

Effective chaos for the Kirchhoff equation on tori

Pietro Baldi, Filippo Giuliani, Marcel Guardia, and Emanuele Haus

Abstract. We consider the Kirchhoff equation on tori of any dimension and we construct solutions whose Sobolev norms oscillate in a chaotic way on certain long timescales. The chaoticity is encoded in the time between oscillations of the norm, which can be chosen in any prescribed way. This phenomenon, which we name *effective chaos* (it occurs over a long, but finite, timescale), is a consequence of the existence of symbolic dynamics for an effective system. Since the first-order resonant dynamics has been proved to be essentially stable, we need to perform a second-order analysis to find an effective model displaying chaotic dynamics. More precisely, after some nontrivial reductions, this model behaves as two weakly coupled pendulums.

Contents

1. Introduction and main result	1
2. Effective dynamics for the Kirchhoff equation	7
3. The truncated effective system with two triplets	13
4. Chaos for two weakly coupled pendulums	21
5. Back to the truncated effective system	29
6. Approximation argument	39
References	48

1. Introduction and main result

We consider the Kirchhoff equation

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

on the torus \mathbb{T}^d , $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$, in any dimension $d \geq 1$ (periodic boundary conditions), where the unknown $u = u(t, x)$, $x \in \mathbb{T}^d$, is a real-valued function.

Equation (1.1) was first introduced by Kirchhoff [25] in 1876, to model nonlinear transverse oscillations of elastic structures such as strings and plates ($d = 1, 2$). It is a

Mathematics Subject Classification 2020: 74H65 (primary); 35B34, 35A24, 37D45 (secondary).

Keywords: chaotic dynamics, effective chaos, Kirchhoff equation, quasilinear PDEs, resonances for Hamiltonian PDEs.

quasilinear wave equation, with cubic, nonlocal nonlinearity and Hamiltonian structure. Given its physical relevance, equation (1.1) has been well studied throughout the years; nonetheless, its study is still challenging, because several basic questions remain open.

While it has long been known (Dickey [11], Arosio–Panizzi [1]) that the Cauchy problem for (1.1) is locally well posed with initial data $(u(0), \partial_t u(0))$ 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 solutions with initial data of any given Sobolev regularity are global in time or not. In particular, it is not even known whether C^∞ (or even Gevrey) initial data of small amplitude produce solutions that are global in time. For initial data of analytic class instead, global well-posedness has been known since the work of Bernstein [6] in 1940.

Moreover, below the regularity threshold $H^{\frac{3}{2}} \times H^{\frac{1}{2}}$, neither local well-posedness nor ill-posedness have been established. A partial, interesting result in this direction was recently obtained by Ghisi and Gobino [12].

More general questions regard the lifespan of the solutions and their behavior as time evolves, at least close to the equilibrium $u = 0$. First of all, 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 equation, it is not a priori obvious that one can obtain better estimates. For instance, in the well-known example by Klainerman and Majda [26] *all* nontrivial space-periodic solutions of size ε blow up in a time of order ε^{-2} . On the other hand, in the papers [3–5], using techniques from normal form theory, it is proved that for the Kirchhoff equation the situation is more favorable. More precisely, in [3], performing one step of quasilinear normal form, it is proved that the lifespan of all solutions of small amplitude is at least of order ε^{-4} . This is a consequence of the fact that the only resonant cubic terms that cannot be erased in the first step of normal form give no contribution to the energy estimates. In [4] the second step of quasilinear normal form is computed, and it is proved that there are resonant terms of degree five that cannot be erased and that give a nontrivial contribution to the time evolution of Sobolev norms.

This is a starting point for describing interesting long-time dynamics for the Kirchhoff equation. The qualitative behavior of solutions of the Kirchhoff equation over long timescales is poorly understood, even for small, compactly Fourier supported initial data, which obey finite-dimensional systems.

Broadly speaking, for the dynamics of small data we can look for two different types of regimes:

- **Stable regime:** This is the case in which the long-time behavior of Fourier modes resembles the dynamics of the linearized equation, namely the energy of the modes remains almost constant over long time. We mention [5] where stable motions of equation (1.1) are obtained for a suitable set of nonresonant initial data, for which the effect of the resonant terms of degree five remain small on a longer timescale of order ε^{-6} . We also mention [2] and [10], where the existence of invariant tori is proved for a forced version of (1.1).

- Unstable regime: In this case the nonlinear terms lead to a new type of dynamics, very different from the linear one. Of particular interest is understanding how the nonlinear effects create exchanges of energy among different modes.

Concerning the unstable regime, some remarkable results in the literature regard the “energy cascade” for nonlinear Schrödinger equations, where the energy travels from low to high modes (or vice versa), in strong connection with weak turbulence theory. Such a phenomenon, which can be measured in terms of an arbitrarily large growth of Sobolev norms, was considered by Bourgain one of the most important problems in Hamiltonian PDEs; see [7], and also [8, 9, 18–21].

Other interesting dynamical behaviors are also based on the mechanism of energy exchange among Fourier modes. Such exchanges can be recurrent (i.e., periodic or quasi-periodic in time) or chaotic. Recurrent energy exchanges are obtained, for instance, in [14–16, 22, 23]. To the best of our knowledge, the only paper in the literature in which chaotic exchanges of energy are constructed for PDEs is [13] (see also [24, 27] for previous results about chaotic motions for PDEs with external parameters and dissipation). In [13] the authors consider cubic wave and beam equations and prove the existence of solutions essentially Fourier supported on a finite number of resonant modes that exchange energy among themselves in a chaotic way. The chaoticity reflects the fact that it is possible to provide energy exchanges among modes at a sequence of prescribed times (randomness of exchanging time) or among modes belonging to a prescribed resonant tuple (randomness of active and inactive modes).

Both in [13] and in the present paper the existence of chaotic solutions is due to the presence of chaotic dynamics for the normal form of the equation, up to a certain degree. More precisely, the normalized system leaves invariant a finite-dimensional subspace; then the chaotic behavior arises from the existence of a Smale horseshoe, which gives rise to symbolic dynamics. The orbits of the normalized system are globally defined in time, and the chaotic behavior is displayed for an infinitely long time. However, this does not imply the existence of chaotic solutions of the full PDE for an infinitely long time. The chaotic behavior for the full PDE is obtained by proving the vicinity of certain solutions of it to the chaotic orbits of the normalized system, and this approximation only holds over a long, but finite, time interval. We call this behavior *effective chaoticity*, in the sense that the dynamics behaves as chaotic in rather long timescales (in analogy with the stability over long timescales, often called *effective stability* in Hamiltonian dynamics).

1.1. Main result

We denote by \mathbb{N}_0 the set $\{0, 1, 2, \dots\}$ of nonnegative integers. The next theorem, which shows the existence of solutions of the Kirchhoff equation displaying chaotic-like, small amplitude, oscillations in the Sobolev norms, is the main result of the paper.

Theorem 1.1. *There exist universal positive constants M , τ , ε_* , C , r_0 , b , K , K_0 with the following property. Let $d \geq 1$. For every sequence $(m_j)_{j \in \mathbb{N}_0} = (m_0, m_1, m_2, \dots)$ of*

integers such that $m_j \geq M$ for all $j \in \mathbb{N}_0$, there exists a sequence

$$0 = s_0 < \bar{s}_0 < s_1 < \bar{s}_1 < s_2 \leq \bar{s}_2 < \dots, \quad s_{j+1} = s_j + \tau(m_j + \theta_j), \quad 0 \leq \theta_j < 1,$$

such that for every $\varepsilon \in (0, \varepsilon_*]$ there exists a solution $u(t, x)$ of the Kirchhoff equation (1.1) on \mathbb{T}^d , global in time, with finite Fourier support, whose norm

$$\mathcal{N}(t) := \left(\|u(t)\|_{H^{\frac{3}{2}}(\mathbb{T}^d)}^2 + \|\partial_t u(t)\|_{H^{\frac{1}{2}}(\mathbb{T}^d)}^2 \right)^{\frac{1}{2}}$$

satisfies

$$\mathcal{N}(t) \leq C\varepsilon \quad \forall t \in \mathbb{R},$$

and it oscillates around the central value $A_\varepsilon := \varepsilon + \varepsilon^2 r_0$ with oscillations described in terms of the amplitude $B_\varepsilon := \varepsilon^2 r_0$ and the error $\delta_\varepsilon := \frac{1}{10} B_\varepsilon$ as

$$-\delta_\varepsilon \leq \mathcal{N}(t) - A_\varepsilon \leq B_\varepsilon + \delta_\varepsilon \quad \forall t \in I_j = [t_j, \bar{t}_j], \quad \max_{t \in I_j} \mathcal{N}(t) - A_\varepsilon \geq B_\varepsilon - \delta_\varepsilon,$$

$$-B_\varepsilon - \delta_\varepsilon \leq \mathcal{N}(t) - A_\varepsilon \leq \delta_\varepsilon \quad \forall t \in E_j = [\bar{t}_j, t_{j+1}], \quad \min_{t \in E_j} \mathcal{N}(t) - A_\varepsilon \leq -B_\varepsilon + \delta_\varepsilon,$$

where

$$t_j = \frac{s_j}{b\varepsilon^3}, \quad \bar{t}_j = \frac{\bar{s}_j}{b\varepsilon^3},$$

for all intervals I_j, E_j contained in the time interval $[0, T_\varepsilon]$, where

$$T_\varepsilon = K\varepsilon^{-3} \log(\varepsilon^{-1}).$$

One has $I_j, E_j \subset [0, T_\varepsilon]$ for all $j = 0, \dots, N$, where the integer N satisfies

$$\sum_{j=0}^N m_j \leq K_0 \log(\varepsilon^{-1}).$$

In other words, Theorem 1.1 says that, around the equilibrium $u = 0$, the Kirchhoff equation possesses solutions whose norm $\mathcal{N}(t)$ exhibits oscillations that follow any prescribed sequence of times on the time interval $[0, T_\varepsilon]$, and the number N of oscillations within that interval, or more generally the sum of the time lengths of the oscillations, is arbitrarily large for ε small enough. Roughly speaking, the norm $\mathcal{N}(t)$ is centered at A_ε and it oscillates with amplitude B_ε , staying in the upper half $[A_\varepsilon, A_\varepsilon + B_\varepsilon]$ over the time interval I_j and in the lower half $[A_\varepsilon - B_\varepsilon, A_\varepsilon]$ over the time interval E_j , up to small errors δ_ε . These oscillations can be seen as a *chaotic-like modulation* of a stable motion, meaning that the oscillating solutions are of size ε , they are ε^2 -close to effectively stable solutions (over long timescales), but they exhibit chaotic-like exchanges of size ε^2 between the amplitude of different Fourier modes. The chaotic-like modulations are obtained thanks to the existence of a Smale's horseshoe in a universal model, obtained by suitable scalings and certain transformations (see Section 3), which gives rise to chaotic oscillations of amplitude of size $O(1)$ for all times $t \in \mathbb{R}$ (see Proposition 4.1).

Remark 1.2. The solution $u(t, x)$ in Theorem 1.1 is Fourier supported on the set $\{k \in \mathbb{Z}^d : |k| \in \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}\}$, where

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

and m, p are integers with $2 \leq m < p$ and ratio $\sigma = m/p \leq \sigma_*$, where σ_* is a universal constant.

In fact, the ratio σ is the perturbation parameter we use in the entire construction. In principle, the constant M in Theorem 1.1 depends on the ratio $\sigma = m/p$ and it is of the order $M \sim \log(\sigma^{-1})$; see (4.14). Theorem 1.1 is stated after fixing m, p with $m = 2$ and p the minimum integer such that $p > 2$ and $2/p \leq \sigma_*$.

Remark 1.3. In Theorem 1.1 the Sobolev norm $\mathcal{N}(t)$ is used to describe the transfer of energy between Fourier modes, because $H^{\frac{3}{2}}(\mathbb{T}^d) \times H^{\frac{1}{2}}(\mathbb{T}^d)$ is the space of the standard local well-posedness for the Kirchhoff equation. Since the solution $u(t, x)$ in Theorem 1.1 has a fixed, finite Fourier support for all times, all the Sobolev norms of $(u, \partial_t u)$ are equivalent, and all are equally able to describe the chaotic transfer of energy among the Fourier modes – all except the norm of the energy space $H^1(\mathbb{T}^d) \times L^2(\mathbb{T}^d)$, which corresponds to a conserved quantity of the approximating system that we use in the construction; see Remark 3.2.

Remark 1.4. The factor $1/10$ in the definition of δ_ε in Theorem 1.1 comes from an arbitrary choice. We could replace $1/10$ by any other positive number; in that case, the constants $M, \varepsilon_*, C, r_0, b, K, K_0$ must be chosen accordingly.

Remark 1.5. For simplicity, Theorem 1.1 and its proof are stated entirely in terms of nonnegative times. However, with only minor changes, one proves that the result holds over the time interval $[-T_\varepsilon, T_\varepsilon]$.

Remark 1.6. Adapting the formulation of the symbolic dynamics for the approximating system (see Proposition 4.7), one can prove an alternative version of Theorem 1.1, where the prescribed random behavior of the norm $\mathcal{N}(t)$ is not only given by the sequence of the time lengths of its oscillations, but also by any sequence (a_0, a_1, a_2, \dots) with $a_j \in \{0, 1\}$ prescribing the ordered sequence of “up” and “down” movements of $\mathcal{N}(t)$. In that case, $\mathcal{N}(t)$ still makes oscillations of order ε^2 around a central value of order ε , varying in a range, say, $[\ell_\varepsilon, h_\varepsilon]$; the difference with respect to Theorem 1.1 is that, in the j th time interval, $\mathcal{N}(t)$ gets close to the low value ℓ_ε , and it remains in the slightly enlarged lower half of the range if $a_j = 0$, while $\mathcal{N}(t)$ gets close to the high value h_ε , and it remains in the slightly enlarged upper half of the range if $a_j = 1$.

In other words, around the equilibrium $u = 0$, the Kirchhoff equation possesses solutions whose norms $\mathcal{N}(t)$ exhibit oscillations that follow any prescribed sequence of “up” and “down” on the time interval $[0, T_\varepsilon]$.

Remark 1.7. The result in [3] shows that there are *no* transfers of energy of size ε between Fourier spheres in a time interval of length ε^{-4} . This could make one think that, on such a timescale, between Fourier spheres there are no energy transfers at all. Theorem 1.1 shows that this is not true; in particular, it proves the existence of chaotic transfers of energy of smaller size on a shorter timescale, i.e., transfers of size ε^2 on a timescale $\varepsilon^{-3} \log(\varepsilon^{-1})$.

1.2. Main ideas of the proof

The main steps of the proof of Theorem 1.1 can be summarized as follows:

- (1) Derive an effective resonant model for small solutions of (1.1). This reduced system is obtained by using normal form arguments and introducing some “macroscopic” variables describing the collective behavior of Fourier frequencies with the same modulus.
- (2) Show that, carefully choosing a finite set of Fourier frequencies, one can make the effective system nearly integrable.
- (3) Prove the existence of chaotic dynamics (a Smale horseshoe) for the effective system.
- (4) Show that certain solutions of the Kirchhoff equation (1.1) closely follow those in the Smale horseshoe of the effective system for a sufficiently long time interval.

The effective system is obtained with a normal form analysis. To this end, in Section 2, we perform two steps of *quasilinear normal form* (following [3, 4]), and introduce a set of special variables, found in [4], which allow one to reduce the dimension of the problem without losing information on the time evolution of the Sobolev norm of the solution. Thus, the resulting reduced model can be seen as a “macroscopic” effective system, where we do not distinguish the evolution of the energy of each single Fourier mode.

The reduction to a finite-dimensional effective system is done in Section 3. We restrict the Fourier support to two coupled resonant triplets. The space of functions supported on these modes is invariant for equation (1.1), thanks to the particular form of its nonlinearity. Relying on symmetries of the problem, we are able to further reduce the model to obtain a four-dimensional system.

The next step is to construct chaotic motions for such a system. Since we rely on perturbative techniques, we want the system to be nearly integrable; this is obtained by choosing resonant triplets with Fourier modes as explained in Remark 1.2. In particular, the system behaves as a pair of weakly coupled pendulums.

Then, in Section 4, we apply the classical Poincaré–Melnikov theory [28] to prove that the system has a hyperbolic periodic orbit with transverse homoclinic orbits. By the classical Smale–Birkhoff theorem, this implies the existence of a Smale horseshoe, which is a hyperbolic invariant set with symbolic dynamics. Note that the set is invariant and therefore one can describe the dynamics of its orbits *for all times*.

Finally, it remains to translate the dynamics of the effective system to the original equation (1.1). We prove that there exist solutions of the full PDE that closely follow those of the effective system. Even if the approximation argument is done through a Grönwall estimate (see Section 6), this is a rather delicate procedure. Indeed, since we have performed several reductions, rescalings, and two steps of normal form, we have to ensure that the solutions of the Kirchhoff equation shadowing those of the effective system satisfy all the required constraints over a sufficiently long timescale. This final part of the proof is done in Sections 5 and 6.

The general strategy of the proof is similar to that developed in [13] for the cubic wave and beam equations. The proof of Theorem 1.1, however, is based on a higher-order normal form analysis, which is needed to consider systems which are integrable at first order. This is the typical situation for PDEs on one-dimensional spatial domains. Indeed, resonant Hamiltonian monomials of low degree, which provide the dominant dynamics close to the origin, usually do not drastically change the Fourier actions (and so, the Sobolev norms). Main examples are given by the KdV, Klein–Gordon, and Schrödinger equations and the pure gravity water waves equation in infinite depth, under Dirichlet or periodic boundary conditions. This is somewhat the case for equation (1.1) also, even if the spatial domain is the torus \mathbb{T}^d of any dimension $d \geq 1$, and even if the integrability property of the equation at cubic order only holds for the macroscopic variables. This makes the implementation of the above strategy rather delicate. One of the issues in performing this kind of analysis is that interesting instability phenomena only occur after a longer time.

Another relevant difference with respect to the equations considered in [13] is that equation (1.1) is quasilinear, namely the nonlinearity contains derivatives of the same order as the linear part. This fact is not trivial, because, even if one is able to construct normal form transformations for the quasilinear equation (1.1) (as done in [3,4]), here one has to be able to provide a result of approximation between the effective model and the full PDE for a long timescale. This requires one to consider an equation for the difference of a special orbit of the effective system and a solution of (1.1). This equation is quasilinear itself and presents a time-dependent linear part. For such an equation one has to provide a result of long-time stability.

Another difference with respect to [13] regards a quantitative aspect in the energy exchange between Fourier frequencies: in [13] a large portion of the energy is exchanged among Fourier frequencies having similar modulus; here, on the contrary, a very small portion of the energy is exchanged among Fourier frequencies of modulus $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ (see Remark 1.2), where $\alpha_2, \alpha_3, \alpha_4$ are much larger than α_1 .

2. Effective dynamics for the Kirchhoff equation

In this section we recall how the “macroscopic” quantities S_λ, B_λ in (2.12) are derived from the Kirchhoff equation (1.1), starting with the normal form procedure.

2.1. A quasilinear partial normal form

Written as a first-order evolution equation, (1.1) becomes

$$\begin{cases} \partial_t u = v, \\ \partial_t v = \left(1 + \int_{\mathbb{T}^d} |\nabla u|^2 dx\right) \Delta u. \end{cases} \quad (2.1)$$

It is proved in [3] and [4] (see also the shorter, unified description in [5, Appendix A]) that system (2.1) can be transformed, after two steps of a quasilinear, partial normal form procedure, into another system, where the cubic and the quintic terms are in normal form (up to harmless terms that do not contribute to energy estimates). More precisely, it is proved that, renaming (\tilde{u}, \tilde{v}) the original, “physical” variables of system (2.1), there exists a change of variable $(\tilde{u}, \tilde{v}) = \Phi(u, v)$ (see (2.5)) such that system (2.1) becomes

$$\partial_t(u, v) = W(u, v), \quad (2.2)$$

where

$$W(u, v) = (1 + \mathcal{P}(u, v))(\mathcal{D}_1(u, v) + \mathcal{Z}_3(u, v) + \mathcal{Z}_5(u, v)) + W_{\geq 7}(u, v). \quad (2.3)$$

The unknown (u, v) for the transformed system (2.2) is a pair of complex conjugate functions with zero average over \mathbb{T}^d . The term $(1 + \mathcal{P}(u, v))$ is a scalar multiplicative factor, close to 1, depending on (u, v) , and it is a function of time, independent of the space variable x . Also, \mathcal{D}_1 is the linear operator $\mathcal{D}_1(u, v) := (-i|D_x|u, i|D_x|v)$, where $|D_x|$ is the Fourier multiplier $e^{ik \cdot x} \mapsto |k|e^{ik \cdot x}$, $k \in \mathbb{Z}^d$. Note that $|k|$ are the linear frequencies of oscillation of system (2.2). The solutions of equations as

$$\sigma_1|k_1| + \cdots + \sigma_n|k_n| = 0, \quad \sigma_i \in \{\pm 1\}, \quad i = 1, \dots, n,$$

give the resonances of order n and identify the resonant Hamiltonian terms. Next, \mathcal{Z}_3 is the cubic resonant operator $\mathcal{Z}_3(u, v) = ((\mathcal{Z}_3)_1(u, v), (\mathcal{Z}_3)_2(u, v))$ with components

$$\begin{aligned} (\mathcal{Z}_3)_1(u, v) &:= -\frac{i}{4} \sum_{\substack{j, k \in \mathbb{Z}^d \setminus \{0\} \\ |j|=|k|}} u_j u_{-j} |j|^2 v_k e^{ik \cdot x}, \\ (\mathcal{Z}_3)_2(u, v) &:= \frac{i}{4} \sum_{\substack{j, k \in \mathbb{Z}^d \setminus \{0\} \\ |j|=|k|}} v_j v_{-j} |j|^2 u_k e^{ik \cdot x}, \end{aligned}$$

where u_j, v_j are the Fourier coefficients of u, v . The entire term $(1 + \mathcal{P}(u, v))(\mathcal{D}_1(u, v) + \mathcal{Z}_3(u, v))$ gives no contribution to the energy estimates. The term \mathcal{Z}_5 is the resonant

quintic operator $\mathcal{Z}_5(u, v) = ((\mathcal{Z}_5)_1(u, v), (\mathcal{Z}_5)_2(u, v))$, where

$$\begin{aligned}
 (\mathcal{Z}_5)_1(u, v) &= \frac{i}{32} \sum_{\substack{j, \ell, k \in \mathbb{Z}^d \setminus \{0\} \\ |j| = |\ell|}} u_j u_{-j} v_\ell v_{-\ell} u_k e^{ik \cdot x} |j|^2 |\ell|^2 \left(\frac{1}{|j| + |k|} - \frac{(1 - \delta_{|k|}^{|j|})}{|\ell| - |k|} \right) \\
 &+ \frac{3i}{32} \sum_{\substack{j, \ell, k \in \mathbb{Z}^d \setminus \{0\} \\ |k| = |j| + |\ell|}} u_j u_{-j} u_\ell u_{-\ell} v_k e^{ik \cdot x} |j| |\ell| |k| \\
 &+ \frac{i}{16} \sum_{\substack{j, \ell, k \in \mathbb{Z}^d \setminus \{0\} \\ |j| = |k|}} u_j u_{-j} u_\ell v_{-\ell} v_k e^{ik \cdot x} |j|^2 |\ell| \\
 &\quad \times \left(6 + \frac{|\ell|}{|\ell| + |j|} + \frac{|\ell|(1 - \delta_{|\ell|}^{|j|})}{|\ell| - |j|} \right) \\
 &+ \frac{3i}{16} \sum_{\substack{j, \ell, k \in \mathbb{Z}^d \setminus \{0\} \\ |k| = |j| - |\ell|}} u_j u_{-j} v_\ell v_{-\ell} v_k e^{ik \cdot x} |j| |\ell| |k|, \tag{2.4}
 \end{aligned}$$

and $(\mathcal{Z}_5)_2(u, v)$ is obtained from $(\mathcal{Z}_5)_1(u, v)$ by complex conjugation. For the coefficients in (2.4) we adopt the convention that $\frac{0}{0} = 0$, where δ is the usual Kronecker delta. The term $W_{\geq 7}(u, v)$ in (2.3) contains only terms of homogeneity at least 7 in (u, v) , and it is estimated in [4] and [5]. We note that the monomials appearing in (2.4) are resonant, because the small denominators appearing in the second step of the Birkhoff procedure are of the form

$$2(|j| - |\ell|) + |k| - |k| = 0, \quad 2(|j| + |k| - |\ell|) = 0.$$

The map Φ that transforms (2.1) into (2.2) is obtained by composition, and it is

$$\Phi = \Phi_1 \circ \Phi_2 \circ \Phi_3 \circ \Phi_4 \circ \Phi_5. \tag{2.5}$$

The maps Φ_1 and Φ_2 are simply the linear operators that symmetrize and diagonalize the linear part of system (2.1) (i.e., the linear wave equation); see (6.27). The map Φ_3 is a nonlinear, preparatory transformation, which is required because the problem is quasi-linear. The map Φ_4 is the transformation of the first step of the normal form procedure, and Φ_5 is that of the second step. The explicit expressions of Φ_3 and Φ_4 are given in [3] and that of Φ_5 in [4]. Unlike Φ_1 and Φ_2 , the transformations Φ_3, Φ_4, Φ_5 are all close to the identity map. In the present paper we do not use the explicit formula for Φ_3, Φ_4, Φ_5 , but only the following properties of their composition.

For $s \in \mathbb{R}$, let $H_0^s(\mathbb{T}^d, \mathbb{C})$ be the Sobolev space of zero-average, complex-valued functions

$$H_0^s(\mathbb{T}^d, \mathbb{C}) := \{u : \mathbb{T}^d \rightarrow \mathbb{C} : u_0 = 0, \|u\|_s < \infty\}, \quad \|u\|_s^2 := \sum_{k \in \mathbb{Z}^d \setminus \{0\}} |u_k|^2 |k|^{2s},$$

where $u_k, k \in \mathbb{Z}^d$, are the Fourier coefficients of u , and let

$$H_0^s(\mathbb{T}^d, \text{c.c.}) := \{(u, v) : u, v \in H_0^s(\mathbb{T}^d, \mathbb{C}), v = \bar{u}\} = \{(u, \bar{u}) : u \in H_0^s(\mathbb{T}^d, \mathbb{C})\},$$

where the notation ‘‘c.c.’’ reminds one that they are pairs of complex conjugate functions. Given $d \in \mathbb{N}$, let

$$m_1 := 1 \text{ if } d = 1, \quad m_1 := 2 \text{ if } d \geq 2. \quad (2.6)$$

Lemma 2.1 (From [5, Lemma 2.9]). *There exist universal constants $\delta, C > 0$ such that for all $(u, v) \in H_0^{m_1}(\mathbb{T}^d, \text{c.c.})$ in the ball $\|u\|_{m_1} \leq \delta$, for all $k \in \mathbb{Z}^d$, the k th Fourier coefficient $f_k = \bar{g}_{-k}$ of $(f, g) := (\Phi_3 \circ \Phi_4 \circ \Phi_5)(u, v)$ satisfies*

$$|f_k - u_k| \leq C \|u\|_{m_1}^2 (|u_k| + |u_{-k}|). \quad (2.7)$$

From (2.7) it follows that $\|f - u\|_s \leq 2C \|u\|_{m_1}^2 \|u\|_s$ for all $s \in \mathbb{R}$, and that the map $\Phi_3 \circ \Phi_4 \circ \Phi_5$ is well defined in the ball $\|u\|_{m_1} \leq \delta$, with $\|f\|_s \leq \|u\|_s (1 + 2C \|u\|_{m_1}^2)$, $s \in \mathbb{R}$.

Inequality (2.7) also implies the invariance of the Fourier support: if $u_k = u_{-k} = 0$ for some $k \in \mathbb{Z}^d$, then also $f_k = f_{-k} = 0$, and vice versa (since $C\delta^2 < 1/2$).

Lemma 2.2 (From [5, Lemma 2.3]). *There exist universal constants $\delta_1, C_1, C_0 > 0$ with the following properties. Let $(u_0, v_0) \in H_0^{m_1}(\mathbb{T}^d, \text{c.c.})$ and*

$$\|u_0\|_{m_1} \leq \delta_1. \quad (2.8)$$

Then the Cauchy problem of system (2.2) with initial condition

$$(u(0), v(0)) = (u_0, v_0) \quad (2.9)$$

has a unique solution $(u, v) \in C([0, T_{\text{NF}}], H_0^{m_1}(\mathbb{T}^d, \text{c.c.}))$ on the time interval $[0, T_{\text{NF}}]$. The solution satisfies

$$\|u(t)\|_{m_1} \leq C_1 \|u_0\|_{m_1} \leq \delta \quad \forall t \in [0, T_{\text{NF}}], \quad T_{\text{NF}} = C_0 \|u_0\|_{m_1}^{-4}, \quad (2.10)$$

where δ is the constant in Lemma 2.1. As a consequence, for all $t \in [0, T_{\text{NF}}]$ the solution $(u(t), v(t))$ remains in the ball $\|u(t)\|_{m_1} \leq \delta$ where $\Phi_3 \circ \Phi_4 \circ \Phi_5$ is well defined, the function

$$(\tilde{u}(t), \tilde{v}(t)) := \Phi(u(t), v(t)) \quad (2.11)$$

(where Φ is the map in (2.5)) solves the original system (2.1) on the time interval $[0, T_{\text{NF}}]$, and $\tilde{u}(t)$ solves the Kirchhoff equation (1.1) on $[0, T_{\text{NF}}]$.

As the notation suggests, T_{NF} is the existence time we obtain by the normal form procedure. For more details on the map Φ and on the transformed vector field $W(u, v)$ see [3–5].

2.2. The effective system

We recall the derivation of the *effective system* (or *effective equation*) from [4,5]. Let

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

For any pair $(u, v) \in L^2(\mathbb{T}^d, \text{c.c.})$ of complex conjugate functions, for any $\lambda \in \Gamma$ we define

$$S_\lambda := \sum_{\substack{k \in \mathbb{Z}^d \\ |k|=\lambda}} |u_k|^2 = \sum_{\substack{k \in \mathbb{Z}^d \\ |k|=\lambda}} u_k v_{-k}, \quad B_\lambda := \sum_{\substack{k \in \mathbb{Z}^d \\ |k|=\lambda}} u_k u_{-k}, \quad (2.12)$$

and note that

$$S_\lambda \geq 0, \quad B_\lambda \in \mathbb{C}, \quad |B_\lambda| \leq S_\lambda.$$

The quantity S_λ is called the ‘‘superaction’’ of u on the sphere $|k| = \lambda$. Its evolution on the time interval $[0, T_{\text{NF}}]$ remains confined between two multiples of its initial value, as is observed in the next lemma.

Lemma 2.3 (From [5, Lemma 2.4]). *Let $(u_0, v_0) \in H_0^{m_1}(\mathbb{T}^d, \text{c.c.})$, with u_0 in the ball (2.8). Let $(u(t), v(t))$ be the solution of the Cauchy problem (2.2), (2.9) on the time interval $[0, T_{\text{NF}}]$, with T_{NF} in (2.10), given by Lemma 2.2. For every $t \in [0, T_{\text{NF}}]$, let $S_\lambda(t)$ be the sum defined in (2.12). Then*

$$C_1 S_\lambda(0) \leq S_\lambda(t) \leq C_2 S_\lambda(0) \quad (2.13)$$

for all $t \in [0, T_{\text{NF}}]$, for all $\lambda \in \Gamma$, where $C_1, C_2 > 0$ are universal constants.

By (2.13), for every $\lambda \in \Gamma$, either $S_\lambda(t) > 0$ for all $t \in [0, T_{\text{NF}}]$, or $S_\lambda(t) = 0$ for all $t \in [0, T_{\text{NF}}]$. Hence, we decompose Γ as the disjoint union of

$$\Gamma_0 := \{\lambda \in \Gamma : S_\lambda = 0 \forall t\}, \quad \Gamma_1 := \{\lambda \in \Gamma : S_\lambda > 0 \forall t\}. \quad (2.14)$$

It is observed in [4,5] that, if $(u(t), v(t))$ solves (2.2) on some time interval, then, for every $\lambda \in \Gamma$, calculating the Fourier coefficients of $W(u, v)$ in (2.3) and taking the sum over all indices $k \in \mathbb{Z}^d$ on the sphere $|k| = \lambda$, the corresponding quantities $S_\lambda(t), B_\lambda(t)$ in (2.12) satisfy the equations

$$\begin{aligned} \partial_t S_\lambda &= \frac{3i}{32} \sum_{\substack{\alpha, \beta \in \Gamma \\ \alpha + \beta = \lambda}} (B_\alpha B_\beta \overline{B_\lambda} - \overline{B_\alpha} \overline{B_\beta} B_\lambda) \alpha \beta \lambda \\ &\quad + \frac{3i}{16} \sum_{\substack{\alpha, \beta \in \Gamma \\ \alpha - \beta = \lambda}} (B_\alpha \overline{B_\beta} \overline{B_\lambda} - \overline{B_\alpha} B_\beta B_\lambda) \alpha \beta \lambda + R_{S_\lambda}, \\ \partial_t B_\lambda &= -2i(1 + \mathcal{P}) \left(\lambda + \frac{1}{4} \lambda^2 S_\lambda \right) B_\lambda + R_{B_\lambda} \end{aligned} \quad (2.15)$$

on the same time interval, where the terms $R_{S_\lambda}, R_{B_\lambda}$ satisfy the following estimates.

Lemma 2.4 ([5, Lemma 2.2]). *Let $(u, v) \in H_0^{m_1}(\mathbb{T}^d, \text{c.c.})$ with $\|u\|_{m_1} \leq \delta$, where δ is the constant given by Lemma 2.1 and appearing in (2.10). Then, for all $\lambda \in \Gamma$, the terms R_{S_λ} , R_{B_λ} in (2.15) satisfy*

$$|R_{S_\lambda}| \leq C \|u\|_{m_1}^6 S_\lambda, \quad |R_{B_\lambda}| \leq C \|u\|_{m_1}^4 S_\lambda, \quad (2.16)$$

where $C > 0$ is a universal constant.

Define

$$Z_{\alpha\beta\lambda} := B_\alpha B_\beta \overline{B_\lambda}, \quad \vartheta_{\alpha\beta\lambda} := \text{Im}(Z_{\alpha\beta\lambda}) = \frac{B_\alpha B_\beta \overline{B_\lambda} - \overline{B_\alpha B_\beta} B_\lambda}{2i}.$$

If S_λ, B_λ satisfy system (2.15), then $S_\lambda, Z_{\alpha\beta\lambda}$ satisfy

$$\partial_t S_\lambda = -\frac{3}{16} \sum_{\substack{\alpha, \beta \in \Gamma \\ \alpha + \beta = \lambda}} \vartheta_{\alpha\beta\lambda} \alpha \beta \lambda + \frac{3}{8} \sum_{\substack{\alpha, \beta \in \Gamma \\ \beta + \lambda = \alpha}} \vartheta_{\beta\lambda\alpha} \alpha \beta \lambda + R_{S_\lambda}, \quad (2.17)$$

$$\partial_t Z_{\alpha\beta\lambda} = -2i(1 + \mathcal{P}) \left(\alpha + \beta - \lambda + \frac{1}{4}(\alpha^2 S_\alpha + \beta^2 S_\beta - \lambda^2 S_\lambda) \right) Z_{\alpha\beta\lambda} + \tilde{R}_{Z_{\alpha\beta\lambda}},$$

where

$$\tilde{R}_{Z_{\alpha\beta\lambda}} := R_{B_\alpha} B_\beta \overline{B_\lambda} + B_\alpha R_{B_\beta} \overline{B_\lambda} + B_\alpha B_\beta \overline{R_{B_\lambda}}.$$

For $\alpha + \beta = \lambda$, isolating the first nontrivial contribution from terms of higher homogeneity orders, one has

$$\partial_t Z_{\alpha\beta\lambda} = -\frac{i}{2}(\alpha^2 S_\alpha + \beta^2 S_\beta - \lambda^2 S_\lambda) Z_{\alpha\beta\lambda} + R_{Z_{\alpha\beta\lambda}}, \quad (2.18)$$

where

$$R_{Z_{\alpha\beta\lambda}} := -\frac{i}{2} \mathcal{P}_2(\alpha^2 S_\alpha + \beta^2 S_\beta - \lambda^2 S_\lambda) Z_{\alpha\beta\lambda} + \tilde{R}_{Z_{\alpha\beta\lambda}}. \quad (2.19)$$

The remainder R_{S_λ} in (2.17) has been bounded in Lemma 2.4. The remainder $R_{Z_{\alpha\beta\lambda}}$ in (2.19) is estimated in the next lemma.

Lemma 2.5 ([5, Lemma 2.5]). *Assume the hypotheses of Lemma 2.4. Then for all $\alpha, \beta, \lambda \in \Gamma$ with $\alpha + \beta = \lambda$ the remainder $R_{Z_{\alpha\beta\lambda}}$ defined in (2.19) satisfies*

$$|R_{Z_{\alpha\beta\lambda}}| \leq C \|u\|_{m_1}^4 S_\alpha S_\beta S_\lambda, \quad (2.20)$$

where $C > 0$ is a universal constant.

2.3. The truncated effective system

If we remove the remainders from equations (2.17) and (2.18), we obtain a system that we call a *truncated effective system*, which is

$$\partial_t S_\lambda = -\frac{3}{16} \sum_{\substack{\alpha, \beta \in \Gamma \\ \alpha + \beta = \lambda}} \vartheta_{\alpha\beta\lambda} \alpha \beta \lambda + \frac{3}{8} \sum_{\substack{\alpha, \beta \in \Gamma \\ \beta + \lambda = \alpha}} \vartheta_{\beta\lambda\alpha} \beta \lambda \alpha, \quad (2.21)$$

$$\partial_t Z_{\alpha\beta\lambda} = -\frac{i}{2} \omega_{\alpha\beta\lambda} Z_{\alpha\beta\lambda}, \quad (2.22)$$

where

$$\omega_{\alpha\beta\lambda} := \alpha^2 S_\alpha + \beta^2 S_\beta - \lambda^2 S_\lambda.$$

For all $\alpha, \beta, \lambda \in \Gamma$ such that $\alpha + \beta = \lambda$, let

$$r_{\alpha\beta\lambda} := \operatorname{Re}(Z_{\alpha\beta\lambda}).$$

Since $\omega_{\alpha\beta\lambda}$ is real, the real and imaginary parts of equation (2.22) are given by the system

$$\partial_t r_{\alpha\beta\lambda} = \frac{1}{2} \omega_{\alpha\beta\lambda} \vartheta_{\alpha\beta\lambda}, \quad \partial_t \vartheta_{\alpha\beta\lambda} = -\frac{1}{2} \omega_{\alpha\beta\lambda} r_{\alpha\beta\lambda}. \quad (2.23)$$

The solutions $(r_{\alpha\beta\lambda}(t), \vartheta_{\alpha\beta\lambda}(t))$ of (2.23) remain on a circle, because they satisfy

$$\partial_t (|Z_{\alpha\beta\lambda}|^2) = \partial_t (r_{\alpha\beta\lambda}^2 + \vartheta_{\alpha\beta\lambda}^2) = 0.$$

Thus $|Z_{\alpha\beta\lambda}|$ is a prime integral of the truncated effective system (2.21)–(2.22). Therefore, if $S_\lambda, Z_{\alpha\beta\lambda}$ solve (2.21)–(2.22), then the real and imaginary parts of $Z_{\alpha\beta\lambda}$ satisfy

$$r_{\alpha\beta\lambda}(t) = \rho_{\alpha\beta\lambda} \cos(\varphi_{\alpha\beta\lambda}(t)), \quad \vartheta_{\alpha\beta\lambda}(t) = \rho_{\alpha\beta\lambda} \sin(\varphi_{\alpha\beta\lambda}(t)), \quad (2.24)$$

where

$$\rho_{\alpha\beta\lambda} = |Z_{\alpha\beta\lambda}| \geq 0$$

is a constant, and $\varphi_{\alpha\beta\lambda}(t)$ is an angle. Moreover, plugging (2.24) into (2.23) gives the equation for the evolution of the angle. Hence the truncated effective system (2.21)–(2.22) becomes

$$\partial_t S_\lambda = -\frac{1}{2} \sum_{\substack{\alpha, \beta \in \Gamma \\ \alpha + \beta = \lambda}} c_{\alpha\beta\lambda} \sin(\varphi_{\alpha\beta\lambda}) + \sum_{\substack{\alpha, \beta \in \Gamma \\ \beta + \lambda = \alpha}} c_{\beta\lambda\alpha} \sin(\varphi_{\beta\lambda\alpha}), \quad (2.25)$$

$$\partial_t \varphi_{\alpha\beta\lambda} = -\frac{1}{2} (\alpha^2 S_\alpha + \beta^2 S_\beta - \lambda^2 S_\lambda), \quad (2.26)$$

where

$$c_{\alpha\beta\lambda} := \frac{3}{8} \rho_{\alpha\beta\lambda} \alpha \beta \lambda = \frac{3}{8} |Z_{\alpha\beta\lambda}| \alpha \beta \lambda = \frac{3}{8} |B_\alpha| |B_\beta| |B_\lambda| \alpha \beta \lambda \quad (2.27)$$

is a constant. We note that (2.25) and (2.26) form a closed system for the variables $S_\lambda, \varphi_{\alpha\beta\lambda}$.

3. The truncated effective system with two triplets

We consider the case in which the Fourier support Γ_1 in (2.14) has only four distinct elements, forming two resonant triplets with two elements in common, in the following way:

$$\begin{aligned} \Gamma_1 &= \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}, \quad \alpha_1 < \alpha_2 < \alpha_3 < \alpha_4, \\ \alpha_1 + \alpha_2 &= \alpha_3, \quad \alpha_2 + \alpha_3 = \alpha_4, \quad 2\alpha_1 \neq \alpha_2. \end{aligned} \quad (3.1)$$

We can assume, without loss of generality, that the four elements of Γ_1 are natural numbers. Examples of such sets are any four consecutive elements of the Fibonacci sequence greater than 1, like $\{2, 3, 5, 8\}$, or, more generally, any set of the form

$$\alpha_1 := m, \quad \alpha_2 := m + p, \quad \alpha_3 := 2m + p, \quad \alpha_4 := 3m + 2p, \quad (3.2)$$

where m, p are distinct positive integers with

$$2 \leq m < p.$$

Lemma 3.1. *If $\alpha, \beta, \lambda \in \Gamma_1$ in (3.1) satisfy $\alpha + \beta = \lambda$, then the ordered triplet (α, β, λ) must be of the form*

$$(\alpha_1, \alpha_2, \alpha_3) \text{ or } (\alpha_2, \alpha_1, \alpha_3) \text{ or } (\alpha_2, \alpha_3, \alpha_4) \text{ or } (\alpha_3, \alpha_2, \alpha_4),$$

with $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ in (3.2), and there are no other options.

Proof. For example, one has

$$\alpha_3 = \alpha_1 + \alpha_2 < \alpha_2 + \alpha_2 < \alpha_2 + \alpha_3 = \alpha_4,$$

therefore $2\alpha_2 \notin \Gamma_1$, and the triplet $(\alpha_2, \alpha_2, 2\alpha_2)$ is not admissible; the other cases can be checked similarly. ■

To slightly shorten the notation, we denote

$$S_1 := S_{\alpha_1}, \quad \varphi_{123} := \varphi_{\alpha_1 \alpha_2 \alpha_3}, \quad c_{123} := c_{\alpha_1 \alpha_2 \alpha_3},$$

and so on. Hence, system (2.25)–(2.26) becomes

$$\begin{aligned} \partial_t S_1 &= c_{123} \sin(\varphi_{123}), \\ \partial_t S_2 &= c_{123} \sin(\varphi_{123}) + c_{234} \sin(\varphi_{234}), \\ \partial_t S_3 &= -c_{123} \sin(\varphi_{123}) + c_{234} \sin(\varphi_{234}), \\ \partial_t S_4 &= -c_{234} \sin(\varphi_{234}), \\ \partial_t \varphi_{123} &= -\frac{1}{2}(\alpha_1^2 S_1 + \alpha_2^2 S_2 - \alpha_3^2 S_3), \\ \partial_t \varphi_{234} &= -\frac{1}{2}(\alpha_2^2 S_2 + \alpha_3^2 S_3 - \alpha_4^2 S_4), \end{aligned} \quad (3.3)$$

which is a system of six equations in six unknowns.

3.1. Meaningfulness condition for the solutions

Our strategy is this: we want to find solutions of the truncated effective system (3.3) with a prescribed, interesting dynamical behavior, and to show that the solution of the effective system (2.17)–(2.18) is so close to the solution of the truncated effective system (3.3)

that the dynamical behaviors of the two solutions are very similar, on a sufficiently long interval of time. Later, in Section 6, we show that there exists a solution of the original PDE (2.1) which is very close (up to the change of coordinates Φ) to the solution of the effective system.

Hence we look for solutions of the truncated effective system (3.3) that satisfy the natural meaningfulness condition required by system (2.17)–(2.18), which is simply this: if a solution is defined on a time interval $[0, T]$, then it must satisfy

$$S_n(t) > 0 \quad \forall n = 1, 2, 3, 4, \quad \forall t \in [0, T]. \quad (3.4)$$

So we reject any solution of (3.3) such that some of the S_n become nonpositive at some time t (recall definitions (2.12) and (2.14)).

In the following analysis, we first ignore the constraint (3.4); later, we will select only solutions satisfying it. Analogously, we first consider the coefficients c_{123}, c_{234} as any two given constants; later, we will go back to the identities (2.27).

3.2. First integrals and a linear change of coordinates

Given any linear combination $E := \mu_1 S_1 + \mu_2 S_2 + \mu_3 S_3 + \mu_4 S_4$ of S_1, \dots, S_4 with constant real coefficients μ_1, \dots, μ_4 , we have

$$\partial_t E = (\mu_1 + \mu_2 - \mu_3)c_{123} \sin(\varphi_{123}) + (\mu_2 + \mu_3 - \mu_4)c_{234} \sin(\varphi_{234})$$

along the solutions of system (3.3). Hence any E with coefficients μ_1, \dots, μ_4 satisfying

$$\mu_1 + \mu_2 - \mu_3 = 0, \quad \mu_2 + \mu_3 - \mu_4 = 0$$

is a first integral. We choose the two functionally independent first integrals

$$E_1 := S_1 + S_3 + S_4, \quad E_2 := S_2 + S_3 + 2S_4. \quad (3.5)$$

Remark 3.2. One has

$$\alpha_1 E_1 + \alpha_2 E_2 = \alpha_1(S_1 + S_3 + S_4) + \alpha_2(S_2 + S_3 + 2S_4) = \sum_{n=1}^4 \alpha_n S_n, \quad (3.6)$$

because $\alpha_1 + \alpha_2 = \alpha_3$ and $\alpha_1 + 2\alpha_2 = \alpha_2 + (\alpha_1 + \alpha_2) = \alpha_2 + \alpha_3 = \alpha_4$. Hence, when S_n are given by (2.12), identity (3.6) implies that the Sobolev norm $\|u\|_{\frac{1}{2}}^2 = \sum_{n=1}^4 \alpha_n S_n$ is also a first integral of (3.3).

At each time t , the values $S_1(t), S_2(t)$ can be obtained from $E_1, E_2, S_3(t), S_4(t)$ by (3.5), i.e.,

$$S_1(t) = E_1 - S_3(t) - S_4(t), \quad S_2(t) = E_2 - S_3(t) - 2S_4(t). \quad (3.7)$$

Hence, system (3.3) can be reduced to a system of four equations in the four unknowns $S_3, S_4, \varphi_{123}, \varphi_{234}$, obtained by replacing S_1, S_2 by (3.7) in the last two equations of (3.3). We get

$$\begin{aligned} \partial_t S_3 &= -c_{123} \sin(\varphi_{123}) + c_{234} \sin(\varphi_{234}), \\ \partial_t S_4 &= -c_{234} \sin(\varphi_{234}), \\ \partial_t \varphi_{123} &= -\frac{1}{2}(\alpha_1^2 E_1 + \alpha_2^2 E_2) + \frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \alpha_3^2)S_3 + \frac{1}{2}(\alpha_1^2 + 2\alpha_2^2)S_4, \\ \partial_t \varphi_{234} &= -\frac{1}{2}\alpha_2^2 E_2 - \frac{1}{2}(\alpha_3^2 - \alpha_2^2)S_3 + \frac{1}{2}(2\alpha_2^2 + \alpha_4^2)S_4. \end{aligned} \quad (3.8)$$

We summarize the observations above in the following lemma.

Lemma 3.3. *Let c_{123}, c_{234} be any two constants. The following properties hold:*

- (i) *Let $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ be a solution of system (3.3) on some time interval I . Then E_1, E_2 defined by (3.5) are constant in time and $(S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ solves system (3.8) on I .*
- (ii) *Let E_1, E_2 be constants, and let $(S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ be a solution of system (3.8) on some time interval I . Define the functions $S_1(t), S_2(t)$ by the identities (3.7). Then $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ solves system (3.3) on I .*

We note that the sum of the first two equations in (3.8) does not contain the angle φ_{234} . Hence, we consider a linear change of variable that treats the sum $S_3 + S_4$ as a new variable.

Lemma 3.4. *Let c_{123}, c_{234} be any two constants. The following properties hold:*

- (i) *Let E_1, E_2 be constants, and let $(S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ be a solution of system (3.8) on some time interval I . Define the functions $x_1(t), x_2(t), y_1(t), y_2(t)$ by the change of coordinates*

$$\varphi_{123} = x_1, \quad \varphi_{234} = x_2, \quad S_3 = y_1 - y_2, \quad S_4 = y_2. \quad (3.9)$$

Then $(x_1(t), x_2(t), y_1(t), y_2(t))$ solves

$$\begin{aligned} \partial_t x_1 &= -\frac{1}{2}b_1 + \frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \alpha_3^2)y_1 - \frac{1}{2}(\alpha_3^2 - \alpha_2^2)y_2, \\ \partial_t x_2 &= -\frac{1}{2}b_2 - \frac{1}{2}(\alpha_3^2 - \alpha_2^2)y_1 + \frac{1}{2}(\alpha_2^2 + \alpha_3^2 + \alpha_4^2)y_2, \\ \partial_t y_1 &= -c_{123} \sin(x_1), \\ \partial_t y_2 &= -c_{234} \sin(x_2), \end{aligned} \quad (3.10)$$

on I , where b_1, b_2 are the constants

$$b_1 := \alpha_1^2 E_1 + \alpha_2^2 E_2, \quad b_2 := \alpha_2^2 E_2.$$

- (ii) Let b_1, b_2 be constants, and let $(x_1(t), x_2(t), y_1(t), y_2(t))$ be a solution of system (3.10) on some time interval I . Define the constants E_1, E_2 as

$$E_1 = \frac{b_1 - b_2}{\alpha_1^2}, \quad E_2 = \frac{b_2}{\alpha_2^2}, \quad (3.11)$$

and define $S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t)$ by (3.9). Then $(S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ solves system (3.8) on I .

3.3. The Hamiltonian structure

System (3.10) is the two-dimensional Hamiltonian system

$$\dot{x}_n = \partial_{y_n} H(x, y), \quad \dot{y}_n = -\partial_{x_n} H(x, y), \quad n = 1, 2,$$

with Hamiltonian

$$H(x, y) = -c_{123} \cos(x_1) - c_{234} \cos(x_2) - \frac{1}{2} b \cdot y + \frac{1}{4} A y \cdot y$$

where $b = (b_1, b_2) \in \mathbb{R}^2$ and

$$A := \begin{pmatrix} (\alpha_1^2 + \alpha_2^2 + \alpha_3^2) & -(\alpha_3^2 - \alpha_2^2) \\ -(\alpha_3^2 - \alpha_2^2) & (\alpha_2^2 + \alpha_3^2 + \alpha_4^2) \end{pmatrix}.$$

The matrix A is symmetric, positive definite, and invertible with

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} (\alpha_2^2 + \alpha_3^2 + \alpha_4^2) & (\alpha_3^2 - \alpha_2^2) \\ (\alpha_3^2 - \alpha_2^2) & (\alpha_1^2 + \alpha_2^2 + \alpha_3^2) \end{pmatrix}$$

and

$$\det A = \alpha_1^2 \alpha_4^2 + (\alpha_1^2 + \alpha_4^2)(\alpha_2^2 + \alpha_3^2) + 4\alpha_2^2 \alpha_3^2 > 0.$$

The invertibility of A is the so-called *twist condition* for the Hamiltonian H ; thanks to it, we can eliminate the linear term $b \cdot y$ from the Hamiltonian by a translation of the y -variables

$$x_1 = \tilde{x}_1, \quad x_2 = \tilde{x}_2, \quad y_1 = q_1 + \tilde{y}_1, \quad y_2 = q_2 + \tilde{y}_2, \quad (3.12)$$

with $q = A^{-1}b$, namely

$$q_1 = \frac{\alpha_2^2 + \alpha_3^2 + \alpha_4^2}{\det A} b_1 + \frac{\alpha_3^2 - \alpha_2^2}{\det A} b_2, \quad q_2 = \frac{\alpha_3^2 - \alpha_2^2}{\det A} b_1 + \frac{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}{\det A} b_2. \quad (3.13)$$

This change of coordinates is symplectic and the new Hamiltonian is just $\tilde{H}(\tilde{x}, \tilde{y}) = H(\tilde{x}, q + \tilde{y})$, whose equations are given by

$$\begin{aligned} \partial_t \tilde{x}_1 &= \frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \alpha_3^2) \tilde{y}_1 - \frac{1}{2}(\alpha_3^2 - \alpha_2^2) \tilde{y}_2, \\ \partial_t \tilde{y}_1 &= -c_{123} \sin(\tilde{x}_1), \\ \partial_t \tilde{x}_2 &= -\frac{1}{2}(\alpha_3^2 - \alpha_2^2) \tilde{y}_1 + \frac{1}{2}(\alpha_2^2 + \alpha_3^2 + \alpha_4^2) \tilde{y}_2, \\ \partial_t \tilde{y}_2 &= -c_{234} \sin(\tilde{x}_2). \end{aligned} \quad (3.14)$$

The equivalence of systems (3.10) and (3.14) is described in the following lemma.

Lemma 3.5. *Let c_{123}, c_{234} be any two constants. The following properties hold:*

- (i) *Let b_1, b_2 be constants, and let $(x_1(t), x_2(t), y_1(t), y_2(t))$ be a solution of system (3.10) on some time interval I . Define the constants q_1, q_2 by (3.13), and define the functions $\tilde{x}_1(t), \tilde{x}_2(t), \tilde{y}_1(t), \tilde{y}_2(t)$ by (3.12). Then $(\tilde{x}_1(t), \tilde{x}_2(t), \tilde{y}_1(t), \tilde{y}_2(t))$ solves (3.14).*
- (ii) *Let $(\tilde{x}_1(t), \tilde{x}_2(t), \tilde{y}_1(t), \tilde{y}_2(t))$ be a solution of system (3.14) on some interval I . Let q_1, q_2 be any two real numbers. Define constants b_1, b_2 by the identity $b = Aq$, i.e., define*

$$\begin{aligned} b_1 &= (\alpha_1^2 + \alpha_2^2 + \alpha_3^2)q_1 - (\alpha_3^2 - \alpha_2^2)q_2, \\ b_2 &= -(\alpha_3^2 - \alpha_2^2)q_1 + (\alpha_2^2 + \alpha_3^2 + \alpha_4^2)q_2, \end{aligned} \quad (3.15)$$

and define $x_1(t), x_2(t), y_1(t), y_2(t)$ by (3.12). Then $(x_1(t), x_2(t), y_1(t), y_2(t))$ solves (3.10) on I .

3.4. Normalization of coefficients by rescaling

Now we want to normalize the leading coefficients of system (3.14), using a rescaling of the time variable and dilations of the \tilde{y} -variables. We consider the change of variables

$$\tilde{x}_1(t) = \xi_1(Bt), \quad \tilde{x}_2(t) = \xi_2(Bt), \quad \tilde{y}_1(t) = A_1\eta_1(Bt), \quad \tilde{y}_2(t) = A_2\eta_2(Bt), \quad (3.16)$$

where A_1, A_2, B are defined as follows. We assume that

$$c_{123} > 0$$

and we fix

$$A_1 := \left(\frac{2c_{123}}{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} \right)^{\frac{1}{2}}, \quad B := \left(\frac{c_{123}(\alpha_1^2 + \alpha_2^2 + \alpha_3^2)}{2} \right)^{\frac{1}{2}}, \quad (3.17)$$

$$A_2 := \frac{2B}{\alpha_2^2 + \alpha_3^2 + \alpha_4^2}. \quad (3.18)$$

Note that A_1 and A_2 are related by

$$A_1 = A_2\gamma, \quad \text{with } \gamma := \frac{\alpha_2^2 + \alpha_3^2 + \alpha_4^2}{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}. \quad (3.19)$$

Thus, system (3.14) becomes

$$\begin{aligned} \dot{\xi}_1 &= \eta_1 - \mu_1\eta_2, \\ \dot{\eta}_1 &= -\sin(\xi_1), \\ \dot{\xi}_2 &= \eta_2 - \mu_2\eta_1, \\ \dot{\eta}_2 &= -\lambda \sin(\xi_2), \end{aligned} \quad (3.20)$$

where

$$\mu_1 := \frac{\alpha_3^2 - \alpha_2^2}{\alpha_2^2 + \alpha_3^2 + \alpha_4^2} = \frac{(\alpha_3^2 - \alpha_2^2)A_2}{2B}, \quad (3.21)$$

$$\mu_2 := \frac{\alpha_3^2 - \alpha_2^2}{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} = \mu_1 \gamma = \frac{(\alpha_3^2 - \alpha_2^2)A_1}{2B}, \quad (3.22)$$

$$\lambda := \frac{c_{234}(\alpha_2^2 + \alpha_3^2 + \alpha_4^2)}{c_{123}(\alpha_1^2 + \alpha_2^2 + \alpha_3^2)} = \frac{c_{234}}{c_{123}} \gamma = \frac{c_{234}}{A_2 B}. \quad (3.23)$$

We observe that the system with $\mu_1 = 0$ is given by the sum of two uncoupled Hamiltonians, while for $\mu_1 \neq 0$ the Hamiltonian structure is lost. The equivalence of systems (3.14) and (3.20) is described in the following lemma.

Lemma 3.6. *The following statements are satisfied:*

- (i) *Let c_{123}, c_{234} be any two constants, with $c_{123} > 0$, and let $(\tilde{x}_1(t), \tilde{x}_2(t), \tilde{y}_1(t), \tilde{y}_2(t))$ be a solution of system (3.14) on some time interval $[0, T]$. Let A_1, A_2, B be the constants defined in (3.17), (3.18). Define the functions $\xi_1, \xi_2, \eta_1, \eta_2$ as*

$$\xi_1(t) = \tilde{x}_1\left(\frac{t}{B}\right), \quad \xi_2(t) = \tilde{x}_2\left(\frac{t}{B}\right), \quad \eta_1(t) = \frac{1}{A_1} \tilde{y}_1\left(\frac{t}{B}\right), \quad \eta_2(t) = \frac{1}{A_2} \tilde{y}_2\left(\frac{t}{B}\right).$$

Define the constants μ_1, μ_2, λ by (3.21), (3.22), (3.23). Then $(\xi_1(t), \xi_2(t), \eta_1(t), \eta_2(t))$ solves (3.20) on the time interval $[0, BT]$.

- (ii) *Let λ be any constant. Let μ_1, μ_2 be the constants defined in (3.21), (3.22). Let $(\xi_1(t), \xi_2(t), \eta_1(t), \eta_2(t))$ be a solution of system (3.20) on some time interval $[0, T]$. Let A_1 be any positive constant. Define the constant c_{123} as*

$$c_{123} = \frac{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}{2} A_1^2, \quad (3.24)$$

and define the constant c_{234} by means of (3.23), i.e., $c_{234} = c_{123} \lambda / \gamma$, where γ is defined in (3.19). Define the constant B by the second identity in (3.17), and the constant A_2 by (3.18). Define the functions $\tilde{x}_1, \tilde{x}_2, \tilde{y}_1, \tilde{y}_2$ by (3.16). Then $(\tilde{x}_1(t), \tilde{x}_2(t), \tilde{y}_1(t), \tilde{y}_2(t))$ solves (3.14) on the time interval $[0, B^{-1}T]$.

3.5. Large Fourier frequency as a perturbation parameter

Recall definition (3.2) of $\alpha_1, \dots, \alpha_4$ as functions of the two integer parameters m, p . Then

$$\begin{aligned} \alpha_3^2 - \alpha_2^2 &= (2m + p)^2 - (m + p)^2 = 3m^2 + 2mp, \\ \alpha_1^2 + \alpha_2^2 + \alpha_3^2 &= m^2 + (m + p)^2 + (2m + p)^2 = 6m^2 + 6mp + 2p^2, \\ \alpha_2^2 + \alpha_3^2 + \alpha_4^2 &= (m + p)^2 + (2m + p)^2 + (3m + 2p)^2 = 14m^2 + 18mp + 6p^2. \end{aligned}$$

We note that the monomial p^2 cancels out in the difference $\alpha_3^2 - \alpha_2^2$, while it is present in the other two sums. For this reason, taking p large with respect to m gives a small

parameter, which we will use in our perturbation analysis (the other small parameter of the problem is the size of the solution, i.e., the size of the initial data of the Kirchhoff equation). Denoting

$$\sigma := \frac{m}{p},$$

one has

$$\begin{aligned} \mu_1 &= \frac{2\sigma + 3\sigma^2}{6 + 18\sigma + 14\sigma^2} = \sigma \left(\frac{1}{3} + \tilde{\mu}_1(\sigma) \right), & \gamma &= \frac{6 + 18\sigma + 14\sigma^2}{2 + 6\sigma + 6\sigma^2} = 3 + O(\sigma), \\ \mu_2 &= \mu_1 \gamma = \frac{2\sigma + 3\sigma^2}{2 + 6\sigma + 6\sigma^2} & \tilde{\mu}_1(\sigma), \tilde{\mu}_2(\sigma) &= O(\sigma) \quad \text{as } \sigma \rightarrow 0, \\ &= \sigma(1 + \tilde{\mu}_2(\sigma)), \end{aligned} \quad (3.25)$$

where $\tilde{\mu}_1(\sigma)$, $\tilde{\mu}_2(\sigma)$ are defined by the identities (3.25). We also note that $1 < \gamma < 3$ for all $\sigma > 0$.

Thus, for $\sigma = m/p$ small, the ‘‘coupling’’ terms $\mu_1\eta_2$ and $\mu_2\eta_1$ in system (3.20) can be considered as perturbations of the ‘‘unperturbed’’ system of two uncoupled pendulums

$$\begin{cases} \dot{\xi}_1 = \eta_1, & \dot{\xi}_2 = \eta_2, \\ \dot{\eta}_1 = -\sin(\xi_1), & \dot{\eta}_2 = -\lambda \sin(\xi_2). \end{cases} \quad (3.26)$$

We also want to normalize the coefficient λ appearing in the last equation of system (3.20). Later we will see that this normalization corresponds to a constraint on the initial data for equation (2.22); at this stage, however, we simply observe that the parameter λ in Lemma 3.6(ii) is not subject to any constraint. Thus, in the following analysis we fix $\lambda = 1$ and simply do not consider other values of that parameter. For $\lambda = 1$, system (3.20) becomes

$$\begin{cases} \dot{\xi}_1 = \eta_1 - \mu_1\eta_2, \\ \dot{\eta}_1 = -\sin(\xi_1), \\ \dot{\xi}_2 = \eta_2 - \mu_2\eta_1, \\ \dot{\eta}_2 = -\sin(\xi_2). \end{cases} \quad (3.27)$$

System (3.27) has a conserved quantity (obtained expressing the old Hamiltonian in terms of the new variables), which is

$$\mu_2 \left(\frac{1}{2} \eta_1^2 + 1 - \cos \xi_1 \right) + \mu_1 \left(\frac{1}{2} \eta_2^2 + 1 - \cos \xi_2 \right) - \mu_1 \mu_2 \eta_1 \eta_2. \quad (3.28)$$

The only parameters in system (3.27) are the constants μ_1 , μ_2 defined in (3.21), (3.22), which depend only on the ratio $\sigma = m/p$, and tend to zero as $\sigma \rightarrow 0$ (see (3.25)). The ‘‘unperturbed’’ part of system (3.27) is (3.26) with $\lambda = 1$, that is,

$$\begin{cases} \dot{\xi}_1 = \eta_1, & \dot{\xi}_2 = \eta_2, \\ \dot{\eta}_1 = -\sin(\xi_1), & \dot{\eta}_2 = -\sin(\xi_2). \end{cases} \quad (3.29)$$

System (3.29) is fully normalized and it is the two-dimensional Hamiltonian system of two uncoupled normalized pendulums

$$\dot{\xi}_n = \partial_{\eta_n} H, \quad \dot{\eta}_n = -\partial_{\xi_n} H, \quad n = 1, 2,$$

with Hamiltonian

$$H = H_1(\xi_1, \eta_1) + H_2(\xi_2, \eta_2), \quad H_n(\xi_n, \eta_n) = \frac{1}{2}\eta_n^2 + (1 - \cos(\xi_n)), \quad n = 1, 2. \quad (3.30)$$

The constant term 1 in the formula for H_n has been added just to give zero energy to the elliptic equilibrium.

It is convenient to consider system (3.27) as a perturbed double-pendulum system where the perturbative parameter is given by σ , instead of μ_1, μ_2 . Namely, using (3.25),

$$\begin{aligned} \dot{\xi}_1 &= \eta_1 - \sigma \left(\frac{1}{3} + \tilde{\mu}_1(\sigma) \right) \eta_2, & \dot{\xi}_2 &= \eta_2 - \sigma (1 + \tilde{\mu}_2(\sigma)) \eta_1, \\ \dot{\eta}_1 &= -\sin(\xi_1), & \dot{\eta}_2 &= -\sin(\xi_2). \end{aligned} \quad (3.31)$$

The conserved quantity in (3.28), divided by σ , is

$$\begin{aligned} \mathcal{E}(\xi_1, \eta_1, \xi_2, \eta_2) &:= \left(\frac{1}{3} + \tilde{\mu}_1(\sigma) \right) H_1(\xi_1, \eta_1) + (1 + \tilde{\mu}_2(\sigma)) H_2(\xi_2, \eta_2) \\ &\quad - \sigma \left(\frac{1}{3} + \tilde{\mu}_1(\sigma) \right) (1 + \tilde{\mu}_2(\sigma)) \eta_1 \eta_2. \end{aligned} \quad (3.32)$$

4. Chaos for two weakly coupled pendulums

In this section we prove the following result about chaotic solutions of system (3.31). We denote by $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ the set of nonnegative integers. Given an energy parameter $a \in (0, 2)$, we denote by $(\xi_1^*(t), \eta_1^*(t))$ the periodic solution of the pendulum satisfying

$$H_1(\xi_1^*(t), \eta_1^*(t)) = a, \quad \xi_1^*(0) = 0 \quad \eta_1^*(0) > 0. \quad (4.1)$$

To emphasize its dependence on a we will also denote $\xi_1^*(t) = \xi_1^*(t; a)$ and $\eta_1^*(t) = \eta_1^*(t; a)$.

Proposition 4.1. *There exists a universal constant $a_0 \in (0, 2)$ (see Lemma 4.6) such that the following holds. Let $a \in [a_0, 2)$ and let T_a be the period of $(\xi_1^*(t; a), \eta_1^*(t; a))$. There exist universal constants $\sigma_0 \in (0, 1)$, $C_1, C_2 > 0$ such that for every $\sigma \in (0, \sigma_0)$ there exists $M_0(\sigma)$ in the interval*

$$C_1 \log(\sigma^{-1}) \leq M_0(\sigma) \leq C_2 \log(\sigma^{-1})$$

such that the following properties hold. Let $(m_0, m_1, m_2, \dots) = (m_j)_{j \in \mathbb{N}_0}$ be any sequence of integers with $m_j \geq M_0(\sigma)$ for all $j \in \mathbb{N}$. Then there exists a solution $(\xi_1, \eta_1, \xi_2, \eta_2)(t)$ of system (3.31) such that the following holds:

(i) The function (ξ_1, η_1) satisfies

$$\sup_{t \in \mathbb{R}} |\xi_1(t) - \xi_1^*(t)| \leq C\sigma, \quad \sup_{t \in \mathbb{R}} |\eta_1(t) - \eta_1^*(t)| \leq C\sigma,$$

for some universal constant $C > 0$.

(ii) There exists a sequence of times $(t_0, t_1, t_2, \dots) = (t_j)_{j \in \mathbb{N}_0}$ with

$$t_0 = 0, \quad t_{j+1} = t_j + T_a(m_j + \theta_j), \quad 0 \leq \theta_j < 1, \quad \forall j \in \mathbb{N}_0 \quad (4.2)$$

such that

$$\eta_2(t_j) = 1 \quad \forall j \in \mathbb{N}_0.$$

Moreover, there exists another sequence of times $(\bar{t}_j)_{j \in \mathbb{N}_0}$ satisfying $t_j < \bar{t}_j < t_{j+1}$ such that

$$1 < \eta_2(t) \leq 2 + C\sigma \quad \forall t \in (t_j, \bar{t}_j), \quad -C\sigma \leq \eta_2(t) < 1 \quad \forall t \in (\bar{t}_j, t_{j+1}) \quad (4.3)$$

and

$$\max_{t \in [t_j, \bar{t}_j]} \eta_2(t) \geq 2 - C\sigma, \quad \min_{t \in [\bar{t}_j, t_{j+1}]} \eta_2(t) \leq C\sigma \quad (4.4)$$

for some universal constant $C > 0$.

We remark that $\eta_2 = 1$ is the value around which $\eta_2(t)$ is oscillating up and down with the randomly chosen sequence of times.

In order to prove Proposition 4.1 we will find a partially hyperbolic periodic orbit of the full system (3.31) and show that, for $\sigma > 0$ small enough, its stable and unstable invariant manifolds intersect transversally. This will imply the existence of a Smale horseshoe and the existence of symbolic chaotic dynamics.

Remark 4.2. The pendulum energy a in (4.1) is used as a free parameter only in this section. Proposition 4.1 will be applied in Sections 5 and 6 only for $a = a_0$. Since a_0 is a universal constant, for $a = a_0$ any quantity depending only on a becomes a universal constant.

4.1. Partially hyperbolic periodic orbit

We start by searching for the partially hyperbolic periodic solution. The unperturbed system ($\sigma = 0$) has plenty of partially hyperbolic periodic orbits, which are given for instance by the product of librations in the first pendulum (in the plane (ξ_1, η_1)) and the saddle of the second pendulum (in the plane (ξ_2, η_2)). We select one of these orbits and we apply an implicit function theorem argument to prove the existence of a nearby periodic orbit with the same period.

We consider the periodic orbit

$$\mathcal{P} = \mathcal{P}(t; a) := (\xi_1^*(t; a), \eta_1^*(t; a), \pi, 0) \quad (4.5)$$

of the unperturbed system (3.29) (which is system (3.31) with $\sigma = 0$), where (ξ_1^*, η_1^*) is defined in (4.1) and $a \in (0, 2)$ will be fixed at the end of Section 4.2. Note that the elliptic equilibrium $(\xi_1, \eta_1) = (0, 0)$ of the first pendulum has energy $H_1(0, 0) = 0$, and its saddle $(\xi_1, \eta_1) = (\pi, 0)$, as well as its homoclinic orbits, has energy $H_1(\pi, 0) = 2$. The solution \mathcal{P} is supported on the curve

$$\mathbb{T}_0 := \{(\xi_1, \eta_1, \xi_2, \eta_2) : H_1(\xi_1, \eta_1) = a, \xi_2 = \pi, \eta_2 = 0\}. \quad (4.6)$$

We denote by T the period T_a of the orbit \mathcal{P} and by $\omega := 2\pi/T$ its frequency. Of course \mathbb{T}_0 , T , ω depend on the energy parameter a ; in fact, all the quantities in the present subsection and in the next one (including, in particular, the smallness radius σ_1 given by Proposition 4.3) depend on a . Nonetheless, in general we do not indicate the dependence on a explicitly; we just underline that, after fixing a , every quantity appearing in Sections 4.1 and 4.2 will be determined, with no dependence on any other hidden parameter.

The unperturbed ($\sigma = 0$) homoclinic manifold of \mathcal{P} is

$$W_0(\mathcal{P}) = \{(\xi_1, \eta_1, \xi_2, \eta_2) : H_1(\xi_1, \eta_1) = a, H_2(\xi_2, \eta_2) = 2\}. \quad (4.7)$$

We consider its time parametrization

$$\Gamma_0^\pm := \{(\xi_1^*(\tau_1), \eta_1^*(\tau_1), q_h(\tau_2), p_h^\pm(\tau_2)) : \tau_1, \tau_2 \in \mathbb{R}\}, \quad (4.8)$$

where

$$(q_h(s), p_h^\pm(s)) = \left(2 \arcsin(\tanh(s)), \pm \frac{2}{\cosh(s)}\right).$$

Now we prove that the periodic orbit \mathcal{P} persists when $0 < \sigma \ll 1$. More precisely, we prove the following result.

Proposition 4.3. *Let \mathcal{P} be the T -periodic orbit of (3.29) defined in (4.5). Then there exist constants $\sigma_1 > 0$, $C > 0$ such that, for all $0 < \sigma < \sigma_1$, there exists a T -periodic solution $\mathcal{P}_\sigma(t)$ of (3.31) which is σ -close to \mathcal{P} in the C^1 -topology, namely*

$$\|\mathcal{P}_\sigma - \mathcal{P}\|_{C^1(\mathbb{R})} \leq C\sigma.$$

Moreover, \mathcal{P}_σ possesses one stable and one unstable hyperbolic direction.

Recall that system (3.31) has the energy (3.32) as first integral. Then the existence of a hyperbolic periodic orbit at each energy level is a consequence of classical perturbation theory. However, Proposition 4.3 gives the existence of a periodic orbit for a fixed period. This could be shown by proving that the period is monotone with \mathcal{E} . Below, to make this paper self-contained, we give an alternative proof of Proposition 4.3 based on a symmetry argument.

Proof of Proposition 4.3. To prove the persistence of the periodic orbit \mathcal{P} we use the fact that system (3.31) is reversible with respect to the involution

$$\rho: (\mathbb{T} \times \mathbb{R})^2 \rightarrow (\mathbb{T} \times \mathbb{R})^2, \quad \rho(\xi_1, \eta_1, \xi_2, \eta_2) = (-\xi_1, \eta_1, -\xi_2, \eta_2).$$

This means that, if we denote by X the vector field of (3.31), then $X \circ \rho = -\rho_* X$, where ρ_* is the differential of ρ , which acts on the tangent space \mathbb{R}^4 . To apply an implicit function theorem argument it is convenient to pass to action-angle coordinates on the first pendulum (plane (ξ_1, η_1)). This will simplify the analysis of the linearized problem in the tangential directions at the periodic orbit \mathcal{P} . In the domain

$$S := \{(x_1, y_1) \in \mathbb{T} \times \mathbb{R} : H_1(x_1, y_1) \in (0, 2)\},$$

we consider the action-angle variable transformation

$$\Phi: \mathbb{T} \times \mathcal{I} \rightarrow S, \quad (\xi_1, \eta_1) = \Phi(\theta, I) = (f(\theta, I), g(\theta, I))$$

for some open interval $\mathcal{I} \subset \mathbb{R}$. If we fix (ξ_1, η_1) and call $2\kappa = H_1(\xi_1, \eta_1)$, we can express f and g using elliptic functions in the following way:

$$\begin{cases} f(\theta, I) = 2 \arcsin\left(\sqrt{\kappa} \operatorname{sn}\left(\frac{2K(\kappa)}{\pi}\theta \mid \kappa\right)\right), \\ g(\theta, I) = 2\sqrt{\kappa} \operatorname{cn}\left(\frac{2K(\kappa)}{\pi}\theta \mid \kappa\right), \end{cases}$$

where $K(\kappa)$ is the complete elliptic integral of the first kind and sn and cn are the elliptic sine and the elliptic cosine respectively (see e.g., [17]).

We now drop the sub-index from ξ_2, η_2 . Let us denote

$$\Psi = (\Phi, \operatorname{Id}): \mathbb{T} \times \mathcal{I} \times \mathbb{T} \times \mathbb{R} \rightarrow S \times \mathbb{T} \times \mathbb{R}, \quad \Psi(\theta, I, \xi, \eta) = (\Phi(\theta, I), \xi, \eta).$$

The involution ρ expressed in these new coordinates is just given by $\nu = \Psi^{-1} \circ \rho \circ \Psi$. Since the elliptic sine is odd and the elliptic cosine is even (with respect to its first variable), it is straightforward to see that system (3.31) in the new coordinates is reversible with respect to the involution

$$\nu(\theta, I, \xi, \eta) = (-\theta, I, -\xi, \eta).$$

We denote by $I \mapsto \Omega(I)$ the action-to-frequency map of the unperturbed pendulum.

Remark 4.4. Note that $\Omega'(I) \neq 0$ for all $I \in \mathcal{I}$.

The unperturbed periodic orbit \mathcal{P} now reads

$$\theta(t) = \Omega(I_0)t, \quad I(t) = I_0, \quad \xi(t) = \pi, \quad \eta(t) = 0$$

for some $I_0 \in \mathcal{I}$. We remark that $\Omega(I_0) = \omega$, where $\omega = 2\pi/T$ was defined below (4.6). We consider the scaled time $t \rightsquigarrow \omega t$, and the system (3.31) becomes

$$\begin{cases} \omega \dot{\theta} = \Omega(I) + \sigma \mathcal{R}_1(\theta, I, \eta), \\ \omega \dot{I} = \sigma \mathcal{R}_2(\theta, I, \eta), \\ \omega \dot{\xi} = \eta + \sigma \mathcal{R}_3(\theta, I), \\ \omega \dot{\eta} = -\sin(\xi), \end{cases} \quad (4.9)$$

where the functions \mathcal{R}_i , $i = 1, 2, 3$, are determined by the relation

$$[D\Psi(\theta, I, \xi, \eta)]^{-1} \begin{pmatrix} -(\frac{1}{3} + \tilde{\mu}_1(\sigma))\eta \\ 0 \\ -(1 + \tilde{\mu}_2(\sigma))g(\theta, I) \\ 0 \end{pmatrix} = \begin{pmatrix} \mathcal{R}_1(\theta, I, \eta) \\ \mathcal{R}_2(\theta, I, \eta) \\ \mathcal{R}_3(\theta, I) \\ 0 \end{pmatrix}.$$

We look for 2π -periodic, smooth, reversible solutions of (4.9), namely $u(t) = (\theta(t), I(t), \xi(t), \eta(t))$ such that $\nu u(-t) = u(t)$. In other words, we look for solutions in the invariant set

$$C_{\text{odd, even}}^1(\mathbb{T}) := \{(\theta(t), I(t), \xi(t), \eta(t)) \in C^1(\mathbb{T}; \mathbb{T} \times \mathbb{I} \times \mathbb{T} \times \mathbb{R}) : \theta, \xi \text{ odd}, I, \eta \text{ even}\}.$$

Similarly, we define the space

$$C_{\text{even, odd}}^0(\mathbb{T}) := \{(g_1(t), g_2(t), g_3(t), g_4(t)) \in C^0(\mathbb{T}; \mathbb{R}^4) : g_1, g_3 \text{ even}, g_2, g_4 \text{ odd}\}.$$

Remark 4.5. We note that 2π -solutions of (4.9) correspond to T -periodic solutions of (3.31).

Let us define

$$\mathcal{F}(\sigma; \cdot): C_{\text{odd, even}}^1(\mathbb{T}) \rightarrow C_{\text{even, odd}}^0(\mathbb{T}), \quad \mathcal{F}(\sigma; \theta, I, \xi, \eta) = \begin{pmatrix} \omega\dot{\theta} - \Omega(I) - \sigma\mathcal{R}_1(\theta, I, \eta) \\ \omega\dot{I} - \sigma\mathcal{R}_2(\theta, I, \eta) \\ \omega\dot{\xi} - \eta - \sigma\mathcal{R}_3(\theta, I) \\ \omega\dot{\eta} + \sin(\xi) \end{pmatrix}.$$

Since the maps Ω , Ψ and the vector field X are analytic, we have that $\mathcal{F}(\sigma; \cdot)$ is at least C^1 . Then

$$\mathcal{F}(0; \omega t, I_0, \pi, 0) = 0.$$

We now study the linearized system at the unperturbed solution. We fix $(g_1, g_2, g_3, g_4) \in C_{\text{even, odd}}^0(\mathbb{T})$ and we look for solutions of the linear system

$$\begin{cases} \omega\partial_t\theta - \Omega'(I_0)I = g_1, \\ \omega\partial_t I = g_2, \\ \omega\partial_t\xi - \eta = g_3, \\ \omega\partial_t\eta - \xi = g_4. \end{cases} \quad (4.10)$$

We observe that the above system is decoupled, hence we can study the equations for (θ, I) and those for (ξ, η) separately. Concerning the former, we first solve the equation for the actions. Since g_2 is odd we have $\langle g_2 \rangle = 0$ and

$$I - \langle I \rangle = (\omega\partial_t)^{-1}g_2,$$

where we denote by $\langle \cdot \rangle$ the time average over \mathbb{T} and we denote by $(\omega \partial_t)^{-1} g_2$ the primitive of g_2 with zero average. Hence I is determined up to its average. Substituting in the equation for the angle we obtain

$$\omega \partial_t \theta = \Omega'(I_0)(\omega \partial_t)^{-1} g_2 + \Omega'(I_0) \langle I \rangle + g_1. \quad (4.11)$$

Equation (4.11) can be solved only if the right-hand side has zero average. Therefore we fix

$$\langle I \rangle = -\frac{\langle g_1 \rangle}{\Omega'(I_0)}, \quad \theta = (\omega \partial_t)^{-1} [\Omega'(I_0)(\omega \partial_t)^{-1} g_2 + \Omega'(I_0) \langle I \rangle + g_1]$$

(note that $\Omega'(I_0) \neq 0$; see Remark 4.4). Concerning the equations for the (ξ, η) -variables we have the following: by setting $v := \xi + \eta$, $w := \xi - \eta$, $h := g_3 + g_4$, and $\tilde{h} := g_3 - g_4$ we have that

$$\omega \partial_t v - v = h, \quad \omega \partial_t w + w = \tilde{h}. \quad (4.12)$$

By Fourier series, one has that

$$v(t) = \sum_{k \in \mathbb{Z}} \frac{h_k}{i\omega k - 1} e^{ikt}, \quad w(t) = \sum_{k \in \mathbb{Z}} \frac{\tilde{h}_k}{i\omega k + 1} e^{ikt}$$

are the unique solutions of (4.12). Since $\xi = (v + w)/2$ and $\eta = (v - w)/2$, we recover the ξ, η components of the solution of (4.10). From the explicit expression of the solutions we have that for $g_i \equiv 0$, $i = 1, \dots, 4$, the only solution of (4.10) is zero. Moreover, the solutions are of class C^1 . This implies that $d\mathcal{F}(0; \omega t, I_0, \pi, 0)$ is invertible and by the implicit function theorem there exists $\sigma_1 > 0$ and a C^1 function $g: (-\sigma_1, \sigma_1) \rightarrow C^1_{\text{odd, even}}(\mathbb{T})$ such that

$$\mathcal{F}(\sigma; g(\sigma)) = 0 \quad \forall \sigma \in (-\sigma_1, \sigma_1), \quad g(0; t) = (\omega t, I_0, \pi, 0).$$

We call \mathcal{P}_σ the periodic orbit $g(\sigma)$ written in the original coordinates $(\xi_1, \eta_1, \xi_2, \eta_2)$. We also notice that by (4.12) this orbit is hyperbolic in the (ξ_2, η_2) directions. This concludes the proof. \blacksquare

By the classical theory of persistence of invariant manifolds, \mathcal{P}_σ has stable and unstable invariant manifolds $W_\sigma^{s,u}(\mathcal{P}_\sigma)$ that depend differentiably on σ . Moreover, these manifolds can be locally parametrized as C^1 graphs over the unperturbed invariant manifold (4.7).

4.2. Transverse intersection of invariant manifolds

In this section we prove that the stable and unstable invariant manifolds $W_\sigma^{s,u}(\mathcal{P}_\sigma)$ intersect transversally at some point. Since the invariant manifolds have dimension 2 and we look for intersections within a three-dimensional energy level (see (3.32)), it is sufficient to construct a one-dimensional section Λ and measure the distance between the manifolds on the projection of Λ .

We recall the time parametrization of the unperturbed separatrix Γ_0^\pm given in (4.8). By symmetry we can consider just a single branch of the unperturbed homoclinic manifold, say $\Gamma_0 = \Gamma_0^+$. Let us consider a point $z_0 = z_0(\tau_1, \tau_2) \in \Gamma_0 \subset \{H_1 = a, H_2 = 2\}$ and define the section

$$\Lambda := \{z_0 + \lambda \nabla H_1(z) : \lambda \in \mathbb{R}\},$$

where we use the notation $H_1(\xi_1, \eta_1, \xi_2, \eta_2) := H_1(\xi_1, \eta_1)$. The line Λ passes through z_0 and it is normal to Γ_0 .

By the continuous dependence of the invariant manifolds on the parameters, for σ small enough, Λ also transversally intersects $W_\sigma^{s,u}(\mathcal{P}_\sigma)$ at two points $z_\sigma^{s,u} = z_\sigma^{s,u}(\tau_1, \tau_2)$. We use the unperturbed energy of the first pendulum H_1 to measure the distance between z_σ^s and z_σ^u . Note that the gradient of H_1 never vanishes on Γ_0 , hence it is a good measure of a displacement in the normal directions of Γ_0 . We define the distance

$$H_1(z_\sigma^s) - H_1(z_\sigma^u) = \sigma M + O(\sigma^2), \quad (4.13)$$

where the first order of this distance is given by

$$M = M(\tau_1, \tau_2) := \frac{d}{d\sigma} \Big|_{\sigma=0} (H_1(z_\sigma^s) - H_1(z_\sigma^u)).$$

By classical arguments (see for instance [30]) we have that the first order is given by the Melnikov integral

$$M(\tau_1, \tau_2) = \int_{-\infty}^{+\infty} DH_1(\Phi_H^t(\Gamma_0(\tau_1, \tau_2))) [Y(\Phi_H^t(\Gamma_0(\tau_1, \tau_2)))] dt,$$

where Y is the first order in σ of the perturbation of system (3.31), namely

$$Y(\xi_1, \eta_1, \xi_2, \eta_2) := (-\eta_2/3, 0, -\eta_1, 0),$$

and Φ_H^t is the Hamiltonian flow of H in (3.30). We observe that (recall (4.8))

$$\Phi_H^t(\Gamma_0(\tau_1, \tau_2)) = (\xi_1^*(\tau_1 + t), \eta_1^*(\tau_1 + t), q_h(\tau_2 + t), p_h^+(\tau_2 + t)).$$

By the autonomous nature of system (3.31), the Melnikov integral depends just on one parameter. We define $\tau := \tau_2 - \tau_1$ and we consider the reduced Melnikov integral

$$\mathcal{M}(\tau) := -\frac{1}{3} \int_{-\infty}^{+\infty} p_h^+(\tau + s) \sin(\xi_1^*(s)) ds.$$

Note that, for all $\tau_1 \in \mathbb{R}$, one has $\mathcal{M}(\tau) = M(\tau_1, \tau_1 + \tau) = M(0, \tau)$. We now prove the following.

Lemma 4.6. *There exists a universal constant $a_0 \in (0, 2)$ such that, for all $a \in [a_0, 2)$, the reduced Melnikov integral $\mathcal{M}(\tau)$ has a nondegenerate zero at $\tau = 0$.*

Proof. We observe that $\mathcal{M}(0) = 0$ because $\xi_1^*(t)$ is an odd function, while $p_h^+(t)$ is even. The derivative of the reduced Melnikov integral at $\tau = 0$ is

$$\frac{d}{d\tau}\mathcal{M}(0) = \frac{1}{3} \int_{-\infty}^{+\infty} \sin(q_h(s)) \sin(\xi_1^*(s)) ds = \frac{2}{3} \int_0^{+\infty} \sin(q_h(s)) \sin(\xi_1^*(s)) ds.$$

We recall that (ξ_1^*, η_1^*) in (4.5) depends on the parameter a , and we explicitly indicate the dependence on a of the integral we want to study, denoting

$$J(a) := \int_0^{+\infty} \sin(q_h(s)) \sin(\xi_1^*(s; a)) ds.$$

We have to prove that $J(a)$ is nonzero for some value of $a \in (0, 2)$. By the classical theorem of continuous dependence on initial data for ODEs, one has the following pointwise convergence:

$$\text{for every } s \in [0, \infty), \quad \lim_{a \rightarrow 2} \xi_1^*(s; a) = q_h(s)$$

(furthermore, the convergence is uniform on compact intervals). Hence, for every $s \in [0, \infty)$, $f_a(s)$ converges to $f_2(s)$ as $a \rightarrow 2$, where

$$f_a(s) := \sin(q_h(s)) \sin(\xi_1^*(s; a)), \quad f_2(s) := \sin^2(q_h(s)).$$

Moreover, since $q_h(s) \in [0, \pi)$ for all $s \in [0, \infty)$, one has

$$|f_a(s)| \leq |\sin(q_h(s))| = \sin(q_h(s)) = \frac{2 \sinh(s)}{\cosh^2(s)} =: g(s) \quad \forall s \in [0, \infty),$$

and $g \in L^1(0, \infty)$. Hence, by the dominated convergence theorem,

$$J_* := \lim_{a \rightarrow 2} J(a) = \int_0^{\infty} f_2(s) ds = \int_0^{\infty} g^2(s) ds.$$

The limit J_* is finite by the exponential decay of g^2 , and it is positive because g^2 is positive. Hence there exists $a_0 \in (0, 2)$ such that $|J(a) - J_*| \leq J_*/2$ for all $a \in [a_0, 2)$, and the lemma is proved. \blacksquare

By the above lemma and the implicit function theorem, for $\sigma > 0$ small enough there exists at least one zero of the distance (4.13), with transverse intersection.

4.3. Symbolic dynamics

We introduce the section

$$\Pi := \{(\xi_1, \eta_1, \xi_2, \eta_2) : \xi_2 = 0, \eta_2 > 0, \mathcal{E}(\xi_1, \eta_1, \xi_2, \eta_2) = \mathcal{E}(\mathcal{P}_\sigma)\},$$

where $\mathcal{E}(\mathcal{P}_\sigma)$ is the value of the prime integral \mathcal{E} in (3.32) at the solution \mathcal{P}_σ . This section is transverse to the unperturbed flow ($\sigma = 0$) at a point of $W_0(\mathcal{P})$ and so, for $\sigma > 0$

small enough, also to the perturbed one. Moreover, by Lemma 4.6, it contains points of $W_\sigma^s(\mathcal{P}_\sigma) \pitchfork W_\sigma^u(\mathcal{P}_\sigma)$ (where \pitchfork means transverse intersection).

Denote by Φ^t the flow of system (3.31). Fixing a point $z \in \Pi$, we define $\mathcal{T}(z) > 0$ as the first (forward) return time to Π . For those points z that do not hit back the section Π (for instance, the points of $W_\sigma^s(\mathcal{P}_\sigma)$), we set $\mathcal{T}(z) = +\infty$. We define the open set $\mathcal{U} \subset \Pi$ as

$$\mathcal{U} := \{z \in \Pi : \mathcal{T}(z) < +\infty\}$$

and the associated Poincaré map $P: \mathcal{U} \subset \Pi \rightarrow \Pi$ by $P(z) = \Phi^{\mathcal{T}(z)}|_{t=\mathcal{T}(z)}$.

Proposition 4.7 (Smale horseshoe). *There exist universal constants $\sigma_0, C_1, C_2 > 0$ such that for all $\sigma \in (0, \sigma_0)$ there exists a positive integer $M_0(\sigma)$ in the interval*

$$C_1 \log(\sigma^{-1}) \leq M_0(\sigma) \leq C_2 \log(\sigma^{-1}) \quad (4.14)$$

such that the Poincaré map P possesses an invariant set $Y \subset \mathcal{U}$ whose dynamics is conjugated to the infinite symbols shift. Namely, there exists a homeomorphism $h: \mathcal{A} \rightarrow Y$, where

$$\mathcal{A} := \{\omega = \{\omega_k\}_{k \in \mathbb{Z}} : \omega_k \in \mathbb{N}, \omega_k \geq M_0(\sigma) \forall k \in \mathbb{Z}\},$$

such that $P|_Y = h \circ \mathfrak{d} \circ h^{-1}$, where $\mathfrak{d}: \mathcal{A} \rightarrow \mathcal{A}$ is the shift

$$(\mathfrak{d}\omega)_k = \omega_{k+1}, \quad k \in \mathbb{Z}.$$

Moreover, h^{-1} can be defined as follows. Associated to $z \in Y$ one can define the sequence of hitting times

$$t_0 = 0, \quad t_k = t_{k-1} + \mathcal{T}(P^{k-1}z) \text{ for } k \geq 1, \quad t_k = t_{k+1} - \mathcal{T}(P^k z) \text{ for } k \leq -1,$$

and $h^{-1}(z) := \omega = (\omega_k)_{k \in \mathbb{Z}}$, with

$$\omega_k = \left\lfloor \frac{t_k - t_{k-1}}{T} \right\rfloor,$$

where $T = T_a$ is the period of the periodic orbit \mathcal{P}_σ , and $\lfloor \cdot \rfloor$ denotes the integer part.

Proof. The proof follows the same lines as the construction of symbolic dynamics done by Moser in [29, Chapter 3]. ■

Proposition 4.7 concludes the proof of Proposition 4.1. Note that the return times to the section are large since orbits get close to the hyperbolic periodic orbit.

5. Back to the truncated effective system

In Section 4 we proved the existence of chaotic solutions for system (3.31). These solutions are global in time. In the next lemma we obtain the corresponding solutions of the truncated effective system (3.3).

Lemma 5.1. *Let $a = a_0$ in Proposition 4.1, and let σ_0 be the corresponding universal constant given by Proposition 4.1. Let m, p be two integers, $2 \leq m < p$, with ratio $\sigma := m/p$ in the interval $(0, \sigma_0)$. Define $\alpha_1, \dots, \alpha_4$ by (3.2), $\tilde{\mu}_1(\sigma), \tilde{\mu}_2(\sigma)$ by (3.25), and γ by the second identity in (3.19). Assume the hypotheses of Proposition 4.1, and consider the solution $(\xi_1(t), \eta_1(t), \xi_2(t), \eta_2(t))$ of system (3.31) obtained in Proposition 4.1.*

Consider any three real numbers A_1, q_1, q_2 , with $A_1 > 0$, and define the following constants: define c_{123} by (3.24), define $c_{234} = c_{123}/\gamma$, define B by the second identity in (3.17), define A_2 by (3.18), define b_1, b_2 by (3.15), define E_1, E_2 by (3.11). Define the functions

$$\begin{aligned} S_1(t) &:= E_1 - q_1 - A_1 \eta_1(Bt), \\ S_2(t) &:= E_2 - q_1 - q_2 - A_1 \eta_1(Bt) - A_2 \eta_2(Bt), \\ S_3(t) &:= q_1 - q_2 + A_1 \eta_1(Bt) - A_2 \eta_2(Bt), \\ S_4(t) &:= q_2 + A_2 \eta_2(Bt), \\ \varphi_{123}(t) &:= \xi_1(Bt), \\ \varphi_{234}(t) &:= \xi_2(Bt). \end{aligned} \tag{5.1}$$

Then $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t), c_{123}, c_{234})$ satisfies (3.3) for all $t \in \mathbb{R}$.

Proof. We apply Lemma 3.6 (ii) with $\lambda = 1$ to go back from system (3.31) to system (3.14), then Lemma 3.5 (ii) to go back from system (3.14) to system (3.10), then Lemma 3.4 (ii) to go back from system (3.10) to system (3.8), and finally Lemma 3.3 (ii) to go back from system (3.8) to system (3.3). ■

5.1. Positivity of the superactions

Now we come to the question of whether the solutions of system (3.3) obtained in Lemma 5.1 satisfy the inequalities (3.4). As a first step, we study the constant terms

$$E_1 - q_1, \quad E_2 - q_1 - q_2, \quad q_1 - q_2, \quad q_2 \tag{5.2}$$

appearing in the definition of S_1, \dots, S_4 in Lemma 5.1. We compute the formulas for E_1, E_2 as functions of $q_1 R, Rq_2$: from (3.11) and (3.15) we get

$$E_1 = \frac{\alpha_1^2 + 2\alpha_3^2}{\alpha_1^2} q_1 - \frac{2\alpha_3^2 + \alpha_4^2}{\alpha_1^2} q_2, \quad E_2 = -\frac{\alpha_3^2 - \alpha_2^2}{\alpha_2^2} q_1 + \frac{\alpha_2^2 + \alpha_3^2 + \alpha_4^2}{\alpha_2^2} q_2. \tag{5.3}$$

Next we observe in the following lemma that q_1, q_2 can be chosen such that the constant terms (5.2) are all positive.

Lemma 5.2. *Let q_1, q_2 be positive real numbers with ratio q_1/q_2 in the interval $(1 + \frac{r}{2}, 1 + r)$, $r = \alpha_4^2/\alpha_3^2$, and let E_1, E_2 be defined by (5.3). Then the constants (5.2) are all positive. In particular, if*

$$q_1 = \left(1 + \frac{2}{3}r\right)q_2, \tag{5.4}$$

then

$$E_1 - q_1 = \frac{1}{3} \frac{\alpha_4^2}{\alpha_1^2} q_2, \quad E_2 - q_1 - q_2 = \frac{1}{3} \frac{\alpha_4^2}{\alpha_2^2} q_2, \quad q_1 - q_2 = \frac{2}{3} \frac{\alpha_4^2}{\alpha_3^2} q_2. \quad (5.5)$$

Proof. By (5.3),

$$\begin{aligned} E_1 - q_1 &= \left(\frac{2\alpha_3^2}{\alpha_1^2} \frac{q_1}{q_2} - \frac{2\alpha_3^2 + \alpha_4^2}{\alpha_1^2} \right) q_2, \\ E_2 - q_1 - q_2 &= \left(-\frac{\alpha_3^2}{\alpha_2^2} \frac{q_1}{q_2} + \frac{\alpha_3^2 + \alpha_4^2}{\alpha_2^2} \right) q_2, \\ q_1 - q_2 &= \left(\frac{q_1}{q_2} - 1 \right) q_2. \end{aligned}$$

We write the ratio q_1/q_2 as $1 + \vartheta r$, where $\vartheta \in \mathbb{R}$ is a free parameter and $r = \alpha_4^2/\alpha_3^2$. Then

$$E_1 - q_1 = \frac{\alpha_4^2 q_2}{\alpha_1^2} (2\vartheta - 1), \quad E_2 - q_1 - q_2 = \frac{\alpha_4^2 q_2}{\alpha_2^2} (1 - \vartheta), \quad q_1 - q_2 = \frac{\alpha_4^2 q_2}{\alpha_3^2} \vartheta.$$

The minimum

$$\min\{2\vartheta - 1, 1 - \vartheta, \vartheta, 1\} = \min\{2\vartheta - 1, 1 - \vartheta\}$$

is positive for $\vartheta \in (\frac{1}{2}, 1)$, and it reaches its maximum value at $\vartheta = 2/3$. \blacksquare

Note that, by (3.2), the ratio $r = \alpha_4^2/\alpha_3^2$ tends to 4 as $\sigma = m/p \rightarrow 0$. By (3.19), (3.25), the constant A_2 in Lemma 5.1 satisfies

$$A_2 = \frac{A_1}{\gamma} = \frac{1 + 3\sigma + 3\sigma^2}{3 + 9\sigma + 7\sigma^2} A_1 \leq A_1. \quad (5.6)$$

The solutions η_1, η_2 , constructed in Proposition 4.1 and appearing in Lemma 5.1, satisfy

$$\sup_{t \in \mathbb{R}} |\eta_1(t)| \leq 3, \quad \sup_{t \in \mathbb{R}} |\eta_2(t)| \leq 3 \quad (5.7)$$

(more accurate estimates about η_1, η_2 have been obtained in Proposition 4.1). Thus, we prove the following bound for S_n from below.

Lemma 5.3. *Let q_2, A_1 be any two positive real numbers, and define q_1 by (5.4). If*

$$A_1 \leq \frac{q_2}{9}, \quad (5.8)$$

then the functions $S_n(t)$ defined in Lemma 5.1 satisfy

$$S_n(t) \geq \frac{q_2}{2} > 0 \quad \forall t \in \mathbb{R}, \quad \forall n = 1, 2, 3, 4. \quad (5.9)$$

Proof. By (5.5) and (5.6), the functions $S_n(t)$ defined in Lemma 5.1 satisfy for all $t \in \mathbb{R}$,

$$\begin{aligned} S_1(t) &\geq E_1 - q_1 - 3A_1 = \frac{1}{3} \frac{\alpha_4^2}{\alpha_1^2} q_2 - 3A_1, \\ S_2(t) &\geq E_2 - q_1 - q_2 - 3A_1 - 3A_2 \geq \frac{1}{3} \frac{\alpha_4^2}{\alpha_2^2} q_2 - 6A_1, \\ S_3(t) &\geq q_1 - q_2 - 3A_1 - 3A_2 \geq \frac{2}{3} \frac{\alpha_4^2}{\alpha_3^2} q_2 - 6A_1, \\ S_4(t) &\geq q_2 - 3A_2 \geq q_2 - 3A_1. \end{aligned}$$

By (3.2), one has

$$\frac{\alpha_4^2}{\alpha_1^2} > 4, \quad \frac{\alpha_4^2}{\alpha_2^2} > 4, \quad \frac{2\alpha_4^2}{\alpha_3^2} > 4,$$

which, together with (5.8), implies (5.9). \blacksquare

By Lemma 5.3, condition (5.8) implies the meaningfulness condition (3.4) for the solutions of system (3.3). By Lemma 5.3, now we have two free parameters, which are q_2 and A_1 , related by the inequality $0 < A_1 \leq q_2/9$, while q_1 is now constrained by formula (5.4).

5.2. Initial data in the normal form ball

The special solutions of system (3.3) defined in Lemma 5.1 will be compared, by a Grönwall argument, with those of the full (i.e., nontruncated) effective system (2.17), (2.18) starting at the same initial data at time $t = 0$. The initial data we are interested in correspond to functions $u_0(x)$ in the ball (2.8), because every u_0 in that ball produces a solution of the Cauchy problem (2.2), (2.9) that remains, for a sufficiently long interval of time, in the domain where the normal form transformation is well defined, as is explained quantitatively in Lemma 2.2. Recall that $\delta_1 > 0$ in (2.8) is a universal constant, and m_1 is defined in (2.6). The following lemma deals with the ball (2.8) written in terms of S_n .

Lemma 5.4. *Consider the solutions of system (3.3) given by Lemma 5.1, and assume (5.4), (5.8). Then*

$$\sum_{n=1}^4 \alpha_n^{2s} S_n(t) \leq 8\alpha_4^{2s} q_2 \quad \forall t \in \mathbb{R}, \quad s \in [1, \infty). \quad (5.10)$$

Proof. For all $t \in \mathbb{R}$, if the smallness condition (5.8) is satisfied, then the solutions of system (3.3) obtained in Lemma 5.1 satisfy

$$\begin{aligned} \sum_{n=1}^4 \alpha_n^{2s} S_n &\leq \alpha_1^{2s} (E_1 - q_1 + 3A_1) + \alpha_2^{2s} (E_2 - q_1 - q_2 + 6A_1) + \alpha_3^{2s} (q_1 - q_2 + 6A_1) \\ &\quad + \alpha_4^{2s} (q_2 + 3A_1) \end{aligned}$$

$$\begin{aligned} &\leq \alpha_1^{2s} \frac{2\alpha_4^2 q_2}{\alpha_1^2} + \alpha_2^{2s} \frac{2\alpha_4^2 q_2}{\alpha_2^2} + \alpha_3^{2s} \frac{2\alpha_4^2 q_2}{\alpha_3^2} + \alpha_4^{2s} 2q_2 \\ &\leq 8\alpha_4^{2s} q_2, \end{aligned}$$

where we have used the bounds in (5.7) for η_1, η_2 , the identities (5.5) for the constant terms (5.2), the bound (5.8) for A_1 , the bound (5.6) for A_2 , and the fact that $s \geq 1, \alpha_1 < \dots < \alpha_4$. \blacksquare

A consequence of this lemma is the following. If one chooses q_2 such that

$$8\alpha_4^{2m_1} q_2 \leq \delta_1^2, \quad (5.11)$$

where m_1, δ_1 are the universal constants of the ball (2.8), then

$$\sum_{n=1}^4 \alpha_n^{2m_1} S_n(t) \leq \delta^2 \quad \forall t \in \mathbb{R}. \quad (5.12)$$

5.3. Construction of a compatible initial datum

Given a trigonometric polynomial $u \in C(\mathbb{T}^d, \mathbb{C})$, Fourier supported on the set $\{k \in \mathbb{Z}^d : |k| \in \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}\}$, with Fourier coefficients u_k , we use the superscript u to denote

$$S_n^{(u)} := \sum_{|k|=\alpha_n} |u_k|^2, \quad B_n^{(u)} := \sum_{|k|=\alpha_n} u_k u_{-k}, \quad n = 1, 2, 3, 4, \quad (5.13)$$

$$Z_{123}^{(u)} := B_1^{(u)} B_2^{(u)} \overline{B_3^{(u)}}, \quad \rho_{123}^{(u)} := |Z_{123}^{(u)}|, \quad c_{123}^{(u)} := \frac{3}{8} \rho_{123}^{(u)} \alpha_1 \alpha_2 \alpha_3, \quad (5.14)$$

analogous definitions for $Z_{234}^{(u)}, \rho_{234}^{(u)}, c_{234}^{(u)}$, and, if $\rho_{123}^{(u)}, \rho_{234}^{(u)}$ are positive, we define the angles $\varphi_{123}^{(u)}, \varphi_{234}^{(u)} \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ by the identities

$$Z_{123}^{(u)} = \rho_{123}^{(u)} \exp(i\varphi_{123}^{(u)}), \quad Z_{234}^{(u)} = \rho_{234}^{(u)} \exp(i\varphi_{234}^{(u)}). \quad (5.15)$$

We consider the following question:

Given a solution $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ of system (3.3) obtained in Lemma 5.1, and taking, in particular, its value at time $t = 0$, does there exist a function $u_0(x)$ in the ball (2.8), Fourier supported on the spheres of radius $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, such that

$$\begin{cases} S_n^{(u_0)} = S_n(0), & n = 1, 2, 3, 4, \\ \varphi_{123}^{(u_0)} = \varphi_{123}(0), & \varphi_{234}^{(u_0)} = \varphi_{234}(0), \\ c_{123}^{(u_0)} = c_{123}, & c_{234}^{(u_0)} = c_{234}? \end{cases} \quad (5.16)$$

The equations for the angles must be interpreted as identities of elements of $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. The affirmative answer to this question is given in Lemma 5.8 below, whose proof uses the next three simple, preparatory lemmas.

Lemma 5.5. *Let $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ be a solution of system (3.3) obtained in Lemma 5.1. Assume that q_1, q_2, A_1 satisfy (5.4), (5.8). If, in addition, A_1, q_2 satisfy*

$$A_1 \leq \left(\frac{3}{32} \frac{\alpha_1^3}{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} \right)^{\frac{1}{2}} q_2^{\frac{3}{2}}, \quad (5.17)$$

then there exist real numbers $r_n, \psi_n, n = 1, 2, 3, 4$, with

$$0 < r_n \leq S_n(0), \quad (5.18)$$

such that

$$\psi_1 + \psi_2 - \psi_3 = \varphi_{123}(0), \quad \frac{3}{8}(r_1\alpha_1)(r_2\alpha_2)(r_3\alpha_3) = c_{123}, \quad (5.19)$$

$$\psi_2 + \psi_3 - \psi_4 = \varphi_{234}(0), \quad \frac{3}{8}(r_2\alpha_2)(r_3\alpha_3)(r_4\alpha_4) = c_{234}. \quad (5.20)$$

Proof. Regarding $(\psi_1, \psi_2, \psi_3, \psi_4)$, there are infinitely many solutions, because they are four unknowns that have to satisfy just two linear constraints. For example, we can fix $(\psi_1, \psi_2, \psi_3, \psi_4) = (\varphi_{123}(0), 0, 0, -\varphi_{234}(0))$. Regarding r_n , we first recall that, from Lemma 5.1, the constants c_{123}, c_{234} are

$$c_{123} = \frac{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}{2} A_1^2, \quad c_{234} = \frac{c_{123}}{\gamma},$$

where γ is defined in (3.19) and satisfies $1 < \gamma < 3$. Therefore c_{123}, c_{234} are positive, because A_1 is positive. We have to choose r_n such that

$$\begin{aligned} (\alpha_1 r_1)(\alpha_2 r_2)(\alpha_3 r_3) &= \frac{8}{3} c_{123} = \frac{4}{3} (\alpha_1^2 + \alpha_2^2 + \alpha_3^2) A_1^2, \\ (\alpha_2 r_2)(\alpha_3 r_3)(\alpha_4 r_4) &= \frac{8}{3} \frac{c_{123}}{\gamma} = \frac{4}{3} (\alpha_1^2 + \alpha_2^2 + \alpha_3^2) A_1^2 \frac{1}{\gamma}. \end{aligned}$$

Hence we fix

$$r_1 = \frac{r_0}{\alpha_1}, \quad r_2 = \frac{r_0}{\alpha_2}, \quad r_3 = \frac{r_0}{\alpha_3}, \quad r_4 = \frac{r_0}{\gamma \alpha_4}, \quad r_0 := \left(\frac{4}{3} (\alpha_1^2 + \alpha_2^2 + \alpha_3^2) A_1^2 \right)^{\frac{1}{3}}.$$

Thus, r_n are all positive and satisfy the required identities. It only remains to check that $r_n \leq S_n(0)$. By (5.9), we know that $q_2/2 \leq S_n(0)$ for all $n = 1, 2, 3, 4$. Since $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$ and $\gamma > 1$, we have $r_4 < r_3 < r_2 < r_1$. Hence it is sufficient to check that $r_1 \leq q_2/2$, and, by the definition of $r_1 = r_0/\alpha_1$, this holds if A_1, q_2 satisfy (5.17). \blacksquare

Lemma 5.6. *Let s, r, ψ be real numbers such that $0 < r \leq s$. Then there exists $z_1, z_2 \in \mathbb{C} \setminus \{0\}$ such that $|z_1|^2 + |z_2|^2 = s$ and $2z_1 z_2 = r e^{i\psi}$.*

Lemma 5.7. *Let $s_n, r_n, \psi_n, n = 1, 2, 3, 4$, be real numbers such that $0 < r_n \leq s_n$. Then, in any dimension $d \geq 1$, there exists a trigonometric polynomial $u_0 \in C(\mathbb{T}^d, \mathbb{C})$, Fourier supported on the set $\{k \in \mathbb{Z}^d : |k| \in \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}\}$, such that, recalling the notation (5.13),*

$$S_n^{(u_0)} = s_n, \quad B_n^{(u_0)} = r_n e^{i\psi_n}, \quad n = 1, 2, 3, 4. \quad (5.21)$$

Proof. For each $n = 1, 2, 3, 4$, fix an integer vector $k_n \in \mathbb{Z}^d$ with $|k_n| = \alpha_n$, and apply Lemma 5.6 to determine two nonzero complex numbers $z_{1,n}, z_{2,n}$ such that

$$|z_{1,n}|^2 + |z_{2,n}|^2 = s_n, \quad 2z_{1,n}z_{2,n} = r_n e^{i\psi_n}.$$

We define u_0 as the trigonometric polynomial having $z_{1,n}, z_{2,n}$ as Fourier coefficients for the frequencies $k_n, -k_n$, and having no other frequencies in its support, i.e.,

$$u_0(x) = \sum_{n=1}^4 (z_{1,n} e^{ik_n \cdot x} + z_{2,n} e^{-ik_n \cdot x}).$$

Then

$$S_n^{(u_0)} = |z_{1,n}|^2 + |z_{2,n}|^2 = s_n, \quad B_n^{(u_0)} = 2z_{1,n}z_{2,n} = r_n e^{i\psi_n}. \quad \blacksquare$$

Lemma 5.7 deals with trigonometric polynomials supported on just one pair $(k_n, -k_n)$ of points on the sphere $\{k : |k| = \alpha_n\}$; this is the minimal situation, valid in any dimension $d \geq 1$. Of course, in dimension $d \geq 2$ the Fourier support can contain more than one pair of opposite frequencies on the same sphere, and therefore the construction of u_0 with prescribed $S_n^{(u_0)}, B_n^{(u_0)}$ has even more free parameters at its disposal.

The following lemma gives the answer to question (5.16).

Lemma 5.8. *Let $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ be a solution of system (3.3) obtained in Lemma 5.1. Also assume that q_1, q_2, A_1 satisfy (5.4), (5.8), (5.17). Then there exists a trigonometric polynomial $u_0 \in C(\mathbb{T}^d, \mathbb{C})$, Fourier supported on the set $\{k \in \mathbb{Z}^d : |k| \in \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}\}$, satisfying all the identities in (5.16). If, in addition, q_2 satisfies (5.11), then u_0 belongs to the ball (2.8).*

Proof. By Lemma 5.5, there exist constants r_n, ψ_n satisfying (5.18), (5.19), (5.20). Define $s_n := S_n(0)$, so that (5.18) becomes $0 < r_n \leq s_n$. By Lemma 5.7, there exists a trigonometric polynomial u_0 , with the desired Fourier support, satisfying (5.21). By the first identity in (5.21) we directly have

$$S_n^{(u_0)} = s_n = S_n(0), \quad n = 1, 2, 3, 4.$$

By the second identity in (5.21), the first definition in (5.14), and the first identity in (5.19), we obtain

$$Z_{123}^{(u_0)} = r_1 r_2 r_3 \exp(i(\psi_1 + \psi_2 - \psi_3)) = r_1 r_2 r_3 \exp(i\varphi_{123}(0)). \quad (5.22)$$

By (5.22), by the second and third definitions in (5.14), and by the second identity in (5.19), we get

$$\rho_{123}^{(u_0)} = r_1 r_2 r_3, \quad c_{123}^{(u_0)} = \frac{3}{8} r_1 r_2 r_3 \alpha_1 \alpha_2 \alpha_3 = c_{123}.$$

Moreover, since $r_1 r_2 r_3 > 0$, (5.22) is a polar representation of $Z_{123}^{(u_0)}$, and hence $\varphi_{123}^{(u_0)} = \varphi_{123}(0)$ as elements of $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. A similar proof applies for $c_{234}, \varphi_{234}(0)$.

Finally, if q_2 satisfies (5.11), then, by Lemma 5.4, bound (5.12) holds; in particular, this bound at time $t = 0$ implies that u_0 belongs to the ball (2.8). \blacksquare

5.4. Joining the two amplitude parameters

In Lemma 5.1 the solutions of system (3.3) obtained from Proposition 4.1 are described by the three independent parameters q_1, q_2, A_1 . Then, to get the positivity of the functions S_n , it is enough to use that the ratio q_1/q_2 is bounded from below and from above by (5.4). After that, only two independent parameters remain, which are q_2 and A_1 . Then, to obtain the lower bound (5.9), we need (5.8), which is a bound of the form $A_1 \leq C q_2$ for some universal constant C . Also, q_2 itself must satisfy the smallness condition (5.11), which is an inequality of the form $q_2 \leq K$, for some constant K depending on m, p . Next, A_1, q_2 must also satisfy condition (5.17), which is an inequality of the form $A_1 \leq K q_2^{3/2}$, for some constant K depending on m, p .

We would like to obtain values of A_1 as large as possible, because A_1 and its multiple A_2 are the amplitudes of the chaotic movements we want to construct. We fix A_1 as the largest value compatible with (5.8) and (5.17). Thus, we define

$$\varepsilon := q_2^{\frac{1}{2}}, \quad A_1 := \left(\frac{3}{32} \frac{\alpha_1^3}{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} \right)^{\frac{1}{2}} \varepsilon^3, \quad (5.23)$$

so that (5.17) is satisfied. Note that (5.8) becomes

$$\left(\frac{3}{32} \frac{\alpha_1^3}{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} \right)^{\frac{1}{2}} \varepsilon^3 \leq \frac{1}{9} \varepsilon^2. \quad (5.24)$$

We define

$$\varepsilon_0 := \min \left\{ 1, \frac{1}{9} \left(\frac{32}{3} \frac{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}{\alpha_1^3} \right)^{\frac{1}{2}}, \frac{\delta_1}{\sqrt{8\alpha_4^{m_1}}} \right\}, \quad (5.25)$$

where δ_1 is the universal constant in (5.11) and in (2.8), so that both (5.24) and (5.11) are satisfied for all $0 < \varepsilon \leq \varepsilon_0$. The constant ε_0 depends only on m, p . For $\sigma = m/p$ small enough, the minimum in (5.25) is the third element of the set, as we note in the following lemma.

Lemma 5.9. *There exists a universal constant $\sigma_1 \in (0, 1)$ such that, if $\sigma = m/p \leq \sigma_1$, then ε_0 defined in (5.25) is $\varepsilon_0 = \delta_1/(\sqrt{8\alpha_4^{m_1}})$.*

Proof. It is enough to recall that $\alpha_1^2 + \alpha_2^2 + \alpha_3^2 > 3\alpha_1^2, \alpha_4^{m_1} \geq \alpha_4$, and $\alpha_4 > \alpha_1 > \sqrt{\alpha_1}$. ■

Lemma 5.10. *Let $a, m, p, \sigma, \alpha_1, \dots, \alpha_4, \bar{\mu}_1(\sigma), \bar{\mu}_2(\sigma), \gamma, (\xi_1(t), \eta_1(t), \xi_2(t), \eta_2(t))$ be like in Lemma 5.1. Consider any $\varepsilon \in (0, \varepsilon_0)$, where ε_0 is the constant, depending only on m, p , defined in (5.25), and define $q_2 := \varepsilon^2, A_1$ given by the second identity in (5.23), q_1 by (5.4), E_1, E_2 by (5.3), and $c_{123}, c_{234}, B, A_2, S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t)$ like in Lemma 5.1. Also define $\rho_{123} = \frac{8}{3}c_{123}/(\alpha_1\alpha_2\alpha_3), \rho_{234} = \frac{8}{3}c_{234}/(\alpha_2\alpha_3\alpha_4)$.*

Then $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ is a solution of system (3.3) for all $t \in \mathbb{R}$. Moreover, identity (5.4) is satisfied, and therefore (5.5) holds; inequality (5.8) is

satisfied, and therefore the lower bound (5.9) holds; inequality (5.11) is satisfied, and therefore the bound (5.12) holds; inequality (5.17) is satisfied, and therefore the thesis of Lemma 5.8 holds.

Proof. The proof follows from (5.23), (5.25), and the results of the previous subsections. ■

Recalling that the function η_2 given by Proposition 4.1 satisfies (4.4), we also have the following lemma, where $\varepsilon^2, \varepsilon^3$ are isolated from the other parts of the coefficients. The reason to consider the quantity \mathcal{N}_1 in (5.28) is that it corresponds to the square of the Sobolev norm $H^1(\mathbb{T}^d)$ of the solution of the Kirchhoff equation.

To simplify the exposition of the lemma, we rewrite (5.1) as

$$\begin{aligned} S_1(t) &= \varepsilon^2 s_1 - \varepsilon^3 a_1 \eta_1(b\varepsilon^3 t), \\ S_2(t) &= \varepsilon^2 s_2 - \varepsilon^3 a_1 \eta_1(b\varepsilon^3 t) - \varepsilon^3 a_2 \eta_2(b\varepsilon^3 t), \\ S_3(t) &= \varepsilon^2 s_3 + \varepsilon^3 a_1 \eta_1(b\varepsilon^3 t) - \varepsilon^3 a_2 \eta_2(b\varepsilon^3 t), \\ S_4(t) &= \varepsilon^2 s_4 + \varepsilon^3 a_2 \eta_2(b\varepsilon^3 t), \end{aligned}$$

where

$$\begin{aligned} s_1 &:= \frac{\alpha_4^2}{3\alpha_1^2}, & s_2 &:= \frac{\alpha_4^2}{3\alpha_2^2}, & s_3 &:= \frac{2\alpha_4^2}{3\alpha_3^2}, & s_4 &:= 1, \\ a_1 &:= \left(\frac{3}{32} \frac{\alpha_1^3}{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} \right)^{\frac{1}{2}}, & a_2 &:= \frac{1}{\gamma} a_1, & b &:= \frac{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}{2} a_1, \end{aligned} \quad (5.26)$$

so that

$$A_1 = a_1 \varepsilon^3, \quad A_2 = a_2 \varepsilon^3, \quad B = b \varepsilon^3. \quad (5.27)$$

Note that the constants $s_1, s_2, s_3, s_4, a_1, a_2, b$ depend only on m, p . We also define the associated function

$$\begin{aligned} \mathcal{N}_1(t) &:= \sum_{n=1}^4 \alpha_n^2 S_n(t) \\ &= \varepsilon^2 \frac{7\alpha_4^2}{3} + \varepsilon^3 (2\alpha_1 \alpha_2) a_1 \eta_1(b\varepsilon^3 t) + \varepsilon^3 (2\alpha_2 \alpha_3) a_2 \eta_2(b\varepsilon^3 t). \end{aligned} \quad (5.28)$$

Lemma 5.11. *There exists a universal constant $\sigma_2 \in (0, 1)$ with the following property. Let t_j, \bar{t}_j be given by Proposition 4.1, and define*

$$t_j^* := \frac{t_j}{B} = \frac{t_j}{b\varepsilon^3}, \quad \bar{t}_j^* := \frac{\bar{t}_j}{B} = \frac{\bar{t}_j}{b\varepsilon^3}, \quad I_j^* := [t_j^*, \bar{t}_j^*], \quad E_j^* := [\bar{t}_j^*, t_{j+1}^*]. \quad (5.29)$$

Assume the hypotheses of Lemma 5.10. If, in addition, the ratio $\sigma = m/p$ satisfies $\sigma \leq \sigma_2$, then $\mathcal{N}_1(t)$ defined in (5.28) satisfies

$$\begin{aligned} \varepsilon^2 c_1 + \frac{99}{100} \varepsilon^3 r_1 &\leq \mathcal{N}_1(t) \leq \varepsilon^2 c_1 + \frac{201}{100} \varepsilon^3 r_1 \quad \forall t \in I_j^*, \\ \varepsilon^2 c_1 - \frac{1}{100} \varepsilon^3 r_1 &\leq \mathcal{N}_1(t) \leq \varepsilon^2 c_1 + \frac{101}{100} \varepsilon^3 r_1 \quad \forall t \in E_j^*, \\ \max_{t \in I_j^*} \mathcal{N}_1(t) &\geq \varepsilon^2 c_1 + \frac{199}{100} \varepsilon^3 r_1, \quad \min_{t \in E_j^*} \mathcal{N}_1(t) \leq \varepsilon^2 c_1 + \frac{1}{100} \varepsilon^3 r_1, \end{aligned} \quad (5.30)$$

for all $j \geq 0$, where $c_1 := \frac{7}{3} \alpha_4^2$, $r_1 := (2\alpha_2 \alpha_3) a_2$ depend only on m, p .

Proof. To prove the first inequality in the last line of (5.30), consider the interval I_j^* , and let t'_j be a point in that interval where the function $\eta_2(b\varepsilon^3 t)$ achieves its maximum value. Since by (4.4) and (5.7),

$$\eta_2(b\varepsilon^3 t'_j) \geq 2 - C\sigma \quad \text{and} \quad |\eta_1(t)| \leq 3 \quad \forall t \in \mathbb{R},$$

from (5.28) and the identity $a_1 = \gamma a_2$ (see (5.26)) we get

$$\begin{aligned} \max_{t \in [t'_j, t_j^*]} \mathcal{N}_1(t) &\geq \mathcal{N}_1(t'_j) \geq \varepsilon^2 \frac{7\alpha_4^2}{3} + \varepsilon^3 (2\alpha_2 \alpha_3) a_2 (2 - C\sigma) - \varepsilon^3 (2\alpha_1 \alpha_2) 3a_1 \\ &= \varepsilon^2 \frac{7\alpha_4^2}{3} + \varepsilon^3 (2\alpha_2 \alpha_3) a_2 (2 - C\sigma) \left(1 - \frac{3\gamma\alpha_1}{\alpha_3(2 - C\sigma)}\right). \end{aligned}$$

The difference in the last parentheses tends to 1 as $\sigma \rightarrow 0$, because $\gamma \rightarrow 3$ and the ratio $\alpha_1/\alpha_3 = m/(2m+p)$ tends to 0 as $\sigma = m/p \rightarrow 0$. The other inequalities in (5.30) are proved similarly, using (4.3), (4.4), and the fact that

$$\begin{aligned} \mathcal{N}_1(t) &= \varepsilon^2 \frac{7\alpha_4^2}{3} + \varepsilon^3 (2\alpha_2 \alpha_3) a_2 \left(\eta_2(b\varepsilon^3 t) + \frac{\alpha_1 \gamma}{\alpha_3} \eta_1(b\varepsilon^3 t) \right) \\ &= \varepsilon^2 c_1 + \varepsilon^3 r_1 (\eta_2(b\varepsilon^3 t) + O(\sigma)). \end{aligned} \quad (5.31)$$

■

Remark 5.12. By (5.31), the oscillations of \mathcal{N}_1 at the main order in σ are fully described by those of η_2 , i.e., of S_4 .

Notation. From now on, we will sometimes be much less accurate than before in keeping track of the explicit dependence of the various constants on $\alpha_1, \alpha_2, \alpha_3, \alpha_4$; we will denote generically by K (or sometimes K' , or K'') any constant, possibly different from line to line, that depends only on the integers m, p .

With the new notation, the constants defined in Lemma 5.10 become

$$c_{123} = K\varepsilon^6, \quad c_{234} = K\varepsilon^6, \quad \rho_{123} = K\varepsilon^6, \quad \rho_{234} = K\varepsilon^6,$$

where the four constants K denote four (possibly different) values.

6. Approximation argument

Consider a solution $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ of system (3.3) obtained in Lemma 5.10, and let u_0 be the corresponding trigonometric polynomial in the H^{m_1} ball (2.8) constructed in Lemma 5.8. Let $v_0 := \bar{u}_0$ be the complex conjugate of u_0 , and let $(u(t), v(t))$ be the solution of the Cauchy problem for the transformed Kirchhoff equation (2.2) with initial condition (2.9). Since u_0 is a trigonometric polynomial, the solution (u, v) is global in time. Moreover, since u_0 belongs to the ball (2.8), the solution (u, v) satisfies (2.10) on the time interval $[0, T_{\text{NF}}]$; see Lemma 2.2.

The estimate in Lemma 5.4 computed at time $t = 0$ gives the inequality

$$\|u_0\|_s^2 \leq 8\alpha_4^{2s} q_2 \quad \forall s \geq 1,$$

and therefore, recalling (5.23) and the notation at the end of the previous section, T_{NF} in (2.10) can also be estimated in terms of ε as

$$T_{\text{NF}} \geq \frac{C_0}{(8\alpha_4^{2m_1} q_2)^2} = K\varepsilon^{-4}. \quad (6.1)$$

For each time $t \in [0, T_{\text{NF}}]$, to slightly simplify the notation (5.13), (5.14), (5.15), we denote by

$$S_n^u(t) := S_n^{(u(t))}$$

the superactions of the function $u(t, \cdot)$, and we also introduce analogous notation for all the other quantities in (5.13), (5.14), (5.15).

By Lemma 2.3, one has

$$C_1 S_n^u(0) \leq S_n^u(t) \leq C_2 S_n^u(0) \quad \forall t \in [0, T_{\text{NF}}], n = 1, 2, 3, 4. \quad (6.2)$$

By construction, one has $S_n^u(0) = S_n(0)$. Hence, by (6.2), (5.10), (5.23),

$$\begin{aligned} \|u(t)\|_s^2 &= \sum_{n=1}^4 \alpha_n^{2s} S_n^u(t) \leq C_2 \sum_{n=1}^4 \alpha_n^{2s} S_n^u(0) = C_2 \sum_{n=1}^4 \alpha_n^{2s} S_n(0) = C_2 \|u_0\|_s^2 \\ &\leq 8C_2 \alpha_4^{2s} q_2 = 8C_2 \alpha_4^{2s} \varepsilon^2 \end{aligned} \quad (6.3)$$

for all $t \in [0, T_{\text{NF}}]$, all $s \geq 1$. For $s \in \{1, m_1\}$, one has $\alpha_4^{2s} \leq \alpha_4^{2m_1}$, and therefore we can simply write

$$\|u(t)\|_s \leq K\varepsilon \quad \forall t \in [0, T_{\text{NF}}], s \in \{1, m_1\}. \quad (6.4)$$

Since $(u(t), v(t))$ solves (2.2) on $[0, T_{\text{NF}}]$, the functions $S_n^u(t), B_n^u(t)$ solve the effective equations (2.15) on the same time interval, and $Z_{123}^u(t), Z_{234}^u(t)$ solve (2.18) on the same time interval. Since $\rho_{123}^u(t), \rho_{234}^u(t)$ are continuous and they are positive at time $t = 0$, there exists $T_{\text{polar}} > 0$ such that $\rho_{123}^u(t), \rho_{234}^u(t)$ are positive for all $t \in [0, T_{\text{polar}}]$.

Remark 6.1. Note that, in general, T_{polar} could be smaller than T_{NF} .

On the time interval $[0, T_{\text{polar}}]$, the angles $\varphi_{123}^u(t)$, $\varphi_{234}^u(t)$ are well defined by the identities

$$\begin{aligned} Z_{123}^u(t) &= \rho_{123}^u(t) \exp(i\varphi_{123}^u(t)), \\ Z_{234}^u(t) &= \rho_{234}^u(t) \exp(i\varphi_{234}^u(t)), \quad \forall t \in [0, T_{\text{polar}}]. \end{aligned}$$

Since $Z_{123}^u(t)$, $Z_{234}^u(t)$ solve (2.18), the functions $\rho_{123}^u(t)$, $\rho_{234}^u(t)$ solve the equations

$$\partial_t \rho_{123}^u = R_{\rho_{123}^u}, \quad \partial_t \rho_{234}^u = R_{\rho_{234}^u} \quad (6.5)$$

on $[0, T_{\text{polar}}]$, and the functions $\varphi_{123}^u(t)$, $\varphi_{234}^u(t)$ solve the equations

$$\begin{aligned} \partial_t \varphi_{123}^u &= -\frac{1}{2}(\alpha_1^2 S_1^u + \alpha_2^2 S_2^u - \alpha_3^2 S_3^u) + R_{\varphi_{123}^u}, \\ \partial_t \varphi_{234}^u &= -\frac{1}{2}(\alpha_2^2 S_2^u + \alpha_3^2 S_3^u - \alpha_4^2 S_4^u) + R_{\varphi_{234}^u} \end{aligned} \quad (6.6)$$

on $[0, T_{\text{polar}}]$, where

$$R_{\rho_{123}^u} := \text{Re}(\exp(-i\varphi_{123}^u) R_{Z_{123}^u}), \quad R_{\varphi_{123}^u} := \frac{1}{\rho_{123}^u} \text{Im}(\exp(-i\varphi_{123}^u) R_{Z_{123}^u}), \quad (6.7)$$

and analogous definitions for $R_{\rho_{234}^u}$, $R_{\varphi_{234}^u}$ (just replace 123 with 234 everywhere in (6.7)). The remainders of type R_Z are defined in (2.19) and estimated in (2.20). Moreover, S_n^u , $n = 1, 2, 3, 4$, solve

$$\begin{aligned} \partial_t S_1^u &= \frac{3}{8} \rho_{123}^u \alpha_1 \alpha_2 \alpha_3 \sin(\varphi_{123}^u) + R_{S_1^u}, \\ \partial_t S_2^u &= \frac{3}{8} \rho_{123}^u \alpha_1 \alpha_2 \alpha_3 \sin(\varphi_{123}^u) + \frac{3}{8} \rho_{234}^u \alpha_2 \alpha_3 \alpha_4 \sin(\varphi_{234}^u) + R_{S_2^u}, \\ \partial_t S_3^u &= -\frac{3}{8} \rho_{123}^u \alpha_1 \alpha_2 \alpha_3 \sin(\varphi_{123}^u) + \frac{3}{8} \rho_{234}^u \alpha_2 \alpha_3 \alpha_4 \sin(\varphi_{234}^u) + R_{S_3^u}, \\ \partial_t S_4^u &= -\frac{3}{8} \rho_{234}^u \alpha_2 \alpha_3 \alpha_4 \sin(\varphi_{234}^u) + R_{S_4^u} \end{aligned} \quad (6.8)$$

on $[0, T_{\text{polar}}]$, where the remainders of the type R_S appear in (2.15) and are estimated in (2.16). In the following lemma we prove a formula for T_{polar} .

Lemma 6.2. *Assume the hypotheses of Lemma 5.10. There exists a universal constant $\sigma_3 \in (0, 1)$ such that, if $\sigma = m/p$, in addition to the hypotheses of Lemma 5.10, also satisfies $\sigma \leq \sigma_3$, then $Z_{123}^u(t) \neq 0$, $Z_{234}^u(t) \neq 0$ for all $t \in [0, T_{\text{polar}}]$, with*

$$T_{\text{polar}} = \frac{C \alpha_1 A_1^2}{\alpha_4^{4m_1+2} q_2^5} = K \varepsilon^{-4} \leq T_{\text{NF}}, \quad (6.9)$$

where $C > 0$ is a universal constant and $K > 0$ is a constant depending only on m , p . As a consequence, on $[0, T_{\text{polar}}]$, ρ_{123}^u , ρ_{234}^u are positive and φ_{123}^u , φ_{234}^u are well defined. Moreover,

$$|\rho_{123}^u(t) - \rho_{123}| \leq \frac{1}{2} \rho_{123}, \quad |\rho_{234}^u(t) - \rho_{234}| \leq \frac{1}{2} \rho_{234}. \quad (6.10)$$

Proof. In this proof, C (and C' , C'') denotes a universal constant, possibly different from line to line. Suppose that ρ_{123}^u is positive on some time interval $[0, T] \subset [0, T_{\text{NF}}]$. Then, by (6.5) and (6.7),

$$|\partial_t \rho_{123}^u| = |\operatorname{Re}(\exp(-i\varphi_{123}^u) R_{Z_{123}^u})| \leq |\exp(-i\varphi_{123}^u) R_{Z_{123}^u}| = |R_{Z_{123}^u}|.$$

By (2.20), for all $s \in \mathbb{R}$,

$$|R_{Z_{123}^u}| \leq C \|u\|_{m_1}^4 S_1^u S_2^u S_3^u = \frac{C \|u\|_{m_1}^4}{(\alpha_1 \alpha_2 \alpha_3)^{2s}} (\alpha_1^{2s} S_1^u) (\alpha_2^{2s} S_2^u) (\alpha_3^{2s} S_3^u) \leq \frac{C \|u\|_{m_1}^4 \|u\|_s^6}{(\alpha_1 \alpha_2 \alpha_3)^{2s}}$$

and, by (6.3), for $s \geq 1$ and $t \in [0, T]$,

$$\|u(t)\|_{m_1}^4 \|u(t)\|_s^6 \leq C \alpha_4^{4m_1+6s} q_2^5.$$

Hence, since $\alpha_4 \leq 2\alpha_3$ and $\alpha_4 \leq 3\alpha_2$, for $s = 1$, we obtain

$$|\partial_t \rho_{123}^u| \leq |R_{Z_{123}^u}| \leq \frac{C \alpha_4^{4m_1+6} q_2^5}{(\alpha_1 \alpha_2 \alpha_3)^2} \leq \frac{C' \alpha_4^{4m_1+2} q_2^5}{\alpha_1^2} \quad \forall t \in [0, T]. \quad (6.11)$$

At time $t = 0$ we have, by construction, and by the definition (3.24) of c_{123} ,

$$\rho_{123}^u(0) = \rho_{123} = \frac{8c_{123}}{3\alpha_1\alpha_2\alpha_3} = \frac{4(\alpha_1^2 + \alpha_2^2 + \alpha_3^2)}{3\alpha_1\alpha_2\alpha_3} A_1^2, \quad (6.12)$$

which is positive because $A_1 > 0$. Hence

$$|\rho_{123}^u(t) - \rho_{123}| \leq \int_0^t |\partial_s \rho_{123}^u(s)| ds \leq \frac{C' \alpha_4^{4m_1+2} q_2^5}{\alpha_1^2} t \quad \forall t \in [0, T],$$

and the last quantity is strictly less than ρ_{123} for all $t \in [0, T]$ if

$$T < \frac{C \alpha_1 A_1^2}{\alpha_4^{4m_1+2} q_2^5} =: T_{123}^*.$$

This implies that $\rho_{123}^u > 0$ on $[0, T_{123}^*]$. Moreover, by (5.17) and (6.1),

$$T_{123}^* = \frac{C \alpha_1 A_1^2}{\alpha_4^{4m_1+2} q_2^5} \leq C' \frac{\alpha_1^4}{\alpha_4^4} \frac{1}{\alpha_4^{2m_1} q_2^2} \leq C'' \sigma^4 T_{\text{NF}},$$

where $\sigma = m/p$. Then, for $C'' \sigma^4 \leq 1$, one has $T_{123}^* \leq T_{\text{NF}}$.

Proceeding similarly for ρ_{234}^u , we obtain the estimate

$$|\partial_t \rho_{234}^u| \leq |R_{Z_{234}^u}| \leq \frac{C \alpha_4^{4m_1+6} q_2^5}{(\alpha_2 \alpha_3 \alpha_4)^2} \leq C' \alpha_4^{4m_1} q_2^5, \quad (6.13)$$

and we deduce that $\rho_{234}^u > 0$ on $[0, T_{234}^*]$, with

$$T_{234}^* = \frac{C A_1^2}{\alpha_4^{4m_1+1} q_2^5} \leq C' \sigma^3 T_{\text{NF}}.$$

Also, $T_{123}^* \leq T_{234}^* C \sigma \leq T_{234}^*$ for $C \sigma \leq 1$.

Finally, the estimates already proved also give that $|\rho_{123}^u(t) - \rho_{123}| \leq \frac{1}{2}\rho_{123}$ for all $t \in [0, T_{123}^*/2]$. Thus, we fix $T_{\text{polar}} = T_{123}^*/2$. ■

We recall that the superscript u indicates quantities related to the solution (u, v) of the Cauchy problem (2.2), (2.9), while the absence of that superscript corresponds to the solution of the truncated effective system (3.3). By construction, we have

$$S^u(0) = S(0), \quad \rho^u(0) = \rho(0), \quad \varphi^u(0) = \varphi(0).$$

Moreover, ρ is constant in time, and there are no remainders R_S, R_ρ, R_φ without the superscript u because (3.3) is the truncated effective system.

The next lemma gives estimates for the difference between $S^u(t)$ and $S(t)$ in a certain time interval.

Lemma 6.3. *Assume the hypotheses of Lemma 5.10, and consider the associated solution $(S_1(t), S_2(t), S_3(t), S_4(t), \varphi_{123}(t), \varphi_{234}(t))$ of system (3.3) given by Lemma 5.10. Let u_0 be the corresponding trigonometric polynomial in the ball (2.8) constructed in Lemma 5.8. Let $v_0 := \bar{u}_0$ be the complex conjugate of u_0 , and let $(u(t), v(t))$ be the solution of the Cauchy problem for the transformed Kirchhoff equation (2.2) with initial condition (2.9). Assume that the ratio $\sigma = m/p$ satisfies $\sigma \leq \sigma_3$, where σ_3 is defined in Lemma 6.2.*

There exists $\varepsilon_1 \in (0, \varepsilon_0)$, depending only on m, p , where ε_0 is the constant in (5.25), such that, for $0 < \varepsilon \leq \varepsilon_1$, one has

$$\varepsilon^{-3}|S_n^u(t) - S_n(t)| \leq \varepsilon^{\frac{3}{4}} \quad \forall t \in [0, T_\varepsilon], \quad n = 1, 2, 3, 4, \quad (6.14)$$

where

$$T_\varepsilon = K\varepsilon^{-3} \log(\varepsilon^{-1}) \quad (6.15)$$

for some positive constant K depending only on m, p .

Proof. From the difference of equations (6.8) for S_n^u and equations (3.3) for S_n , we obtain that for all $t \in [0, T_{\text{polar}}]$, $n = 1, 2, 3, 4$,

$$\begin{aligned} |\partial_t(S_n^u - S_n)| &\leq \alpha_1\alpha_2\alpha_3(\rho_{123}|\varphi_{123}^u - \varphi_{123}| + |\rho_{123}^u - \rho_{123}|) \\ &\quad + \alpha_2\alpha_3\alpha_4(\rho_{234}|\varphi_{234}^u - \varphi_{234}| + |\rho_{234}^u - \rho_{234}|) + |R_{S_n^u}|. \end{aligned}$$

By (2.16) one has $|R_{S_n^u}| \leq C\|u\|_{m_1}^6|S_n^u|$. By (6.4),

$$\|u\|_{m_1} \leq K\varepsilon, \quad |S_n^u| \leq \alpha_n|S_n^u| \leq \|u\|_1^2 \leq K\varepsilon^2,$$

and therefore $|R_{S_n^u}| \leq K\varepsilon^8$ on $[0, T_{\text{polar}}]$. Hence, for $t \in [0, T_{\text{polar}}]$ and $n = 1, 2, 3, 4$,

$$\begin{aligned} |\partial_t(S_n^u - S_n)| &\leq K(\rho_{123}|\varphi_{123}^u - \varphi_{123}| + |\rho_{123}^u - \rho_{123}|) \\ &\quad + K(\rho_{234}|\varphi_{234}^u - \varphi_{234}| + |\rho_{234}^u - \rho_{234}|) + K\varepsilon^8. \end{aligned} \quad (6.16)$$

Since ρ_{123}, ρ_{234} are constants, by equations (6.5) for $\rho_{123}^u, \rho_{234}^u$, one has

$$\partial_t(\rho_{123}^u - \rho_{123}) = \partial_t \rho_{123}^u = R_{\rho_{123}^u}, \quad \partial_t(\rho_{234}^u - \rho_{234}) = \partial_t \rho_{234}^u = R_{\rho_{234}^u}.$$

The time derivatives $\partial_t \rho_{123}^u, \partial_t \rho_{234}^u$ were estimated in (6.11) and (6.13); hence, integrating in time, we get

$$|\rho_{123}^u(t) - \rho_{123}| \leq K\varepsilon^{10}t, \quad |\rho_{234}^u(t) - \rho_{234}| \leq K\varepsilon^{10}t \quad \forall t \in [0, T_{\text{polar}}].$$

Also, by (6.12) and (5.27), $\rho_{123} \leq K\varepsilon^6, \rho_{234} \leq K\varepsilon^6$. We plug these estimates into (6.16) and we get

$$|\partial_t(S_n^u - S_n)| \leq K\varepsilon^6(|\varphi_{123}^u - \varphi_{123}| + |\varphi_{234}^u - \varphi_{234}|) + K\varepsilon^{10}t + K\varepsilon^8 \quad (6.17)$$

for all $t \in [0, T_{\text{polar}}], n = 1, 2, 3, 4$.

Subtracting the equations (6.6) for $\varphi_{123}^u, \varphi_{234}^u$ and those for $\varphi_{123}, \varphi_{234}$ in (3.3), we have

$$\begin{aligned} |\partial_t(\varphi_{123}^u - \varphi_{123})| &\leq \alpha_1^2|S_1^u - S_1| + \alpha_2^2|S_2^u - S_2| + \alpha_3^2|S_3^u - S_3| + |R_{\varphi_{123}^u}|, \\ |\partial_t(\varphi_{234}^u - \varphi_{234})| &\leq \alpha_2^2|S_2^u - S_2| + \alpha_3^2|S_3^u - S_3| + \alpha_4^2|S_4^u - S_4| + |R_{\varphi_{234}^u}|. \end{aligned}$$

By (6.10), $\rho_{123}^u \geq \frac{1}{2}\rho_{123}$ on $[0, T_{\text{polar}}]$. Hence, by (6.7), $|R_{\varphi_{123}^u}| \leq (2/\rho_{123})|R_{Z_{123}^u}|$, and $R_{Z_{123}^u}$ was estimated in (6.11). Also, ρ_{123} is given in (6.12). Therefore,

$$|R_{\varphi_{123}^u}| \leq \frac{2}{\rho_{123}}|R_{Z_{123}^u}| \leq \frac{Kq_2^5}{A_1^2} = K\varepsilon^4$$

and, similarly, $|R_{\varphi_{234}^u}| \leq K\varepsilon^4$. Thus,

$$|\partial_t(\varphi_{123}^u - \varphi_{123})| + |\partial_t(\varphi_{234}^u - \varphi_{234})| \leq K \sum_{n=1}^4 |S_n^u - S_n| + K\varepsilon^4. \quad (6.18)$$

We apply Grönwall's inequality to (6.17) and (6.18). It is convenient to introduce a factor ε^β , with β to be determined, because the factors in (6.17) and (6.18) contain different powers of ε . We define the vector

$$\begin{aligned} \psi &:= (\varepsilon^\beta(S_1^u - S_1), \varepsilon^\beta(S_2^u - S_2), \varepsilon^\beta(S_3^u - S_3), \varepsilon^\beta(S_4^u - S_4), \\ &\quad (\varphi_{123}^u - \varphi_{123}), (\varphi_{234}^u - \varphi_{234})) \\ &=: (\psi_1, \psi_2, \psi_3, \psi_4, \psi_5, \psi_6), \end{aligned}$$

which is a function of $t \in [0, T_{\text{polar}}]$ taking values in \mathbb{R}^6 . From (6.17) we obtain

$$\begin{aligned} |\partial_t \psi_n| = \varepsilon^\beta |\partial_t(S_n^u - S_n)| &\leq \varepsilon^\beta [K\varepsilon^6(|\psi_5| + |\psi_6|) + K\varepsilon^{10}t + K\varepsilon^8] \\ &\leq K\varepsilon^{6+\beta}|\psi| + K\varepsilon^{10+\beta}t + K\varepsilon^{8+\beta}, \quad n = 1, 2, 3, 4, \end{aligned}$$

and from (6.18) we get

$$|\partial_t \psi_5| + |\partial_t \psi_6| \leq K\varepsilon^{-\beta} (|\psi_1| + |\psi_2| + |\psi_3| + |\psi_4|) + K\varepsilon^4 \leq K\varepsilon^{-\beta} |\psi| + K\varepsilon^4,$$

where $|\cdot|$ is the usual Euclidean norm of \mathbb{R}^6 . Therefore,

$$|\partial_t \psi| \leq K(\varepsilon^{6+\beta} + \varepsilon^{-\beta})|\psi| + K\varepsilon^{10+\beta} t + K\varepsilon^{8+\beta} + K\varepsilon^4. \quad (6.19)$$

We fix the value of β that minimizes the sum $\varepsilon^{6+\beta} + \varepsilon^{-\beta}$, that is, $\beta = -3$. Hence (6.19) becomes

$$|\partial_t \psi| \leq K\varepsilon^3 |\psi| + K\varepsilon^7 t + K\varepsilon^4.$$

Moreover, by the choice of the initial conditions, $\psi(0) = 0$. Then, by Grönwall's inequality,

$$|\psi(t)| \leq \exp(K_0\varepsilon^3 t) K_1(\varepsilon^7 t^2 + \varepsilon^4 t) \quad \forall t \in [0, T_{\text{polar}}], \quad (6.20)$$

for some positive constants K_0, K_1 depending only on m, p .

Now we consider the time $T_{\varepsilon, h} := K_0^{-1} \varepsilon^{-3} \log(\varepsilon^{-h})$, where h is any positive real constant and $K_0 > 0$ is the constant appearing in (6.20); we note that $T_{\varepsilon, h}$ depends on h, ε, m, p . By (6.9), one has $T_{\varepsilon, h} \leq T_{\text{polar}}$ for $0 < \varepsilon \leq C$, for some positive constant C depending only on h, m, p . By (6.20), for all $t \in [0, T_{\varepsilon, h}]$ one has

$$\begin{aligned} |\psi(t)| &\leq \exp(K_0\varepsilon^3 t) K_1(\varepsilon^7 t^2 + \varepsilon^4 t) \\ &\leq \varepsilon^{-h} K_1(\varepsilon^7 K_0^{-2} \varepsilon^{-6} \log^2(\varepsilon^{-h}) + \varepsilon^4 K_0^{-1} \varepsilon^{-3} \log(\varepsilon^{-h})) \\ &\leq \varepsilon^{1-h} K_1(K_0^{-2} \log^2(\varepsilon^{-h}) + K_0^{-1} \log(\varepsilon^{-h})). \end{aligned}$$

We also note that

$$\varepsilon^{1-h} K_1(K_0^{-2} \log^2(\varepsilon^{-h}) + K_0^{-1} \log(\varepsilon^{-h})) \leq \varepsilon^{1-2h} \quad \text{for } 0 < \varepsilon \leq C',$$

for some positive constant C' depending on h, m, p . We can fix, for example, $h = 1/8$. ■

6.1. Motion of the Sobolev norms

We use the $H^1(\mathbb{T}^d)$ Sobolev norm $\|u(t)\|_1$ to describe the transfer of energy, namely the exchanges in amplitude of the superactions $S_n^u(t)$, $n = 1, 2, 3, 4$, as time evolves. Except for the $H^{\frac{1}{2}}$ norm, which is constant in time for the solutions of the approximating system (3.31) given by Lemma 5.1 (see Remark 3.6), any other H^s norms could be used to capture the transfer of energy among the superactions. We decide to use the H^1 norm because H^1 here corresponds to the space $H^{\frac{3}{2}} \times H^{\frac{1}{2}}$ for the ‘‘physical variables’’, which is the space of the standard local well-posedness for the Kirchhoff equation.

Lemma 6.4. *Assume all the hypotheses of Lemma 6.3, and also let $\sigma \leq \sigma_2$, where σ_2 is given by Lemma 5.11. Let $t_j^*, \bar{t}_j^*, I_j^*, E_j^*$ be defined in (5.29). There exists a constant $\varepsilon_2 \in (0, \varepsilon_1)$, depending only on m, p , where ε_1 is defined in Lemma 6.3, such that if, in*

addition to the hypotheses of Lemma 6.3, $0 < \varepsilon \leq \varepsilon_2$, then the $H^1(\mathbb{T}^d)$ Sobolev norm of $u(t)$ satisfies

$$\begin{aligned} \varepsilon^2 c_1 + \frac{98}{100} \varepsilon^3 r_1 &\leq \|u(t)\|_1^2 \leq \varepsilon^2 c_1 + \frac{202}{100} \varepsilon^3 r_1 \quad \forall t \in I_j^*, \\ \varepsilon^2 c_1 - \frac{2}{100} \varepsilon^3 r_1 &\leq \|u(t)\|_1^2 \leq \varepsilon^2 c_1 + \frac{102}{100} \varepsilon^3 r_1 \quad \forall t \in E_j^*, \\ \max_{t \in I_j^*} \|u(t)\|_1^2 &\geq \varepsilon^2 c_1 + \frac{198}{100} \varepsilon^3 r_1, \quad \min_{t \in E_j^*} \|u(t)\|_1^2 \leq \varepsilon^2 c_1 + \frac{2}{100} \varepsilon^3 r_1, \end{aligned} \quad (6.21)$$

for all the indices $j \geq 0$ for which the intervals I_j^* , E_j^* are contained in $[0, T_\varepsilon]$, where T_ε is defined in (6.15) and c_1 , r_1 are defined in Lemma 5.11.

Moreover, there exists a constant $K > 0$, depending only on m , p , such that, if the integers m_j introduced in Proposition 4.1 satisfy

$$\sum_{j=0}^N m_j \leq K \log(\varepsilon^{-1}) \quad (6.22)$$

for some N , then the intervals I_0^* , E_0^* , \dots , I_N^* , E_N^* are all contained in $[0, T_\varepsilon]$.

Proof. Consider an interval $I_j^* \subseteq [0, T_\varepsilon]$, and a point t'_j in that interval where the function \mathcal{N}_1 , defined in (5.28), achieves its maximum value over I_j^* . By (5.30) and (6.14),

$$\begin{aligned} \max_{t \in I_j^*} \|u(t)\|_1^2 &\geq \|u(t'_j)\|_1^2 = \sum_{n=1}^4 \alpha_n^2 S_n^u(t'_j) = \mathcal{N}_1(t'_j) + \sum_{n=1}^4 \alpha_n^2 (S_n^u(t'_j) - S_n(t'_j)) \\ &\geq \varepsilon^2 c_1 + \frac{199}{100} \varepsilon^3 r_1 - \sum_{n=1}^4 \alpha_n^2 \varepsilon^{3+\frac{3}{4}} \geq \varepsilon^2 c_1 + \frac{198}{100} \varepsilon^3 r_1. \end{aligned}$$

The other inequalities in (6.21) are proved similarly.

From (4.2) one has $t_{j+1} = \tau(m_0 + \theta_0 + \dots + m_j + \theta_j)$. Hence, by (5.29),

$$t_{j+1}^* = \frac{\tau}{b\varepsilon^3} \sum_{k=0}^j (m_k + \theta_k) \quad \forall j = 0, 1, 2, \dots$$

Since $(m_k + \theta_k) \leq 2m_k$, recalling (6.15), one has $t_{N+1}^* \leq T_\varepsilon$ if (6.22) holds. \blacksquare

Lemma 6.5. Assume the hypotheses of Lemma 6.4. There exists a constant $\varepsilon_3 \in (0, \varepsilon_2)$, depending only on m , p , where ε_2 is defined in Lemma 6.4, such that, if $0 < \varepsilon \leq \varepsilon_3$, then

$$\begin{aligned} \varepsilon \tilde{c}_1 + \frac{48}{100} \varepsilon^2 \tilde{r}_1 &\leq \|u(t)\|_1 \leq \varepsilon \tilde{c}_1 + \frac{101}{100} \varepsilon^2 \tilde{r}_1 \quad \forall t \in I_j^*, \\ \varepsilon \tilde{c}_1 - \frac{2}{100} \varepsilon^2 \tilde{r}_1 &\leq \|u(t)\|_1 \leq \varepsilon \tilde{c}_1 + \frac{51}{100} \varepsilon^2 \tilde{r}_1 \quad \forall t \in E_j^*, \\ \max_{t \in I_j^*} \|u(t)\|_1 &\geq \varepsilon \tilde{c}_1 + \frac{98}{100} \varepsilon^2 \tilde{r}_1, \quad \min_{t \in E_j^*} \|u(t)\|_1 \leq \varepsilon \tilde{c}_1 + \frac{1}{100} \varepsilon^2 \tilde{r}_1, \end{aligned} \quad (6.23)$$

where $\tilde{c}_1 := \sqrt{c_1}$ and $\tilde{r}_1 := r_1/\sqrt{c_1}$. The inequalities in (6.23) hold for the indices j described in Lemma 6.4, that is, for $j = 0, \dots, N$, where N satisfies (6.22).

Proof. The inequalities in (6.23) are obtained from (6.21) by the Taylor expansion $(1+x)^{1/2} = 1 + \frac{1}{2}x + O(x^2)$ as $x \rightarrow 0$ and the inequality $(1+x)^{1/2} \leq 1 + \frac{1}{2}x$, which holds for all $x \geq -1$. ■

6.2. Back to the solutions of the Kirchhoff equation

As is observed in Lemma 2.2, if (u_0, v_0) is in the ball (2.8), then, for all $t \in [0, T_{\text{NF}}]$, the solution (u, v) of the Cauchy problem (2.2), (2.9) remains in the ball $\|u\|_{m_1} \leq \delta$ where the transformation Φ is well defined, and $(\tilde{u}, \tilde{v}) = \Phi(u, v)$ in (2.11) solves the original system (2.1) on the same time interval.

We want to prove that the solution (\tilde{u}, \tilde{v}) has a dynamical behavior similar to that of (u, v) in Lemma 6.5. We underline that $v = \bar{u}$, hence $\|v\|_1 = \|u\|_1$, and, in fact, the inequalities in Lemma 6.5 concern the solution (u, v) . On the contrary, the ‘‘physical’’ solution (\tilde{u}, \tilde{v}) is a pair of real-valued functions solving (2.1), and therefore $\tilde{v} = \partial_t \tilde{u}$. Thus, $\|u\|_1$ appearing in Lemma 6.5 corresponds to \mathcal{N} in Lemma 6.9.

The transformation Φ is defined in (2.5). We consider the map $\Phi_3 \circ \Phi_4 \circ \Phi_5$ first, and then Φ_1, Φ_2 . From Lemma 2.1 we deduce the following property.

Lemma 6.6. *Let δ, C be the universal constants in Lemma 2.1. Let $(u, v) \in H_0^{m_1}(\mathbb{T}^d, \text{c.c.})$, with $\|u\|_{m_1} \leq \delta$, and let $(f, g) := \Phi_3 \circ \Phi_4 \circ \Phi_5(u, v)$. Then, for every $s \in \mathbb{R}$,*

$$\|f - u\|_s \leq 2C \|u\|_{m_1}^2 \|u\|_s. \quad (6.24)$$

Proof. By (2.7),

$$\begin{aligned} \|f - u\|_s^2 &= \sum_{k \in \mathbb{Z}^d} |k|^{2s} |f_k - u_k|^2 \leq \sum_k |k|^{2s} C^2 \|u\|_{m_1}^4 (|u_k| + |u_{-k}|)^2 \\ &\leq 2C^2 \|u\|_{m_1}^4 \sum_k |k|^{2s} (|u_k|^2 + |u_{-k}|^2) = 4C^2 \|u\|_{m_1}^4 \|u\|_s^2. \quad \blacksquare \end{aligned}$$

We apply estimate (6.24) to the solution (u, v) of (2.2) constructed in the previous sections.

Lemma 6.7. *Assume the hypotheses of Lemma 6.5. Let v be the complex conjugate of u , and let $(f, g) := \Phi_3 \circ \Phi_4 \circ \Phi_5(u, v)$. There exists a constant $\varepsilon_4 \in (0, \varepsilon_3)$, depending only on m, p , where ε_3 is given by Lemma 6.5, such that, if $0 < \varepsilon \leq \varepsilon_4$, then*

$$\begin{aligned} \varepsilon \tilde{c}_1 + \frac{47}{100} \varepsilon^2 \tilde{r}_1 &\leq \|f(t)\|_1 \leq \varepsilon \tilde{c}_1 + \frac{102}{100} \varepsilon^2 \tilde{r}_1 \quad \forall t \in I_j^*, \\ \varepsilon \tilde{c}_1 - \frac{3}{100} \varepsilon^2 \tilde{r}_1 &\leq \|f(t)\|_1 \leq \varepsilon \tilde{c}_1 + \frac{52}{100} \varepsilon^2 \tilde{r}_1 \quad \forall t \in E_j^*, \\ \max_{t \in I_j^*} \|f(t)\|_1 &\geq \varepsilon \tilde{c}_1 + \frac{97}{100} \varepsilon^2 \tilde{r}_1, \quad \min_{t \in E_j^*} \|f(t)\|_1 \leq \varepsilon \tilde{c}_1 + \frac{2}{100} \varepsilon^2 \tilde{r}_1, \end{aligned} \quad (6.25)$$

where \tilde{c}_1, \tilde{r}_1 are defined in Lemma 6.5. Inequalities (6.25) hold for the indices j described in Lemma 6.4, that is, for $j = 0, \dots, N$, where N satisfies (6.22).

Proof. By (6.24) and (6.4), one has

$$\|f - u\|_1 \leq 2C \|u\|_{m_1}^2 \|u\|_1 \leq K\varepsilon^3 \quad (6.26)$$

on the time interval $[0, T_{\text{NF}}]$, where the constant K depends only on m, p . All the inequalities in the lemma are proved by (6.23) and (6.26). ■

The transformations Φ_1, Φ_2 are simply these:

$$\begin{aligned} (q, p) &= \Phi_2(f, g) = \left(\frac{f+g}{\sqrt{2}}, \frac{f-g}{i\sqrt{2}} \right), \\ (\tilde{u}, \tilde{v}) &= \Phi_1(q, p) = (|D_x|^{-\frac{1}{2}}q, |D_x|^{\frac{1}{2}}p), \end{aligned} \quad (6.27)$$

where $|D_x|^s$ is the Fourier multiplier $|D_x|^s e^{ik \cdot x} = |k|^s e^{ik \cdot x}$, $s \in \mathbb{R}$, $k \in \mathbb{Z}^d \setminus \{0\}$ (the frequency $k = 0$ can be ignored here, because only zero-average functions are involved).

Lemma 6.8. *Let $s \in \mathbb{R}$, let f be a zero-average, complex-valued function in $H^s(\mathbb{T}^d, \mathbb{C})$, and let g be its complex conjugate, i.e., $(f, g) \in H_0^s(\mathbb{T}^d, \text{c.c.})$. Then (q, p) defined in (6.27) is a pair of zero-average, real-valued functions in $H_0^s(\mathbb{T}^d, \mathbb{R})$, with*

$$\|q\|_s^2 + \|p\|_s^2 = \|f\|_s^2 + \|g\|_s^2 = 2\|f\|_s^2,$$

and (\tilde{u}, \tilde{v}) defined in (6.27) is a pair of zero-average, real-valued functions in $H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$, with

$$\|\tilde{u}\|_{s+\frac{1}{2}} = \|q\|_s, \quad \|\tilde{v}\|_{s-\frac{1}{2}} = \|p\|_s.$$

Proof. Elementary calculations with Fourier coefficients. ■

The next lemma concerns the solutions (\tilde{u}, \tilde{v}) of system (2.1), that is, the solutions \tilde{u} of the original Kirchhoff equation (1.1).

Lemma 6.9. *Assume the hypotheses of Lemma 6.7, and let $(\tilde{u}, \tilde{v}) := \Phi_1 \circ \Phi_2(f, g)$. If $\varepsilon \leq \varepsilon_4$, then the function*

$$\mathcal{N}(t) := (\|\tilde{u}(t)\|_{\frac{3}{2}}^2 + \|\tilde{v}(t)\|_{\frac{1}{2}}^2)^{\frac{1}{2}}$$

satisfies

$$\begin{aligned} \varepsilon c_0 + \frac{47}{100} \varepsilon^2 r_0 &\leq \mathcal{N}(t) \leq \varepsilon c_0 + \frac{102}{100} \varepsilon^2 r_0 \quad \forall t \in I_j^*, \\ \varepsilon c_0 - \frac{3}{100} \varepsilon^2 r_0 &\leq \mathcal{N}(t) \leq \varepsilon c_0 + \frac{52}{100} \varepsilon^2 r_0 \quad \forall t \in E_j^*, \\ \max_{t \in I_j^*} \mathcal{N}(t) &\geq \varepsilon c_0 + \frac{97}{100} \varepsilon^2 r_0, \quad \min_{t \in E_j^*} \mathcal{N}(t) \leq \varepsilon c_0 + \frac{2}{100} \varepsilon^2 r_0, \end{aligned} \quad (6.28)$$

where $c_0 := \sqrt{2}\tilde{c}_1$, $r_0 := \sqrt{2}\tilde{r}_1$, and \tilde{c}_1, \tilde{r}_1 are defined in Lemma 6.5. Inequalities (6.28) hold for the indices j described in Lemma 6.4, that is, for $j = 0, \dots, N$, where N satisfies (6.22).

Proof. By Lemma 6.8, one has $\mathcal{N}(t) = \sqrt{2}\|f(t)\|_1$. Hence (6.28) follows directly from (6.25). ■

Proof of Theorem 1.1. All the previous smallness conditions on σ are satisfied for $\sigma \leq \sigma_4 := \min\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$, where $\sigma_0 = \sigma_0(a_0)$ is given by Proposition 4.1, σ_1 by Lemma 5.9, σ_2 by Lemma 5.11, σ_3 by Lemma 6.2. Note that $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ are all universal constants, and therefore σ_4 is universal too. All the previous smallness conditions on ε are satisfied for $0 < \varepsilon \leq \varepsilon_4$, where ε_4 is defined in Lemma 6.7. Also note that ε_4 depends only on m, p .

Let $m = 2$, and let p be the minimum integer such that $p > m = 2$ and $\sigma = m/p = 2/p \leq \sigma_*$. Since σ_* is a universal constant, the integers m, p are universal constants too. Then all the constants depending only on m, p now become universal constants. In particular, $M_0(\sigma)$ given by Proposition 4.1 is now a universal constant.

By Lemma 6.9, renaming s_j, \bar{s}_j the times t_j, \bar{t}_j in (4.2), renaming τ the period $T = T_a$, at $a = a_0$, in Proposition 4.1, renaming t_j, \bar{t}_j the times t_j^*, \bar{t}_j^* in (5.29), renaming I_j, E_j the intervals I_j^*, E_j^* in (5.29), renaming ε the product εc_0 where c_0 is defined in Lemma 6.9, renaming r_0 the constant $r_0/(2c_0^2)$ where c_0, r_0 are defined in Lemma 6.9, renaming b the ratio b/c_0^3 where c_0 is defined in Lemma 6.9 and b in (5.26), and also renaming u the solution \tilde{u} in Lemma 6.9, the proof of Theorem 1.1 is complete. ■

Acknowledgments. The authors wish to thank the anonymous referees for very useful suggestions.

Funding. P.B. and E.H. are supported by the Italian Project PRIN 2020XB3EFL *Hamiltonian and dispersive PDEs*. F.G. and E.H. have received funding from INdAM-GNAMPA, Project CUP_E55F22 000270001. M.G. is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 757802). M.G. is also supported by the Catalan Institution for Research and Advanced Studies via an ICREA Academia Prize 2019. This work is also supported by the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M).

References

- [1] A. Arosio and S. Panizzi, On the well-posedness of the Kirchhoff string. *Trans. Amer. Math. Soc.* **348** (1996), no. 1, 305–330 Zbl [0858.35083](#) MR [1333386](#)
- [2] P. Baldi, Periodic solutions of forced Kirchhoff equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **8** (2009), no. 1, 117–141 Zbl [1180.35040](#) MR [2512203](#)

- [3] P. Baldi and E. Haus, On the existence time for the Kirchhoff equation with periodic boundary conditions. *Nonlinearity* **33** (2020), no. 1, 196–223 Zbl [1437.35506](#) MR [4037808](#)
- [4] P. Baldi and E. Haus, On the normal form of the Kirchhoff equation. *J. Dynam. Differential Equations* **33** (2021), no. 3, 1203–1230 Zbl [1471.35208](#) MR [4299967](#)
- [5] P. Baldi and E. Haus, Longer lifespan for many solutions of the Kirchhoff equation. *SIAM J. Math. Anal.* **54** (2022), no. 1, 306–342 Zbl [7453677](#) MR [4359274](#)
- [6] S. N. Bernstein, Sur une classe d'équations fonctionnelles aux dérivées partielles. *Izv. Akad. Nauk SSSR Ser. Mat.* **4** (1940), 17–26 Zbl [0026.01901](#) MR [0002699](#)
- [7] J. Bourgain, Problems in Hamiltonian PDEs. GAFA 2000 (Tel Aviv, 1999). *Geom. Funct. Anal.*, Special Volume, Part I (2000), 32–56 Zbl [1050.35016](#) MR [1826248](#)
- [8] T. Buckmaster, P. Germain, Z. Hani, and J. Shatah, Onset of the wave turbulence description of the longtime behavior of the nonlinear Schrödinger equation. *Invent. Math.* **225** (2021), no. 3, 787–855 Zbl [1484.35345](#) MR [4296350](#)
- [9] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. *Invent. Math.* **181** (2010), no. 1, 39–113 Zbl [1197.35265](#) MR [2651381](#)
- [10] L. Corsi and R. Montalto, Quasi-periodic solutions for the forced Kirchhoff equation on \mathbb{T}^d . *Nonlinearity* **31** (2018), no. 11, 5075–5109 Zbl [1403.37083](#) MR [3867228](#)
- [11] R. W. Dickey, Infinite systems of nonlinear oscillation equations related to the string. *Proc. Amer. Math. Soc.* **23** (1969), 459–468 Zbl [0218.34015](#) MR [0247189](#)
- [12] M. Ghisi and M. Gobino, Global solutions to the Kirchhoff equation with spectral gap data in the energy space. 2022, arXiv:[2208.05400](#)
- [13] F. Giuliani, M. Guardia, P. Martin, and S. Pasquali, Chaotic-like transfers of energy in Hamiltonian PDEs. *Comm. Math. Phys.* **384** (2021), no. 2, 1227–1290 Zbl [1472.37079](#) MR [4259386](#)
- [14] B. Grébert, É. Paturel, and L. Thomann, Beating effects in cubic Schrödinger systems and growth of Sobolev norms. *Nonlinearity* **26** (2013), no. 5, 1361–1376 Zbl [1480.35352](#) MR [3056129](#)
- [15] B. Grébert and L. Thomann, Resonant dynamics for the quintic nonlinear Schrödinger equation. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **29** (2012), no. 3, 455–477 Zbl [1259.37045](#) MR [2926244](#)
- [16] B. Grébert and C. Villegas-Blas, On the energy exchange between resonant modes in nonlinear Schrödinger equations. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **28** (2011), no. 1, 127–134 Zbl [1216.35137](#) MR [2765514](#)
- [17] A. G. Greenhill, *The applications of elliptic functions*. Macmillan, London, 1892. Republication: Dover Publications, New York, 1959 Zbl [0087.08501](#) MR [0111864](#)
- [18] M. Guardia, Z. Hani, E. Haus, A. Maspero, and 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. (JEMS)* **25** (2023), no. 4, 1497–1551 Zbl [1518.35574](#) MR [4577970](#)
- [19] M. Guardia, E. Haus, and M. Procesi, Growth of Sobolev norms for the analytic NLS on \mathbb{T}^2 . *Adv. Math.* **301** (2016), 615–692 Zbl [1353.35260](#) MR [3539385](#)
- [20] M. Guardia and V. Kaloshin, Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation. *J. Eur. Math. Soc. (JEMS)* **17** (2015), no. 1, 71–149 Zbl [1311.35284](#) MR [3312404](#)
- [21] Z. Hani, B. Pausader, N. Tzvetkov, and N. Visciglia, Modified scattering for the cubic Schrödinger equation on product spaces and applications. *Forum Math. Pi* **3** (2015), article no. e4 Zbl [1326.35348](#) MR [3406826](#)

- [22] E. Haus and M. Procesi, KAM for beating solutions of the quintic NLS. *Comm. Math. Phys.* **354** (2017), no. 3, 1101–1132 Zbl [1375.35488](#) MR [3668616](#)
- [23] E. Haus and L. Thomann, Dynamics on resonant clusters for the quintic non linear Schrödinger equation. *Dyn. Partial Differ. Equ.* **10** (2013), no. 2, 157–169 Zbl [1346.35187](#) MR [3071705](#)
- [24] P. Holmes and J. Marsden, A partial differential equation with infinitely many periodic orbits: Chaotic oscillations of a forced beam. *Arch. Rational Mech. Anal.* **76** (1981), no. 2, 135–165 Zbl [0507.58031](#) MR [0629701](#)
- [25] G. Kirchhoff, *Vorlesungen über mathematische Physik: Mechanik*. Ch. 29, Teubner, Leipzig, 1876
- [26] S. Klainerman and A. Majda, Formation of singularities for wave equations including the nonlinear vibrating string. *Comm. Pure Appl. Math.* **33** (1980), no. 3, 241–263 Zbl [0443.35040](#) MR [0562736](#)
- [27] Y. Li, Smale horseshoes and symbolic dynamics in perturbed nonlinear Schrödinger equations. *J. Nonlinear Sci.* **9** (1999), no. 4, 363–415 Zbl [0960.37031](#) MR [1700670](#)
- [28] V. K. Melnikov, On the stability of a center for time-periodic perturbations. *Trans. Moscow Math. Soc.* **12** (1963), 1–57 MR [0156048](#)
- [29] J. Moser, *Stable and random motions in dynamical systems*. Ann. of Math. Stud. 77, Princeton University Press, Princeton, New Jersey, 1973. Reprinted: Princeton Landmarks Math., Princeton University Press, Princeton, NJ, 2001 Zbl [0991.70002](#) MR [1829194](#)
- [30] C. Robinson, *Dynamical systems. Stability, symbolic dynamics, and chaos*. Stud. Adv. Math., CRC Press, Boca Raton, Florida, 1999 Zbl [0914.58021](#) MR [1792240](#)

Received 18 March 2023; revised 30 August 2023; accepted 2 November 2023.

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

Filippo Giuliani

Dipartimento di Matematica – MUR Excellence Department 2023-2027, Politecnico di Milano, Piazza Leonardo da Vinci, 32, Campus Bonardi, 20133 Milano, Italy; filippo.giuliani@polimi.it

Marcel Guardia

Departament de Matemàtiques i Informàtica, Universitat de Barcelona, Gran Via, 585, 08007 Barcelona; Centre de Recerca Matemàtica, Campus de Bellaterra, Edifici C, 08193 Bellaterra, Spain; guardia@ub.edu

Emanuele Haus

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