be 0 in order to count, so when k is odd $\lambda_n^{(k)} = 0$. When k is even, the mode $e^{i(n+s)x}$ can only be shifted as far as $e^{i(n+s\pm k/2)x}$; otherwise it can never be shifted back. Hence when $k \le 2|n| + 2 = 2|n+s|$, the frequency always has the same sign as n or becomes 0. In the former case we can take $\operatorname{sgn} m = \operatorname{sgn} n$ in (3-20), while in the latter case the derivative in L kills it, so it does not hurt if we still take $\operatorname{sgn} m = \operatorname{sgn} n$ in (3-20). Either way we can take $\operatorname{sgn} m = \operatorname{sgn} n$ in (3-20). Thus the action of S_n is the same as that of S_n', where

$$\mathcal{F}(S_n' g)(m) = \frac{\hat{g}(m)}{i(m - n - \operatorname{sgn} n)}, \quad m \ne n + \operatorname{sgn} n.$$

For $n > 0$, the operator S_n' is the analog of S_n for L^+, with

$$\mathcal{F}(L^+ g)(m) = i(m-1)\hat{g}(m),$$

i.e., $L^+ g = g' - ig$. Hence $\lambda_n^{(k)}$ remains the same if we replace L with L^+. Now we have

$$L_\epsilon^+ g := L^+ g + \sum_{n=1}^{\infty} \epsilon^n L^{(n)} g = -v_\epsilon g' - ig + (u_\epsilon g)' = ((u_\epsilon - v_\epsilon)g)' - ig,$$

whose eigenvalue problem is

$$((u_\epsilon - v_\epsilon)g)' - ig = \lambda^+(\epsilon) g.$$

Using the same change of variable as in Section 3B, the problem above can be transformed to

$$\tilde{h}' - ic_\epsilon^{-1}\tilde{h} = c_\epsilon^{-1}\lambda^+(\epsilon)\tilde{h},$$

whose eigenvalues are

$$\lambda_{n'}^+(\epsilon) = n' c_\epsilon i - i.$$

Since when $\epsilon \to 0$, $\lambda_n(\epsilon) \to ni$ and $c_\epsilon \to 1$, we must have $n' = n + 1$, and so

$$\lambda_n(\epsilon) = (n+1)c_\epsilon i - i + O_n(\epsilon^{2n+4}).$$

For $n < 0$, note that since L preserves real-valued functions, its eigenvalues come in conjugate pairs, so $\lambda_n(\epsilon) = \overline{\lambda_{|n|}(\epsilon)} = -\lambda_{|n|}(\epsilon)$ has the same property. □

Corollary 3.10. *When $|\epsilon|$ is small enough,*

$$|\lambda_n(\epsilon) - (n + \operatorname{sgn} n)c_\epsilon i + i \operatorname{sgn} n| < |n|(C\epsilon)^{2|n|+4} < C'\epsilon^6, \quad n \in \mathbb{Z}\backslash\{0\},$$
$$|\lambda_n'(\epsilon) - (n + \operatorname{sgn} n)\partial_\epsilon c_\epsilon i| < |n|(C\epsilon)^{2|n|+3} < C'\epsilon^5, \quad n \in \mathbb{Z}\backslash\{0\},$$

for some constant C, $C' > 0$ independent of n.

Proof. By Proposition 3.9. the Taylor expansions of $\lambda_n(\epsilon)$ and $(n+\operatorname{sgn} n)c_\epsilon i - i \operatorname{sgn} n$ differ only from the term $\epsilon^{2|n|+4}$. By Corollary 3.5, the error terms of the former sum up to $O\big(|n|\sum_{k=2|n|+4}^{\infty}(C\epsilon)^k\big) = O(|n|(C\epsilon)^{2|n|+4})$ if, say, $C|\epsilon| < \frac{1}{2}$. The error term of the latter clearly also satisfy this bound.

To extend the chain of inequalities it suffices to note that $|n|(C\epsilon)^{2|n|-2}$ is uniformly bounded for $n \neq 0$ if $|C\epsilon| < \frac{1}{2}$. $\square$

3G. ***Time resonance analysis.*** For m, n and $l \in \mathbb{Z}$ we consider

$$\lambda_m(\epsilon) + \lambda_n(\epsilon) + \lambda_l(\epsilon) = (m+n+l)c_\epsilon i + (\operatorname{sgn} m + \operatorname{sgn} n + \operatorname{sgn} l)(c_\epsilon - 1)i + O(\epsilon^6).$$

Proposition 3.11. *If $m, n, l \in \mathbb{Z}$ and $mnl \neq 0$, then when ϵ is small enough, $|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)| > \frac{1}{5}\epsilon^2$.*

Proof. By (3-7) and (2-7),

$$\begin{aligned}
\frac{c_\epsilon}{2\pi} &= \left(\int_0^{2\pi} \frac{dy}{u_\epsilon(y) - v_\epsilon}\right)^{-1} = \left(\int_0^{2\pi} \frac{dy}{1 + \epsilon\cos y - \frac{1}{2}\epsilon^2\cos 2y + \frac{1}{4}\epsilon^2}\right)^{-1} + O(\epsilon^3)\\
&= \left(\int_0^{2\pi} \big(1 - \epsilon\cos y + \epsilon^2\cos^2 y + \tfrac{1}{2}\epsilon^2\cos 2y - \tfrac{1}{4}\epsilon^2\big)\,dy\right)^{-1} + O(\epsilon^3),\\
c_\epsilon &= \big(1 + \tfrac{1}{4}\epsilon^2\big)^{-1} + O(\epsilon^3) = 1 - \tfrac{1}{4}\epsilon^2 + O(\epsilon^3).
\end{aligned}$$

We distinguish three cases.

Case 1: $m+n+l \neq 0$. Then $|m+n+l| \geq 1$. Since $c_\epsilon - 1 \lesssim \epsilon^2$,

$$\lambda_m(\epsilon) + \lambda_n(\epsilon) + \lambda_l(\epsilon) = (m+n+l)c_\epsilon i + O(\epsilon^2).$$

Since $c_\epsilon \to 1$ as $\epsilon \to 0$, we have $|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)| > \frac{1}{2}|m+n+l|$ for small ϵ.

Case 2: $m+n+l = 0$ and $mnl \neq 0$. Then

$$\lambda_m(\epsilon) + \lambda_n(\epsilon) + \lambda_l(\epsilon) = -\tfrac{1}{4}(\operatorname{sgn} m + \operatorname{sgn} n + \operatorname{sgn} l)\epsilon^2 i + O(\epsilon^3).$$

Since $|\operatorname{sgn} m| = |\operatorname{sgn} n| = |\operatorname{sgn} l| = 1$, we have $|\operatorname{sgn} m + \operatorname{sgn} n + \operatorname{sgn} l| \geq 1$, so $|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)| > \frac{1}{5}\epsilon^2$ when ϵ is small enough. $\square$

When $m+n+l = 0$ and $mnl = 0$, since $\lambda_n(\epsilon)$ is odd in n, it follows that $\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon) = 0$. We do have time resonance in this case. We will eliminate this case by choosing a new frame of reference.

4. A new frame of reference

Recall that the traveling wave solution

$$f_\epsilon(x,t) = u_\epsilon(x + v_\epsilon t)$$

satisfies

$$\partial_t f_\epsilon = H f_\epsilon + f_\epsilon \partial_x f_\epsilon,$$

i.e.,

$$v_\epsilon u_\epsilon' = H u_\epsilon + u_\epsilon u_\epsilon'.$$

Now we aim to find a new reference frame. Let $P_0^\pm(\epsilon)$ be the projection on the 1-dimensional space spanned by the eigenvector $\varphi_0^+(\epsilon) = \partial_\epsilon u_\epsilon$ and $\varphi_0^-(\epsilon) = -\epsilon^{-1} u_\epsilon'$, respectively. Then we aim to rewrite

$$f(x,t) = u_{\epsilon(t)}(x + a(t)) + g(x + a(t), t),$$

where $\epsilon, a \in \mathbb{R}$ and $P_0^\pm(\epsilon(t)) g = 0$. We first show that it is always possible, provided that f is close to a traveling wave.

Proposition 4.1. *Let $k \geq 2$. Then there is $r = r(k) > 0$ such that if $|\epsilon_0| < r$ and $\|f - u_{\epsilon_0}\|_{H^k} < r|\epsilon_0|$, then there is $\epsilon \in \mathbb{R}$, $a \in \mathbb{R}/2\pi\mathbb{Z}$ and $g \in H^k$ such that*

$$f(x) = u_\epsilon(x + a) + g(x + a), \tag{4-1}$$

$$P_0^\pm(\epsilon) g = 0, \tag{4-2}$$

$$|\epsilon - \epsilon_0| + \|g\|_{H^k} \lesssim \|f - u_{\epsilon_0}\|_{H^k}. \tag{4-3}$$

Moreover, ϵ, a and g depend smoothly on f.

Proof. Define the map $F : (-r, r)^2 \to \mathbb{R}^2$, $(\epsilon, a) \mapsto (y^+, y^-)$, with

$$P_0^\pm(\epsilon)(f(x - a) - u_\epsilon(x)) = y^\pm \varphi_0^\pm(\epsilon). \tag{4-4}$$

We now find the solution to the equation $F(\epsilon, a) = 0$. Since $P_0^\pm(\epsilon)$ is uniformly bounded in L^2 and $\|\varphi_0^\pm(\epsilon)\|$ is uniformly bounded from below,

$$|F(\epsilon, a)| \lesssim \|f(x - a) - u_\epsilon\|_{L^2}. \tag{4-5}$$

Summing the two equations in (4-4) and taking the total derivative yields

$$-(P_0^+(\epsilon) + P_0^-(\epsilon))(f'(x - a))da - \varphi_0^+(\epsilon)d\epsilon + (\partial_\epsilon P_0^+(\epsilon) + \partial_\epsilon P_0^-(\epsilon))(f(x - a) - u_\epsilon(x))d\epsilon \tag{4-6}$$

$$= \varphi_0^+(\epsilon)dy^+ + \varphi_0^-(\epsilon)dy^- + y^+\partial_\epsilon\varphi_0^+(\epsilon)d\epsilon + y^-\partial_\epsilon\varphi_0^-(\epsilon)d\epsilon. \tag{4-7}$$

Since $\|f\|_{H^2} \leq \|u_{\epsilon_0}\|_{H^2} + r|\epsilon_0| \lesssim |\epsilon_0|$, we have

$$\begin{aligned} \|f(x - a) - u_\epsilon\|_{H^1} &\leq \|f(x - a) - f(x)\|_{H^1} + \|f - u_{\epsilon_0}\|_{H^1} + \|u_\epsilon - u_{\epsilon_0}\|_{H^1} \\ &\lesssim |a\epsilon_0| + r|\epsilon_0| + |\epsilon - \epsilon_0|. \end{aligned} \tag{4-8}$$

Since both $P_0^\pm(\epsilon)$ and $\partial_\epsilon P_0^\pm(\epsilon)$ are uniformly bounded on L^2, and $u_\epsilon' = -\epsilon\varphi_0^-(\epsilon)$,

$$\|(4\text{-}6) - \epsilon\varphi_0^-(\epsilon)da + \varphi_0^+(\epsilon)d\epsilon\|_{L^2} \lesssim (|a\epsilon_0| + r|\epsilon_0| + |\epsilon - \epsilon_0|)(|da| + |d\epsilon|).$$

By (4-5) and (4-8),

$$\|y^\pm\partial_\epsilon\varphi_0^\pm(\epsilon)\|_{L^2} \lesssim |F(\epsilon, a)| \lesssim |a\epsilon_0| + r|\epsilon_0| + |\epsilon - \epsilon_0|,$$

so

$$\|(4\text{-}7) - \varphi_0^+(\epsilon)dy^+ - \varphi_0^-(\epsilon)dy^-\|_{L^2} \lesssim (|a\epsilon_0| + r|\epsilon_0| + |\epsilon - \epsilon_0|)|d\epsilon|.$$

Hence the equality between (4-6) and (4-7) gives an estimate of the differential

$$\left\| dF(\epsilon, a) - \begin{pmatrix} 1 & 0 \\ 0 & -\epsilon \end{pmatrix} \right\| \lesssim |a\epsilon_0| + r|\epsilon_0| + |\epsilon - \epsilon_0|.$$

We assume that the solution (ϵ, a) satisfies $|\epsilon - \epsilon_0| + |a\epsilon_0| < r_0|\epsilon_0|$, where r_0 is small enough. This in particular implies $\frac{1}{2}|\epsilon_0| < |\epsilon| < 2|\epsilon_0|$. Then

$$\left\| dF(\epsilon, a) - \begin{pmatrix} 1 & 0 \\ 0 & -\epsilon_0 \end{pmatrix} \right\| \lesssim (r_0 + r)|\epsilon_0|$$

is also small enough. Let

$$G = \mathbb{I} + \left(dF(\epsilon, a) - \begin{pmatrix} 1 & 0 \\ 0 & -\epsilon_0 \end{pmatrix} \right) \begin{pmatrix} 1 & 0 \\ 0 & -1/\epsilon_0 \end{pmatrix}.$$

Then

$$dF = G \begin{pmatrix} 1 & 0 \\ 0 & -\epsilon_0 \end{pmatrix}$$

and

$$\|G - \mathbb{I}\| \lesssim r_0 + r.$$

If r_0 and r are small enough, then $\|G\|$ and $\|G^{-1}\| < 2$.

Let $(\epsilon_1, a_1) = (\epsilon_0, 0) - dF(\epsilon_0, 0)^{-1}F(\epsilon_0, 0)$. Then (recalling (4-5))

$$|\epsilon_1 - \epsilon_0| + |a_1\epsilon_0| \lesssim |G^{-1}F(\epsilon_0, 0)| \lesssim |F(\epsilon_0, 0)| \lesssim \|f - u_{\epsilon_0}\|_{L^2} \lesssim r|\epsilon_0|.$$

Since $|\partial_\epsilon^2 F|$ and $|\partial_{a\epsilon} F| \lesssim 1$, and $|\partial_a^2 F| \lesssim \|f\|_{H^2} \lesssim |\epsilon_0|$, by Taylor's theorem,

$$|F(\epsilon_1, a_1)| \lesssim |\epsilon_1 - \epsilon_0|^2 + |\epsilon_1 - \epsilon_0|\,|a_1| + |\epsilon_0|\,|a_1|^2 \lesssim r|F(\epsilon_0, 0)|.$$

Hence the iteration $(\epsilon_{n+1}, a_{n+1}) = (\epsilon_n, a_n) + dF(\epsilon_n, a_n)^{-1}F(\epsilon_n, a_n)$ converges when r is small enough. Moreover $|\epsilon_n - \epsilon_0| + |a_n\epsilon_0| \lesssim |F(\epsilon_0, 0)|$. Then $(\epsilon, a) := \lim_{n\to\infty}(\epsilon_n, a_n)$ satisfies $F(\epsilon, a) = 0$ and $|\epsilon - \epsilon_0| + |a\epsilon_0| \lesssim |F(\epsilon_0, 0)| \lesssim r|\epsilon_0| < r_0|\epsilon_0|$ if r is small compared to r_0.

Let $g = f(x - a) - u_\epsilon$. Then (4-1) and (4-2) clearly hold. Moreover,

$$\begin{aligned}
\|g\|_{H^k} = \|g(x + a)\|_{H^k} &= \|f(x) - u_\epsilon(x + a)\|_{H^k} \\
&\le \|f - u_{\epsilon_0}\|_{H^k} + \|u_\epsilon(x + a) - u_{\epsilon_0}(x)\|_{H^k} \\
&\lesssim \|f - u_{\epsilon_0}\|_{H^k} + |\epsilon - \epsilon_0| + |a\epsilon_0| \\
&\lesssim \|f - u_{\epsilon_0}\|_{H^k} + |F(\epsilon_0, 0)| \lesssim \|f - u_{\epsilon_0}\|_{H^k}
\end{aligned}$$

showing (4-3). The smooth dependence of ϵ, a and g on f is also clear. □

By translation symmetry, if f is $r|\epsilon_0|$-close to $u_{\epsilon_0}(x + a)$ for some $a \in \mathbb{R}/2\pi\mathbb{Z}$, we can reach a similar conclusion. Then we can write

$$f(x, t) = u_{\epsilon(t)}(x + a(t)) + g(x + a(t), t).$$

We will obtain an energy estimate for g. Combined with local well-posedness of (1-1) and Proposition 4.1, we can show that the solution extends as long as the energy estimate closes, see the end of Section 5B.

To get the energy estimate, we first need to derive an evolution equation for g. Since f is differentiable in t, so are $\epsilon(t)$, $a(t)$ and g, and we get

$$f_t(x,t) = a'(t)(u'_\epsilon + g_x)(x+a(t)) + \epsilon'(t)\partial_\epsilon u_\epsilon(x+a(t)) + g_t(x+a(t),t)$$

and

$$\begin{aligned}(Hf + ff_x)(x,t) = (Hu_\epsilon + u_\epsilon u'_\epsilon)(x+a(t)) + Hg(x+a(t),t)\\ + \partial_x(u_\epsilon(x+a(t))g(x+a(t),t)) + (gg_x)(x+a(t),t).\end{aligned}$$

The equation for g is then

$$\begin{aligned} g_t &= v_\epsilon u'_\epsilon - a'(t)(u'_\epsilon + g_x) - \epsilon'(t)\partial_\epsilon u_\epsilon + Hg + (u_\epsilon g)_x + gg_x \\ &= L_\epsilon g + (v_\epsilon - a'(t))(u'_\epsilon + g_x) - \epsilon'(t)\partial_\epsilon u_\epsilon + gg_x. \end{aligned}$$

Since $P_0^\pm(\epsilon)g(t) = 0$, we have $P_0^\pm(\epsilon)g_t = -\epsilon'(t)\partial_\epsilon P_0^\pm(\epsilon)g$, so the action of the projections $P_0^\pm(\epsilon)$ on the above equation is

$$\begin{aligned} (v_\epsilon - a'(t))P_0^+(\epsilon)g_x + \epsilon'(t)(\partial_\epsilon P_0^+(\epsilon)g - \partial_\epsilon u_\epsilon) + P_0^+(\epsilon)(gg_x) = 0, \\ (v_\epsilon - a'(t))(u'_\epsilon + P_0^-(\epsilon)g_x) + \epsilon'(t)\partial_\epsilon P_0^-(\epsilon)g + P_0^-(\epsilon)(gg_x) = 0. \end{aligned}$$

Since $P_0^\pm(\epsilon)$ are bounded on L^2, we have $\|P_0^\pm(\epsilon)g_x\|_{L^2} \lesssim \|g\|_{H^1}$. Since $P_0^\pm(\epsilon)$ are analytic in ϵ, we have $\|\partial_\epsilon P_0^\pm(\epsilon)g\|_{L^2} \lesssim \|g\|_{L^2}$. Since $P_0^\pm(\epsilon)$ is a projection, we have $P_0^\pm(\epsilon)^2 = P_0^\pm(\epsilon)$. Taking the derivative in ϵ and using the constraint $P_0^\pm(\epsilon)g = 0$, we have $P_0^\pm(\epsilon)\partial_\epsilon P_0^\pm(\epsilon)g = \partial_\epsilon P_0^\pm(\epsilon)g$, i.e., $\partial_\epsilon P_0^\pm(\epsilon)g$ is in the 1-dimensional space spanned by $\varphi_0^\pm(\epsilon)$. Hence

$$|P_0^\pm(\epsilon)g_x/\varphi_0^\pm(\epsilon)| \lesssim \|g\|_{H^1}, \quad |\partial_\epsilon P_0^\pm(\epsilon)g/\varphi_0^\pm(\epsilon)| \lesssim \|g\|_{L^2}.$$

Thus, dividing the two equations by $\varphi_0^\pm(\epsilon)$ we get

$$\left| \left(\begin{pmatrix} 0 & 1 \\ \epsilon & 0 \end{pmatrix} + O(\|g\|_{H^1}) \right) \begin{pmatrix} v_\epsilon - a'(t) \\ \epsilon'(t) \end{pmatrix} \right| = \left| \begin{pmatrix} P_0^+(\epsilon)(gg_x)/\varphi_0^+(\epsilon) \\ P_0^-(\epsilon)(gg_x)/\varphi_0^-(\epsilon) \end{pmatrix} \right| \lesssim \|g(t)\|_{H^1}^2.$$

Assuming $\|g(t)\|_{H^1}/|\epsilon|$ is small enough we have

$$\begin{pmatrix} v_\epsilon - a'(t) \\ \epsilon'(t) \end{pmatrix} = \begin{pmatrix} O(\|g(t)\|_{H^1}^2/|\epsilon|) \\ O(\|g(t)\|_{H^1}^2) \end{pmatrix}. \tag{4-9}$$

4A. *Diagonalization.* To find the evolution of other modes, we diagonalize the equation for g. Let $g = h_x$ and $h = \tilde{h} \circ \phi_\epsilon$, where ϕ_ϵ satisfies (3-3). Recall from (3-1) that $L_\epsilon g = -v_\epsilon g_x + Hg + (u_\epsilon g)_x$, so

$$h_t = -v_\epsilon h_x + Hh + u_\epsilon h_x - \epsilon'(t)\partial_\epsilon U_\epsilon + (v_\epsilon - a'(t))(u_\epsilon + h_x) + \tfrac{1}{2}h_x^2 \pmod 1,$$

where U_ϵ is a primitive of u_ϵ. Differentiating $h = \tilde{h} \circ \phi_\epsilon$ with respect to ϵ we get

$$h_t = \tilde{h}_t \circ \phi_\epsilon + \epsilon'(t)(\partial_\epsilon \phi_\epsilon)(\tilde{h}_x \circ \phi_\epsilon).$$

On the other hand,

$$(-v_\epsilon h_x + Hh + u_\epsilon h_x)_x = L_\epsilon g = (((c_\epsilon \partial_x + H + R_\epsilon)\tilde{h}) \circ \phi_\epsilon)_x,$$

so

$$\tilde{h}_t = (c_\epsilon \partial_x + H + R_\epsilon)\tilde{h} - \epsilon'(t)(\partial_\epsilon \phi_\epsilon \circ \phi_\epsilon^{-1})\tilde{h}_x - \epsilon'(t)\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1} \\ + (v_\epsilon - a'(t))(u_\epsilon + h_x) \circ \phi_\epsilon^{-1} + \tfrac{1}{2} h_x^2 \circ \phi_\epsilon^{-1} \pmod 1.$$

By the chain rule, $h_x = \phi_\epsilon'(\tilde{h}_x \circ \phi_\epsilon)$, so $h_x \circ \phi_\epsilon^{-1} = (\phi_\epsilon' \circ \phi_\epsilon^{-1})\tilde{h}_x$, and

$$\tilde{h}_t = (c_\epsilon \partial_x + H + R_\epsilon)\tilde{h} + \Phi_\epsilon \tilde{h}_x - \epsilon'(t)\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1} + (v_\epsilon - a'(t))u_\epsilon \circ \phi_\epsilon^{-1} + \tfrac{1}{2}(\phi_\epsilon' \circ \phi_\epsilon^{-1})^2 \tilde{h}_x^2 \pmod 1,$$
$$\Phi_\epsilon = -\epsilon'(t)(\partial_\epsilon \phi_\epsilon \circ \phi_\epsilon^{-1}) + (v_\epsilon - a'(t))(\phi_\epsilon' \circ \phi_\epsilon^{-1}).$$

Using the operator W_ϵ from Lemma 3.7 we have

$$(1 + W_\epsilon)(c_\epsilon \partial_x + H + R_\epsilon) = \Lambda_\epsilon (1 + W_\epsilon),$$

where Λ_ϵ is a Fourier multiplier whose action on the Fourier mode $e^{i(n+\operatorname{sgn} n)x}$ is multiplication by $\lambda_n(\epsilon)$. Since W_ϵ/ϵ is of class $\mathcal{S}$, uniformly in ϵ, for any smooth function F, the operator

$$\tilde{h} \mapsto \mathcal{R}_\epsilon(F)\tilde{h} := (1 + W_\epsilon)(F\tilde{h}_x) - F((1 + W_\epsilon)\tilde{h})_x$$

is of class $\mathcal{S}$, with the implicit constants depending on the C^k norms of F.

Let $\mathfrak{h} = (1 + W_\epsilon)\tilde{h}$. Then

$$(1 + W_\epsilon)\tilde{h}_t = \Lambda_\epsilon \mathfrak{h} + \Phi_\epsilon \mathfrak{h}_x - \epsilon'(t)(1 + W_\epsilon)(\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1}) \\ + (v_\epsilon - a'(t))(1 + W_\epsilon)(u_\epsilon \circ \phi_\epsilon^{-1}) + N_\epsilon[\mathfrak{h}, \mathfrak{h}] + \mathcal{R}_\epsilon(\Phi_\epsilon)\tilde{h} \pmod 1,$$

where

$$N_\epsilon[\mathfrak{h}, \mathfrak{h}] = \tfrac{1}{2}(\phi_\epsilon' \circ \phi_\epsilon^{-1})^2((1 + W_\epsilon)^{-1}\mathfrak{h})_x^2. \tag{4-10}$$

Both $\mathcal{R}_\epsilon(\partial_\epsilon \phi_\epsilon \circ \phi_\epsilon^{-1})$ and $\mathcal{R}_\epsilon(\phi_\epsilon' \circ \phi_\epsilon^{-1} - 1)/\epsilon$ are of class $\mathcal{S}$, uniformly in ϵ when ϵ is small. Moreover, since W_ϵ is analytic in ϵ with $W_0 = 0$, so is $\mathcal{R}_\epsilon(1)$ with $\mathcal{R}_0(1) = 0$. Hence $R_\epsilon(1)/\epsilon$ is of class $\mathcal{S}$ uniformly in ϵ, and so is $\mathcal{R}_\epsilon(\phi_\epsilon' \circ \phi_\epsilon^{-1})/\epsilon$.

Since $\partial_\epsilon u_\epsilon$ and u_ϵ' are in the generalized eigenspace of L_ϵ associated with the eigenvalue 0, we have $\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1}$ and $u_\epsilon \circ \phi_\epsilon^{-1}$ are in the corresponding space of $c_\epsilon \partial_x + H + R_\epsilon$, so $(1 + W_\epsilon)(\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1})$ and $(1 + W_\epsilon)(u_\epsilon \circ \phi_\epsilon^{-1})$ are in the space spanned by $\sin x$ and $\cos x$, according to Lemma 3.7.

Now we have

$$\begin{aligned} \mathfrak{h}_t &= (1 + W_\epsilon)\tilde{h}_t + \epsilon'(t)\partial_\epsilon W_\epsilon \tilde{h} \\ &= \Lambda_\epsilon \mathfrak{h} + \Phi_\epsilon \mathfrak{h}_x + N_\epsilon[\mathfrak{h}, \mathfrak{h}] + \text{Rest} \pmod{1, \sin x, \cos x}, \end{aligned} \tag{4-11}$$

where $N_\epsilon[\mathfrak{h}, \mathfrak{h}]$ is given by (4-10) and

$$\text{Rest} = \epsilon'(t)(\partial_\epsilon W_\epsilon)\tilde{h} + \mathcal{R}_\epsilon(\Phi_\epsilon)\tilde{h}$$

is also of class $\mathcal{S}$ uniformly in ϵ when ϵ is small.

Recall that $\epsilon'(t)$ and $a'(t)$ are chosen such that $P_0(\epsilon)g(t)=0$ for all t, where $P_0(\epsilon)$ is the projection onto the span of $\partial_\epsilon u_\epsilon$ and u'_ϵ. This implies $Q_0(\epsilon)\tilde{h}(t)=0$ for all t, where $Q_0(\epsilon)$ is the projection onto the span of $\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1}$ and $u_\epsilon \circ \phi_\epsilon^{-1}$. Since $1+W_\epsilon$ maps the span of $\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1}$ and $u_\epsilon \circ \phi_\epsilon^{-1}$ to the span of $\sin x$ and $\cos x$, we have $\hat{\mathfrak{h}}(1)=\hat{\mathfrak{h}}(-1)=0$ for all t.

5. Energy estimates

Since $\hat{\mathfrak{h}}(1)=\hat{\mathfrak{h}}(-1)=0$ for all t, for $k=0,1,\dots$ we define the energy

$$E_k=\tfrac{1}{2}\|\mathfrak{h}\|_{\dot{H}^k}^2=\tfrac{1}{2}\|\mathfrak{h}\|_{H^k/(1,\sin x,\cos x)}^2$$

and aim to control its growth.

Using the evolution equation (4-11) for $\mathfrak{h}$ and the anti-self-adjointness of Λ_ϵ we get

$$\begin{aligned}\frac{d}{dt}E_k(t)&=E_\Phi(t)+E_N(t)+E_{\mathrm{Rest}}(t),\\ E_\Phi(t)&=\langle \Phi_\epsilon \mathfrak{h}_x,\mathfrak{h}\rangle_{\dot{H}^k},\\ E_N(t)&=\langle N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)],\mathfrak{h}(t)\rangle_{\dot{H}^k},\\ E_{\mathrm{Rest}}(t)&=\langle \epsilon'(t)\partial_\epsilon W_\epsilon \tilde{h}(t)+\mathcal{R}_\epsilon(\Phi_\epsilon)\tilde{h}(t),\mathfrak{h}(t)\rangle_{\dot{H}^k}.\end{aligned}$$

Recall that $g=h_x$, $h=\tilde{h}\circ\phi_\epsilon$ and $\mathfrak{h}=(1+W_\epsilon)\tilde{h}$. When ϵ is small enough, the last two are bounded operators with bounded inverse between $\dot{H}^k$, $k=0,1,\dots$, so

$$\|g\|_{H^k}\approx_k \|h\|_{\dot{H}^{k+1}}\approx_k \|\tilde{h}\|_{\dot{H}^{k+1}}\approx_k \|\mathfrak{h}\|_{\dot{H}^{k+1}}. \tag{5-1}$$

Since $\mathcal{R}_\epsilon(\partial_\epsilon\phi_\epsilon\circ\phi_\epsilon^{-1})$, $\mathcal{R}_\epsilon(\phi'_\epsilon\circ\phi_\epsilon^{-1})/\epsilon$ and $\partial_\epsilon W_\epsilon$ are of class $\mathcal{S}$ uniformly in ϵ,

$$\begin{gathered}\|(\epsilon'(t)\partial_\epsilon W_\epsilon-\epsilon'(t)\mathcal{R}_\epsilon(\partial_\epsilon\phi_\epsilon\circ\phi_\epsilon^{-1}))\tilde{h}(t)\|_{\dot{H}^k}\lesssim_k \|g(t)\|_{H^1}^2\|\tilde{h}(t)\|_{\dot{H}^1}\lesssim_k E_2(t)^{3/2},\\ \|(v_\epsilon-a'(t))\mathcal{R}_\epsilon(\phi'_\epsilon\circ\phi_\epsilon^{-1})\tilde{h}(t)\|_{\dot{H}^k}\lesssim_k (\|g(t)\|_{H^1}^2/\epsilon)\epsilon\|\tilde{h}(t)\|_{\dot{H}^1}\lesssim_k E_2(t)^{3/2}\end{gathered}$$

so

$$|E_{\mathrm{Rest}}(t)|\lesssim_k E_2(t)^{3/2}E_k(t)^{1/2}. \tag{5-2}$$

To bound E_Φ we use (4-9) and (5-1) to get

$$\|\Phi'_\epsilon\|_{C^k}\lesssim_k \|g(t)\|_{H^1}^2+(\|g(t)\|_{H^1}^2/|\epsilon|)|\epsilon|\lesssim_k E_2(t).$$

Since E_Φ loses only one derivative in $\mathfrak{h}$, we have

$$|E_\Phi(t)-\langle \Phi_\epsilon\partial_x^{k+1}\mathfrak{h}(t),\partial_x^k\mathfrak{h}(t)\rangle_{L^2/(1)}|\lesssim_k E_2(t)E_k(t). \tag{5-3}$$

For the sake of bounding this term, since the inner product is taken in the space $L^2/(1)$, we can without loss of generality assume that $\hat{\mathfrak{h}}(0)=0$ (which is not true in general) and integrate by parts to get

$$2\langle \Phi_\epsilon\partial_x^{k+1}\mathfrak{h}(t),\partial_x^k\mathfrak{h}(t)\rangle_{L^2/(1)}=\int_0^{2\pi}\Phi_\epsilon\partial_x(\partial_x^k\mathfrak{h}(t))^2\,dx=-\int_0^{2\pi}\Phi'_\epsilon(\partial_x^k\mathfrak{h}(t))^2\,dx$$

so again by (4-9) and (5-1),

$$|E_\Phi(t)|\lesssim_k E_2(t)E_k(t). \tag{5-4}$$

Combining (5-2), (5-3) and (5-4) shows that

$$\left|\frac{d}{dt}E_k(t)-E_N(t)\right|\lesssim_k E_2(t)E_k(t). \tag{5-5}$$

5A. *Normal form transformation.* To bound E_N we recall the expression of N_ϵ from (4-10). Since N_ϵ does not depend on the constant mode of $\mathfrak{h}$, we can also assume without loss of generality that $\hat{\mathfrak{h}}(0)=0$. We further have the decompositions

$$\begin{aligned} E_N(t)&=E_{N1}(t)+E_{N2}(t),\\ E_{N1}(t)&=\frac{1}{2}\int_0^{2\pi}\partial_x^k\mathfrak{h}(t)\partial_x^k(\partial_x\mathfrak{h}(t))^2\,dx=\sum_{j=2}^{[k/2]+1}c_{kj}\int_0^{2\pi}\partial_x^k\mathfrak{h}(t)\partial_x^{k+2-j}\mathfrak{h}(t)\partial_x^j\mathfrak{h}(t), \end{aligned}\tag{5-6}$$

where $c_{kj}\in\mathbb{R}$ are constants and we integrated by parts to get rid of the terms with $k+1$ derivatives falling on a single factor of $\mathfrak{h}$.

We use the normal form transformation to bound them. Define the trilinear map

$$\begin{aligned} D_\epsilon[f_1,f_2,f_3]=\sum_{mnl\neq 0}\frac{1}{\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)}\int_0^{2\pi}&\hat{f}_1(m+\operatorname{sgn}m)e^{i(m+\operatorname{sgn}m)x}\\ &\times\hat{f}_2(n+\operatorname{sgn}n)e^{i(n+\operatorname{sgn}n)x}\hat{f}_3(l+\operatorname{sgn}l)e^{i(l+\operatorname{sgn}l)x}\,dx \end{aligned}$$

and put

$$D_{1,k,j}(t)=D_{\epsilon(t)}[\partial_x^k\mathfrak{h}(t),\partial_x^{k+2-j}\mathfrak{h}(t),\partial_x^j\mathfrak{h}(t)].$$

Then

$$\begin{aligned}\frac{d}{dt}D_{1,k,j}(t)=\epsilon'(t)(\partial_\epsilon D_\epsilon)&[\partial_x^k\mathfrak{h}(t),\partial_x^{k+2-j}\mathfrak{h}(t),\partial_x^j\mathfrak{h}(t)]\\ &+D_\epsilon[\partial_x^k\partial_t\mathfrak{h}(t),\partial_x^{k+2-j}\mathfrak{h}(t),\partial_x^j\mathfrak{h}(t)]\\ &\quad+D_\epsilon[\partial_x^k\mathfrak{h}(t),\partial_x^{k+2-j}\partial_t\mathfrak{h}(t),\partial_x^j\mathfrak{h}(t)]\\ &\qquad+D_\epsilon[\partial_x^k\mathfrak{h}(t),\partial_x^{k+2-j}\mathfrak{h}(t),\partial_x^j\partial_t\mathfrak{h}(t)].\end{aligned}$$

Note that $E_{N1}(t)$ is a linear combination of the last three lines on the right-hand side, with ∂_t replaced with Λ_ϵ, so $(d/dt)\sum_{j=2}^{[k/2]+1}c_{jk}D_{1,k,j}(t)-E_{N1}(t)$ is a linear combination of

$$\epsilon'(t)(\partial_\epsilon D_\epsilon)[\partial_x^k\mathfrak{h}(t),\partial_x^{k+2-j}\mathfrak{h}(t),\partial_x^j\mathfrak{h}(t)], \tag{5-7}$$

$$D_\epsilon[\partial_x^k(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t),\partial_x^{k+2-j}\mathfrak{h}(t),\partial_x^j\mathfrak{h}(t)], \tag{5-8}$$

$$D_\epsilon[\partial_x^k\mathfrak{h}(t),\partial_x^{k+2-j}(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t),\partial_x^j\mathfrak{h}(t)], \tag{5-9}$$

$$D_\epsilon[\partial_x^k\mathfrak{h}(t),\partial_x^{k+2-j}\mathfrak{h}(t),\partial_x^j(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t)]. \tag{5-10}$$

We estimate these terms one by one.

By the definition of D_ϵ,

$$\begin{aligned}(5\text{-}7)=\epsilon'(t)\sum_{mnl\neq 0}&\frac{(\lambda_m'(\epsilon)+\lambda_n'(\epsilon)+\lambda_l'(\epsilon))}{2(\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon))^2}\\ &\times\int_0^{2\pi}\hat{\mathfrak{h}}(m+\operatorname{sgn}m,t)\partial_x^k e^{i(m+\operatorname{sgn}m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn}n,t)\partial_x^{k+2-j}e^{i(n+\operatorname{sgn}n)x}\hat{\mathfrak{h}}(l+\operatorname{sgn}l,t)\partial_x^j e^{i(l+\operatorname{sgn}l)x}\,dx.\end{aligned}$$

We first bound the fraction. By Corollary 3.10, when ϵ is small enough,

$$\begin{aligned}\lambda'_m(\epsilon)+\lambda'_n(\epsilon)+\lambda'_l(\epsilon) &= (m+n+l+\operatorname{sgn} m+\operatorname{sgn} n+\operatorname{sgn} l)\partial_\epsilon c_\epsilon i+O(\epsilon^5)\\ &\lesssim (|m+n+l|+1)|\epsilon|.\end{aligned} \tag{5-11}$$

On the other hand, the integral vanishes unless

$$m+n+l+\operatorname{sgn} m+\operatorname{sgn} n+\operatorname{sgn} l=0, \tag{5-12}$$

in which case $m+n+l$ is an odd number, and so is nonzero. Then by Case 1 of Proposition 3.11,

$$|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|>\tfrac{1}{2}|m+n+l|, \tag{5-13}$$

so

$$\frac{|\lambda'_m(\epsilon)+\lambda'_n(\epsilon)+\lambda'_l(\epsilon)|}{|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|^2}\lesssim|\epsilon|. \tag{5-14}$$

Then for $k\ge 3$,

$$\begin{aligned}|(5\text{-}7)| &\lesssim |\epsilon(t)\epsilon'(t)|\sum_{\substack{mnl\neq 0\\ (5\text{-}12)}}|(m+\operatorname{sgn} m)^k\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)\\ &\qquad\qquad\times(n+\operatorname{sgn} n)^{k+2-j}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)(l+\operatorname{sgn} l)^j\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)|\\ &\approx |\epsilon(t)\epsilon'(t)|\left|\int_0^{2\pi}\partial_x^k H(x,t)\partial_x^{k+2-j}H(x,t)\partial_x^j H(x,t)\right|dx\\ &\lesssim_k |\epsilon(t)\epsilon'(t)|\,\|H(x,t)\|_{H_x^k}^2\,\|H(x,t)\|_{W_x^{[k/2]+1,\infty}}\lesssim_k |\epsilon(t)\epsilon'(t)|\,\|H(x,t)\|_{H_x^k}^3\end{aligned} \tag{5-15}$$

since $k\ge[k/2]+2$, where

$$H(x,t)=\sum_{m\neq 0}|\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)|e^{i(m+\operatorname{sgn} m)x}$$

satisfies

$$\|H(x,t)\|_{H_x^k}=\|\mathfrak{h}(t)\|_{\dot H^k}\lesssim E_k(t)^{1/2}$$

so by (4-9) and (5-1),

$$|(5\text{-}7)|\lesssim_k|\epsilon|E_2(t)E_k(t)^{3/2}. \tag{5-16}$$

To bound the other terms (5-8), (5-9) and (5-10), we use the evolution equation (4-11) of $\mathfrak{h}$, which loses one derivative in $\mathfrak{h}$, so

$$\|(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t)\|_{\dot H^{k-1}}\lesssim(\|g\|_{H^1}^2/|\epsilon|)\|\mathfrak{h}(t)\|_{\dot H^k}+\|\mathfrak{h}(t)\|_{\dot H^k}^2.$$

If $\|g(t)\|_{H^1}/|\epsilon|$ is small enough and $k\ge 2$, the first term is dominated by the second term thanks to (5-1). Since in the summation of D_ϵ it holds that $m+n+l\neq 0$, the denominator is uniformly bounded from below thanks to (5-13). Unless $j=2$ in (5-8) and (5-9), we can integrate by parts if necessary to ensure that at most $k-1$ derivatives in x hit each factor of j. Then similarly to (5-15) it follows that for $k\ge 5$,

$$|((5\text{-}8),\,j\ge 3)+((5\text{-}9),\,j\ge 3)+(5\text{-}10)|\lesssim_k E_k(t)^2. \tag{5-17}$$

For $j=2$, by symmetry of D_ϵ it is clear that

$$((5\text{-}9),\ j=2) = ((5\text{-}8),\ j=2), \tag{5-18}$$

which according to (4-11) equals

$$D_\epsilon[\partial_x^k(\Phi_\epsilon \mathfrak{h}_x(t) + N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)] + R(t)),\ \partial_x^k\mathfrak{h}(t),\ \partial_x^2\mathfrak{h}(t)].$$

Similarly to (5-15),

$$|D_\epsilon[\partial_x^k R(t),\ \partial_x^k\mathfrak{h}(t),\ \partial_x^2\mathfrak{h}(t)]| \lesssim_k E_3(t)^2 E_k(t)^{1/2}. \tag{5-19}$$

Similarly to (5-3),

$$|D_\epsilon[\partial_x^k(\Phi_\epsilon \mathfrak{h}_x(t)) - \Phi_\epsilon\partial_x^{k+1}\mathfrak{h}(t),\ \partial_x^k\mathfrak{h}(t),\ \partial_x^2\mathfrak{h}(t)]| \lesssim_k E_3(t)^{3/2} E_k(t). \tag{5-20}$$

By the definition of D_ϵ,

$$\begin{aligned} &D_\epsilon[\Phi_\epsilon\partial_x^{k+1}\mathfrak{h}_x(t),\ \partial_x^k\mathfrak{h}(t),\ \partial_x^2\mathfrak{h}(t)] \\ &= \sum_{mm'nl\neq 0} \frac{1}{\lambda_{m'}(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)} \int_0^{2\pi} \widehat{\Phi}_\epsilon(p)e^{ipx}\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)\partial_x^{k+1}e^{i(m+\operatorname{sgn} m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t) \\ &\qquad\qquad \times \partial_x^k e^{i(n+\operatorname{sgn} n)x}\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\partial_x^2 e^{i(l+\operatorname{sgn} l)x}\,dx, \end{aligned} \tag{5-21}$$

where $m'+\operatorname{sgn} m' = p+m+\operatorname{sgn} m \neq 0, \pm 1$. We break the summation into several parts.

<u>Part 1</u>: $|p| \geq \frac{1}{3}|m+\operatorname{sgn} m|$. Then we can transfer the extra derivative from $\mathfrak{h}$ to Φ_ϵ, and compute as in (5-3) to get

$$|\text{Part 1}| \lesssim_k E_3(t)^{3/2}E_k(t). \tag{5-22}$$

<u>Part 2</u>: $|p| < |m+\operatorname{sgn} m|/3$ but $|p| \geq |n+\operatorname{sgn} n|/3$. If $|n+\operatorname{sgn} n| \geq |m|/3$ then $|p| \geq |m|/9$, and we get the same bound as before. Otherwise, since the integral vanishes unless

$$p+m+n+l+\operatorname{sgn} m+\operatorname{sgn} n+\operatorname{sgn} l = 0 \tag{5-23}$$

in which case we have $|l+\operatorname{sgn} l| > |n+\operatorname{sgn} n|/3$, we can transfer the extra derivative to the factor $\partial_x^2\mathfrak{h}$ to get (note that $\|\Phi_\epsilon\|_{C^k} \lesssim_k \|g\|_{H^1}^2/|\epsilon|$)

$$|\text{Part 2}| \lesssim_k (\|g(t)\|_{H^1}^2/|\epsilon|)E_4(t)^{1/2}E_k(t) \lesssim E_4(t)E_k(t) \tag{5-24}$$

provided that $\|g(t)\|_{H^1}/|\epsilon|$ is small enough.

<u>Part 3</u>: $|p| < \frac{1}{3}|m+\operatorname{sgn} m|$ and $|p| < \frac{1}{3}|n+\operatorname{sgn} n|$. Then $\operatorname{sgn}(m'+\operatorname{sgn} m') = \operatorname{sgn}(m+\operatorname{sgn} m)$, i.e., $\operatorname{sgn} m' = \operatorname{sgn} m$, so $m' = m+p$. By symmetry,

$$\begin{aligned} \text{Part 3} = &\sum_{\substack{mnl\neq 0\\ |p|<|m+\operatorname{sgn} m|/3\\ |p|<|n+\operatorname{sgn} n|/3}} \frac{\frac{1}{2}}{\lambda_{m+p}(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)} \int_0^{2\pi} \hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\partial_x^2 e^{i(l+\operatorname{sgn} l)x}\widehat{\Phi}_\epsilon(p)e^{ipx}\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t) \\ &\qquad\qquad \times\partial_x^{k+1}e^{i(m+\operatorname{sgn} m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)\partial_x^k e^{i(n+\operatorname{sgn} n)x}\,dx \\ &+ \sum_{\substack{mnl\neq 0\\ |p|<|m+\operatorname{sgn} m|/3\\ |p|<|n+\operatorname{sgn} n|/3}} \frac{\frac{1}{2}}{\lambda_m(\epsilon)+\lambda_{n+p}(\epsilon)+\lambda_l(\epsilon)} \int_0^{2\pi} \hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\partial_x^2 e^{i(l+\operatorname{sgn} l)x}\widehat{\Phi}_\epsilon(p)e^{ipx}\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t) \\ &\qquad\qquad \times\partial_x^k e^{i(m+\operatorname{sgn} m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)\partial_x^{k+1}e^{i(n+\operatorname{sgn} n)x}\,dx. \end{aligned}$$

Note that the two denominators are uniformly bounded from below. Also, $\operatorname{sgn}(m+p) = \operatorname{sgn} m$ and $|m+p| > \frac{1}{3}(2|m|-1)$, and similarly for l. Then by Corollary 3.10, the two denominators differ by $O(|m|\epsilon^{4|m|/3+3} + |n|\epsilon^{4|n|/3+3})$, so

$$\begin{aligned}\text{Part } 3 = &\sum_{\substack{mnl\neq 0\\ |p|<|m+\operatorname{sgn} m|/3\\ |p|<|n+\operatorname{sgn} n|/3}} \frac{\frac{1}{2}}{\lambda_{m+p}(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)} \int_0^{2\pi} \hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\partial_x^2 e^{i(l+\operatorname{sgn} l)x}\widehat{\Phi}_\epsilon(p)e^{ipx}\partial_x(\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)\\ &\qquad\times \partial_x^k e^{i(m+\operatorname{sgn} m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)\partial_x^k e^{i(n+\operatorname{sgn} n)x})\,dx\\ &+\sum_{(5\text{-}23)} O(|m|\epsilon^{4|m|/3+3}+|n|\epsilon^{4|n|/3+3})\int_0^{2\pi} |(l+\operatorname{sgn} l)^2\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\widehat{\Phi}_\epsilon(p)(m+\operatorname{sgn} m)^k\\ &\qquad\times\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)(n+\operatorname{sgn} n)^{k+1}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)|\\ = &\sum_{\substack{mnl\neq 0\\ |p|<|m+\operatorname{sgn} m|/3\\ |p|<|n+\operatorname{sgn} n|/3}} \frac{-\frac{1}{2}}{\lambda_{m+p}(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)} \int_0^{2\pi} \partial_x(\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\partial_x^2 e^{i(l+\operatorname{sgn} l)x}\widehat{\Phi}_\epsilon(p)e^{ipx})\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)\\ &\qquad\times \partial_x^k e^{i(m+\operatorname{sgn} m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)\partial_x^k e^{i(n+\operatorname{sgn} n)x}\,dx\\ &+\sum_{(5\text{-}23)} O(\epsilon^4)\int_0^{2\pi} |(l+\operatorname{sgn} l)^2\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\\ &\qquad\times\widehat{\Phi}_\epsilon(p)(m+\operatorname{sgn} m)^k\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)(n+\operatorname{sgn} n)^k\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)|,\end{aligned}$$

where we integrated by parts in the first integral and used the bounds $|m|\epsilon^{4|m|/3+3}$ and

$$|n(n+\operatorname{sgn} n)\epsilon^{4|n|/3+3}| \lesssim \epsilon^4$$

in the second. Then as in (5-15) it follows that

$$\begin{aligned}|\text{Part } 3| &\lesssim_k (\|g(t)\|_{H^1}^2/|\epsilon|)E_4(t)^{1/2}E_k(t) + \epsilon^4(\|g(t)\|_{H^1}^2/|\epsilon|)E_3(t)^{1/2}E_k(t)\\ &\lesssim E_4(t)E_k(t)\end{aligned} \tag{5-25}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

Combining (5-20), (5-22), (5-24) and (5-25) shows that

$$|D_\epsilon[\partial_x^k(\Phi_\epsilon\mathfrak{h}_x(t)),\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]| \lesssim_k E_4(t)(1+E_4(t)^{1/2})E_k(t) \tag{5-26}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

We now turn to $D_\epsilon[\partial_x^k N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)],\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]$. Similarly to (5-3),

$$\begin{aligned}&\big|D_\epsilon[\partial_x^k N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)],\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]\\ &-D_\epsilon[(\phi_\epsilon'\circ\phi_\epsilon^{-1})^2\partial_x((1+W_\epsilon)^{-1}\mathfrak{h}(t))(\partial_x^{k+1}(1+W_\epsilon)^{-1}\mathfrak{h}(t)),\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]\big| \lesssim_k |\epsilon|E_3(t)^{1/2}E_k(t)^{3/2}.\end{aligned} \tag{5-27}$$

Since W_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ, so is $((1+W_\epsilon)^{-1}-1)/\epsilon$, so

$$\begin{aligned}&|D_\epsilon[(\phi_\epsilon'\circ\phi_\epsilon^{-1})^2(\partial_x(1+W_\epsilon)^{-1}\mathfrak{h}(t))(\partial_x^{k+1}((1+W_\epsilon)^{-1}-1)\mathfrak{h}(t)),\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]|\\ &\qquad\lesssim_k |\epsilon|E_3(t)^{3/2}E_k(t)^{1/2}.\end{aligned} \tag{5-28}$$

Finally, $D_\epsilon[(\phi'_\epsilon\circ\phi_\epsilon^{-1})^2(\partial_x(1+W_\epsilon)^{-1}\mathfrak{h}(t))(\partial_x^{k+1}\mathfrak{h}(t)),\partial_x^k\mathfrak{h}(t),\partial_x^2\mathfrak{h}(t)]$ is of the same form as the left-hand side of (5-21), so we trace the same argument to get

$$\begin{aligned}|\text{Part }1| &\lesssim_k E_3(t)E_k(t),\\ |\text{Part }2| &\lesssim_k E_4(t)E_k(t),\\ |\text{Part }3| &\lesssim_k E_4(t)E_k(t)+\epsilon^4E_3(t)E_k(t)\lesssim E_4(t)E_k(t)\end{aligned}$$

provided that ϵ is small enough. Hence

$$|D_\epsilon[(\phi'_\epsilon\circ\phi_\epsilon^{-1})^2(\partial_x(1+W_\epsilon)^{-1}\mathfrak{h}(t))(\partial_x^{k+1}\mathfrak{h}(t)),\partial_x^k\mathfrak{h}(t),\partial_x^2\mathfrak{h}(t)]|\lesssim_k E_4(t)E_k(t). \tag{5-29}$$

Combining (5-27), (5-28) and (5-29) shows that, for $k\ge 4$,

$$|D_\epsilon[\partial_x^kN_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)],\partial_x^k\mathfrak{h}(t),\partial_x^2\mathfrak{h}(t)]|\lesssim_k E_4(t)^{1/2}E_k(t)^{3/2} \tag{5-30}$$

provided that ϵ is small enough.

Combining (5-19), (5-26) and (5-30) shows that, for $k\ge 4$,

$$|((5\text{-}8),\ j=2)|\lesssim_k E_4(t)^{1/2}(1+E_4(t)^{1/2})E_k(t)^{3/2} \tag{5-31}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

Finally, combining (5-16), (5-17), (5-18) and (5-31) shows that, for $k\ge 5$,

$$\left|\frac{d}{dt}\sum_{j=2}^{[k/2]+1}c_{jk}D_{1,k,j}(t)-E_{N1}(t)\right|\lesssim_k (1+E_4(t)^{1/2})E_k(t)^2 \tag{5-32}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

5B. *Lifespan when $\delta\ll\epsilon$*. In this section we will obtain a preliminary bound for $E_{N2}=E_N-E_{N1}$ and show a lifespan of $1/(\epsilon\delta)$ when $\|g_0\|_{H^5(\mathbb{T})}=\delta\ll\epsilon$, i.e., $\delta\le c\epsilon$ for some $c>0$ independent of ϵ.

Recall from (5-6) that

$$E_N(t)=\frac12\int_0^{2\pi}\partial_x^k\mathfrak{h}(t)\partial_x^k((\phi'_\epsilon\circ\phi_\epsilon^{-1})^2((1+W_\epsilon)^{-1}\mathfrak{h}(t))_x^2)\,dx.$$

Similarly to (5-3), for $k\ge 3$,

$$\left|E_N(t)-\int_0^{2\pi}(\phi'_\epsilon\circ\phi_\epsilon^{-1})^2\partial_x^k\mathfrak{h}(t)\partial_x^k(((1+W_\epsilon)^{-1}\mathfrak{h}_x(t))^2)\,dx\right|\lesssim_k|\epsilon|E_k(t)^{3/2}.$$

Since $((1+W_\epsilon)^{-1}-1)/\epsilon$ is of class $\mathcal{S}$ uniformly in ϵ,

$$\left|\int_0^{2\pi}(\phi'_\epsilon\circ\phi_\epsilon^{-1})^2\partial_x^k\mathfrak{h}(t)\partial_x^k(((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t))^2)\,dx\right|\lesssim_k\epsilon^2E_k(t)^{3/2},$$

$$\begin{aligned}2\bigg|\int_0^{2\pi}&(\phi'_\epsilon\circ\phi_\epsilon^{-1})^2\partial_x^k\mathfrak{h}(t)\partial_x^k(((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t))\mathfrak{h}_x(t))\,dx\\ &-\int_0^{2\pi}(\phi'_\epsilon\circ\phi_\epsilon^{-1})^2((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t))\partial_x^k\mathfrak{h}(t)\partial_x^{k+1}\mathfrak{h}(t)\,dx\bigg|\lesssim_k|\epsilon|E_k(t)^{3/2}.\end{aligned}$$

Finally, by integration by parts,

$$\left|2\int_0^{2\pi}(\phi_\epsilon'\circ\phi_\epsilon^{-1})^2((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t))\partial_x^k\mathfrak{h}(t)\partial_x^{k+1}\mathfrak{h}(t)\,dx\right|$$
$$=\left|\int_0^{2\pi}\partial_x((\phi_\epsilon'\circ\phi_\epsilon^{-1})^2((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t)))(\partial_x^k\mathfrak{h}(t))^2\,dx\right|\lesssim_k|\epsilon|E_k(t)^{3/2}.$$

Combining the bounds above shows that, for $k\ge 3$,

$$|E_{N_2}(t)|=|E_N(t)-E_{N_1}(t)|\lesssim_k|\epsilon|E_k(t)^{3/2} \tag{5-33}$$

provided that ϵ is small enough.

Now combining (5-5), (5-32) and (5-33) shows that, for $k\ge 5$,

$$\frac{d}{dt}\left|\sum_{j=2}^{[k/2]+1}c_{jk}D_{1,k,j}(t)-E_k(t)\right|\lesssim_k(1+E_4(t)^{1/2})E_k(t)^2+|\epsilon|E_k(t)^{3/2}. \tag{5-34}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough. Hence

$$E_k(t)-E_k(0)=\sum_{j=2}^{[k/2]+1}c_{jk}(D_{1,k,j}(t)-D_{1,k,j}(0))+O_k(\|(1+E_4^{1/2})E_k^2+|\epsilon|E_k^{3/2}\|_{L^1([0,t])}).$$

Similarly to (5-15),

$$|D_{1,k,j}(t)|=|D_{\epsilon(t)}[\partial_x^k\mathfrak{h}(t),\,\partial_x^{k+2-j}\mathfrak{h}(t),\,\partial_x^j\mathfrak{h}(t)]|\lesssim_k E_k(t)^{3/2}.$$

Now we are able to show a lifespan longer than what follows from local well-posedness. Assume that the initial data is

$$f(x,0)=u_\epsilon(x)+g(x),$$

where $|\epsilon|\le\epsilon_0$ is small enough, the energy $E_k(0)$ computed from g is $E_k(0)=\delta^2$, and $|\delta/\epsilon|$ is also small enough. Let

$$T^*=\sup\bigl\{T:\text{there exists a solution } f(x,t)=u_{\epsilon(t)}(x+a(t))+g(x+a(t),t), \tag{5-35}$$
$$t\in[0,T]\text{ such that }\tfrac{1}{2}|\epsilon|\le|\epsilon(t)|\le 2|\epsilon|,\ \ E_k(t)\le 4\delta^2\bigr\}. \tag{5-36}$$

Then the above conditions hold for all $t<T^*$. Moreover, the energy estimate implies

$$E_k(t)=\delta^2+O_k(\delta^3+t(\delta^4+|\epsilon|\delta^3))=\delta^2+O_k(\delta^3(1+t|\epsilon|)).$$

Then there is $c_k>0$ such that if $T^*\le c_k/|\epsilon|\delta$, then $E_k(t)\le 2\delta^2$. Also,

$$|\|f(x,t)\|_{L^2}-\|u_\epsilon\|_{L^2}|=|\|f(x,0)\|_{L^2}-\|u_\epsilon\|_{L^2}|\le\|g\|_{L^2}\lesssim\delta$$

by conservation of the L^2 norm. Meanwhile $|\|f(x,t)\|_{L^2}-\|u_{\epsilon(t)}\|_{L^2}|\lesssim\delta$, so $|\|u_{\epsilon(t)}\|_{L^2}-\|u_\epsilon\|_{L^2}|\lesssim\delta$. When $|\epsilon|$ is small enough, $\|u_\epsilon\|_{L^2}$ is differentiable in ϵ with nonzero derivative at $\epsilon=0$. Since $|\delta/\epsilon|$ is small enough, $|\epsilon(t)-\epsilon|\lesssim\delta$.

By local well-posedness, the solution can be extended to a time $t^* > T^*$, with

$$\|f(x,t) - f(x,T^*)\|_{H^2} \lesssim (t^* - T^*)(\|f(x,t)\|_{H^3} + \|f(x,t)\|_{H^3}^2) \le (t^* - T^*)|\epsilon|$$

for $t \in [T^*, t^*]$. Then $\|f(x,t) - u_{\epsilon(T^*)}(x + a(T^*))\|_{H^2} \lesssim (t^* - T^*)|\epsilon| + \delta$. Take $t^* = T^* + \delta/|\epsilon|$. Then $f(x,t)$ satisfies the conditions in Proposition 4.1, so (5-35) holds up to time t^*. Since $f(x,T^*)$ is small in H^4, $f(x,t)$ is uniformly bounded in H^4 on $[T^*, t^*]$, so it stays within a compact set in H^2. Since ϵ is differentiable in $f \in H^2$, $|\epsilon(t) - \epsilon(T^*)| \lesssim (t^* - T^*)|\epsilon| \lesssim \delta$, so $|\epsilon(t) - \epsilon| \lesssim \delta$, so $|\epsilon|/2 \le |\epsilon(t)| \le 2|\epsilon|$ holds up to time t^*. The energy estimate then implies $E_k \le 3\delta^2$ also up to time t^*, so (5-36) holds up to time t^*, contradicting the definition of T^*. Hence the lifespan $T^* \gtrsim_k 1/|\epsilon|\delta$.

5C. ***Longer lifespan when $\delta \ll \epsilon^2$.*** When the perturbation g is very small compared to ϵ^2, that is, $\|g_0\|_{H^5(\mathbb{T})} = \delta \ll \epsilon^2$, we can obtain a longer lifespan by applying the normal form transformation to

$$E_{N2} = E_N - E_{N1} = E_{N21} + E_{N22} + E_{N23} + E_{N24},$$

where

$$E_{N21} = \sum_{j=1}^{[k/2]+1} c'_{kj} \int_0^{2\pi} \partial_x^k \mathfrak{h}(t) \partial_x^{k+2-j}((\phi'_\epsilon \circ \phi_\epsilon^{-1})((1+W_\epsilon)^{-1} - 1)\mathfrak{h}(t)) \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\mathfrak{h}(t))\, dx,$$

$$E_{N22} = \sum_{j=1}^{[k/2]+1} \sum_{i=1}^{k+2-j} c_{kji} \int_0^{2\pi} \partial_x^k \mathfrak{h}(t) \partial_x^i(\phi'_\epsilon \circ \phi_\epsilon^{-1}) \partial_x^{k+2-i-j}\mathfrak{h}(t) \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\mathfrak{h}(t))\, dx,$$

$$E_{N23} = \sum_{j=2}^{[k/2]+1} c_{kj} \int_0^{2\pi} \partial_x^k \mathfrak{h}(t)(\phi'_\epsilon \circ \phi_\epsilon^{-1} - 1) \partial_x^{k+2-j}\mathfrak{h}(t) \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\mathfrak{h}(t))\, dx,$$

$$E_{N24} = \sum_{j=2}^{[k/2]+1} c_{kj} \int_0^{2\pi} \partial_x^k \mathfrak{h}(t) \partial_x^{k+2-j}\mathfrak{h}(t) \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\mathfrak{h} - \mathfrak{h})\, dx,$$

where c_{kj}, c'_{kj} and $c_{kji} \in \mathbb{R}$ are constants and we integrated by parts to get rid of the terms with $k+1$ derivatives falling on a single factor of $\mathfrak{h}$, except for the term with $j = 1$ in E_{N21}, in which the $k+1$ derivatives do not matter in view of the fact that the operator $(\phi'_\epsilon \circ \phi_\epsilon^{-1})((1+W_\epsilon)^{-1} - 1)$ is of class $\mathcal{S}$.

Now we define

$$\begin{aligned} D_{\epsilon,21}[f_1, f_2, f_3] &= \sum_{\substack{j=1 \\ mnl \neq 0}}^{[k/2]+1} \frac{c'_{kj}}{\lambda_m(\epsilon) + \lambda_n(\epsilon) + \lambda_l(\epsilon)} \int_0^{2\pi} \hat{f}_1(m + \operatorname{sgn} m) e^{i(m + \operatorname{sgn} m)x} \\ &\quad \times \partial_x^{k+2-j}((\phi'_\epsilon \circ \phi_\epsilon^{-1})((1+W_\epsilon)^{-1} - 1)\hat{f}_2(n + \operatorname{sgn} n) e^{i(n + \operatorname{sgn} n)x}) \\ &\quad \times \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\hat{f}_3(l + \operatorname{sgn} l) e^{i(l + \operatorname{sgn} l)x}), \end{aligned}$$

and

$$D_{21}(t) = D_{\epsilon,21}[\mathfrak{h}(t), \mathfrak{h}(t), \mathfrak{h}(t)],$$

and similarly define D_{22}, D_{23} and D_{24}. Then

$$\begin{aligned}\frac{d}{dt}D_{21}(t)-E_{N21}(t)=\epsilon'(t)(\partial_\epsilon D_{\epsilon,21})[\mathfrak{h}(t),\mathfrak{h}(t),\mathfrak{h}(t)] &\quad (5\text{-}37)\\ +D_{\epsilon,21}[(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t),\mathfrak{h}(t),\mathfrak{h}(t)] &\quad (5\text{-}38)\\ +D_{\epsilon,21}[\mathfrak{h}(t),(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t),\mathfrak{h}(t)] &\quad (5\text{-}39)\\ +D_{\epsilon,21}[\mathfrak{h}(t),\mathfrak{h}(t),(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t)]. &\quad (5\text{-}40)\end{aligned}$$

We estimate these terms one by one.

For (5-37), (5-11) still holds, but there are nontrivial actions on $\mathfrak{h}$ in the slots, so no frequency restriction such as (5-12) exists. When $m+n+l\neq 0$, we are in Case 1 of Proposition 3.11, so (5-13), and hence (5-14), still hold. When $m+n+l=0$, by Case 2 of Proposition 3.11, when ϵ is small enough,

$$|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|>\tfrac{1}{5}\epsilon^2, \tag{5-41}$$

which, combined with (5-11), shows that the multiplier in $\partial_\epsilon D_\epsilon$ is bounded by

$$\frac{|\lambda_m'(\epsilon)+\lambda_n'(\epsilon)+\lambda_l'(\epsilon)|}{|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|^2}\lesssim|\epsilon|^{-3} \tag{5-42}$$

instead of (5-14). Since both $(\phi_\epsilon'\circ\phi_\epsilon^{-1})((1+W_\epsilon)^{-1}-1)/\epsilon$ and $\partial_\epsilon(\phi_\epsilon'\circ\phi_\epsilon^{-1})((1+W_\epsilon)^{-1}-1)$ are of class $\mathcal{S}$ uniformly in ϵ, it follows that, for $k\geq 3$,

$$|(5\text{-}37)|\lesssim_k|\epsilon'(t)|\epsilon^{-2}E_k(t)^{3/2}\lesssim\epsilon^{-2}E_2(t)E_k(t)^{3/2} \tag{5-43}$$

provided that ϵ is small enough.

The terms (5-38), (5-39) and (5-40) are like (5-8), (5-9) and (5-10) respectively, except that instead of the uniform lower bound of $\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)$ we now have (5-41), which loses two factors of ϵ, but we are helped by the ϵ-smallness of $(\phi_\epsilon'\circ\phi_\epsilon^{-1})((1+W_\epsilon)^{-1}-1)$, which wins back a factor of ϵ. All told we lose a factor of ϵ compared to (5-32), so, for $k\geq 5$,

$$|(5\text{-}38)+(5\text{-}39)+(5\text{-}40)|\lesssim_k|\epsilon|^{-1}(1+E_4(t)^{1/2})E_k(t)^2 \tag{5-44}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

Combining (5-43) and (5-44) shows that, for $k\geq 5$,

$$\left|\frac{d}{dt}D_{21}(t)-E_{N21}(t)\right|\lesssim_k|\epsilon|^{-1}(1+E_4(t)^{1/2})E_k(t)^2 \tag{5-45}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough. We can also save a factor of ϵ in the other terms E_{N22}, E_{N23} and E_{N24} thanks to the ϵ-smallness of $(\phi_\epsilon'\circ\phi_\epsilon^{-1})'$ and $\phi_\epsilon'\circ\phi_\epsilon^{-1}-1$. Hence the bound (5-45) also holds for E_{N22}, E_{N23} and E_{N24}.

Combining (5-5), (5-32) and (5-45) shows that, for $k\geq 5$,

$$E_k(t)-E_k(0)=\sum_{j=2}^{[k/2]+1}c_{jk}(D_{1,k,j}(t)-D_{1,k,j}(0))+\sum_{j=1}^{4}(D_{2j}(t)-D_{2j}(0))\\ +O_k(|\epsilon|^{-1}\|(1+E_4^{1/2})E_k^2\|_{L^1([0,t])})$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough. Similarly to (5-33), for $k \geq 3$,

$$|D_{2,k,j}(t)| \lesssim_k \epsilon(t)^{-2}|\epsilon(t)|E_k(t)^{3/2} = E_k(t)^{3/2}/|\epsilon|.$$

Hence if $E_k(0) = \delta^2 \lesssim 1$ and $E_k \leq 2\delta^2$ on $[0, t]$ then

$$E_k(t) = \delta^2 + |\epsilon|^{-1}\delta^3 + O_k(t|\epsilon|^{-1}\delta^4).$$

Assume δ/ϵ^2 is small. Then the second term on the right-hand side is $\lesssim \delta^{5/2}$, so we close the estimate for a time $t \lesssim_k |\epsilon|/\delta^2$, which is also the lifespan in this case.

Acknowledgements

This work is supported in part by the Spanish Ministry of Science and Innovation, through the "Severo Ochoa Programme for Centres of Excellence in R&D" (CEX2019-000904-S) and MTM2017-89976-P. Castro was partially supported by the Europa Excelencia program ERC2018-092824. Córdoba and Zheng were partially supported by the ERC Advanced Grant 788250.

References

[Agmon 1975] S. Agmon, "Spectral properties of Schrödinger operators and scattering theory", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **2**:2 (1975), 151–218. MR Zbl

[Arnold 1983] V. I. Arnold, *Geometrical methods in the theory of ordinary differential equations*, Grundlehren der Math. Wissenschaften **250**, Springer, 1983. MR Zbl

[Biello and Hunter 2010] J. Biello and J. K. Hunter, "Nonlinear Hamiltonian waves with constant frequency and surface waves on vorticity discontinuities", *Comm. Pure Appl. Math.* **63**:3 (2010), 303–336. MR Zbl

[Bressan and Nguyen 2014] A. Bressan and K. T. Nguyen, "Global existence of weak solutions for the Burgers–Hilbert equation", *SIAM J. Math. Anal.* **46**:4 (2014), 2884–2904. MR Zbl

[Bressan and Zhang 2017] A. Bressan and T. Zhang, "Piecewise smooth solutions to the Burgers–Hilbert equation", *Commun. Math. Sci.* **15**:1 (2017), 165–184. MR Zbl

[Castro et al. 2010] A. Castro, D. Córdoba, and F. Gancedo, "Singularity formations for a surface wave model", *Nonlinearity* **23**:11 (2010), 2835–2847. MR Zbl

[Chen and Pusateri 2022] G. Chen and F. Pusateri, "The 1-dimensional nonlinear Schrödinger equation with a weighted L^1 potential", *Anal. PDE* **15**:4 (2022), 937–982. MR Zbl

[Crandall and Rabinowitz 1971] M. G. Crandall and P. H. Rabinowitz, "Bifurcation from simple eigenvalues", *J. Funct. Anal.* **8** (1971), 321–340. MR Zbl

[Ehrnström and Wang 2019] M. Ehrnström and Y. Wang, "Enhanced existence time of solutions to the fractional Korteweg–de Vries equation", *SIAM J. Math. Anal.* **51**:4 (2019), 3298–3323. MR Zbl

[Germain et al. 2012] P. Germain, N. Masmoudi, and J. Shatah, "Global solutions for the gravity water waves equation in dimension 3", *Ann. of Math.* (2) **175**:2 (2012), 691–754. MR Zbl

[Germain et al. 2015] P. Germain, N. Masmoudi, and J. Shatah, "Global existence for capillary water waves", *Comm. Pure Appl. Math.* **68**:4 (2015), 625–687. MR Zbl

[Germain et al. 2018] P. Germain, F. Pusateri, and F. Rousset, "The nonlinear Schrödinger equation with a potential", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **35**:6 (2018), 1477–1530. MR Zbl

[Hunter 2018] J. K. Hunter, "The Burgers–Hilbert equation", pp. 41–57 in *Theory, numerics and applications of hyperbolic problems, II* (Aachen, Germany, 2016), edited by C. Klingenberg and M. Westdickenberg, Springer Proc. Math. Stat. **237**, Springer, 2018. MR Zbl

[Hunter and Ifrim 2012] J. K. Hunter and M. Ifrim, "Enhanced life span of smooth solutions of a Burgers–Hilbert equation", *SIAM J. Math. Anal.* **44**:3 (2012), 2039–2052. MR Zbl

[Hunter et al. 2015] J. K. Hunter, M. Ifrim, D. Tataru, and T. K. Wong, "Long time solutions for a Burgers–Hilbert equation via a modified energy method", *Proc. Amer. Math. Soc.* **143**:8 (2015), 3407–3412. MR Zbl

[Hunter et al. 2022] J. K. Hunter, R. C. Moreno-Vasquez, J. Shu, and Q. Zhang, "On the approximation of vorticity fronts by the Burgers–Hilbert equation", *Asymptot. Anal.* **129**:2 (2022), 141–177. MR Zbl

[Ionescu and Pausader 2022] A. D. Ionescu and B. Pausader, *The Einstein–Klein–Gordon coupled system: global stability of the Minkowski solution*, Ann. of Math. Stud. **213**, Princeton Univ. Press, 2022. MR Zbl

[Ionescu and Pusateri 2015] A. D. Ionescu and F. Pusateri, "Global solutions for the gravity water waves system in 2D", *Invent. Math.* **199**:3 (2015), 653–804. MR Zbl

[Ionescu and Pusateri 2018] A. D. Ionescu and F. Pusateri, *Global regularity for 2D water waves with surface tension*, Mem. Amer. Math. Soc. **1227**, Amer. Math. Soc., Providence, RI, 2018. MR Zbl

[Kato 1976] T. Kato, *Perturbation theory for linear operators*, 2nd ed., Grundlehren der Math. Wissenschaften **132**, Springer, 1976. MR Zbl

[Klein and Saut 2015] C. Klein and J.-C. Saut, "A numerical approach to blow-up issues for dispersive perturbations of Burgers' equation", *Phys. D* **295/296** (2015), 46–65. MR Zbl

[Krupa and Vasseur 2020] S. G. Krupa and A. F. Vasseur, "Stability and uniqueness for piecewise smooth solutions to a nonlocal scalar conservation law with applications to Burgers–Hilbert equation", *SIAM J. Math. Anal.* **52**:3 (2020), 2491–2530. MR Zbl

[Marsden and Weinstein 1983] J. Marsden and A. Weinstein, "Coadjoint orbits, vortices, and Clebsch variables for incompressible fluids", *Phys. D* **7**:1-3 (1983), 305–323. MR Zbl

[Saut and Wang 2022] J.-C. Saut and Y. Wang, "The wave breaking for Whitham-type equations revisited", *SIAM J. Math. Anal.* **54**:2 (2022), 2295–2319. MR Zbl

[Shatah 1985] J. Shatah, "Normal forms and quadratic nonlinear Klein–Gordon equations", *Comm. Pure Appl. Math.* **38**:5 (1985), 685–696. MR Zbl

[Yang 2021] R. Yang, "Shock formation of the Burgers–Hilbert equation", *SIAM J. Math. Anal.* **53**:5 (2021), 5756–5802. MR Zbl

Received 3 Mar 2021. Revised 22 Mar 2022. Accepted 29 Apr 2022.

ÁNGEL CASTRO: angel_castro@icmat.es
Instituto de Ciencias Matemáticas ICMAT-CSIC-UAM-UCM-UC3M, Madrid, Spain

DIEGO CÓRDOBA: dcg@icmat.es
Instituto de Ciencias Matemáticas ICMAT-CSIC-UAM-UCM-UC3M, Madrid, Spain

FAN ZHENG: fan.zheng@icmat.es
Instituto de Ciencias Matemáticas ICMAT-CSIC-UAM-UCM-UC3M, Madrid, Spain

Analysis & PDE
msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 16 No. 9 2023