

Galerkin averaging method and Poincaré normal form for some quasilinear PDEs

DARIO BAMBUSI

Abstract. We use the Galerkin averaging method to construct a coordinate transformation putting a nonlinear PDE in Poincaré normal form up to finite order. We also give a rigorous estimate of the remainder showing that it is small as a differential operator of very high order. The abstract theorem is then applied to a quasilinear wave equation, to the water wave problem and to a nonlinear heat equation. The normal form is then used to construct approximate solutions whose difference from true solutions is estimated. In the case of hyperbolic equations we obtain an estimate of the error valid over time scales of order ϵ^{-1} (ϵ being the norm of the initial datum), as in averaging theorems. For parabolic equations we obtain an estimate of the error valid over infinite time.

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

1. Introduction

1.1. Presentation of the results

In this paper we study the dynamics of a partial differential equation of evolution type in the neighborhood of an equilibrium solution. In particular we prove a theorem allowing to put the equation in Poincaré normal form up to a remainder 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 of a part which is in normal form and of 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.

Pervenuto alla Redazione il 19 gennaio 2005 e in forma definitiva il 28 ottobre 2005.

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. This result shows that, in parabolic problems, normal form provides a very efficient tool in order to study the dynamics. We will discuss further this point below.

In the hyperbolic semilinear 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} . Under an additional assumption a similar result is obtained also in the quasilinear case.

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 three concrete cases: a quasilinear wave equation on \mathbb{T}^n , the water wave problem in dimension 2 with periodic boundary conditions, and a nonlinear heat equation on a segment. In particular, in the case of the wave equation and of the heat equation we deduce dynamical consequence according to the above theory, while in the case of the water wave problem we just give the normal form result. 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.

From a technical point of view the proof is rather simple, and it is quite surprising that such a simple technique could lead to such a general theory. The main point consists in performing the normalization procedure on the finite-dimensional truncation of the system keeping into account the dependence of all the constants on the dimension N of the truncated system. This requires a control of the rate at which the small denominators decay as $N \rightarrow \infty$. This is simply obtained by assuming a suitable condition on non-vanishing small denominators. Such a condition was already recognized to be fulfilled in very general situations in [Bou96, Bam03b, BG04, Bam03a]. Then some care is needed in order to obtain estimates allowing to get the final result. Indeed, a trivial application of standard finite-dimensional algorithms would produce a transformation defined in a neighborhood of the origin whose size decreases exponentially with N . This would make impossible the completion of the proof. Indeed the final result can be obtained only if one is able to prove that the radius of the domain of definition of the transformation decreases at most as $N^{-\beta}$ with some positive β . Moreover, in order to get the final result we exploit the fact that H^s functions have Fourier coefficients that decay fast if s is large; in turn the use of such a property is possible only if the constant β is independent of the regularity index s . To prove this fact we proceed by introducing a suitable norm in order to measure polynomial functions and in showing that it has some good properties (see Section 4.2). Furthermore it turns out that it is necessary

to develop the iterative proof of the normalization lemma using techniques typical of the theory of infinite-dimensional dynamical systems.

It also has to be emphasized that the use of a normal form having a remainder which is very singular operator is far from trivial. Here we use normal form to construct approximate solutions of the original system. Then we develop different techniques in order to control the error with respect to a true solution. Apart from the very simple semilinear case we develop a method based on Kato's theory for the quasilinear hyperbolic case and a method based on the use of the Hamiltonian as a Lyapunov function for the Hamiltonian fully nonlinear case. Actually a large part of the paper is devoted to the construction of these methods.

1.2. Discussion

We recall now some previous results which are closely related to the present one. They will be divided into three groups, the first one containing the papers [Sha85, MS03, SV87, FS87, FS91, Cra96, Pal96, Kro89, Bam03a], the second one containing the papers [Bam03b, BG03, BG04], and the third one containing the papers [Nik86, Zeh78].

The papers of the first group have in common the fact that all their normal form results are either included in the present one or could be re-obtained by simple modifications of the present scheme so that one can say that the present paper extends and unifies the results of all these papers.

We come to a more detailed description starting from the paper [Bam03a]. Here the author generalized to PDEs the technique of direct construction of the integrals of motions due to Whittaker, Cherry and Giorgilli. This has been done by introducing a Galerkin cutoff and by applying the finite-dimensional construction to the truncated system. Then such a technique was used to prove that any small amplitude solution is close to a finite-dimensional torus. Such a technique is not a normal form technique, and while on the one hand it is particularly simple, on the other hand it has the drawback of being applicable only to Hamiltonian systems which moreover must have *nonresonant frequencies*. This last limitation in particular is due to a consistency problem which is already present in the formal construction of the approximate integrals. The technique of the present paper (a normal form technique) on the contrary applies also to the resonant case and to the nonhamiltonian case. In particular we point out that all examples dealt with in the present paper are outside the applicability of the theory of [Bam03a]. Moreover it turns out that, while in the context of Hamiltonian systems the kind of normal form obtained here only allows a control of the dynamics over time scales of order ϵ^{-1} (as in averaging theorems), in the case of parabolic equations it allows to control the dynamics over infinite times. Thus we think that in the parabolic case such a result is quite satisfactory, while in the case of Hamiltonian systems some stronger results would be very suited (see below the discussion of [Bam03a, BG04]).

In the papers [SV87, Pal96, Kro89] the authors introduced and used a Galerkin averaging method, namely they used a Galerkin truncation together with a normal-

ization of the finite-dimensional approximation in order to study perturbations of linear PDEs. In those papers, however the authors did not realize that the small denominator can be controlled in a quite general way, and they only dealt either with completely resonant cases (on which much stronger results can be obtained by the methods of [BN98]) or with particular nonresonant cases for which an ad hoc study of the small denominators was done. Furthermore in these papers only a first order normalization has been performed and no iteration scheme was developed. Finally their techniques do not apply neither to quasilinear problems nor to the deduction of modulation equations. This is due to the fact that the authors did not realize the usefulness of normal form for the construction of approximate solution. A point of view which is essential in the present paper.

The papers [FS87, FS91] use a method different from ours in order to obtain a result on the dynamics of the Navier Stokes equation which is very close to our Proposition 3.3. In our opinion the theory of the present paper gives a new insight in these results. Here we did not make a precise statement on the Navier Stokes case since it would have required an ad hoc nontrivial study of its flow and this is outside the scope of the present paper where we just aim to present a general method.

The paper [Cra96] contains a formal computation of the fourth order Birkhoff normal form of the water wave problem. It contains also a rigorous construction of the normalizing transformation eliminating third order terms from the Hamiltonian. In particular such a construction was obtained by using an abstract Cauchy Kovalewski theorem. However the estimate of the remainder thus obtained only ensures that it is small as a differential operator of infinite order, thus making impossible the construction of the subsequent normalizing transformation. The technique of the present paper (which is much simpler than that of [Cra96]) allows to put the Hamiltonian of the water waves problem in normal form up to any finite order with the remainder estimated as a differential operator of finite order (see Section 2.3.2).

In the paper [MS03, Mat01] a method very similar to the present one was used in order to eliminate the time dependence from a rapidly forced PDE, a problem different from ours. We point out that the paper [Mat01] was the first one where it was shown that the method of Galerkin truncation followed by some perturbative procedure is particularly effective in parabolic equations. It had a great influence on the present paper.

Finally we quote the paper [Sha85]: its normal form result, based on the ideas of reabsorbing small denominators by allowing a loss of derivatives, could also be re-obtained by a generalization of the present method.

We come now to the papers [Bam03b, BG03, BG04]. In those papers (see in particular [BG04]) a result much stronger than the present one was obtained for a smaller class of systems. The theory of those papers only allows to deal with semilinear equations and requires a condition on the small denominators that is much stronger than our one. As a consequence the theory of [BG04] applies to quite general equations in one space dimensions, but only to some particular models in more than one space dimensions. The result of [BG04] is stronger than the present one in the sense that the remainder is shown to be small as a bounded operator and this has the consequence of allowing the use of the normal form to prove a priori

estimates on the solution and to describe the dynamics for time scales of order ϵ^{-r} (with arbitrary r) also in Hyperbolic systems.

Finally, in the paper [Nik86, Zeh78] a result of convergence of normal form generalizing Poincaré theorem to PDEs is proved. The main limitation of these papers is due to the nonresonance assumption which is very restrictive thus making the theory applicable only to very special models.

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

ACKNOWLEDGEMENTS. 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’. I thank Bert Van Geemen for suggesting the proof of Proposition 2.18.

2. Normal form

2.1. Poincaré normal form

For $s \geq 0$ consider the real Hilbert space ℓ_s^2 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_s^2 . In the following $\mathbb{N} := \{0, 1, 2, 3, 4 \dots\}$ will denote the set of non negative integers.

In ℓ_s^2 , 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 $\mathcal{U}_{s+d} \subset \ell_{s+d}^2$ such that $X \in C^{r+2}(\mathcal{U}_{s+d}, \ell_s^2)$. □

Write

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

with

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

Denote by $\mathbf{e}_j \in \ell_s^2$ 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 $\text{Span}(\mathbf{e}_{-j}, \mathbf{e}_j)$ 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 $\text{Span}(\mathbf{e}_{-j}, \mathbf{e}_j)$. Remark also that $\lambda_j^* = \lambda_{-j}$. \square

For simplicity we will also assume that L is diagonalizable.¹ Concerning the eigenvalues we assume

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

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

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

In order to define the normal form consider the complexification of the phase space and introduce a basis in which the operator L is diagonal. Then, using this basis, define the vector monomials

$$P_{k,i}(z) := z^k \mathbf{e}_i , \quad k = (\dots k_{-l}, \dots, k_{-1}, k_1, \dots, k_l, \dots) , \quad (2.5)$$

$$z^k := \dots z_{-l}^{k_{-l}} \dots z_{-1}^{k_{-1}} z_1^{k_1} \dots z_l^{k_l} \dots \quad (2.6)$$

which form a basis for the space of polynomials.

Definition 2.2. A polynomial map Z will be said to be in normal form if, writing $Z(z) = \sum_{k,i} Z_i^k P_{k,i}(z)$, one has that $Z_i^k \neq 0$ implies

$$\sum_j \lambda_j k_j - \lambda_i = 0 . \quad (2.7)$$

Remark 2.3. If a vector field Z is in normal form then

$$[Z, Lx] = 0 , \quad (2.8)$$

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

$$[F, G] := dF G - dG F . \quad (2.9)$$

\square

¹ For the extension to the non-diagonalizable case see e.g. [GP88]

Theorem 2.4. Fix r , assume $(r\text{-}S, DL, r\text{-}NR)$, then there exist constants 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 $T : B_s(R) \rightarrow B_s(2R)$ that puts the system in normal form up to order r . Precisely, in the coordinates y defined by $x = T(y)$, 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

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

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

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

Remark 2.5. We remark that the theorem ensures the existence of a function $s \mapsto R_s$ with the stated properties. Thus in particular in equation (2.12) $R_{s+s'}$ denotes the value of this map at $s + s'$. □

Remark 2.6. In the proof we will give an algorithm allowing to construct explicitly the normal form. It coincides with the classical algorithm by Poincaré applied to a $2N$ -dimensional Galerkin truncation of the system with a suitable R dependent integer N . □

Remark 2.7. Even if the transformation is defined on $B_s(R)$, the remainder \mathcal{R} is estimated on a smaller domain, namely on $B_{s+s'}(R)$, so it is bounded only as an operator extracting s' derivatives. □

Remark 2.8. 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 (2.11) is substituted by

$$\sup_{\|x\|_s \leq R} \|Z(x)\|_s \leq C_s R^{\theta+1 - \frac{1}{2r}}, \quad \forall R < R_s. \tag{2.14} \quad \square$$

Remark 2.9. The quantity R_s also depends on the number of steps r . One could try to estimate such a dependence and then to optimize the number of steps in order to obtain an estimate of the remainder better than a power of R , possibly exponentially small. This is possible (the main point consists in estimating the dependence on r of the constants in condition $(r\text{-}NR)$), however we decided not to proceed in this direction since we expect that the result would be valid for a class of phase spaces in which the Fourier coefficients decay faster than exponentially (spaces of entire analytic functions with finite order). □

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_s^2 with the symplectic structure $\sum_l dx_l \wedge dx_{-l}$ and consider a Hamiltonian system with Hamiltonian function H . We also fix a positive κ and assume that H is smooth as a function from a neighborhood of the origin in ℓ_κ^2 to \mathbb{R} .

Remark 2.10. As a consequence one has that H is also smooth in a neighborhood of the origin in ℓ_s^2 for any $s > \kappa$. □

Assume that H 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 ($H_0(x) = \frac{1}{2}d^2H(0)(x, x)$). We assume that the coordinates are such that

$$H_0(x) := \sum_{l>0} \omega_l \frac{x_l^2 + x_{-l}^2}{2}, \quad \omega_l \in \mathbb{R} \tag{2.14}$$

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

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

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

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

Theorem 2.11. *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 2.12. 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. In such a particular case the result was already proved by a direct method in [Bam03a]. □

2.3. Applications

2.3.1. Wave equation

Fix an n -dimensional vector $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ 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

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

and the nonlinear wave equation

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

where we used the summation convention for the indexes $i, j = 1, \dots, n$, we denoted by $\partial u \equiv (\partial_1 u, \dots, \partial_n u)$ 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 $u = \partial u = 0$ and g has a zero of second order at the same point.

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

$$u(\mathbf{x}, t) = \sum_{\mathbf{j} \in \widetilde{\mathbb{Z}}^n} u_{\mathbf{j}}(t)\phi_{\mathbf{j}}(\mathbf{x}), \quad (2.18)$$

where $\{\phi_{\mathbf{j}}(\mathbf{x})\}$ are the normalized eigenfunction of the Laplacian, which coincide with the normalization of the real part of

$$\tilde{\varphi}(x) := \prod_{i=1}^n e^{i\sqrt{a_i}j_i x_i}$$

and $\widetilde{\mathbb{Z}}^n \subset \mathbb{Z}^n$ is a any set of integer vectors with the property that

$$\widetilde{\mathbb{Z}}^n \cup (-\widetilde{\mathbb{Z}}^n) = \mathbb{Z}^n \text{ and } \widetilde{\mathbb{Z}}^n \cap (-\widetilde{\mathbb{Z}}^n) = \{0\}.$$

Then (2.17) is converted into the infinite system

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

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

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

Remark 2.13. The linear system is Hamiltonian and has the form (2.14). For any index $\mathbf{j} \equiv (j_1, \dots, j_n) \neq 0$ 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, and therefore the theory of [Bam03a] never applies. \square

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 $\mu_{\mathbf{j}}$'s form a non decreasing sequence. Passing to the corresponding first order system one gets a system of the form (2.2). Remark that the space ℓ_s^2 is isomorphic to $H^{\tilde{s}+1} \oplus H^{\tilde{s}}$ with $\tilde{s} = ns/2$.

The nonresonance 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

$$\text{either } a \cdot \mathbf{j} = 0, \quad \text{or } |a \cdot \mathbf{j}| \geq \frac{\gamma_1}{|\mathbf{j}|^{\tau_1}}. \tag{2.20}$$

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 2.14. *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.*

Assumption (r-S) 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 2.11 the normal form coincides with the Birkhoff normal form of the system.

Remark 2.15. Define the quantities J_l by

$$J_l := \sum_{\mathbf{j}: \omega_{|\mathbf{j}|} = \omega_j} I_j. \tag{2.20}$$

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

Remark 2.16. The present result could be adapted also to the case of Dirichlet boundary conditions thus re-obtaining the main result of [Bam03a]. On the contrary, due to the presence of resonances the technique of [Bam03a] does not apply to the case of periodic boundary conditions. \square

2.3.2. The water wave problem

The water wave problem consists in describing the motion of the free surface of a fluid subjected to the gravitational force. Here we will consider the case of a fluid lying in a two dimensional domain of infinite depth. Moreover we will restrict to the irrotational case and we will study space periodic solutions. In terms of the velocity potential $\varphi(x, y)$ and of the profile of the surface $\eta(x)$, the equations of motion are given by

$$\Delta\varphi = 0, \quad -\pi \leq x \leq \pi, \quad -\infty < y < \eta(x) \tag{2.21}$$

$$\varphi \rightarrow 0, \quad y \rightarrow -\infty \tag{2.22}$$

$$\eta_t = \varphi_y - \eta_x \varphi_x, \quad \varphi_t = -g\eta - (\partial\varphi)^2, \quad y = \eta(x) \tag{2.23}$$

$$\eta(x - \pi, y) = \eta(x + \pi, y), \quad \varphi(x - \pi, y) = \varphi(x + \pi, y). \tag{2.24}$$

We also add the normalization condition

$$\int_{-\pi}^{\pi} \eta(x) dx = 0$$

which is invariant under the dynamics. Zakharov[Zak68] pointed out that this is a Hamiltonian system. The corresponding Hamiltonian function is the energy of the fluid, and conjugated variables are given by the wave profile $\eta(x)$ and by $\xi(x) := \varphi(x, \eta(x))$, namely the velocity potential at the free surface. In canonical variables the Hamiltonian of the system is given by

$$H(\eta, \xi) = \frac{1}{2} \int_{-\pi}^{\pi} \left(\xi G(\eta)\xi + g\eta^2 \right) dx \tag{2.25}$$

where $G(\eta)$ is the Dirichlet–Neumann operator. We come now to the definition of the Dirichlet–Neumann operator. Having fixed a 2π –periodic function ξ , consider the boundary value problem for φ

$$\Delta\varphi = 0, \quad -\pi \leq x \leq \pi, \quad -\infty < y < \eta(x) \tag{2.26}$$

$$\varphi \rightarrow 0, \quad y \rightarrow -\infty \tag{2.27}$$

$$\varphi(x - \pi, y) = \varphi(x + \pi, y), \tag{2.28}$$

$$\varphi(x, \eta(x)) = \xi(x) \tag{2.29}$$

which has a unique solution provided η and ξ are smooth enough. Denote by $\partial_n\varphi(x)$ the normal derivative of φ at the point $(x, \eta(x))$ (upper boundary of the domain). Then G is defined by

$$G(\eta)\xi := \sqrt{1 + \eta_x^2} \partial_n \varphi. \tag{2.30}$$

The following theorem was proved in [CSS97] and plays an essential role in the theory:

Theorem 2.17 (Theorem 4.2 of [CSS97]). *There is a constant R_0 such that the Dirichlet–Neumann operator G is analytic in the neighborhood $\{\eta : |\eta|_{C^1} < R_0, |\eta|_{C^{s+1}} < \infty\}$ as a mapping in the space of bounded linear operators from H^{s+1} to H^s .*

Denoting by $\nabla_\xi H$ the L^2 gradient of H with respect to ξ , and by $\nabla_\eta H$ the L^2 gradient of H with respect to η , one has that the equations of motion of H are given by

$$\dot{\eta} = \nabla_\xi H = G(\eta)\xi, \tag{2.31}$$

$$\dot{\xi} = -\nabla_\eta H = -\frac{1}{1 + \eta_x^2} \left(\xi_x^2 - (G(\eta)\xi)^2 - 2\eta_x \xi_x G(\eta)\xi \right) \tag{2.32}$$

(see *e.g.* [Cra96]). The rest solution $\xi \equiv \eta \equiv 0$ thus appears as an equilibrium point, and the linearized equations are the Hamilton equations of

$$H_0(\eta, \xi) = \frac{1}{2} \int_{-\pi}^\pi \left(\xi G_0 \xi + g \eta^2 \right) dx \tag{2.33}$$

where $G_0 := G(0)$ is given by $G_0 \equiv |\partial_x|$, namely the operator that in Fourier space acts as a multiplier by $|k|$. Thus, expanding ξ and η in (rescaled) Fourier series, *i.e.* putting,

$$\xi(x) = \frac{1}{\sqrt{\pi}} \sum_{k>0} \frac{\sqrt{g}}{k^{1/4}} (p_k \cos(kx) + p_{-k} \sin kx) \tag{2.34}$$

$$\eta(x) = \frac{1}{\sqrt{\pi}} \sum_{k>0} \frac{k^{1/4}}{\sqrt{g}} (q_k \cos(kx) + q_{-k} \sin kx) \tag{2.35}$$

one gets

$$H_0(p, q) = \sum_{k \in \mathbb{Z}} \omega_k \frac{p_k^2 + q_k^2}{2}$$

with

$$\omega_k := \sqrt{|k|g}. \tag{2.36}$$

By Theorem 2.17 the smoothness assumption (*r*-S) is fulfilled for any *r*. Concerning the nonresonance assumption we have the following

Proposition 2.18. *For any positive r the frequencies (2.36) fulfill the nonresonance assumption (*r*-NRH).*

Proof. Consider a linear combination with length r and coefficients ± 1 of the frequencies, namely an expression of the form

$$\omega_{k_1} \pm \omega_{k_2} \pm \dots \pm \omega_{k_r} \tag{2.37}$$

(possibly with repetitions) and remark that the polynomial (in the ω 's)

$$\mathcal{P}(\omega_{k_1}, \dots, \omega_{k_r}) := \prod_{\pm} (\omega_{k_1} \pm \omega_{k_2} \pm \dots \pm \omega_{k_r}) \tag{2.38}$$

is even in any one of the ω_{k_l} , and therefore it is a function of $\omega_{k_l}^2 \equiv g|k_l|$ only. It follows that (2.38) either vanishes or it is an integer multiple of g . To start with consider the case where (2.38) does not vanish. If all the indexes are smaller than N , the modulus of (2.37) is bounded from below by

$$\frac{g}{|\prod_{\sigma} (\omega_{k_1} \pm \omega_{k_2} \pm \dots \pm \omega_{k_r})|} \geq \frac{g}{(\sqrt{gN})^{2^{r-1}}} \tag{2.39}$$

where the product in the denominator is taken over all the permutations σ of the signs except that of (2.37). Thus the thesis is proved in such a particular case. Consider now the case where (2.38) vanishes. Then there is at least one of the factors that vanishes. Take one of them, use it to express ω_{k_r} in terms of the other frequencies and substitute in (2.37). Due to the structure of the equations one gets an expression of the form

$$c_1 \omega_{k_1} \pm c_2 \omega_{k_2} \pm \dots \pm c_{r-1} \omega_{k_{r-1}} \tag{2.40}$$

where the c_j 's are either 0 or ± 2 . Thus one is reduced to bound from below the modulus of an expression of the form

$$2(\omega_{k_1} \pm \omega_{k_2} \pm \dots \pm \omega_{k_l})$$

where the indexes were rearranged so that the frequencies with a zero coefficient are $\omega_{k_{l+1}} \dots \omega_{k_{r-1}}$. Iterate until there are no more vanishing factors when taking the products of all permutations of indexes. This requires at most $r - 1$ steps, and therefore produces at most a factor 2^{r-1} in front of the expression. Then the application of the same argument as for (2.39) gives the result. \square

Thus, Theorems 2.4 and 2.11 apply to the Hamiltonian of the water wave problem which can be put in normal form up to any finite order with remainder estimated by (2.12).

Remark 2.19. The first terms of the normal form of the Hamiltonian were computed explicitly in a series of paper[DZ94, Cra96, CS93, CW95], where it was shown that the third order term vanishes, and that, surprisingly enough the fourth order term is integrable. The dynamics of this integrable system has also been studied in detail. However the rigorous existence of the normalizing transformation was, up to now, established only for the transformation putting the Hamiltonian in third order normal form[Cra96]. \square

Remark 2.20. At present no rigorous results on the relation between the dynamics of the normal form and the dynamics of the complete systems are available. In order to deduce dynamical consequences one should know something on the Lyapunof exponents of the system (in the spirit of Section 3.2 below). We expect that it should be possible to deduce such properties from known results on the local well posedness of the Cauchy problem. Thus we expect that our theory should allow to show that the normal form controls the dynamics over a time scale R^{-1} (if R is the size of the initial datum). However, since the third order normal form vanishes, the nonlinear effects are relevant only over a time scale R^{-2} , and a control of the dynamics over such a time scale seems to be outside the possibility of the theory developed in the present paper. \square

2.3.3. A heat equation

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

$$u_t = u_{xx} - V(x)u + f(x, u) , \tag{2.41}$$

$$u(0) = u(\pi) = 0 . \tag{2.42}$$

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 $H^s([-\pi, \pi])$. 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}_ρ of the potentials defined by

$$\mathcal{V}_\rho := \left\{ V(x) = \sum_{k \geq 1} v_k \rho e^{-\sigma k} \cos kx \mid v_k \in \left[-\frac{1}{2}, \frac{1}{2} \right] \text{ for } k \geq 1 \right\} \tag{2.43}$$

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 2.21. *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*

$$\sum_{j=1}^N \lambda_j k_j - \lambda_i \neq 0 , \quad \forall k, |k| \geq 2 .$$

Assuming that the function f is smooth and that $f(x, u(x))$ extends to a C^∞ skew-symmetric function whenever u is C^∞ and skew, 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.

3. Dynamics

We will use the normal form to construct an approximate solution of the original system. The approximate solution is quite easy to be used in the semilinear case, while in the quasilinear or fully nonlinear case this is nontrivial.

Thus we will consider the Cauchy problem for the normalized equations, namely

$$\dot{y} = Ly + Z(y) , \tag{3.1}$$

$$y(0) = y_0 , \tag{3.2}$$

and the general idea is that one should be able to describe quite well its solutions (or at least some of their properties like existence of integrals of motions or of periodic orbits, otherwise the normal form is useless). Then one wants to compare $y(t)$ with a true solution. The intermediate step is to go back to the original variables, *i.e.* to define the approximate solution

$$\zeta(t) := \mathcal{T}(y(t)) . \tag{3.3}$$

Since \mathcal{T} is defined only in a neighborhood of the origin it is important to know the amount of time y remains small. Precisely the relevant quantity is the escape time $T_e(s, R)$ of $y(t)$ from $B_s(R)$ defined as

$$T_e(s, R) := \sup \{t \geq 0 ; y(t) \in B_s(R)\} . \tag{3.4}$$

Then remark that the approximate solution fulfills the equation

$$\dot{\zeta} = X(\zeta) - \tilde{\mathcal{R}}(t) \tag{3.5}$$

where $\tilde{\mathcal{R}}(t) := (\mathcal{T}^{-1*}\mathcal{R})(\zeta(t))$ fulfills the estimate

$$\sup_{t \in [0, T]} \left\| \tilde{\mathcal{R}}(t) \right\|_s \leq C_T R^{r+3/2} , \quad T < T_e(s, R) . \tag{3.6}$$

Then we consider an exact solution $x(t)$ with initial datum x_0 close to ζ_0 and we want to bound $\|x(t) - \zeta(t)\|_s$ for times as long as possible (and for s as large as possible). Sometimes the escape time of $x(t)$ from $B_s(R)$ plays a role, thus we define

$$\tilde{T}_e(s, R) := \sup \{t ; x(t) \in B_s(R)\} . \tag{3.7}$$

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 $P \in C^\infty(\mathcal{U}_s, \ell_s^2)$ for any $s \geq s_0$ and $\mathcal{U}_s \subset \ell_s^2$ a neighborhood of the origin.

Furthermore, in order to ensure that the solution 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 L , we assume that

$$\|e^{Lt}\|_{\ell_s^2, \ell_s^2} \leq e^{-\nu t}, \quad \forall s \geq s_0, \tag{3.8}$$

with $\nu = \nu(s) \geq 0$. Moreover, we assume that P has a zero of order $\theta + 1$ at the origin, namely that

$$\|P(x)\|_s \leq C_s \|x\|_s^{\theta+1}$$

in a neighborhood of the origin.

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

Remark 3.2. In the parabolic case $\nu > 0$, by parabolic estimates, one has that, if the initial datum y_0 for the normalized equations is in $B_s(R)$, then the escape time $T_e(s, R)$ is infinite. \square

We have the following

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

$$\|x_0 - \zeta_0\|_s \leq R^{r+3/2} \tag{3.9}$$

then for all positive t one has

$$\|x(t) - \zeta(t)\|_s \leq 2R^{r+3/2}. \tag{3.10}$$

Proposition 3.4. *Assume $\nu = 0$; fix arbitrary $T_0 > 0$ and $T < T_e(s + s', R)$; assume R small enough. If*

$$\|x_0 - \zeta_0\|_s \leq R^{r-\theta+1/2} \tag{3.11}$$

then for all t fulfilling

$$|t| < \min \left\{ T, \frac{T_0}{R^\theta} \right\}$$

one has

$$\|x(t) - \zeta(t)\|_s \leq 2R^{r-\theta+1/2}. \tag{3.12}$$

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

$$\dot{\delta} = L\delta + [P(\zeta(t) + \delta) - P(\zeta(t))] + \tilde{\mathcal{R}}(t). \tag{3.13}$$

Then, using the formula of variation of the arbitrary constants the smoothness of P , the fact that it has a zero of order $\theta + 1$ at the origin and the estimate (3.6) one gets

$$\|\delta(t)\|_s \leq e^{-\nu t} \|\delta_0\| + C \int_0^t R^\theta e^{-\nu(t-s)} \left(\|\delta(s)\| + R^{r-\theta+1/2} \right) ds \quad (3.14)$$

which, using Gronwall Lemma gives

$$\|\delta(t)\|_s \leq e^{-(\nu-CR^\theta)t} \|\delta_0\|_s + \frac{1 - e^{-(\nu-CR^\theta)t}}{\nu - R^\theta C} CR^{r+3/2} \quad (3.15)$$

from which the thesis of both corollaries follows. □

Proposition 3.3 directly applies to the nonlinear heat equation (2.41). Proposition 3.4 directly applies to the nonlinear wave equation (2.17) when the coefficients b_{ij} vanish identically.

In the conservative case the following result plays a relevant role

Proposition 3.5. *Let $\Phi : \ell_\kappa^2 \supset \mathcal{U} \rightarrow \mathbb{R}$ (κ a positive parameter) be a C^1 integral of motion for the system in normal form, and let $x_0 \in B_{s+s'}(R)$, with $s \geq \kappa, s_0$; define $w(t) := \mathcal{T}^{-1}(x(t))$, and fix $\tilde{T} < \tilde{T}_e(s + s', R)$. If R is small enough, then one has*

$$|\Phi(w(t)) - \Phi(w(0))| \leq C \sup_{y \in B_\kappa(R)} \|d\Phi(y)\| |t| R^{r+3/2}, \quad \forall t < \tilde{T}. \quad (3.16)$$

Remark 3.6. By Gronwall lemma, for any $T_0 > 0$ there exists $C \geq 0$ such that, if $x_0 \in B_{s+s'}(R/C)$ then $\tilde{T}_e(s + s', R) \geq T_0/R^\theta$. □

Proof. Remark that w satisfies

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

and that $w(t) \in B_{s+s'}(2R) \forall |t| < T$. Then compute

$$\frac{d}{dt} \Phi(w(t)) = d\Phi(w)[Lw + Z(w) + \mathcal{R}(w)] = d\Phi(w)\mathcal{R}(w)$$

which is estimated by

$$|d\Phi(w)\mathcal{R}(w)| \leq \sup_{\|y\|_s \leq R} \|d\Phi(y)\| \sup_{\|y\|_{s+s'} \leq R} \|\mathcal{R}(y)\| \leq C \sup_{\|y\|_s \leq R} \|d\Phi(y)\| R^{r+3/2},$$

from which the thesis immediately follows. □

3.2. Quasilinear case

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

$$\dot{z} = A(z)z + g(z) , \tag{3.17}$$

where

- (Q1) There exists d , and, for any $s \geq s_0$ a neighborhood $\mathcal{U}_{s+d} \subset \ell_{s+d}^2$ of the origin, such that the map

$$\mathcal{U}_{s+d} \ni z \mapsto A(z) \in \mathcal{B}(\ell_{s+d}^2, \ell_s^2)$$

is of class C^∞ . Moreover g is smooth, *i.e.* $g \in C^\infty(\mathcal{U}_{s+d}, \ell_s^2)$. □

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

$$\sup_{t \in [0, T]} \|\zeta(t)\|_{s+d} + \sup_{t \in [0, T]} \|\dot{\zeta}(t)\|_s \leq R \tag{3.18}$$

and the linear time dependent equation

$$\dot{x} = A(\zeta(t))x . \tag{3.19}$$

- (Q2) There exists $\theta \geq 1$ such that the evolution operator $U(t, s)$ associated to equation (3.19) exists and fulfills the estimate

$$\sup_{0 \leq t \leq \tau \leq T} \|U(t, \tau)\|_{\ell_{s+d}^2 \rightarrow \ell_{s+d}^2} \leq M e^{\beta R^\theta T} , \tag{3.20}$$

with some constants M, β independent of ζ, T, R . □

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

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. □

Here we prefer to assume well posedness, thus we add

- (Q4) The Cauchy problem is well posed in ℓ_s^2 for any s large enough and any initial datum small enough. □

Proposition 3.8. *Assume (Q1–Q4) and R small enough. Let $x_0 \in B_{s+s'}(R)$ ($s \geq s_0$) be such that*

$$\|x_0 - \zeta_0\|_s \leq R^{t+1/2} . \tag{3.21}$$

Then, for any $T_0 > 0$ there exists K s.t.

$$\|x(t)\|_{s+s'} \leq KR, \quad \text{for } |t| \leq T_0/R^\theta \tag{3.22}$$

moreover, for any $T_1 > 0$ one has

$$\|x(t) - \zeta(t)\|_s \leq 2R^{r-\theta+1/2}, \quad \forall |t| \leq \frac{T_1}{R^{\theta-1/2r}}. \tag{3.23}$$

Remark 3.9. The time scale of (3.23) is not very satisfactory. Indeed in such a case the normal form has a zero of order $\theta + 1$ at the origin, and therefore the first corrections to the dynamics are visible over a time scale $R^{-\theta}$ which is longer than that covered by (3.23). Thus over the time scale (3.23) the solution is up to errors of order $R^{1+1/2r}$ equal to the linear dynamics. Anyway, the estimate of the error given in (3.23) is much stronger than this trivial one. Below, in the case of Hamiltonian systems, we will give a result valid over the time scale R^θ . \square

Remark 3.10. The theory of this section applies to the nonlinear wave equation (2.17). Indeed in ref. [Bam03a] it was proved that it fulfills assumptions (Q1-Q4). \square

Proof. First (following [Bam03a]) we prove that $x(t) \in B_{s+s'}(KR)$ for $|t| \leq T_0/R^\theta$.

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

$$\|x(t)\|_{s+s'} < KR \tag{3.24}$$

holds. Let \bar{T} be the first time at which (3.24) is violated, then one has $\|x(\bar{T})\|_{s+s'} = 2R$. Denote $g^x(t) := g(x(t))$. So, $x(t)$ fulfills the “linear” equation

$$\dot{x} = A^x(t)x + g^x(t), \quad 0 \leq t \leq \bar{T} \tag{3.25}$$

where the estimate (3.20) holds until time \bar{T} . By Theorem 2 of [Kat75] (which follows from the formula of variation of constants) the solution of (3.25) satisfies the estimate

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

which, provided K is large enough, implies $\bar{T} > T_0/R^\theta$. Thus (3.22) is proved.

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

$$\dot{w} = Lw + Z(w) + \mathcal{R}(t) \tag{3.26}$$

with

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

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

$$\dot{\delta} = L\delta + Z(y(t) + \delta) + \mathcal{R}(t) \tag{3.27}$$

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

$$\|\delta(t)\|_s \leq M \|\delta_0\|_s + \int_0^t M \left(CR^{\theta-1/2r} + CR^{r+3/2} \right) ds \tag{3.28}$$

which implies

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

and the thesis. □

It is also easy to see that Proposition 3.5 holds exactly in the same form also in this quasilinear case.

3.3. Hamiltonian case

As anticipated above in the Hamiltonian case one can get a result stronger than Proposition 3.8. This is given by the forthcoming Proposition 3.12, whose proof is an abstract formulation of a technique of [SW00]. Here the assumptions we need are quite different from those of the previous subsection. In particular we do not assume the quasilinear structure of the previous subsection.

So consider a Hamiltonian system as in Subsection 2.2 with Hamiltonian function $H = H_0 + H_1$. Concerning the structure of the Hamiltonian we assume the so called potential well condition, while for the vector field we only assume that its Cauchy problem is well posed.

(QH1) The Hamiltonian is C^∞ in a neighborhood of the origin of ℓ_κ^2 with some positive κ . Moreover

$$C \|x\|_\kappa^2 \leq H_0(x) , \quad |H_1(x)| \leq C \|x\|_\kappa^{\theta+2} \tag{3.29}$$

(QH2) The Cauchy problem is well posed in a neighborhood of the origin of ℓ_s^2 for some $s \geq \kappa$. □

For simplicity we will assume $\kappa < s_1$.

Remark 3.11. If the vector field is quasilinear as in the previous section and fulfills the assumptions (Q1-Q4), one has that, having fixed an arbitrary T_0 , there exists a $C > 0$ such that, if the initial datum fulfills

$$\|x_0\|_s \leq \frac{R}{C}$$

for some $s \geq \kappa$ then

$$\tilde{T}_e(\kappa, R) > \frac{T_0}{R^\theta};$$

this is relevant for the interpretation of the next proposition. □

Proposition 3.12. *Assume (QH1, QH2), fix arbitrary $T_0 > 0$, $\tilde{T} < \tilde{T}_e(\kappa, R)$, $T < T_e(s_1 + s', R)$; assume R small enough. If*

$$\|x_0 - \zeta_0\|_\kappa \leq R^{r+1/2} \tag{3.30}$$

then for all t fulfilling

$$|t| < \min \left\{ T, \tilde{T}, \frac{T_0}{R^\theta} \right\}$$

one has

$$\|x(t) - \zeta(t)\|_\kappa \leq 2R^{r-\theta+1/2}. \tag{3.31}$$

Proof. It is useful to introduce some notations. First we denote by $\langle \cdot, \cdot \rangle$ the ℓ^2 scalar product and by ∇ the gradient with respect to the ℓ^2 metric. Moreover *only for this section* we will denote by J the Poisson tensor, namely the linear operator defined by

$$J(x_k, x_{-k}) = (x_{-k}, -x_k) \quad \forall k > 0.$$

So the Hamilton equations of our system coincide with

$$\dot{x} = J\nabla H(x).$$

Remark also that the approximate solution $\zeta(t)$ fulfills

$$\dot{\zeta} = J\nabla H(\zeta) - \tilde{\mathcal{R}}(\zeta(t)). \tag{3.32}$$

Consider the equations of motion for $\delta := x - \zeta(t)$, namely equations (3.13). Interpreting the transformation $x \mapsto \delta$ as a time dependent coordinate transformation one immediately has that the equation for δ , namely (3.13), is Hamiltonian with a time dependent Hamiltonian function given by

$$K(\delta, t) := H(\zeta(t) + \delta) + \langle J\dot{\zeta}, \delta \rangle. \tag{3.33}$$

The same equations of motions are also generated by the Hamiltonian function $K(\delta, t) - H(\zeta(t))$, so that in particular the time derivative of $K(\delta(t), t) - H(\zeta(t))$ coincides with its partial derivative with respect to time. Remark also that, using (3.32) to compute $\dot{\zeta}$, one has

$$K(\delta, t) - H(\zeta(t)) = H(\zeta(t) + \delta) - H(\zeta(t)) - dH(\zeta(t))\delta - \langle J\tilde{\mathcal{R}}, \delta \rangle \tag{3.34}$$

$$= H_0(\delta) + H_1(\zeta(t) + \delta) - H_1(\zeta(t)) - dH_1(\zeta(t))\delta - \langle J\tilde{\mathcal{R}}, \delta \rangle \tag{3.35}$$

from which

$$K(\delta, t) - H(\zeta(t)) \geq C \left| \|\delta\|_{\kappa}^2 - R^{r+3/2} \|\delta\|_{\kappa} \right| \tag{3.36}$$

and also

$$\begin{aligned} \left| \frac{\partial(K-H)}{\partial t} \right| &\leq \left| \left(dH_1(\zeta(t) + \delta) - dH_1(\zeta(t)) - d^2H_1(\zeta(t))\delta \right) \dot{\zeta} \right| + \left| \langle J\check{\mathcal{R}}, \delta \rangle \right| \\ &\leq CR^{\theta-1} \|\delta\|_{\kappa}^2 \left\| J\nabla H(\zeta(t)) - \check{\mathcal{R}}(t) \right\|_{\kappa} \\ &\leq CR^{\theta} \|\delta\|_{\kappa}^2 + CR^{r+3/2} \|\delta\|_{\kappa} . \end{aligned} \tag{3.37}$$

using these equations one immediately gets

$$\begin{aligned} \|\delta(t)\|_{\kappa}^2 &\leq C [K(\delta(t), t) - H(\zeta(t))] + CR^{r+3/2} \|\delta(t)\|_{\kappa} \\ &\leq C [K(\delta(0), 0) - H(\zeta(0))] + CR^{r+3/2} \|\delta(t)\|_{\kappa} \\ &\quad + \int_0^t \left(R^{\theta} \|\delta(s)\|_{\kappa}^2 + CR^{r+3/2} \|\delta(s)\|_{\kappa} \right) ds . \end{aligned}$$

Finally, taking $\|\delta(0)\|_{\kappa} \leq R^{r+1/2}$, as far as $\|\delta(t)\|_{\kappa} \leq \tilde{K}R^{r-\theta+3/2}$ with any fixed constant \tilde{K} one has that the above inequality gives

$$\|\delta(t)\|_{\kappa}^2 \leq \tilde{K}R^{2r-\theta+3/2} + \int_0^t \left(R^{\theta} \|\delta(s)\|_{\kappa}^2 + CR^{r+3/2} \|\delta(s)\|_{\kappa} \right) ds \tag{3.38}$$

which, using Gronwall lemma immediately gives the thesis. □

This theory applies to equation (2.17) in the Hamiltonian case. In this case the energy norm of a phase space point $x = (u, \dot{u})$ is the $H^1 \times L^2$ norm. One has the following

Theorem 3.13. *Consider equation (2.17), assume that it is Hamiltonian, that the vector a is Diophantine and that the mass m is chosen so that different frequencies are nonresonant (see Theorem 2.14). Fix an arbitrary T_0 and assume R small enough. Then there exists C s.t. if*

$$\|x_0 - \zeta_0\|_{H^1 \times L^2} \leq R^{r+1/2}, \quad \|x_0\|_{H^{s+1} \times H^s} \leq \frac{R}{C}, \quad \|\zeta_0\|_{H^{s+1} \times H^s} \leq \frac{R}{C}, \tag{3.39}$$

with s large enough, then for all t fulfilling

$$|t| < \frac{T_0}{R^{\theta}}$$

one has

$$\|x(t) - \zeta(t)\|_{H^1 \times L^2} \leq 2R^{r-\theta+1/2}. \tag{3.40}$$

Proof. The proof amounts in bounding from below T_e and \tilde{T}_e . Actually, since the equation fulfills the assumptions (Q1-Q4) one has the lower bound of $\tilde{T}_e > T_0/R^\theta$ given in remark 3.11. In order to bound from below the escape time of y just remark that the resonant form only allows exchanges of energies among modes of the same frequency, which have Fourier coefficients with the same $|k|$. Thus in particular one has that for any s the norm

$$\|y(t)\|_{H^{s+1} \times H^s}$$

is independent of time. So its escape time is actually infinite. □

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 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_{s+d}^2 to ℓ_s^2 . □

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

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

□

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

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

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*

$$\|\mathcal{R}_*(x)\|_s \leq \|x\|_{s+d}^{r_*+1}, \quad \forall x \in \mathcal{U}_{s+d}^{(0)} \tag{4.4}$$

$$\|X_*(x) - X_{\sharp}(x)\|_s \leq \frac{\|x\|_{s+\sigma+d}}{N^\sigma}, \quad \forall x \in \ell_{s+d+\sigma}^2.$$

Proof. First remark that

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

and that

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

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

$$\begin{aligned} & \sup_{x \in \mathcal{U}_{s+d}^{(0)}} \|X_*(x) - X_*(\Pi_N x)\|_s \\ & \leq \sup_{x \in \mathcal{U}_{s+d}} \|dX_*(x)\|_{s+d \rightarrow s} \|\mathbb{1} - \Pi_N\|_{s+d+\sigma \rightarrow s+d} \|x\|_{s+d+\sigma} \leq \frac{1}{N^\sigma} \|x\|_{s+\sigma+d} \end{aligned}$$

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 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 complexified 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* $\lfloor F \rfloor$ of F by

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

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

$$\sup_{\|z\|_s \leq 1} \|\lfloor F \rfloor(z)\|_{s_1} \leq (2N)^{\frac{3r+1}{2}} \sup_{\|z\|_s \leq 1} \|F(z)\|_{s_1} . \tag{4.7}$$

Remark 4.6. For the rest of the proof it is essential that the exponent of N in (4.7) does not depend neither on s nor on s_1 , and that the operator at l.h.s and the one at r.h.s. of (4.7) are estimated in the same spaces. □

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 $k! := k_{-N}! \dots k_N!$. 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 $\|z\|_s^\infty \leq R$ is equivalent to

$$|z_j| \leq R |j|^{-s} =: \rho_j .$$

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

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

which implies

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

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

$$\sup_{\|z\|_s^\infty \leq 1} \|\lfloor F \rfloor(z)\|_{s_1}^\infty \leq (2N)^r \sup_{\|z\|_s^\infty \leq 1} \|F(z)\|_{s_1}^\infty. \tag{4.9}$$

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

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

one gets the thesis. □

By Remarks 4.1 and 4.2 one has

Corollary 4.7. *The polynomials $\{X_l\}_{l=0}^{r_*}$ (cf. (4.2)) fulfill the estimate*

$$\sup_{\|z\|_s \leq R} \|\lfloor X_l \rfloor(z)\|_s \leq N^{d + \frac{3l+4}{2}} R^{l+1}, \quad l \geq 1. \tag{4.10}$$

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

Definition 4.8. 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

$$|F|_R := \sup_{\|z\|_s \leq R} \|\lfloor F \rfloor(x)\|_s. \tag{4.11}$$

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

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

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 $\phi := \phi_1 \equiv \phi_t|_{t=1}$.

Definition 4.9. The map ϕ will be called the *Lie transform* generated by W .

Remark 4.10. Given an analytic vector field F , consider the differential equation

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

Define ϕ^*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 4.11. Using the relation

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

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

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

□

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

Lemma 4.12. *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*

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

Proof. Simply remark that

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

Then the Cauchy inequality gives $\|d [F] [G](z)\|_s \leq |F|_R/d$ for any z with $\|z\|_s \leq R - d$. Thus one gets

$$\|d [F] [G]\|_s \leq \frac{1}{d} |F|_R |G|_R .$$

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 4.13. *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*

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

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

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

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

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

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

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

□

Lemma 4.15. Consider W as above and let $F \in \mathcal{A}_R$ be an analytic map. Fix $0 < d < R$ assume $\sup_{\|z\|_s \leq R} \|W(z)\|_s < d/3$, then one has

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

Lemma 4.16. 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

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

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

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

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

Lemma 4.17. *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*

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

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 4.13 one gets the result. \square

In the statement of the forthcoming iterative lemma we will use the following notations: Define $\delta = \delta(R) := R/2r_*$ where R is a positive number.

Proposition 4.18. Iterative Lemma. *Consider the system (4.1). For any $r \leq r_*$ there exists a positive $R_{*r} \ll 1$ and, for any $N > 1$ there exists an analytic transformation*

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

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

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

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

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

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

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

$$\sup_{z \in B_s(R-r\delta)} \left\| \mathcal{R}_T^{(r)}(z) \right\| \leq (RN^{\alpha+\alpha_1})^{r_*+2}. \tag{4.31}$$

Proof. We proceed by induction. First remark that the theorem is trivially true when $r = 0$ with $T^{(0)} = I$, $Z^{(0)} = 0$, $F^{(0)} = X^{(0)}$ 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 formulas (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 4.16 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

$$|W_r|_{R_r} \leq N^\alpha R^2 N^{\alpha_1} (N^{\alpha+\alpha_1} R)^r, \quad |Z_r|_{R_r} \leq N^{\alpha_1} R^2 (N^{\alpha+\alpha_1} R)^r. \tag{4.36}$$

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 4.13, denoting

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

provided $R_{*(r+1)}$ is small enough. Using (4.21), (4.29), (4.30) and Lemma 4.17 one gets

$$\begin{aligned} \left| F^{(r+1)} \right|_{R_r - \delta} &\leq \sum_{l \geq 1} N^{\alpha_1} R^2 \varsigma^l + \sum_{l \geq 1} N^{\alpha_1} R^2 \varsigma^l (RN^{\alpha+\alpha_1})^r \\ &\quad + \sum_{l \geq 2} N^{\alpha_1} R^2 \varsigma^{l-1} (RN^{\alpha+\alpha_1})^r \\ &\leq \varsigma N^{\alpha_1} R^2 \leq N^{\alpha_1} R^2 (RN^{\alpha+\alpha_1})^{r+1}. \end{aligned}$$

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:

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

Define now

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

By Lemma 4.15 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 4.15 to get

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

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

Proof of Theorem 2.11. 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 4.16 are Hamiltonian vector fields. The simplest way to verify such a property consists in introducing the variables

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

in which the symplectic form becomes $\sum_l i d\xi_l \wedge d\eta_l$ 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

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

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

and $H_Z(\xi, \eta) := \sum_{kj} H_{Zkj} \xi^k \eta^j$ 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 4.16 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. □

References

[Bam03a] D. BAMBUSI, *An averaging theorem for quasilinear Hamiltonian PDEs*, Ann. Inst. H. Poincaré Anal. Non Linéaire **4** (2003), 685–712.
 [Bam03b] D. BAMBUSI, *Birkhoff normal form for some nonlinear PDEs*, Commun. Math. Phys. **234** (2003), 253–285.
 [Bam03c] D. BAMBUSI, *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*, Discrete Contin. Dyn. Syst. Ser. B **2** (2002), 109–128.
 [BG03] D. BAMBUSI and B. GRÉBERT, *Forme normale pour NLS en dimension quelconque*, C. R. Math. Acad. Sci. Paris **337** (2003), 409–414.
 [BG04] D. BAMBUSI and B. GRÉBERT, *Birkhoff normal form for PDEs with tame modulus*, Duke Math. J. (2004), to appear.
 [BN98] D. BAMBUSI and N. N. NEKHOROSHEV, *A property of exponential stability in nonlinear wave equations near the fundamental linear mode*, Phys. D **122** (1998), 73–104.
 [Bou96] J. BOURGAIN, *Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations*, Geom. Funct. Anal. **6** (1996), 201–230.

- [Cra96] W. CRAIG, *Birkhoff normal forms for water waves*, In: “Mathematical problems in the theory of water waves” (Luminy, 1995), Vol. 200, Contemp. Math., Amer. Math. Soc., Providence, RI, 1996, 57–74.
- [CS93] W. CRAIG and C. SULEM *Numerical simulation of gravity waves*, J. Comput. Phys. **108** (1993), 73–83.
- [CSS97] W. CRAIG, U. SCHANZ and C. SULEM, *The modulational regime of three-dimensional water waves and the Davey-Stewartson system*, Ann. Inst. H. Poincaré Anal. Non Linéaire **14** (1997), 615–667.
- [CW95] W. CRAIG and P. A. WOLFOLK, *An integrable normal form for water waves in infinite depth*, Phys. D **84** (1995), 513–531.
- [DZ94] A. I. D’YACHENKO and V. E. ZAKHAROV, *Is free-surface hydrodynamics an integrable system?* Phys. Lett. A **190** (1994), 144–148.
- [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] C. FOIAS and J.-C. SAUT, *Asymptotic integration of Navier-Stokes equations with potential forces. I*, Indiana Univ. Math. J. **40** (1991), 305–320.
- [GP88] A. GIORGILLI and A. POSILICANO, *Estimates for normal forms of differential equations near an equilibrium point*, Z. Angew. Math. Phys. **39** (1988), 713–732.
- [Kat75] T. KATO, *Quasi-linear equations of evolution, with applications to partial differential equations*, In: “Spectral Theory and Differential Equations” (Proc. Sympos., Dundee, 1974; dedicated to Konrad Jörgens), Lecture Notes in Math., Vol. 448, Springer, Berlin, 1975, 25–70.
- [Kro89] M. S. KROL, *On a Galerkin-averaging method for weakly nonlinear wave equations*, Math. Methods Appl. Sci. **11** (1989), 649–664.
- [Mat01] K. MATTHIES, *Time-averaging under fast periodic forcing of parabolic partial differential equations: exponential estimates*, J. Differential Equations **174** (2001), 133–180.
- [MS03] K. MATTHIES and A. SCHEEL, *Exponential averaging for Hamiltonian evolution equations*, Trans. Amer. Math. Soc. **355** (2003), 747–773.
- [Nik86] N. V. NIKOLENKO, *The method of Poincaré normal forms in problems of integrability of equations of evolution type*, Uspekhi Mat. Nauk **41** (1986), 109–152, 263.
- [Pal96] H. PALS, *The Galerkin-averaging method for the Klein-Gordon equation in two space dimensions*, Nonlinear Anal. **27** (1996), 841–856.
- [PB05] A. PONNO and D. BAMBUSI, *Korteweg-de Vries equation and energy sharing in Fermi-Pasta-Ulam*, Chaos **15** (2005), 015107, 5.
- [Sha85] J. SHATAH, *Normal forms and quadratic nonlinear Klein-Gordon equations*, Comm. Pure Appl. Math. **38** (1985), 685–696.
- [SV87] A. C. J. STROUCKEN and F. VERHULST, *The Galerkin-averaging method for nonlinear, undamped continuous systems*, Math. Methods Appl. Sci. **9** (1987), 520–549.
- [SW00] G. SCHNEIDER and C. E. WAYNE, *Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model* In: “International Conference on Differential Equations”, Vol. 1, 2 (Berlin, 1999), World Sci. Publishing, River Edge, NJ, 2000, 390–404.
- [Zak68] V. E. ZAKHAROV *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, J. Appl. Mech. Tech. Phys. **2** (1968), 190–194.
- [Zeh78] E. ZEHNDER, *C. L. Siegel’s linearization theorem in infinite dimensions*, Manuscripta Math. **23** (1977/78), 363–371.