# Perturbation theory for PDEs

30/11/07

Dario Bambusi[a]

[a] Dipartimento di Matematica, Università degli Studi di Milano, Milano, Italia

## Contents

# 1 Glossary

**Perturbation theory.** The study of a dynamical systems which is a perturbation of a system whose dynamics is known. Typically the unperturbed system is linear or integrable.
**Normal Form.** The method of normal form consists in constructing a coordinate transformation changing the equations of a dynamical system into new ones which are as simple as possible. In Hamiltonian systems the theory is particularly effective and typically leads to a very precise description of the dynamics.
**Hamiltonian PDE.** A Hamiltonian PDE is a partial differential equations (abbreviated, PDE) which is equivalent to the Hamilton equations of a suitable Hamiltonian function. Classical examples are the nonlinear wave equation, the Nonlinear Schrödinger equation, the Kortweg de Vries equation.
**Resonance-non resonance.** A frequency vector $\{\omega_k\}_{k=1}^n$ is said to be non resonant if its components are independent over the relative integers. If on the contrary there exists a non vanishing $K \in \mathbb{Z}^n$ such that $\omega \cdot K = 0$ the frequency vector is said to be resonant. Such a property plays a fundamental role in normal form theory. Nonresonance typically implies stability.
**Actions.** The action of a harmonic oscillator is its energy divided by its frequency. It is usually denoted by $I$. The typical issue of normal form theory is that in nonresonant systems the actions remain approximatively unchanged for very long times. In resonant systems there are linear combinations of the actions with such properties.
**Sobolev space.** Space of functions which have weak derivatives enjoying suitable integrability properties. Here we will use the spaces $H^s$, $s \in \mathbb{N}$ of the functions which are square integrable together with their first $s$ weak derivatives.

# 2 Definition of the subject and its importance.

Perturbation theory for PDEs is a part of qualitative theory of differential equations. One of the most effective methods of perturbation theory is normal form theory which consists in using coordinate transformations in order to come to the description of the qualitative features of a given or generic equation. Classical normal form theory for ordinary differential equations has been used all along the last century in many different domains leading to important results in pure mathematics, celestial mechanics, plasma physics, biology, solid state physics, chemistry and many other fields.

The development of effective methods to understand the dynamics of partial differential equations is relevant in pure mathematics as well as in all the fields in which partial differential equations play an important role. Fluidodynamics, oceanography, meteorology, quantum mechanics, electromagnetic theory are just example of potential applications. More precisely the normal form theory allows to understand whether a small nonlinearity can change the dynamics of a linear PDE or not. Moreover, it allows to understand how the changes can be avoided or forced. Finally, when the changes are possible it allows to predict the behavior of the perturbed system.

# 3 Introduction

The method of normal form has been developed by Poincaré and Birkhoff between the end of the $19^{th}$ century and the beginning of the $20^{th}$ century. During the last 20 years the method has been successfully generalized to a suitable class of partial differential equations (PDEs) in finite volume (in the case of infinite volume dispersive effects appear and the theory is very different. See e.g. [58]). In this article we will give an introduction to this recent field. We will almost only deal with Hamiltonian PDEs since, on the one hand the theory for non Hamiltonian systems is a small variant of the one we will present here, and on the other hand most models are Hamiltonian.

We will start by a generalization of the Hamiltonian formalism to PDEs, followed by a review of the classical theory and by the actual generalization of normal form theory to PDEs.

In the next section we give a generalization the Hamiltonian formalism to PDEs. The main new fact is that in PDEs the Hamiltonian is usually a smooth function, but the corresponding vector field is nonsmooth (it is an operator extracting derivatives). So the standard formalism has to be slightly modified [29, 52, 61, 49, 10, 45]. Here we will present a version of the Hamiltonian formalism which is enough to cover the models of interest for local perturbation theory. To clearly illustrate the situation we will start the article with an introduction to the Lagrangian and Hamiltonian formalism for the wave equation. This will lead to the introduction of the paradigm Hamiltonian which is usually studied in this context. This will be followed by a few results on the Hamiltonian formalism that are needed for perturbation theory.

Subsequently we shortly present the standard Birkhoff normal form theory for finite dimensional systems. This is useful since all the formal aspects are equal in the classical and in the PDEs case.

Then we come to the generalization of normal form theory to PDEs. In the present paper we will concentrate almost only on the case of 1-dimensional semilinear equations. This is due to the fact that the theory of higher dimensional and quasilinear equations is still quite unsatisfactory.

In PDEs one meets essentially two kinds of difficulties. The first one is related to the existence of non smooth vector fields. The second difficulty is due to the fact that in the infinite dimensional case there are small denominators which are much worse than in the finite dimensional one.

We first present the theory for completely resonant systems [10, 14] in which the difficulties related to small denominators do not appear. It turns out that it is quite easy to obtain a normal form theorem for resonant PDEs, but the kind of normal form one gets is usually quite poor. In order to extract dynamical informations from the normal form one can only compute and study it explicitly.

Usually this is very difficult. Nevertheless in some cases it is possible and leads to quite strong results. We will illustrate such a situation by studying a nonlinear Schrödinger equation [2, 25].

For the general case there is a theorem ensuring that a generic system admits at least one family of "periodic like trajectories" which are stable over exponentially long times [15]. We will give its statement and an application to the nonlinear wave equation

$$u_{tt} - u_{xx} + \mu^2 u + f(u) = 0 \ , \tag{1}$$

with $\mu = 0$ and Dirichlet boundary conditions on a segment[55].

Then we turn to the case of nonresonant PDEs. The main difficulty is that small denominators accumulate to zero already at order 3. Such a problem has been overcome in [4, 12, 9, 43, 6] by taking advantage of the fact that the nonlinearities appearing in PDEs typically have a special form. In this case one can deduce a very precise description of the dynamics and also some interesting results of the kind of almost global existence of smooth solutions [46]. To illustrate the theory we will make reference to the nonlinear wave equation (1) with almost any $\mu$ and to the nonlinear Schrödinger equation.

Another aspect of the theory of close to integrable Hamiltonian PDEs concerns the extension of KAM theory to PDEs. We will not present it here. We just recall the most celebrated results which are those due to Kuksin [48], Wayne [60], Craig–Wayne [34], Bourgain [24, 26], Kuksin-Pöschel [50], Eliasson–Kuksin [39], Yuan [62]. All these results ensure the existence of families of quasiperiodic solutions, i.e. solutions lying on finite dimensional manifolds. We also mention the papers [21, 57] where some Cantor families of full dimension tori are constructed. We point out that in the dynamics on such $\infty$–dimensional tori the amplitude of oscillation of the linear modes decreases super exponentially with their index. A remarkable exception is provided by the paper [27] where the tori constructed are more "thick" (even if of course they lie on Cantor families).

On the contrary the results of normal form theory describe solutions starting on opens subsets of the phase space, and do not have particularly strong localizations properties with respect to the index. The price one has to pay is that the description one gets turns out to be valid only over long but finite times.

Finally we point out a related research stream that has been carried on by Bourgain [21, 22, 23, 25] who studied intensively the behavior of high Sobolev norms in close to integrable Hamiltonian PDEs (see also [28]).

# 4 The Hamiltonian formalism for PDEs

## 4.1 The gradient of a functional

**Definition 4.1.** *Consider a function $f \in C^\infty(\mathcal{U}_s, \mathbb{R})$, $\mathcal{U}_s \subset H^s(\mathbb{T})$ open, $s \geq 0$ a fixed parameter and $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ is the 1 dimensional torus. We will denote by $\nabla f(u)$ the gradient of $f$ with respect to the $L^2$ metric, namely the unique* *function such that*

$$\langle \nabla f(u), h \rangle_{L^2} = \mathrm{d}f(u)h \ , \quad \forall h \in H^s \tag{2}$$

*where*

$$\langle u, v \rangle_{L^2} := \int_{-\pi}^{\pi} u(x)v(x)\mathrm{d}x \tag{3}$$

*is the $L^2$ scalar product and $\mathrm{d}f(u)$ is the differential of $f$ at $u$. The gradient is a smooth map from $H^s$ to $H^{-s}$ (see e.g. [3]).*

*Example* 4.2. Consider the function

$$f(u) := \int_{-\pi}^{\pi} \frac{u_x^2}{2} \mathrm{d}x \ , \tag{4}$$

which is differentiable as a function from $H^s \to \mathbb{R}$ for any $s \geq 1$. One has

$$\mathrm{d}f(u)h = \int_{-\pi}^{\pi} u_x h_x \mathrm{d}x = \int_{-\pi}^{\pi} -u_{xx} h \mathrm{d}x = \langle -u_{xx}, h \rangle_{L^2} \tag{5}$$

and therefore in this case one has $\nabla f(u) = -u_{xx}$.

*Example* 4.3. Let $\mathcal{F} : \mathbb{R}^2 \to \mathbb{R}$ be a smooth function and define

$$f(u) = \int_{-\pi}^{\pi} \mathcal{F}(u, u_x)\mathrm{d}x \tag{6}$$

then the gradient of $f$ coincides with the so called functional derivative of $\mathcal{F}$:

$$\nabla f \equiv \frac{\delta \mathcal{F}}{\delta u} := \frac{\partial \mathcal{F}}{\partial u} - \frac{\partial}{\partial x}\frac{\partial \mathcal{F}}{\partial u_x} \ . \tag{7}$$

## 4.2 Lagrangian and Hamiltonian formalism for the wave equation

Until subsection 4.4 we will work at a formal level, without specifying the function spaces and the domains.

**Definition 4.4.** *Let $L(u, \dot{u})$ be a Lagrangian function, then the corresponding Lagrange equations are given*

$$\nabla_u L - \frac{\mathrm{d}}{\mathrm{d}t} \nabla_{\dot{u}} \ L = 0 \tag{8}$$

*where $\nabla_u L$ is the gradient with respect to $u$ only, and similarly $\nabla_{\dot{u}}$ is the gradient with respect to $\dot{u}$.*

*Example* 4.5. Consider the Lagrangian

$$L(u, \dot{u}) := \int_{-\pi}^{\pi} \left( \frac{\dot{u}^2}{2} - \frac{u_x^2}{2} - \mu^2 \frac{u^2}{2} - F(u) \right) \mathrm{d}x \ . \tag{9}$$

then the corresponding Lagrange equations are given by (1) with $f = -F'$.

Given a Lagrangian system with Lagrangian function $L$ one defines the corresponding Hamiltonian system as follows.

**Definition 4.6.** *Consider the momentum $v := \nabla_{\dot u} L$ conjugated to $u$; assume that $L$ is convex with respect to $\dot u$, then the Hamiltonian function associated to $L$ is defined by*

$$H(v, u) := \left[\langle v; \dot u\rangle_{L^2} - L(u, \dot u)\right]_{\dot u = \dot u(u,v)} \ . \tag{10}$$

**Definition 4.7.** *Let $H(v, u)$ be a Hamiltonian function, then the corresponding Hamilton equations are given by*

$$\dot v = -\nabla_u H \ , \quad \dot u = \nabla_v H \ . \tag{11}$$

As in the finite dimensional case one has that the Lagrange equations are equivalent to the Hamilton equation of $H$.

An elementary computation shows that for the wave equation one has $v = \dot u$ and

$$H(v, u) = \int_{-\pi}^{\pi} \left( \frac{v^2 + u_x^2 + \mu^2 u^2}{2} + F(u) \right) \mathrm{d}x \tag{12}$$

## 4.3 Canonical coordinates

Consider a Lagrangian system and let $\mathbf{e}_k$ be an orthonormal basis of $L^2$, write $u = \sum_k q_k \mathbf{e}_k$ and $\dot u = \sum_k \dot q_k \mathbf{e}_k$, then one has the following proposition

*Proposition* 4.8. The Lagrange equations (8) are equivalent to

$$\frac{\partial L}{\partial q_k} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot q_k} = 0 \tag{13}$$

*Proof.* Taking the scalar product of (8) with $\mathbf{e}_k$ one gets

$$\langle \mathbf{e}_k; \nabla_u L\rangle_{L^2} - \frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{e}_k; \nabla_{\dot u} L\rangle_{L^2} = 0$$

but one has $\langle \mathbf{e}_k; \nabla_u L\rangle = \frac{\partial L}{\partial q_k}$ and similarly for the other term. Thus the thesis follows. □

This proposition shows that, once a basis has been introduced, the Lagrange equations have the same form as in the finite dimensional case.

In the Hamiltonian case exactly the same result holds. Precisely, denoting $v := \sum_k p_k \mathbf{e}_k$ one has

*Proposition* 4.9. The Hamilton equations of a Hamiltonian function $H$, are equivalent to

$$\dot p_k = -\frac{\partial H}{\partial q_k} \ , \quad \dot q_k = \frac{\partial H}{\partial p_k} \ . \tag{14}$$

In the case of the nonlinear wave equation, in order to get a convenient form of the equations, one can choose the Fourier basis. Such a basis is defined by

$$\hat{e}_k := \begin{cases} \frac{1}{\sqrt{\pi}} \cos kx & k > 0 \\ \frac{1}{\sqrt{2\pi}} & k = 0 \\ \frac{1}{\sqrt{\pi}} \sin -kx & k < 0 \end{cases} \tag{15}$$

Thus the Hamiltonian (12) takes the form

$$H(p,q) = \sum_{k \in \mathbb{Z}} \frac{p_k^2 + \omega_k^2 q_k^2}{2} + \int_{-\pi}^{\pi} F\left(\sum_k q_k \hat{e}_k(x)\right) \mathrm{d}x \ , \tag{16}$$

where $\omega_k^2 := k^2 + \mu^2$. For the forthcoming developments it is worth to rescale the variables by defining

$$p'_k := \frac{p_k}{\sqrt{\omega_k}} \ , \quad q'_k := \sqrt{\omega_k} q_k \ , \tag{17}$$

so that, omitting primes, the Hamiltonian takes the form

$$H(p,q) = \sum_k \omega_k \frac{p_k^2 + q_k^2}{2} + H_P(p,q) \tag{18}$$

where $H_P$ has a zero of order higher than 2. *In the following we will always study systems of the form* (18). Moreover, possibly by relabeling the variables and the frequencies it is possible to reduce to the case where $k$ varies in $\mathbb{N} \equiv \{1, 2, 3...\}$. This is what we will assume in developing the abstract theory.

*Example* 4.10. An example of a different nature in which the Hamiltonian takes the form (18) is the nonlinear Schrödinger equation

$$-\mathrm{i}\dot{\psi} = \psi_{xx} + f(|\psi|^2)\psi \ , \tag{19}$$

where $f$ is a smooth function. Eq. (19) has the conserved energy functional

$$H(\psi, \bar{\psi}) := \int_{-\pi}^{\pi} \left(|\psi|^2 + F(|\psi|^2)\right) \mathrm{d}x \ , \tag{20}$$

where $F$ is such that $F' = f$. Introduce canonical coordinates $(p_k, q_k)$ by

$$\psi = \sum_{k \in \mathbb{Z}} \frac{p_k + \mathrm{i} q_k}{\sqrt{2}} \hat{e}_k \ , \tag{21}$$

then the energy takes the form (18) with $\omega_k = k^2$ and the NLS is equivalent to the corresponding Hamilton equations.

*Example* 4.11. Consider the Kortweg de Vries equation

$$u_t + u_{xxx} + u u_x = 0 \ , \tag{22}$$

in the space of functions with zero mean value. The conserved energy is given by

$$H(u) = \int_{-\pi}^{\pi} \left( \frac{u_x^2}{2} + \frac{u^3}{6} \right) \mathrm{d}x \ , \tag{23}$$

which again is also the Hamiltonian of the system. Canonical coordinates are here introduced by

$$u(x) = \sum_{k>0} \sqrt{k}(p_k \hat{e}_k + q_k \hat{e}_{-k}) \ , \tag{24}$$

in which the Hamiltonian takes the form (18) with $\omega_k = k^3$.

*Remark* 4.12. It is also interesting to study some of these equations with Dirichlet boundary conditions (DBC) typically on $[0, \pi]$. This will always be done by identifying the space of the functions fulfilling DBC with the space of the function fulfilling periodic boundary conditions on $[-\pi, \pi]$ which are skew symmetric. Similarly, Neumann boundary conditions will be treated by identifying the corresponding functions with periodic even functions. In some cases (e.g. in equation (1) with DBC and an $f$ which does not have particular symmetries) the equations do not extend naturally to the space of skew symmetric and this has some interesting consequences (see [13, 7]).

## 4.4 Basic elements of Hamiltonian formalism for PDEs

A suitable topology in the phase space is given by a Sobolev like topology.

For any $s \in \mathbb{R}$, define the Hilbert space $\ell^2_s$ of the sequences $x \equiv \{x_k\}_{k\geq 1}$ with $x_k \in \mathbb{R}$ such that

$$\|x\|_s^2 := \sum_k |k|^{2s} |x_k|^2 < \infty \tag{25}$$

and the phase spaces $\mathcal{P}_s := \ell^2_s \oplus \ell^2_s \equiv z \ni (p, q) \equiv (\{p_k\}, \{q_k\})$. In $\mathcal{P}_s$ we will sometimes use the scalar product

$$\langle (p, q), (p^1, q^1) \rangle_s := \langle p, p^1 \rangle_{\ell^2_s} + \langle q, q^1 \rangle_{\ell^2_s} \ . \tag{26}$$

In the following we will always assume that

$$|\omega_k| \leq C|k|^d \tag{27}$$

for some $d$.

*Remark* 4.13. Defining the operator $A_0 : D(A_0) \to \mathcal{P}_s$ by $A_0(p, q) = (\omega_k p_k, \omega_k q_k)$ one can write $H_0 = \frac{1}{2}\langle A_0 z; z\rangle_0$, $D(A_0) \supset \mathcal{P}_{s+d}$.

Given a smooth Hamiltonian function $\chi : \mathcal{P}_s \supset \mathcal{U}_s \to \mathbb{R}$, $\mathcal{U}_s$ being an open neighborhood of the origin, we define the corresponding Hamiltonian vector field $X_\chi : \mathcal{U}_s \mapsto \mathcal{P}_{-s}$ by

$$X_\chi \equiv \left( -\frac{\partial \chi}{\partial q_k}, \frac{\partial \chi}{\partial p_k} \right) \ . \tag{28}$$

*Remark* 4.14. Corresponding to a function $\chi$ as above we will denote by $\nabla\chi$ its gradient with respect to the $\ell^2 \equiv \ell^2_0$ metric. Defining the operator $J$ by $J(p,q) := (-q,p)$ one has $X_\chi = J\nabla\chi$.

**Definition 4.15.** *The Poisson Bracket of two smooth functions $\chi_1$, $\chi_2$ is formally defined by*

$$\{\chi_1;\chi_2\} := d\chi_1 X_{\chi_2} \equiv \langle\nabla\chi_1; J\nabla\chi_2\rangle_0 \tag{29}$$

*Remark* 4.16. As the example $\chi_1 = \sum_k kq_k$, $\chi_2 := \sum_k kp_k$ shows, there are cases where the Poisson Bracket of two functions is not defined.

For this reason a crucial role is played by the functions whose vector field is smooth.

**Definition 4.17.** *A function $\chi \in C^\infty(\mathcal{U}_s, \mathcal{P}_s)$, $\mathcal{U}_s \subset \mathcal{P}_s$ open, is said to be of class* $\mathrm{Gen}_s$*, if the corresponding Hamiltonian vector field $X_\chi$ is a smooth map from $\mathcal{U}_s \to \mathcal{P}_s$. In this case we will write $\chi \in \mathrm{Gen}_s$*

*Proposition* 4.18. Let $\chi_1 \in \mathrm{Gen}_s$. If $\chi_2 \in C^\infty(\mathcal{U}_s,\mathbb{R})$ then $\{\chi_1,\chi_2\} \in C^\infty(\mathcal{U}_s,\mathbb{R})$. If $\chi_2 \in \mathrm{Gen}_s$ then $\{\chi_1,\chi_2\} \in \mathrm{Gen}_s$.

**Definition 4.19.** *A smooth coordinate transformation $\mathcal{T} : \mathcal{P}_s \supset \mathcal{U}_s \to \mathcal{P}_s$ is said to be canonical if for any Hamiltonian function $H$ one has $X_{H\circ\mathcal{T}} = \mathcal{T}^* X_H \equiv \mathrm{d}\mathcal{T}^{-1} X_H \circ \mathcal{T}$, i.e. it transforms the Hamilton equations of $H$ into the Hamilton equations of $H \circ \mathcal{T}$.*

*Proposition* 4.20. Let $\chi_1 \in \mathrm{Gen}_s$, and let $\Phi^t_{\chi_1}$ be the corresponding time $t$ flow (which exists by standard theory). Then for $\Phi^t_{\chi_1}$ is a canonical transformation.

# 5 Normal form for finite dimensional Hamiltonian systems

Consider a system of the form (18), but with finitely many degrees of freedom, namely a system with Hamiltonian of the form (18) with

$$H_0(p,q) = \sum_{k=1}^{n} \omega_k \frac{p_k^2 + q_k^2}{2} \ , \quad \omega_k \in \mathbb{R} \tag{30}$$

and $H_P$ which is a smooth function having a zero of order at least 3 at the origin.

**Definition 5.1.** *A polynomial $Z$ will be said to be in normal form if $\{H_0; Z\} \equiv 0$.*

**Theorem 5.2.** *(Birkhoff) For any positive integer $r \geq 0$, there exist a neighborhood $\mathcal{U}^{(r)}$ of the origin and a canonical transformation $\mathcal{T}_r : \mathbb{R}^{2n} \supset \mathcal{U}^{(r)} \to \mathbb{R}^{2n}$ which puts the system (18) in Birkhoff Normal Form up to order $r$, namely s.t.*

$$H^{(r)} := H \circ \mathcal{T}_r = H_0 + Z^{(r)} + \mathcal{R}^{(r)} \tag{31}$$

*where $Z^{(r)}$ is a polynomial of degree $r+2$ which is in normal form, $\mathcal{R}^{(r)}$ is small, i.e.*

$$|\mathcal{R}^{(r)}(z)| \leq C_r \|z\|^{r+3} \ , \quad \forall z \in \mathcal{U}^{(r)} \ ; \tag{32}$$

*moreover, one has*

$$\|z - \mathcal{T}_r(z)\| \leq C'_r \|z\|^2 \ , \quad \forall z \in \mathcal{U}^{(r)} \ . \tag{33}$$

*An inequality identical to (33) is fulfilled by the inverse transformation $\mathcal{T}_r^{-1}$.*

*If the frequencies are nonresonant at order $r+2$, namely if*

$$\omega \cdot K \neq 0 \ , \quad \forall K \in \mathbb{Z}^n \ , \quad 0 < |K| \leq r+2 \tag{34}$$

*the function $Z^{(r)}$ depends on the actions*

$$I_j := \frac{p_j^2 + q_j^2}{2}$$

*only.*

*Remark* 5.3. If the nonlinearity is analytic and the frequencies are Diophantine, i.e. there exist $\gamma > 0$ and $\tau$ such that

$$|\omega \cdot K| \geq \frac{\gamma}{|K|^\tau} \ , \quad \forall K \in \mathbb{Z}^n - \{0\} \ . \tag{35}$$

then one can compute the dependence of the constant $C_r$ (cf. eq.(32)) on $r$ and optimize the value of $r$ as a function of $\|z\|$. This allows to improve (32) and to show that there exists and $r_{opt}$ such that (see e.g. [40])

$$|\mathcal{R}^{(r_{opt})}(z)| \leq C \exp\left(-\frac{c}{\|z\|^{1/(\tau+1)}}\right) \ . \tag{36}$$

In turn such an estimate is the starting point for the proof of the celebrated Nekhoroshev's theorem [53].

The idea of the proof is to construct a canonical transformation putting the system in a form which is as simple as possible, namely the normal form. More precisely one constructs a canonical transformation $\mathcal{T}^{(1)}$ pushing the non normalized part of the Hamiltonian to order four followed by a transformation $\mathcal{T}^{(2)}$ pushing it to order five and so on, thus getting $\mathcal{T}_r = \mathcal{T}^{(1)} \circ \mathcal{T}^{(2)} \circ ... \circ \mathcal{T}^{(r)}$. Each of the transformations $\mathcal{T}^{(j)}$ is constructed as the time one flow of a suitable auxiliary Hamiltonian function say $\chi_j$ (Lie transform method). It turns out that $\chi_j$ is determined as the solution of the Homological equation

$$Z_j := \{\chi_j, H_0\} + H^{(j)} \tag{37}$$

where $H^{(j)}$ is constructed recursively and $Z_j$ has to be determined together with $\chi_j$ in such a way that $\{Z_j; H_0\} = 0$ and (37) holds. In particular $H^{(1)}$ coincides with the first non vanishing term in the Taylor expansion of $H_P$.

The algorithm of solution of the Homological equation (37) involves division by the so called small denominators $\mathrm{i}\omega\cdot K$, where $K\in\mathbb{Z}^n-\{0\}$, fulfills $|K|\le j+2$ and $\omega\cdot K\neq 0$.

The above construction is more or less explicit: provided one has at disposal enough time, he can explicitly compute $Z^{(r)}$ up to any given order. In the case of nonresonant frequencies this is not needed if one wonts to understand the dynamics over long times. Indeed its features are an easy consequence of the fact that $Z^{(r)}$ depends on the actions only. A precise statement will be given in the case of PDEs. It has to be noticed that the normal form can be used also as a starting point for the construction of invariant tori through KAM theory. To this end however one has to verify a nondegeneracy condition and this requires the explicit computation of the normal form.

In the resonant case the situation is more complicated, however, it is often enough to compute the first non vanishing term of $Z^{(r)}$ in order to get relevant informations on the dynamics. This usually requires only to be able to compute the function $Z_1$ defined by (37) with $H^{(1)}$ coinciding with the first non vanishing term of the Taylor expansion of $H_P$. For a detailed analysis we referr to other sections of the Encyclopedia.

A particular case where one can use a coordinate independent formula for the computation of $Z_j$ and $\chi_j$ is the one in which the frequencies are completely resonant.

Assume that there exists $\nu>0$ and integer numbers $\ell_1,..,\ell_n$ such that

$$\omega_k=\nu\ell_k \quad \forall k=1,...,n\ . \tag{38}$$

*Remark* 5.4. Denote by $\Psi^t$ the flow of the linear system with Hamiltonian $H_0$, then one has

$$\Psi^{t+T}=\Psi^t\ ,\quad T:=\frac{2\pi}{\nu}\ ,\quad t\in\mathbb{R}\ . \tag{39}$$

Moreover in this case one has $\omega\cdot K\neq 0\Longrightarrow|\omega\cdot K|\ge\nu>0$, so there are no small denominators.

In this case one has an interesting coordinate independent formula for the solution of the homological equation (37).

*Lemma* 5.5. Let $f$ be smooth function, defined in neighborhood of the origin. Define

$$Z(z)\equiv\langle f\rangle(z):=\frac{1}{T}\int_0^T f(\Psi^t(z))\mathrm{d}t\ ,\quad \chi(z):=\frac{1}{T}\int_0^T t\left[f(\Psi^t(z))-Z(\Psi^t(z))\right]\mathrm{d}t\ , \tag{40}$$

then such quantities fulfill the equation $\{H_0,\chi\}+f=Z$.

# 6 Normal form for Hamiltonian PDEs: general comments

As anticipated in the introduction there are two problems in order to generalize Birkhoff's theorem to PDEs: (1) the existence of nonsmooth vector fields and

(2) the appearance of small denominators accumulating at zero already at order 3.

There are two reasons why (1) is a problem. The first one is that if the vector field of $\chi_1$ were not smooth then it would be nontrivial to ensure that it generates a flow, and thus that the normalizing transformation exists. The second related problem is that, if a transformation could be generated, then the Taylor expansion of the transformed Hamiltonian would contain a term of the form $\{H_1;\chi_1\} = \mathrm{d}H_1 X_{\chi_1}$, which is typically not smooth if $X_\chi$ is not smooth. Thus one has to show that the construction involves only functions which are of class $\mathrm{Gen}_s$ for some $s$ (see def. 4.17).

The difficulty related to small denominators is the following: In the finite dimensional case $\{\omega\cdot K \neq 0\ ,\ |K| \leq r+2\}$ implies $|\omega\cdot K| \geq \gamma > 0$. Thus division by $\omega \cdot K$ is a harmless operation in the finite dimensional case. In the infinite dimensional case this is no more true. For example when $\omega_k = \sqrt{k^2+\mu^2}$ one has already

$$\inf_{0\neq|K|\leq 3} |\omega \cdot K| = 0\ .$$

In order to solve such a problem one has to take advantage of a property of the nonlinearity which typically holds in PDEs and is called of having *localized coefficients*. Assuming also a suitable nonresonance property for the frequency vector one can deduce a normal form theorem identical to 5.2. The main difficulty consists in verifying the assumptions of the theorem. We will show how to verify such assumptions by the application to some typical examples.

# 7 Normal form for resonant Hamiltonian PDEs and its consequences

In the case of resonant frequencies and smooth vector field it is possible to obtain a normal form up to an exponentially small remainder.

Consider the system (18) in the phase space $\mathcal{P}_s$ with some fixed $s$. Assume that the frequencies are completely resonant, namely that (38) holds (with $k \in \mathbb{N}$); assume that $H_P \in \mathrm{Gen}_s$ and that its vector field extends to a complex analytic function in a neighborhood of the origin. Finally assume that $H_P$ has a zero of order $n \geq 3$ at the origin. Then we have the following theorem

**Theorem 7.1.** *[10, 14] There exists a neighborhood of the origin $\mathcal{U}_s \subset \mathcal{P}_s$ and an analytic canonical transformation $\mathcal{T} : \mathcal{U}_s \to \mathcal{P}_s$ with the following properties: $\mathcal{T}$ is close to identity, namely it satisfies*

$$\left\|z - \mathcal{T}(z)\right\|_s \leq C \left\|z\right\|_s^{n-1}\ . \tag{41}$$

*$\mathcal{T}$ puts the Hamiltonian in resonant normal form up to an exponentially small remainder, namely one has*

$$H \circ \mathcal{T} = H_0 + \langle H_P \rangle + Z_2 + \mathcal{R} \tag{42}$$

*where $\langle H_P \rangle$ is the average (defined by (40)) of $H_P$ with respect to the unperturbed flow; $Z_2$ is in normal form, namely $\{Z_2; H_0\} \equiv 0$, and has a zero of order $2n-2$ at the origin; $\mathcal{R}$ is an exponentially small remainder whose vector field is estimated by*

$$\|X_{\mathcal{R}}(z)\|_s \leq C \|z\|_s^{n-1} \exp\left(-\frac{C}{\|z\|_s^{n-2}}\right) .$$

*Example* 7.2. The nonlinear Schrödinger equation (19). Here one has $\ell_k = k^2$ and $\nu = 1$. The Sobolev embedding theorems ensure that the vector field of the nonlinearity is analytic if $f$ is analytic in a neighborhood of the origin. Thus theorem 7.1 applies to the NLS. To deduce dynamical consequences it is convenient to compute explicitly $\langle H_P \rangle$. Assuming $f(0) = 0$ and $f'(0) = 1$ this was done in [2] using formula (40) which gives

$$\langle H_P \rangle(z) = \frac{1}{2}\left(\sum_k I_k\right)^2 - \frac{1}{8}\sum_k |I_k|^2 \tag{43}$$

where $I_k = (p_k^2 + q_k^2)/2$ are the linear actions. Thus one has that $H_0 + \langle H_P \rangle$ is a function of the actions only, and thus it is an integrable system. It is thus natural to study the system (42) as a perturbation of such an integrable system. This was done in [2] and [25] obtaining the results we are going to state. For simplicity we will concentrate here on the case of Dirichlet boundary conditions, thus the function $\psi$ will always be assumed to be skew symmetric with respect to the origin. Define

$$\epsilon_s := \left(\frac{1}{2}\int_{-\pi}^{\pi} |\partial_x^s \psi^{(0)}(x)|^2\right)^{1/2} \tag{44}$$

i.e. the $H^s$ norm of the initial datum $\psi^{(0)}$, $s \geq 0$, and denote by $I_k(0)$ the initial value of the linear actions.

*Theorem* 7.3. *[2] Fix $N \geq 1$, then there exists a constant $\epsilon_*$, with the property that, if the initial datum $\psi^{(0)}$ is such that*

$$\epsilon_1 < \epsilon_* \ , \qquad \sum_{k \geq N+1} I_k(0)^2 \leq C \epsilon_1^4 \epsilon_1^{2-1/N} \ , \tag{45}$$

*then along the corresponding solution of* (19) *one has*

$$\sum_{k \geq 1} |I_k(t) - I_k(0)|^2 \leq C' \epsilon_1^{4+1/N} \tag{46}$$

*for the times $t$ fulfilling*

$$|t| \leq C'' \exp\left(\frac{\epsilon_*}{\epsilon_1}\right)^{1/N} .$$

This result in particular allows to control the distance in energy norm of the solution from the torus given by the intersection of the level surfaces of the actions taken at the initial time.

*Theorem* 7.4. *[25] Fix an arbitrarily large $r$, then there exists $s_r$ such that for any $s \geq s_r$ there exists $\epsilon_{*s}$, such that the following holds true: most of the initial data with $\epsilon_s < \epsilon_{*s}$ give rise to solutions with*

$$\left\|\psi(t)\right\|_s \leq C\epsilon_s \ , \quad \forall |t| \leq \frac{C}{\epsilon_s^r} \tag{47}$$

For the precise meaning of "most of the initial data" we refer to the original paper. The result is based on the proof that the considered solutions remain close in the $H^s$ topology to an infinite dimensional torus. In particular the uniform estimate of the Sobolev norm is relevant for applications to numerical analysis [30, 44].

*Example* 7.5. Consider the nonlinear wave equation (1) with $\mu = 0$. Here the frequencies are given by $|k|$ and thus they are completely resonant. Again the smoothness of the nonlinearity is ensured by Sobolev embedding theorem. In the case of DBC in order to ensure smoothness one has also to assume that the nonlinearity is odd, namely that $f(-u) = -f(u)$, then theorem 7.1 applies. However in this case the computation of $\langle H_P \rangle$ is nontrivial. It has been done (see [55]) in the case of $f(u) = \pm u^3 + O(u^4)$ and Dirichlet boundary conditions. The result however is that the function $\langle H_P \rangle$ does not have a particularly simple structure, and thus it is not easy to extract informations on the dynamics.

In order to extract informations on the dynamics consider the simplified system in which the remainder is neglected, namely the system with Hamiltonian

$$H_S := H_0 + \langle H_P \rangle + Z_2 \ . \tag{48}$$

Such a system has two integrals of motion, namely $H_0$ and $\langle H_P \rangle + Z_2$. Let $\gamma_\epsilon$ be the set of the $z$'s at which $\langle H_P \rangle + Z_2$ restricted to the surface $\mathcal{S}_\epsilon := \{z : H_0(z) = \epsilon^2\}$ has an extremum, say a maximum. Then $\gamma_\epsilon$ is an invariant set for the dynamics of $H_S$. By the invariance under the flow of $H_0$, one has that $\gamma_\epsilon$ is the union of one dimensional closed curves, but generically it is just a single closed curve. In such a situation it is also a stable periodic orbit of (48) (see [35]). Actually it is very difficult to compute $(\langle H_P \rangle + Z_2)|_{\mathcal{S}_\epsilon}$, but a maximum of such a function can be easily constructed by applying the implicit function theorem to a non degenerate maximum of $\langle H_P \rangle|_{\mathcal{S}_\epsilon}$. The addition of the remainder then modifies the dynamics only after an exponentially long time. We are now going to state the corresponding theorem.

First remark that a critical point of $\langle H_P \rangle|_{\mathcal{S}_1}$ is a solution $z_a$ of the system

$$\lambda_a A_0 z_a + \nabla \langle H_P \rangle (z_a) = 0 \ , \quad H_0(z_a) = 1 \tag{49}$$

where we used the notations of remarks 4.13 and 4.14. Here $\lambda_a$ is clearly the Lagrange multiplier. The closed curve $\gamma_a := \bigcup_t \Psi^t(z_a)$ is (the trajectory of) a periodic solution of $H_0 + \langle H_P \rangle$. Consider now the linear operator $B_a := \mathrm{d}(\nabla \langle H_P \rangle)(z_a)$.

**Definition 7.6.** *The critical point $z_a$ will be said to be non degenerate if the system*

$$\lambda_a A_0 h + B_a h = 0 \ , \quad \langle A_0 z_a ; h \rangle_0 = 0 \tag{50}$$

*has at most one solution.*

Under the assumptions of theorem 7.7 below it is easy to prove that $\gamma_a$ is a smooth curve and that its tangent vector $h_a := \frac{\mathrm{d}}{\mathrm{d}t}\Psi^t(z_a)\big|_{t=0}$ is a solution of (50).

**Theorem 7.7.** *([14, 15]) Assume that $X_{H_P} \in \mathrm{Gen}$ for any $s$ large enough, assume also that there exists a non degenerate maximum $z_a$ of $\langle H_P \rangle|_{\mathcal{S}_1}$, then there exists a constant $\epsilon_*$, such that the following holds true: consider a solution $z(t)$ of the Hamilton equation of (18) with initial datum $z_0$; if there exists $\epsilon < \epsilon_*$, such that*

$$d_E(\epsilon\gamma_a, z_0) \leq C\epsilon^n \ , \tag{51}$$

*then one has*

$$d_E(\epsilon\gamma_a, z(t)) \leq C'\epsilon^n \ , \tag{52}$$

*for all times $t$ with $|t| \leq \frac{C}{\epsilon^{r-1}} \exp\left(\frac{\epsilon_*}{\epsilon}\right)^{n-1}$. Here $d_E$ is the distance in the energy norm.*

Such a theorem does not ensure that there exist periodic orbits of the complete system, but just a family of closed curves with the property that starting close to it one remains close to it for exponentially long times. Some results concerning the existence of true periodic orbits close to such periodic like trajectories can also be proved (see e.g. [51, 16, 19, 42, 20]).

*Example* 7.8. In the paper [55] it has been proved that the non degeneracy assumption (50) of theorem 7.7 holds for the equation (1) with $f(u) = \pm u^3$+higher order terms and Dirichlet boundary conditions. In the case of such an equation an extremum of $\langle H_P \rangle|_{\mathcal{S}_1}$ is given by

$$u(x) = V_m \mathrm{sn}(\mathtt{w}x, \mathtt{m}) \ , \quad v(x) \equiv 0 \ ,$$

with $V_m$, $\mathtt{w}$ and $\mathtt{m}$ suitable constants, and sn the Jacobi elliptic sine. Therefore the curve $\gamma_a$ is the phase space trajectory of the solution of the linear wave equation with such an initial datum. There are no other extrema of $\langle H_P \rangle|_{\mathcal{S}_1}$. Thus the theorem 7.7 ensures that solutions starting close to a rescaling of such a curve remain close to it for very long times.

# 8 Normal form for Hamiltonian nonresonant PDEs

## 8.1 A statement

We turn now to the nonresonant case. The theory we will present has been developed in [9, 43, 6], and is closely related to the one of [4, 11, 37]. First we introduce the class of equations to which the theory applies. To this end it is useful to treat the $p$'s and the $q$'s exactly on an equal footing so we will denote by $z \equiv (z_k)_{k\in\bar{\mathbb{Z}}}$, $\bar{\mathbb{Z}} := \mathbb{Z} - \{0\}$ the set of all the variables, where

$$z_{-k} := p_k \ , \quad z_k := q_k \quad k \geq 1 \ .$$

Given a polynomial function $f : \mathcal{P}_\infty \to \mathbb{R}$ of degree $r$ one can decompose it as follows

$$f(z) = \sum_{k_1,\ldots,k_r} f_{k_1,\ldots,k_r} z_{k_1} \ldots z_{k_r} \tag{53}$$

We will assume suitable localization properties for the coefficients $f_{k_1,\ldots,k_r}$ as functions of the indexes $k_1, ..., k_r$.

**Definition 8.1.** *Given a multi-index $k \equiv (k_1, ..., k_r)$, let $(k_{i_1}, k_{i_2}, k_{i_3} ..., k_{i_r})$ be a reordering of $k$ such that*

$$|k_{i_1}| \geq |k_{i_2}| \geq |k_{i_3}| \geq \ldots \geq |k_{i_r}| \ .$$

*We define $\mu(k) := |k_{i_3}|$ and*

$$S(k) := \mu(k) + ||k_{i_1}| - |k_{i_2}|| \ . \tag{54}$$

**Definition 8.2.** *Let $f : \mathcal{P}_\infty \to \mathbb{R}$ be a polynomial of degree $r$. We will say that $f$ has localized coefficients if there exists $\nu \in [0, +\infty)$ such that $\forall N \geq 1$ there exists $C_N$ such that for any choice of the indexes $k_1, ..., k_r$ following inequality holds*

$$|f_{k_1,\ldots,k_r}| \leq C_N \frac{\mu(k)^{\nu+N}}{S(k)^N} \ . \tag{55}$$

**Definition 8.3.** *A function $f \in \mathrm{Gen}_s$ for any $s$ large enough, will be said to have localized coefficients if all the terms of its Taylor expansion have localized coefficients.*

Some important properties of functions with localized coefficients are given by

**Theorem 8.4.** *Let $f : \mathcal{P}_\infty \to \mathbb{R}$ be a polynomial of degree $r$ with localized coefficients, then there exists $s_0$ such that for any $s \geq s_0$ the vector field $X_f$ extends to a smooth map from $\mathcal{P}_s$ to itself; moreover the following estimate holds*

$$\|X_f(z)\|_s \leq C \|z\|_s \|z\|_{s_0}^{r-2} \ . \tag{56}$$

In particular it follows that a function with localized coefficients is of class $\mathrm{Gen}_s$ for any $s \geq s_0$.

**Theorem 8.5.** *The Poisson Bracket of two functions with localized coefficients has localized coefficients.*

In order to develop perturbation theory we need also a quantitative nonresonance condition.

**Definition 8.6.** *Fix a positive integer $r$. The frequency vector $\omega$ is said to fulfill the* property ($r$–NR) *if there exist $\gamma > 0$, and $\alpha \in \mathbb{R}$ such that for any $N$ large enough one has*

$$\left| \sum_{k \geq 1} \omega_k K_k \right| \geq \frac{\gamma}{N^\alpha} \ , \tag{57}$$

*for any $K \in \mathbb{Z}^\infty$, fulfilling $0 \neq |K| := \sum_k |K_k| \leq r+2$, $\sum_{k>N} |K_k| \leq 2$.*

It is easy to see that under this condition one can solve the homological equation and that, if the known term of the equation has localized coefficients, then also the solution has localized coefficients.

**Theorem 8.7.** *([12, 9]) Fix $r \geq 1$, assume that the frequencies fulfill the non-resonance condition (r-NR); assume that $H_P$ has localized coefficients. Then there exists a finite $s_r$ a neighborhood $\mathcal{U}_{s_r}^{(r)}$ of the origin in $\mathcal{P}_{s_r}$ and a canonical transformation $\mathcal{T} : \mathcal{U}_{s_r}^{(r)} \to \mathcal{P}_{s_r}$ which puts the system in normal form up to order $r+3$, namely*

$$H^{(r)} := H \circ \mathcal{T} = H_0 + Z^{(r)} + \mathcal{R}^{(r)} \tag{58}$$

*where $Z^{(r)}$ has localized coefficients and is a function of the actions $I_k$ only; $\mathcal{R}^{(r)}$ has a small vector field, i.e.*

$$\left\| X_{\mathcal{R}^{(r)}}(z) \right\|_{s_r} \leq C \left\| z \right\|_{s_r}^{r+2} \quad , \quad \forall z \in \mathcal{U}_{s_r}^{(r)} \ ; \tag{59}$$

*one has*

$$\left\| z - \mathcal{T}_r(z) \right\|_{s_r} \leq C \left\| z \right\|_{s_r}^{2} \quad , \quad \forall z \in \mathcal{U}_{s_r}^{(r)} \ . \tag{60}$$

*An inequality identical to* (60) *is fulfilled by the inverse transformation $\mathcal{T}_r^{-1}$. Finally for any $s \geq s_r$ there exists a subset $\mathcal{U}_s^{(r)} \subset \mathcal{U}_{s_r}^{(r)}$ open in $\mathcal{P}_s$ such that the restriction of the canonical transformation to $\mathcal{U}_s^{(r)}$ is analytic also as a map from $\mathcal{P}_s \to \mathcal{P}_s$ and the inequalities* (59) *and* (60) *hold with $s$ in place of $s_r$.*

This theorem allows to give a very precise description of the dynamics.

*Proposition* 8.8. Under the same assumptions of theorem 8.7, $\forall s \geq s_r$ there exists $\epsilon_{*s}$ such that, if the initial datum fulfills $\epsilon := \left\| z_0 \right\|_s < \epsilon_{*s}$, then one has

$$\left\| z(t) \right\|_s \leq 4\epsilon \ , \quad \sum_k k^{2s} \left| I_k(t) - I_k(0) \right| \leq C\epsilon^3 \tag{61}$$

for all the times $t$ fulfilling $|t| \leq \epsilon^{-r}$. Moreover there exists a smooth torus $\mathbb{T}_0$ such that, $\forall M \leq r$

$$d_s(z(t), \mathbb{T}_0) \leq C\epsilon^{(M+3)/2} \quad , \quad \text{for} \quad |t| \leq \frac{1}{\epsilon^{r-M}} \tag{62}$$

where $d_s(.,.)$ is the distance in $\mathcal{P}_s$.

A generalization to the resonant or partially resonant case is easy to be obtained and can be found e.g. in [11].

## 8.2 Verification of the property of localization of coefficients

The property of localization of coefficients is quite abstract. We illustrate through a few examples some ways to verify it.

*Example* 8.9. Consider the nonlinear wave equation (1) with Neumann boundary conditions on $[0, \pi]$. We recall that the corresponding space of functions will be considered as a subset of the space of periodic functions.

Consider the Taylor expansion of the nonlinearity, i.e. write $F(u) = \sum_{r\geq 3} c_r \int_{-\pi}^{\pi} u^r$. Then one has to prove that the functions $f_r(u) \equiv \int_{-\pi}^{\pi} u^r$ have localized coefficients. The coefficients $f_{k_1,...,k_r}$ are given by

$$f_{k_1,...,k_r} = \int_{-\pi}^{\pi} \cos(k_1 x) \cos(k_2 x) ... \cos(k_r x) \mathrm{d}x \ . \tag{63}$$

One could compute and estimate such a quantity directly, but it is easier to proceed in a different way: to show that $f_3$ has localized coefficients and then to use theorem 8.5 to show that each $f_r$ has localized coefficients for any $r$. This is the path we will follow. Consider

$$f_{k_1,k_2,k_3} = \int_{-\pi}^{\pi} \cos(k_1 x) \cos(k_2 x) \cos(k_3 x) \mathrm{d}x \tag{64}$$

since the estimate (55) is symmetric with respect to the indexes, we can assume that they are ordered, $k_1 \geq k_2 \geq k_3$, so that (64) $= \pi \delta_{k_1}^{k_2+k_3}/2$, $\mu(k) = k_3$, $S(k) = k_3 + k_1 - k_2$ from which one immediately sees that (55) holds with $\nu = 0$. As a consequence one also gets that the function $g_3 := \int_{-\pi}^{\pi} v u^2$ has localized coefficients. Since $\{g_3; f_r\} = r f_{r+1}$, by induction theorem 8.5 ensures that $f_r$ has localized coefficients for any $r$.

Often it is impossible to explicitly compute the coefficients $f_{k_1,...,k_r}$, so one needs a different way to verify the property.

*Example* 8.10. Consider the nonlinear wave equation

$$u_{tt} - u_{xx} + V u = f(u) \tag{65}$$

with Neumann boundary conditions. Here $V$ is a smooth, even, periodic potential. The Hamiltonian reduces to the form (18) by introducing the variables $q_k$ by $u(x) = \sum_k q_k \varphi_k(x)$ where $\varphi_k(x)$ are the eigenfunctions of the Sturm Liouville operator $-\partial_{xx} + V$, and similarly for $v$. In such a case one has

$$f_{k_1,k_2,k_3} = \int_{-\pi}^{\pi} \varphi_{k_1}(x) \varphi_{k_2}(x) \varphi_{k_3}(x) \mathrm{d}x \ . \tag{66}$$

Here the idea is to consider (66) as the matrix element $L_{k_1,k_2}$ of the operator $L$ of multiplication by $\varphi_{k_3}(x)$ on the basis of the eigenfunctions of the operator $S := -\partial_{xx} + V$. The key idea is to proceed as follows.

Let $L$ be a linear operator which maps $D(S^r)$ into itself for all $r \geq 0$, and define the sequence of operators

$$L_N := [S, L_{N-1}] \ , \quad L_0 := L \ . \tag{67}$$

*Lemma* 8.11. Let $S$ be as above, then, for any $N \geq 0$ one has

$$|L_{k_1,k_2}| = |\langle L\varphi_{k_1}; \varphi_{k_2}\rangle| \leq \frac{1}{|\lambda_{k_1} - \lambda_{k_2}|^N} |\langle L_N \varphi_{k_1}; \varphi_{k_2}\rangle| \tag{68}$$

where $\lambda_{k_j}$ is the eigenvalue of $S$ corresponding to $\varphi_{k_j}$.

Then, in order to conclude the verification of the property of localization of the coefficients one has just to compute $L_N$ and to estimate the scalar product product at r.h.s. All the computations can be found in [6].

*Example* 8.12. A third example where the verification of the property of localization of coefficients goes almost in the same way is the nonlinear Schrödinger equation

$$\mathrm{i}\dot{\psi} = -\psi_{xx} + V\psi + \frac{\partial F(\psi, \bar{\psi})}{\partial \bar{\psi}} \tag{69}$$

with Dirichlet Boundary conditions on $[0, \pi]$. Here one has to assume that $V$ is a smooth even potential and that $F$ is smooth and fulfills $F(-\psi, -\bar{\psi}) = F(\psi, \bar{\psi})$ (this is required in order to leave invariant the space of skew symmetric functions, see remark 4.12). Here the variables $(p, q)$ are introduced by

$$\psi = \sum_{k \in \mathbb{Z}} \frac{p_k + \mathrm{i}q_k}{\sqrt{2}} \varphi_k \ , \tag{70}$$

with $\varphi_k$ the eigenfunctions of $S$ with Dirichlet boundary conditions. Here the Taylor expansion of the nonlinearity has only even terms. Thus the building block for the proof of the property of localization of coefficients is the operator $L$ of multiplication by $\varphi_{k_3}\varphi_{k_4}$. Then, mutatis mutandis the proof goes as in the previous case.

## 8.3 Verification of the nonresonance property

Finally in order to apply theorem 8.7 one has to verify the nonresonance property ($r$–NR). As usual in dynamical system this is done by tuning the frequencies using parameters. In the case of the nonlinear wave equation (1) one can use the mass $\mu$.

**Theorem 8.13.** *[4, 12, 36] There exists a zero measure set $S \subset \mathbb{R}$ such that, if $\mu \in \mathbb{R} - S$, then the frequencies $\omega_k = \sqrt{k^2 + \mu^2}$, $k \geq 1$ fulfill the condition (r-NR) for any $r$.*

Thus the theorem 8.7 applies to the equation (1) with almost any mass.

A similar result holds for the equation (65), where the role of the mass is played by the average of the potential.

The situation of the nonlinear Schrödinger is more difficult. Here one can use the Fourier coefficients of the potential as parameters.

Fix $\sigma > 0$ and, for any positive $R$ define the space of the potentials, by

$$\mathcal{V}_R := \left\{ V(x) = \sum_{k\geq 1} v_k \cos kx \mid v_k' := v_k R^{-1} e^{\sigma k} \in \left[ -\frac{1}{2}, \frac{1}{2} \right] \text{ for } k \geq 1 \right\} \quad (71)$$

Endow such a space with the product normalized probability measure.

**Theorem 8.14.** *([12], see also [21]) For any $r$ there exists a positive $R$ and a set $\mathcal{S} \subset \mathcal{V}_R$ such that property (r-NR) holds for any potential $V \in \mathcal{S}$ and $|\mathcal{V}_R - \mathcal{S}| = 0$.*

So, provided the potential is chosen in the considered set theorem 8.7 applies also to the equation (69).

We point out that the proof of theorem 8.13 and of theorem 8.14 consists essentially of two steps. First one proves that for most values of the parameters one has

$$\left| \sum_{k\geq 1} \omega_k K_k \right| \geq \frac{\gamma}{N^\alpha} \ , \quad (72)$$
$$\forall K \in \mathbb{Z}^\infty \ , \quad \text{with} \quad 0 \neq |K| := \sum_k |K_k| \leq r+2$$

and then one uses the asymptotic of the frequencies, namely $\omega_k \sim a k^d$ with $d \geq 1$ in order to get the result.

# 9 Non Hamiltonian PDEs

In this section we will present some results for the non Hamiltonian case.

It is useful to complexify the phase space. So all along this section we will denote by $\mathcal{P}_s$ the space of the complex sequences $\{z_k\}$ whose norm (defined by (25)) is finite.

In the space $\mathcal{P}_s$ consider a system of differential equations of the form

$$\dot{z}_k = \lambda_k z_k + P_k(z) \ , \quad k \in \mathbb{Z} - \{0\} \quad (73)$$

where $\lambda_k$ are complex numbers and $P(z) \equiv \{P_k(z)\}$ has a zero of order at least 2 at the origin. Moreover we will assume $P$ to be a complex analytic map from a neighborhood of the origin of $\mathcal{P}_s$ to $\mathcal{P}_s$. The quantities $\lambda_k$ are clearly the eigenvalues of the linear operator describing the linear part of the system, for this reason they will be called "the eigenvalues".

*Example* 9.1. A system of the form (18) with $H_P$ having a vector field which is analytic. The corresponding Hamilton equations have the form (73) with $\lambda_k = -\lambda_{-k} = \mathrm{i}\omega_k$, $k \geq 1$.

*Example* 9.2. Consider the following nonlinear heat equation with periodic boundary conditions on $[-\pi, \pi]$

$$u_t = u_{xx} - V(x)u + f(u) \ . \tag{74}$$

If $f$ is analytic then it can be given the form (73) by introducing the basis of the eigenfunctions $\varphi_k$ of the Sturm Liouville operator $S := -\partial_{xx} + V$, i.e. denoting $u = \sum_k z_k \varphi_k$. In this case the the eigenvalues $\lambda_k$ are the opposite of the periodic eigenvalues of $S$. Thus in particular one has $\lambda_k \in \mathbb{R}$ and $\lambda_k \sim -k^2$.

In this context one has to introduce a suitable concept of nonresonance:

**Definition 9.3.** *A sequence of eigenvalues is said to be resonant if there exists a sequence of integer numbers $K_k \geq 0$ and an index $i$ such that*

$$\sum_k \lambda_k K_k - \lambda_i = 0 \ . \tag{75}$$

In the finite dimensional case the most celebrated results concerning systems of the form (73) are the Poincaré theorem, the Poincaré–Dulac theorem and the Siegel theorem [1]. The Poincaré theorem is of the form of Birkhoff's theorem 5.2, while Poincaré Dulac and Siegel theorem guarantee (under suitable assumptions) that there exists an analytic coordinate transformation reducing the system to its normal form or linear part (no remainder!).

At present there is not a satisfactory extension of Poincaré Dulac theorem to PDEs (some partial results have been given in [41]). We are now going to state a known extension of Siegel theorem to PDEs.

**Theorem 9.4.** *[54, 64] Assume that the eigenvalues fulfill the Diophantine type condition*

$$\left| \sum_k \lambda_k K_k - \lambda_i \right| \geq \frac{\gamma}{|K|^\tau} \ , \quad \forall i, K \ \textit{with} \ 2 \leq |K| \ , \tag{76}$$

*where $\gamma > 0$ and $\tau \in \mathbb{R}$ are suitable parameters; then there exists an analytic coordinate transformation defined in a neighborhood of the origin, such that the system* (73) *is transformed into its linear part*

$$\dot{z}_k = \lambda_k z_k \tag{77}$$

The main remark concerning this theorem is that the condition (76) is only exceptionally satisfied. If $\lambda \in \mathbb{C}^n$ then condition (76) is generically satisfied only if $\tau > (n-2)/2$. Nevertheless some examples where such an equation is satisfied are known [54].

The formalism of sect. 8 can be easily generalized to the non Hamiltonian case giving rise to a generalization of Poincaré's theorem that we are going to state.

Given a polynomial map $P : \mathcal{P}_\infty \to \mathcal{P}_{-\infty}$ one can expand it on the canonical basis $\mathbf{e}_k$ of $\mathcal{P}_0$ as follows

$$P(z) = \sum_{k_1,..,k_r,i} P^i_{k_1,..,k_r} z_{k_1} ... z_{k_r} \mathbf{e}_i \ , \quad P^i_{k_1,..,k_r} \in \mathbb{C} \tag{78}$$

**Definition 9.5.** *A polynomial map $P$ is said to have localized coefficients if there exists $\nu \in [0, +\infty)$ such that $\forall N \geq 1$ there exists $C_N$ such that for any choice of the indexes $k_1, ..., k_r, i$ following the inequality holds*

$$\left| P^i_{k_1,...,k^i_r} \right| \leq C_N \frac{\mu(k,i)^{\nu+N}}{S(k,i)^N} \ , \tag{79}$$

*where $(k, i) = (k_1, .., k_r, i)$. A map is said to have localized coefficients if for any $s$ large enough it is smooth as a map from $\mathcal{P}_s$ to itself and if each term of its Taylor expansion has localized coefficients.*

**Definition 9.6.** *The eigenvalues are said to be strongly nonresonant at order $r$ if for any $N$ large enough , any $K = (K_{k_1}, ..., K_{k_r})$ and any index $i$ such that $|K| \leq r$ and there are at most two of the indexes $k_1, .., k_r, i$ larger than $N$ the following inequality holds*

$$\left| \sum_k \lambda_k K_k - \lambda_i \right| \geq \frac{\gamma}{N^\alpha} \ . \tag{80}$$

**Theorem 9.7.** *Assume that the nonlinearity has localized coefficients and that the eigenvalues are strongly nonresonant at order $r$, then there exists $s_r$ and an analytic coordinate transformation $\mathcal{T}_r : \mathcal{U}_{s_r} \to \mathcal{P}_{s_r}$ which transforms the system (73) to the form*

$$\dot{z}_k = \lambda_k z_k + \mathcal{R}_k(z) \ , \tag{81}$$

*where the following inequality holds*

$$\|\mathcal{R}(z)\|_{s_r} \leq C \|z\|^r_{s_r} \ . \tag{82}$$

# 10 Extensions and related results

The theory presented here applies to quite general semilinear equations in one space dimension. At present a satisfactory theory applying to quasilinear equations and/or equations in more than one space dimensions is not available. The main difficulty for the extension of the theory to semilinear equations in higher space dimension is related to the nonresonance condition. The general theory can be easily extended to the case where the differences $\omega_k - \omega_l$ between frequencies accumulate only at a set constituted by isolated points.

*Example* 10.1. Consider the nonlinear wave equation on the $d$-dimensional sphere

$$u_{tt} - \Delta_g u + \mu^2 u = f(x, u) \ , \quad x \in S^d \tag{83}$$

with $\Delta_g$ the Laplace Beltrami operator; the frequencies are given by $\omega_k = \sqrt{k(k + d - 1) + \mu^2}$ and their differences accumulate only at integers. A version of theorem 8.7 applicable to (83) was proved in [9]. As a consequence in particular one has a lower bound on the existence time $t$ of the small amplitude solutions of the form $|t| \geq \epsilon^{-r}$, where $\epsilon$ is proportional to a high Sobolev norm of the initial datum. An extension to $u_{tt} - \Delta_g u + V u = f(x, u)$ is also known.

*Example* 10.2. A similar result was proved in [11] for the nonlinear Schrdinger equation

$$-\mathrm{i}\dot{\psi} = -\Delta\psi + V * \psi + f(|\psi|^2)\psi \ , \quad x \in \mathbb{T}^d \tag{84}$$

where the star denotes convolution.

The only general result available at present for quasilinear system is the following theorem.

**Theorem 10.3.** *[5] Fix $r \geq 1$, assume that the frequency vector fulfills condition* (72) *and that there exists $d_1$ such that the vector field of $H_P$ is smooth as a map from $\mathcal{P}_{s+d_1}$ to $\mathcal{P}_s$ for any $s$ large enough. Then the same result of theorem 8.7 holds, but the functions do not necessarily have localized coefficients and, for any $s$ large enough* **the remainder is estimated by**

$$\left\| X_{\mathcal{R}^{(r)}}(z) \right\|_s \leq C \left\| z \right\|_{s+d_r}^{r+2} \ , \quad \forall z \in \mathcal{U}_s^{(r)} \ ; \tag{85}$$

*where $d_r$ is a large positive number.*

The estimate (85) shows that the remainder is small only when considered as an operator extracting a lot of derivatives. In particular it is non trivial to use such a theorem in order to extract informations on the dynamics. Following the approach of [8] and [5] this can be done using the normal form to construct approximate solutions and suitable versions of the Gronwall lemma to compare it with solutions of the true system. This however allows to control the dynamics only over times of the order of $\epsilon^{-1}$, $\epsilon$ being again a measure of the size of the initial datum. Such a theory has been applied to quasilinear wave equations in [5] and to the Fermi Pasta Ulam problem in [17]. Among the large number of papers containing related results we recall [41, 47, 59, 56]. A stronger result for the nonlinear wave equation valid over times of order $\epsilon^{-2}$ can be found in [36].

# 11 Future Directions.

Future directions of research include both purely theoretical aspects and applications to other sciences.

From a purely theoretical point of view the most important open problems pertain the validity of normal form theory for equation in which the nonlinearity involves derivatives, and for general equations in more than one space dimensions.

These results would be particularly important since the kind of equations appearing in most domains of physics are quasilinear and higher dimensional.

Concerning applications we would like to describe a few of them which would be of interest.

- Water wave problem. The problem of description of the free surface of a fluid has been shown to fit in the scheme of Hamiltonian dynamics [63]. Normal form theory could allow to extract the relevant informations on the dynamics in different situations [38, 31], ranging from the theory of

Tsunamis [32] to the theory of fluid interface, which is relevant e.g. to the construction of oil platforms [33].

- Quantum mechanics. A Bose condensate is known to be well described by the Gross Pitaevskii equation. When the potential is confining, such an equation, is of the form (18). Normal form theory has already been used for some preliminary results [18], but a systematic investigation could lead to interesting new results.

- Electromagnetic theory and magnetohydrodynamics. The equations have a Hamiltonian form; normal form theory could help to describe some instability arising in plasmas.

- Elastodynamics. Here one of the main theoretical open problem is that of the stability of equilibria which are a minimum of the energy. The problem is that in higher dimensions the conservation of energy does not ensure enough smoothness of the solution to ensure stability. Such a problem is of the same kind of the one solved in [9] when dealing with the existence times of the nonlinear wave equation.

# References


[1] V. Arnold, *Chapitres supplémentaires de la théorie des équations différentielles ordinaires*, "Mir", Moscow, 1984.

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

[3] ———, *On the Darboux theorem for weak symplectic manifolds*, Proc. Amer. Math. Soc. **127** (1999), no. 11, 3383–3391.

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

[5] ———, *Galerkin averaging method and Poincaré normal form for some quasilinear PDEs*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **4** (2005), no. 4, 669–702.

[6] ———, *A Birkhoff normal form theorem for some semilinear PDEs*, Preprint. (2007).

[7] D. Bambusi, A. Carati, T. Penati, *On the relevance of boundary conditions for the FPU paradox*, preprint 2007. Istit. Lombardo Accad. Sci. Lett. Rend. A, to appear.

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

[9] D. Bambusi, J.M. Delort, B. Grébert, and J. Szeftel, *Almost global existence for Hamiltonian semi-linear Klein-Gordon equations with small Cauchy data on Zoll manifolds*, Comm. Pure Appl. Math. **60** (2007), no. 11, 1665,1690.

[10] D. Bambusi and A. Giorgilli, *Exponential stability of states close to resonance in infinite-dimensional Hamiltonian systems*, J. Statist. Phys. **71** (1993), no. 3-4, 569–606.

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

[12] D. Bambusi and B. Grébert, *Birkhoff normal form for partial differential equations with tame modulus*, Duke Math. J. **135** (2006), no. 3, 507–567.

[13] D. Bambusi, D. Muraro, T. Penati, *Numerical studies on boundary effects on the FPU paradox*, Preprint 2007. Phys. Lett. A. to appear.

[14] D. Bambusi and N. N. Nekhoroshev, *A property of exponential stability in the nonlinear wave equation close to main linear mode*, Physica D **122** (1998), 73–104.

[15] D. Bambusi and N. N. Nekhoroshev, *Long time stability in perturbations of completely resonant PDE's*, Acta Appl. Math. **70** (2002), no. 1-3, 1–22, Symmetry and perturbation theory.

[16] D. Bambusi and S. Paleari, *Families of periodic orbits for resonant PDE's*, J. Nonlinear Science **11** (2001), 69–87.

[17] D. Bambusi and A. Ponno, *On metastability in FPU*, Comm. Math. Phys. **264** (2006), no. 2, 539–561.

[18] D. Bambusi and A. Sacchetti, *Exponential times in the one-dimensional Gross–Petaevskii equation with multiple well potential*, Commun. Math. Phys. **234** (2007), no. 2, 1,36.

[19] M. Berti and P. Bolle, *Periodic solutions of nonlinear wave equations with general nonlinearities*, Commun. Math. Phys. **243** (2003), 315–328.

[20] ———, *Cantor families of periodic solutions for completely resonant nonlinear wave equations*, Duke Math. J. **134** (2006), no. 2, 359–419.

[21] J. Bourgain, *Construction of approximative and almost-periodic solutions of perturbed linear Schrödinger and wave equations*, Geometric and Functional Analysis **6** (1996), 201–230.

[22] ———, *On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE*, Internat. Math. Res. Notices (1996), no. 6, 277–304.

[23] J. Bourgain, *On growth in time of Sobolev norms of smooth solutions of nonlinear Schrödinger equations in* $\mathbf{R}^D$, J. Anal. Math. **72** (1997), 299–310.

[24] J. Bourgain, *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Shödinger equation*, Ann. of Math. **148** (1998), 363–439.

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

[26] ———, *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies, vol. 158, Princeton University Press, Princeton, NJ, 2005.

[27] J. Bourgain, *On invariant tori of full dimension for 1D periodic NLS*, J. Funct. Anal. **229** (2005), no. 1, 62–94.

[28] J. Bourgain and V. Kaloshin, *On diffusion in high-dimensional Hamiltonian systems*, J. Funct. Anal. **229** (2005), no. 1, 1–61.

[29] P. R. Chernoff and J. E. Marsden, *Properties of infinite dimensional Hamiltonian systems*, Springer-Verlag, Berlin, 1974, Lecture Notes in Mathematics, Vol. 425.

[30] D. Cohen, E. Hairer, and C. Lubich, *Long-time analysis of nonlinearly perturbed wave equations via modulated fourier expansions*, (2006).

[31] W. Craig, *Birkhoff normal forms for water waves*, Mathematical problems in the theory of water waves (Luminy, 1995), Contemp. Math., vol. 200, Amer. Math. Soc., Providence, RI, 1996, pp. 57–74.

[32] ———, *Surface water waves and tsunamis*, J. Dynam. Differential Equations **18** (2006), no. 3, 525–549.

[33] W. Craig, P. Guyenne, and H. Kalisch, *Hamiltonian long-wave expansions for free surfaces and interfaces*, Comm. Pure Appl. Math. **58** (2005), no. 12, 1587–1641.

[34] W. Craig and C. E. Wayne, *Newton's method and periodic solutions of nonlinear wave equations*, Comm. Pure Appl. Math. **46** (1993), 1409–1498.

[35] G. F. Dell'Antonio, *Fine tuning of resonances and periodic solutions of Hamiltonian systems near equilibrium*, Comm. Math. Phys. **120** (1989), no. 4, 529–546.

[36] J. M. Delort and J. Szeftel, *Long–time existence for small data nonlinear Klein–Gordon equations on tori and spheres*, Internat. Math. Res. Notices **37** (2004), 1897–1966.

[37] J.-M. Delort and J. Szeftel, *Long-time existence for semi-linear Klein-Gordon equations with small Cauchy data on Zoll manifolds*, Amer. J. Math. **128** (2006), no. 5, 1187–1218.

[38] A. I. Dyachenko and V. E. Zakharov, *Is free-surface hydrodynamics an integrable system?*, Phys. Lett. A **190** (1994), 144–148.

[39] H. L. Eliasson and S. B. Kuksin, *KAM for non-linear Schroedinger equation*, Preprint (2006). Ann. of Math, to appear.

[40] F. Fassò, *Lie series method for vector fields and Hamiltonian perturbation theory*, Z. Angew. Math. Phys. **41** (1990), no. 6, 843–864.

[41] 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.

[42] G. Gentile, V. Mastropietro, and M. Procesi, *Periodic solutions for completely resonant nonlinear wave equations with Dirichlet boundary conditions*, Comm. Math. Phys. **256** (2005), no. 2, 437–490.

[43] B. Grébert, *Birkhoff normal form and Hamiltonian PDES*, Preprint. (2006).

[44] E. Hairer and C. Lubich, *Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations*, (2006).

[45] T. Kappeler and J. Pöschel, *KdV & KAM*, Springer-Verlag, Berlin, 2003.

[46] S. Klainerman, *On "almost global" solutions to quasilinear wave equations in three space dimensions*, Comm. Pure Appl. Math. **36** (1983), 325–344.

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

[48] S. B. Kuksin, *Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum*, Funct. Anal. Appl. **21** (1987), 192–205.

[49] ———, *Nearly integrable infinite-dimensional Hamiltonian systems*, Springer-Verlag, Berlin, 1993.

[50] S. B. Kuksin and J. Pöschel, *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Ann. Math. **143** (1996), 149–179.

[51] B. V. Lidskij and E. I. Shulman, *Periodic solutions of the equation* $u_{tt} - u_{xx} + u^3 = 0$, Funct. Anal. Appl. **22** (1988), 332–333.

[52] J. Marsden, *Darboux's theorem fails for weak symplectic forms*, Proc. Amer. Math. Soc. **32** (1972), 590–592.

[53] N. N. Nekhoroshev, *Exponential estimate of the stability of near integrable Hamiltonian systems*, Russ. Math. Surveys **32** (1977), no. 6, 1–65.

[54] 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.

[55] S. Paleari, D. Bambusi, and S. Cacciatori, *Normal form and exponential stability for some nonlinear string equations*, ZAMP **52** (2001), 1033–1052.

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

[57] J. Pöschel, *On the construction of almost-periodic solutions for a nonlinear Schrödinger equation*, Ergod. Th. & Dynam. Syst. **22** (2002), 1–22.

[58] A. Soffer and M. I. Weinstein, *Resonances, radiation damping and instability in Hamiltonian nonlinear wave equations*, Invent. Math. **136** (1999), no. 1, 9–74.

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

[60] C. E. Wayne, *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Physics **127** (1990), 479–528.

[61] A. Weinstein, *Symplectic structures on Banach manifolds*, Bull. Amer. Math. Soc. **75** (1969), 1040–1041.

[62] X. Yuan, *A KAM theorem with applications to partial differential equations of higher dimensions*, Commun. Math. Phys. **275** (2007), 97–137.

[63] V. E. Zakharov, *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, Appl. Mech. Tech. Physics **2** (1968), 190–194.

[64] E. Zehnder, *C. L. Siegel's linearization theorem in infinite dimensions*, Manuscripta Math. **23** (1978), 363–371.