Normal form and dynamics of the Kirchhoff equation

Pietro Baldi, Emanuele Haus

Abstract. We summarize some recent results on the Cauchy problem for the Kirchhoff equation

$$\partial_{tt}u - \Delta u \left(1 + \int_{\mathbb{T}^d} |\nabla u|^2 \right) = 0$$

on the *d*-dimensional torus \mathbb{T}^d , with initial data u(0, x), $\partial_t u(0, x)$ of size ε in Sobolev class. While the standard local theory gives an existence time of order ε^{-2} , a quasilinear normal form allows to give a lower bound on the existence time of the order of ε^{-4} for all initial data, improved to ε^{-6} for initial data satisfying a suitable nonresonance condition. We also use such a normal form in an ongoing work with F. Giuliani and M. Guardia to prove existence of chaotic-like motions for the Kirchhoff equation.

1 Introduction

The aim of this note is to summarize some recent results ([3], [4], [5]) and to introduce a work in progress ([2]) on the Kirchhoff equation on the *d*-dimensional torus \mathbb{T}^d , $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ (periodic boundary conditions)

$$\partial_{tt}u - \Delta u \left(1 + \int_{\mathbb{T}^d} |\nabla u|^2 \, dx \right) = 0, \tag{1.1}$$

where the unknown $u = u(t, x), x \in \mathbb{T}^d$, is a real-valued function, with initial data at time t = 0

$$u(0,x) = a(x), \quad \partial_t u(0,x) = b(x).$$
 (1.2)

Equation (1.1) was first introduced by Kirchhoff [29] in 1876, to model nonlinear transversal oscillations of strings and plates (d = 1, 2).

While it has long been known (Dickey [18], Arosio-Panizzi [1]) that such a Cauchy problem is locally wellposed for initial data (a, b) in the Sobolev space $H^{\frac{3}{2}}(\mathbb{T}^d, \mathbb{R}) \times H^{\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$, it is still an open problem whether the solutions of (1.1)-(1.2) of any given Sobolev regularity are global in time or not. In particular, it is not even known if C^{∞} (or even Gevrey) initial data of small amplitude produce solutions that are global in time (for initial data in analytic class, instead, global wellposedness is known since the work of Bernstein [10] in 1940). On the other hand, below the regularity threshold $H^{\frac{3}{2}} \times H^{\frac{1}{2}}$, neither local wellposedness nor illposedness have been established. A partial, interesting result in this direction has been recently obtained by Ghisi and Gobbino [20].

As a consequence of the linear theory, for initial data of size ε in $H^{\frac{3}{2}} \times H^{\frac{1}{2}}$, the existence of the solution is guaranteed at least for a time of the order ε^{-2} . Since (1.1) is a quasilinear wave equation, it is not a priori obvious that one can obtain better estimates. For instance, in the well-known example by Klainerman and Majda [30] *all* nontrivial space-periodic solutions of size ε blow up in a time of order ε^{-2} . In the papers [3], [4], [5] we have proved that for the Kirchhoff equation the situation is more favorable. More precisely, in [3], performing one step of quasilinear normal form, we have proved that the lifespan of all solutions of small amplitude is at least of order ε^{-4} , since the only resonant cubic terms that cannot be erased in the first step of normal form give no contribution to the energy estimates (Theorem 2.1 below). In [4] we have computed the second step of quasilinear normal form for the Kirchhoff equation and showed that there are resonant terms of degree five that cannot be erased and give a nontrivial contribution to the time evolution of Sobolev norms. Finally, in [5] we have shown that for a suitable set of nonresonant initial data the effect of such terms of degree five can be neglected on a longer timescale and the lifespan of the corresponding solution is at least of order ε^{-6} (Theorem 2.4 below). Since equation (1.1) is set on the compact manifold \mathbb{T}^d , dispersion mechanisms that hold on \mathbb{R}^d are not available, and the main tool to prove existence beyond the time of the standard local theory is the *normal form* method (see, for instance, the results of Kuksin-Pöschel [31], Bourgain [13], and Bambusi-Delort-Grébert-Szeftel [6]-[7]). Some of the difficulties and achievements in this active research field regard the extension of the theory to to *quasilinear* PDEs (see e.g. Delort [16]-[17] on quasilinear Klein-Gordon equations, Craig-Sulem [15], Ifrim-Tataru [28], Berti-Delort [11], Berti-Feola-Pusateri [12] on water waves, Feola-Iandoli [19] on quasilinear NLS) and to *resonant* equations without the help of external parameters (see e.g. Berti-Feola-Pusateri [12] on pure gravity water waves, Bernier-Faou-Grébert [8]-[9] on resonant NLS, KdV and Benjamin-Ono with rational normal forms).

The Kirchhoff equation (1.1), despite its simple structure, contains these difficulties: (i) it is a quasilinear PDE, because the nonlinear term $\Delta u \int |\nabla u|^2$ has the same order of derivatives as the linear part of the equation. (ii) It is a resonant equation: the linear frequencies of oscillation, namely the eigenvalues of the linear wave $\partial_{tt} - \Delta$, are square roots $|k| = \sqrt{k_1^2 + \ldots + k_d^2}$, $k \in \mathbb{Z}^d$, of natural numbers, and therefore equations like $|k| + |j| - |\ell| = 0$ and similar, which one encounters along a normal form procedure, have infinitely many nontrivial solutions. (iii) Finally, there are no external parameters that could help to avoid the resonances.

In Section 2 we shall give a formal statement of our main results in [3]-[5] and in Section 3 we shall briefly describe the strategy of the proof. Finally, in Section 4, we shall spend a few words about the ongoing work [2] with Giuliani and Guardia, where we prove the existence of small solutions of the Kirchhoff equation exhibiting a chaotic-like behavior over long timescales.

2 Main results

On the torus \mathbb{T}^d , it is not restrictive to assume that both the initial data a(x), b(x) and the unknown function u(t, x) have zero average in the space variable x (because the space average and the zero-mean component of any a, b, u satisfy two uncoupled Cauchy problems; the problem for the averages is elementary).

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

$$H_0^s(\mathbb{T}^d,\mathbb{C}) := \left\{ u(x) = \sum_{j \in \mathbb{Z}^d \setminus \{0\}} u_j e^{ij \cdot x} : u_j \in \mathbb{C}, \ \|u\|_s < \infty \right\}, \qquad \|u\|_s^2 := \sum_{j \neq 0} |u_j|^2 |j|^{2s}, \quad (2.1)$$

and its subspace of real-valued functions

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

The main result of [3] is the following.

Theorem 2.1 ([3]). For $d \in \mathbb{N}$, let

$$m_0 = 1$$
 if $d = 1$, $m_0 = \frac{3}{2}$ if $d \ge 2$. (2.2)

There exist universal constants $\varepsilon_0, C, C_1 > 0$ with the following properties. If $(a, b) \in H_0^{m_0 + \frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{m_0 - \frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$ with

 $\varepsilon := \|a\|_{m_0 + \frac{1}{2}} + \|b\|_{m_0 - \frac{1}{2}} \le \varepsilon_0,$

then the Cauchy problem (1.1)-(1.2) 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

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

and

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

If, in addition, $(a,b) \in H_0^{s+\frac{1}{2}}(\mathbb{T},\mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T},\mathbb{R})$ 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

$$\max_{t \in [0,T]} (\|u(t)\|_{s+\frac{1}{2}} + \|\partial_t u(t)\|_{s-\frac{1}{2}}) \le C(\|a\|_{s+\frac{1}{2}} + \|b\|_{s-\frac{1}{2}}).$$
(2.3)

Remark 2.2 (Evolution of higher norms). The constant C in (2.3) 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 $\partial_{tt}v - a(t)\Delta v = 0$. \Box

Remark 2.3 (Why m_0 in (2.2) is different in dimension d = 1 and $d \ge 2$). The proof of Theorem 2.1 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. On the 1-dimensional torus \mathbb{T} , the difference ||j| - |k|| is either zero or ≥ 1 , while on \mathbb{T}^d , $d \ge 2$, 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 $||j| - |k|| \ge \frac{1}{|j|+|k|}$. Note that, in dimension d = 1, the definition $m_0 = 1$ means that the threshold $H^{\frac{3}{2}} \times H^{\frac{1}{2}}$, below which local wellposedness is unknown, is perfectly matched, and Theorem 2.1 extends the time of existence of the solution from ε^{-2} to ε^{-4} , without requiring any higher regularity on the initial data; this is a pure improvement, without additional assumptions, with respect to the standard local theory.

The main result of [5] further extends the time of existence from ε^{-4} to ε^{-6} , but only for initial data satisfying a suitable nonresonance condition. We now introduce some notation in order to give a precise statement.

Define

$$m_1 := 1$$
 if $d = 1$, $m_1 := 2$ if $d \ge 2$ (2.4)

and

$$\Gamma := \{ |k| : k \in \mathbb{Z}^d, \ k \neq 0 \} \subseteq \{ \sqrt{n} : n \in \mathbb{N} \} \subset [1, \infty),$$
(2.5)

where $|k| = (k_1^2 + \ldots + k_d^2)^{\frac{1}{2}}$ is the usual Euclidean norm, and $\mathbb{N} := \{1, 2, \ldots\}$. Given a pair (a, b) of functions, with

$$a(x) = \sum_{k \in \mathbb{Z}^d \setminus \{0\}} a_k e^{ik \cdot x}, \quad b(x) = \sum_{k \in \mathbb{Z}^d \setminus \{0\}} b_k e^{ik \cdot x}, \tag{2.6}$$

for each $\lambda \in \Gamma$ we define

$$U_{\lambda} := U_{\lambda}(a, b) := \sum_{|k|=\lambda} (\lambda^3 |a_k|^2 + \lambda |b_k|^2).$$
(2.7)

We denote

$$\Gamma_1 := \Gamma_1(a, b) := \{ \lambda \in \Gamma : U_\lambda(a, b) > 0 \} = \Gamma \setminus \Gamma_0.$$
(2.8)

Theorem 2.4 ([5]). There exist universal constants $\delta \in (0, 1)$, C, A > 0 with the following properties. Let ε , c_0 be real numbers with

$$0 < \varepsilon \le \delta c_0, \qquad 0 < c_0 \le 1, \tag{2.9}$$

and let

$$(a,b) \in H_0^{m_1 + \frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{m_1 - \frac{1}{2}}(\mathbb{T}^d, \mathbb{R}), \qquad \|a\|_{m_1 + \frac{1}{2}} + \|b\|_{m_1 - \frac{1}{2}} \le \varepsilon.$$
(2.10)

Let $U_{\lambda} = U_{\lambda}(a,b)$, $\lambda \in \Gamma$, be the sums defined in (2.7), and let $\Gamma_1 = \Gamma_1(a,b)$ be the set in (2.8). Assume that (a,b) satisfy

$$|U_{\alpha} + U_{\beta} - U_{\lambda}| \ge c_0(U_{\alpha} + U_{\beta} + U_{\lambda}) \quad \forall \alpha, \beta, \lambda \in \Gamma_1 \ s.t. \ \alpha + \beta = \lambda.$$
(2.11)

Then the solution u of the Cauchy problem (1.1)-(1.2) is defined on the time interval [0,T], where $T = Ac_{\perp}^{3} \varepsilon^{-6}$

$$\begin{aligned} u(t)\|_{m_1+\frac{1}{2}} &= H_0 e^{-\tau}, \\ with \ u \in C^0([0,T], H_0^{m_1+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})) \cap C^1([0,T], H_0^{m_1-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})) \ and \\ \|u(t)\|_{m_1+\frac{1}{2}} + \|\partial_t u(t)\|_{m_1-\frac{1}{2}} \leq C\varepsilon \quad \forall t \in [0,T]. \end{aligned}$$

Remark 2.5 (Why m_1 in (2.4) is higher than m_0 in dimension $d \ge 2$). As explained in Remark 2.3, the higher regularity threshold in dimension $d \ge 2$ comes from the small divisors that one encounters along the normal form procedure. Theorem 2.1 is a consequence of the first step of normal form, where the small divisors are of the form |j| - |k|, $j, k \in \mathbb{Z}^d$. Theorem 2.4, instead, relies on the second step of normal form, where one finds small divisors of the form $|j| + |k| - |\ell|$, $j, k, \ell \in \mathbb{Z}^d$, which in dimension $d \ge 2$ accumulate to zero faster than |j| - |k|. The definition of m_1 is motivated by the loss of regularity caused by such small divisors.

Assumption (2.11) is specifically designed to avoid the triple resonances of the Kirchhoff equation, and it deserves some comments, showing that the set of functions satisfying (2.11) is nonempty, and in fact it contains several nontrivial examples. A more detailed discussion can be found in [5].

Remark 2.6. (*Decreasing sequences*). Any decreasing sequence $(\sigma_{\lambda})_{\lambda \in \Gamma}$ of nonnegative real numbers satisfies

$$|\sigma_{\alpha} + \sigma_{\beta} - \sigma_{\lambda}| \ge \frac{1}{3}(\sigma_{\alpha} + \sigma_{\beta} + \sigma_{\lambda})$$

for all $\alpha, \beta, \lambda \in \Gamma$ with $\alpha + \beta = \lambda$.

Remark 2.7. (*Fixed power decay*). The observation of Remark 2.6 applies, for example, to the sequence $\sigma_{\lambda} = \lambda^{-2\sigma}$, which is decreasing for $\sigma \geq 0$. Hence any pair (a, b) of functions such that $U_{\lambda}(a, b) = \lambda^{-2\sigma}$ with $\sigma \geq 0$ satisfies (2.11) with $c_0 = 1/3$. Concerning the Sobolev regularity of such functions, one has that $(a, b) \in H_0^{m_1 + \frac{1}{2}} \times H_0^{m_1 - \frac{1}{2}}$ for $\sigma > m_1$.

Remark 2.8. (Absence of triplets: odd integers). If the set Γ_1 does not contain any triplet (α, β, λ) with $\alpha + \beta = \lambda$, then (2.11) is trivially satisfied. For example, this holds if $\Gamma_1 \subseteq \{n \in \mathbb{N} : n \text{ odd}\}$. Other examples can be constructed as lacunary subsets of \mathbb{N} .

Remark 2.9. (Arithmetic decomposition of Γ). The set Γ can be decomposed as the disjoint union $\cup_p \Gamma(\sqrt{p})$ of the sets

$$\Gamma(\sqrt{p}) := \{n\sqrt{p} : n \in \mathbb{N}\} \cap \Gamma,\$$

where p runs over the square-free positive integers. Now, if $\alpha, \beta, \lambda \in \Gamma$ are such that $\alpha + \beta = \lambda$, then there exists a square-free p such that α, β, λ all belong to the same $\Gamma(\sqrt{p})$. This means that condition (2.11) can be verified independently on each of the sets $\Gamma(\sqrt{p})$.

Remark 2.10. (*Perturbations of* (2.11)). Condition (2.11) displays some stability under perturbation. In fact, if (a, b) satisfies (2.11) and (f, g) is another pair of functions, with

$$U_{\lambda}(f,g) \le \mu^2 U_{\lambda}(a,b) \quad \text{for some } \mu \in \left(0, \frac{c_0}{4}\right),$$

$$(2.12)$$

then the pair (a + f, b + g) satisfies (2.11) with the constant c_0 replaced by $\frac{c_0 - 2\mu - \mu^2}{1 + 2\mu + \mu^2}$.

Applying this observation to the example in Remark 2.7, one proves that if (a, b) is a pair of functions such that $U_{\lambda}(a, b) = \lambda^{-2\sigma}$ with $\sigma > m_1$ and if

$$(f,g) \in H_0^{\sigma+\frac{3}{2}} \times H_0^{\sigma+\frac{1}{2}}, \quad \|f\|_{\sigma+\frac{3}{2}}^2 + \|g\|_{\sigma+\frac{1}{2}}^2 \le \frac{1}{576},$$

then (a + f, b + g) satisfies (2.11) with $c_0 = \frac{1}{6}$. We remark, however, that the set

$$\mathcal{B}(a,b) := \left\{ (a,b) + (f,g) : \|f\|_{\sigma+\frac{3}{2}}^2 + \|g\|_{\sigma+\frac{1}{2}}^2 \le \frac{1}{576} \right\}$$

is not a ball in the Sobolev space $H_0^{\sigma+\frac{3}{2}} \times H_0^{\sigma+\frac{1}{2}}$, since (a, b) only belongs to $H_0^{s+\frac{1}{2}} \times H_0^{s-\frac{1}{2}}$ for $s < \sigma$.

3 Description of the proof

3.1 Theorem 2.1

As mentioned above, 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

$$\begin{cases} \partial_t u = \nabla_v H(u, v) = v, \\ \partial_t v = -\nabla_u H(u, v) = \left(1 + \int_{\mathbb{T}^d} |\nabla u|^2 dx\right) \Delta u, \end{cases}$$
(3.1)

where the Hamiltonian is

$$H(u,v) = \frac{1}{2} \int_{\mathbb{T}^d} v^2 dx + \frac{1}{2} \int_{\mathbb{T}^d} |\nabla u|^2 dx + \left(\frac{1}{2} \int_{\mathbb{T}^d} |\nabla u|^2 dx\right)^2,$$
(3.2)

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

$$\langle f,g \rangle := \int_{\mathbb{T}^d} f(x)g(x) \, dx \quad \forall f,g \in L^2(\mathbb{T}^d,\mathbb{R}),$$
(3.3)

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 (3.2), 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 $\lesssim \|(u,v)\|_{H^s \times H^{s-1}}^3$, 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, so that the linear wave operator becomes diagonal, and system (3.1) becomes

$$\begin{cases} \partial_t u = -i\Lambda u - \frac{i}{4} \langle \Lambda(u + \overline{u}), u + \overline{u} \rangle \Lambda(u + \overline{u}), \\ \partial_t \overline{u} = i\Lambda \overline{u} + \frac{i}{4} \langle \Lambda(u + \overline{u}), u + \overline{u} \rangle \Lambda(u + \overline{u}), \end{cases}$$
(3.4)

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 (3.3), even for complex-valued functions f, g. We note that the cubic nonlinearity in (3.4) already has a "paralinear" structure, in the sense that, for all functions u, v, h, all $s \geq 0$, one has

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

Hence (3.4) can be interpreted as a linear system whose operator coefficients depend on (u, \overline{u}) , namely

$$\partial_t \begin{pmatrix} u \\ \overline{u} \end{pmatrix} = \begin{pmatrix} -A(u,\overline{u}) & -B(u,\overline{u}) \\ B(u,\overline{u}) & A(u,\overline{u}) \end{pmatrix} \begin{pmatrix} u \\ \overline{u} \end{pmatrix}, \tag{3.5}$$

where

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

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 $\leq 2||u||_{\frac{1}{2}}^2||u||_{s+\frac{1}{2}}^2$ 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 $\|\Phi(u, \overline{u}) - (u, \overline{u})\|_{H^s \times H^s} \leq \|(u, \overline{u})\|_{H^s \times H^s}^3$. 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 [16], [17] 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 [11] and [12] about water waves equations on \mathbb{T} .

The off-diagonal unbounded terms of (3.4) are eliminated by constructing a nonlinear bounded transformation $\Phi^{(3)}$ (see Section 3 of [3]) that conjugates system (3.4) to a new system of the form

$$\begin{cases} \partial_t u = -i\sqrt{1+2P(u,\overline{u})}\,\Lambda u + \frac{i}{4(1+2P(u,\overline{u}))} \Big(\langle\Lambda\overline{u},\Lambda\overline{u}\rangle - \langle\Lambda u,\Lambda u\rangle\Big)\overline{u},\\ \partial_t\overline{u} = i\sqrt{1+2P(u,\overline{u})}\,\Lambda\overline{u} + \frac{i}{4(1+2P(u,\overline{u}))}\Big(\langle\Lambda\overline{u},\Lambda\overline{u}\rangle - \langle\Lambda u,\Lambda u\rangle\Big)u, \end{cases}$$
(3.6)

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 $y \mapsto y\sqrt{1+2y}$, $y \ge 0$. System (3.6) still has the structure (3.5), with the improvement that the off-diagonal part $B(u, \overline{u})$ is now a bounded operator, satisfying

$$||B(u,\overline{u})h||_{s} \le ||u||_{1}^{2} ||h||_{s}$$

for all $s \ge 0$, all u, h. Thanks to the special structure of the Kirchhoff equation, and in particular to the lower bound $\frac{1}{4}\langle \Lambda(u+\overline{u}), u+\overline{u}\rangle = \int_{\mathbb{T}^d} (\Re(\Lambda^{\frac{1}{2}}u))^2 dx \ge 0$, the transformation $\Phi^{(3)}$ is global, namely it is defined for all $u \in H_0^1(\mathbb{T}^d, \mathbb{C})$, and not only for small u. In (3.4) 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 (3.6) 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 $u \in H_0^1(\mathbb{T}^d, \mathbb{C})$. 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 (3.6) could be normalized to 1; however, this is not needed to prove our results, because these coefficients are independent of x, and therefore the (unbounded) diagonal terms cancel out in the energy estimates.

At this point one can start with a standard Poincarè-Dulac normal form. Actually, we perform a "partial" normal form because we do not modify the harmless cubic diagonal terms. The first step consists in a close to identity (Id. + cubic terms) transformation, removing all the nonresonant monomials from the cubic nonlinearity. After such a transformation (see Section 4 of [3]), system (3.6) is conjugated to

$$\begin{pmatrix} \partial_t u\\ \partial_t \overline{u} \end{pmatrix} = \begin{pmatrix} -i\sqrt{1+2P(u,\overline{u})}\,\Lambda u\\ i\sqrt{1+2P(u,\overline{u})}\,\Lambda \overline{u} \end{pmatrix} + X_3^+(u,\overline{u}) + X_{\geq 5}(u,\overline{u}).$$
(3.7)

Here, $X_3^+(u, \overline{u})$ is the cubic normal form, while $X_{\geq 5}(u, \overline{u})$ is a remainder of homogeneity ≥ 5 . The first component of the vector field X_3^+ has the form

$$(X_3^+)_1(u,\overline{u}) = -\frac{i}{4} \sum_{j,k\neq 0, \ |k|=|j|} u_j u_{-j} |j|^2 (\overline{u})_k e^{ik \cdot x}.$$
(3.8)

(the second component of X_3^+ can be obtained by conjugation, thanks to the real structure of the system). The relevant feature here is that only monomials with |j| = |k| have survived in the resonant part of the vector field, which means that the transfer of energy caused by these monomials could happen only between modes on the same sphere centered at zero in Fourier space. This is why the contribution of the cubic resonant terms in the energy estimates vanishes, and therefore one gets an improved estimate of the form

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

for the transformed system (3.9), whence one deduces that the lifespan of the solutions of the original Cauchy problem (1.1) is at least of order $(||a||_{s+\frac{1}{2}} + ||b||_{s-\frac{1}{2}})^{-4}$.

3.2 Theorem **2.4**

The strategy of the proof of Theorem 2.4 is based on performing a further step of normal form; the computation is rather long and it is the core of the paper [4]. After the second step of normal form, the system is in the form

$$\begin{pmatrix} \partial_t u\\ \partial_t \overline{u} \end{pmatrix} = \begin{pmatrix} -i\sqrt{1+2P(u,\overline{u})}\Lambda u\\ i\sqrt{1+2P(u,\overline{u})}\Lambda \overline{u} \end{pmatrix} + X_3^+(u,\overline{u}) + X_5^+(u,\overline{u}) + X_{\geq 7}(u,\overline{u}),$$
(3.9)

11-1

where $X_5^+(u, \overline{u})$ only contains resonant terms of degree five, while $X_{\geq 7}(u, \overline{u})$ is a remainder of homogeneity ≥ 7 . The formula for the first component of the vector field X_5^+ is

$$(X_{5}^{+})_{1}(u,\bar{u}) = \frac{i}{32} \sum_{\substack{j,\ell,k\\|j|=|\ell|}} u_{j}u_{-j}(\bar{u})_{\ell}(\bar{u})_{-\ell}u_{k}e^{ik\cdot x}|j|^{2}|\ell|^{2} \left(\frac{1}{|j|+|k|} - \frac{1-\delta_{|\ell|}^{|\kappa|}}{|\ell|-|k|}\right) + \frac{3i}{32} \sum_{\substack{j,\ell,k\\|k|=|j|+|\ell|}} u_{j}u_{-j}u_{\ell}(\bar{u})_{k}e^{ik\cdot x}|j||\ell||k| + \frac{i}{16} \sum_{\substack{j,\ell,k\\|j|=|k|}} u_{j}u_{-j}u_{\ell}(\bar{u})_{-\ell}(\bar{u})_{k}e^{ik\cdot x}|j|^{2}|\ell| \left(6 + \frac{|\ell|}{|\ell|+|j|} + \frac{|\ell|(1-\delta_{|\ell|}^{|j|})}{|\ell|-|j|}\right) + \frac{3i}{16} \sum_{\substack{j,\ell,k\\|k|=|j|-|\ell|}} u_{j}u_{-j}(\bar{u})_{\ell}(\bar{u})_{-\ell}(\bar{u})_{k}e^{ik\cdot x}|j||\ell||k|,$$
(3.10)

where, in expressions such as $\frac{1-\delta_{|\ell|}^{|k|}}{|\ell|-|k|}$, we adopt the convention $\frac{0}{0} = 0$, i.e. $\frac{1-\delta_{|\ell|}^{|k|}}{|\ell|-|k|} = 0$ if $|k| = |\ell|$. The second and the fourth line of (3.10) contain monomials that produce an interaction between frequencies that are not on the same sphere in Fourier space. Because of this, at homogeneity degree 5, resonant terms give a nontrivial contribution to the energy estimates and it is not possible to use the same strategy as for the cubic terms to obtain a result of existence of solutions on a timescale ε^{-6} for all initial data.

However, the special structure of the vector field allows one to write down an effective system, with less degrees of freedom, governing the time evolution of the Sobolev norms of solutions. Such an effective system can be used to prove that "nonresonant" initial data (in the sense specified by Theorem 2.4) give rise to solutions that exist over a timescale of order at least ε^{-6} . The analysis of the effective system leading to such a result is the core of the paper [5]: in the next lines, we shall summarize the main ideas.

By defining the "macroscopic quantities"

$$S_{\lambda} := \sum_{k:|k|=\lambda} |u_k|^2 = \sum_{k:|k|=\lambda} u_k(\bar{u})_{-k}, \quad B_{\lambda} := \sum_{k:|k|=\lambda} u_k u_{-k}, \quad (3.11)$$

and ignoring higher order remainder terms, one gets the system

$$\partial_t S_{\lambda} = \frac{3i}{32} \sum_{\substack{\alpha,\beta\in\Gamma\\\alpha+\beta=\lambda}} (B_{\alpha} B_{\beta} \overline{B_{\lambda}} - \overline{B_{\alpha} B_{\beta}} B_{\lambda}) \alpha\beta\lambda + \frac{3i}{16} \sum_{\substack{\alpha,\beta\in\Gamma\\\alpha-\beta=\lambda}} (B_{\alpha} \overline{B_{\beta} B_{\lambda}} - \overline{B_{\alpha}} B_{\beta} B_{\lambda}) \alpha\beta\lambda, \quad (3.12)$$

$$\partial_t B_\lambda = -2i \left(\lambda + \frac{1}{4} \lambda^2 S_\lambda\right) B_\lambda. \tag{3.13}$$

The growth of Sobolev norms of solutions of the Kirchhoff equation only depends on the time evolution of the "superactions" S_{λ} : as equation (3.12) shows, such an evolution is governed by the imaginary part of products of the form $B_{\alpha}B_{\beta}\overline{B_{\lambda}}$, where $\alpha + \beta = \lambda$. It is therefore very natural to define

$$Z_{\alpha\beta\lambda} := B_{\alpha}B_{\beta}\overline{B_{\lambda}}, \quad \alpha + \beta = \lambda, \tag{3.14}$$

and use (3.13) to write down the equation for the evolution of $Z_{\alpha\beta\lambda}$. Ignoring again terms of higher homogeneity, one gets, for $\alpha + \beta = \lambda$,

$$\partial_t Z_{\alpha\beta\lambda} = -\frac{i}{2} (\alpha^2 S_\alpha + \beta^2 S_\beta - \lambda^2 S_\lambda) Z_{\alpha\beta\lambda}.$$
(3.15)

The heuristics behind the proof of Theorem 2.4, then, is quite simple. Equation (3.15) describes a rotational dynamics, meaning that $\rho_{\alpha\beta\lambda} := |Z_{\alpha\beta\lambda}|$ is a constant of motion and the velocity of rotation of $Z_{\alpha\beta\lambda}$ on the circle of radius $\rho_{\alpha\beta\lambda}$ in the complex plane is given by the factor $\omega_{\alpha\beta\lambda} :=$ $\alpha^2 S_{\alpha} + \beta^2 S_{\beta} - \lambda^2 S_{\lambda}$. If some of the terms $\omega_{\alpha\beta\lambda}$ are zero, or if they are very close to zero, then the corresponding variables $Z_{\alpha\beta\lambda}$ are going to stay almost constant in time. Looking again at equation (3.12), this means that some of the terms $B_{\alpha}B_{\beta}\overline{B_{\lambda}} - \overline{B_{\alpha}B_{\beta}}B_{\lambda}$ would be almost constant, which might create a "secular" drift of the superactions, visible already on a timescale of the order of ε^{-4} . However, if all the $\omega_{\alpha\beta\lambda}$'s are bounded away from zero, then the time evolution of the superactions will benefit from a further averaging effect and will be much slower. Since the quantities $\omega_{\alpha\beta\lambda}$ are linear combinations of the superactions, the fact that the evolution of the superactions S_{λ} is slow will trigger a virtuous circle, keeping the $\omega_{\alpha\beta\lambda}$'s bounded away from zero. Making this quantitative, and letting the higher order remainder terms into the picture, one proves stability of the superactions, and therefore existence of the solution, on a longer timescale (of order ε^{-6}), provided that the quantites $\omega_{\alpha\beta\lambda}$ are bounded away from zero at the time t = 0. The assumption that we make is that, at t = 0, one has $|\omega_{\alpha\beta\lambda}| = |\alpha^2 S_{\alpha} + \beta^2 S_{\beta} - \lambda^2 S_{\lambda}| \ge c_0(\alpha^2 S_{\alpha} + \beta^2 S_{\beta} + \lambda^2 S_{\lambda}).$ In the original variables, this is precisely assumption (2.11) in Theorem 2.4.

4 Chaotic-like solutions

It is a natural question whether the effective system (3.12)-(3.13) can be used not only to prove the stable behavior and long time existence of some solutions of (1.1), but also the existence of some solutions exhibiting an unstable behavior as a genuinely nonlinear phenomenon.

One may be interested in looking, e.g., for solutions with a large growth of Sobolev norms (like the ones constructed in [14], [25], [24], [23] for the nonlinear Schrödinger equation), or for nonlinear beatings (see [22], [27], [26] for the NLS), or for chaotic-like solutions (see [21] for wave, Hartree and beam equations). In the ongoing work [2], we use the effective system (3.12)-(3.13) to prove the existence of solutions of (1.1) that exhibit a chaotic-like behavior for a long time, when the equation is set on \mathbb{T}^d with $d \geq 2$. As far as we know, this is only the second instance of such a type of result for PDEs, and it differs from the already existing result [21] in several aspects. In particular, one of the most relevant differences is that the equations studied in [21] are semilinear, while the Kirchhoff equation (1.1) is quasilinear. Another dynamically relevant feature is that the chaotic-like motions in [21] are produced by the first step of normal form, thus the energy exchanges between modes are of the same size as the amplitude of the solution. For equation (1.1), instead, the first step of normal form does not produce any chaotic behavior, so we prove the existence of such special solutions as a perturbation of the effective system (3.12)-(3.13), which is deduced by the second step of normal form. The consequence of this feature is that the solutions constructed in [2] exhibit a *chaotic-like modulation* of a stable motion, meaning that they are solutions of size ε^2 between the amplitude of different Fourier modes. Such an exchange of energy is chaotic-like, in the sense that the time spent in each transfer of energy can be chosen randomly.

The general strategy of the construction in [2] is based on using (3.12)-(3.13) and making a suitable choice of the activated Fourier modes, so that the effective dynamics can be reduced to a system of two weakly coupled pendulums; an interesting and possibly unexpected fact is that the weakness in the coupling of the pendulums is a consequence of the choice of the Fourier sites $(m \ll p \text{ in } (4.1) \text{ below})$. From this point, we construct a symbolic dynamics (Smale horseshoe) and we have to show that the constructed solutions persist on a relevant timescale, after adding to (3.12)-(3.13) the higher order remainder terms of the normal form. This last perturbative step is a priori far from obvious, especially since the chaotic-like motion is a smaller (size ε^2) perturbation of a much larger (size ε) stable solution.

The following is an attempt to write a precise, and nonetheless not too detailed, statement of the main result of [2]. To keep the statement shorter, instead of describing the motion of the single Fourier modes, we describe here the cumulative effect of the energy transfers via the variation of the Sobolev norms of the solution.

Theorem 4.1. Denote $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. There exist two universal constants $\sigma_* \in (0, 1)$, $C_* > 0$ with the following properties. Let m, p be two integers, with $2 \leq m < p$ and $m/p \leq \sigma_*$. Then there exist constants $M_* \geq 1$, $\varepsilon_* \in (0, 1)$, $K, K', K'', K_1, K_2 > 0$, all depending only on m, p, with the following properties.

Let $(m_j)_{j \in \mathbb{N}_0}$ be a sequence of integers with $m_j \ge M_*$ for all $j \ge 0$, and let $0 < \varepsilon \le \varepsilon_*$. Then there exist two sequences of real numbers (t_j^*) , (\bar{t}_j^*) , $j \in \mathbb{N}_0$, with

$$t_0^* = 0, \quad t_{j+1}^* = t_j^* + K\varepsilon^{-3}(m_j + \theta_j), \quad 0 \le \theta_j < 1, \quad t_j^* < \bar{t}_j^* < t_{j+1}^*,$$

and there exists a solution u(t,x) of the Kirchhoff equation (1.1), global in time, Fourier supported on the set $\{k \in \mathbb{Z}^d : |k| \in \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}\}$, where

$$\alpha_1 = m, \quad \alpha_2 = m + p, \quad \alpha_3 = 2m + p, \quad \alpha_4 = 3m + 2p,$$
(4.1)

whose Sobolev norms satisfy

$$\max_{t \in [t_j^*, t_j^*]} \left(\|u(t)\|_{\frac{3}{2}} + \|\partial_t u(t)\|_{\frac{1}{2}} \right) \ge V_1, \qquad \min_{t \in [t_j^*, t_{j+1}^*]} \left(\|u(t)\|_{\frac{3}{2}} + \|\partial_t u(t)\|_{\frac{1}{2}} \right) \le V_0,$$

for all j = 0, ..., N, where N satisfies

$$\sum_{j=0}^{N} m_j \le K' \log(\varepsilon^{-1}).$$

The values V_0, V_1 are

$$V_0 = \varepsilon K_1 + \varepsilon^2 K_2, \quad V_1 = \varepsilon K_1 + 3\varepsilon^2 K_2,$$

and the Sobolev norms of the solution satisfy $\left(\|u(t)\|_{\frac{3}{2}} + \|\partial_t u(t)\|_{\frac{1}{2}}\right) \leq K'' \varepsilon$ for all $t \in [0, t^*_{N+1}]$.

Acknowledgements. We would like to thank Claudio Bonanno and the DinAmici group for the participation to the DAI days. This work has been supported by PRIN 2020XB3EFL "Hamiltonian and dispersive PDEs".

Conflict of interest. We declare that there is no conflict of interest.

References

- A. Arosio, S. Panizzi, On the well-posedness of the Kirchhoff string, Trans. Amer. Math. Soc. 348 (1996), 305-330.
- [2] P. Baldi, F. Giuliani, M. Guardia, E. Haus, Effective chaos for the Kirchhoff equation on tori, work in progress.
- [3] P. Baldi, E. Haus, On the existence time for the Kirchhoff equation with periodic boundary conditions, Nonlinearity 33 (2020), no. 1, 196-223.
- [4] P. Baldi, E. Haus, On the normal form of the Kirchhoff equation, J. Dyn. Diff. Equat. 33 (2021), 1203-1230, special issue in memory of Walter Craig.
- [5] P. Baldi, E. Haus, Longer lifespan for many solutions of the Kirchhoff equation, SIAM J. Math. Anal. 54 (2022), no. 1, 306-342.
- [6] D. Bambusi, B. Grébert, Birkhoff normal form for partial differential equations with tame modulus, Duke Math. J. 135 (2006), 507-567.
- [7] 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), no. 11, 1665-1690.
- [8] J. Bernier, E. Faou, B. Grébert, Rational normal forms and stability of small solutions to nonlinear Schrödinger equations, Annals of PDE 6 (2020), no. 2, art. 14.
- [9] J. Bernier, B. Grébert, Long time dynamics for generalized Korteweg-de Vries and Benjamin-Ono Equations, Arch. Ration. Mech. Anal. 241 (2021), no. 3, 1139-1241.
- [10] S.N. Bernstein, Sur une classe d'équations fonctionnelles aux dérivées partielles, Izv. Akad. Nauk SSSR Ser. Mat. 4 (1940), 17-26.
- [11] 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.
- [12] M. Berti, R. Feola, F. Pusateri, Birkhoff normal form and long time existence for periodic gravity water waves, Comm. Pure Appl. Math. (2022), online first, doi:10.1002/cpa.22041.
- J. Bourgain, On diffusion in high-dimensional Hamiltonian systems and PDE, J. Analyse Math. 80 (2000), 1-35.
- [14] 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.
- [15] W. Craig, C. Sulem, Mapping properties of normal forms transformations for water waves, Boll. Unione Mat. Ital. 9 (2016), 289-318.
- [16] 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.
- [17] 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).
- [18] R. W. Dickey, Infinite systems of nonlinear oscillation equations related to the string, Proc. Amer. Math. Soc. 23 (1969), 459-468.
- [19] R. Feola, F. Iandoli, Long time existence for fully nonlinear NLS with small Cauchy data on the circle, Ann. Sc. Norm. Super. Pisa Cl. Sci., Vol. XXII (2021), no. 1, 109-182.
- [20] M. Ghisi, M. Gobbino, Global solutions to the Kirchhoff equation with spectral gap data in the energy space, arXiv:2208.05400.
- [21] F. Giuliani, M. Guardia, P. Martin, S. Pasquali, Chaotic-like transfers of energy in Hamiltonian PDEs, Comm. Math. Phys. 384 (2021), 1227-1290.
- [22] B. Grébert, L. Thomann, Resonant dynamics for the quintic nonlinear Schrödinger equation, Ann. Inst. H. Poincaré Anal. Non Linéaire 29 (2012), 455-477.
- [23] M. Guardia, Z. Hani, E. Haus, A. Maspero, M. Procesi, Strong nonlinear instability and growth of Sobolev norms near quasiperiodic finite gap tori for the 2D cubic NLS equation, J. Eur. Math. Soc. (2022), online first, doi:10.4171/JEMS/1200.
- [24] M. Guardia, E. Haus, M. Procesi, Growth of Sobolev norms for the analytic NLS on T², Adv. Math. 301 (2016), 615-692.
- [25] M. Guardia, V. Kaloshin, Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation, J. Eur. Math. Soc. (JEMS) 17 (2015), 71-149.
- [26] E. Haus, M. Procesi, KAM for beating solutions of the quintic NLS, Comm. Math. Phys. 354 (2017), 1101-1132.
- [27] E. Haus, L. Thomann, Dynamics on resonant clusters for the quintic nonlinear Schrödinger equation, Dyn. Partial Differ. Equ. 10 (2013), 157-169.
- [28] 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.
- [29] G. Kirchhoff, Vorlesungen über mathematische Physik: Mechanik, ch.29, Teubner, Leipzig, 1876.
- [30] S. Klainerman, A. Majda, Formation of singularities for wave equations including the nonlinear vibrating string, Comm. Pure Appl. Math. 33 (1980), 241-263.

[31] S. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation, Ann. of Math. (2) 143 (1996), no. 1, 149-179.

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

Emanuele Haus Dipartimento di Matematica e Fisica Università Roma Tre Largo San Leonardo Murialdo 1, 00146 Roma ehaus@mat.uniroma3.it