GALERKIN AVERAGING METHOD AND POINCARÉ NORMAL FORM FOR SOME QUASILINEAR PDEs

D. Bambusi

13.1.05

Abstract

We use the Galerkin averaging method to construct a coordinate transformation putting a nonlinear PDE in Poincaré normal form up to a remainder of arbitrary order. The abstract theorem we obtain applies to quite general quasilinear equations in one or more space dimensions. Applications to nonlinear wave and heat equations are given. Dynamical consequences are obtained.

Mathematics Subject Classification (2000): 35B20, 37K55, 37L10.

1 Introduction

In this paper we study the dynamics of a partial differential equation of evolution type in the neighbourhood of an equilibrium solution. In particular we prove a theorem allowing to put the equation in Poincaré normal form up to a reminder of any given order. The key idea is to make a Galerkin cutoff, i.e. to approximate the original system by a finite dimensional one, to put in normal form the cut-offed system, and then to choose the dimension of the cut-offed system in such a way that the error due to the Galerkin cutoff and the error due to the truncation of the normalization procedure are of the same order of magnitude. The system one gets is composed by a part which is in normal form and by a remainder which is small when considered as an operator from a Sobolev space to a Sobolev space of much smaller order. As a consequence the remainder is a small but very singular perturbation. Therefore the transformed equations cannot be used directly to study the dynamics. Neglecting the remainder one gets a normalized system whose solutions are *approximate solutions* of the complete system.

Concerning true solutions the situation is different in the parabolic and in the hyperbolic case. In the parabolic semilinear case we show that any true solution with small initial data remains forever close to an approximate solution. In the hyperbolic case we show that a solution with an initial datum of size $R \ll 1$ remains very close to an approximate solution for times of order R^{-1}.

We also consider the case where the equations are Hamiltonian. We show that the normalization is compatible with the Hamiltonian structure and that the normal form obtained is the Birkhoff normal form of the system. The same ideas can be used to deal with the case of different preserved structures, for example with volume preserving equations. The result on Hamiltonian systems was announced in [Bam03c].

The abstract theorem is applied to two concrete cases: a quasilinear wave equation in dimension n with periodic or Dirichlet boundary conditions on a parallelepiped, and a nonlinear heat equation on a segment. The abstract theorem could also be used to deal with modulation equations (see [BCP02, PB05]). Here however we avoid such an application since a precise discussion would require a detailed analysis of the particular models dealt with. Applications to further specific models will be presented elsewhere.

The present result is closely related to the main result of [Bam03a] where some approximate integrals of motion where constructed for nonresonant Hamiltonian PDEs. The results of [Bam03a] are also corollaries of our main theorem; on the contrary the techniques of [Bam03a] cannot be used to deal neither with the non Hamiltonian case nor with the Hamiltonian resonant case (e.g. wave equation with periodic boundary conditions).

A result stronger than the present one was obtained in [BG04] (see also [Bam03b, BG03]), where e quite general normal form theorem for Hamiltonian PDEs was proved. The theorem of [BG04] allows to control the dynamics for time scales much longer than those considered in the present paper. However the theory of [BG04] applies to a more particular class of PDEs. In particular it does not apply to the case of PDEs in which the nonlinearity contains differential operators and to the general case of equations in more than one space dimensions.

Finally we recall the papers [Sha85, MS02, SV87, Pal96, Kro89] which have a strong relation with the present paper.

Plan of the paper. In sect. 2 we state our main normal form theorem and present its applications to concrete models. In sect. 3 we present the dynamical consequences of the theory, and also its applications to concrete models. Finally in sect. 4 we give the proof of the normal form theorem.

Acknowledgement. This work was partially supported by the MIUR project 'Sistemi dinamici di dimensione infinita con applicazioni ai fondamenti dinamici della meccanica statistica e alla dinamica dell'interazione radiazione materia'.

2 Normal Form

2.1 Poincaré normal form

For $s \geq 0$ consider the real Hilbert space ℓ^2_s of the sequences $x \equiv \{x_j\}_{j\in\bar{\mathbb{Z}}}$, $\bar{\mathbb{Z}} := \mathbb{Z} - \{0\}$, such that

$$\|x\|_s^2 := \sum_{j\in\bar{\mathbb{Z}}} |j|^{2s}|x_j|^2 < \infty \ , \tag{2.1}$$

and let $B_s(R)$ be the open ball of radius R and center 0 in ℓ^2_s.

In ℓ^2_s, with $s \geq s_0$, consider the system

$$\dot{x} = X(x) \tag{2.2}$$

where X is a vector field having an equilibrium point at 0, i.e. fulfilling $X(0) = 0$.

Having fixed a positive integer r, we assume

(r-S) There exists $d = d(r)$ with the following properties: for any $s \geq s_0$ there exists an open neighborhood of the origin $\mathcal{U}_{s+d} \subset \ell^2_{s+d}$ such that $X \in C^{r+2}(\mathcal{U}_{s+d}, \ell^2_s)$. $\square$

Write

$$X(x) = Lx + P(x) \ , \tag{2.3}$$

with

$$L := dX(0) \ .$$

Denote by $\mathbf{e}_j \in \ell^2_s$ the vector with all components equal to zero but the j-th one which is equal to 1. We assume

(DL) The linear operator L leaves invariant the spaces $\mathrm{Span}(\mathbf{e}_{-j}, \mathbf{e}_j)$ for all j's. $\square$

Remark 2.1. The operator L has pure point spectrum. We will denote by λ_{-j}, λ_j the eigenvalues of the restriction of L to $\mathrm{Span}(\mathbf{e}_{-j}, \mathbf{e}_j)$. Remark also that $\lambda_j^* = \lambda_{-j}$. $\square$

For simplicity we will also assume that L is diagonalizable. Concerning the eigenvalues we assume

(r-NR) There exist $\alpha = \alpha(r)$ and $\gamma = \gamma(r) > 0$ such that

$$\text{either} \quad \sum_{j=-N}^{N} \lambda_j k_j - \lambda_i = 0 \quad \text{or} \quad \left| \sum_{j=-N}^{N} \lambda_j k_j - \lambda_i \right| \geq \frac{\gamma}{N^\alpha} \ , \tag{2.4}$$

for any N, any $k \in \mathbb{N}^{2N}$ with $1 \leq |k| \leq r+1$ and any $i \in \bar{\mathbb{Z}}$ with $|i| \leq N$. $\square$

Definition 2.2. A vector field Z will be said to be in normal form if

$$[Z, Lx] = 0 \ , \tag{2.5}$$

with the Lie brackets $[F, G]$ of two vector fields defined in the usual way, namely by

$$[F, G] := dF\, G - dG\, F \ . \tag{2.6}$$

Remark 2.3. Complexifying the phase space one can introduce a basis in which the operator L is diagonal. After the introduction of such a basis consider the basis for the space of polynomials given by the monomials

$$P_{k,i}(z) := z^k \mathbf{e}_i \ , \quad k = (...k_{-l}, ..., k_{-1}, k_1, ..., k_l, ...) \ , \tag{2.7}$$

$$z^k := ...z_{-l}^{k_{-l}}...z_{-1}^{k_{-1}} z_1^{k_1}...z_l^{k_l}... \tag{2.8}$$

Then, a polynomial map Z is in normal form if and only if, writing $Z(z) = \sum_{k,i} Z_i^k P_{k,i}(z)$ one has that $Z_i^k \neq 0$ implies

$$\sum_j \lambda_j k_j - \lambda_i = 0 \ . \tag{2.9}$$

□

Theorem 2.4. *Fix r, assume (r-S,DL,r-NR), then there exist constant s', s_1, with the following properties: for any $s \geq s_1$ there exists $R_s > 0$ such that for any $R < R_s$ there exists an analytic transformation $\mathcal{T} : B_s(R/2) \to B_s(R)$ that puts the system in normal form up to order r. Precisely, in the coordinates y defined by $x = \mathcal{T}(y)$, the system turns out to be*

$$\dot{y} = Ly + Z(y) + \mathcal{R}(y) \tag{2.10}$$

where $Z(y)$ is a smooth polynomial of degree $r + 1$ which is in normal form. Moreover, the following estimates hold

$$\sup_{\|x\|_s \leq R} \|Z(x)\|_s \leq C_s R^{2-\frac{1}{2r}} \ , \quad \forall R < R_s \tag{2.11}$$

$$\sup_{\|x\|_{s+s'} \leq R} \|\mathcal{R}(x)\|_s \leq C_s R^{r+3/2} \ , \quad \forall R < R_{s+s'} \tag{2.12}$$

$$\sup_{\|x\|_s \leq R} \|x - \mathcal{T}(x)\|_s \leq C_s R^{2-\frac{1}{2r}} \ , \quad \forall R < R_s \tag{2.13}$$

□

In the proof we will give an algorithm allowing to construct explicitly the normal form. It coincides with the classical algorithm by Poincare applied to a $2N$–dimensional Galerkin truncation of the system with a suitable R dependent integer N.

Remark 2.5. If P has a zero of order $\theta+1$ at $x=0$ then it turns out that also the normal form has a zero of the same order. Moreover in such a case the estimate (2.10) is substituted by

$$\sup_{\|x\|_s \le R} \|Z(x)\|_s \le C_s R^{\theta+1-\frac{1}{2r}} \ , \quad \forall R < R_s$$

□

2.2 Hamiltonian case: Birkhoff normal form

In this section we consider the case where the system is Hamiltonian and the equilibrium point is elliptic. We will show that the normal form obtained is actually the Birkhoff normal form of the system, i.e. the normalizing transformation is canonical and the normalized system is Hamiltonian. With the same ideas one can deal with the case of different structures, for example with the case of volume preserving systems.

Thus endow the space ℓ^2_s with the symplectic structure $\sum_l dx_l \wedge dx_{-l}$ and consider a Hamiltonian system with Hamiltonian function H which is smooth as a function from a neighbourhood of the origin in ℓ^2_s to $\mathbb{R}$. Assume that it has a zero of second order at the origin, and that the corresponding Hamiltonian vector field fulfills assumption (r-S). Let H_0 be the quadratic part of the Hamiltonian ($H_0(x) = \frac{1}{2} < d^2_x H; x, x >$). We assume that the coordinates are such that

$$H_0(x) := \sum_{l \in \bar{\mathbb{Z}}} \omega_l \frac{x_l^2 + x_{-l}^2}{2} \ , \quad \omega_l \in \mathbb{R} \tag{2.14}$$

in particular one has that the equilibrium point is elliptic. Then it is easy to see that assumption (DL) is fulfilled by the Hamiltonian vector field X_{H_0} of H_0. Assumption (r-NR) is equivalent to

(r-NRH) There exist $\gamma > 0$, and $\alpha \in \mathbb{R}$ such that for any N large enough one has

$$\text{either} \quad \sum_{l=1}^{N} \omega_l k_l = 0 \quad \text{or} \quad \left| \sum_{l=1}^{N} \omega_l k_l \right| \ge \frac{\gamma}{N^\alpha} \ , \tag{2.15}$$

for any $k \in \mathbb{Z}^N$, with $|k| \le r+2$. □

Then theorem 2.4 applies and the Hamiltonian vector field can be put in normal form.

Theorem 2.6. *Under the above assumption the normalizing transformation $\mathcal{T}$ is canonical, and one has*

$$H \circ \mathcal{T} = H_0 + H_Z + H_{\mathcal{R}} \tag{2.16}$$

where $\{H_Z, H_0\} = 0$, the Hamiltonian vector fields of H_Z and $H_{\mathcal{R}}$ are the fields Z and $\mathcal{R}$ of (2.10). □

The proof consists in showing that all the steps of the proof of theorem 2.4 are compatible with the Hamiltonian structure (for the details see section 4).

Remark 2.7. In the nonresonant case $\omega \cdot k \neq 0$ for $k \neq 0$ it is well known that the function H_Z depends on the actions

$$I_l = \frac{x_l^2 + x_{-l}^2}{2}$$

only. Therefore such quantities are integrals of motion for the normalized system. □

2.3 Applications

2.3.1 Wave Equation

Fix an n-dimensional vector $a = (a_1, ..., a_n) \in \mathbb{R}^n$ with $a_i > 0$ and consider the n dimensional torus $\mathbb{T}_a^n$ with sides of length $L_i \equiv 2\pi/\sqrt{a_i}$, namely

$$\mathbb{T}_a^n := \frac{\mathbb{R}}{L_1\mathbb{Z}} \times \frac{\mathbb{R}}{L_2\mathbb{Z}} ... \times \frac{\mathbb{R}}{L_n\mathbb{Z}}$$

and the nonlinear wave equation

$$u_{tt} - \Delta u + mu - b_{ij}(\mathbf{x}, u, \nabla u)\partial_i\partial_j u + g(\mathbf{x}, u, \nabla u) = 0 \ . \quad \mathbf{x} \in \mathbb{T}_a^n \quad (2.17)$$

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 functions of class C^∞ (and periodic in the $\mathbf{x}$ variables). Moreover we assume that the b_{ij}'s vanish for $u = \nabla u = 0$ and g has a zero of second order at the same point.

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

$$u(\mathbf{x}, t) = \sum_{\mathbf{j} \in \mathbb{Z}^n} u_{\mathbf{j}}(t) \left(\prod_{i=1}^n \frac{a_i^{1/4}}{\sqrt{2\pi}} \right) e^{i\sqrt{a_i}\mathbf{j}\cdot\mathbf{x}} \ , \quad (2.18)$$

so that (2.17) is converted into the infinite system

$$\ddot{u}_{\mathbf{j}} + \omega_{\mathbf{j}}^2 u_{\mathbf{j}} = P_{\mathbf{j}}(u) \ , \ \mathbf{j} \in \bar{\mathbb{Z}}$$

where P has a zero of order at least 2 at $u = 0$, and

$$\omega_{\mathbf{j}} := \sqrt{\mu_{\mathbf{j}} + m} \ , \quad \mu_{\mathbf{j}} := a_1 j_1^2 + a_2 j_2^2 + ... + a_n j_n^2 \ . \quad (2.19)$$

Remark 2.8. The linear system is Hamiltonian and has the form (2.14). For any index $\mathbf{j} \neq 0$ the corresponding frequency has a multiplicity w between 2 and 2^n, in the sense that there are at least w oscillators with such a frequency. Indeed the frequency does not change if j_i is changed into $-j_i$ thus the system is always resonant also in a Hamiltonian sense. □

To fit the abstract scheme we enumerate the eigenvalues $\mu_{\mathbf{j}}$ of the Laplacian by integers indexes $j \in \mathbb{N}$ in such a way that the μ_j's form a non decreasing sequence. Passing to the corresponding first order system one gets a system of the form (2.2).

It is clear that the space $\mathcal{P}_s$ is isomorphic to $H^{\tilde{s}+1} \oplus H^{\tilde{s}}$ with $\tilde{s} = ns/2$.

The nonresonant condition depends on the choice of the vector a. Assume that a fulfills the condition

- There exist τ_1 and $\gamma_1 > 0$ such that, for any $\mathbf{j} \in \mathbb{Z}^n$, one has

$$\text{either} \quad a \cdot \mathbf{j} = 0 \ , \quad \text{or} \quad |a \cdot \mathbf{j}| \geq \frac{\gamma_1}{|\mathbf{j}|^{\tau_1}}$$

Which holds for example in the case where all the a_j's are equal or in the case where they form a Diophantine vector. From now on we assume that a is fixed and fulfills the above condition. The following theorem is a simple variant of theorem 3.1 of [Bam03a].

Theorem 2.9. *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 $\omega_{\mathbf{j}}$ fulfill the assumption (r-NRH) for any r. Moreover, if a is Diophantine then the frequencies $\{\omega_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{N}^n}$ are nonresonant.* □

Concerning assumption (r–S) it is an immediate consequence of Sobolev embedding theorem. Then the system can be put in normal form up to any finite order. Moreover, if the system is Hamiltonian, according to theorem 2.6 the normal form coincides with the Birkhoff normal form of the system.

Remark 2.10. Define the quantities

$$J_{\mathbf{l}} := \sum_{\mathbf{j}:\omega_{\mathbf{l}}=\omega_{\mathbf{j}}} I_{\mathbf{j}} \ . \tag{2.20}$$

If a is Diophantine and the system is Hamiltonian such quantities are integrals of motion for the normalized system. □

An interesting subcase in is that of Dirichlet boundary conditions. Denote by Ω the parallelepiped of sides $L_i/2$, and consider equation (2.17) with Dirichlet boundary conditions in Ω.

It is well known [Bre83] that the solution $(u, v = \dot{u})$ of the *linear* wave equation belongs to $H^{\tilde{s}+1} \times H^{\tilde{s}}$ if and only if the initial datum fulfills the compatibility conditions

$$(-\Delta)^j \qquad u\big|_{\partial\Omega} = 0 \ , \quad 0 \leq j \leq \left[\tfrac{\tilde{s}}{2}\right] \ , \tag{2.21}$$

$$(-\Delta)^j \quad v\big|_{\partial\Omega} = 0 \ , \quad 0 \leq j \leq \left[\tfrac{\tilde{s}+1}{2}\right] - 1 \ . \tag{2.22}$$

Thus we consider the phase space $\mathcal{F}_s$ of the functions $H^{\tilde{s}+1} \times H^{\tilde{s}}$ fulfilling the compatibility conditions (2.21,2.22), which coincides with the space of the functions of class $H^{\tilde{s}+1} \times H^{\tilde{s}}$ on the torus $\mathbb{T}^n_a$ which are skew symmetric with respect to each of the variables.

By introducing the (normalization) of the basis formed by the functions $\prod_{i=1}^n \sin(\sqrt{a_i} j_i x_i)$ and relabeling the indexes by integer numbers one gets an isomorphism of $\mathcal{F}_s$ with ℓ^2_s.

In this case it is particularly interesting to deal with the Hamiltonian case. We assume that the functions b_{ij} and g are independent of $\mathbf{x}$ and 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

$$H(u, v) = \int_\Omega \left[\frac{v^2}{2} + \frac{\|\nabla u\|^2}{2} + \frac{m u^2}{2} + W(u, \nabla u) \right] d^n \mathbf{x} \ , \tag{2.23}$$

where we denoted by $\| \, . \, \|$ the Euclidean norm of a vector of $\mathbb{R}^n$.

It is easy to see that if W is even in each of its arguments then the smoothness assumption (r–S) is fulfilled for each r. Moreover, if the vector a is Diophantine then, by theorem 2.9 the frequencies are nonresonant (in a Hamiltonian sense) for m belonging to a set of full measure. Thus the theorem 2.6 applies and the system can be put in Birkhoff normal form up to any finite order. Moreover the normal form depends on the actions only, which therefore are expected to be approximate constant of motion.

2.3.2 A heat equation

On the segment $[0, \pi]$ consider the nonlinear heat equation

$$u_t = u_{xx} - V(x) u + f(x, u) \ , \tag{2.24}$$

$$u(0) = u(\pi) = 0 \tag{2.25}$$

The appropriate phase space for the system is the space $\mathcal{F}_s$ of the functions $u \in H^s([0, \pi])$ that extend to 2π periodic skew symmetric functions of class H^s. Similarly it can be defined as the space of the H^s functions fulfilling compatibility conditions of the kind of (2.21,2.22). The resonance relations fulfilled by the eigenvalues of $\partial_{xx} - V$ strongly depend on the potential V. Corresponding to most of the small amplitude smooth potentials there are no resonances among the frequencies. To give the precise statement fix $\sigma > 0$ and, for any positive $\rho \ll 1$ consider the space $\mathcal{V}_\rho$ of the potentials defined by

$$\mathcal{V}_\rho := \left\{ V(x) = \sum_{k \geq 1} v_k \cos kx \mid v'_k := \rho^{-1} e^{\sigma k} \in \left[-\frac{1}{2}, \frac{1}{2} \right] \text{ for } k \geq 1 \right\} \tag{2.26}$$

that we endow with the product probability measure. Thus the potentials we consider are small and analytic. In [BG04] the following theorem was proved

Theorem 2.11. *For any r there exists a positive ρ and a set $\mathcal{S} \subset \mathcal{V}_\rho$ of measure one such that property (r-NR) holds for any potential $V \in \mathcal{S}$, and moreover*

$$\sum_{j=1}^{N} \lambda_j k_j - \lambda_i \neq 0 \ , \quad k \neq 0$$

□

Assuming that the function f is smooth and that $f(x, u(x))$ extends to C^∞ odd function whenever u is C^∞ and odd, also the smoothness condition is fulfilled, and thus the system can be put in normal form up to a remainder of arbitrary order. Moreover the normal form just coincides with the linear part of the system.

Remark 2.12. If $V \equiv 0$ and f is independent of x one can show that the same result holds. In this case, using the techniques by Nikolenko [Nik86], one can show that the normal form actually converges, i.e. the remainder can be completely removed. □

Remark 2.13. In the present case the eigenvalues of the linearized system are in the so called Poincaré domain, and therefore in the finite dimensional case one has that the system can be reduced to its normal form by an analytic transformation. Similar results have been obtained also for some PDEs (see [FS87, FS91]), however a general theorem is still missing. Work is in progress in this direction. □

3 Dynamics

We will use the normal form to construct an approximate solution of the system. The approximate solution is quite easy to be used in the semilinear case, while in the quasilinear case this is nontrivial. Thus in the quasilinear case we will present a scheme leading to a quite general but weaker result.

3.1 Semilinear case

By semilinear case we mean the case in which the system has the form (2.3), i.e.

$$\dot{x} = Lx + P(x)$$

with $P \in C^\infty(\ell^2_s, \ell^2_s)$ for any $s \geq s_0$.

Furthermore, in order to ensure that the system does not leave the domain of validity of the normal form in a time of order 1 we assume that L generates a semigroup of contractions, precisely, denoting by e^{Lt} the semigroup generated by A, we assume that

$$\left\| e^{Lt} \right\|_{\ell^2_s, \ell^2_s} \leq \mathrm{e}^{-\nu t} \ , \quad \forall s \geq s_0 \ , \tag{3.1}$$

with $\nu = \nu(s) \geq 0$.

Remark 3.1. By Segal theory the Cauchy problem for the complete nonlinear system is well posed. □

Fix a large s and consider the Cauchy problem for the normalized equations

$$\dot{y} = Ly + Z(y) \ , \tag{3.2}$$

$$y(0) = y_0 \in \ell^2_{s+s'} \tag{3.3}$$

let $\{y(t)\}_{t\in[0,T]}$ be the corresponding solution, and assume that

$$\|y(t)\|_{s+s'} \le R < R_{s+s'} \ , \quad \forall t \in [0,T] \tag{3.4}$$

(where $R_{s+s'}$ is defined in theorem 2.4).

Remark 3.2. If $\nu > 0$ then it is easy to see, by parabolic estimates, that if (3.4) is fulfilled at $t = 0$ then it is also fulfilled for all positive times. When $\nu = 0$ there always exists a $T_0 > 0$ such that if $\|y_0\|_{s+s'} \le R/2$ then (3.4) is fulfilled for $t \in [0, T_0/R]$. In general, when $\nu = 0$, it is worth to study the time of validity of (3.4) case by case in order to improve T_0. □

Define the approximate solution $\zeta(t) := \mathcal{T}(y(t))$ and an exact solution $x(t)$ of the original system with initial datum x_0 close to $\zeta_0 := \mathcal{T}(y(0))$.

We have the following

Corollary 3.3. *Assume $\nu > 0$, and R small enough, and*

$$\|x_0 - \zeta_0\|_s \le R^{r+3/2} \tag{3.5}$$

then for all positive t one has

$$\|x(t) - \zeta(t)\|_s \le R^{r+3/2} \tag{3.6}$$

□

Corollary 3.4. *Assume $\nu = 0$, and R small enough. If*

$$\|x_0 - \zeta_0\|_s \le R^{r+1/2} \tag{3.7}$$

then for all t fulfilling $|t| < T$ one has

$$\|x(t) - \zeta(t)\|_s \le 2R^{r+1/2} \tag{3.8}$$

where T is any positive time such that (3.4) holds. □

Proof. The function $\zeta(t)$ fulfills the equation

$$\dot{\zeta} = X(\zeta) - \tilde{\mathcal{R}}(t) \tag{3.9}$$

where $\tilde{\mathcal{R}}(t) := (\mathcal{T}^{-1*})(\zeta(t))$ and therefore

$$\sup_{t\in[0,T]} \left\|\tilde{\mathcal{R}}(t)\right\|_s \le CR^{r+3/2} \ . \tag{3.10}$$

Denote $\delta(t) := x(t) - \zeta(t)$. Then δ fulfills the equation

$$\dot{\delta} = A\delta + [P(\zeta(t) + \delta) - P(\zeta(t))] + \tilde{\mathcal{R}}(t)$$

Then, using the formula of variation of the arbitrary constants the smoothness of P, the fact that it has a zero of second order at the origin and the estimate (3.10) one gets

$$\|\delta(t)\|_s \leq \mathrm{e}^{-\nu t} \|\delta_0\| + C \int_0^t R\mathrm{e}^{-\nu(t-s)} \left(\|\delta(s)\| + R^{r+1/2}\right) ds \tag{3.11}$$

which, using Gronwall Lemma gives

$$\|\delta(t)\|_s \leq \mathrm{e}^{-(\nu - CR)t} \|\delta_0\|_s + \frac{1 - \mathrm{e}^{-(\nu - CR)t}}{\nu - RC} CR^{r+3/2} \tag{3.12}$$

from which the thesis of both corollaries follows. □

Corollary 3.3 directly applies to the nonlinear heat equation (2.24). A further result one can deduce on this equation is the following

Theorem 3.5. *Fix $r \geq 1$, consider the Cauchy problem for the system (2.24,2.25) with initial datum u_0,*

$$\|u_0\|_s := \epsilon \ll 1$$

and s large enough. Then, for all positive times, the corresponding solution fulfills

$$\|u(t) - \zeta(t)\|_{s-s'} \leq C\epsilon^{r+3/2}$$

where $\zeta(t)$ is the approximate solution constructed above. □

Corollary 3.4 directly applies to the nonlinear wave equation (2.17) when the coefficients b_{ij} vanish identically.

In the Hamiltonian case the following result plays a relevant role

Proposition 3.6. *Let $\Phi \in C^\infty(\ell^2_s)$ (s large enough) be an integral of motion for the system in normal form, and let $x(t)$ be a solution of the complete system with initial datum in $B_{s+s'}(R)$, R small enough and $w(t) := \mathcal{T}^{-1}(x(t))$, then one has*

$$|\Phi(w(t)) - \Phi(w(0))| \leq C \sup_{y \in B_s(R)} \|d\Phi(y)\| \, |t| R^{r+3/2} \tag{3.13}$$

for all times t smaller than the escape time of $x(t)$ from $B_{s+s'}(R)$. □

In general one can bound the escape time from below by T_0/R^θ, if the perturbation P has a zero of order $\theta + 1$ at the origin.

3.2 Quasilinear case

Denote by $\mathcal{B}(\ell^2_{s+d},\ell^2_s)$ the space of bounded linear operators from ℓ^2_{s+d} to ℓ^2_s and assume that the system is quasilinear, namely that it has the form

$$\dot{z} = A(z)z + g(z) \ , \tag{3.14}$$

where

(Q1) There exists d, and, for any $s \geq s_0$ a positive R_{s+d}, such that the map

$$B_{s+d}(R_{s+d}) \ni z \mapsto A(z) \in \mathcal{B}(\ell^2_{s+d},\ell^2_s)$$

is of class C^∞. Moreover g is smooth, i.e. $g \in C^\infty(B_{s+d}(R_{s+d}),\ell^2_{s+d})$. $\square$

For any R small enough and any positive T, consider the set of the functions $\zeta \in C^0([0,T],\ell^2_{s+d}) \cap C^1([0,T],\ell^2_s)$ fulfilling

$$\sup_{t\in[0,T]} \left\|\zeta(t)\right\|_{s+d} + \sup_{t\in[0,T]} \left\|\dot{\zeta}(t)\right\|_s \leq R \tag{3.15}$$

and the linear time dependent equation

$$\dot{x} = A(\zeta(t))x \tag{3.16}$$

(Q2) There exists $\theta \geq 1$ such that the evolution operator $U(t,s)$ associated to equation (3.16) exists and fulfills the estimate

$$\sup_{0\leq t\leq\tau\leq T} \left\|U(t,\tau)\right\|_{\ell^2_{s+d}\to\ell^2_{s+d}} \leq Me^{\beta R^\theta T} \ , \tag{3.17}$$

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

(Q3) g has a zero of order at least $\theta+1$ at the origin. $\square$

Remark 3.7. If one adds some technical assumptions then it becomes possible to apply Kato's theory [Kat75] in order to ensure well posedness of the Cauchy problem. $\square$

Here we prefer to assume well posedness, thus we add

(Q4) The Cauchy problem is well posed in ℓ^2_s for any s large enough and any initial datum small enough. $\square$

As in the previous subsection let $y(t)$ be the solution of the Cauchy problem for the normalized system (3.2,3.3), and denote $\zeta(t) := \mathcal{T}(y(t))$ the approximate solution.

Proposition 3.8. *Assume (Q1–Q4) and R small enough. Let $x_0 \in B_{s+s'}(R)$ be such that*

$$\left\|x_0-\zeta_0\right\|_s \leq R^{r+1/2} \ . \tag{3.18}$$

Then, there exists $T_0 > 0$ s.t.

$$\|x(t)\|_{s+s'} \le 2R \ , \quad \textit{for} \quad |t| \le T_0/R^{\theta} \tag{3.19}$$

moreover, for any $T_1 > 0$ one has

$$\|x(t) - \zeta(t)\|_s \le R^{r+1/2} \ , \quad \forall |t| \le \frac{T_1}{R^{\theta - 1/2r}} \tag{3.20}$$

□

Proof. First (following [Bam03a]) we prove the existence of a T_1 such that $x(t) \in B_{s+s'}(2R)$ for $|t| \le T_1/R^{\theta}$.

By standard continuation argument $x(t)$ can be continued at least until

$$\|x(t)\|_{s+s'} < 2R \tag{3.21}$$

holds. Let $\bar{T}$ be the first time at which (3.21) is violated, then one has $\left\|x(\bar{T})\right\|_{s+s'} = 2R$. Denote $g^x(t) := g(x(t))$. So, $x(t)$ fulfills the "linear" equation

$$\dot{x} = A^x(t)x + g^x(t), \quad 0 \le t \le \bar{T} \tag{3.22}$$

where the estimate (3.17) holds until time $\bar{T}$. By theorem 2 of [Kat75] (which follows from the formula of variation of constants) the solution of (3.22) satisfies the estimate

$$2R = \left\|x(\bar{T})\right\|_{s+s'} \le M e^{\beta(2R)^{\theta}\bar{T}} \left(\|x_0\|_{s+s'} + CR^{\theta+1}\bar{T}\right)$$

which, provided $\|x_0\|_{s+s'}/R$ is small enough, implies $\bar{T} > T_0/R^{\theta}$. Thus (3.19) is proved.

Denote now $w(t) := \mathcal{T}^{-1}(x(t))$; by the normal form theorem one has $w(t) \in B_{s+s'}(3R)$ for the considered times, thus it fulfills

$$\dot{w} = Lw + Z(w) + \mathcal{R}(t) \tag{3.23}$$

with

$$\|\mathcal{R}(t)\|_s \le CR^{r+3/2} \ .$$

denote now $\delta := w - y$, it fulfills the equation

$$\dot{\delta} = L\delta + Z(y(t) + \delta) + \mathcal{R}(t) \tag{3.24}$$

from which, using the smoothness of Z and the estimate of e^{Lt} (which follows from (Q2)) one gets

$$\|\delta(t)\|_s \le M \|\delta_0\|_s + \int_0^t M \left(CR^{\theta-1/2r} + CR^{r+3/2}\right) ds \tag{3.25}$$

which implies

$$\|\delta(t)\|_s \le M \|\delta_0\|_s e^{CR^{\theta-1/2r}t} + CR^{r+1/2} e^{CR^{\theta-1/2r}t}$$

and the thesis. □

It is also easy to see that proposition 3.6 holds also in this quasilinear case. Precisely one has the following

Proposition 3.9. *Let $\Phi \in C^\infty(\ell^2_s)$ (s large enough) be an integral of motion for the system in normal form, and let $x(t)$ be a solution of the complete system with initial datum in $B_{s+s'}(R)$, R small enough and $w(t) := \mathcal{T}^{-1}(x(t))$, then one has*

$$|\Phi(w(t)) - \Phi(w(0))| \le C \sup_{y \in B_s(R)} \|d\Phi(y)\| \, |t| R^{r+3/2} \tag{3.26}$$

for all times t smaller than the escape time of $x(t)$ from $B_{s+s'}(R)$. Moreover if the initial datum belongs to $B_{s+s'}(R/2)$ there exists T_0 such that the escape time T_e is bounded by

$$|T_e| > \frac{T_0}{R^\theta} \ .$$

□

The theory of this section applies to the nonlinear wave equation (2.17). Indeed in ref. [Bam03a] the following theorem was proved

Theorem 3.10. *Under the assumptions of section 2.3 the system (2.17) fulfills assumptions (Q1-Q4) of subsection 3.2.* □

Moreover if the system is Hamiltonian then in the case of periodic boundary conditions one can show that the quantities (2.20) are approximate integrals of motion for the system in the sense of proposition (3.9). In the case of Dirichlet boundary conditions all the actions are approximate integrals of motion. In this case a further conclusion can be extracted on the dynamics, indeed the level surface of the approximate integrals of motion is an infinite dimensional torus and the following result holds

Corollary 3.11. *Consider the system with Hamiltonian (2.23), under the assumptions of subsection 2.3.1. Fix r, then there exist s', s_* and, for any $s \ge s_*$ there exist constants ϵ_s, T_0 such that the following holds true. For any initial datum $\zeta_0 = (u_0, v_0)$ fulfills*

$$\epsilon := \|\zeta_0\|_s \le \epsilon_s \ , \tag{3.27}$$

then there exists an infinite dimensional torus $\mathbb{T}$ such that one has

$$d_{s-s'}(\zeta(t), \mathbb{T}) \le C\epsilon^{r+3/2-\theta} \tag{3.28}$$

for all times

$$|t| \le \frac{T_0}{\epsilon^\theta} \tag{3.29}$$

□

This theorem was already proved, with a different techniques in [Bam03a].

4 Proofs

Denote by r_*, a number for which assumptions (r-S) and (r-NR) are fulfilled with $r = r_*$ and fix it once for all (it represents the number of normal form steps we will to perform).

In what follows we will use the notation

$$a \preceq b$$

to mean: There exists a positive constant C independent of R and of N such that

$$a \leq Cb \ .$$

4.1 Cutoffs

Expand the perturbation P in Taylor series up to order $r_* + 1$,

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

where P_l are homogeneous polynomial of degree $l+1$, and $\mathcal{R}_*$ is the remainder. We will denote

$$X_*(x) := Lx + \sum_{l=0}^{r_*} P_l(x) \ .$$

Remark 4.1. Due to assumption (r–S) the linear operator L is bounded from ℓ^2_{s+d} to ℓ^2_s. □

Remark 4.2. The polynomials P_l are entire analytic functions, and therefore, introducing the complexification $\ell^{2,\mathbb{C}}_s$ of ℓ^2_s, one has that for any s large enough there exists a constant C_s such that

$$\left\|P_l(x)\right\|_s \leq C_s \left\|x\right\|^{l+1}_{s+d} \ , \ \forall x \in \ell^{2,\mathbb{C}}_s \ ,$$

□

To make the Galerkin cutoff fix a large N that will eventually be related to R, introduce the projector Π_N defined by having fixed a positive integer number N,

$$\Pi_N(x_{-\infty}, ..., x_{-1}, x_1, ..., x_\infty) := (..., 0, 0, x_{-N}, ..., x_{-1}, x_1, ..., x_N, 0, 0, ...) \ ,$$

and put

$$X_\sharp := L_\sharp + \sum_{l=1}^{r_*} X_l \equiv L_\sharp + F^{(0)} \ , \tag{4.1}$$

with

$$L_\sharp := \Pi_N L \Pi_N \ , \quad X_l := \Pi_N P_l \circ \Pi_N \ . \tag{4.2}$$

The system (4.1) is the one we will put in normal form. Remark that

$$X = X_\sharp + (X_* - X_\sharp) + \mathcal{R}_* \ . \tag{4.3}$$

We estimate now the terms we neglect.

Lemma 4.3. *For any $s \geq s_0$ there exists a domain $\mathcal{U}_s^{(0)} \subset \ell_s^{2,\mathbb{C}}$ such that, for any $\sigma > 0$, and any $N \geq 0$, one has*

$$\begin{aligned} \|\mathcal{R}_*(x)\|_s &\preceq \|x\|_{s+d}^{r_*+1} \quad , \quad \forall x \in \mathcal{U}_{s+d}^{(0)} \\ \|X_*(x) - X_\sharp(x)\|_s &\preceq \frac{\|x\|_{s+\sigma+d}}{N^\sigma} \quad , \quad \forall x \in \ell^2_{s+d+\sigma} \ . \end{aligned} \tag{4.4}$$

□

Proof. First remark that

$$\| \mathbb{1} - \Pi_N\|_{s+\sigma\to s} = \frac{1}{N^\sigma} \ . \tag{4.5}$$

and that

$$X_* - X_\sharp = (\mathbb{1} - \Pi_N)X_* + \Pi_N \left(X_* - X_* \circ \Pi_N\right) \ .$$

The first term at right hand side is easily estimated by (4.5). The second term is estimated by the remark that X_* is Lipschitz (with Lipschitz constant estimated by the norm of its differential), thus

$$\sup_{x\in\mathcal{U}_{s+d}^{(0)}} \|X_*(x) - X_*(\Pi_N x)\|_s$$

$$\leq \sup_{x\in\mathcal{U}_{s+d}} \|dX_*(x)\|_{s+d\to s} \| \mathbb{1} - \Pi_N\|_{s+d+\sigma\to s+d} \|x\|_{s+d+\sigma} \preceq \frac{1}{N^\sigma} \|x\|_{s+\sigma+d}$$

The estimate of $\mathcal{R}_*$ is obtained by applying Lagrange estimate of the remainder of the Taylor expansion. □

System (4.1) is finite dimensional, so the standard theory applies. However in order to be able to deduce meaningful results on the original infinite dimensional system the estimates we need have to be quite precise. To this end some tools are needed. They will be introduced in the next subsection.

4.2 The modulus of a polynomial and its norm

As a preliminary step we complexify the space and introduce the basis in which the operator $L_\sharp$ is diagonal and the corresponding basis for the space of the polynomials.

From now on we will use only the complexifyed phase space and the basis in which the operator $L_\sharp$ is diagonal. A point of the phase space will be denoted by z. Moreover, in $\mathbb{C}^{2N}$ we will continue to use the norms defined by (2.1). We will denote by $B_s(R) \subset \mathbb{C}^{2N}$ the ball of radius R in the norm of $\ell_s^{2,\mathbb{C}}$.

To introduce the norm we will use, consider an $\ell_s^{2,\mathbb{C}}$ valued polynomial function F, and expand it on the basis $P_{k,i}$, namely write

$$F(z) = \sum_{k,i} F_i^k P_{k,i}(z) = \sum_{k,i} F_i^k z^k \mathbf{e}_i \ .$$

Definition 4.4. Following Nikolenko [Nik86] we define the *modulus* $\lfloor F \rceil$ of F by

$$\lfloor F \rceil := \sum_{k,i} |F_i^k| P_{k,i} \ . \tag{4.6}$$

Lemma 4.5. *Let $F : \mathbb{C}^{2N} \to \mathbb{C}^{2N}$ be a homogeneous polynomial of degree r. For any couple of positive numbers s, s_1 one has*

$$\sup_{\|z\|_s \le 1} \left\| \lfloor F \rceil (z) \right\|_{s_1} \le (2N)^{\frac{3r+1}{2}} \sup_{\|z\|_s \le 1} \|F(z)\|_{s_1} \tag{4.7}$$

□

Proof. First remark that

$$F_i^k = \frac{1}{k!} \frac{\partial^{|k|} F_i}{\partial z^k} \Big|_{z=0} \ ,$$

$F_i(z)$ being the i-th component of $F(z)$ and $k! := k_{-N}! ... k_N!$. We aim to use Cauchy inequality to estimate F_i^k. To this end it is useful to introduce a family of auxiliary norms in C^{2N}. They are given by

$$\|z\|_s^\infty := \sup_j |j|^s |z_j| \ . \tag{4.8}$$

Remark that $\|z\|_s^\infty \le R$ is equivalent to

$$|z_j| \le R|j|^{-s} =: R_j \ .$$

Thus, from Cauchy inequality applied to a function from C^{2N} endowed by the norm (4.8) to $\mathbb{C}$ one has

$$\left| F_i^k \right| \le \frac{1}{\prod_j R_j^{k_j}} \sup_{\|z\|_s^\infty \le R} |F_i(z)|$$

which implies

$$\left| F_i^k \right| |z^k| \le \sup_{\|z\|_s^\infty \le R} |F_i(z)| \ , \qquad \forall k, i$$

Summing over $k \in \mathbb{N}^{2N}$, and taking into account that there are $(2N)^r$ different values of k with $|k| = r$, one immediately has

$$\sup_{\|z\|_s^\infty \le 1} \left\| \lfloor F \rceil (z) \right\|_{s_1}^\infty \le (2N)^r \sup_{\|z\|_s^\infty \le 1} \|F(z)\|_{s_1}^\infty \ . \tag{4.9}$$

Using the relation between the auxiliary norm and the ℓ_s^2 norm, namely

$$\|z\|_s^\infty \le \|x\|_s \le \sqrt{2N} \, \|z\|_s^\infty \ ,$$

one gets the thesis. □

By remarks 4.1 and 4.2 one has

Corollary 4.6. *The polynomials* $\{X_l\}_{l=0}^{r_*}$ *(cf. (4.2)) fulfill the estimate*

$$\sup_{\|z\|_s \le R} \|\lfloor X_l \rceil (z)\|_s \preceq N^{d+\frac{3l+4}{2}} R^{l+1} \ , \quad l \ge 1 \tag{4.10}$$

□

From now on we fix the index s of the norm.

Definition 4.7. Let F be a vector field analytic in a ball of radius R in the norm $\ell_s^{2,\mathbb{C}}$; we will use the notation

$$|F|_R := \sup_{\|z\|_s \le R} \|\lfloor F \rceil (x)\|_s \tag{4.11}$$

The space of the $\mathbb{C}^{2N}$ valued functions which are analytic and bounded on the ball $\|z\|_s < R$ will be denoted by $\mathcal{A}_R$. The norm (4.11) makes it a Banach space.

By corollary 4.6 one has $F^{(0)} \in \mathcal{A}_R$ for all R small enough, and

$$|F^{(0)}|_R \preceq N^{\alpha_1} R^2 \ , \tag{4.12}$$

with

$$\alpha_1 = d(r_*) + \frac{3r_* + 7}{2} \ . \tag{4.13}$$

4.3 Normalization

In order to normalize the system we will use the method of Lie transform which enables to deal with the case of a preserved structure.

Thus, given an auxiliary vector field W we consider the auxiliary differential equation

$$\dot{z} = W(z) \tag{4.14}$$

and denote by ϕ_t the flow it generates. Moreover denote $\phi := \phi_1 \equiv \phi_t\big|_{t=1}$.

Definition 4.8. The map ϕ will be called the *Lie transform* generated by W. □

Remark 4.9. Given an analytic vector field F, consider the differential equation

$$\dot{z} = F(z) \ . \tag{4.15}$$

Define $\phi^* F$ by

$$(\phi^* F)(y) := d\phi^{-1}(\phi(y))F(\phi(y)) \tag{4.16}$$

then, in the variables y defined by $z = \phi(y)$, the differential equation (4.15) takes the form

$$\dot{y} = (\phi^* F)(y) \tag{4.17}$$

□

Remark 4.10. Using the relation

$$\frac{d}{dt}\phi_t^* F = \phi_t^* \left[W, F\right]$$

it is easy to see that, at least formally, one has

$$\phi^* F = \sum_{l=0}^{\infty} F_l \ , \tag{4.18}$$

with F_l defined by

$$F_0 := F \ , F_l := \frac{1}{l}\left[W, F_{l-1}\right] \ , \qquad l \geq 1 \ . \tag{4.19}$$

□

To come to an estimate of the terms involved in the series (4.18) we start with the following

Lemma 4.11. *Let $F, G \in \mathcal{A}_R$ be two analytic maps, then, for any positive $d < R$, one has $[F,G] \in \mathcal{A}_{R-d}$ and*

$$\left|[F,G]\right|_{R-d} \leq \frac{2}{d}\left|F\right|_R \left|G\right|_R \tag{4.20}$$

□

Proof. Simply remark that

$$\begin{aligned} \left\|\lfloor [F,G] \rceil\right\|_s = \left\|\lfloor dF\, G - dG\, F \rceil\right\|_s &\leq \left\|\lfloor dF\, G \rceil\right\|_s + \left\|\lfloor dG\, F \rceil\right\|_s \\ &\leq \left\|d\lfloor F \rceil \ \lfloor G \rceil\right\|_s + \left\|d\lfloor G \rceil \ \lfloor F \rceil\right\|_s \ . \end{aligned}$$

Then the Cauchy inequality gives $\left\|d\lfloor F \rceil(z)\right\|_s \leq |F|_R/d$ for any z with $\left\|z\right\|_s \leq R - d$. Thus one gets

$$\left\|d\lfloor F \rceil \ \lfloor G \rceil\right\|_s \leq \frac{1}{d}|F|_R|G|_R \ .$$

estimating the other term in the same way one gets the thesis. □

We estimate now the terms of the series (4.18,4.19) defining the Lie transform.

Lemma 4.12. *Let $F \in \mathcal{A}_R$ and $W \in \mathcal{A}_R$ be two analytic maps; denote by F_n the maps defined recursively by (4.19); then, for any positive $d < R$, one has $F_n \in \mathcal{A}_{R-d}$, and the following estimate holds*

$$\left|F_n\right|_{R-d} \leq \left|F\right|_R \left(\frac{2e}{d}\left|W\right|_R\right)^n \ . \tag{4.21}$$

□

Proof. Fix n, and denote $\delta := d/n$, we look for a sequence $C_l^{(n)}$ such that

$$|F_l|_{R-\delta l} \le C_l^{(n)} \ , \quad \forall l \le n \ .$$

By (4.20) this sequence can be defined by

$$C_0^{(n)} = |F|_R \ , \quad C_l^{(n)} = \frac{2}{l\delta} C_{l-1}^{(n)} |\chi|_R = \frac{2n}{ld} C_{l-1}^{(n)} |\chi|_R \ .$$

So one has

$$C_n^{(n)} = \frac{1}{n!} \left(\frac{2n |\chi|_R}{d} \right)^n |F|_R \ .$$

Using the inequality $n^n < n! e^n$, which is easily verified by writing the iterative definition of $n^n/n!$, one has the thesis. ☐

Remark 4.13. Let $W \in \mathcal{A}_R$ be an analytic map. Fix $d < R$. Assume

$$\sup_{\|z\|_s \le R} \|W(z)\|_s < d$$

and consider the corresponding time t flow ϕ_t. Then, for $|t| \le 1$, one has

$$\sup_{\|z\|_s \le R-d} \|\phi_t(z) - z\|_s \le \sup_{\|z\|_s \le R} \|W(z)\|_s < d \tag{4.22}$$

☐

Lemma 4.14. *Consider W as above and let $F \in \mathcal{A}_R$ be an analytic map. Fix $0 < d < R$ assume $\sup_{\|z\|_s \le R} \|W(z)\|_s < d/3$, then one has*

$$\sup_{\|z\|_s \le R-d} \|\phi^* F(z)\|_s \le 2 \sup_{\|z\|_s \le R} \|F(z)\|_s$$

☐

Lemma 4.15. *Let $F \in \mathcal{A}_R$ be a polynomial map of degree $r+1$. There exists $W, Z \in \mathcal{A}_R$ with Z in normal form such that*

$$[L_\sharp, W] + F = Z \tag{4.23}$$

Moreover Z and W fulfill the estimates

$$|W|_R \le \frac{N^\alpha}{\gamma} |F|_R \, , \quad |Z|_R \le |F|_R \tag{4.24}$$

☐

Proof. Since we are using the coordinates in which the operator $L_\sharp$ is diagonal, one has

$$[L_\sharp, P_{k,i}] = (\lambda \cdot k - \lambda_i) P_{k,i} \tag{4.25}$$

thus, writing $F = \sum_{k,i} F_i^k P_{k,i}$, one defines

$$Z := \sum_{RS} F_i^k P_{k,i} \ , \quad W = \sum_{RS^c} \frac{F_i^k}{\lambda \cdot k - \lambda_i} P_{k,i}$$

where the resonant set RS is defined by

$$RS := \left\{(k,i) \in \mathbb{N}^{2N} \times (-N, ..., N) \ : \ \lambda \cdot k - \lambda_i = 0\right\}$$

and RS^c is its complement. Then the thesis immediately follows from the definition of the norm. $\square$

Lemma 4.16. *Let $W \in \mathcal{A}_R$ be the solution of the homological equation (4.23) with $F \in \mathcal{A}_R$. Denote by L_j the functions defined recursively as in (4.19) from $L_\sharp$; for any positive $d < R$, one has $L_j \in \mathcal{A}_{R-d}$, and the following estimate holds*

$$|L_j|_{R-d} \le 2\,|F|_R \left(\frac{2e}{d}\,|W|_R\right)^j \ . \tag{4.26}$$

$\square$

Proof. The idea of the proof is that, using the homological equation one gets $L_1 = Z - F \in \mathcal{A}_R$. Then proceeding as in the proof of lemma 4.12 one gets the result. $\square$

In the statement of the forthcoming iterative lemma we will use the following notations: For any positive R, define $\delta := R/2r_*$ and $R_r := R - r\delta$.

Proposition 4.17. Iterative Lemma. *Consider the system (4.1). For any $r \le r_*$ there exists a positive $R_{*r} \ll 1$ and, for any $N > 1$ there exists an analytic transformation*

$$\mathcal{T}^{(r)} : B_s\left(\frac{R_{*r}(2r_* - r)}{2N^{\alpha+\alpha_1} r_*}\right) \to \ell_s^{2,\mathbb{C}}$$

which puts (4.1) in the form

$$X^{(r)} := \mathcal{T}^{(r)*} X_\sharp = L_\sharp + Z^{(r)} + F^{(r)} + \mathcal{R}_T^{(r)} \ . \tag{4.27}$$

Assume $R < R_{*r}/N^{\alpha+\alpha_1}$, then the following properties hold

1) the transformation $\mathcal{T}^{(r)}$ satisfies

$$\sup_{z \in B_s(R_r)} \left\| z - \mathcal{T}^{(r)}(z) \right\|_s \preceq N^{\alpha+\alpha_1} R^2 \tag{4.28}$$

2) $Z^{(r)}$ is a polynomial of degree $r+1$, it is in normal form, and has a zero of order 2 at the origin; $F^{(r)}$ is a polynomial of degree r_*+1 having a zero of order $r+2$ at the origin. Moreover the following estimates hold

$$\left|Z^{(r)}\right|_{R_r} \ \preceq \ N^{\alpha_1} R^2 \ , \quad \forall r \ge 1 \tag{4.29}$$

$$\left|F^{(r)}\right|_{R_r} \ \preceq \ N^{\alpha_1} R^2 (R N^{\alpha+\alpha_1})^r \tag{4.30}$$

3) the remainder term, $\mathcal{R}_T^{(r)}$ *satisfies*

$$\sup_{z\in B(R_r)} \left\| \mathcal{R}_T^{(r)}(z) \right\| \preceq \left(R N^{\alpha+\alpha_1} \right)^{r_*+2} \ . \tag{4.31}$$

□

Proof. We proceed by induction. First remark that the theorem is trivially true when $r = 0$ with $\mathcal{T}^{(0)} = I$, $\mathcal{Z}^{(0)} = 0$, $F^{(0)} = X^{(0)}$ and $\mathcal{R}_T^{(0)} = 0$.

Then we look for a Lie transform, $\mathcal{T}_r$, eliminating the non normalized part of order $r+1$ from $X^{(r)}$. Let W_r be the generating field of $\mathcal{T}_r$. Using the formulae (4.18,4.19) one writes

$$\mathcal{T}_r^* \left(L_\sharp + Z^{(r)} + F^{(r)} \right) = L_\sharp + Z^{(r)} \tag{4.32}$$

$$+ [W_r, L_\sharp] + F^{(r)} \tag{4.33}$$

$$+ \sum_{l\geq 1} Z_l^{(r)} + \sum_{l\geq 1} F_l^{(r)} + \sum_{l\geq 2} L_l \tag{4.34}$$

where $Z_l^{(r)}$ are the terms of the expansion (4.19) of $Z^{(r)}$ and similarly for the other quantities. Then it is easy to see that (4.32) is the already normalized part of the transformed system, (4.33) is the non normalized part of order $r+2$ that has to be eliminated by a suitable choice of W_r, (4.34) contains all the terms of degree higher than $r+2$.

We first use lemma 4.15 to determine W_r as the solution of the equation

$$[W_r, L_\sharp] + F^{(r)} = Z_r \tag{4.35}$$

with Z_r in normal form. By (4.24) and (4.30) one has the estimates

$$|W_r|_{R_r} \preceq N^\alpha R^2 N^{\alpha_1} \left(N^{\alpha+\alpha_1} R \right)^r , \quad |\mathcal{Z}_r|_{R_r} \preceq N^{\alpha_1} R^2 \left(N^{\alpha+\alpha_1} R \right)^r \ . \tag{4.36}$$

In particular, in view of (4.22) and of the remark that $RN^{\alpha+\alpha_1} < R_*$, the estimate (4.28) is proved at level $r+1$.

Define now $Z^{(r+1)} := Z^{(r)} + Z_r$, and $F_C^{(r+1)}$:=(4.34). From (4.24) the estimate (4.29) holds at level $r+1$. By lemma 4.12, denoting

$$\varsigma := \frac{2e}{\delta} |W_r|_{R_r} \preceq (N^{\alpha+\alpha_1} R)^{r+1} \leq \frac{1}{2},$$

provided $R_{*(r+1)}$ is small enough. Using (4.21), (4.29), (4.30) and lemma 4.16 one gets

$$\left| F^{(r+1)} \right|_{R_r-\delta} \preceq \sum_{l\geq 1} N^{\alpha_1} R^2 \varsigma^l + \sum_{l\geq 1} N^{\alpha_1} R^2 \varsigma^l \left(R N^{\alpha+\alpha_1} \right)^r + \sum_{l\geq 2} N^{\alpha_1} R^2 \varsigma^{l-1} \left(R N^{\alpha+\alpha_1} \right)^r$$

$$\preceq \varsigma N^{\alpha_1} R^2 \preceq N^{\alpha_1} R^2 \left(R N^{\alpha+\alpha_1} \right)^{r+1} \ .$$

Write now

$$F_C^{(r+1)} = F^{(r+1)} + \mathcal{R}_{r,T}$$

where $F^{(r+1)}$ is the Taylor polynomial of degree $r_* + 1$ of $F_C^{(r+1)}$ and $\mathcal{R}_{r,T}$ is the remainder which therefore has a zero of order $r_* + 2$ at the origin. Since $F^{(r+1)}$ is a truncation of $F_C^{(r+1)}$ the previous estimate holds also for it. Then the remainder $\mathcal{R}_{r,T}$ is estimated using Lagrange and Cauchy estimates:

$$\sup_{\|z\|_s \le R} \|\mathcal{R}_{r,T}(z)\|_s \le \frac{R^{r_*+2}}{(r_*+2)!} \sup_{\|z\|_s \le R_{*r}/2N^{\alpha+\alpha_1}} \left\| d^{r_*+2} F_C^{(r+1)}(z) \right\|$$

$$\le R^{r_*+2} \left(\frac{2N^{\alpha+\alpha_1}}{R_{*r}} \right)^{r_*+2} \sup_{\|z\|_s \le R_{*r}/N^{\alpha+\alpha_1}} \left\| F_C^{(r+1)}(z) \right\|_s \preceq (N^{\alpha+\alpha_1} R)^{r_*+2}$$

Define now

$$\mathcal{R}_T^{(r+1)} := \mathcal{T}_r^* \mathcal{R}_T^{(r)} + \mathcal{R}_{r,T} \ . \tag{4.37}$$

By lemma 4.14 one gets the the estimate (4.31) at level $r+1$. □

Proof of theorem 2.4. Consider the transformation $\mathcal{T}^{r_*}$ defined by the iterative lemma, then one has

$$\mathcal{T}^{(r_*)*} X = L_\sharp + Z^{(r_*)} + \mathcal{R}_T + \mathcal{R}_N \tag{4.38}$$

with

$$\mathcal{R}_T := F^{(r_*)} + \mathcal{R}_T^{(r_*)}$$

(see the iterative lemma) and

$$\mathcal{R}_N := \mathcal{T}^{r_* *}(X_* - X_\sharp + \mathcal{R}_*)$$

(see (4.3)). Then use the iterative lemma to estimate $\mathcal{R}_T$ and lemmas 4.3 and 4.14 to get

$$\sup_{\|z\|_{s+d+\sigma} \le R/2} \|\mathcal{R}_N(z)\|_s \le C \left(\frac{R}{N^\sigma} + R^{r_*+2} \right) \ , \quad \forall R < \frac{R_{s+d+\sigma}}{N^{\alpha+\alpha_1}} \tag{4.39}$$

Finally choose $N = R^{-\beta}$ with $\beta = [2\alpha_2(r_*+1)]^{-1}$, $\alpha_2 := \alpha + \alpha_1$, $\sigma := 2\alpha_2^2(r_* + 2)^2 - 1$ and define $s' := d + \sigma$. □

Proof of theorem 2.6. We just show that all the steps of the construction are compatible with the Hamiltonian structure.

The cutoffs: the Taylor cutoff of the vector field is clearly equivalent to the Taylor cutoff of the Hamiltonian. Thus the vector field X_* is Hamiltonian with Hamiltonian function H_* given by the truncation at degree $r_* + 2$ of the Taylor expansion of the Hamiltonian. The Galerkin cut-offed vector field $X_\sharp$ is the Hamiltonian vector field of $H_\sharp(x) := H_*(\Pi_N x)$.

The one step transformation. The key remark is that the solutions Z and W of the homological equation as constructed by lemma 4.15 are Hamiltonian vector fields. The simplest way to verify such a property consists in introducing the variables

$$\xi_l := \frac{1}{\sqrt{2}}(x_l + \mathrm{i}x_{-l}) \ ; \ \eta_l := \frac{1}{\sqrt{2}}(x_l - \mathrm{i}x_{-l}) \quad l \geq 1 \ , \tag{4.40}$$

in which the symplectic form becomes $\sum_l \mathrm{i}\ d\xi_l \wedge d\eta_l$ and the operator $L = X_{H_0}$ is diagonal. Then given a polynomial Hamiltonian function $f(\xi, \eta)$ decompose it as

$$f(\xi, \eta) = \sum_{kj} f_{kj} \xi^k \eta^j$$

and define

$$H_{Z\,kj} := f_{kj} \ , \quad j, l \quad \text{such that} \quad \omega \cdot (j - l) = 0 \tag{4.41}$$

$$H_{W\,kj} := \frac{f_{kj}}{\mathrm{i}\omega \cdot (k - j)} \ , \quad j, k \quad \text{such that} \quad \omega \cdot (j - k) \neq 0 \ , \tag{4.42}$$

and $H_Z(\xi, \eta) := \sum_{kj} f_{kj} H_{Z\,kj} \xi^k \eta^j$ and similarly for H_W. Then consider the homological equation (4.23) with $F \equiv X_f$ (the Hamiltonian vector field of f). It is very easy to verify that the fields W and Z constructed in lemma 4.15 are the Hamiltonian vector fields of the function H_Z and H_W just constructed. Then it turns out that the Lie transform generated by W is a canonical transformation and therefore the statement follows. $\square$

References

[Bam03a] D. Bambusi, *An averaging theorem for quasilinear Hamiltonian PDEs*, Ann. Henri Poincaré **4** (2003), 685–712.

[Bam03b] ______, *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Physics **234** (2003), 253–283.

[Bam03c] ______, *Birkhoff normal form for some quasilinear Hamiltonian PDEs*, Preprint (2003).

[BCP02] D. Bambusi, A. Carati, and A. Ponno, *The nonlinear Schrödinger equation as a resonant normal form*, DCDS-B **2** (2002), 109–128.

[BG03] D. Bambusi and B. Grebert, *Forme normale pour NLS en dimension quelconque*, Compt. Rendu. Acad. Sciences Paris **337** (2003), 409–414.

[BG04] ______, *Birkhoff normal form for PDEs with tame modulus*, Preprint (2004).

[Bre83] H. Brezis, *Periodic solutions of nonlinear vibrating string and duality principle*, Bull. A.M.S. **8** (1983), 409–426.

[FS87] C. Foias and J. C. Saut, *Linearization and normal form of the Navier-Stokes equations with potential forces*, Ann. Inst. H. Poincaré Anal. Non Linéaire **4** (1987), 1–47.

[FS91] ______, *Asymptotic integration of Navier-Stokes equations with potential forces*, Indiana Univ. Math. J. **40** (1991), 305–320.

[Kat75] T. Kato, *Quasi-linear equations of evolutions, with applications to partial differential equations*, Lecture Notes in Mathematics, vol. 448, pp. 25–70, Springer, Berlin, 1975.

[Kro89] M.S. Krol, *On Galerkin–averaging method for weakly nonlinear wave equations*, Math. meth. in the appl. Sci. **11** (1989), 649–664.

[MS02] K. Matthies and A. Scheel, *Exponential averaging for Hamiltonian evolution equations*, Trans. Amer. Math. Soc. **355** (2002), 747–773.

[Nik86] N.V. Nikolenko, *The method of Poincaré normal form in problems of integrability of equations of evolution type*, Russ. Math. Surveys **41** (1986), 63–114.

[Pal96] H. Pals, *The Galerkin–averaging method for the Klein Gordon equation in two space dimensions*, Nonlinear Anal. TMA **27** (1996), 841–856.

[PB05] A. Ponno and D. Bambusi, *KdV equation and energy sharing in FPU*, Chaos (to appear) (2005).

[Sha85] J. Shatah, *Normal forms and quadratic nonlinear Klein Gordon equations*, Comm. Pure and Appl. Math. **38** (1985), 685–696.

[SV87] A.C.J. Stroucken and F. Verhulst, *The Galerkin–averaging method for nonlinear, undanped continuous systems*, Math. meth. in the appl. Sci. **335** (1987), 520–549.

Dario Bambusi
Dipartimento di Matematica
Via Saldini 50, 20133 Milano, Italy
E-mail: `dario.bambusi@unimi.it`
web-address: `http://users.mat.unimi.it/users/bambusi/`