On the existence time for the Kirchhoff equation with periodic boundary conditions

Pietro Baldi and Emanuele Haus

Abstract. We consider the Cauchy problem for the Kirchhoff equation on $\mathbb{T}^d$ with initial data of small amplitude ε in Sobolev class. We prove a lower bound ε^{-4} for the existence time, which improves the bound ε^{-2} given by the standard local theory. The proof relies on a normal form transformation, preceded by a nonlinear transformation that diagonalizes the operator at the highest order, which is needed because of the quasilinear nature of the equation.

Keywords. Kirchhoff equation, quasilinear wave equations, Cauchy problems, normal forms, quasilinear normal forms, long-time existence.
MSC2010: 35L72, 35L15, 35Q74, 37J40, 70K45.

Contents.

1 Introduction

This paper deals with an old open problem, concerning the global wellposedness of the Kirchhoff equation

$$\partial_{tt} u - \Big(1 + \int_\Omega |\nabla u|^2 \, dx\Big) \Delta u = 0 \tag{1.1}$$

with periodic boundary conditions $\Omega = \mathbb{T}^d$ or Dirichlet boundary conditions $u|_{\partial\Omega} = 0$ on a bounded domain $\Omega \subset \mathbb{R}^d$. In 1940 Bernstein [13] proved that, in the 1-dimensional case $\Omega = [0, \pi]$ with zero boundary conditions $u(t, 0) = u(t, \pi) = 0$, the Cauchy problem for (1.1) with initial data

$$u(0, x) = \alpha(x), \quad \partial_t u(0, x) = \beta(x) \tag{1.2}$$

is globally wellposed for (α, β) analytic, and locally wellposed for (α, β) in the Sobolev space $H^2 \times H^1$. Later on, these results have been extended to higher dimension, also including the periodic setting $\Omega = \mathbb{T}^d$, proving global wellposedness in larger spaces containing the analytic functions, and local wellposedness in the Sobolev space $H^{\frac{3}{2}} \times H^{\frac{1}{2}}$, with existence time $T \sim (\|\alpha\|_{\frac{3}{2}} + \|\beta\|_{\frac{1}{2}})^{-2}$, see Section 1.3 for a short overview. Nonetheless, the basic question about the existence time for the Cauchy problem with C^∞ data, even of small amplitude, is still open. In particular, it is still not known whether the maximal existence time is finite or infinite (notice that the quasilinear wave equation $u_{tt} - (1 + u_x^2)u_{xx} = 0$ on the circle $\mathbb{T}$, which looks like (1.3) in one dimension without the integral sign, has a finite blowup time $T \sim (\|\alpha\|_{C^2} + \|\beta\|_{C^1})^{-2}$, as proved by Klainerman and Majda [37]).

In this paper we prove that in the periodic setting $\Omega = \mathbb{T}^d$, $d \geq 1$, for small amplitude initial data $(\alpha, \beta) \in H^{\frac{3}{2}} \times H^{\frac{1}{2}}$ if $d = 1$, or $(\alpha, \beta) \in H^2 \times H^1$ if $d \geq 2$, the existence time is at least $T \sim (\|\alpha\| + \|\beta\|)^{-4}$ (Theorem 1.1), which is longer than the time $(\|\alpha\| + \|\beta\|)^{-2}$ provided by the classical local theory. The same result also holds in the case of zero Dirichlet boundary conditions on the cube $\Omega = [0, \pi]^d$ (Remark 1.4). To give a precise statement of our main result, we first introduce some notation.

On the torus $\mathbb{T}^d$, it is not restrictive to set the problem in the space of functions with zero average in space, for the following reason. Given initial data $\alpha(x), \beta(x)$, we split both them and the unknown $u(t, x)$ into the sum of a zero-mean function and the average term,

$$\alpha(x) = \alpha_0 + \tilde{\alpha}(x), \quad \beta(x) = \beta_0 + \tilde{\beta}(x), \quad u(t, x) = u_0(t) + \tilde{u}(t, x),$$

where

$$\int_{\mathbb{T}^d} \tilde{\alpha}(x)\, dx = 0, \quad \int_{\mathbb{T}^d} \tilde{\beta}(x)\, dx = 0, \quad \int_{\mathbb{T}^d} \tilde{u}(t, x)\, dx = 0 \quad \forall t.$$

Then the Cauchy problem

$$\partial_{tt} u - \Big(1 + \int_{\mathbb{T}^d} |\nabla u|^2\, dx\Big)\Delta u = 0, \quad u(0, x) = \alpha(x), \quad \partial_t u(0, x) = \beta(x) \tag{1.3}$$

splits into two distinct, uncoupled Cauchy problems: one is the problem for the average $u_0(t)$, which is

$$u_0''(t) = 0, \quad u_0(0) = \alpha_0, \quad u_0'(0) = \beta_0$$

and has the unique solution $u_0(t) = \alpha_0 + \beta_0 t$; the other one is the problem for the zero-mean component $\tilde{u}(t, x)$, which is

$$\partial_{tt} \tilde{u} - \Big(1 + \int_{\mathbb{T}^d} |\nabla \tilde{u}|^2\, dx\Big)\Delta \tilde{u} = 0, \quad \tilde{u}(0, x) = \tilde{\alpha}(x), \quad \partial_t \tilde{u}(0, x) = \tilde{\beta}(x).$$

Thus one has to study the Cauchy problem for the zero-mean unknown $\tilde{u}(t, x)$ with zero-mean initial data $\tilde{\alpha}(x), \tilde{\beta}(x)$; this means to study (1.3) in the class of functions with zero average in x.

For any real $s \geq 0$, we consider the Sobolev space of zero-mean functions

$$H_0^s(\mathbb{T}^d, \mathbb{C}) := \Big\{ u(x) = \sum_{j \in \mathbb{Z}^d \setminus \{0\}} u_j e^{ij\cdot x} : u_j \in \mathbb{C},\ \|u\|_s < \infty \Big\}, \tag{1.4}$$

$$\|u\|_s^2 := \sum_{j \in \mathbb{Z}^d \setminus \{0\}} |u_j|^2 |j|^{2s},$$

and its subspace

$$H_0^s(\mathbb{T}^d, \mathbb{R}) := \{u \in H_0^s(\mathbb{T}^d, \mathbb{C}) : u(x) \in \mathbb{R}\}$$

of real-valued functions. For $s = 0$, we write L_0^2 instead of H_0^0 the space of square-integrable functions with zero average.

The main result of the paper is the following theorem.

Theorem 1.1. *For $d \in \mathbb{N}$, let*

$$m_0 = 1 \quad \text{if} \;\; d = 1, \qquad m_0 = \frac{3}{2} \quad \text{if} \;\; d \geq 2. \tag{1.5}$$

There exist universal constants $\varepsilon_0, C, C_1 > 0$ with the following properties.

If $(\alpha, \beta) \in H_0^{m_0+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{m_0-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$ with

$$\varepsilon := \|\alpha\|_{m_0+\frac{1}{2}} + \|\beta\|_{m_0-\frac{1}{2}} \leq \varepsilon_0,$$

then the Cauchy problem (1.3) has a unique solution $u \in C^0([0,T], H_0^{m_0+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})) \cap C^1([0,T], H_0^{m_0-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}))$ on the time interval $[0,T]$, where

$$T = \frac{C_1}{\varepsilon^4},$$

and

$$\max_{t \in [0,T]} (\|u(t)\|_{m_0+\frac{1}{2}} + \|\partial_t u(t)\|_{m_0-\frac{1}{2}}) \leq C\varepsilon.$$

If, in addition, $(\alpha, \beta) \in H_0^{s+\frac{1}{2}}(\mathbb{T}, \mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}, \mathbb{R})$ for some $s \geq m_0$, then u belongs to $C^0([0,T], H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})) \cap C^1([0,T], H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}))$, with

$$\max_{t \in [0,T]} (\|u(t)\|_{s+\frac{1}{2}} + \|\partial_t u(t)\|_{s-\frac{1}{2}}) \leq C(\|\alpha\|_{s+\frac{1}{2}} + \|\beta\|_{s-\frac{1}{2}}). \tag{1.6}$$

Remark 1.2 (Evolution of higher norms)**.** The constant C in (1.6) does not depend on s. This unusual property is a consequence of the special structure of the Kirchhoff equation: if u is a solution of (1.1), then u also solves the linear wave equation with time-dependent coefficient $\partial_{tt} u - a(t)\Delta u = 0$, with $a(t) = 1 + \int_\Omega |\nabla u|^2 \, dx$, and therefore $v := |D_x|^s u$ also solves $\partial_{tt} v - a(t)\Delta v = 0$. □

Remark 1.3 (Why m_0 in (1.5) is different in dimension $d = 1$ and $d \geq 2$)**.** The proof of Theorem 1.1 is based on a normal form transformation. In the construction of such a normal form, one encounters the differences of the linear eigenvalues $|j|$, $j \in \mathbb{Z}^d$, as denominators of the transformation coefficients (see (4.12)-(4.13)). On the 1-dimensional torus $\mathbb{T}$, the difference $||j| - |k||$ is either zero or ≥ 1, while on $\mathbb{T}^d$, $d \geq 2$, the differences $||j| - |k|| = |\sqrt{j_1^2 + \ldots + j_d^2} - \sqrt{k_1^2 + \ldots + k_d^2}|$ accumulate to zero, with lower bounds $||j| - |k|| \geq \frac{1}{|j|+|k|}$. This is the reason for the different regularity threshold we obtain in dimension 1 or higher. □

Remark 1.4 (Dirichlet boundary conditions on the cube)**.** Theorem 1.1 immediately implies a similar result for the Cauchy problem with zero Dirichlet boundary conditions on the cube $\Omega := [0,\pi]^d$. Given any function $u : \Omega \to \mathbb{R}$, let

$$U : [-\pi,\pi]^d \to \mathbb{R}, \quad U(x) := \text{sign}(x_1 x_2 \cdots x_d) u(|x_1|, \ldots, |x_d|)$$

be its extension by odd reflection, and let $u_{ext} : \mathbb{T}^d \to \mathbb{R}$ be the periodic extension of U.

A function u belongs to $H^s(\Omega)$, $s = 1$ or $s = 2$ (i.e. the weak partial derivatives of order $\leq s$ belong to $L^2(\Omega)$) with Dirichlet boundary condition $u = 0$ on the boundary $\partial\Omega$ if and only if (see, e.g., [27], [3]) u belongs to the domain $V_s(\Omega)$ of the fractional Laplacian $(-\Delta)^{s/2}$ on Ω with zero Dirichlet boundary conditions (a spectrally defined Sobolev space). In such a case, the extension u_{ext} belongs to the Sobolev space $H^s_0(\mathbb{T}^d)$ defined in (1.4). Hence, for initial data $\alpha \in H^2(\Omega)$, $\beta \in H^1(\Omega)$ with $\alpha = \beta = 0$ on $\partial\Omega$, one consider the periodic odd extensions $\alpha_{ext} \in H^2_0(\mathbb{T}^d)$, $\beta_{ext} \in H^1_0(\mathbb{T}^d)$, and Theorem 1.1 applies.

In dimension $d = 1$, Theorem 1.1 requires less regularity, and it is sufficient that $\alpha \in V_{\frac{3}{2}}(0,\pi)$ and $\beta \in V_{\frac{1}{2}}(0,\pi)$. One has $\alpha \in V_{\frac{3}{2}}(0,\pi)$ if and only if α belongs to the fractional Sobolev space $H^{\frac{3}{2}}(0,\pi)$ on the interval, with $\alpha(0) = \alpha(\pi) = 0$, while $\beta \in V_{\frac{1}{2}}(0,\pi)$ if and only if $\beta \in H^{\frac{1}{2}}(0,\pi)$ with $\int_0^\pi \frac{|\beta(x)|^2}{x(\pi - x)}\, dx < \infty$ (see [27], [3]). □

1.1 Strategy of the proof

Since the problem is set on the torus $\mathbb{T}^d$, which is a compact manifold, no dispersive estimates are available to study the long-time dynamics, and the main point is the analysis of the resonances, for which the key tool is the normal form theory.

The main difficulty in the application of the normal form theory to the Kirchhoff equation is due to the fact that it is a *quasilinear* PDE. Let us explain this point in more detail. The Kirchhoff equation has the Hamiltonian structure

$$\begin{cases} \partial_t u = \nabla_v H(u,v) = v, \\ \partial_t v = -\nabla_u H(u,v) = \Big(1 + \int_{\mathbb{T}^d} |\nabla u|^2 dx\Big) \Delta u, \end{cases} \tag{1.7}$$

where the Hamiltonian is

$$H(u,v) = \frac{1}{2} \int_{\mathbb{T}^d} v^2 dx + \frac{1}{2} \int_{\mathbb{T}^d} |\nabla u|^2 dx + \Big(\frac{1}{2} \int_{\mathbb{T}^d} |\nabla u|^2 dx \Big)^2, \tag{1.8}$$

and $\nabla_u H$, $\nabla_v H$ are the gradients with respect to the real scalar product

$$\langle f, g \rangle := \int_{\mathbb{T}^d} f(x) g(x)\, dx \quad \forall f, g \in L^2(\mathbb{T}^d, \mathbb{R}), \tag{1.9}$$

namely $H'(u,v)[f,g] = \langle \nabla_u H(u,v), f \rangle + \langle \nabla_v H(u,v), g \rangle$ for all u, v, f, g. As a consequence, the first natural attempt is trying to construct the Birkhoff normal form, using close-to-identity, symplectic transformations that are the time one flow of auxiliary Hamiltonians, with the goal of removing the nonresonant terms from the Hamiltonian (1.8), proceeding step by step with respect to the homogeneity orders. When one calculates (at least formally) the first step of this procedure, one finds a transformation Φ that is bounded on a ball of $H^s(\mathbb{T}^d, \mathbb{R}) \times H^{s-1}(\mathbb{T}^d, \mathbb{R})$ around the origin, but it is not close to the identity as a bounded operator, in the sense that $\|\Phi(u,v) - (u,v)\|_{H^s \times H^{s-1}}$ is not $\lesssim \|(u,v)\|^3_{H^s \times H^{s-1}}$, as one needs for the application of the Birkhoff normal form method. Hence the transformed Hamiltonian $H(\Phi(u,v))$ cannot be Taylor expanded in homogeneous orders without paying a loss of derivative, and the Birkhoff normal form procedure fails. This is ultimately a consequence of the quasilinear nature of the Kirchhoff equation. Also, even working with more general close-to-identity transformations of vector fields, not necessarily preserving the Hamiltonian structure, the direct application of the Poincaré normal form procedure encounters the same obstacle.

Thus, one has to look at the equation more carefully, distinguishing some terms that are harmless and some other terms that are responsible for the failure of the normal form construction. To this aim, it is convenient to introduce symmetrized complex coordinates (see Section 2), so that the linear wave operator becomes diagonal, and system (1.7) becomes (see (2.6))

$$\begin{cases} \partial_t u = -i\Lambda u - \frac{i}{4}\langle \Lambda(u+\overline{u}), u+\overline{u}\rangle \Lambda(u+\overline{u}), \\ \partial_t \overline{u} = i\Lambda \overline{u} + \frac{i}{4}\langle \Lambda(u+\overline{u}), u+\overline{u}\rangle \Lambda(u+\overline{u}), \end{cases} \tag{1.10}$$

where $\overline{u}$ is the complex conjugate of u, $\Lambda := |D_x|$ is the Fourier multiplier of symbol $|\xi|$, and $\langle f, g\rangle := \int_{\mathbb{T}^d} f(x) g(x)\, dx$ is the same as in (1.9), even for complex-valued functions f, g. We note that the cubic nonlinearity in (1.10) already has a "paralinear" structure, in the sense that, for all functions u, v, h, all $s \geq 0$, one has

$$\|\langle \Lambda u, v\rangle \Lambda h\|_s = |\langle \Lambda u, v\rangle|\, \|h\|_{s+1} \leq \|u\|_{\frac{1}{2}} \|v\|_{\frac{1}{2}} \|h\|_{s+1}.$$

Hence (1.10) can be interpreted as a linear system whose operator coefficients depend on $(u, \overline{u})$, namely

$$\partial_t \begin{pmatrix} u \\ \overline{u} \end{pmatrix} = \begin{pmatrix} -A(u,\overline{u}) & -B(u,\overline{u}) \\ B(u,\overline{u}) & A(u,\overline{u}) \end{pmatrix} \begin{pmatrix} u \\ \overline{u} \end{pmatrix}, \tag{1.11}$$

where

$$B(u,\overline{u}) = \frac{i}{4}\langle \Lambda(u+\overline{u}), u+\overline{u}\rangle \Lambda, \quad A(u,\overline{u}) = i\Lambda + B(u,\overline{u}).$$

Since our goal is the analysis of the existence time of the solutions, we calculate the time derivative $\partial_t(\|u\|_s^2)$ of the Sobolev norms and observe that the diagonal terms $A(u,\overline{u})$ give a zero contribution, while the off-diagonal terms $B(u,\overline{u})$, which couple u with $\overline{u}$, give terms that are $\leq 2\|u\|^2_{\frac{1}{2}} \|u\|^2_{s+\frac{1}{2}}$ only. Thus, on the one hand, this energy estimate has a loss of half a derivative and cannot be used for the existence theory; on the other hand, this observation suggests that $A(u,\overline{u})$ can be left untouched by the normal form transformation.

Hence the next natural attempt is the construction of a "partial" normal form transformation Φ that eliminates the cubic nonresonant terms only from $B(u,\overline{u})$ and does not modify $A(u,\overline{u})$. Indeed, such a transformation exists, it is bounded, and, unlike the full normal form, is close to the identity as a bounded transformation, namely $\|\Phi(u,\overline{u})-(u,\overline{u})\|_{H^s\times H^s} \lesssim \|(u,\overline{u})\|^3_{H^s\times H^s}$. Moreover, the cubic resonant terms of $B(u,\overline{u})$ that remain in the transformed system give zero contribution to the energy estimate. However, the transformed system contains unbounded off-diagonal terms of quintic and higher homogeneity order, which produce in the energy estimate the same loss of half a derivative as above.

At this point it becomes clear that one has to eliminate the off-diagonal unbounded terms *before* the normal form construction. This is at the base of the method developed by Delort in [22], [23] to construct a normal form for quasilinear Klein-Gordon equations on the circle. Roughly speaking, such a method consists in paralinearizing the equation, diagonalizing its principal symbol, so that one can obtain quasilinear energy estimates, and then starting with the normal form procedure. Further developments of this approach can be found in [14] and [15] about water waves equations on $\mathbb{T}$.

The off-diagonal unbounded terms of (1.10) are eliminated in Section 3, where we construct a nonlinear bounded transformation $\Phi^{(3)}$ that conjugates system (1.10) to a new system (see (3.13)) of the form

$$\begin{cases} \partial_t u = -i\sqrt{1+2P(u,\overline{u})}\,\Lambda u + \dfrac{i}{4(1+2P(u,\overline{u}))}\Big(\langle \Lambda\overline{u},\Lambda\overline{u}\rangle - \langle \Lambda u,\Lambda u\rangle\Big)\overline{u}, \\ \partial_t \overline{u} = i\sqrt{1+2P(u,\overline{u})}\,\Lambda \overline{u} + \dfrac{i}{4(1+2P(u,\overline{u}))}\Big(\langle \Lambda\overline{u},\Lambda\overline{u}\rangle - \langle \Lambda u,\Lambda u\rangle\Big)u, \end{cases} \tag{1.12}$$

where $P(u,\overline{u})$ is a real, nonnegative function of time only, defined as $P(u,\overline{u}) = \varphi(\frac{1}{4}\langle\Lambda(u+\overline{u}),u+\overline{u}\rangle)$, and φ is the inverse of the real map $x\mapsto x\sqrt{1+2x}$, $x\geq 0$. System (1.12) still has the structure (1.11), with the improvement that the off-diagonal part $B(u,\overline{u})$ is now a *bounded* operator, satisfying

$$\|B(u,\overline{u})h\|_s \leq \|u\|_1^2\|h\|_s$$

for all $s\geq 0$, all u,h. Thanks to the special structure of the Kirchhoff equation, and in particular to the lower bound $\frac{1}{4}\langle\Lambda(u+\overline{u}),u+\overline{u}\rangle = \int_{\mathbb{T}^d}(\mathrm{Re}\,(\Lambda^{\frac{1}{2}}u))^2\,dx \geq 0$, the transformation $\Phi^{(3)}$ is *global*, namely it is defined for all $u\in H^1_0(\mathbb{T}^d,\mathbb{C})$, and not only for small u. In (1.10) the off-diagonal term is an operator of order one with coefficient $\langle\Lambda(u+\overline{u}),u+\overline{u}\rangle$ defined for $u\in H_0^{\frac{1}{2}}(\mathbb{T}^d,\mathbb{C})$, while, after $\Phi^{(3)}$, the new off-diagonal term in (1.12) is an operator of order zero where the coefficient $(\langle\Lambda\overline{u},\Lambda\overline{u}\rangle - \langle\Lambda u,\Lambda u\rangle)$ is defined for $u\in H^1_0(\mathbb{T}^d,\mathbb{C})$. Thus the price to pay for removing the unbounded off-diagonal terms is an increase of $\frac{1}{2}$ in the regularity threshold for u (as if we had integrated by parts).

We remark that, reparametrizing the time variable, the coefficient $\sqrt{1+2P(u,\overline{u})}$ of the diagonal part in (1.12) could be normalized to 1; however, this is not needed to prove our result, because these coefficients are independent of x, and therefore the (unbounded) diagonal terms cancel out in the energy estimate.

In Section 4 we perform one step of normal form. It is a "partial" normal form because it does not modify the harmless cubic diagonal terms. The construction involves the differences $|j|-|k|$, $j,k \in \mathbb{Z}^d$, $j \neq k$, as denominators, which accumulate to zero in dimension $d \geq 2$. This produces the different regularity thresholds m_0 in Theorem 1.1, see Remark 1.3. The normal form transformation $\Phi^{(4)}$ is a bounded cubic correction of the identity map, and the off-diagonal terms of the transformed system (4.8), (4.41) remain bounded (unlike in the discussion above). The resonant cubic terms that remain after $\Phi^{(4)}$ create a nonlinear interaction between all Fourier coefficients $u_j(t)$ with Fourier modes $j \in \mathbb{Z}^d$ on a sphere $|j| = $ constant, while any two Fourier coefficients $u_j(t), u_k(t)$ with $|j| \neq |k|$ are uncoupled at the cubic homogeneity order. This, together with the conservation of the Hamiltonian, implies that there is no growth of Sobolev norms at the cubic homogeneity order. Therefore all the possible nonlinear effects of growth of Sobolev norms come from the terms of quintic and higher homogeneity order. This leads to the improved energy estimate (see (4.47))

$$\partial_t(\|u(t)\|_s^2) \leq C\|u(t)\|_{m_0}^4 \|u(t)\|_s^2$$

for the transformed system, whence we deduce that the lifespan of the solutions of the original Cauchy problem (1.3) is $T \sim (\|\alpha\|_{s+\frac{1}{2}} + \|\beta\|_{s-\frac{1}{2}})^{-4}$.

Preliminary further calculations suggest that, after performing the next step of normal form to remove the off-diagonal nonresonant quintic terms, some remaining quintic resonant terms could produce a nonlinear interaction between modes $|j| \neq |k|$, so that, in principle, a transfer of energy from low to high Fourier modes, and a growth of Sobolev norms (as in [18], [32], [33], [31] for the semilinear Schrödinger equation on $\mathbb{T}^2$) cannot be excluded. The analysis of the quintic order is the objective of a further investigation.

As a final comment, we observe that the general strategy developed in [22], [23], [14], [15] and also adopted in the present paper has a strong analogy with the technique developed for KAM theory for quasilinear PDEs in [6], [7], [26], [16], [8]: the first part of these methods uses pseudo-differential or paradifferential calculus to reduce the linearized or paralinearized operator to some more convenient diagonal form up to a sufficiently smoothing remainder, and it is a reduction with respect to the order of differentiation; then the second part uses normal forms or KAM reducibility schemes to reduce the size of the nonconstant remainders in the operator. In short: first reduce in $|D_x|$, then in ε.

1.2 Reversible Hamiltonian structure and prime integrals

In this section we make some observations about the structure of the Kirchhoff equation. We do not use them directly in the proof of Theorem 1.1, but they could be interesting *per se*.

As is well-known, the Kirchhoff equation has a Hamiltonian structure, which is (1.7)-(1.8). Also, since the Hamiltonian (1.8) is even in v, namely $H(u,-v) = H(u,v)$, the Hamiltonian vector field $X(u,v) = (\nabla_v H(u,v), -\nabla_u H(u,v))$ satisfies $X \circ S + S \circ X = 0$, where S is the involution $S(u,v) = (u,-v)$. Therefore system (1.7) is time-reversible with respect to S, which simply means that if $u(t,x)$ is a solution of (1.1), then $u(-t,x)$ is also a solution of the same equation.

Another observation is that the space of functions $u(t,x) = u(t,-x)$ that are even in x is an invariant subspace for the Kirchhoff equation, as well as the space of odd functions $u(t,x) = -u(t,-x)$. The Fourier support is also invariant for the flow: since the Kirchhoff equation for $u(t,x) = \sum_{j\in\mathbb{Z}^d} u_j(t)e^{ij\cdot x}$ is the system of equations

$$u_j'' + |j|^2 u_j \Big(1 + \sum_{k\in\mathbb{Z}^d} |k|^2 |u_k|^2 \Big) = 0 \quad \forall j \in \mathbb{Z}^d, \tag{1.13}$$

if $u_j(0) = u_j'(0) = 0$ for some j, then $u_j(t) = 0$ for all t. In particular, if the initial data (α, β) have finite Fourier support, then the solution exists for all times, and a simple application of finite-dimensional KAM theory shows that some of them are quasi-periodic in time.

In addition to the Hamiltonian, the momentum

$$M = \int_{\mathbb{T}^d} (\partial_t u) \nabla u \, dx$$

is also a conserved quantity. Even more, because of the special structure of the Kirchhoff equation, the momentum is the sum $M = \sum_{j\in\mathbb{Z}^d} M_j$ of infinitely many prime integrals M_j, defined in the following way. If $u(t,x) = \sum_{j\in\mathbb{Z}^d} u_j(t)e^{ij\cdot x}$, then

$$M_j = \frac{1}{2} ij(u_j \partial_t u_{-j} - u_{-j}\partial_t u_j), \quad j \in \mathbb{Z}^d,$$

and one has

$$\partial_t M_j = \frac{1}{2} ij(u_j \partial_{tt} u_{-j} - u_{-j}\partial_{tt} u_j) = 0$$

because each u_j satisfies (1.13). This observation seems to be new. Since $M_{-j} = M_j$, only "a half" of these prime integrals are independent.

The standard linear changes of coordinates of Section 2 preserve both the Hamiltonian and the reversible structure (S becomes $S_1(u, \overline{u}) = (\overline{u}, u)$ in complex coordinates). The nonlinear transformation of Section 3 is not symplectic, but it preserves the reversible structure, which is also preserved by the normal form of Section 4.

1.3 Related literature and open questions

Equation (1.1) was introduced by Kirchhoff [36] (and, in one dimension, independently rediscovered in [17] and [47]) to model the transversal oscillations of a clamped string or plate, taking into account nonlinear elastic effects. The first results on the Cauchy problem (1.1)-(1.2) are due to Bernstein. In his 1940 pioneering paper [13], he studied the Cauchy problem on an interval, with Dirichlet boundary conditions, and proved global wellposedness for analytic initial data (α, β), and local wellposedness for $(\alpha, \beta) \in H^2 \times H^1$.

After that, the research on the Kirchhoff equation has been developed in various directions, with a different kind of results on compact domains (bounded subset $\Omega \subset \mathbb{R}^d$ with Dirichlet boundary conditions, or periodic boundary conditions $\Omega = \mathbb{T}^d$) or non compact domains ($\Omega = \mathbb{R}^d$ or "exterior domains" $\Omega = \mathbb{R}^d \setminus K$, with $K \subset \mathbb{R}^d$ compact domain).

For $\Omega = \mathbb{R}^d$, Greenberg and Hu [30] in dimension $d = 1$ and D'Ancona and Spagnolo [21] in higher dimension proved global wellposedness with scattering for small initial data in weighted Sobolev spaces. Further improvements, dealing with spectrally characterized initial data in larger subsets of the Sobolev spaces, and also including the case of exterior domains, have been more recently obtained, for example, by Yamazaki, Matsuyama and Ruzhansky, see [54], [41], [43] and the many references therein. For global solutions that do not scatter see [40]. Still open is the main question whether the solutions with small initial data in the standard (not weighted) Sobolev spaces $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$ are globally defined.

Another research direction regards the extension of global wellposedness, on both compact and non compact domains, to non small initial data that are in a larger space than analytic functions: see, for example, Pokhozhaev [50], Arosio and Spagnolo [4], Nishihara [48], Manfrin [39], Ghisi and Gobbino [29], and the references therein. Still open is the question whether the solutions with initial data of arbitrary size and Gevrey regularity (on any domain) are globally defined.

On compact domains, dispersion, scattering and time-decay mechanisms are not available, and there are no results of global existence, nor of finite time blowup, for initial data (α, β) of Sobolev, or C^∞, or Gevrey regularity. The local wellposedness in the Sobolev class $H^{\frac{3}{2}} \times H^{\frac{1}{2}}$ has been proved by Dickey [24], Medeiros and Milla Miranda [44] and Arosio and Panizzi [3], with existence time of order $(\|\alpha\|_{\frac{3}{2}} + \|\beta\|_{\frac{1}{2}})^{-2}$. Beyond the question about the global wellposedness for small data in Sobolev class, another open question concerns the local wellposedness in the energy space $H^1 \times L^2$ or in $H^s \times H^{s-1}$ for $1 < s < \frac{3}{2}$.

For more details, generalizations (degenerate Kirchhoff equations, Kirchhoff systems, forced and/or damped Kirchhoff equations, etc.) and other open questions, we refer to Lions [38] and the surveys of Arosio [2], Spagnolo [52], and Matsuyama and Ruzhansky [42].

We also mention the recent results [5], [45], [46], [19], which prove the existence of time periodic or quasi-periodic solutions of time periodically or quasi-periodically forced Kirchhoff equations on $\mathbb{T}^d$, using Nash-Moser and KAM techniques.

Concerning the normal form theory for quasilinear PDEs, we mention the pioneering work of Shatah [51] and the subsequent development by Ozawa, Tsutaya and Tsutsumi [49] on quasilinear Klein-Gordon equations on $\mathbb{R}^d$, the abstract result of Bambusi [9], the aforementioned papers of Delort [22], [23] on quasilinear Klein-Gordon on $\mathbb{T}$, and the recent literature on water waves by Wu [53], Germain, Masmoudi and Shatah [28], Alazard and Delort [1], Ionescu and Pusateri [35], Craig and Sulem [20], Ifrim and Tataru [34], Berti and Delort [14], Berti, Feola and Pusateri [15]. Other applications of normal form techniques to get long-time existence for nonlinear PDEs on compact domains can be found in the work of Bambusi, Nekhoroshev, Grébert, Delort and Szeftel, see e.g. [12], [11], [10], and Feola, Giuliani and Pasquali [25].

Acknowledgements. This research was supported by ERC under FP7, Project no. 306414 *Hamiltonian PDEs and small divisor problems: a dynamical systems approach* and by PRIN 2015 *Variational methods, with applications to problems in mathematical physics and geometry.*

2 Linear transformations

In this section we make two elementary, standard linear changes of variables to transform system (1.7) into another one (see (2.6)) where the linear part is diagonal, preserving both the real and the Hamiltonian structure of the problem. These standard transformations are the symmetrization of the highest order (section 2.1) and then the diagonalization of the linear terms (section 2.2).

2.1 Symmetrization of the highest order

In the Sobolev spaces (1.4) of zero-mean functions, the Fourier multiplier

$$\Lambda := |D_x| : H^s_0 \to H^{s-1}_0, \quad e^{ij\cdot x} \mapsto |j| e^{ij\cdot x}$$

is invertible. System (1.7) writes

$$\begin{cases} \partial_t u = v \\ \partial_t v = -(1 + \langle \Lambda u, \Lambda u \rangle) \Lambda^2 u, \end{cases} \tag{2.1}$$

where $\langle \cdot, \cdot \rangle$ is defined in (1.9); the Hamiltonian (1.8) is

$$H(u,v) = \frac{1}{2}\langle v, v\rangle + \frac{1}{2}\langle \Lambda u, \Lambda u \rangle + \frac{1}{4}\langle \Lambda u, \Lambda u\rangle^2.$$

To symmetrize the system at the highest order, we consider the linear, symplectic transformation

$$(u,v) = \Phi^{(1)}(q,p) = (\Lambda^{-\frac{1}{2}} q, \Lambda^{\frac{1}{2}} p). \tag{2.2}$$

System (2.1) becomes

$$\begin{cases} \partial_t q = \Lambda p \\ \partial_t p = -(1 + \langle \Lambda^{\frac{1}{2}} q, \Lambda^{\frac{1}{2}} q \rangle) \Lambda q, \end{cases} \tag{2.3}$$

which is the Hamiltonian system $\partial_t(q,p) = J \nabla H^{(1)}(q,p)$ with Hamiltonian $H^{(1)} = H \circ \Phi^{(1)}$, namely

$$H^{(1)}(q,p) = \frac{1}{2}\langle \Lambda^{\frac{1}{2}} p, \Lambda^{\frac{1}{2}} p\rangle + \frac{1}{2}\langle \Lambda^{\frac{1}{2}} q, \Lambda^{\frac{1}{2}} q\rangle + \frac{1}{4}\langle \Lambda^{\frac{1}{2}} q, \Lambda^{\frac{1}{2}} q\rangle^2, \quad J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \tag{2.4}$$

Note that the original problem requires the "physical" variables (u,v) to be real-valued; this corresponds to (q,p) being real-valued too. Also note that $\langle \Lambda^{\frac{1}{2}} p, \Lambda^{\frac{1}{2}} p \rangle = \langle \Lambda p, p \rangle$.

2.2 Diagonalization of the highest order: complex variables

To diagonalize the linear part $\partial_t q = \Lambda p$, $\partial_t p = -\Lambda q$ of system (2.3), we introduce complex variables.

System (2.3) and the Hamiltonian $H^{(1)}(q,p)$ in (2.4) are also meaningful, without any change, for *complex* functions q, p. Thus we define the change of complex variables $(q,p) = \Phi^{(2)}(f,g)$ as

$$(q,p) = \Phi^{(2)}(f,g) = \Big(\frac{f+g}{\sqrt{2}}, \frac{f-g}{i\sqrt{2}} \Big), \qquad f = \frac{q+ip}{\sqrt{2}}, \quad g = \frac{q-ip}{\sqrt{2}}, \tag{2.5}$$

so that system (2.3) becomes

$$\begin{cases} \partial_t f = -i\Lambda f - i\frac{1}{4}\langle \Lambda(f+g), f+g\rangle \Lambda(f+g) \\ \partial_t g = i\Lambda g + i\frac{1}{4}\langle \Lambda(f+g), f+g\rangle \Lambda(f+g) \end{cases} \tag{2.6}$$

where the pairing $\langle \cdot, \cdot \rangle$ denotes the integral of the product of any two complex functions

$$\langle w, h \rangle := \int_{\mathbb{T}^d} w(x) h(x)\, dx = \sum_{j \in \mathbb{Z}^d \setminus \{0\}} w_j h_{-j}, \quad w, h \in L^2(\mathbb{T}^d, \mathbb{C}). \tag{2.7}$$

The map $\Phi^{(2)} : (f,g) \mapsto (q,p)$ in (2.5) is a $\mathbb{C}$-linear isomorphism of the space $L^2_0(\mathbb{T}^d, \mathbb{C}) \times L^2_0(\mathbb{T}^d, \mathbb{C})$ of pairs of complex functions. When (q,p) are real, (f,g) are complex conjugate. The restriction of $\Phi^{(2)}$ to the space

$$L^2_0(\mathbb{T}^d, c.c.) := \{(f,g) \in L^2_0(\mathbb{T}^d, \mathbb{C}) \times L^2_0(\mathbb{T}^d, \mathbb{C}) : g = \overline{f}\}$$

of pairs of complex conjugate functions is an $\mathbb{R}$-linear isomorphism onto the space $L^2_0(\mathbb{T}^d, \mathbb{R}) \times L^2_0(\mathbb{T}^d, \mathbb{R})$ of pairs of real functions. For $g = \overline{f}$, the second equation in (2.6) is redundant, being the complex conjugate of the first equation. In other words, system (2.6) has the following "real structure": it is of the form

$$\partial_t \begin{pmatrix} f \\ g \end{pmatrix} = \mathcal{F}(f,g) = \begin{pmatrix} \mathcal{F}_1(f,g) \\ \mathcal{F}_2(f,g) \end{pmatrix}$$

where the vector field $\mathcal{F}(f,g)$ satisfies

$$\mathcal{F}_2(f, \overline{f}) = \overline{\mathcal{F}_1(f, \overline{f})}. \tag{2.8}$$

Under the transformation $\Phi^{(2)}$, the Hamiltonian system (2.3) for complex variables (q,p) becomes (2.6), which is the Hamiltonian system $\partial_t(f,g) = iJ\nabla H^{(2)}(f,g)$ with Hamiltonian $H^{(2)} = H^{(1)} \circ \Phi^{(2)}$, namely

$$H^{(2)}(f,g) = \langle \Lambda f, g \rangle + \frac{1}{16}\langle \Lambda(f+g), f+g\rangle^2,$$

where J is defined in (2.4), $\langle \cdot, \cdot \rangle$ is defined in (2.7), and $\nabla H^{(2)}$ is the gradient with respect to $\langle \cdot, \cdot \rangle$. System (2.3) for real (q,p) (which corresponds to the original Kirchhoff equation) becomes system (2.6) restricted to the subspace $L^2_0(\mathbb{T}^d, c.c.)$ where $g = \overline{f}$.

To complete the definition of the function spaces, for any real $s \geq 0$ we define

$$H^s_0(\mathbb{T}^d, c.c.) := \{(f,g) \in L^2_0(\mathbb{T}^d, c.c.) : f, g \in H^s_0(\mathbb{T}^d, \mathbb{C})\}.$$

3 Diagonalization of the order one

Following a "para-differential approach", we note that the term $\langle \Lambda(f+g), f+g\rangle$ in (2.6) plays the rôle of a coefficient, while Λ outside the scalar product is an operator of order one, in the sense that

$$\|\langle \Lambda f, g\rangle \Lambda h\|_s = \|h\|_{s+1} |\langle \Lambda f, g\rangle| \le \|h\|_{s+1} \|f\|_{\frac12} \|g\|_{\frac12} \quad \forall s \ge 0,\ h \in H^{s+1},\ f, g \in H^{\frac12}.$$

Thus we write system (2.6) as

$$\partial_t \begin{pmatrix} f \\ g \end{pmatrix} = i \begin{pmatrix} -1 - Q(f,g) & -Q(f,g) \\ Q(f,g) & 1 + Q(f,g) \end{pmatrix} \Lambda \begin{pmatrix} f \\ g \end{pmatrix} \tag{3.1}$$

where

$$Q(f,g) := \frac14 \langle \Lambda(f+g), f+g\rangle. \tag{3.2}$$

The aim of this section is to diagonalize system (3.1) up to a bounded remainder, dealing with $Q(f,g)$ as a coefficient (even if it depends nonlinearly on the variables (f,g)). On the real subspace $L^2_0(\mathbb{T}^d, c.c.)$ one has $g = \overline{f}$, and therefore

$$Q(f,g) = \frac14 \langle \Lambda(f+g), f+g\rangle = \frac14 \langle \Lambda^{\frac12}(f+\overline{f}), \Lambda^{\frac12}(f+\overline{f})\rangle = \int_{\mathbb{T}^d} \big(\Lambda^{\frac12} \mathrm{Re}\,(f)\big)^2\, dx \ge 0,$$

where $\mathrm{Re}\,(f)$ is the real part of f. Since $Q(f,g) \ge 0$, the matrix of the coefficients in (3.1) has purely imaginary eigenvalues. For any $x \ge 0$, one has

$$\begin{pmatrix} -1-x & -x \\ x & 1+x \end{pmatrix} \begin{pmatrix} 1 & \rho(x) \\ \rho(x) & 1 \end{pmatrix} = \begin{pmatrix} 1 & \rho(x) \\ \rho(x) & 1 \end{pmatrix} \begin{pmatrix} -\sqrt{1+2x} & 0 \\ 0 & \sqrt{1+2x} \end{pmatrix} \tag{3.3}$$

where

$$\rho(x) := \frac{-x}{1+x+\sqrt{1+2x}}\,. \tag{3.4}$$

Note that $-1 < \rho(x) \le 0$ for $x \ge 0$, so that the matrix $\left(\begin{smallmatrix} 1 & \rho(x) \\ \rho(x) & 1 \end{smallmatrix}\right)$ is invertible. We define

$$\begin{pmatrix} f \\ g \end{pmatrix} = \mathcal{M} \begin{pmatrix} \eta \\ \psi \end{pmatrix}, \quad \mathcal{M} = \mathcal{M}(\rho) := \frac{1}{\sqrt{1-\rho^2}} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \tag{3.5}$$

where $\rho = \rho(Q(f,g))$, with ρ defined in (3.4), and $Q(f,g)$ in (3.2). The presence of the factor $(1-\rho^2)^{-1/2}$ in the definition of $\mathcal{M}$ is discussed in Remark 3.2 below. To define a nonlinear change of variable expressing (f,g) in terms of (η,ψ) by using (3.5), we have to express the matrix $\mathcal{M}$ as a function of η, ψ. Using (3.5), we calculate

$$Q(f,g) = \frac14 \langle \Lambda(f+g), f+g\rangle = \frac{1+\rho(Q(f,g))}{4(1-\rho(Q(f,g)))} \langle \Lambda(\eta+\psi), \eta+\psi\rangle.$$

From definition (3.4), for any $x \ge 0$ one has

$$\frac{1-\rho(x)}{1+\rho(x)} = \sqrt{1+2x},$$

whence

$$Q(f,g)\sqrt{1+2Q(f,g)} = \frac{1}{4}\langle \Lambda(\eta+\psi), \eta+\psi\rangle = Q(\eta,\psi). \tag{3.6}$$

The function $x \mapsto x\sqrt{1+2x}$ is invertible, and we denote by φ its inverse,

$$x\sqrt{1+2x} = y \quad \Leftrightarrow \quad x = \varphi(y). \tag{3.7}$$

Hence we can express $Q(f,g)$ in terms of (η,ψ) as

$$Q(f,g) = \varphi\Big(\frac{1}{4}\langle \Lambda(\eta+\psi), \eta+\psi\rangle\Big) = \varphi(Q(\eta,\psi)) =: P(\eta,\psi). \tag{3.8}$$

As a consequence, the matrix $\mathcal{M}$ in (3.5) can also be expressed as a function of (η,ψ). In short, we denote it by $\mathcal{M}(\eta,\psi)$, so that $\mathcal{M}(\eta,\psi)$ is $\mathcal{M}(\rho)$ where $\rho = \rho(\varphi(Q(\eta,\psi))) = \rho(P(\eta,\psi))$, namely

$$\mathcal{M}(\eta,\psi) := \frac{1}{\sqrt{1-\rho^2(P(\eta,\psi))}} \begin{pmatrix} 1 & \rho(P(\eta,\psi)) \\ \rho(P(\eta,\psi)) & 1 \end{pmatrix}. \tag{3.9}$$

We define the transformation $(f,g) = \Phi^{(3)}(\eta,\psi)$ by formula (3.5) where $\mathcal{M} = \mathcal{M}(\eta,\psi)$.

Lemma 3.1. *Let $\Phi^{(3)}$ be the map*

$$\Phi^{(3)}(\eta,\psi) = \mathcal{M}(\eta,\psi)\begin{pmatrix} \eta \\ \psi \end{pmatrix}, \tag{3.10}$$

where $\mathcal{M}(\eta,\psi)$ is defined in (3.9)*, ρ is defined in* (3.4) *and P in* (3.8)*. Then, for all real $s \geq \frac{1}{2}$, the nonlinear map $\Phi^{(3)} : H_0^s(\mathbb{T}^d, c.c.) \to H_0^s(\mathbb{T}^d, c.c.)$ is invertible, continuous, with continuous inverse*

$$(\Phi^{(3)})^{-1}(f,g) = \frac{1}{\sqrt{1-\rho^2(Q(f,g))}} \begin{pmatrix} 1 & -\rho(Q(f,g)) \\ -\rho(Q(f,g)) & 1 \end{pmatrix}\begin{pmatrix} f \\ g \end{pmatrix}.$$

Moreover, for all $s \geq \frac{1}{2}$, all $(\eta,\psi) \in H_0^s(\mathbb{T}^d, c.c.)$, one has

$$\|\Phi^{(3)}(\eta,\psi)\|_s \leq C(\|\eta,\psi\|_{\frac{1}{2}})\|\eta,\psi\|_s$$

for some increasing function C. The same estimate is satisfied by $(\Phi^{(3)})^{-1}$.

Proof. The regularity $H^{\frac{1}{2}}$ guarantees that $Q(f,g)$ and $Q(\eta,\psi)$ are finite. The only point to prove is that $\Phi^{(3)}$ and its inverse map pairs of complex conjugate functions into pairs of complex conjugate functions. Let $(\eta,\psi) \in H_0^{\frac{1}{2}}(\mathbb{T}^d, c.c.)$. Then $Q(\eta,\psi)$, and therefore also $P(\eta,\psi) = \varphi(Q(\eta,\psi))$, are real and ≥ 0. Let $(f,g) = \Phi^{(3)}(\eta,\psi)$, namely

$$f = \frac{\eta+\rho\psi}{\sqrt{1-\rho^2}}, \quad g = \frac{\rho\eta+\psi}{\sqrt{1-\rho^2}}$$

where $\rho = \rho(P(\eta,\psi))$. Since $\psi = \overline{\eta}$ and ρ is real, we deduce that $\overline{f} = g$, and therefore $(f,g) \in H_0^{\frac{1}{2}}(\mathbb{T}^d, c.c.)$. □

Now we calculate how system (2.6), i.e. (3.1), transforms under the change of variable $(f,g) = \Phi^{(3)}(\eta,\psi) = \mathcal{M}(\eta,\psi)[\eta,\psi]$. We calculate

$$\partial_t(f,g) = \partial_t\{\mathcal{M}(\eta,\psi)[\eta,\psi]\} = \mathcal{M}(\eta,\psi)[\partial_t\eta,\partial_t\psi] + \partial_t\{\mathcal{M}(\eta,\psi)\}[\eta,\psi],$$

and

$$\partial_t\{\mathcal{M}(\eta,\psi)\} = \frac{1}{(1-\rho^2)^{3/2}}\begin{pmatrix}\rho & 1\\ 1 & \rho\end{pmatrix}\partial_t\rho,$$

$$\partial_t\rho = \partial_t\{\rho(\varphi(Q(\eta,\psi)))\} = \rho'\big(\varphi(Q(\eta,\psi))\big)\,\varphi'\big(Q(\eta,\psi)\big)\,\frac{1}{2}\langle\Lambda(\eta+\psi),\partial_t\eta+\partial_t\psi\rangle.$$

By (3.3) and (3.8), we have

$$\begin{pmatrix}-i(1+Q(f,g)) & -iQ(f,g)\\ iQ(f,g) & i(1+Q(f,g))\end{pmatrix}\begin{pmatrix}\Lambda f\\ \Lambda g\end{pmatrix} = \mathcal{M}(\eta,\psi)\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix} i\sqrt{1+2P(\eta,\psi)}\begin{pmatrix}\Lambda\eta\\ \Lambda\psi\end{pmatrix}.$$

Thus, applying $\mathcal{M}(\eta,\psi)^{-1}$ from the left, (3.1) becomes

$$\partial_t\begin{pmatrix}\eta\\ \psi\end{pmatrix} + \mathcal{M}(\eta,\psi)^{-1}\partial_t\{\mathcal{M}(\eta,\psi)\}\begin{pmatrix}\eta\\ \psi\end{pmatrix} = \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix} i\sqrt{1+2P(\eta,\psi)}\begin{pmatrix}\Lambda\eta\\ \Lambda\psi\end{pmatrix}. \quad (3.11)$$

We calculate

$$\mathcal{M}(\eta,\psi)^{-1} = \frac{1}{\sqrt{1-\rho^2}}\begin{pmatrix}1 & -\rho\\ -\rho & 1\end{pmatrix},$$

$$\frac{1}{\sqrt{1-\rho^2}}\begin{pmatrix}1 & -\rho\\ -\rho & 1\end{pmatrix}\frac{1}{(1-\rho^2)^{\frac{3}{2}}}\begin{pmatrix}\rho & 1\\ 1 & \rho\end{pmatrix} = \frac{1}{1-\rho^2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},$$

$$\rho'(x) = \frac{-1}{\sqrt{1+2x}\,(1+x+\sqrt{1+2x})},$$

$$\frac{1}{1-\rho^2(x)}\cdot\frac{-1}{\sqrt{1+2x}\,(1+x+\sqrt{1+2x})} = \frac{-1}{2(1+2x)},$$

$$\varphi'(y) = \frac{\sqrt{1+2\varphi(y)}}{1+3\varphi(y)}.$$

Hence

$$\mathcal{M}(\eta,\psi)^{-1}\partial_t\{\mathcal{M}(\eta,\psi)\}\begin{pmatrix}\eta\\ \psi\end{pmatrix} = \mathcal{K}(\eta,\psi)\begin{pmatrix}\partial_t\eta\\ \partial_t\psi\end{pmatrix}$$

where $\mathcal{K}(\eta,\psi)$ is the operator

$$\mathcal{K}(\eta,\psi)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} := \begin{pmatrix}\psi\\ \eta\end{pmatrix} F(\eta,\psi)\langle\Lambda(\eta+\psi),\alpha+\beta\rangle$$

and $F(\eta,\psi)$ is the scalar factor

$$F(\eta,\psi) := \frac{-1}{4(1+3P(\eta,\psi))\sqrt{1+2P(\eta,\psi)}}. \quad (3.12)$$

By induction, for all $n = 1, 2, 3, \ldots$ one has

$$\mathcal{K}^n(\eta,\psi)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\psi\\ \eta\end{pmatrix} F(\eta,\psi)^n \langle \Lambda(\eta+\psi), \eta+\psi\rangle^{n-1} \langle \Lambda(\eta+\psi), \alpha+\beta\rangle.$$

Thus, by geometric series,

$$\sum_{n=1}^{\infty} (-\mathcal{K}(\eta,\psi))^n \begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\psi\\ \eta\end{pmatrix} \langle \Lambda(\eta+\psi), \alpha+\beta\rangle \frac{-F(\eta,\psi)}{1+F(\eta,\psi)\langle \Lambda(\eta+\psi), \eta+\psi\rangle}$$

provided that $|F(\eta,\psi)\langle \Lambda(\eta+\psi), \eta+\psi\rangle| < 1$. Since $\langle \Lambda(\eta+\psi), \eta+\psi\rangle = 4Q(\eta,\psi)$, using (3.12), (3.8), (3.6) and (3.7), we have

$$F(\eta,\psi)\langle \Lambda(\eta+\psi), \eta+\psi\rangle = \frac{-Q(\eta,\psi)}{(1+3P(\eta,\psi))\sqrt{1+2P(\eta,\psi)}} = \frac{-Q(f,g)}{1+3Q(f,g)},$$

whence $|F(\eta,\psi)\langle \Lambda(\eta+\psi), \eta+\psi\rangle| < 1/3$ for all $Q(f,g) \geq 0$, and the geometric series converges. Using the same identities, we also obtain that

$$\frac{-F(\eta,\psi)}{1+F(\eta,\psi)\langle \Lambda(\eta+\psi), \eta+\psi\rangle} = \frac{1}{4(1+2P(\eta,\psi))^{\frac{3}{2}}}.$$

Hence

$$(I+\mathcal{K}(\eta,\psi))^{-1}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \begin{pmatrix}\psi\\ \eta\end{pmatrix} \langle \Lambda(\eta+\psi), \alpha+\beta\rangle \frac{1}{4(1+2P(\eta,\psi))^{\frac{3}{2}}}.$$

Then system (3.11) becomes

$$\partial_t \begin{pmatrix}\eta\\ \psi\end{pmatrix} = (I+\mathcal{K}(\eta,\psi))^{-1} \begin{pmatrix}-i\sqrt{1+2P(\eta,\psi)} & 0\\ 0 & i\sqrt{1+2P(\eta,\psi)}\end{pmatrix} \begin{pmatrix}\Lambda\eta\\ \Lambda\psi\end{pmatrix},$$

which is

$$\partial_t \begin{pmatrix}\eta\\ \psi\end{pmatrix} = \begin{pmatrix}-i\sqrt{1+2P(\eta,\psi)}\,\Lambda\eta\\ i\sqrt{1+2P(\eta,\psi)}\,\Lambda\psi\end{pmatrix} + \begin{pmatrix}\psi\\ \eta\end{pmatrix} \langle \Lambda(\eta+\psi), \Lambda(\psi-\eta)\rangle \frac{i}{4(1+2P(\eta,\psi))},$$

namely

$$\begin{cases} \partial_t \eta = -i\sqrt{1+2P(\eta,\psi)}\,\Lambda\eta + \dfrac{i}{4(1+2P(\eta,\psi))}\Big(\langle \Lambda\psi, \Lambda\psi\rangle - \langle \Lambda\eta, \Lambda\eta\rangle\Big)\psi \\ \partial_t \psi = i\sqrt{1+2P(\eta,\psi)}\,\Lambda\psi + \dfrac{i}{4(1+2P(\eta,\psi))}\Big(\langle \Lambda\psi, \Lambda\psi\rangle - \langle \Lambda\eta, \Lambda\eta\rangle\Big)\eta. \end{cases} \tag{3.13}$$

We remark that system (3.13) is diagonal at the order one, i.e. the coupling of η and ψ (except for the coefficients) is confined to terms of order zero. Also note that the coefficients of (3.13) are finite for $\eta, \psi \in H^1_0$, while the coefficients in (2.6) are finite for $f, g \in H^{\frac{1}{2}}_0$: the regularity threshold of the transformed system is $\frac{1}{2}$ higher than before.

We also note that the real structure is preserved, namely the second equation in (3.13) is the complex conjugate of the first one, or, in other words, the vector field in (3.13) satisfies property (2.8).

Remark 3.2. It would be tempting to use the transformation

$$\begin{pmatrix} f \\ g \end{pmatrix} = \frac{1}{1+\rho} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \eta \\ \psi \end{pmatrix} \tag{3.14}$$

instead of (3.5), because (3.14) preserves the formula of Q:

$$Q(f,g) = \frac{1}{4}\langle \Lambda(f+g), f+g\rangle = \frac{1}{4}\langle \Lambda(\eta+\psi), \eta+\psi\rangle = Q(\eta,\psi),$$

avoiding the use of the inverse function φ. However, using (3.14) would produce a diagonal term of order zero in the transformed system which does not cancel out in the energy estimate (in fact, on the real subspace $\psi = \bar\eta$ such a diagonal term has a real coefficient). The factor $(1-\rho^2)^{-\frac{1}{2}}$ in (3.5) is the only (up to constant factors) choice that eliminates those diagonal terms of order zero. This property is related to the symplectic structure of (2.6). □

4 Normal form transformation

Let (η, ψ) be a solution of (3.13), with $\psi = \overline{\eta}$. Then its a priori energy estimate is

$$\begin{aligned}
\partial_t(\|\eta\|_s^2) &= \partial_t\langle \Lambda^s\eta, \Lambda^s\psi\rangle = \langle \partial_t\eta, \Lambda^{2s}\psi\rangle + \langle \Lambda^{2s}\eta, \partial_t\psi\rangle \\
&= \langle -i\sqrt{1+2P(\eta,\psi)}\,\Lambda\eta + \frac{i(\langle\Lambda\psi,\Lambda\psi\rangle - \langle\Lambda\eta,\Lambda\eta\rangle)}{4(1+2P(\eta,\psi))}\psi, \Lambda^{2s}\psi\rangle \\
&\quad + \langle \Lambda^{2s}\eta, i\sqrt{1+2P(\eta,\psi)}\,\Lambda\psi + \frac{i(\langle\Lambda\psi,\Lambda\psi\rangle - \langle\Lambda\eta,\Lambda\eta\rangle)}{4(1+2P(\eta,\psi))}\eta\rangle \\
&= \frac{i(\langle\Lambda\psi,\Lambda\psi\rangle - \langle\Lambda\eta,\Lambda\eta\rangle)}{4(1+2P(\eta,\psi))}(\langle\psi,\Lambda^{2s}\psi\rangle + \langle\Lambda^{2s}\eta,\eta\rangle) \\
&\le \|\eta\|_1^2\|\eta\|_s^2.
\end{aligned}$$

This gives the local existence in H_0^1 in a time interval $[0,T]$ with $T = O(\|\eta(0)\|_1^{-2})$. We note that the terms $(-i\sqrt{1+2P(\eta,\psi)}\,\Lambda\eta, i\sqrt{1+2P(\eta,\psi)}\,\Lambda\psi)$ give no contribution to the energy estimate, thanks to their diagonal structure, which was obtained in the previous section. Hence, to improve the energy estimate and to extend the existence time, there is no need to modify those terms. In fact, reparametrizing the time variable, the coefficient $\sqrt{1+2P(\eta,\psi)}$ could be normalized to 1; however, as just noticed, this is not needed to our purposes.

The next step in our proof is the cancellation of the cubic terms contributing to the energy estimate. We write (3.13) as

$$\partial_t(\eta,\psi) = X(\eta,\psi) = \mathcal{D}_1(\eta,\psi) + \mathcal{D}_{\ge 3}(\eta,\psi) + \mathcal{B}_3(\eta,\psi) + \mathcal{R}_{\ge 5}(\eta,\psi) \tag{4.1}$$

where $\mathcal{D}_1(\eta,\psi)$ is the linear component of the unbounded diagonal operator $\mathcal{D}(\eta,\psi) = i\sqrt{1+2P(\eta,\psi)}(-\Lambda\eta, \Lambda\psi)$, namely

$$\mathcal{D}_1(\eta,\psi) = \begin{pmatrix} -i\Lambda\eta \\ i\Lambda\psi \end{pmatrix},$$

$\mathcal{D}_{\geq 3}(\eta,\psi)$ is the difference $\mathcal{D}-\mathcal{D}_1$, namely

$$\mathcal{D}_{\geq 3}(\eta,\psi) = \Big(\sqrt{1+2P(\eta,\psi)}\,-1\Big)\begin{pmatrix}-i\Lambda\eta\\ i\Lambda\psi\end{pmatrix}, \tag{4.2}$$

$\mathcal{B}_3(\eta,\psi)$ is the cubic component of the bounded, off-diagonal term

$$\mathcal{B}_3(\eta,\psi) = \frac{i}{4}\Big(\langle\Lambda\psi,\Lambda\psi\rangle - \langle\Lambda\eta,\Lambda\eta\rangle\Big)\begin{pmatrix}\psi\\ \eta\end{pmatrix} \tag{4.3}$$

and $\mathcal{R}_{\geq 5}(\eta,\psi)$ is the bounded remainder of higher homogeneity degree

$$\mathcal{R}_{\geq 5}(\eta,\psi) = \frac{-iP(\eta,\psi)}{2(1+2P(\eta,\psi))}\Big(\langle\Lambda\psi,\Lambda\psi\rangle - \langle\Lambda\eta,\Lambda\eta\rangle\Big)\begin{pmatrix}\psi\\ \eta\end{pmatrix}. \tag{4.4}$$

The aim of this section is to remove $\mathcal{B}_3$ ($\mathcal{D}$ gives no contribution to the energy estimate, and $\mathcal{R}_{\geq 5}(\eta,\psi) = O((\eta,\psi)^5)$ gives a contribution of higher order).

We consider a transformation $(\eta,\psi) = \Phi^{(4)}(w,z)$ of the form

$$\begin{pmatrix}\eta\\ \psi\end{pmatrix} = \Phi^{(4)}(w,z) = (I+M(w,z))\begin{pmatrix}w\\ z\end{pmatrix}, \tag{4.5}$$

$$M(w,z) = \begin{pmatrix}M_{11}(w,z) & M_{12}(w,z)\\ M_{21}(w,z) & M_{22}(w,z)\end{pmatrix},$$

$$M_{ij}(w,z) = A_{ij}[w,w] + B_{ij}[w,z] + C_{ij}[z,z], \quad i,j\in\{1,2\},$$

where A_{ij}, B_{ij}, C_{ij} are bilinear maps. We also denote

$$A[w,w] = \begin{pmatrix}A_{11}[w,w] & A_{12}[w,w]\\ A_{21}[w,w] & A_{22}[w,w]\end{pmatrix}$$

and similarly for $B[w,z]$ and $C[z,z]$. We assume that

$$A[w_1,w_2] = A[w_2,w_1], \quad C[z_1,z_2] = C[z_2,z_1] \quad \forall w_1,w_2,z_1,z_2.$$

We calculate how system (3.13) transforms under the change of variable $(\eta,\psi) = \Phi^{(4)}(w,z)$. One has

$$\partial_t\begin{pmatrix}\eta\\ \psi\end{pmatrix} = (I+M(w,z))\begin{pmatrix}\partial_t w\\ \partial_t z\end{pmatrix} + \{\partial_t M(w,z)\}\begin{pmatrix}w\\ z\end{pmatrix}$$

and

$$\begin{aligned}\partial_t M(w,z) &= \partial_t(A[w,w] + B[w,z] + C[z,z])\\ &= 2A[w,\partial_t w] + B[\partial_t w,z] + B[w,\partial_t z] + 2C[z,\partial_t z].\end{aligned}$$

Thus

$$\partial_t\begin{pmatrix}\eta\\ \psi\end{pmatrix} = (I+K(w,z))\begin{pmatrix}\partial_t w\\ \partial_t z\end{pmatrix},$$

where

$$K(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = M(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \{2A[w,\alpha]+B[\alpha,z]+B[w,\beta]+2C[z,\beta]\}\begin{pmatrix}w\\ z\end{pmatrix}. \tag{4.6}$$

System (3.13), namely (4.1), becomes

$$(I+K(w,z))\begin{pmatrix}\partial_t w\\ \partial_t z\end{pmatrix} = X(\Phi^{(4)}(w,z)). \tag{4.7}$$

Assume that, by Neumann series, $I+K(w,z)$ is invertible (this will be proved below, after the choice of $M(w,z)$). Thus (4.7) becomes

$$\partial_t\begin{pmatrix}w\\ z\end{pmatrix} = (I+K(w,z))^{-1}X(\Phi^{(4)}(w,z)) =: X^+(w,z). \tag{4.8}$$

Since $X = \mathcal{D}_1 + \mathcal{D}_{\geq 3} + \mathcal{B}_3 + \mathcal{R}_{\geq 5}$ and $(I+K(w,z))^{-1} = I - K(w,z) + \tilde{K}(w,z)$, where $\tilde{K}(w,z) := \sum_{n=2}^{\infty}(-K(w,z))^n$, we calculate

$$\begin{aligned} X^+(w,z) = {} & \mathcal{D}_1(w,z) + \mathcal{D}_1\Big(M(w,z)\begin{pmatrix}w\\ z\end{pmatrix}\Big) - K(w,z)\mathcal{D}_1(w,z) \\ & - K(w,z)\mathcal{D}_1\Big(M(w,z)\begin{pmatrix}w\\ z\end{pmatrix}\Big) + \tilde{K}(w,z)\mathcal{D}_1(\Phi^{(4)}(w,z)) + \mathcal{B}_3(w,z) \\ & + (I+K(w,z))^{-1}\mathcal{D}_{\geq 3}(\Phi^{(4)}(w,z)) + (I+K(w,z))^{-1}\mathcal{R}_{\geq 5}(\Phi^{(4)}(w,z)) \\ & + [\mathcal{B}_3(\Phi^{(4)}(w,z)) - \mathcal{B}_3(w,z)] + \big(-K(w,z)+\tilde{K}(w,z)\big)\mathcal{B}_3(\Phi^{(4)}(w,z)). \end{aligned} \tag{4.9}$$

We look for $M(w,z)$ such that the cubic terms

$$X_3^+(w,z) := \mathcal{D}_1\Big(M(w,z)\begin{pmatrix}w\\ z\end{pmatrix}\Big) - K(w,z)\mathcal{D}_1(w,z) + \mathcal{B}_3(w,z) \tag{4.10}$$

give no contribution to the energy estimate. Note that X_3^+ is not the entirety of the cubic terms of X^+, because a cubic term also arises from $(I+K(w,z))^{-1}\mathcal{D}_{\geq 3}(\Phi^{(4)}(w,z))$; however, this cubic term is diagonal, it does not contribute to the energy estimate, and it does not interact with the off-diagonal cubic term $\mathcal{B}_3(w,z)$, therefore we do not include it in (4.10).

The first component $(X_3^+)_1(w,z)$ of the vector $X_3^+(w,z)$ in (4.10) is

$$\begin{aligned} (X_3^+)_1(w,z) = {} & -i\Lambda M_{11}(w,z)w - i\Lambda M_{12}(w,z)z + iM_{11}(w,z)\Lambda w - iM_{12}(w,z)\Lambda z \\ & - \{2A_{11}[w,-i\Lambda w] + B_{11}[-i\Lambda w,z] + B_{11}[w,i\Lambda z] + 2C_{11}[z,i\Lambda z]\}w \\ & - \{2A_{12}[w,-i\Lambda w] + B_{12}[-i\Lambda w,z] + B_{12}[w,i\Lambda z] + 2C_{12}[z,i\Lambda z]\}z \\ & + \frac{i}{4}\Big(\langle\Lambda z,\Lambda z\rangle - \langle\Lambda w,\Lambda w\rangle\Big)z. \end{aligned}$$

We choose

$$M_{11} = 0, \quad B_{12} = 0,$$

because M_{11} is not involved in the calculation to remove the off-diagonal terms (those ending with z), and there are no terms of the form [coefficient $O(wz)$ times z] to remove. It remains

$$(X_3^+)_1(w,z) = -i\Lambda A_{12}[w,w]z - i\Lambda C_{12}[z,z]z - iA_{12}[w,w]\Lambda z - iC_{12}[z,z]\Lambda z \\ + 2iA_{12}[w,\Lambda w]z - 2iC_{12}[z,\Lambda z]z + \frac{i}{4}\big(\langle \Lambda z, \Lambda z\rangle - \langle \Lambda w, \Lambda w\rangle\big) z.$$

We look for A_{12}, C_{12} of the form

$$A_{12}[u,v]h = \sum_{j,k\in\mathbb{Z}^d\setminus\{0\}} u_j v_{-j} a_{12}(j,k) h_k e^{ik\cdot x} \quad \forall u,v,h,$$
$$C_{12}[u,v]h = \sum_{j,k\in\mathbb{Z}^d\setminus\{0\}} u_j v_{-j} c_{12}(j,k) h_k e^{ik\cdot x} \quad \forall u,v,h,$$

for some coefficients $a_{12}(j,k), c_{12}(j,k)$ to be determined, where u_j, v_j, h_k are the Fourier coefficients of any functions $u(x), v(x), h(x)$. Hence

$$(X_3^+)_1(w,z) = \sum_{j,k\neq 0} w_j w_{-j} z_k e^{ik\cdot x}\Big(2i(|j|-|k|)a_{12}(j,k) - \frac{i}{4}|j|^2\Big) \\ + \sum_{j,k\neq 0} z_j z_{-j} z_k e^{ik\cdot x}\Big(-2i(|j|+|k|)c_{12}(j,k) + \frac{i}{4}|j|^2\Big).$$

We fix

$$a_{12}(j,k) := \begin{cases} \dfrac{|j|^2}{8(|j|-|k|)} & \text{if } |j|\neq|k|, \\ 0 & \text{if } |j| = |k|, \end{cases} \qquad c_{12}(j,k) := \frac{|j|^2}{8(|j|+|k|)}. \tag{4.11}$$

Thus the operators A_{12}, C_{12} are

$$A_{12}[u,v]h = \sum_{j,k\neq 0,\, |j|\neq|k|} u_j v_{-j} \frac{|j|^2}{8(|j|-|k|)} h_k e^{ik\cdot x}, \tag{4.12}$$

$$C_{12}[u,v]h = \sum_{j,k\neq 0} u_j v_{-j} \frac{|j|^2}{8(|j|+|k|)} h_k e^{ik\cdot x}, \tag{4.13}$$

and

$$(X_3^+)_1(w,z) = -\frac{i}{4} \sum_{j,k\neq 0,\, |k|=|j|} w_j w_{-j} |j|^2 z_k e^{ik\cdot x}. \tag{4.14}$$

The analogous calculation for the second component $(X_3^+)_2(w,z)$ of the vector in (4.10) leads to the choice

$$M_{22} = 0, \quad B_{21} = 0, \quad A_{21} = C_{12}, \quad C_{21} = A_{12}, \tag{4.15}$$

and it remains

$$(X_3^+)_2(w,z) = \frac{i}{4} \sum_{j,k\neq 0,\, |k|=|j|} z_j z_{-j} |j|^2 w_k e^{ik\cdot x}. \tag{4.16}$$

We will see below (see (4.43)) that the remaining cubic terms $(X_3^+)_1(w,z)$ and $(X_3^+)_2(w,z)$ do not contribute to the growth of the Sobolev norms in the energy estimate.

Now that M has been fixed, we have to prove the invertibility of $(I + K(w,z))$ by Neumann series. Since

$$M(w,z) = \begin{pmatrix} 0 & A_{12}[w,w] + C_{12}[z,z] \\ A_{21}[w,w] + C_{21}[z,z] & 0 \end{pmatrix}, \tag{4.17}$$

recalling (4.6) one has

$$K(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix} A_{12}[w,w]\beta + C_{12}[z,z]\beta + 2A_{12}[w,\alpha]z + 2C_{12}[z,\beta]z \\ A_{21}[w,w]\alpha + C_{21}[z,z]\alpha + 2A_{21}[w,\alpha]w + 2C_{21}[z,\beta]w \end{pmatrix}, \tag{4.18}$$

namely

$$K(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = M(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} + E(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

where $M(w,z)$ is given in (4.17) and

$$E(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} := \begin{pmatrix} 2A_{12}[w,\alpha]z + 2C_{12}[z,\beta]z \\ 2A_{21}[w,\alpha]w + 2C_{21}[z,\beta]w \end{pmatrix}.$$

To estimate matrix operators and vectors in $H_0^s(\mathbb{T}^d, c.c.)$, we define $\|(w,z)\|_s := \|w\|_s = \|z\|_s$ for every pair $(w,z) = (w,\overline{w})$ of complex conjugate functions.

Lemma 4.1. *Let A_{12}, C_{12} be the operators defined in* (4.12), (4.13), *and let m_0 be defined in* (1.5). *For all complex functions u,v,h, all real $s \geq 0$,*

$$\|A_{12}[u,v]h\|_s \leq \frac{3}{8}\|u\|_{m_0}\|v\|_{m_0}\|h\|_s, \quad \|C_{12}[u,v]h\|_s \leq \frac{1}{16}\|u\|_1\|v\|_1\|h\|_s. \tag{4.19}$$

Proof. In dimension $d = 1$, one has $||j| - |k|| \geq 1$ for $|j| \neq |k|$. Therefore, by Hölder's inequality,

$$\begin{aligned}\|A_{12}[u,v]h\|_s^2 &= \sum_{k\neq 0}\Big|\sum_{j\neq 0,\,|j|\neq|k|} u_j v_{-j}\frac{|j|^2}{8(|j|-|k|)}h_k\Big|^2|k|^{2s} \\ &\leq \frac{1}{64}\sum_{k\neq 0}\Big(\sum_{j\neq 0,\,|j|\neq|k|}|u_j||j||v_{-j}||j|\Big)^2|h_k|^2|k|^{2s} \leq \frac{1}{64}\|u\|_1^2\|v\|_1^2\|h\|_s^2.\end{aligned}$$

In dimension $d \geq 2$, we observe that

$$\frac{1}{||j|-|k||} \leq 3|j| \quad \forall j,k \in \mathbb{Z}^d \setminus \{0\},\ |j| \neq |k|. \tag{4.20}$$

If $||j| - |k|| \geq 1$, then (4.20) holds because $|j| \geq 1$. Let $||j| - |k|| < 1$, with $|j| \neq |k|$. Then $|k| < |j| + 1$, and, since $(|j| - |k|)(|j| + |k|) = |j|^2 - |k|^2$ is a nonzero integer, one has

$$\frac{1}{||j|-|k||} = \frac{|j|+|k|}{||j|^2-|k|^2|} \leq |j| + |k| < 2|j| + 1 \leq 3|j|.$$

Hence

$$\|A_{12}[u,v]h\|_s^2 = \sum_{k\neq 0}\Big|\sum_{j\neq 0,\,|j|\neq|k|} u_j v_{-j}\frac{|j|^2}{8(|j|-|k|)}h_k\Big|^2|k|^{2s}$$
$$\le \sum_{k\neq 0}\Big(\sum_{j\neq 0,\,|j|\neq|k|}|u_j||v_{-j}|\frac{3}{8}|j|^3\Big)^2|h_k|^2|k|^{2s} \le \frac{9}{64}\|u\|_{\frac{3}{2}}^2\|v\|_{\frac{3}{2}}^2\|h\|_s^2.$$

To estimate C_{12}, we use the bound $8(|j|+|k|)\ge 16$, which holds in any dimension. ☐

Lemma 4.2. *For all $s\ge 0$, all $(w,z)\in H_0^{m_0}(\mathbb{T}^d, c.c.)$, $(\alpha,\beta)\in H_0^s(\mathbb{T}^d, c.c.)$ one has*

$$\Big\|M(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\Big\|_s \le \frac{7}{16}\|w\|_{m_0}^2\|\alpha\|_s, \tag{4.21}$$
$$\Big\|K(w,z)\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\Big\|_s \le \frac{7}{16}\|w\|_{m_0}^2\|\alpha\|_s + \frac{7}{8}\|w\|_{m_0}\|w\|_s\|\alpha\|_{m_0}, \tag{4.22}$$

where m_0 is defined in (1.5). *For $\|w\|_{m_0}<\frac{1}{2}$, the operator $(I+K(w,z)) : H_0^{m_0}(\mathbb{T}^d, c.c.) \to H_0^{m_0}(\mathbb{T}^d, c.c.)$ is invertible, with inverse*

$$(I+K(w,z))^{-1} = I - K(w,z) + \tilde K(w,z), \quad \tilde K(w,z) := \sum_{n=2}^{\infty}(-K(w,z))^n,$$

satisfying

$$\Big\|(I+K(w,z))^{-1}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\Big\|_s \le C(\|\alpha\|_s + \|w\|_{m_0}\|w\|_s\|\alpha\|_{m_0}),$$

for all $s\ge 0$, where C is a universal constant.

Proof. Use (4.17), (4.18), (4.19) and Neumann series. ☐

By contraction lemma, we prove that the nonlinear, continuous map $\Phi^{(4)}$ is invertible in a ball around the origin.

Lemma 4.3. *For all $(\eta,\psi)\in H_0^{m_0}(\mathbb{T}^d, c.c.)$ in the ball $\|\eta\|_{m_0}\le\frac{1}{4}$, there exists a unique $(w,z)\in H_0^{m_0}(\mathbb{T}^d, c.c.)$ such that $\Phi^{(4)}(w,z)=(\eta,\psi)$, with $\|w\|_{m_0}\le 2\|\eta\|_{m_0}$. If, in addition, $\eta\in H_0^s$ for some $s>m_0$, then w also belongs to H_0^s, and $\|w\|_s\le 2\|\eta\|_s$. This defines the continuous inverse map $(\Phi^{(4)})^{-1} : H_0^s(\mathbb{T}^d, c.c.)\cap\{\|\eta\|_{m_0}\le\frac{1}{4}\} \to H_0^s(\mathbb{T}^d, c.c.)$.*

Proof. Existence. Given (η,ψ), the problem of finding (w,z) such that $\Phi^{(4)}(w,z)=(\eta,\psi)$ is the fixed point problem $\Psi(w,z)=(w,z)$, where

$$\Psi(w,z) := \begin{pmatrix}\eta\\ \psi\end{pmatrix} - M(w,z)\begin{pmatrix}w\\ z\end{pmatrix}.$$

Let $B_R := \{(w,z) \in H_0^{m_0}(\mathbb{T}^d, c.c.) : \|w\|_{m_0} \le R\}$. By (4.21), Ψ maps $B_R \to B_R$ if $\|\eta\|_{m_0} + \frac{7}{16}R^3 \le R$. Since

$$\begin{aligned} & M(w_1,z_1)\begin{pmatrix} w_1 \\ z_1 \end{pmatrix} - M(w_2,z_2)\begin{pmatrix} w_2 \\ z_2 \end{pmatrix} \\ & \quad = \int_0^1 K\big(w_2 + \vartheta(w_1-w_2), z_2 + \vartheta(z_1-z_2)\big)\, d\vartheta \begin{pmatrix} w_1 - w_2 \\ z_1 - z_2 \end{pmatrix}, \end{aligned} \tag{4.23}$$

by (4.22) Ψ is a contraction if $\frac{21}{16}R^2 < 1$. We choose $R = 2\|\eta\|_{m_0}$, so that Ψ is a contraction in B_R if $\|\eta\|_{m_0} \le \frac{1}{4}$. As a consequence, there exists a unique fixed point $(w,z) = \Psi(w,z)$ in B_R, with $\|w\|_{m_0} \le R = 2\|\eta\|_{m_0}$.

Regularity. Assume, in addition, that $\eta \in H^s$. The fixed point w is the limit in H^{m_0} of the sequence $w_n := \Psi(w_{n-1})$, $w_0 := 0$. We write w as the sum of the telescoping series $\sum_{n=0}^\infty h_n$, which converges in H^{m_0}, where $h_n := w_{n+1} - w_n$. Since $\eta \in H^s$ and Ψ maps $H^s \to H^s$, then $w_n \in H^s$ for all n. By (4.23),

$$\|h_n\|_{m_0} \le B^n \|h_0\|_{m_0} \quad \forall n \ge 0, \tag{4.24}$$

where $B := \frac{21}{16}R^2$. Note that $h_0 = w_1 = \eta$. By induction, we prove that

$$(i)\ \|w_n\|_s \le \rho_s; \qquad (ii)\ \|h_n\|_s \le B^n\|h_0\|_s + nB^{n-1}A_s\|h_0\|_{m_0} \tag{4.25}$$

for some constants ρ_s, A_s to determine.

At $n=0$ (4.25) trivially holds. At $n=1$, (i) holds if $\rho_s \ge \|\eta\|_s$, and (ii) holds because, by (4.21), $\|h_1\|_s = \|M(\eta,\psi)\binom{\eta}{\psi}\|_s \le \frac{7}{16}\|\eta\|_{m_0}^2\|\eta\|_s$ and $h_0 = \eta$.

Assume that (4.25) holds for all $k \le n$, for some $n \ge 1$. Using (4.23), (4.22), $(i)_n$ and $(i)_{n-1}$, we deduce that $\|h_{n+1}\|_s \le \frac{7}{16}R^2\|h_n\|_s + \frac{7}{8}R\rho_s\|h_n\|_{m_0}$. Using $(ii)_n$ and (4.24), this is $\le (\frac{7}{16}R^2B^n)\|h_0\|_s + (\frac{7}{16}R^2 nB^{n-1}A_s + \frac{7}{8}R\rho_s B^n)\|h_0\|_{m_0}$. Since $B = \frac{21}{16}R^2$, $(ii)_{n+1}$ holds provided that $\frac{7}{8}R\rho_s \le A_s$. We fix $A_s = \frac{7}{8}R\rho_s$.

To prove $(i)_{n+1}$, we use $(ii)_k$ for $k = 0,\ldots,n$, and we estimate $\|w_{n+1}\|_s \le \sum_{k=0}^n \|h_k\|_s \le \sum_{k=0}^n B^k\|h_0\|_s + \sum_{k=0}^n kB^{k-1}A_s\|h_0\|_{m_0} \le \frac{1}{1-B}\|h_0\|_s + \frac{1}{(1-B)^2}\frac{7}{8}R\rho_s\|h_0\|_{m_0}$. Hence $(i)_{n+1}$ holds by choosing $\rho_s = 2\|\eta\|_s$. The proof of (4.25) is complete.

As a consequence, w_n is a Cauchy sequence in H^s, and its limit w satisfies $\|w\|_s \le \rho_s = 2\|\eta\|_s$.

Continuity. The inverse map $(\Phi^{(4)})^{-1}$ is Lipschitz-continuous because it is constructed as a solution of the fixed point problem (recall (4.23)). ☐

Lemma 4.4. *For all complex functions u, v, y, h, one has*

$$\langle A_{12}[u,v]y, h\rangle = \langle y, A_{12}[u,v]h\rangle, \qquad \langle C_{12}[u,v]y, h\rangle = \langle y, C_{12}[u,v]h\rangle, \tag{4.26}$$

$$\overline{A_{12}[u,v]y} = A_{12}[\overline{u},\overline{v}]\overline{y}, \qquad \overline{C_{12}[u,v]y} = C_{12}[\overline{u},\overline{v}]\overline{y}, \tag{4.27}$$

$$A_{12}[u,v]\Lambda^s y = \Lambda^s A_{12}[u,v]y, \qquad C_{12}[u,v]\Lambda^s y = \Lambda^s C_{12}[u,v]y \tag{4.28}$$

where $\overline{u}$ is the complex conjugate of u, and so on. As a consequence, for all complex functions w, z, y, h, one has

$$\langle M_{12}(w,z)y, h\rangle = \langle y, M_{12}(w,z)h\rangle, \quad \langle M_{21}(w,z)y, h\rangle = \langle y, M_{21}(w,z)h\rangle, \tag{4.29}$$

$$\overline{M_{12}(w,z)h} = M_{12}(\overline{w},\overline{z})\overline{h}, \qquad \overline{M_{21}(w,z)h} = M_{21}(\overline{w},\overline{z})\overline{h}, \tag{4.30}$$

$$[M_{12}(w,z), \Lambda^s] = 0, \qquad [M_{21}(w,z), \Lambda^s] = 0. \tag{4.31}$$

Moreover, for all complex w, z, h*,*

$$M_{12}(w,z)h = M_{21}(z,w)h \tag{4.32}$$

and

$$M(w,z)\mathcal{D}_1 + \mathcal{D}_1 M(w,z) = 0. \tag{4.33}$$

Proof. All (4.26)-(4.32) directly follow from the definition (4.11) of the coefficients $a_{12}(j,k), c_{12}(j,k)$ and from (4.15), (4.17). The anti-commutator identity (4.33) follows from (4.31). □

Lemma 4.5. *The maps* $M(w,\overline{w})$*,* $K(w,\overline{w})$*, and the transformation* $\Phi^{(4)}$ *preserve the structure of real vector field* (2.8)*. Hence* X^+ *defined in* (4.8) *satisfies* (2.8)*.*

Proof. It follows from Lemma 4.4. □

For a system $\partial_t(w,\overline{w}) = \mathcal{F}(w,\overline{w})$ where the vector field $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2)$ satisfies (2.8), the Sobolev norm of any solution evolves in time according to

$$\begin{aligned}\partial_t(\|w\|_s^2) &= \langle \Lambda^s \mathcal{F}_1(w,\overline{w}), \Lambda^s \overline{w}\rangle + \langle \Lambda^s w, \Lambda^s \mathcal{F}_2(w,\overline{w})\rangle \\ &= 2\mathrm{Re}\,\langle \Lambda^s \mathcal{F}_1(w,\overline{w}), \Lambda^s \overline{w}\rangle. \end{aligned} \tag{4.34}$$

The vector field X^+ in (4.9) is

$$\begin{aligned} X^+(w,z) = \mathcal{D}_1(w,z) - K(w,z)\mathcal{D}_1\Big(M(w,z)\begin{pmatrix} w \\ z\end{pmatrix}\Big) + \tilde{K}(w,z)\mathcal{D}_1(\Phi^{(4)}(w,z)) \\ + X_3^+(w,z) + (I + K(w,z))^{-1}\mathcal{D}_{\geq 3}(\Phi^{(4)}(w,z)) + \mathcal{R}^+_{\geq 5}(w,z) \end{aligned} \tag{4.35}$$

where

$$\begin{aligned}\mathcal{R}^+_{\geq 5}(w,z) := (I + K(w,z))^{-1}\mathcal{R}_{\geq 5}(\Phi^{(4)}(w,z)) + [\mathcal{B}_3(\Phi^{(4)}(w,z)) - \mathcal{B}_3(w,z)] \\ + \big(-K(w,z) + \tilde{K}(w,z)\big)\mathcal{B}_3(\Phi^{(4)}(w,z)).\end{aligned}$$

By (4.33), equation (4.10) becomes

$$\big(M(w,z) + K(w,z)\big)\mathcal{D}_1\begin{pmatrix} w \\ z\end{pmatrix} = \mathcal{B}_3(w,z) - X_3^+(w,z). \tag{4.36}$$

We use (4.36) to rewrite the terms in (4.35) containing $\mathcal{D}_1$, $\mathcal{D}_{\geq 3}$. At a first glance, these terms seem to be unbounded, as $\mathcal{D}_1, \mathcal{D}_{\geq 3}$ are operators of order one, but, using (4.36), it becomes clear that they are, in fact, bounded. Omitting to write $\binom{w}{z}$ and (w,z), identity (4.36) is $(M+K)\mathcal{D}_1 = \mathcal{B}_3 - X_3^+$, the anti-commutator formula (4.33) is $M\mathcal{D}_1 + \mathcal{D}_1 M = 0$, and therefore we have

$$\begin{aligned}
-K\mathcal{D}_1 M + \tilde{K}\mathcal{D}_1(I+M) &= KM\mathcal{D}_1 + \sum_{n=2}^{\infty}(-K)^n\mathcal{D}_1 + \sum_{n=2}^{\infty}(-K)^n\mathcal{D}_1 M \\
&= KM\mathcal{D}_1 + \sum_{n=2}^{\infty}(-K)^n\mathcal{D}_1 - \sum_{n=2}^{\infty}(-K)^n M\mathcal{D}_1 \\
&= -\sum_{n=1}^{\infty}(-K)^n M\mathcal{D}_1 - \sum_{n=1}^{\infty}(-K)^n K\mathcal{D}_1 \\
&= -\sum_{n=1}^{\infty}(-K)^n (M+K)\mathcal{D}_1 \\
&= K(I+K)^{-1}(\mathcal{B}_3 - X_3^+). \qquad (4.37)
\end{aligned}$$

Regarding the terms with $\mathcal{D}_{\geq 3}$, recalling (4.2) one has

$$\mathcal{D}_{\geq 3}(\Phi^{(4)}(w,z)) = \mathcal{P}(w,z)\mathcal{D}_1(\Phi^{(4)}(w,z)) \qquad (4.38)$$

where

$$\mathcal{P}(w,z) := \sqrt{1 + 2P(\Phi^{(4)}(w,z))} - 1. \qquad (4.39)$$

We recall that P is defined in (3.8), (3.2), (3.7), and it is a function of t only (i.e., it does not depend on x). We write (4.38) as $\mathcal{D}_{\geq 3}(I+M) = \mathcal{P}\mathcal{D}_1(I+M)$, where $\mathcal{P}$ is the multiplication operator $\mathcal{P}h = \mathcal{P}(w,z)h$. Using the identities $(M+K)\mathcal{D}_1 = \mathcal{B}_3 - X_3^+$ and $M\mathcal{D}_1 + \mathcal{D}_1 M = 0$, namely (4.36) and (4.33), and the fact that $\mathcal{P}K = K\mathcal{P}$ (because $\mathcal{P}$ is a function of time only), we calculate

$$\begin{aligned}
(I+K)^{-1}\mathcal{D}_{\geq 3}\Phi^{(4)} &= (I+K)^{-1}\mathcal{P}\mathcal{D}_1(I+M) \\
&= \mathcal{P}\Big(\sum_{n=0}^{\infty}(-K)^n\mathcal{D}_1 + \sum_{n=0}^{\infty}(-K)^n\mathcal{D}_1 M\Big) \\
&= \mathcal{P}\mathcal{D}_1 - \mathcal{P}\Big(\sum_{n=0}^{\infty}(-K)^n K\mathcal{D}_1 + \sum_{n=0}^{\infty}(-K)^n M\mathcal{D}_1\Big) \\
&= \mathcal{P}\mathcal{D}_1 - \mathcal{P}(I+K)^{-1}(K+M)\mathcal{D}_1 \\
&= \mathcal{P}\mathcal{D}_1 - \mathcal{P}(I+K)^{-1}(\mathcal{B}_3 - X_3^+). \qquad (4.40)
\end{aligned}$$

By (4.37) and (4.40), the vector field X^+ in (4.35) becomes

$$X^+(w,z) = \big(1 + \mathcal{P}(w,z)\big)\mathcal{D}_1(w,z) + X_3^+(w,z) + X_{\geq 5}^+(w,z) \qquad (4.41)$$

where

$$\begin{aligned}
X_{\geq 5}^+(w,z) &:= K(w,z)\big(I + K(w,z)\big)^{-1}\big(\mathcal{B}_3(w,z) - X_3^+(w,z)\big) + \mathcal{R}_{\geq 5}^+(w,z) \\
&\quad - \mathcal{P}(w,z)\big(I + K(w,z)\big)^{-1}\big(\mathcal{B}_3(w,z) - X_3^+(w,z)\big). \qquad (4.42)
\end{aligned}$$

To analyze the energy estimate (4.34) for $\mathcal{F} = X^+$, we prove that the contribution of $(1+\mathcal{P})\mathcal{D}_1$ and X_3^+ is zero, and the one of all the other terms is quintic. Since $\mathcal{P} = \mathcal{P}(w,z)$ is a function of time only, one simply has

$$\langle \Lambda^s(1+\mathcal{P})(-i\Lambda w), \Lambda^s z\rangle + \langle \Lambda^s w, \Lambda^s(1+\mathcal{P})i\Lambda z\rangle = 0.$$

Next, recalling (4.14) and (4.16), one has

$$\begin{aligned}&\langle \Lambda^s(X_3^+)_1, \Lambda^s z\rangle + \langle \Lambda^s w, \Lambda^s(X_3^+)_2\rangle \\ &= -\frac{i}{4}\sum_{\substack{j,k\in\mathbb{Z}^d\setminus\{0\}\\ |k|=|j|}} w_j w_{-j}|j|^2 z_k |k|^{2s} z_{-k} + \frac{i}{4}\sum_{\substack{j,k\in\mathbb{Z}^d\setminus\{0\}\\ |k|=|j|}} z_j z_{-j}|j|^2 w_k|k|^{2s}w_{-k} = 0 \end{aligned} \tag{4.43}$$

(rename $j \leftrightarrow k$ in the second sum and use $|j| = |k|$). To estimate the contribution of $X_{\geq 5}^+$, we collect a few elementary estimates in the next lemma.

Lemma 4.6. *For all $s \geq 0$, all pairs of complex conjugate functions (w,z), one has*

$$\|\mathcal{B}_3(w,z)\|_s \leq \frac{1}{2}\|w\|_1^2\|w\|_s, \quad \|X_3^+(w,z)\|_s \leq \frac{1}{4}\|w\|_1^2\|w\|_s, \tag{4.44}$$

and, for $\|w\|_{m_0} \leq \frac{1}{2}$, for all complex functions h,

$$\|\mathcal{P}(w,z)h\|_s = \mathcal{P}(w,z)\|h\|_s, \quad 0 \leq \mathcal{P}(w,z) \leq C\|w\|_{\frac{1}{2}}^2, \tag{4.45}$$

$$\|\mathcal{R}_{\geq 5}(w,z)\|_s \leq 2P(w,z)\|\mathcal{B}_3(w,z)\|_s \leq C\|w\|_{\frac{1}{2}}^2\|w\|_1^2\|w\|_s \tag{4.46}$$

where $\mathcal{R}_{\geq 5}$ is defined in (4.4) *and C is a universal constant.*

Proof. Estimate (4.44) follows from (4.3), (4.14), (4.16). To prove (4.45) and (4.46), recall (4.39), (3.8), (3.2), (3.7), (4.5), (4.21). □

Lemma 4.7. *For all $s \geq 0$, all $(w,z) \in H_0^s(\mathbb{T}^d, c.c.) \cap H_0^{m_0}(\mathbb{T}^d, c.c.)$ with $\|w\|_{m_0} \leq \frac{1}{2}$, one has*

$$\|X_{\geq 5}^+(w,z)\|_s \leq C\|w\|_1^2\|w\|_{m_0}^2\|w\|_s$$

where C is a universal constant.

Proof. Use (4.42), (4.22) and Lemma 4.6. □

As a consequence, we obtain the following improved energy estimate.

Lemma 4.8. *Let $T > 0$, $s \geq m_0$. Any solution $(w,\overline{w}) \in C^0([0,T], H_0^s(\mathbb{T}^d, c.c.))$ of the system $\partial_t(w,\overline{w}) = X^+(w,\overline{w})$ satisfies*

$$\partial_t(\|w\|_s^2) \leq C_*\|w\|_1^2\|w\|_{m_0}^2\|w\|_s^2 \tag{4.47}$$

as long as it remains in the ball $\|w\|_{m_0} \leq \frac{1}{2}$, for some universal constant $C_* > 0$.

5 Proof of Theorem 1.1

We now perform the composition of all the changes of variables defined in the previous sections, namely we define

$$\Phi = \Phi^{(1)} \circ \Phi^{(2)} \circ \Phi^{(3)} \circ \Phi^{(4)},$$

where $\Phi^{(1)}$, $\Phi^{(2)}$, $\Phi^{(3)}$ and $\Phi^{(4)}$ have been defined in (2.2), (2.5), (3.5), (3.10), (4.5). The definitions of $\Phi^{(1)}$ and $\Phi^{(2)}$, together with Lemma 3.1 and Lemma 4.3, directly imply the following lemma.

Lemma 5.1. *There exist universal constants $\delta_0 \in (0, \frac{1}{4})$, $C_0 > 0$ such that, for all $s \geq m_0$ (where m_0 is defined in* (1.5)*), for all pairs of zero mean real functions $(u, v) \in H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$ satisfying*

$$\|u\|_{m_0+\frac{1}{2}} + \|v\|_{m_0-\frac{1}{2}} \leq \delta_0,$$

there exists a unique pair $(w, z) = (w, \overline{w}) \in H_0^s(\mathbb{T}^d, c.c.)$ such that $(u, v) = \Phi(w, \overline{w})$. Moreover, $(w, \overline{w}) = \Phi^{-1}(u, v)$ satisfies the estimate

$$\|w\|_s \leq C_0\big(\|u\|_{s+\frac{1}{2}} + \|v\|_{s-\frac{1}{2}}\big).$$

Conversely, if $w \in H_0^s(\mathbb{T}^d, \mathbb{C})$ satisfies

$$\|w\|_{m_0} \leq \delta_0,$$

then $(u, v) = \Phi(w, \overline{w}) \in H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$ is a pair of zero mean real functions satisfying

$$\|u\|_{s+\frac{1}{2}} + \|v\|_{s-\frac{1}{2}} \leq C_0\|w\|_s.$$

As a consequence, in the following corollary we deduce the equivalence of the Kirchhoff equation (1.3) and the transformed system (4.8).

Corollary 5.2. *Let $\delta_0, C_0 > 0$ be given by Lemma 5.1. Then, for all $s \geq m_0$, if $u \in C^0([0, T], H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})) \cap C^1([0, T], H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}))$ is a solution of equation* (1.3) *on some time interval $[0, T]$ with*

$$\max_{t\in[0,T]} \big(\|u(t)\|_{m_0+\frac{1}{2}} + \|\partial_t u(t)\|_{m_0-\frac{1}{2}}\big) \leq \delta_0,$$

then the pair $(w, z) = (w, \overline{w}) \in C^0([0, T], H_0^s(\mathbb{T}^d, c.c.))$ defined as $(w(t), \overline{w}(t)) = \Phi^{-1}(u(t), \partial_t u(t))$ is a solution of system (4.8)*, satisfying*

$$\max_{t\in[0,T]} \|w(t)\|_s \leq C_0 \max_{t\in[0,T]} \big(\|u(t)\|_{s+\frac{1}{2}} + \|\partial_t u(t)\|_{s-\frac{1}{2}}\big).$$

Conversely, if $(w, \overline{w}) \in C^0([0, T], H_0^s(\mathbb{T}^d, c.c.))$ is a solution of system (4.8) *satisfying*

$$\max_{t\in[0,T]} \|w(t)\|_{m_0} \leq \delta_0,$$

then the pair of real functions $(u,v) \in C^0([0,T], H_0^{s+\frac{1}{2}}(\mathbb{T}^d,\mathbb{R})) \times C^0([0,T], H_0^{s-\frac{1}{2}}(\mathbb{T}^d,\mathbb{R}))$ *defined as* $(u(t),v(t)) = \Phi(w(t),\overline{w}(t))$ *satisfies* $v = \partial_t u$ *and* $u \in C^0([0,T], H_0^{s+\frac{1}{2}}(\mathbb{T}^d)) \cap C^1([0,T], H_0^{s-\frac{1}{2}}(\mathbb{T}^d))$ *is a solution of equation* (1.3) *satisfying*

$$\max_{t\in[0,T]} \big(\|u(t)\|_{s+\frac{1}{2}} + \|\partial_t u(t)\|_{s-\frac{1}{2}} \big) \leq C_0 \max_{t\in[0,T]} \|w(t)\|_s.$$

By a repeated use of Lemma 5.1 and Corollary 5.2 we prove Theorem 1.1.

Proof of Theorem 1.1. The classical local existence and uniqueness theory for the Kirchhoff equation (1.3) (see [3]) and Corollary 5.2 imply the local existence and uniqueness for system (4.8), for every initial data $(w_0, \overline{w_0})$ in the ball $\|w_0\|_{m_0} \leq \delta_0$.

Let $(\alpha,\beta) \in H_0^{m_0+\frac{1}{2}}(\mathbb{T}^d,\mathbb{R}) \times H_0^{m_0-\frac{1}{2}}(\mathbb{T}^d,\mathbb{R})$ with

$$\varepsilon := \|\alpha\|_{m_0+\frac{1}{2}} + \|\beta\|_{m_0-\frac{1}{2}} \leq \varepsilon_0 := \frac{\delta_0}{2C_0}.$$

Let $(w_0, \overline{w_0}) := \Phi^{-1}(\alpha,\beta)$. By Lemma 5.1, one has $\|w_0\|_{m_0} \leq C_0\varepsilon \leq \frac{\delta_0}{2}$, and therefore the Cauchy problem for system (4.8) with initial data $(w_0,\overline{w_0})$ has a (unique) local solution $(w(t),\overline{w}(t))$, whose existence time can be extended as long as $w(t)$ remains in the ball $\|w\|_{m_0} \leq \delta_0$. By Lemma 4.8,

$$\partial_t(\|w(t)\|_{m_0}^2) \leq C_* \|w(t)\|_{m_0}^6.$$

Hence

$$\|w(t)\|_{m_0} \leq \frac{\|w_0\|_{m_0}}{(1 - 2C_*\|w_0\|_{m_0}^4 t)^{\frac{1}{4}}} \leq 2\|w_0\|_{m_0} \leq \delta_0$$

for all $t \in [0,T]$, with

$$T := \frac{C_1}{\varepsilon^4}, \qquad C_1 := \frac{15}{32 C_* C_0^4}.$$

Then $(u,v) := \Phi(w,\overline{w})$ belongs to $C^0([0,T], H_0^{m_0+\frac{1}{2}}(\mathbb{T}^d,\mathbb{R})) \times C^0([0,T], H_0^{m_0-\frac{1}{2}}(\mathbb{T}^d,\mathbb{R}))$ and solves (1.7), so that $u \in C^0([0,T], H_0^{m_0+\frac{1}{2}}(\mathbb{T}^d,\mathbb{R})) \cap C^1([0,T], H_0^{m_0-\frac{1}{2}}(\mathbb{T}^d,\mathbb{R}))$ solves (1.3) with initial data (α,β), and $\|u(t)\|_{m_0+\frac{1}{2}} + \|\partial_t u(t)\|_{m_0-\frac{1}{2}} \leq 2C_0^2\varepsilon$ for all $t \in [0,T]$.

If, in addition, $(\alpha,\beta) \in H_0^{s+\frac{1}{2}}(\mathbb{T}^d,\mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}^d,\mathbb{R})$ for some $s \geq m_0$, then $w_0 \in H_0^s(\mathbb{T}^d,\mathbb{C})$ and, by Lemma 4.8,

$$|\partial_t(\|w(t)\|_s^2)| \leq C_*\|w(t)\|_{m_0}^4 \|w(t)\|_s^2 \leq C_*(2\|w_0\|_{m_0})^4 \|w(t)\|_s^2$$

for all $t \in [0,T]$. Hence

$$\|w(t)\|_s \leq \|w_0\|_s \exp(8C_*\|w_0\|_{m_0}^4 t),$$

whence

$$\|u(t)\|_{s+\frac{1}{2}} + \|\partial_t u(t)\|_{s-\frac{1}{2}} \leq C_0^2 e^{\frac{15}{4}} (\|\alpha\|_{s+\frac{1}{2}} + \|\beta\|_{s-\frac{1}{2}})$$

for all $t \in [0,T]$. The proof of Theorem 1.1 is complete. □

References

[1] T. Alazard, J.-M. Delort, *Sobolev estimates for two dimensional gravity water waves*, Astérisque 374 (2015).

[2] A. Arosio, *Averaged evolution equations. The Kirchhoff string and its treatment in scales of Banach spaces*, in: 2nd Workshop on functional-analytic methods in complex analysis (Trieste, 1993), World Scientific, Singapore.

[3] A. Arosio, S. Panizzi, *On the well-posedness of the Kirchhoff string*, Trans. Amer. Math. Soc. 348 (1996), 305-330.

[4] A. Arosio, S. Spagnolo, *Global solutions of the Cauchy problem for a nonlinear hyperbolic equation*, in: Nonlinear PDE's and their applications, Collége de France Seminar, Vol. VI, 1–26, H. Brezis & J.L. Lions eds., Research Notes Math. 109, Pitman, Boston, 1984.

[5] P. Baldi, *Periodic solutions of forced Kirchhoff equations*, Ann. Sc. Norm. Sup. Pisa, Cl. Sci. (5), Vol. VIII (2009), 117-141.

[6] P. Baldi, M. Berti, R. Montalto, *KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation*, Math. Annalen 359 (2014), no. 1-2, 471-536.

[7] P. Baldi, M. Berti, R. Montalto, *KAM for autonomous quasi-linear perturbations of KdV*, Ann. Inst. H. Poincaré (C) Anal. Non Linéaire 33 (2016), 1589-1638.

[8] P. Baldi, M. Berti, E. Haus, R. Montalto, *Time quasi-periodic gravity water waves in finite depth*, Invent. Math. 214 (2018), no. 2, 739-911.

[9] D. Bambusi, *Galerkin averaging method and Poincaré normal form for some quasilinear PDEs*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 4 (2005), 669-702.

[10] D. Bambusi, J.-M. Delort, B. Grébert, J. Szeftel, *Almost global existence for Hamiltonian semilinear Klein-Gordon equations with small Cauchy data on Zoll manifolds*, Comm. Pure Appl. Math. 60 (2007), 1665-1690.

[11] D. Bambusi, B. Grébert, *Birkhoff normal form for partial differential equations with tame modulus*, Duke Math. J. 135 (2006), 507-567.

[12] D. Bambusi, N.N. Nekhoroshev, *A property of exponential stability in nonlinear wave equation near the fundamental linear mode*, Phys. D 122 (1998), 73-104.

[13] S.N. Bernstein, *Sur une classe d'équations fonctionnelles aux dérivées partielles*, Izv. Akad. Nauk SSSR Ser. Mat. 4 (1940), 17-26.

[14] M. Berti, J.-M. Delort, *Almost global solutions of capillary-gravity water waves equations on the circle*. Lecture Notes Unione Mat. Ital. 24, Springer, 2018.

[15] M. Berti, R. Feola, F. Pusateri, *Birkhoff normal form and long time existence for periodic gravity water waves*, arXiv:1810.11549.

[16] M. Berti, R. Montalto, *KAM for gravity capillary water waves*, Mem. AMS, to appear (arXiv:1602.02411).

[17] G.F. Carrier, *On the nonlinear vibration problem of the elastic string*, Quart. Appl. Math. 3 (1945), 157-165.

[18] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation*, Invent. Math. 181 (2010), 39-113.

[19] L. Corsi, R. Montalto, *Quasi-periodic solutions for the forced Kirchhoff equation on* $\mathbb{T}^d$, Nonlinearity 31 (2018), no. 11, 5075-5109.

[20] W. Craig, C. Sulem, *Mapping properties of normal forms transformations for water waves*, Boll. Unione Mat. Ital. 9 (2016), 289-318.

[21] P. D'Ancona, S. Spagnolo, *A class of nonlinear hyperbolic problems with global solutions*, Arch. Rational Mech. Anal. 124 (1993), 201-219.

[22] J.-M. Delort, *Long-time Sobolev stability for small solutions of quasi-linear Klein-Gordon equations on the circle*, Trans. Amer. Math. Soc. 361 (2009), no. 8, 4299-4365.

[23] J.-M. Delort, *A quasi-linear Birkhoff normal forms method. Application to the quasi-linear Klein-Gordon equation on* $\mathbb{S}^1$, Astérisque 341 (2012).

[24] R. W. Dickey, *Infinite systems of nonlinear oscillation equations related to the string*, Proc. Amer. Math. Soc. 23 (1969), 459-468.

[25] R. Feola, F. Giuliani, S. Pasquali, *On the integrability of Degasperis-Procesi equation: control of the Sobolev norms and Birkhoff resonances*, J. Differential Equations, published online 2018, doi: 10.1016/j.jde.2018.09.003.

[26] R. Feola, M. Procesi, *Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations*, J. Differential Equations 259 (2015), no. 7, 3389-3447.

[27] D. Fujiwara, *Concrete characterization of the domains of fractional powers of some elliptic differential operators of the second order*, Proc. Japan Acad. **43** (1967), 82–86.

[28] P. Germain, N. Masmoudi, J. Shatah, *Global solutions for the gravity water waves equation in dimension 3*, Ann. of Math. (2) 175 (2012), 691-754.

[29] M. Ghisi, M. Gobbino, *Kirchhoff equations from quasi-analytic to spectral-gap data*, Bull. Lond. Math. Soc. 43 (2011), 374-385.

[30] J.M. Greenberg, S.C. Hu, *The initial value problem for a stretched string*, Quart. Appl. Math. 38 (1980/81), 289-311.

[31] M. Guardia, E. Haus, M. Procesi, *Growth of Sobolev norms for the analytic NLS on* $\mathbb{T}^2$, Adv. Math. 301 (2016), 615-692.

[32] M. Guardia, V. Kaloshin, *Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation*, J. Eur. Math. Soc. (JEMS) 17 (2015), 71-149.

[33] E. Haus, M. Procesi, *Growth of Sobolev norms for the quintic NLS on* $\mathbb{T}^2$, Anal. PDE 8 (2015), 883-922.

[34] M. Ifrim, D. Tataru, *The lifespan of small data solutions in two dimensional capillary water waves*, Arch. Ration. Mech. Anal. 225 (2017), no. 3, 1279-1346.

[35] A. Ionescu, F. Pusateri, *Global solutions for the gravity water waves system in 2d*, Invent. Math. 199 (2015), 653-804.

[36] G. Kirchhoff, *Vorlesungen über mathematische Physik: Mechanik*, ch.29, Teubner, Leipzig, 1876.

[37] S. Klainerman, A. Majda, *Formation of singularities for wave equations including the nonlinear vibrating string*, Comm. Pure Appl. Math. 33 (1980), 241-263.

[38] J.L. Lions, *On some questions in boundary value problems of mathematical physics*, in: Contemporary developments in continuum mechanics and PDE's, G.M. de la Penha & L.A. Medeiros eds., North-Holland, Amsterdam, 1978.

[39] R. Manfrin, *On the global solvability of Kirchhoff equation for non-analytic initial data*, J. Differential Equations 211 (2005), 38-60.

[40] T. Matsuyama, *Asymptotic profiles for the Kirchhoff equation*, Rend. Lincei Mat. Appl. 17 (2006), no. 4, 377-395.

[41] T. Matsuyama, M. Ruzhansky, *Global well-posedness of Kirchhoff systems*, J. Math. Pures Appl. 100 (2013), 220-240.

[42] T. Matsuyama, M. Ruzhansky, *Global well-posedness of the Kirchhoff equation and Kirchhoff systems*, Analytic methods in interdisciplinary applications, 81-96, Springer Proc. Math. Stat., 116, Springer, Cham, 2015.

[43] T. Matsuyama, M. Ruzhansky, *On the Gevrey well-posedness of the Kirchhoff equation*, arXiv:1508.05305.

[44] L.A. Medeiros, M. Milla Miranda, *Solutions for the equation of nonlinear vibrations in Sobolev spaces of fractionary order*, Mat. Apl. Comput. 6 (1987), no. 3, 257-276.

[45] R. Montalto, *Quasi-periodic solutions of forced Kirchhoff equation*, NoDEA Nonlinear Differential Equations Appl. 24 (2017), Art. 9.

[46] R. Montalto, *A reducibility result for a class of linear wave equations on* $\mathbb{T}^d$, Int. Math. Res. Notices, rnx167, https://doi.org/10.1093/imrn/rnx167 (2017).

[47] R. Narasimha, *Nonlinear vibration of an elastic string*, J. Sound Vibration 8 (1968), 134-146.

[48] K. Nishihara, *On a global solution of some quasilinear hyperbolic equation*, Tokyo J. Math. 7 (1984), no. 2, 437-459.

[49] T. Ozawa, K. Tsutaya, Y. Tsutsumi, *Remarks on the Klein-Gordon equation with quadratic nonlinearity in two space dimensions*, Nonlinear waves (Sapporo, 1995), 383-392, GAKUTO Internat. Ser. Math. Sci. Appl., Vol. 10, Gakkōtosho, Tokyo, 1997.

[50] S.I. Pokhozhaev, *On a class of quasilinear hyperbolic equations*, Mat. Sbornik 96 (1975), 152-166 (English transl.: Mat. USSR Sbornik 25 (1975), 145-158).
[51] J. Shatah, *Normal forms and quadratic nonlinear Klein-Gordon equations*, Comm. Pure Appl. Math. 38 (1985), 685-696.
[52] S. Spagnolo, *The Cauchy problem for Kirchhoff equations*, Rend. Sem. Mat. Fis. Milano 62 (1994), 17-51.
[53] S. Wu, *Almost global wellposedness of the 2-D full water wave problem*, Invent. Math. 177 (2009), 45-135.
[54] T. Yamazaki, *Global solvability for the Kirchhoff equations in exterior domains of dimension larger than three*, Math. Methods Appl. Sci. 27 (2004), 1893-1916.

Pietro Baldi
Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università di Napoli Federico II, Via Cintia, Monte S. Angelo, 80126 Napoli, Italy
`pietro.baldi@unina.it`,

Emanuele Haus
Dipartimento di Matematica e Fisica, Università di Roma Tre, Largo S.L. Murialdo, 00146 Roma, Italy
`ehaus@mat.uniroma3.it`