AN AVERAGING THEOREM FOR QUASILINEAR HAMILTONIAN PDEs

Dario BAMBUSI

Dipartimento di Matematica,
Università degli studi di Milano
Via Saldini 50, 20133 MILANO, Italy.

Abstract. We study the dynamics of Hamiltonian quasilinear PDEs close to elliptic equilibria. Under a suitable nonresonance condition we prove an averaging theorem according to which any solution corresponding to small amplitude smooth initial data remain very close to a torus up to long times. An application to quasilinear wave equations in an n dimensional paralleliped is given.

1. Introduction

In this paper we study the dynamics of Hamiltonian partial differential equation of quasilinear type (in the sense of Kato [1,2]). In particular we concentrate on small amplitude solutions. Assuming (a) that the frequencies of small oscillation fulfill a new nonresonance condition that is satisfied in very general cases and (b) a natural estimate on the Lyapunov exponents of the system, we will construct infinitely many approximate integrals of motion. We will deduce that solutions corresponding to "smooth" initial data remain $\mathcal{O}(\epsilon^M)$ close to a finite dimensional torus up to times $\mathcal{O}(\epsilon^{-1})$, ϵ being the norm of the initial datum and M an arbitrary integer.

The abstract theorem is then applied to a quasilinear wave equation in an n-dimensional parallelepiped (see eq. (3.1) below). In particular we show that the abstract nonresonance condition is fulfilled for almost all values of the parameters and we use Kato's theory with the addition of quantitative estimates of the constants in order to verify the assumption on the Lyapunov exponent.

The proof of the abstract theorem is obtained in some steps. The first one consists in making a Galerkin cutoff, namely in approximating the infinite dimensional Hamiltonian system by a finite dimensional one. The second consists in applying to such a finite dimensional system a classical recursive algorithm [3,4,5,6] of construction of integrals of motion. It is well known that such an algorithm gives rise to divergent series, but truncating the construction at a given order one obtaines objects (the "integrals") having a small Poisson Brackets with the (finite dimensional) Hamiltonian. Thirdly we show that the Poisson

Brackets of the "integrals" with the complete infinite dimensional Hamiltonian is very small in a small ball of a phase space of "smooth" functions. This is not enough to ensure that the integrals remain approximatively constant along the solutions; indeed one has still to ensure that the norm of the solution remains small up to the times we are interested in. This will be obtained by adding an assumption on the Lyapunov exponents of the system, an assumption which is automatic in semilinear models and that can be verified using Kato's theory in quasilinear models.

The construction of the approximate integrals is quite delicate since one has to keep into account the dependence on the dimension of the truncated system of all the constants entering in the estimates. This is obtained by using a precise computation of the constants due to Giorgilli [6].

As far as we know, the main result of the present paper constitutes the first averaging type theorem applicable to *quasilinear* equations, namely equations in which the nonlinearity contains as many derivatives as the linear part. We also emphasize that this is important in view of physical applications like Magnetohidrodynamics or Elastodynamics, where the equations are quasilinear.

The main limitation of the result of the present paper rests in its time of validity. Indeed known results for finite dimensional systems and for some semilinear partial differential equations (see [7, 8, 9, 10]) describe the dynamics over a time scale much longer than $O(\epsilon^{-1})$, precisely over a time scale of order ϵ^{-M} with arbitrary M. We point out that such results apply only to equations in which the frequencies of small oscillation fulfill a nonresonant condition stronger than that of the present paper (such a stronger condition is only exceptionally satisfied in more than one space dimensions). A further result applicable to quite general *semilinear* models in arbitrary space dimension was obtained by Kuksin [11], but the time scale covered by that paper is of order $O(\epsilon^{-1})$. Moreover Kuksin's result allows to describe only the dynamics corresponding to initial data in which the energy is initially essentially concentrated on a finite (and fixed) number of linear modes.

Finally we recall the papers [12,13,14,15] where Galerkin truncations have been used together with averaging techniques in order to study the dynamics of nonlinear PDEs. In particular the paper [15] had a great influence on the present paper. Related results are also those of [16,17,18,19].

Plan of the paper. In sect. 2 we state the main results of the paper. In Sect. 3 we give the application to the nonlinear wave equation. In sect. 4 we recall the algorithm of construction of the integrals of motion in finite dimensional Hamiltonian systems. In sect. 5 we prove the main result of the paper. In sect. 6 we show that the diophantine condition we use is fulfilled with probability one in the wave equation. In sect. 7 we use Kato's theory to verify the assumptions on Lyapunov exponents in the case of nonlinear wave equation. Finally in the appendix we prove a result on the geometry of level surfaces of the approximate integrals.

Acknowledgement. I thank Antonio Giorgilli for some interesting discussions and for explaining me the key ideas leading to the proof of proposition 5.4. I also thank Alberto Arosio for some discussions on Kato's theory. This work was developed under the partial support of Gruppo Nazionale di Fisica Matematica (INdAM).

2. Abstract results

2.1 The Approximate Integrals

Consider a Hamiltonian system of the form

$$H(p,q)=\sum_{j\geq 1}\omega_j\frac{p_j^2+q_j^2}{2}+P(p,q) \tag{2.1}$$

where $\{\omega_j\}$ is a sequence of strictly positive quantities and (p_j,q_j) are conjugated variables. We will denote by $z\equiv\{(p_j,q_j)\}_{j\geq 1}$ a phase point and by $X_H(z)\equiv\left(-\frac{\partial H}{\partial q_j}(z),\frac{\partial H}{\partial p_j}(z)\right)$ the (formal) Hamiltonian vector field of a function H.

To define precisely the phase space consider the Hilbert space ℓ^2_s of the sequences $\{x_j\}_{j\geq 1}$ such that

$$\|x\|^2_s:=\sum_{j\geq 1}j^{2s}x_j^2<\infty\ ,$$

and denote $\mathcal{P}_s:=\ell^2_s\times\ell^2_s$. If $z\equiv(p,q)\in\mathcal{P}_s$ we will denote by

$$\|z\|_s:=\sqrt{\|p\|^2_s+\|q\|^2_s}\ ,$$

its norm and by $B_s(R)$ the open ball of radius R in $\mathcal{P}_s$.

For any positive (large) N denote by $\omega^{(N)}:=(\omega_1,...,\omega_N)$ the truncation of length N of the frequency vector. Having fixed a positive (large) r_*, we assume:

H1) (Nonresonance) There exist α and $\gamma>0$ such that for any N one has

$$|\omega^{(N)}\cdot k|\geq\frac{\gamma}{N^\alpha}\ ,\ \forall k\in\mathbf{Z}^N\ ,\ 0<|k|\leq r_*+2$$

where $|k|:=|k_1|+...+|k_N|$. □

Remark 2.1. *Consider the case where*

$$\omega_j=j^{d_1}+\frac{\sigma_j}{j^{d_2}}\ ,\quad \sigma_j\in[0,1]\ ,\quad d_1,d_2\geq 0$$

one can show that H1 is fulfilled if the parameters σ_j are chosen in a set of probability one with respect to the product measure.

To ensure that

$$H_0(p,q):=\sum_{j\geq 1}\omega_j\frac{p_j^2+q_j^2}{2} \tag{2.2}$$

describes the linearization of the system at the origin we assume

H2) There exist s_* and a neighbourhood of the origin $\mathcal{U}\subset\mathcal{P}_{s_*}$, a positive C and $\nu\geq 1$ such that

$$|P(z)|\leq C\,\|z\|^{\nu+2}_{s_*}\ ,\quad \forall z\in\mathcal{U}.$$

$\square$

Then we need some smoothness of the Hamiltonian and of its vector field, so we assume that

H3) One has $H \in C^{r_*+3}(\mathcal{U})$. $\square$

H4) There exists a real d with the following property: for any positive $s \geq s_*$ there exists a neighbourhood of the origin $\mathcal{U}_{s+d} \subset \mathcal{P}_{s+d}$ such that $X_H \in C^{r_*+2}(\mathcal{U}_{s+d}, \mathcal{P}_s)$. $\square$

Remark 2.2. *Actually it is enough that the above assumptions are fulfilled for an infinite unbounded set of values of s.*

In order to ensure that the finite dimensional algorithm of construction of the integrals of motion is applicable we assume that

H5) The nonlinearity is reversible, namely one has $P(-p, q) = P(p, q)$. $\square$

Denote the linear actions by $I_j := (p_j^2 + q_j^2)/2$. In the following we will always denote by C a positive (usually large) constant whose value can change from line to line.

Theorem 2.3. *Fix $M \geq 4$, and assume that H1-H5 hold with some $r_* > \frac{4}{3}M - 3$ then there exist positive constants s', ϵ_*, and β with the following properties: for any $0 < \epsilon < \epsilon_*$ define*

$$N := [\epsilon^{-\frac{1}{4\beta}}] \ ,$$

then

1) there exist N approximate integrals of motion $\{\Phi^{(j)}\}_{j=1}^N$, which are analytic in $B_{s_*}(\epsilon_*)$, and fulfill

$$\begin{aligned} \sup_{z \in B_{s_*+s'}(\epsilon)} \left| \left\{ H, \Phi^{(j)} \right\} (z) \right| &\leq C\epsilon^M \\ \sup_{z \in B_{s_*}(\epsilon)} \left| \Phi^{(j)}(z) - I_j(z) \right| &\leq C\epsilon^2 \epsilon^{1/4} \end{aligned} \tag{2.3}$$

2) For any $s \geq s_*$ there exits ϵ_s and C_s such that one has

$$\sup_{z \in B_{s+s'}(\epsilon)} \left| \left\{ H, I^{(s,N)} \right\} (z) \right| \leq C_s \epsilon^M \ , \quad \forall \epsilon < \epsilon_s$$

where $I^{(s,N)} := \sum_{j>N} j^{2s} I_j$.

Remark 2.4. *In the C^∞ completely nonresonant case M can be chosen arbitrarily large; then the quantity s' turns out to be M dependent.*

If we assume also local well posedness of the dynamics then we can add some dynamical properties of the solutions. So, consider the Hamilton equations of motions

$$\dot{z} = X_H(z) \ , \tag{2.4}$$

and assume:

H6) (Local well posedness) $\forall s \geq s_*$ there exits $\epsilon_{s+d} > 0$ with the following property: $\forall z_0 \in B_{s+d}(\epsilon_{s+d})$ there exist $T > 0$ and a unique solution

$$z \in C^0\left([0,T]; \mathcal{P}_{s+d}\right) \cap C^1([0,T], \mathcal{P}_s)$$

of (2.4) with $z_0 = z(0)$. $\square$

Theorem 2.5. *Under the same assumptions of theorem 2.3, assume also H6, then there exists ϵ_s^* with the following properties: Let $z_0 \in \mathcal{B}_{s+s'}(\epsilon)$ with $\epsilon < \epsilon_s^*$, and denote by $z(t)$ the solution of (2.4) with initial datum z_0. Denote also*

$$\mathbf{T}_{z_0} := \left\{ w \in \mathcal{P}_s \ : \ \Phi^{(j)}(w) = \Phi^{(j)}(z_0) \ , \quad j = 1, ..., N \ , I_j(w) = 0 \ \forall j > N \right\}$$

and let $\tau_\epsilon(z_0)$ be the escape time of the solution from $B_{s+s'}(\epsilon)$, namely the infimum of the times for which $z(t) \notin B_{s+s'}(\epsilon)$, then one has

$$\begin{aligned} |\Phi^{(j)}(t) - \Phi^{(j)}(0)| &\leq C\epsilon^{M_1} \ , \ j = 1, ..., N \\ \left| I^{(s,N)}(t) - I^{(s,N)}(0) \right| &\leq C\epsilon^{M_1} \\ d_s(z(t), \mathbf{T}_{z_0}) &\leq C\epsilon^{M_2} \end{aligned} \tag{2.5}$$

for all the times t such that

$$|t| \leq \min \left\{ \frac{1}{C\epsilon^{M_3}}, \tau_\epsilon(z_0) \right\}$$

where M_1, M_2, M_3 are such that $M_1 + M_2 \leq M$, $2M_2 + M_3 + 2\frac{s+1}{\beta} \leq M$, and $d_s(\cdot;\cdot)$ denotes the distance in $\mathcal{P}_s$.

2.2 Quasilinear Equations

We are now going to bound from below the escape time in the case of quasilinear systems.

We begin with some notations and definitions. Let X and Y be Hilbert spaces, with Y densely embedded in X. The space of bounded linear operators from Y to X will be denoted by $\mathcal{B}(Y, X)$.

We assume the Hamilton equations of (2.1) to be quasilinear, i.e. that

K1) (Quasilinearity) there exists s_*, and, for any $s \geq s_*$ a positive R_{s+d}, a map

$$B_{s+d}(R_{s+d}) \ni z \mapsto A(z) \in \mathcal{B}(\mathcal{P}_{s+d}, \mathcal{P}_s)$$

and a map $g : B_{s+d}(R_{s+d}) \to \mathcal{P}_{s+d}$ such that the Hamilton equations of H take the form

$$\dot{z} = A(z)z + g(z) \equiv X_H(z) \ . \tag{2.6}$$

$\square$

The key assumption concerns the properties of the linearized flow. For any function $\zeta \in C^0([0,T],\mathcal{P}_{s+d}) \cap C^1([0,T],\mathcal{P}_s)$ we will consider the family of linear operators $A^{\zeta}(t) := A(\zeta(t))$ and the corresponding linear time dependent differential equation

$$\dot{z} = A^{\zeta}(t)z \ . \tag{2.7}$$

We recall that, under suitable assumptions there exists a corresponding evolution operator $U(t,s)$, strongly continuous as a map in $\mathcal{B}(X,X)$ and also as a map in $\mathcal{B}(Y,Y)$. The operator $U(t,s)$ is defined by the property that the solution of (2.7) fulfilling $z(s) = z_0$ is given by $z(t) = U(t,s)z_0$.

We assume:

K2) (Linear estimate) There exists $\nu \geq 1$ such that, for any $s \geq s_*$, any $\epsilon > 0$, any $T > 0$ and any function $\zeta \in C^0([0,T],\mathcal{P}_{s+d}) \cap C^1([0,T],\mathcal{P}_s)$ fulfilling

$$\sup_{t\in[0,T]} \|\zeta(t)\|_{s+d} + \sup_{t\in[0,T]} \left\|\dot{\zeta}(t)\right\|_s \leq \epsilon \tag{2.8}$$

the evolution operator $U(t,s)$ associated to equation (2.7) exists and fulfills the estimate

$$\sup_{0\leq t\leq\tau\leq T} \|U(t,\tau)\|_{\ell^2_{s+d}\to\ell^2_{s+d}} \leq M e^{\beta\epsilon^{\nu}T} \ , \tag{2.9}$$

with some constants M,β independent of ζ, T, ϵ. $\square$

K3) g is of class $C^1(\mathcal{U}_{s+d},\mathcal{P}_{s+d})$ and has a zero of order $\nu+1$ at the origin, namely

$$\|g(z)\|_s \leq C_s \|z\|_s^{\nu+1} \ .$$

$\square$

Remark 2.6. *In the semilinear case where the linear operator A is independent of z, one has $A = X_{H_0}$. It follows that assumption K2 is automatically true.*

Remark 2.7. *In the true quasilinear case one can use Kato's theory[1,2] to verify assumptions K2. This is what we will do in the application to the nonlinear wave equation.*

Theorem 2.8. *Under the same assumption of theorem 2.3, assume also K1-K3, then there exists s' and, for any $s \geq s_*$ there exists a constant ϵ_s such that, if the initial datum z_0 fulfills*

$$\epsilon := \|z_0\|_{s+s'} \leq \epsilon_s \ , \tag{2.10}$$

then along the corresponding solution one has

$$\begin{aligned} |\Phi^{(j)}(t) - \Phi^{(j)}(0)| &\leq C\epsilon^{M_1} \ , \ j = 1,...,N \\ \left|I^{(s,N)}(t) - I^{(s,N)}(0)\right| &\leq C\epsilon^{M_1} \\ d_s(z(t),\mathbf{T}_{z_0}) &\leq C\epsilon^{M_2} \end{aligned} \tag{2.11}$$

for all times t fulfilling

$$|t| \leq \frac{1}{C\epsilon^{\nu}} \tag{2.12}$$

where $M_1 = M - \nu$, $M_2 = \frac{M-\nu}{2} - \frac{s+1}{\beta}$.

Remark 2.9. *The strongest limitation of the present result rests in its time of validity. Indeed, in the finite dimensional case and also in the case of semilinear equations in one space dimension* [8]*, a similar description of the dynamics has been proved to hold up to times of order ϵ^{-M} for any M. Due to results of the kind of [20] we expect the time (2.12) to be optimal for quasilinear models. However the nonresonance condition H1 is not satisfied by the model considered in [20], so it is not a counterexample to the result of the present paper. We think it would be very interesting to find an example satisfying the assumptions of theorem 2.8 but developing singularities over a the time scale $\epsilon^{-\nu}$.*

Remark 2.10. *The theory of [8] can be extended to some semilinear equations in higher space dimension (see [10]) like the Nonlinear Schrödinger equation thus extending the validity time of the estimates (2.11). The main property needed to obtain such an extension is a nonresonance condition stronger than H1. However such a condition is not fulfilled in general high dimensional examples like the nonlinear wave equation.*

3. Application to quasilinear wave equations

Having fixed a diophantine vector $a = (a_1, ..., a_n) \in \mathbb{R}^n$ with $a_i > 0$[1], consider the n dimensional parallelepiped $\mathcal{R}$ with sides of length $\pi/\sqrt{a_i}$, namely

$$\mathcal{R} := \left\{ x \equiv (x_1, ..., x_n) \in \mathbb{R}^n \ : \ 0 < x_i < \frac{\pi}{\sqrt{a_i}} \right\}$$

and the nonlinear wave equation

$$\begin{aligned} &u_{tt} - \Delta u + mu - b_{ij}(u, \nabla u)\partial_i\partial_j u + g(u, \nabla u) = 0 \ , \quad x \in \mathcal{R} \\ &u\big|_{\partial\mathcal{R}} = 0 \ , \end{aligned} \tag{3.1}$$

where we used the summation convention for the indexes $i, j = 1, ..., n$, we denoted by $\nabla u \equiv (\partial_1 u, ..., \partial_n u)$ the derivatives of u with respect to the space variables, and b_{ij}, g are

[1] We recall that it means that there exist a positive $\bar\gamma$ and a real $\bar\tau$ such that

$$|a_1 k_1 + ... + a_n k_n| \equiv |a \cdot k| \geq \frac{\gamma_1}{|k|^{\bar\tau}} \ , \quad \forall k \in \mathbf{Z}^n \setminus \{0\}$$

and that diophantine vectors form a set of full measure in $\mathbb{R}^n$.

functions of class C^∞. Moreover we assume that the system is Hamiltonian, i.e. that there exists a C^∞ function $W = W(u, s_1, ..., s_n)$ such that

$$b_{ij}(u, s_1, ..., s_n) = \frac{\partial^2 W}{\partial s_i \partial s_j}(u, s_1, ..., s_n) \ , \quad g = \frac{\partial W}{\partial u} - \frac{\partial^2 W}{\partial u \partial s_k}\partial_k u \ ,$$

so that the Hamiltonian function of the system is given by

$$H(u, v) = \int_{\mathcal{R}} \left[\frac{v^2}{2} + \frac{\|\nabla u\|^2}{2} + \frac{mu^2}{2} + W(u, \nabla u) \right] d^n x \ , \tag{3.2}$$

where $v = \dot{u}$ is the momentum conjugated to u, and we denoted by $\| \, . \, \|$ the Euclidean norm of a vector of $\mathbb{R}^n$. Assume that W has a zero of order $\nu + 2$ at the origin, namely that

$$|W(u, s)| \leq C \left(|u| + \|s\| \right)^{\nu+2} \quad , \quad \nu \geq 1 \ .$$

Expand u in Fourier series in the space variable, namely write

$$u(x, t) = \sum_{\mathbf{j} \equiv (j_1, \ldots, j_n)} u_{\mathbf{j}}(t) \prod_{i=1}^{n} a_i^{1/4} \sin(\sqrt{a_i} j_i x_i) \ , \tag{3.3}$$

with

$$\mathbf{j} \equiv (j_1, ..., j_n) \in \bar{\mathbf{N}}^n := \left\{ \mathbf{k} \equiv (k_1, ..., k_n) \in \bar{\mathbf{Z}}^n \ : \ k_i \geq 1 \ , \ i = 1, ..., n \right\} \ ,$$

so that (3.1) is converted into the infinite system

$$\ddot{u}_{\mathbf{j}} + \omega_{\mathbf{j}}^2 u_{\mathbf{j}} = f_{\mathbf{j}}(u) \ , \ \mathbf{j} \in \bar{\mathbf{N}}^n$$

where f is of order at least $\nu + 1$ in u, and

$$\omega_{\mathbf{j}} := \sqrt{\mu_{\mathbf{j}} + m} \ , \quad \mu_{\mathbf{j}} := a_1 j_1^2 + a_2 j_2^2 + ... + a_n j_n^2 \ . \tag{3.4}$$

We come to the nonlinear problem. To fit the abstract scheme *we enumerate the eigenvalues $\mu_{\mathbf{j}}$ of the Laplacian by integers indexes $j \in \mathbf{N}$ in such a way that the μ_j's form an increasing sequence.* This establishes a 1 to 1 correspondence between $\bar{\mathbf{N}}^n$ and $\mathbf{N}$. Define $p_j := \sqrt{\omega_j}\dot{u}_j$, $q_j := u_j/\sqrt{\omega_j}$. Then it is easy to see that the space $\mathcal{P}_s$ is isomorphic to the space $\mathcal{F}_s$ of the functions $(u, v) \in H^{\tilde{s}+1} \oplus H^{\tilde{s}}$ with $\tilde{s} = 2s/n$, fulfilling the compatibility conditions

$$\begin{aligned} (-\Delta)^j u\big|_{\partial\Omega} = 0 \ , \quad 0 \leq j \leq \left[\frac{\tilde{s}}{2} \right] \ , \\ (-\Delta)^j v\big|_{\partial\Omega} = 0 \ , \quad 0 \leq j \leq \left[\frac{\tilde{s}+1}{2} \right] - 1 \ . \end{aligned} \tag{3.5}$$

We will consider only values of s such that $\tilde{s}$ is integer; in view of remark 2.2 this is possible.

Concerning the validity of the diophantine type condition H1 we have the following

Theorem 3.1. *Fix $b > 0$, then there exists a subset $\mathcal{J} \subset [0, b]$ of measure b such that, if $m \in \mathcal{J}$, then the frequencies ω_j fulfill the assumption $H1$.*

This theorem will be proved in section 6.

Assumptions H2 and H3 are a trivial consequence of Sobolev embedding theorems, and also H6 is automatically true.

To fulfill the smoothness assumptions of the vector field we assume that the potential W is even in each of the arguments, namely

$$W((-)^{c_0}u, (-)^{c_1}s_1, ..., (-)^{c_n}s_n) = W(u, s_1, ..., s_n) \ , \quad \forall (c_0, c_1, ..., c_n) \in (Z_2)^{n+1} \ , \tag{3.6}$$

from which in particular it follows $\nu \geq 2$.

To verify assumptions H4,H6 and K1-K3 it is convenient to proceed as follows: First rescale the domain in order to transform it into a standard cube with sides of length π; then identify the space $\mathcal{F}_s$ with the space $\mathcal{A}$ of the functions of class $H^{\tilde{s}+1}(\mathbf{T}^n) \times H^{\tilde{s}}(\mathbf{T}^n)$) (where $\mathbf{T} = \mathbb{R}/2\pi\mathbf{Z}$) which are skewsymmetric in each variable, namely that fulfill

$$u((-)^{c_1}x_1, ..., (-)^{c_n}x_n) = (-)^{c_1+...+c_n}u(x_1, ..., x_n) \ .$$

Clearly one has $\mathcal{F}_s \simeq \mathcal{A} \cap (H^{\tilde{s}+1}(\mathbf{T}^n) \times H^{\tilde{s}}(\mathbf{T}^n))$. Due to the symmetry properties that (3.6) induces on the nonlinearity, the subspace $\mathcal{A}$ is invariant under the dynamics of (3.1) in $\mathbf{T}^n$. So the validity of H4 is a simple consequence of Sobolev inequalities. The same is true for K1 and K3. Local existence and uniqueness (i.e. H6) was proven by Kato (see [2] theorem 14.3).

Theorem 3.2. *Assume $W \in C^\infty$ and (3.6), then property K2 holds.*

For the proof see sect.7.

So the general theory applies and in particular one has

Theorem 3.3. *For any positive large M there exists a set $\mathcal{J} \subset [0, b]$ of measure b such that, for any $m \in \mathcal{J}$ there exists a positive constant s', and for any $s \geq s_*$ a constant ϵ_s, such that if the initial datum $z_0 \equiv (u_0, v_0)$ fulfills*

$$\epsilon := \|z_0\|_{s+s'} \leq \epsilon_s \tag{3.7}$$

then along the corresponding solution of (3.1) one has

$$d_s(z(t), \mathbf{T}_{z_0}) \leq C\epsilon^M \quad \text{for} \quad |t| \leq \frac{1}{C\epsilon^\nu} \tag{3.8}$$

where $\mathbf{T}_{z_0}$ is a finite dimensional smooth torus.

We conclude by some remarks on the possibility of extending the result to more general domains; to this and there are two kinds of difficulties.

The first one is related to the verification of the nonresonance condition. Following the proof of theorem 3.1 one can show that if the Dirichlet eigenvalues μ_j of the Laplacian in the considered domain fulfill the gap estimate

$$\mu_{j+1} - \mu_j > C/j^\delta$$

with a given δ, then the nonresonance condition H1 is fulfilled for most values of the parameter m. However we do not know whether there exist domains different form the parallelepiped in which the gap condition is fulfilled.

The second point concerns smoothness of the solutions. Indeed it is well known that there are some compatibility conditions which are necessary and sufficient for the smoothness of solutions of the wave equation (which in turn is essential for our approach). We have no clear ideas on how to deal with such conditions in the case of general domains.

4. The classical algorithm

Here we recall the classical algorithm of construction of the approximate integrals of motion and a result by Giorgilli [6] that we need.

Consider a Hamiltonian system with N degrees of freedom, of the form

$$H = \sum_{k\geq 0} H_k \tag{4.1}$$

where H_k is a homogeneous polynomial of degree $k+2$ in the phase variables and H_0 is of the form (2.2). Let $I_j = (p_j^2 + q_j^2)/2$ be an action of the linear system, we aim to construct an integral of motion of the form $\Phi^{(j)} := I_j + \sum_{l\geq 1} \Phi_l$ with Φ_l a homogeneous polynomial of degree $l+2$. Clearly $\Phi^{(j)}$ has to fulfill the equation $\left\{H, \Phi^{(j)}\right\} = 0$. Inserting the Taylor expansions of H and of Φ and equating terms of the same degree we get the recursive system

$$L_{H_0}\Phi_l = \Psi_l \ , \ l \geq 1 \ , \quad L_{H_0} := \{H_0, \, . \, \} \tag{4.2}$$

where Ψ_l is determined by

$$\begin{aligned} \Psi_1 &= \{I, H_1\} \\ \Psi_l &= \sum_{j=1}^{l-1} \{\Phi_j, H_{l-j}\} + \{I, H_l\} \end{aligned} \tag{4.3}$$

So, provided one is able to solve the homological equation (4.2), the formal integral of motion can be computed.

In order to solve the homological equation we remark that at the l-th step (4.2) appears as a linear equation in the finite dimensional space of the polynomials of degree $l+2$. So, if one is able to diagonalize the linear operator L_{H_0}, then equation (4.2) is easily solved. To diagonalize L_{H_0} introduce the variables

$$\xi_j := \frac{1}{\sqrt{2}}\,(q_j - \mathrm{i}p_j) \ , \quad \eta_j := \frac{1}{\sqrt{2}}\,(q_j + \mathrm{i}p_j) \ , \tag{4.4}$$

which give the unperturbed Hamiltonian the form $H_0 = \sum \omega_l \xi_l \eta_l$ (and the symplectic form becomes $\mathrm{i} d\xi_l \wedge d\eta_l$), so that one has

$$L_{H_0} \xi^j \eta^k = [\mathrm{i}(k-j)\cdot\omega] \ \xi^j \eta^k \ , \quad \xi^j \eta^k \equiv \xi_1^{j_1} ... \xi_N^{j_N} \eta_1^{k_1} ... \eta_N^{k_N}$$

i.e. the monomial $\xi^j \eta^k$ form the basis of the space of polynomials on which L_{H_0} is diagonal. So, (4.2) is easily solvable provided Ψ_l has no component on the kernel of L_{H_0}, a fact that can be easily verified recursively* in the case where $H_k(p,q)$ is even in p.

As is well known the series defining the integral of motion are in general divergent, so, following [6], one has to stop the construction at a given order and to estimate the Poisson bracket of the so obtained function with the Hamiltonian. To this end remark that, defining

$$\Phi^{(j,r_*)} := I_j + \sum_{l=1}^{r_*} \Phi_l \ ,$$

one has

$$\dot{\Phi}^{(j,r_*)} = \left\{ H, \Phi^{(j,r_*)} \right\} = O(|(p,q)|^{r_*+3}) \tag{4.5}$$

where we denoted for simplicity $\Phi_0 := I_j$. Finally, in the finite dimensional case, one can obtain a long time estimate of the solution repeating the construction for all j's and then using the approximate constant of motion as Lyapunov functions in order to control that the domain of construction of such functions is not left by the solution up to the times one is interested in.

For the application to infinite dimensional systems we have to keep into account the dependence of all the constants on the number N of degrees of freedom. To this end we recall a quantitative result by Giorgilli [6]. We first have to introduce some notations.

Fix $R \equiv (R_1, ..., R_N)$ with $R_j > 0$ and consider a polynomial $f(p,q) = \sum_{j,l} f_{jl} p^j q^l$, then the size of f is measured by the norm

$$\|f\|_R := \sum_{i,l} |f_{jl}| R^{j+l} \ .$$

To control the small denominators assume that there exists a non-increasing sequence $\{\alpha_s\}_{s=1}^{r_*}$ such that

$$|\omega \cdot k| \geq \alpha_s \text{ for } k \in \mathbf{Z}^N \ , \ 0 < |k| \leq s+2 \ . \tag{4.6}$$

Finally consider the complex neighbourhood of the origin

$$\Delta_{\rho R} := \left\{ (p,q) \in \mathbf{C}^{2N} \ : \ (|p_l|^2 + |q_l|^2)^{1/2} \leq \rho R_l \right\}$$

and denote

$$\Lambda := (\min_l R_l)^{-1} \ .$$

---

* Indeed, a component on the Kernel of L_{H_0} is always even in the momenta p, while from the construction of H_k such a function always turns out to be either independent of p or skewsymmetric in it (see [6]).

Theorem 4.1. *[Giorgilli] Consider a Hamiltonian system of the form (4.1), even in the momenta p, and assume that, for a given $R \in \mathbb{R}_+^N$ there exist constants $h > 0$ and $E > 0$ such that $\|H_k\|_R \leq h^{k-1}E$ for $k \geq 1$; assume that the linear frequencies fulfill (4.6). Then, for any integer $0 < r \leq r_*$ there exist N truncated integrals $\Phi^{(l,r)} = I_l + \sum_{s=1}^r \Phi_s^{(l)}$ such that $\left\{H, \Phi^{(l,r)}\right\}$ is a power series starting with terms of degree al least $r+3$. Moreover, for $z \in \Delta_{\rho R}$ and $\rho < 1/h$, one has the bounds*

$$
\begin{aligned}
\left|\left(\Phi^{(l,r)} - I_l\right)(z)\right| &< \frac{24E}{\alpha_1}\rho^3[1-(\sigma_r\rho)^r](1-\sigma_r\rho)^{-1} \\
\left|\left\{H, \Phi^{(l,r)}\right\}(z)\right| &< C_r\rho^{r+3}(1-h\rho)^{-2} \ ,
\end{aligned}
\tag{4.7}
$$

where

$$
\begin{aligned}
\sigma_1 &= 1 \\
\sigma_r &= \left(12\Lambda^2 E + \frac{8}{9}h\alpha_1\right)\left[\frac{(r+1)!}{\prod_{l=2}^r \alpha_l}\right]^{\frac{1}{r-1}} \ , \ r > 1 \ , \\
C_r &= 8E\left(12\Lambda^2 E + \frac{8}{9}h\alpha_1\right)^r \frac{(r+2)!}{\prod_{l=1}^r \alpha_l} \ , \ r \geq 1 \ .
\end{aligned}
\tag{4.8}
$$

5. Proof of the abstract results.

Expand the perturbation P in Taylor series up to order $r_* + 2$,

$$
P = \sum_{l=1}^{r_*} P_l + \mathcal{R}_*
$$

where P_l are homogeneous polynomial of degree $l+2$ and $\mathcal{R}_*$ is the remainder. We will denote

$$
H_* := H_0 + \sum_{l=1}^{r_*} P_l
$$

Remark 5.1. *The polynomials P_l are analytic functions, and therefore, introducing the complexification $\mathcal{P}_{s_*}^{\mathbf{C}}$ of $\mathcal{P}_{s_*}$, one has that there exist constants R_*, C such that*

$$
|P_l(z)| \leq C\|z\|_{s_*}^{l+2} \ , \ \forall z \in \mathcal{P}_{s_*}^{\mathbf{C}} \quad \|z\|_{s_*} \leq R_* \ .
$$

To make the Galerkin cutoff fix a large N that will eventually be related to ϵ, introduce the projector Π_N defined by

$$\Pi_N(p,q) = (p_1, ..., p_N, q_1, ..., q_N) \ , \tag{5.1}$$

and put $H^{(N)} := H_0^{(N)} + \sum_{l=1}^{r_*} H_l$, with

$$H_0^{(N)} := H_0 \circ \Pi_N \ , \quad H_l := P_l \circ \Pi_N \ .$$

One has

$$H = H^{(N)} + (H_* - H^{(N)}) + \mathcal{R}_* \ .$$

Lemma 5.2. *For any $s \geq s_*$ there exists a constant C and a domain $\mathcal{U}_s$ such that, for any $r \geq 0$, and any $N \geq 0$, one has*

$$\begin{aligned} \|X_{\mathcal{R}_*}(z)\|_s \leq C \|z\|_{s+d}^{r_*+2} \ , \quad \forall z \in \mathcal{U}_{s+d} \\ \left\|X_{H_*-H^{(N)}}(z)\right\|_s \leq \frac{C}{N^r} \|z\|_{s+r+d} \ , \quad \forall z \in \mathcal{U}_{s+d+r} \ . \end{aligned} \tag{5.2}$$

Proof. First remark that

$$\| \mathbb{1} - \Pi_N\|_{s+r\to s} = \frac{1}{N^r} \ . \tag{5.3}$$

and that

$$X_{H_*-H^{(N)}}(z) = [X_{H_*}(z) - X_{H_*}(\Pi_N z)] + (\mathbb{1} - \Pi_N) X_{H_*}(\Pi_N z) \ ,$$

By H4, using the norm of the differential to estimate the Lipschitz constant of X_{H_*}, the first square bracket is estimated by

$$\sup_{z\in\mathcal{U}_{s+d}} \|dX_{H_*}(z)\|_{s+d\to s} \| \mathbb{1} - \Pi_N\|_{s+d+r\to s+d} \|z\|_{s+d+r} \leq \frac{C}{N^r} \|z\|_{s+r+d}$$

The second term is trivially estimated using (5.3). The estimate of X_{R_*} is obtained by applying Lagrange estimate of the remainder of the Taylor expansion. $\square$

We apply now theorem 4.1 to $H^{(N)}$.

Lemma 5.3. *There exist N truncated analytic integrals $\Phi^{(l,r_*)} = I_l + \sum_{s=1}^{r_*} \Phi_s^{(l)}$ for the Hamiltonian system $H^{(N)}$. Moreover for any $z \in \mathcal{P}_{s_*}$ such that*

$$\|z\|_\infty := \sup_{k=1,\ldots,N} \{\sqrt{p_k^2 + q_k^2}\} \leq \frac{1}{CN^{\beta+\alpha}} \tag{5.4}$$

the following estimates hold

$$\begin{aligned} \left|\left(\Phi^{(l,r_*)} - I_l\right)(z)\right| < C(\|z\|_\infty N^a)^3 N^\alpha \\ \left|\left\{H^{(N)}, \Phi^{(l,r_*)}\right\}(z)\right| < C\left(\|z\|_\infty N^a\right)^3 \left(\|z\|_\infty N^\beta\right)^{r_*} \ . \end{aligned} \tag{5.5}$$

where $a = s_* + 2$, $\beta := a + \alpha$.

Proof. To compute the Giorgilli's norm of a function f remark that, if $f(z) = \sum_k f_k z^k$ (k being a mutiindex) one has

$$f_k = \frac{1}{k!}\left.\frac{\partial^{|k|} f}{\partial z^k}\right|_{z=0} \quad , \quad k! := k_1! k_2! ... k_N!$$

and that, defining $R_j^* := \mathcal{R}/j^{s_*+1}$, with a positive $\mathcal{R}$, one has

$$\Delta_{R^*} \subset B_{s_*}(2\mathcal{R}) \ ,$$

where we considered complex balls.

Applying Cauchy inequality to the polynomial $H_l(z) = \sum_k H_{l,k} z^k$ one gets

$$\begin{aligned} |H_{l,k}| &\leq \frac{1}{(R^*)^k} \sup_{\Delta_{R^*}} |H_l| \leq \frac{1}{(R^*)^k} \sup_{B_{s_*}(2\mathcal{R})} |H_l| \\ &\leq \frac{1}{(R^*)^k} C\mathcal{R}^{|k|} \leq C(N^{s_*+1})^{|k|} \end{aligned}$$

We apply Giorgilli's theorem in the domain Δ_R with $R_j := \mathcal{R}$ for all j's. Summing over the indices and taking into account that there are at most $(2N)^l$ distinct polynomials of degree l in $2N$ variables one gets

$$\|H_l\|_R \leq C(\mathcal{R}N^{s_*+2})^3(\mathcal{R}N^{s_*+2})^{l-1}$$

which allows to choose $E = C(\mathcal{R}N^{s_*+2})^3$ and $h = \mathcal{R}N^{s_*+2}$. Compute now the various constants of Giorgilli's theorem; provided $\alpha \geq 2a$, one has $\sigma_{r_*} = C\mathcal{R}N^{2\alpha+a}$ and $C_{r_*}(\mathcal{R}N^{a+\alpha})^{r_*}(\mathcal{R}N^a)^3$. Finally take $\rho := \|z\|_\infty / \mathcal{R}$. This gives the result. □

¿From now on we will always denote simply by $\Phi^{(l)}$ the approximate integrals $\Phi^{(l,r_*)}$.

To prove that the level surfaces of the approximate integrals are tori and that the distance from such tori is controlled by the difference between the values of the functions one has to pay some attention to the fact that the action I_l is singular on the hypersurface $p_l = q_l = 0$. We have the following

Proposition 5.4. *Assume* $\|z\|_\infty < \epsilon$ *and* $\|\bar{z}\|_\infty < \epsilon$ *with some*

$$\epsilon < \frac{1}{CN^{3s_*+6+\alpha}}$$

then there exits η_* *such that*

$$\left|\Phi^{(j)}(z) - \Phi^{(j)}(\bar{z})\right| < \epsilon^2\eta \ , \quad \forall j = 1, ..., N \ , \quad \Pi_N(z) = \Pi_N(\bar{z}) = 0$$

with $\eta < \eta_*$ *implies*

$$d_\infty(\bar{z}, \mathbf{T}_z) < C\epsilon\eta^{1/2} \ .$$

For the proof see the appendix.

Proof of theorem 2.3. Compute the Poisson bracket of $\Phi^{(l)}$ with the Hamiltonian: it is given by

$$\left\{H^{(N)},\Phi^{(l)}\right\}+\left\{H_*-H^{(N)},\Phi^{(l)}\right\}+\left\{\mathcal{R}_*,\Phi^{(l)}\right\}$$
$$=\left\{H^{(N)},\Phi^{(l)}\right\}+d\Phi^{(l)}X_{H_*-H^{(N)}}+d\Phi^{(l)}X_{\mathcal{R}_*}\ .$$

By (5.5), using Cauchy inequality one has, with a suitable definition of C,

$$\sup_{\left\|z\right\|_\infty\leq C/N^{\beta+\alpha}}\left\|d\Phi^{(l)}(z)\right\|\leq CN^{\beta+\alpha}\left\|z\right\|_\infty\ ,$$

from which, using lemma 5.2 and equation (5.2) one has

$$\left|\left\{H,\Phi^{(l)}\right\}(z)\right|\leq C\left(\left\|z\right\|_\infty N^\beta\right)^{r_*}\left(\left\|z\right\|_\infty N^a\right)^3$$
$$+\frac{C}{N^r}\left\|z\right\|_{s+r+d}N^{\beta+\alpha}\left\|z\right\|_{s_*}+C_s\left\|z\right\|_{s+d}^{r_*+2}N^\beta\left\|z\right\|_{s_*}$$

choose now N such that $N^\beta=\left\|z\right\|_{s_*}^{-1/4}$ (remark that the s_* norm controls the ∞ norm), then the argument of the first bracket is smaller than $\left\|z\right\|_{s_*}^{3/4}$, and the last term turns out to be much smaller than the first one. We choose now r in such a way that the first two terms are of the same order of magnitude. This leads to the choice

$$r:=3r_*\beta+2\beta+\alpha\text{ and }s':=r+d$$

□

Proof of theorem 2.5. The only thing to be proved is the last estimate (2.5). To obtain the estimate in the s norm remark that one has $\left\|\Pi_Nz\right\|_s\leq N^{s+1}\left\|\Pi_Nz\right\|_\infty$. □

We come now to the proof of theorem 2.8. To this end the key lemma is the following one

Lemma 5.5. *Assume K1-K3, then there exist C,ϵ_* such that for any initial datum fulfilling*

$$\left\|z_0\right\|_{s+d}\leq\epsilon/C\ ,\quad\epsilon<\epsilon_* \tag{5.6}$$

the existence time of the corresponding solution of (2.6) is larger than $1/(C\epsilon^\nu)$ and moreover one has the estimate

$$\sup_{0\leq t\leq 1/C\epsilon^\nu}\left\|z(t)\right\|_{s+d}\leq\epsilon \tag{5.7}$$

Proof. Let z be the solution of the Cauchy problem for (2.6), then by standard continuation argument it can be continued at least until

$$\left\|z(t)\right\|_{s+d}<\epsilon \tag{5.8}$$

holds. Let $\bar{T}$ be the first time at which (5.8) is violated, then one has $\left\|z(\bar{T})\right\|_{s+d}=\epsilon$. Denote $g^z(t):=g(z(t))$. So, $z(t)$ fulfills the "linear" equation

$$\dot{z}=A^z(t)z+g^z(t),\quad 0\leq t\leq\bar{T} \tag{5.9}$$

where the estimate (2.9) holds until time $\bar{T}$. By theorem 2 of [1] (which follows from the formula of variation of constants) the solution of (5.9) satisfies the estimate

$$\epsilon=\left\|z(\bar{T})\right\|_{s+d}\leq M\mathrm{e}^{\beta\epsilon^\nu\bar{T}}\left(\left\|z_0\right\|_{s+d}+C\epsilon^{\nu+1}\bar{T}\right)$$

which, provided $\left\|z_0\right\|_{s+d}/\epsilon$ is small enough, implies $\bar{T}>1/C\epsilon^\nu$ with a suitable C. □

Proof of theorem 2.8. Define $\tilde{\epsilon}:=C\left\|z\right\|_{s+s'}$ with the constant of (5.6), then by lemma 5.5 one has $\left\|z(t)\right\|_{s+s'}\leq\tilde{\epsilon}$. Thus, from theorem 2.3 (item 3)it follows

$$\left|\Phi^{(j)}(t)-\Phi^{(j)}(0)\right|\leq C\tilde{\epsilon}^{M_1}\ ,\quad |t|\leq\frac{1}{\tilde{\epsilon}^\nu}\ .$$

defining $\epsilon:=\left\|z\right\|_{s+s'}$ and redefining the constants one has the thesis. □

6. Proof of theorem 3.1

For $k\in\mathbf{Z}^N\setminus\{0\}$ with $0<|k|\leq r_*+2$ define

$$\mathcal{R}_k(\gamma,\alpha):=\left\{m\in\mathcal{I}\ :\ \left|\omega^{(N)}\cdot k\right|<\frac{\gamma}{N^\alpha}\right\}$$

We will estimate the measure of the union of such sets.

First one has

Lemma 6.1. *The eigenvalues μ_j fulfill the gap condition*

$$\mu_{j+1}-\mu_j\geq\frac{C}{j^\delta}\ ,\quad\forall j\geq 1\ . \tag{6.1}$$

with $\delta=4\bar{\tau}/n$, $\bar{\tau}$ being the exponent in the diophantine condition for a.

Proof. Since a is a diophantine vector one has

$$|\mu_{\mathbf{k}}-\mu_{\mathbf{j}}|=\left|\sum_l a_l(k_l^2-j_l^2)\right|\geq\frac{\gamma_1}{\left(\sum_l|k_l^2-j_l^2|\right)^{\bar{\tau}}}\ ,$$

but

$$\sum_l|k_l^2-j_l^2|<C\left\|\mathbf{k}\right\|^2\left\|\mathbf{j}\right\|^2\leq C\left[\inf\left(a_i\right)\right]^{-2}\mu_{\mathbf{j}}\mu_{\mathbf{k}}\leq Ck^{2/n}j^{2/n}$$

where k and j are the integer indexes corresponding to $\mathbf{k}$ and $\mathbf{j}$ respectively, and we used the well known estimate

$$C_1 j^{2/n} \leq \mu_j \leq C_2 j^{2/n} \ . \tag{6.2}$$

Since the minimum of the difference is realized for $k = j+1$, one has

$$|\mu_{\mathbf{k}} - \mu_{\mathbf{j}}| \geq \frac{C}{j^{4\bar{\tau}/n}}$$

□

All what we will prove from now on will be based only on the gap estimate (6.1). Moreover we will labelling the frequencies only by the scalar indexes.

Lemma 6.2. *For any $K \leq N$, let $j^{(1)}, ..., j^{(K)}$ be K different indexes smaller or equal than N; then one has*

$$\begin{vmatrix} \omega_{j^{(1)}} & \omega_{j^{(2)}} & \cdots & \omega_{j^{(K)}} \\ \frac{d\omega_{j^{(1)}}}{dm} & \frac{d\omega_{j^{(2)}}}{dm} & \cdots & \frac{d\omega_{j^{(K)}}}{dm} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{d^{K-1}\omega_{j^{(1)}}}{dm^{K-1}} & \frac{d^{K-1}\omega_{j^{(2)}}}{dm^{K-1}} & \cdots & \frac{d^{K-1}\omega_{j^{(K)}}}{dm^{K-1}} \end{vmatrix} \geq \frac{C}{N^{K^2(\delta+2/n)-2K/n}}. \tag{6.3}$$

Proof. First remark that, by explicit computation one has

$$\frac{d^j\omega_i}{dm^j} = \frac{(2j-3)!}{2^{j-2}(j-2)!2^j} \frac{(-)^{j+1}}{(\mu_i+m)^{j-\frac{1}{2}}} \ . \tag{6.4}$$

Substituting (6.4) in the l.h.s. of (6.3) we get the determinant to be estimated. To obtain the estimate factorize from the $l-th$ column the term $(\mu_{j^{(l)}}+m)^{1/2}$, and from the $j-th$ row the term $\frac{(2j-3)!}{2^{j-2}(j-2)!2^j}$. Forgetting the inessential powers of -1, we obtain that the determinant to be estimated is given by

$$\left[\prod_{l=1}^{K} \omega_{j^{(l)}}\right] \left[\prod_{j=1}^{K-1} \frac{(2j-3)!}{2^{j-2}(j-2)!2^j}\right] \times \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_{j^{(1)}} & x_{j^{(2)}} & x_{j^{(3)}} & \cdots & x_{j^{(K)}} \\ x^2_{j^{(1)}} & x^2_{j^{(2)}} & x^2_{j^{(3)}} & \cdots & x^2_{j^{(K)}} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x^{K-1}_{j^{(1)}} & x^{K-1}_{j^{(2)}} & x^{K-1}_{j^{(3)}} & \cdots & x^{K-1}_{j^{(K)}} \end{vmatrix} \tag{6.5}$$

where we denoted by $x_j := (\mu_j+m)^{-1} \equiv \omega_j^{-2}$. The last determinant is a Vandermond determinant whose value is given by

$$\prod_{l<k\leq K} (x_{j^{(l)}} - x_{j^{(k)}}) \qquad (6.6)\ .$$

Now one has

$$|x_i - x_j| = \left|\frac{1}{\mu_i + m} - \frac{1}{\mu_j + m}\right| = \frac{|\mu_i - \mu_j|}{\omega_i^2\omega_j^2} \geq \frac{C}{(\min\{i,j\})^\delta \omega_i^2\omega_j^2} > \frac{C}{i^{\delta+2/n}j^{\delta+2/n}} \tag{6.7}$$

Denoting $y_i := i^{-(\delta+2/n)}$, (6.6) is estimated by

$$\prod_{l=2}^{K}\prod_{N=1}^{l-1} C y_{i^{(l)}} y_{i^{(N)}} = C^{\sum_{l=2}^{K}(l-1)} \prod_{l=2}^{K}\left(y_{i^{(l)}}^{l-1}\prod_{N=1}^{l-1} y_{i^{(N)}}\right) = C\prod_{l=1}^{K} y_{i^{(l)}}^{K} \ ,$$

from which the thesis immediately follows. □

The rest of the proof follows very closely the proof of theorem 6.16 of [8]. We repeat the main steps for completeness.

Lemma 6.3. *(Proposition of Appendix B in [21]) Let $u^{(1)}, ..., u^{(K)}$ be K independent vectors with $\left\|u^{(i)}\right\|_{\ell^1} \leq 1$. Let $w \in \mathbb{R}^K$ be an arbitrary vector, then there exist $i \in [1, ..., K]$, such that*

$$|u^{(i)} \cdot w| \geq \frac{\|w\|_{\ell^1} \det(u^{(i)})}{K^{3/2}} \ ,$$

where $\det(u^{(i)})$ is the determinant of the matrix formed by the components of the vectors $u^{(i)}$.

For the proof see [21].

Corollary 6.4. *Let $w \in \mathbb{R}^{\mathbf{N}}$ be a vector with only K components different from zero, namely those with index $i^{(1)}, ..., i^{(K)}$; assume $K \leq N$, $K \leq r_* + 2$. Then, for any $m \in \mathcal{I}$ there exists, an index $i \in [0, ..., K-1]$ such that*

$$\left|w \cdot \frac{d^i\omega}{dm^i}(m)\right| \geq C\frac{\|w\|_{\ell^1}}{N^\rho}$$

where ω is the frequency vector, and $\rho = K^2(\delta + \frac{2}{n}) - \frac{2K}{n} + n - 1/n$

Proof. The proof is obtained by considering the vectors

$$u^{(i)} := \begin{cases} \frac{1}{\left\|\frac{d^i\omega}{dm^i}\right\|}\frac{d^i\omega}{dm^i} & \text{if} \quad \left\|\frac{d^i\omega}{dm^i}\right\| > 1 \\ \frac{d^i\omega}{dm^i} & \text{if} \quad \left\|\frac{d^i\omega}{dm^i}\right\| \leq 1 \end{cases}$$

and applying lemma 6.3. For more details see [8] proof of corollary 6.10. □

Lemma 6.5. *(Lemma 8.4 of [22]) Let $g : \mathcal{I} \to \mathbb{R}$ be N times differentiable, and assume that*

(1) $\forall m \in \mathcal{I}$ there exists $i \leq K-1$ such that $\frac{d^ig}{dm^i}(m) > d$

(2) there exists A such that $|g^{(i)}(m)| \leq A$ $\forall m \in \mathcal{I}$, and $\forall i$ with $1 \leq i \leq K$.

Define

$$\mathcal{I}_h := \{m \in \mathcal{I} \ : \ |g(m)| \leq h\} \ ,$$

then

$$\frac{|\mathcal{I}_h|}{|\mathcal{I}|} \leq \frac{A}{d} 2(2+3+...+r+2d^{-1})h^{1/K}$$

For the proof see [22] and [23].

By combining lemma 6.5 and lemma 6.4 we get the following

Lemma 6.6. *For any choice of k one has*

$$|\mathcal{R}_k(\gamma,\alpha)| \leq |\mathcal{I}| C \frac{\gamma^{1/K}}{N^{\varsigma}} \tag{6.8}$$

with $\varsigma = -2\rho + \alpha/K$.

Proof of theorem 3.1. By (6.8) one has

$$\left| \bigcup_k \mathcal{R}_k(\gamma,\alpha) \right| \leq \sum_k |\mathcal{R}_k(\gamma,\alpha)| < \frac{(2N)^{r_*+3}|\mathcal{I}|C\gamma^{1/K}}{N^{\varsigma}} \ ,$$

which, provided α is chosen large enough, is bounded uniformly with respect to N. Taking the union over γ one obtains the thesis. □

7. Proof of theorem 3.2.

First of all we introduce scaled variables $\tilde{x}_i = \sqrt{a_i}x_i, \tilde{u}$ by $u = \epsilon\tilde{u}$, so that equation (3.1) is converted into

$$u_{tt} - a^{\epsilon}_{ij}\partial_i\partial_j u + mu = \epsilon^{\nu} g_{\epsilon}(u, \nabla u) \tag{7.1}$$

where we omitted tildes, and used the notations

$$a^{\epsilon}_{ij} = a_i\delta_{ij} + \epsilon^{\nu} b^{\epsilon}_{ij} \ ;$$

in what follows we will omit the index ϵ from g, a_{ij} and b_{ij}.

We consider equation (7.1) in $\mathbf{T}^n$; moreover we will assume that the initial datum has zero average and that the vector field lives invariant the space of functions with zero average. Since we are interested in the case of Dirichlet boundary condition this is not restrictive. We will prove the theorem by considering directly equation (7.1) using the spaces $Y = H^{s+1} \times H^s$ and $X := H^s \times H^{s-1}$.

Thus the unbounded part of (7.1) is $\partial_t\zeta = A(\zeta)\zeta$ with $\zeta \equiv (u,v)$, and

$$A(\zeta) := \begin{pmatrix} 0 & \mathbb{1} \\ \mathcal{A}(u) & 0 \end{pmatrix} , \quad \mathcal{A}(u) := a_{ij}(u,\nabla u)\partial_i\partial_j - m$$

Having fixed $w \in C^0(I;H^{s+1}) \cap C^1(I;H^s)$ with $I := [0,T]$, we have to show that $A^w(t) := A(w(t))$ fulfills assumption K2. We begin by estimating the stability constants of the operator family $A^w(t)$ in X. To this end it is useful to make reference to the the second order equation

$$\ddot{u} = \mathcal{A}^w(t)u \tag{7.2}$$

where $\mathcal{A}^w(t) := a_{ij}(w(t),\nabla w(t))\partial_i\partial_j - m$. We remark that since $w(t)$ has the property (2.8), one has $b_{ij} \in C^0(I;H^s) \cap C^1(I;H^{s-1})$ with bounded norms. The proof will be based on proposition 9.3 of [2], so first of all we recall its statement.

Proposition 7.1. *[Proposition 9.3 of [2]]. Assume that:*
(i'a) H and H' have structures of real Hilbert spaces, with inner products $\langle\ .\ \rangle_t$ and $\langle\ .\ \rangle'_t$ depending on $t \in I$, such that the norms induced by them are equivalent to the standard norms $\|\ \|_H$ and $\|\ \|_{H'}$, respectively. These inner products are Lipschitz continuous in t, in the sense that $|\langle u,v\rangle_t - \langle u,v\rangle_s| \le c|t-s|\,\|u\|_H\,\|v\|_H$, and similarly for $\langle\ .\ \rangle'_t$.
(i'b) $\{\mathcal{A}(t)\}$ is a family of (unbounded) linear operators in H, such that $\mathcal{A}(t) = \mathcal{A}_S(t)+\mathcal{A}_R(t)$, where $\mathcal{A}_S(t)$ is selfadjoint and uniformly positive definite with respect to inner product $\langle\ .\ \rangle_t$, while $\mathcal{A}_R(t)$ is uniformly bounded (with bound K, say) from H' to H. Moreover $D(\mathcal{A}_S(t)) \subset H'$ and

$$\langle \mathcal{A}_S(t)u,v\rangle_t = \langle u,v\rangle'_t \qquad (u \in D(\mathcal{A}_S(t)), v \in H') \ . \tag{7.3}$$

Then the family

$$A(t) := \begin{pmatrix} 0 & \mathbb{1} \\ \mathcal{A}(t) & 0 \end{pmatrix}$$

is a stable family in $H' \oplus H$ with stability constants M, $\beta := 2c + K$, where M depends on the relation between the equivalent norms $\|\ \|_H$ and $\langle\ ,\ \rangle_0^{1/2}$.

Lemma 7.2. *The evolution operator of the equation (2.7) fulfills the estimate*

$$\sup_{0\le s\le t\le T} \|U(s,t)\|_{X,X} \le M \mathrm{e}^{C\epsilon^\nu T}$$

with constants M, C independent of T and ϵ.

Proof. We define the spaces $H := H^{s-1}$, and $H' := H^s$. The standard scalar product in H^{s-1} is defined by

$$\langle u,v\rangle_H \equiv \langle u,v\rangle_{H^{s-1}} \equiv \int (D^{s-1}u)(D^{s-1}v) \ , \quad D := (-\Delta)^{1/2}$$

and the integral is taken over the whole torus. The standard scalar product in H' is defined in the same way with s in place of $s-1$. Define the scalar product $\langle\ ,\ \rangle_t := \langle\ ,\ \rangle_H$ (independent of t), and the operator $\mathcal{A}_S$

$$\mathcal{A}_S u := -D^{-(s-1)}\left(\partial_j\left(a_{jk}\partial_k(D^{s-1}u)\right)\right) + mu\ .$$

Then $\langle\ ,\ \rangle_t'$ is defined by (7.3). It is easy to see that the operator $\mathcal{A}_S$ is uniformly elliptic (just use positivity of a_{ij}). We will study the relation of this operator with $\tilde{\mathcal{A}} := \partial_j(a_{jk}\partial_k) + m$, who in turn is related to $\mathcal{A}^w(t)$ by

$$\tilde{\mathcal{A}}u = \mathcal{A}u + (\partial_j a_{jk})(\partial_k u)\ .$$

Remark that the second term defines an operator whose norm, as an operator from H' to H, is bounded by a constant times ϵ^ν.

We now compare the operators $\tilde{\mathcal{A}}$ and $\mathcal{A}_S$. For simplicity we restrict to the case where $s-1=2r$, i.e. s is odd*. One has

$$\mathcal{A}_S = \tilde{\mathcal{A}} + D^{-(s-1)}\left[\tilde{\mathcal{A}}, D^{s-1}\right] = \tilde{\mathcal{A}} + D^{-(s-1)}\partial_j[a_{jk}\partial_k, D^{s-1}]\ .$$

So, it is enough to compute $[a_{jk}\partial_k, (\partial_l\partial_l)^r]u$ (summation convention!). To this end consider

$$(\partial_l\partial_l)^r\,(a_{jk}\partial_k u) = \partial_{l_1}\partial_{l_1}...\partial_{l_r}\partial_{l_r}\,(a_{jk}\partial_k u)$$

which is the sum of terms of the form

$$c_\alpha\left(\frac{\partial^{|\alpha|}}{\partial x^\alpha}a_{jk}\right)\left(\frac{\partial^{|\beta|}}{\partial x^\beta}u\right)$$

where α and β are multindices fulfilling $|\alpha|+|\beta|=2r$. The only term with $|\beta|=2r$ is the one where the Laplacians act entirely on u so it disappears when computing the commutator. It is easy to see that all the other terms define operators which are bounded as operators from H' to H, provided $s>[n/2]$. And moreover the norms of such operators are bounded by a constant times ϵ^ν. Applying the remaining two operators (namely $D^{-(s-1)}$ and ∂_j) one has that

$$\mathcal{A} = \mathcal{A}_S + \mathcal{A}_R\ ,$$

with $\mathcal{A}_R : H' \to H$ having norm bounded by a constant times ϵ^ν. Finally, it is very easy to see that

$$|\,\langle u,v\rangle_t' - \langle u,v\rangle_s'\,| \le C\epsilon^\nu\,\|u\|_{H^s}\,\|v\|_{H^s}\,|t-s|\ .$$

So, proposition 7.1 gives the result. □

To estimate the linear flow in Y we construct an isomorphism $S: Y \to X$ such that $SA^wS^{-1} = A^w + \epsilon^\nu B$ with an operator $B: X \to X$ uniformly bounded.

* The general case can be dealt with using the operators S_i defined in the proof of lemma 7.3

We define

$$S := \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} \tag{7.4}$$

Lemma 7.3. *One has*

$$SA^w S^{-1} = A^w + \epsilon^\nu B$$

with B uniformly bounded as an operator from X to X

Proof. One has

$$SA^w S^{-1} = \begin{pmatrix} 0 & \mathbb{1} \\ D\mathcal{A}^w D^{-1} & 0 \end{pmatrix} ,$$

so it is enough to prove that

$$D\mathcal{A}^w D^{-1} = \mathcal{A}^w + \epsilon^\nu \mathcal{B}$$

with $\mathcal{B} : H^s \to H^{s-1}$ uniformly bounded. To prove this fact remark that

$$\epsilon^\nu \mathcal{B} = [D, \mathcal{A}^w]D^{-1} = [D, \epsilon^\nu b_{ij}\partial_i\partial_j]D^{-1}$$

so the result is equivalent $[D, b_{ij}\partial_i\partial_j] : H^{s+1} \to H^{s-1}$ continuously. But $[D, b_{ij}\partial_i\partial_j]u = [D, b_{ij}]\partial_i\partial_j u$, so it is enough the prove that $\forall b \in H^s$ one has $[D, b] : H^{s-1} \to H^{s-1}$ continuously. To prove this fact we compute the commutator in the Fourier space. Denoting by a "hat" the (space) Fourier coefficients of a function one has

$$\begin{aligned} \{[D, b]u\}^\wedge(\mathbf{k}) &= \|\mathbf{k}\| \sum_{\mathbf{l}} \hat{b}_{\mathbf{k}-\mathbf{l}}\hat{u}_{\mathbf{l}} - \sum_{\mathbf{l}} \hat{b}_{\mathbf{k}-\mathbf{l}} \|\mathbf{l}\| \hat{u}_{\mathbf{l}} \\ &= \sum_{\mathbf{l}} \hat{b}_{\mathbf{k}-\mathbf{l}}\hat{u}_{\mathbf{l}} (\|\mathbf{k}\| - \|\mathbf{l}\|) \end{aligned}$$

The last bracket can be rewritten in the form

$$\frac{k_1^2 + ... + k_n^2 - l_1^2 - ... - l_n^2}{\|\mathbf{k}\| + \|\mathbf{l}\|} = (k_1 - l_1)\frac{k_1 + l_1}{\|\mathbf{k}\| + \|\mathbf{l}\|} + ... + (k_n - l_n)\frac{k_n + l_n}{\|\mathbf{k}\| + \|\mathbf{l}\|} .$$

Define now the bilinear operators S_i by

$$\{S_i(b, u)\}^\wedge(\mathbf{k}) := \sum_{\mathbf{l}} \hat{b}_{\mathbf{k}+\mathbf{l}}\hat{u}_{\mathbf{l}} \frac{k_i + l_i}{\|\mathbf{k}\| + \|\mathbf{l}\|} ,$$

then one has

$$[D, b]u = \sum_i S_i(\partial_i b, u) .$$

We prove now that $S_i : H^{s-1} \times H^{s-1} \to H^{s-1}$ continuously, and from this the thesis will immediately follow. Since S_i is bilinear it is enough to prove that

$$\|S_i(v, u)\|_{H^{s-1}} \leq C \|v\|_{H^{s-1}} \|u\|_{H^{s-1}} .$$

One has

$$
\begin{aligned}
\|S_i(v,u)\|^2_{H^{s-1}} &= \sum_{\mathbf{k}} \|\mathbf{k}\|^{2(s-1)} \left(\sum_{\mathbf{l}} \hat{v}_{\mathbf{k}-\mathbf{l}} \hat{u}_{\mathbf{l}} \frac{k_i + l_i}{\|\mathbf{k}\| + \|\mathbf{l}\|} \right)^2 \\
&\le \sum_{\mathbf{k}} \|\mathbf{k}\|^{2(s-1)} \left(\sum_{\mathbf{l}} |\hat{v}_{\mathbf{k}-\mathbf{l}} \hat{u}_{\mathbf{l}}| \right)^2 \left[\sup_{\mathbf{k},\mathbf{l}} \left| \frac{k_i + l_i}{\|\mathbf{k}\| + \|\mathbf{l}\|} \right| \right]^2 \qquad (7.5) \\
&= C \sum_{\mathbf{k}} \|\mathbf{k}\|^{2(s-1)} \left(\sum_{\mathbf{l}} |\hat{v}_{\mathbf{k}-\mathbf{l}}| \, |\hat{u}_{\mathbf{l}}| \right)^2 \ .
\end{aligned}
$$

Define now $u^+(x) := \sum_{\mathbf{k}} |\hat{u}_{\mathbf{k}}| e^{i\mathbf{k}\cdot x}$, and similarly for v^+. Then one has $\|u^+\|_{H^{s-1}} = \|u\|_{H^{s-1}}$, and therefore the square root of the last line of (7.5) is given by

$$
C \left\| u^+ v^+ \right\|_{H^{s-1}} \le C \left\| u^+ \right\|_{H^{s-1}} \left\| v^+ \right\|_{H^{s-1}} = C \left\| u \right\|_{H^{s-1}} \left\| v \right\|_{H^{s-1}}
$$

(with a different C), which is the thesis. □

Then the estimate (3.3) of [1] gives

Corollary 7.4. *The evolution operator of equation (2.7) fulfills the estimate*

$$
\sup_{0 \le s \le t \le T} \|U(s,t)\|_{Y,Y} \le M e^{C\epsilon^\nu T}
$$

with constants M, C independent of T and ϵ.

Appendix.

In this appendix we will prove proposition 5.4. We will work $\mathbb{R}^{2N}$. In order to keep into account the dependence of all we will do on N it is worth to scale the coordinates, namely to define coordinates $(\tilde{p}, \tilde{q})$ defined by

$$
\epsilon \tilde{p} = p \ , \quad \epsilon \tilde{q} = q
$$

and to consider the integrals

$$
\tilde{\Phi}^{(l)}(\tilde{p}, \tilde{q}, \epsilon) := \frac{1}{\epsilon^2} \Phi^{(l)}(\epsilon \tilde{p}, \epsilon \tilde{q}) \ .
$$

¿From now to the end of the appendix we will use only these new variables and integrals, so will omit tildes. The role of small parameter will be played by

$$
\mu := N^{3a+\alpha} \epsilon \ .
$$

Lemma A.1. *There exist analytic coordinates* (ξ_l, η_l) *defined in the ball*

$$\|(\xi_l, \eta_l)\| < 1/C$$

In which one has

$$\Phi^{(l)} = \frac{\xi_l^2 + \eta_l^2}{2} + O(|\xi_l, \eta_l|^3) \ , \quad \|(\xi, \eta) - (p, q)\|_\infty \le C\mu \ .$$

The big O is uniform with respect to μ.

Proof. Write

$$\Phi^{(l)}(p, q) = \frac{p_l^2 + q_l^2}{2} + \mu g_l(p, q, \mu) \ ,$$

by the first of the estimates (5.5) the functions g_l are smooth and bounded provided μ is small enough.

Denote by $\hat{p}^l$ the set of variables

$$\hat{p}^l := (p_1, ..., p_{l-1}, p_{l+1}, ..., p_N)$$

and similarly for $\hat{q}^l$. First remark that, provided μ is small enough, there exists a surface $\mathcal{S}_l$ where the function $\Phi^{(l)}$, considered as a function of p_l, q_l has a minimum (just use implicit function theorem applied to the differential of $\Phi^{(l)}$). Moreover such a surface is analytic and has equation

$$p_l = \mu G^l_p(\hat{p}^l, \hat{q}^l, \mu) \ , \quad q_l = \mu G^l_q(\hat{p}^l, \hat{q}^l, \mu) \ .$$

Introduce new coordinates by

$$x_l := p_l - \mu G^l_p(\hat{p}^l, \hat{q}^l, \mu) \ , \quad y_l := q_l - \mu G^l_q(\hat{p}^l, \hat{q}^l, \mu) \ .$$

Using the implicit function theorem one can see that, provided $\mu < \mu_*$ with a suitable μ_*, this coordinate system is well defined in a neighbourhood (in the $\|.\|_\infty$ norm) of the origin (independent of μ). In these coordinates one has

$$\Phi^{(l)}(x, y) = \left[B_l(\hat{x}^l, \hat{y}^l)\right](x_l, y_l) + O(|x_l, y_l|^3)$$

where the quadratic form $\left[B_l(\hat{x}^l, \hat{y}^l)\right]$ is positive definite. It follows that there exists a rotation $R_l(\hat{x}^l, \hat{y}^l, \mu)$ in the plane x_l, y_l which diagonalize such a quadratic form. Moreover, for $\mu = 0$ the transformations R_l reduce to the identity. Introduce the coordinates

$$(\xi_l, \eta_l) = \left[R_l(\hat{x}^l, \hat{y}^l, \mu)\right](x_l, y_l) \ .$$

By implicit function theorem they are well defined in a neighbourhood of the origin and reduce the integrals to the form

$$\Phi^{(l)} = \frac{\lambda_l^1(\hat{\xi}^l, \hat{\eta}^l)\xi_l^2 + \lambda_l^2(\hat{\xi}^l, \hat{\eta}^l)\eta_l^2}{2} + O(|\xi_l, \eta_l|^3) \ .$$

scaling the variables one has the thesis. □

Lemma A.2. *There exist C_1, C_2 and $\eta_* > 0$ with the following properties: let $z \in \mathbb{R}^{2N}$ and $\bar{z} \in \mathbb{R}^{2N}$ be in a ball of radius smaller than $1/C_1$ and fulfill*

$$|\Phi_j(z) - \Phi_j(\bar{z})| < \eta \ , \quad \text{with } \eta < \eta_*$$

then

$$d_\infty(\mathbf{T}_z, \bar{z}) \le C_2 \eta^{1/2}$$

(recall that we are using rescaled coordinates).

Proof. Introduce the coordinates of lemma A.1. Since they are analytic they are also Lipschitz and therefore, in such coordinates distances are estimated by a constant times the original distance. Introduce now polar coordinate ρ_j, ϕ_j in the planes ξ_j, η_j. Then one has

$$\Phi^{(j)} = \rho_j^2 + \rho_j^3 f_j(\rho, \phi) \ ,$$

with bounded and smooth f_j's. So it is possible to use the contraction mapping principle to solve the above system with respect to the ρ_j^2 obtaining

$$\rho_j^2 := \Phi^{(j)} + (\Phi^{(j)})^{3/2} \tilde{f}_j(\Phi, \phi) \ .$$

One has

$$d_\infty(\bar{z}, \mathbf{T}_z) = \sup_j |\rho_j - \bar{\rho}_j|$$

with an obvious notation. Fixing j and omitting it, we have

$$\begin{aligned} |\rho - \bar{\rho}| &= \left| \sqrt{\Phi + \Phi^{3/2} \tilde{f}(\Phi, \phi)} - \sqrt{\bar{\Phi} + \bar{\Phi}^{3/2} \tilde{f}_j(\bar{\Phi}, \bar{\phi})} \right| \\ &= \frac{\left| \Phi + \Phi^{3/2} \tilde{f}_j(\Phi, \phi) - \bar{\Phi} + \bar{\Phi}^{3/2} \tilde{f}_j(\bar{\Phi}, \bar{\phi}) \right|}{\sqrt{\Phi + \Phi^{3/2} \tilde{f}(\Phi, \phi)} + \sqrt{\bar{\Phi} + \bar{\Phi}^{3/2} \tilde{f}_j(\bar{\Phi}, \bar{\phi})}} \ . \end{aligned}$$

Now, the function $\Phi + \Phi^{3/2} \tilde{f}_j(\Phi, \phi)$ is Lipschitz in Φ, and therefore the numerator is estimated by a constant times $\Phi - \bar{\Phi}$. Write $\bar{\Phi} = \Phi + \delta$, then the denominator is monotonically increasing function of Φ, and therefore it has its minimum for $\Phi_j = 0$. Evaluating at such a point one gets that it is larger than a constant times $\sqrt{\delta}$. From this the thesis easily follows. □

Going back to the nonrescaled variables one has the thesis of proposition 5.4.

References

[1] T. Kato: *Quasi-linear Equations of Evolution, with Applications to Partial Differential Equations.* in Lect. Notes Math. 448 (Springer, New York 1975).

[2] T. Kato: Abstract Differential Equations and nonlinear Mixed Problems. Lezioni Fermiane, Scuola Normale Superiore, Pisa 1985

[3] E.T. Whittaker: *On the adelphic integral of the differential equations of dynamics,* Proc. Roy. Soc. Edinburgh, Sect. A, **37**, 95–109, (1916).

[4] T.M. Cherry: *On integrals deveopable about a singular point of a Hamiltonian system of differential equations.* Proc. Camb. Phil. Soc. **22**, 325–349 (1924).

[5] T.M. Cherry: *On integrals deveopable about a singular point of a Hamiltonian system of differential equations, II.* Proc. Camb. Phil. Soc. **22**, 510–533 (1924).

[6] A. Giorgilli: *Rigorous results on the power expansions for the integrals of a hamiltonian system near an elliptic equilibrium* Ann. Inst. Henri Poincaré, **48** 423–439 (1988).

[7] J. Bourgain: *On diffusion in high dimensional Hamiltonian systems and PDE* J. Anal. Math. **80**, 1-35 (2000).

[8] D. Bambusi: *Birkhoff normal form for some nonlinear PDEs.* Commun. Math. Phys. **234**, 253-285 (2003)

[9] J. Bourgain: *Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equation* GAFA **6** 201–230 (1995).

[10] D. Bambusi, B. Grebert: *Forme normale pour NLS en dimension quelconque.* Preprint 2003.

[11] S.B. Kuksin: *An Averaging Theorem for Distributed Conservative Systems and its Application to Von Karman's Equation,* PMM U.S.S.R. **53**, 150-157 (1989).

[12] A.C.J. Stroucken, F. Verhulst: *The Galerkin–averaging method for nonlinear, undanped continuous systems.* Math. meth. in the appl. Sci. **9**, 520-549 (1987).

[13] H. Pals: *The Galerkin–averaging method for the Klein Gordon equation in two space dimensions.* Nonlinear Anal. TMA. **27**, 841-856 (1996).

[14] M.S. Krol: *On Galerkin–averaging method for weakly nonlinear wave equations.* Math. meth. in the appl. Sci. **11**, 649–664 (1989).

[15] K. Matthies, A. Scheel: *Exponential Averaging for Hamiltonian Evolution Equations.* Trans. Amer. Math. Soc. **355**, 747–773 (2002).

[16] J. Shatah: *Normal forms and quadratic nonlinear Klein Gordon Equations*, Comm. Pure and Appl. Math. **38**, 685-696 (1985).

[17] W. Craig: *Birkhoff normal form for water waves*, in "Mathematical problems in the theory of water waves, F.Dias, J.M.Ghidaglia, J.C.Saut editors. Contemporary mathematics **200** (1996).

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

[19] D. Bambusi: *Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations.* Math. Z. **130**, 345–387, (1999).

[20] S. Kleinerman, A. Majda: *Formation of Singularities for Wave Equations Including the Nonlinear Vibrating String.* Comm. Pure and Appl. Math. **33**, 241-263 (1980).

[21] G. Benettin, L. Galgani, A. Giorgilli: *A Proof of Nekhoroshev's Theorem for the Stability Times in Nearly Integrable Hamiltonian Systems* Cel. Mech. **37**, 1-25, (1985).

[22] D. Bambusi: *On long time stability in Hamiltonian perturbations of nonresonant linear PDE's,* Nonlinearity **12**, 823-850, (1999).

[23] J. Xu, J. You, Q. Qiu: *Invariant tori for nearly integrable Hamiltonian systems with degeneracy* Math. Z. **226**, 375–387 (1997).