CHAPTER 15

Hamiltonian PDEs

Sergei B. Kuksin

Department of Mathematics, Heriot-Watt University, Edinburgh EH14 4AS, Scotland, UK
and Steklov Institute of Mathematics, 8 Gubkina St., 117966 Moscow, Russia
E-mail: *kuksin@ma.hw.ac.uk*

With an appendix by Dario Bambusi

Dipartimento di Mathematica, Politecnico di Milano, Via Saldini 50, 20133 Milano, Italy
E-mail: *dario.bambusi@unimi.it*

Contents

HANDBOOK OF DYNAMICAL SYSTEMS, VOL. 1B
Edited by B. Hasselblatt and A. Katok

1. Introduction

In this work we discuss qualitative properties of solutions for Hamiltonian partial differential equations in the finite volume case. That is, when the space-variable x belongs to a finite domain and appropriate boundary conditions are specified on the domain's boundary (or x belongs to the whole space, but the equation contains a potential term, where the potential growths to infinity as $|x| \to \infty$, cf. below Example 5.5 in Section 5.2). Most of these properties have analogies in the classical finite-dimensional Hamiltonian mechanics. In the infinite-volume case properties of the equations become rather different due to the phenomenon of radiation, and we do not touch them here.

Our bibliography is by no means complete.

NOTATION. By $\mathbb{T}^n$ we denote the torus $\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$ and write $\mathbb{T}^1 = S^1$; by $\mathbb{R}^n_+$—the open positive octant in $\mathbb{R}^n$; by $\mathbb{Z}_0$—the set of non-zero integers. By $B_\delta(x; X)$ we denote an open δ-ball in a space X, centred at $x \in X$. Abusing notation, we denote by x both the space-variable and an element of an abstract Banach space X. For an invertible linear operator J we set $\bar{J} = -J^{-1}$. The Lipschitz norm of a map f from a metric space M to a Banach space is defined as $\sup_{m\in M} \|f(m)\| + \sup_{m_1\neq m_2} \frac{\|f(m_1)-f(m_2)\|}{\operatorname{dist}(m_1,m_2)}$.

2. Symplectic Hilbert scales and Hamiltonian equations

2.1. *Hilbert scales and their morphisms*

Let X be a real Hilbert space with a scalar product $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_X$ and a Hilbert basis $\{\varphi_k \mid k \in \widetilde{\mathbb{Z}}\}$, where $\widetilde{\mathbb{Z}}$ is a countable subset of some $\mathbb{Z}^n$. Let us take a positive sequence $\{\theta_k \mid k \in \widetilde{\mathbb{Z}}\}$ which goes to infinity with k. For any s we define X_s as a Hilbert space with the Hilbert basis $\{\varphi_k\theta_k^{-s} \mid k \in \widetilde{\mathbb{Z}}\}$. By $\|\cdot\|_s$ and $\langle\cdot,\cdot\rangle_s$ we denote the norm and the scalar product in X_s (in particular, $X_0 = X$ and $\langle\cdot,\cdot\rangle_0 = \langle\cdot,\cdot\rangle$). The totality $\{X_s\}$ is called a *Hilbert scale*, the basis $\{\varphi_k\}$—the *basis of the scale* and the scalar product $\langle\cdot,\cdot\rangle$—the *basic scalar product of the scale*.

A Hilbert scale may be continuous or discrete, depending on whether $s \in \mathbb{R}$ or $s \in \mathbb{Z}$. The objects we define below and the theorems we discuss are valid in both cases.

A Hilbert scale $\{X_s\}$ possesses the following properties:

(1) X_s is compactly embedded in X_r if $s > r$ and is dense there;

(2) the spaces X_s and X_{-s} are conjugated with respect to the scalar product $\langle\cdot,\cdot\rangle$. That is, for any $u \in X_s \cap X_0$ we have

$$\|u\|_s = \sup\bigl\{\langle u, u'\rangle \mid u' \in X_{-s} \cap X_0,\ \|u'\|_{-s} = 1\bigr\};$$

(3) the norms $\|\cdot\|_s$ satisfy the interpolation inequality; linear operators in the spaces X_s satisfy the interpolation theorem.

Concerning these and other properties of the scales see [77] and [59].

For a scale $\{X_s\}$ we denote by $X_{-\infty}$ and X_∞ the linear spaces $X_{-\infty} = \bigcup X_s$ and $X_\infty = \bigcap X_s$.

Scales of Sobolev functions are the most important for this work:

EXAMPLE 2.1. Basic for us is the Sobolev scale of functions on the d-dimensional torus $\{H^s(\mathbb{T}^d;\mathbb{R}) = H^s(\mathbb{T}^d)\}$. A space $H^s(\mathbb{T}^d)$ is formed by functions $u:\mathbb{T}^d \to \mathbb{R}$ such that

$$u = \sum_{l\in\mathbb{Z}^d} u_l e^{il\cdot x}, \quad \mathbb{C} \ni u_l = \bar{u}_{-l}, \ \|u\|_s^2 = \sum_l \bigl(1+|l|\bigr)^{2s} |u_l|^2 < \infty.$$

The basis $\{\varphi_k\}$ is formed by all distinct properly normalised functions $\operatorname{Re} e^{il\cdot x}$ and $\operatorname{Im} e^{il\cdot x}$, $l \in \mathbb{Z}^d$.

We shall also use the sub-scale $\{H^s(\mathbb{T}^d)_0\}$, where a space $H^s(\mathbb{T}^d)_0$ consists of functions from $H^s(\mathbb{T}^d)$ with zero mean-value.

EXAMPLE 2.2. Consider the scale $\{H_0^s(0,\pi)\}$, where a space $H_0^s = H_0^s(0,\pi)$ is formed by the odd 2π-periodic functions $u = \sum_{k=1}^\infty u_k \sin kx$ such that $\|u\|_s^2 = \sum |k|^{2s}|u_k|^2 < \infty$. Since $\{\sin nx\}$ is a complete system of eigenfunctions of the operator $-\Delta$ in $L_2(0,\pi)$ with the domain of definition $\{u \in H^2(0,\pi) \mid u(0) = u(\pi) = 0\}$, then an equivalent definition of these spaces is that $H_0^s = \mathcal{D}(-\Delta)^{s/2}$ (see [77]). In particular,

$$\begin{aligned} &H_0^1 = \bigl\{u \in H^1(0,\pi) \mid u(0) = u(\pi) = 0\bigr\}, \qquad H_0^2 = H^2(0,\pi) \cap H_0^1, \\ &H_0^3 = \bigl\{u \in H^3(0,\pi) \mid u(0) = u_{xx}(0) = u(\pi) = u_{xx}(\pi) = 0\bigr\}. \end{aligned} \tag{2.1}$$

Given two scales $\{X_s\}$, $\{Y_s\}$ and a linear map $L: X_\infty \to Y_{-\infty}$, we denote by $\|L\|_{s_1,s_2} \leqslant \infty$ its norm as a map $X_{s_1} \to Y_{s_2}$. We say that L defines a (*linear*) *morphism of order* d of the two scales for $s \in [s_0, s_1]$, $s_0 \leqslant s_1$,[1] if $\|L\|_{s,s-d} < \infty$ for every $s \in [s_0, s_1]$. If in addition the inverse map L^{-1} exists and defines a morphism of order $-d$ of the scales $\{Y_s\}$ and $\{X_s\}$ for $s \in [s_0+d, s_1+d]$, we say that L defines an *isomorphism of order* d for $s \in [s_0, s_1]$. If $\{X_s\} = \{Y_s\}$, then an isomorphism is called an *automorphism*.

EXAMPLE 2.3. Multiplication by a non-vanishing C^r-smooth function defines a zero-order automorphism of the Sobolev scale $\{H^s(\mathbb{T}^n)\}$ for $-r \leqslant s \leqslant r$.

If L is a morphism of scales $\{X_s\}$, $\{Y_s\}$ of order d for $s \in [s_0, s_1]$, then adjoint maps L^* form a morphism of the scales $\{Y_s\}$ and $\{X_s\}$ of the same order d for $s \in [-s_1+d, -s_0+d]$. It is called the *adjoint morphism*.

If $L = L^*$ ($L = -L^*$) on the space X_∞, then the morphism L is called symmetric (antisymmetric).

If L is a symmetric morphism of $\{X_s\}$ of order d for $s \in [s_0, d-s_0]$, where $s_0 \geqslant d/2$, then the adjoint morphism L^* is defined for $s \in [s_0, d-s_0]$ and coincide with L on X_∞; hence, $L^* = L$. We call L a *selfadjoint morphism*. Anti-selfadjoint morphisms are defined similarly.

EXAMPLE 2.4. The operator Δ defines a selfadjoint morphism of order 2 of the Sobolev scale $\{H^s(\mathbb{T}^n)\}$ for $-\infty < s < \infty$. The operators $\partial/\partial x_j$, $1 \leqslant j \leqslant n$, define anti-selfadjoint morphisms of order one. The automorphism in Example 1.1 is selfadjoint.

[1] Or $s \in (s_0, s_1)$, etc.

Let $\{Y_s\}$, $\{Y_s\}$ be two scales and $O_s \subset X_s$, $s \in [a,b]$, be a system of (open) domains, compatible in the following sense:

$$O_{s_1} \cap O_{s_2} = O_{s_2} \quad \text{if } a \leqslant s_1 \leqslant s_2 \leqslant b.$$

Let $F : O_a \to Y_{-\infty}$ be a map such that for every $s \in [a,b]$ its restriction to O_s defines an analytic (C^k-smooth) map $F : O_s \to Y_{s-d}$. Then F is called an analytic (C^k-smooth) morphism of order d for $s \in [a,b]$.

EXAMPLE 2.5. Let $\{X_s\}$ be the Sobolev scale $\{H^s(\mathbb{T}^d)\}$ and $f(u,x)$ be a smooth function. Then the map $F : u(x) \mapsto f(u(x),x)$, $X_a \to X_a$, is smooth if $a > \frac{d}{2}$, so on these spaces $\operatorname{ord} F = 0$. If f is analytic, then so is F.

Now let us assume that $d = 1$, f is analytic, $f(0,x) \equiv 0$ and consider F as a map in the scale $\{H_0^s = H_0^s(0,\pi), s \in \mathbb{Z}\}$. For $s \geqslant 1$ the map $F : H_0^s \to H^s(0,\pi)$ is analytic. Since $Fu(0) = Fu(\pi) = 0$, then due to (2.1) for $s = 1$ and $s = 2$ $F(H_0^s) \subset H_0^s$. So on the spaces H_0^1 and H_0^2 we have $\operatorname{ord} F = 0$. Since in general for $u \in H_0^\infty$, $F(u) \in H_0^2$ but $\notin H_0^3$ (see (2.1)), then on the spaces H_0^s, $s \geqslant 3$, we have $\operatorname{ord} F > 0$.

If $f(u,x)$ is odd in u and even in x (e.g., is x-independent), or vice versa, then $F(H_0^s) \subset H_0^s$ for $s \geqslant 1$, so $\operatorname{ord} F = 0$ for any $s \geqslant 1$.

Given a C^k-smooth function $H : X_d \supset O_d \to \mathbb{R}$, $k \geqslant 1$, we consider its *gradient map* with respect to the paring $\langle \cdot, \cdot \rangle$:

$$\nabla H : O_d \to X_{-d}, \quad \big\langle \nabla H(u), v \big\rangle = dH(u)v \ \forall v \in X_d.$$

The map ∇H is C^{k-1}-smooth.

If O_d belongs to a system of compatible domains O_s, $a \leqslant s \leqslant b$, and the gradient map ∇H defines a C^{k-1}-smooth morphism of order d_H for $a \leqslant s \leqslant b$, we write that $\operatorname{ord} \nabla H = d_H$.

2.2. *Symplectic structures*

For simplicity we restrict ourselves to constant-coefficient symplectic structures. For the general case see [59].

Let $\{X_s\}$ be a Hilbert scale and J be its anti-selfadjoint automorphism of order d for $-\infty < s < \infty$. Then the operator $\bar{J} = -J^{-1}$ defines an anti-selfadjoint automorphism of order $-d$. We define a two-form α_2 as

$$\alpha_2 = \bar{J}\, dx \wedge dx,$$

where by definition $\bar{J}\, dx \wedge dx\, [\xi, \eta] = \langle \bar{J}\xi, \eta \rangle$. Clearly, $\bar{J}\, dx \wedge dx$ defines a continuous skew-symmetric bilinear form on $X_r \times X_r$ if $r \geqslant -d/2$. Therefore any space X_r, $r \geqslant -d/2$, becomes a *symplectic* (*Hilbert*) *space* and we shall write it as a pair (X_r, α_2).

The pair $(\{X_s\}, \alpha_2)$ is called a *symplectic* (*Hilbert*) *scale*.

EXAMPLE 2.6. Let us take the index-set $\mathcal{Z}$ to be the union of non-intersecting subsets $\mathcal{Z}_+$ and $\mathcal{Z}_-$, provided with an involution $\mathcal{Z} \to \mathcal{Z}$ which will be denoted $j \mapsto -j$, such that $-\mathcal{Z}_\pm = \mathcal{Z}_\mp$. Let us consider a Hilbert scale $\{X_s\}$ with a basis $\{\phi_k, k \in \mathcal{Z}\}$ and a sequence $\{\theta_k, k \in \mathcal{Z}\}$, such that $\theta_{-j} \equiv \theta_j$. Take J to be the linear operator, defined by the relations

$$J\phi_k = \phi_{-k} \quad \forall k \in \mathcal{Z}_+, \qquad J\phi_k = -\phi_{-k} \quad \forall k \in \mathcal{Z}_-.$$

It defines an anti-selfadjoint automorphism of the scale of zero order, and $\bar{J} = J$. The symplectic scale $(\{X_s\}, \alpha_2 = \bar{J}\, dx \wedge dx = J\, dx \wedge dx)$ will be called a *Darboux scale*.

Let $(\{X_s\}, \alpha_2 = \bar{J}\, dx \wedge dx)$ and $(\{Y_s\}, \beta_2 = \overline{\Upsilon}\, dy \wedge dy)$ be two symplectic Hilbert scales and $O_s \subset X_s$, $a \leqslant s \leqslant b$, be a system of compatible domains. A C^1-smooth morphism of order d_1

$$F : O_s \to Y_{s-d_1}, \quad a \leqslant s \leqslant b,$$

is *symplectic* if $F^*\beta_2 = \alpha_2$. That is, if $\langle \overline{\Upsilon} F_*(x)\xi, F_*(x)\eta\rangle_Y \equiv \langle \bar{J}\xi, \eta\rangle_X$, or

$$F^*(x)\overline{\Upsilon} F_*(x) = \bar{J} \quad \forall x.$$

A symplectic morphism F as above is called a *symplectomorphism* if it is a diffeomorphism.

2.3. *Hamiltonian equations*

To a C^1-smooth function h on a domain $O_d \subset X_d$, the symplectic form α_2 as above corresponds the *Hamiltonian vector field* V_h, defined by the usual relation (cf. [2,43]):

$$\alpha_2\big[V_h(x), \xi\big] = -dh(x)\xi \quad \forall \xi.$$

That is, $\langle \bar{J} V_h(x), \xi\rangle \equiv -\langle \nabla h(x), \xi\rangle$ and

$$V_h(x) = J\nabla h(x).$$

The vector field V_h defines a continuous map $O_d \to X_{-d-d_J}$. Usually we shall assume that V_h is smoother than that and defines a smooth morphism of order $d_1 \leqslant 2d + d_J$ for all s from some segment.

For any C^1-smooth function h on $O_d \times \mathbb{R}$ we denote by V_h the non-autonomous vector field $V_h(x,t) = J\nabla_x h(x,t)$, where ∇_x is the gradient in x, and consider the corresponding *Hamiltonian equation* (or *Hamiltonian system*)

$$\dot{x} = J\nabla_x h(x,t) = V_h(x,t). \tag{2.2}$$

A partial differential equation, supplemented by some boundary conditions, is called a *Hamiltonian partial differential equation*, or an *HPDE*, if under a suitable choice of a symplectic Hilbert scale $(\{X_s\}, \alpha_2)$, a domain $O_d \subset X_d$ and a Hamiltonian h, it can be written in the form (2.2). In this case the vector field V_h is unbounded, $\operatorname{ord} V_h = d_1 > 0$. That is,

$$V_h : O_d \times \mathbb{R} \to X_{d-d_1}.$$

Usually O_d belongs to a system of compatible domains O_s, $s \geqslant d_0$, and V_h (as a function of x) defines an analytic morphism of order d_1 for $s \geqslant d_0$.

A continuous curve $x : [t_0, t_1] \to O_d$ is called a *solution of* (2.2) *in the space* X_d if it defines a C^1-smooth map $x : [t_0, t_1] \to X_{d-d_1}$ and both parts of (2.2) coincide as curves in X_{d-d_1}. A solution x is called *smooth* if it defines a smooth curve in each space X_s.

If a solution $x(t)$, $t \geqslant t_0$, of (2.2) such that $x(t_0) = x_0$ exists and is unique, we write $x(t_1) = S_{t_0}^{t_1} x_0$, or $x(t_1) = S^{t_1 - t_0} x_0$ if the equation is autonomous (we do not assume that $t_1 \geqslant t_0$). The operators $S_{t_0}^{t_1}$ and S^t are called *flow-maps* of the equation. Clearly, $S_{t_0}^{t_1}$ equals $(S_{t_1}^{t_0})^{-1}$ on a joint domain of definition of the two operators.

A non-linear PDE is called *strongly non-linear* if its non-linear part contains as many derivatives as the linear part. Strongly non-linear Hamiltonian PDEs may possess rather unpleasant properties. In particular, for some of them, every non-zero solution develops a singularity in finite time, see an example in Section 1.4 of [59].

We shall call a non-linear PDE *quasilinear* if its non-linear part contains less derivatives then the linear one. A quasilinear equation can be written in the form (2.2) with

$$h(x,t) = \frac{1}{2}\langle Ax, x\rangle + h_0(x,t), \tag{2.3}$$

where A is a linear operator which defines a selfadjoint morphism of the scale (so $\nabla h(x,t) = Ax + \nabla h_0(x,t)$) and $\operatorname{ord} \nabla h_0 < \operatorname{ord} A$.

The class of Hamiltonian PDEs contains many important equations of mathematical physics, some of them are discussed below. The first difficulty one comes across when studies this class is absence of a general theorem which would guarantee that (locally in time) an equation has a unique solution.[2] Such a theorem exists for semilinear equations, where Equation (2.2) will be called *semilinear* if its Hamiltonian has the form (2.3) and $\operatorname{ord} J\nabla h_0 \leqslant 0$ (see [69] and Section 1.4 of [59]).

EXAMPLE 2.7 (*Equations of the Korteweg–de Vries type*). Let us take for $\{X_s\}$ the scale of zero mean-value Sobolev spaces $H^s(S^1)_0$ as in Example 2.1 and choose $J = \partial/\partial x$, so $d_J = 1$. For a Hamiltonian h we take $h(u) = \int_0^{2\pi} (-\frac{1}{8}u'(x)^2 + f(u))\,dx$ with some analytic function $f(u)$. Then $\nabla h(u) = \frac{1}{4}u'' + f'(u)$ and the equation takes the form

$$\dot{u}(t,x) = \frac{1}{4}u''' + \frac{\partial}{\partial x} f'(u).$$

For $f(u) = \frac{1}{4}u^3$ we get the classical Korteweg–de Vries (KdV) equation. The map V_h defines an analytic morphism of order 3 of the scale $\{X_s\}$, for $s > 1/2$. The equation

[2] Still, see [47] for a theory which applies to some classes of quasilinear HPDEs.

has the form (2.2), (2.3), where $\operatorname{ord} JA = 3$ and $\operatorname{ord} J\nabla h_0 = 1$. It is quasilinear, but not semilinear.

EXAMPLE 2.8 (*NLS—non-linear Schrödinger equation*). Let $X_s = H^s(\mathbb{T}^n; \mathbb{C})$, where this Sobolev space is treated as a real Hilbert space, and the basic scalar product of the scale is $\langle u, v\rangle = \operatorname{Re}\int u\bar{v}\,dx$. For J we take the operator $Ju(x) = iu(x)$, so $\operatorname{ord} J = 0$ and $(\{X_s\}, \bar{J}\,du \wedge du)$ is a Darboux scale. We choose

$$h(u) = \frac{1}{2}\int_{\mathbb{T}^n} \big(|\nabla u|^2 + V(x)|u|^2 + g(x, u, \bar{u})\big)\,dx,$$

where V is a smooth real function and $g(x, u, v)$ is a smooth function, real if $v = \bar{u}$. Then $\nabla h(u) = -\Delta u + V(x)u + \frac{\partial}{\partial \bar{u}} g$ and (2.2) takes the form

$$\dot{u} = i\left(-\Delta u + V(x)u + \frac{\partial}{\partial \bar{u}}\, g(x, u, \bar{u})\right), \quad u = u(t, x),\ x \in \mathbb{T}^n. \tag{2.4}$$

This is a semilinear Hamiltonian equation in any space X_{d_0}, $d_0 > n/2$, with $\operatorname{ord} A = 2$ and $\operatorname{ord} \nabla h_0 = 0$.

Non-linear Schrödinger equation (2.4) with $n = 1$, $V(x) = \text{const}$ and $g = \gamma|u|^4$, $\gamma \neq 0$, is called the *Zakharov–Shabat equation*. The equation with $\gamma > 0$ is called *defocusing* and with $\gamma < 0$—*focusing*.

EXAMPLE 2.9 (*1D NLS with Dirichlet boundary conditions*). Let us choose for X_s the space $H_0^s(0, \pi; \mathbb{C})$ (see Example 2.2), $Ju(x) = iu(x)$ and

$$h(u) = \frac{1}{2}\int_0^\pi \big(|u_x|^2 + V(x)|u|^2 + g\big(x, |u|^2\big)\big)\,dx,$$

where g is smooth and 2π-periodic in x. Now $\nabla h(u) = -u_{xx} + V(x)u + f(x, |u|^2)u$, where $f = \frac{\partial g}{\partial |u|^2}$, and (2.2) becomes

$$\dot{u} = i\big(-u_{xx} + V(x)u + f\big(x, |u|^2\big)u\big), \qquad u(0) = u(\pi) = 0. \tag{2.5}$$

For $s = 1$ and 2 the non-linear term defines a smooth map $X_s \to X_s$ (see Example 2.5), so in these spaces this is a semilinear equation with $\operatorname{ord} A = 2$ and $\operatorname{ord} \nabla h_0 = 0$. If in addition f is even in x, then the non-linear term defines a smooth map for every $s \geqslant 1$. This map is analytic if f is.

EXAMPLE 2.10 (*Non-linear wave equations*). Now let $X_s = H^s(\mathbb{T}^n) \times H^s(\mathbb{T}^n)$ and $\alpha_2 = \bar{J}\,d\eta \wedge d\eta$, where $\eta = (u, v)$ and $J(u, v) = \bar{J}(u, v) = (-v, u)$. Let

$$h(u, v) = \int_{\mathbb{T}^n} \left(\frac{1}{2}v^2 + \frac{1}{2}|\nabla u|^2 - f(x, u)\right)dx. \tag{2.6}$$

The corresponding Hamiltonian equation is

$$\dot{u} = -v, \qquad \dot{v} = -\Delta u - {f'}_u(x, u). \tag{2.7}$$

Or

$$\ddot{u} = \Delta u + {f'}_u(x, u), \quad u = u(t, x), \ x \in \mathbb{T}^n. \tag{2.8}$$

This is a *non-linear wave equation* (with the periodic boundary conditions). We have seen that this equation can be re-written as the system (2.7) which is an HPDE. This Hamiltonian form of the equation is inconvenient since the quadratic part of the Hamiltonian (2.6) corresponds to the linear operator $(u, v) \to \frac{1}{2}(-\Delta u, v)$ which does not define an isomorphism of the scale $\{X_s\}$ (of some order m). It turns out that the non-linear wave equation (2.8) admits another Hamiltonian representation (2.2), where the Hamiltonian h has the form (2.3), the operator A defines an isomorphism of the scale and $\operatorname{ord} A < \operatorname{ord} \nabla h_0$ (so the equation is quasilinear). We note that the corresponding linear operator JA *is not* differential. See [52] and [59], also see below Section 4.3, where the non-linear wave equation $\ddot{u} = u_{xx} - \sin u$ (the Sine-Gordon equation) is considered in details.

3. Basic theorems on Hamiltonian systems

Basic theorems from the classical Hamiltonian formalism (see [2,43]) remain true for Hamiltonian equations (2.2) in Hilbert scales, provided that the theorems are properly formulated. In this section we present three corresponding results. Their proofs can be found in [52,59].

Let $(\{X_s\}, \alpha_2 = \bar{J}\, dx \wedge dx)$ and $(\{Y_s\}, \beta_2 = \overline{\Upsilon}\, dy \wedge dy)$ be two symplectic scales and (for simplicity) $\operatorname{ord} J = \operatorname{ord} \Upsilon = d_J \geqslant 0$. Let $\Phi : Q \to O$ be a C^1-smooth symplectic map, where Q and O are domains in Y_d and X_d, $d \geqslant 0$. If $d_J > 0$, we have to assume that

(H1) for any $|s| \leqslant d$ linearised maps $\Phi_*(y)$, $y \in Q$, define linear maps $Y_s \to X_s$ which continuously depend on y.

The first theorem states that symplectic maps transform Hamiltonian equations to Hamiltonian:

THEOREM 3.1. *Let $\Phi : Q \to O$ be a symplectic map as above* (*so* (H1) *holds if $d_J > 0$*). *Let us assume that the vector field V_h of Equation* (2.2) *defines a C^1-smooth map $V_h : O \times \mathbb{R} \to X_{d-d_1}$ of order $d_1 \leqslant 2d$ and that this vector field is tangent to the map Φ* (*i.e., for every $y \in Q$ and every t the vector $V_h(\Phi(y), t)$ belong to the range of the linearised map $\Phi_*(y)$*). *Then Φ transforms solutions of the Hamiltonian equation $\dot{y} = \Upsilon \nabla_y H(y, t)$, where $H = h \circ \Phi$, to solutions of* (2.2).

COROLLARY 3.2. *If under the assumptions of Theorem* 3.1 $\{X_s\} = \{Y_s\}$ *and $h \circ \Phi = h$, $\Phi^* \alpha_2 = \alpha_2$, then Φ preserves the class of solutions for* (2.2).

For Hamiltonian PDEs (and for Hamiltonian equations (2.2)) Theorem 2.1 plays the same role as its classical finite-dimensional counterpart plays for usual Hamiltonian equations: it is used to transform an equation to a normal form, usually in the vicinity of an invariant set (e.g., of an equilibrium).

To apply Theorem 3.1 one needs regular ways to construct symplectic transformations. For classical finite-dimensional systems symplectic transformations usually are obtained either via generating functions, or as Lie transformations (i.e., as flow-maps of additional Hamiltonians), see [2,43,40]. For infinite-dimensional symplectic spaces generating functions play negligible role, while the Lie transformations remain an important tool. An easy but important corresponding result is stated in the theorem below.

Let $(\{X_s\}, \alpha_2)$ be a symplectic Hilbert scale as above and O be a domain in X_d.

THEOREM 3.3. *Let f be a C^1-smooth function on $O \times \mathbb{R}$ such that the map $V_f : O \times \mathbb{R} \to X_d$ is Lipschitz in (x, t) and C^1-smooth in x. Let O_1 be a subdomain of O. Then the flow-maps $S_t^\tau : (O_1, \alpha_2) \to (O, \alpha_2)$ are symplectomorphisms (provided that they map O_1 to O). If the map V_f is C^k-smooth or analytic, then the flow-maps are C^k-smooth or analytic as well.*

The assumption that the map V_f is Lipschitz can be replaced by the much weaker assumption that for a solution $x(t)$ of the equation $\dot{x} = V_f(x)$, the linearised equation $\dot{\xi} = V_{f*}(x(t))\xi$ is such that its flow maps are bounded linear transformations of the space X_d. See [59].

Usually Theorem 3.3 is applied in the situation when $|f| \ll 1$, or $|t - \tau| \ll 1$. In these cases the flow-maps are close to the identity and the corresponding transformations of the space of C^1-smooth functions on O, $H \mapsto H \circ S_t^\tau$, can be written as Lie series (cf. [40]). In particular, the following simple result holds:

THEOREM 3.4. *Under the assumptions of Theorem 3.3, let H be a C^1-smooth function on O. Then*

$$\frac{d}{d\tau} H\big(S_t^\tau(x)\big) = \{f, H\}\big(S_t^\tau(x)\big), \quad x \in O_1. \tag{3.1}$$

In this theorem $\{f, H\}$ denotes the *Poisson bracket* of the two functions:

$$\{f, H\}(x) = \big\langle J\nabla f(x), \nabla H(x)\big\rangle.$$

It is well defined since $J\nabla f = V_f \in X_d$ by assumptions.

Theorem 3.3 and formula (3.1) make from symplectic flow-maps S_t^τ a tool which is well suited to prove KAM-theorems for Hamiltonian PDEs, see the scheme of the proof of Theorem 5.1 in Section 5.1 below.

An immediate consequence of Theorem 3.4 is that for an autonomous Hamiltonian equation $\dot{x} = J\nabla f(x)$ such that $\operatorname{ord} J\nabla f = 0$, a C^1-smooth function H is an integral of motion[3] if and only if $\{f, H\} \equiv 0$.

[3]That is, $H(x(t))$ is a time-independent quantity for any solution $x(t)$.

If $d' = \text{ord}\, J\nabla f > 0$ and $O = O_d$ belongs to a system of compatible domains $O_s \subset X_s$, $s \in [d_0, d]$, where $d_0 = d - d'$, then H such that $\{f, H\} \equiv 0$ is an integrable of motion for the equation $\dot{x} = J\nabla f(x)$, provided that

$$\text{ord}\, J\nabla f = d' \quad \text{and} \quad \text{ord}\, \nabla H = d_H \quad \text{for } s \in [d_0, d],$$

where $d' + d_H \leqslant 2d$. Indeed, since $d_0 - d_H \geqslant -d_0$, then H is a C^1-smooth function on O_{d_0}. Since any solution $x(t)$ is a C^1-smooth curve in O_{d_0} by the definition of a solution, then

$$\frac{d}{dt} H(x) = \langle \nabla H(x), \dot{x} \rangle = \langle \nabla H(x), J\nabla f(x) \rangle = \{f, H\}(x) = 0.$$

In particular, f is an integral of motion for the equation $\dot{x} = J\nabla f(x)$ in O_d if we have $\text{ord}\, J = d_J$ and $\text{ord}\, \nabla f = d_f$ for $s = d$ and for $s \in [d, d - d_f - d_J]$, where $d \geqslant d_f + d_J/2$. That is, if the equation is being considered in sufficiently smooth spaces.

EXAMPLE 3.5. Let us consider a non-linear Schrödinger equation (2.5) such that $g(u, \bar{u}) = g_0(|u|^2)$, and take $H(u) = \|u\|_0^2 = |u|_{L_2}^2$. Now $d' := \text{ord}\, J\nabla f = 2$ for $s \in (n/2, \infty)$, and $\text{ord}\, \nabla H = 0$. Elementary calculations show that $\{f, H\} \equiv 0$. So L_2-norm is an integral of motion for solutions of (2.5) in X_s if $s > n/2 + 2$. (In fact this result remains true for solutions of much lower smoothness, see [15].)

4. Lax-integrable equations

4.1. *General discussion*

Let us take a Hamiltonian PDE and write it as a Hamiltonian equation in a suitable symplectic Hilbert scale $(\{X_s\}, \alpha_2 = \bar{J}\, du \wedge du)$:

$$\dot{u} = J\nabla H(u). \tag{4.1}$$

This equation is called Lax-integrable if there exists an additional Hilbert scale $\{Z_s\}$ (real or complex), and finite order linear morphisms $\mathcal{L}_u$ and $\mathcal{A}_u$ of this scale which depend on the parameter $u \in X_\infty$, such that a curve $u(t)$ is a smooth solution for (4.1) if and only if

$$\frac{d}{dt} \mathcal{L}_{u(t)} = [\mathcal{A}_{u(t)}, \mathcal{L}_{u(t)}]. \tag{4.2}$$

The operators $\mathcal{A}_u$ and $\mathcal{L}_u$, treated as morphisms of the scale $\{Z_s\}$, are assumed to depend smoothly on $u \in X_d$ where d is sufficiently large, so the left-hand side of (4.2) is well defined (for details see [59]). The pair of operators $\mathcal{L}$, $\mathcal{A}$ is called the *Lax pair*.[4]

[4] Due to a deep result by Krichever and Phong [48], any Lax-integrable PDE is a Hamiltonian system. The corresponding symplectic structure belongs to a bigger class than that described in Section 2.2, so to apply our techniques we need to assume a priori that the equation has the form (4.1).

In most known examples of Lax-integrable equations the relation between the scales $\{X_s\}$ and $\{Z_s\}$ is the following: spaces X_s are formed by T-periodic Sobolev vector-functions, while $\mathcal{A}$ and $\mathcal{L}$ are differential or integro-differential operators with u-dependent coefficients, acting in a scale $\{Z_s\}$ of TL-periodic Sobolev vector-functions. Here L is some fixed integer.

Let $u(t)$ be a smooth solution for (4.1). We set $\mathcal{L}_t = \mathcal{L}_{u(t)}$ and $\mathcal{A}_t = \mathcal{A}_{u(t)}$.

LEMMA 4.1. *Let $\chi_0 \in Z_\infty$ be a smooth eigenvector of $\mathcal{L}_0$, i.e., $\mathcal{L}_0\chi_0 = \lambda\chi_0$. Let us assume that the initial-value problem*

$$\dot{\chi} = \mathcal{A}_t\chi, \qquad \chi(0) = \chi_0, \tag{4.3}$$

has a unique smooth solution $\chi(t)$. Then

$$\mathcal{L}_t\chi(t) = \lambda\chi(t) \quad \forall t. \tag{4.4}$$

PROOF. Let us denote the left-hand side of (4.4) by $\xi(t)$, the right-hand side—by $\eta(t)$ and calculate their derivatives. We have:

$$\frac{d}{dt}\xi = \frac{d}{dt}\mathcal{L}\chi = [\mathcal{A}, \mathcal{L}]\chi + \mathcal{L}\mathcal{A}\chi = \mathcal{A}\mathcal{L}\chi = \mathcal{A}\xi$$

and

$$\frac{d}{dt}\eta = \frac{d}{dt}\lambda\chi = \lambda\mathcal{A}\chi = A\eta.$$

Thus, both $\xi(t)$ and $\eta(t)$ solve the problem (4.3) with χ_0 replaced by $\lambda\chi_0$ and coincide by the uniqueness assumption. □

Due to this lemma the discrete spectrum of the operator $\mathcal{L}_u$ is an integral of motion for Equation (4.1). In particular, a set $\mathcal{T}$ formed by all smooth vectors $u \in X_\infty$ such that the eigenvalues of the operator $\mathcal{L}_u$ belong to a fixed subset of $\mathbb{C} \times \mathbb{C} \times \cdots$, is invariant for the flow of Equation (4.1). A remarkable discovery, made by Novikov [68] and Lax [61], is that for integrable Hamiltonian PDEs, considered on finite space-intervals with suitable boundary conditions, some sets $\mathcal{T}$ as above are finite-dimensional symplectic submanifolds $\mathcal{T}^{2n}$ of all symplectic spaces (X_s, α_2), and restriction of Equation (4.1) to any $\mathcal{T}^{2n}$ is an integrable Hamiltonian system. Moreover, for some integrable equations it is known that the union of all these manifolds $\mathcal{T}^{2n}$ is dense in every space X_s. Solutions that fill a manifold $\mathcal{T}^{2n}$ are called *finite-gap solutions*, and the manifold itself—a *finite-gap manifold*. See, e.g., [32,83,8,59].

4.2. *Korteweg–de Vries equation*

The KdV equation

$$\dot{u} = \frac{1}{4}\frac{\partial}{\partial x}\big(u_{xx} + 3u^2\big), \quad u(t,x) \equiv u(t, x+2\pi), \ \int_0^{2\pi} u\,dx \equiv 0, \tag{4.5}$$

takes the form (4.1) in the symplectic Hilbert scale $(\{X_s\}, \alpha_2 = \bar{J}\,du \wedge du)$, where X_s is the Sobolev space $H^s(S^1)_0$ and $Ju = (\partial/\partial x)u$, see Example 2.7. Due to Lax himself, this equation is Lax-integrable and the corresponding Lax pair is

$$\mathcal{L}_u = -\frac{\partial^2}{\partial x^2} - u, \qquad \mathcal{A}_u = \frac{\partial^3}{\partial x^3} + \frac{3}{2}u\frac{\partial}{\partial x} + \frac{3}{4}u_x.$$

Taking for $\{Z_s\}$ the Sobolev scale of 4π-periodic functions and applying Lemma 4.1 we obtain that smooth 4π-periodic spectrum of the operator $\mathcal{L}_u$ is an integral of motion. It is well known that the spectrum of $\mathcal{L}_u$ is formed by eigenvalues

$$\lambda_0 < \lambda_1 \leqslant \lambda_2 < \lambda_3 \leqslant \lambda_4 < \cdots \nearrow \infty,$$

and that the corresponding eigenfunctions are smooth, provided that the potential u is. Let us take any integer n-vector $\mathbf{V}$,

$$\mathbf{V} = (V_1, \ldots, V_n) \in \mathbb{N}^n, \quad V_1 < \cdots < V_n.$$

Denoting $\Delta_j = \lambda_{2j} - \lambda_{2j-1} \geqslant 0$, $j = 1, 2, \ldots$, we define the set $\mathcal{T}_{\mathbf{V}}^{2n}$ as

$$\mathcal{T}_{\mathbf{V}}^{2n} = \big\{u(x) \mid \Delta_j \neq 0 \text{ iff } j \in \{V_1, \ldots, V_n\}\big\}.$$

Clearly $\mathcal{T}_{\mathbf{V}}^{2n}$ equals to the union $\mathcal{T}_{\mathbf{V}}^{2n} = \bigcup_{r \in \mathbb{R}_+^n} T_{\mathbf{V}}^n(r)$, where $\mathbb{R}_+^n = \{r \mid r_j > 0\ \forall j\}$ and

$$T_{\mathbf{V}}^n(r) = \big\{u(x) \in \mathcal{T}_{\mathbf{V}}^{2n} \mid \Delta_j = r_j\ \forall j\big\}.$$

Since the 4π-periodic spectrum $\{\lambda_j\}$ is an integral of motion for (KdV), then the sets $T_{\mathbf{V}}^n(r)$ are invariant for the KdV-flow. Due to the classical theory of the Sturm–Liouville operator $\mathcal{L}_u$, the set $\mathcal{T}_{\mathbf{V}}^{2n}$ is a smooth submanifold of any space X_s, foliated to the smooth n-tori $T_{\mathbf{V}}^n(r)$. For all these results see, e.g., [46].

Due to Novikov and Lax, there exist an analytic map $\Phi = \Phi_{\mathbf{V}} : \{(r,\xi)\} = \mathbb{R}_+^n \times \mathbb{T}^n \to X_s$ (s is any integer), and an analytic function $h = h^n(r)$ such that $T_{\mathbf{V}}^n(r) = \Phi(\{r\} \times \mathbb{T}^n)$, and for any $\xi_0 \in \mathbb{T}^n$ the curve $u(t) = \Phi(r, \xi_0 + t\nabla h(r))$ is a smooth solution for (4.5). As a function of t, this solution is quasiperiodic.[5] The celebrated Its–Matveev formula explicitly represents Φ in terms of theta-functions, see in [32,31,8,59].

[5] A continuous curve $u : \mathbb{R} \to X$ is quasiperiodic if there exist $n \in \mathbb{N}$, $\phi \in \mathbb{T}^n$, $\omega \in \mathbb{R}^n$ and a continuous map $U : \mathbb{T}^n \to X$ such that $u(t) = U(\phi + t\omega)$.

4.3. *Other examples*

Sine-Gordon. The Sine-Gordon (SG) equation on the circle

$$\ddot{u} = u_{xx}(t,x) - \sin u(t,x), \quad x \in S^1 = \mathbb{R}/2\pi\mathbb{Z},$$

is another example of a Lax-integrable HPDE.

First the equation has to be written in a Hamiltonian form. The most straightforward way to do this is to write (SG) as the system

$$\dot{u} = -v, \qquad \dot{v} = -u_{xx} + \sin u(t,x).$$

One immediately sees that this system is a semilinear Hamiltonian equation in the symplectic scale $(\{X_s = H^s(S) \times H^s(S)\}, \alpha_2 = \bar{J}\, d\eta \wedge d\eta)$, where $\eta = (u,v)$ and $J(u,v) = (-v,u)$.

Now we derive another Hamiltonian form of (SG), more convenient for its analysis. To do this we consider the shifted Sobolev scale $\{X_s = H^{s+1}(S^1) \times H^{s+1}(S^1)\}$, where the space X_0 is given the scalar product

$$\langle \xi_1, \xi_2 \rangle = \int_{S^1} \big(\xi'_{1x} \cdot \xi'_{2x} + \xi_1 \cdot \xi_2\big)\, dx,$$

and any space X_s—the product $\langle \xi_1, \xi_2 \rangle_s = \langle A^s \xi_1, \xi_2 \rangle$. Here A is the operator $A = -\partial^2/\partial x^2 + 1$. Obviously, A defines a selfadjoint automorphism of the scale of order one. The operator $J(u,w) = (-\sqrt{A}\, w, \sqrt{A}\, u)$ defines an anti-selfadjoint automorphism of the same order. We provide the scale with the symplectic form $\beta_2 = \bar{J}\, d\xi \wedge d\xi$. We note that (SG) can be written as the system

$$\dot{u} = -\sqrt{A}\, w, \qquad \dot{w} = \sqrt{A}\big(u + A^{-1} f'\big(u(x)\big)\big), \tag{4.6}$$

where $f(u) = -\cos u - \frac{1}{2}u^2$, and that (4.6) is a semilinear Hamiltonian equation in the symplectic scale as above with the Hamiltonian $H(\xi) = \frac{1}{2}\langle \xi, \xi \rangle + \int f(u(x))\, dx$, $\xi = (u,w)$.

Let us denote by X_s^o (X_s^e) subspaces of X_s formed by odd (even) vector functions $\xi(x)$. Then $(\{X_s^o\}, \beta_2)$ and $(\{X_s^e\}, \beta_2)$ are symplectic sub-scales of the scale above. The space X_s^o and X_s^e (with $s \geqslant 0$) are invariant for the flow of Equation (4.6). The restricted flows correspond to the SG equation under the odd periodic and even periodic boundary conditions, respectively.

The SG equation is Lax-integrable under periodic, odd periodic and even periodic boundary conditions. That is, Equation (4.6) is Lax-integrable in the all three symplectic scales defined above. See [8,59].

Zakharov–Shabat equation. Let us take the symplectic Hilbert scale ($X_s = H^s(S^1, \mathbb{C})$, $\bar{J}\, du \wedge du$) as in the Example 2.8. The defocusing and focusing Zakharov–Shabat equations

$$\dot{u} = i\big(-u_{xx} + mu \pm \gamma |u|^2 u\big), \quad \gamma > 0, \tag{4.7}$$

both are Lax-integrable, see [83,8].

5. KAM for PDEs

In this section we discuss the 'KAM for PDEs' theory. Here we cover all relevant topics, except the theory of time-periodic solutions of Hamiltonian PDEs. The latter is reviewed in the Appendix, written by Dario Bambusi. We avoid completely the classical finite-dimensional KAM-theory which deals with time-quasiperiodic solutions of finite-dimensional Hamiltonian systems and instead refer the reader to the recent survey [78].

5.1. *An abstract KAM-theorem*

Let $(\{X_s\}, \alpha_2 = \bar{J}\, du \wedge du)$ be a symplectic Hilbert scale, $-d_J = \operatorname{ord} \bar{J} \leqslant 0$; A be an operator which defines a selfadjoint automorphism of the scale of order $d_A \geqslant -d_J$ and H be a Fréchet–analytic functional on X_{d_0}, $d_0 \geqslant 0$, such that $\operatorname{ord} \nabla H = d_H < d_A$:

$$\nabla H : X_{d_0} \to X_{d_0 - d_H}.$$

We assume that $d_A \leqslant 2d_0$, so the quadratic form $\frac{1}{2}\langle Au, u\rangle$ is well defined on the space X_{d_0}.

In this section we consider the quasilinear Hamiltonian equation with the Hamiltonian $H_\varepsilon(u) = \frac{1}{2}\langle Au, u\rangle + \varepsilon H(u)$:

$$\dot{u}(t) = J\big(Au(t) + \varepsilon \nabla H u(t)\big). \tag{5.1}$$

We assume that the scale $\{X_s\}$ admits a basis $\{\varphi_k, k \in \mathbb{Z}_0 = \mathbb{Z} \setminus \{0\}\}$ such that

$$A\varphi_j^\pm = \lambda_j^A \varphi_j^\pm, \qquad J\varphi_j^\pm = \mp \lambda_j^J \varphi_j^\mp \quad \forall j \geqslant 1, \tag{5.2}$$

with some real numbers λ_j^J, λ_j^A. In particular, the spectrum of the operator JA is $\{\pm i\lambda_j \mid \lambda_j = \lambda_j^J \lambda_j^A\}$. The numbers λ_j are called the *frequencies* of the linear system

$$\dot{u} = JAu. \tag{5.3}$$

Let us fix any $n \geqslant 1$. Then the $2n$-dimensional linear space

$$\operatorname{span}\big\{\varphi_j^\pm \mid 1 \leqslant j \leqslant n\big\} \tag{5.4}$$

is invariant for Equation (5.3) and is foliated to the invariant tori

$$T^n = T^n(I) = \left\{ \sum_{j=1}^{n} u_j^{\pm} \varphi_j^{\pm} \mid {u_j}^{+2} + {u_j}^{-2} = 2I_j \ \forall j \right\}. \tag{5.5}$$

If $I \in \mathbb{R}_+^n$, then $T^n(I)$ is an n-torus. Providing it with the coordinates $q = (q_1, \dots, q_n)$, where $q_j = \operatorname{Arg}(u_j^+ + iu_j^-)$, we see that Equation (5.3) defines on $T^n(I)$ the motion

$$\dot{q} = (\lambda_1, \dots, \lambda_n) =: \omega. \tag{5.6}$$

So all solutions for the linear equation in $T^n(I)$ are quasiperiodic curves with the frequency-vector ω. Our goal in this section is to present and discuss a KAM-theorem which implies that under certain conditions 'most of' trajectories of Equation (5.6) on the torus $T^n(I)$ persist as time-quasiperiodic solutions of the perturbed equation (5.1), if $\varepsilon > 0$ is sufficiently small.

To state the result we assume that the operator A and the function H analytically depend on an additional n-dimensional parameter $a \in \mathcal{A}$, where $\mathcal{A}$ is a connected bounded open domain in $\mathbb{R}^n$. Then $\lambda_j = \lambda_j(a)$. We assume that the first n frequencies $\lambda_l = \omega_l$ depend on a in the non-degenerate way:

(H1) $\det\{\partial\omega_l/\partial a_k \mid 1 \leqslant k,\, l \leqslant n\} \not\equiv 0$;

and that the following spectral asymptotic holds:

(H2) $|\lambda_j(a) - K_1 j^{d_1} - K_1^1 j^{d_1^1} - K_1^2 j^{d_1^2} - \cdots| \leqslant K j^{\tilde{d}},\ \operatorname{Lip} \lambda_j \leqslant j^{\tilde{d}}$,

where $d_1 := d_A + d_J \geqslant 1$, $K_1 > 0$, $\tilde{d} < d_1 - 1$ and the dots stand for a finite sum with exponents $d_1 > d_1^1 > d_1^2 > \cdots$.

Let us denote by X_s^c the complexification of a space X_s and assume that Equation (5.1) is quasilinear and analytic:

(H3) the set $X_{d_0} \times \mathcal{A}$ admits in $X_{d_0}^c \times \mathbb{C}^n$ a complex neighbourhood Q such that the map $\nabla_x H : Q \to X_{d_0 - d_H}^c$ is complex-analytic and bounded uniformly on bounded subsets of Q. Moreover, $d_H + d_J \leqslant \tilde{d}$.

Finally, we shall need the following non-resonance condition:

(H4) For all integer n-vectors s and $(M_2 - n)$-vectors l such that $|s| \leqslant M_1$, $1 \leqslant |l| \leqslant 2$ we have,

$$s \cdot \omega(a) + l_{n+1}\lambda_{n+1}(a) + \cdots + l_{M_2}\lambda_{M_2}(a) \not\equiv 0, \tag{5.7}$$

where the integers $M_1 > 0$ and $M_2 > n$ are to be specified.

Relations (5.7) with $|l| = 1$ and $|l| = 2$ are called, respectively, the first and the second *Melnikov condition*.

Let us fix any $I_0 \in \mathbb{R}_+^n$ and denote by Σ_0 the map $\mathbb{T}^n \times \mathcal{A} \to X_{d_0}$ which sends (q, a) to the point of the torus $T^n(I_0)$ with the coordinate q.

THEOREM 5.1. *Suppose the assumptions* (H1)–(H3) *hold. Then there exist integers* $M_1 > 0$ *and* $M_2 > n$ *such that if* (H4) *is fulfilled, then for arbitrary* $\gamma > 0$ *and for sufficiently small* $\varepsilon < \bar{\varepsilon}(\gamma)$, *a Borel subset* $\mathcal{A}_\varepsilon \subset \mathcal{A}$ *and a Lipschitz map* $\Sigma_\varepsilon : \mathbb{T}^n \times \mathcal{A}_\varepsilon \to X_{d_0}$, *analytic in* $q \in \mathbb{T}^n$, *can be found with the following properties*:

(a) $\mathrm{mes}(\mathcal{A}\backslash\mathcal{A}_\varepsilon) \leqslant \gamma$;
(b) *the map* Σ_ε *is* $C\varepsilon$*-close to* $\Sigma_0|_{\mathbb{T}^n\times\mathcal{A}_\varepsilon}$ *in the Lipschitz norm*;
(c) *each torus* $\Sigma_\varepsilon(\mathbb{T}^n \times \{a\})$, $a \in \mathcal{A}_\varepsilon$, *is invariant for the flow of Equation* (5.1) *and is filled with its time-quasiperiodic solutions of the form* $u_\varepsilon(t;q) = \Sigma_\varepsilon(q+\omega' t, a)$, $q \in \mathbb{T}^n$, *where the frequency vector* $\omega'(a)$ *is* $C\varepsilon$*-close to* $\omega(a)$ *in the Lipschitz norm*;
(d) *the solutions* u_ε *are linearly stable.*[6]

If ∇H defines an analytic map of order d_H on every space X_d, $d \geqslant d_0$, then the solutions u_ε, constructed in the theorem, are smooth. Indeed, if $u_\varepsilon(t)$ is a solution, then due to the equation $JAu_\varepsilon(t)$ is a smooth curve in $X_{d_0-d_H-d_J}$. Since JA is an automorphism of the scale of order d_1, then $u_\varepsilon(t)$ is a smooth curve in $X_{d_0-d_H-d_J+d_1} \subset X_{d_0+1}$. Iterating this arguments we see that u_ε is a smooth curve in each space X_s.

In the semilinear case (i.e., when $d_H + d_J \leqslant \tilde{d} < d_1 - 1$ and $\tilde{d} \leqslant 0$) the theorem is proved in [49,50] (see also [52,73]). The semilinearity restriction $\tilde{d} \leqslant 0$ was removed in [57] (see also [59] and [46]). Simultaneously with [49,50] a related KAM-theorem for infinite-dimensional *Hamiltonian systems with short interactions* was proved by Pöschel [71] (following Eliasson's work [33] on lower-dimensional invariant tori for finite-dimensional systems). The systems (5.1), defined by HPDEs, are not short-interacted, but results of [71] apply to some equations from non-equilibrium statistical physics. For systems with short interaction a KAM-theory for infinite-dimensional invariant tori also is available, see [39,72] and references in [72]. We note that [39] was the first work where the KAM theory was applied to infinite-dimensional Hamiltonian systems.

For some specific HPDEs (5.1) the assertions of Theorem 5.1 can be proven for any $n \geqslant 1$ even if the parameter a is only one-dimensional. In particular, this can be done for the non-linear wave equation as in Example 5.3 below, where $V(x) \equiv a$ and the constant a is the one-dimensional parameter. See [16] and [4].

The proof of Theorem 5.1 is rather technical. For its well-written outline in the semilinear case see [28]. Below we present the proof's scheme in the form which suits our further purposes.

THE SCHEME OF THE PROOF OF THEOREM 5.1. We start with the semilinear case and assume for simplicity that $\lambda_j^J \equiv 1$. Then $I = (I_1, \dots, I_n)$ and $q = (q_1, \dots, q_n)$ form a symplectic coordinate system in the space (2.3). We set $Y = \mathrm{span}\{\varphi_j^\pm,\ j > n\} \subset X$, and denote by $y_j^\pm$, $j > n$, the coordinates in Y with respect to the basis $\{\varphi_j^\pm\}$. To study the vicinity of a torus $T^n(I_0)$, we make the substitution $I = I_0 + p$. Then $\bar{J}\,du \wedge du = dp \wedge dq + dy^+ \wedge dy^-$, and $T^n(I_0) = \{p = 0,\ y = 0\}$. In the new variables Equation (2.1) takes the form

$$\dot{q} = \nabla_p \mathcal{H}_\varepsilon, \qquad \dot{p} = -\nabla_q \mathcal{H}_\varepsilon, \qquad \dot{y} = J\nabla_y \mathcal{H}_\varepsilon,$$

with the Hamiltonian

$$\mathcal{H}_\varepsilon = H_0(p,y) + \varepsilon H_1(p,q,y), \quad H_0 = \omega \cdot p + \frac{1}{2}\langle Ay, y\rangle. \tag{5.8}$$

[6]If Equation (5.1) is not semilinear (i.e., if $d_J + d_H > 0$), then this assertion is proved provided that the equation satisfies some mild regularity condition, see Theorem 8.4 in [59].

The vector ω and the operator A depend on the parameter a; the function H_1 depends on a and I_0. We call H_0 the *integrable part of the Hamiltonian* $\mathcal{H}_\varepsilon$.

Retaining the terms of H_1 which are affine in p and quadratic in y, we write H_1 as

$$H_1 = H_1^1 + H_1^3, \quad H_1^1 = h(q) + h^p(q) \cdot p + \langle h^y(q), y\rangle + \langle h^{yy}(q)y, y\rangle,$$
$$H_1^3 = \mathrm{O}\big(|p|^2 + \|y\|^3 + |p|\,\|y\|\big) =: \mathcal{O}(p, q, y).$$

Next in the vicinity of the torus $T^n = \{p = 0,\ y = 0\}$ we make a symplectic change of variable to kill the part εH_1^1 of the perturbation εH_1. This change of variable is a transformation S_1 which is the time-ε shift along trajectories of an additional Hamiltonian F. Here the recipe is that to kill H_1^1, F should be of the same structure, so $F = f(q) + f^p(q) \cdot p + \langle f^y(q), y\rangle + \langle f^{yy}(q)y, y\rangle$. Due to Theorem 3.4 we can write the transformed Hamiltonian $\mathcal{H}_\varepsilon \circ S_1$ as

$$\mathcal{H}_\varepsilon \circ S_1 = H_0 + \varepsilon H_1 + \varepsilon\langle J\nabla_y F,\ \nabla_y H_0\rangle + \varepsilon\nabla_p F \cdot \nabla_q H_0 - \varepsilon\nabla_q F \cdot \nabla_p H_0$$
$$+ \mathrm{O}\big(\varepsilon^2\big) + \mathcal{O}.$$

Since $\nabla_p H_0 = \omega$, $\nabla_q H_0 = 0$ and $\nabla_y H_0 = Ay$, then the linear in ε term vanishes if the following relations hold:

$$(\omega \cdot \nabla) f = h, \qquad (\omega \cdot \nabla) f^p = h^p,$$
$$(\omega \cdot \nabla) f^y - JAf^y = h^y, \qquad (\omega \cdot \nabla) f^{yy} + [f^{yy},\ JA] = h^{yy}.$$

We take these relations as equations for f, f^p, f^y and f^{yy} (called *'the homological equations'*) and try to solve them.

Since the equations have constant coefficients, then decomposing $f, f^p, \ldots$ in Fourier series in q, we find for their components (and for matrix components of the operator f^{yy}) explicit formulae. Certain terms in these formulae contain small divisors, which vanish for some values of the vector $\omega = \omega(a)$. Careful analysis of these divisors show that all of them are bounded away from zero if $a \notin \mathcal{A}_1$, where $\mathcal{A}_1$ is a Borel subset of $\mathcal{A}$ of small measure. When the equations are solved, we get a symplectic transformation which in a sufficiently small neighbourhood of T^n transforms the Hamiltonian $\mathcal{H}_\varepsilon$ to a Hamiltonian which differs from its integrable part by $\mathrm{O}(\varepsilon^2)$.

The explanation above has some flows. The most important one is that the first and the second homological equations can be solved only if the mean values of h and h^p vanish. To fulfil the first condition we change the Hamiltonian εH_1 by a constant (this change is irrelevant since it does not affect the equations of motion), while to fulfil the second we subtract from εH_1 the average $\varepsilon\langle h^p\rangle \cdot p$ and add it to the integrable part H_0, thus changing the term $\omega \cdot p$ to $\omega^2 \cdot p$, where $\omega^2 = \omega + \varepsilon\langle h^p\rangle$. Similar, to solve the last homological equation we subtract from the operator h^{yy} the average of its diagonal part and add the corresponding quadratic form to H_0. Thus, the transformed Hamiltonian becomes

$$\mathcal{H}_2 := \mathcal{H}_\varepsilon \circ S_1 = \omega_2 \cdot p + \frac{1}{2}\langle A_2 y, y\rangle + \varepsilon^2 H_2(p, q, y) + \mathcal{O}(p, q, y).$$

This transformation is called *the KAM-step*.

Next we perform the second KAM-step. Under the condition that $a \notin \mathcal{A}_2$ we find a transformation S_2 which sends the Hamiltonian $\mathcal{H}_2$ to $\mathcal{H}_3 = \mathcal{H}_2 \circ S_2 = \omega_3 \cdot p + \frac{1}{2}\langle A_3 y, y\rangle + (\varepsilon^2)^2 H_2 + \mathcal{O}(p, q, y)$, etc. After m steps we find transformations $S_1, \ldots, S_m$ such that

$$\mathcal{H}_\varepsilon \circ S_1 \circ \cdots \circ S_m = \omega_m \cdot p + \frac{1}{2}\langle A_m y,\ y\rangle + \varepsilon^{2^m} H_m + \mathcal{O}(p, q, y) =: \mathcal{H}_m.$$

The torus $T^n = \{p = 0, y = 0\}$ is 'almost invariant' for the equation with the Hamiltonian $\mathcal{H}_m$. Hence, the torus $S_1 \circ \cdots \circ S_m(T^n)$ is 'almost invariant' for the original one. Since the sequence ε^{2^m} converges to zero super-exponentially fast, we can choose the sets $\mathcal{A}_1, \mathcal{A}_2, \ldots$ in such a way that $\operatorname{mes}(\mathcal{A}_\infty = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots) < \gamma$, for any $a \notin \mathcal{A}_\infty$ the vectors $\omega_m(a)$ converge to a limiting vector $\omega'(a)$, and the transformations $S_1 \circ \cdots \circ S_m$ converge to a limiting map $\Sigma_\varepsilon(\cdot, a)$, defined on T^n. Then the torus $\Sigma_\varepsilon(T^n,\ a)$ is invariant for Equation (5.1) and is filled with its quasiperiodic solutions $t \to \Sigma_\varepsilon(q + \omega' t, a)$. □

If the equation is not semilinear, then the situation is more complicated since to solve the forth homological equation we have to remove from the operator h^{yy} the whole of its diagonal part (not only its average). Because of that the operator A in the integrable part of the Hamiltonian gets terms which form a small q-dependent diagonal operator of a positive order. Accordingly, the forth homological equation becomes more difficult and cannot be solved by the direct Fourier method. Its resolution follows from a non-trivial lemma, based on properties of fast-oscillating Fourier integrals, proved in [57] (see also [59,46]).

5.2. *Applications to 1D HPDEs*

Theorem 1 well applies to parameter-depending quasilinear HPDEs with one-dimensional space variable in a finite interval, supplemented by boundary conditions such that spectrum of the linear operator JA is not multiple. Indeed, for such equations assumption (H2) follows from usual spectral asymptotics, (H3) is obvious if the non-linearity is analytic, while (H1) and (H4) hold if the equation depends on the additional parameter in a non-degenerate way. More explicitly it means the following. In the examples which we consider below, the equations depend on a potential $V(x;\ a)$, which is analytic in a and smooth in x. The non-degeneracy means that in a functional space, formed by functions of x and a of the required smoothness, the potential V should not belong to some analytic subset of infinite codimension.

Below we just list the examples. In each case application of Theorem 5.1 is straightforward. The theorem applies if dimension of the parameter a is $\geqslant n$ and dependence of the potential V on a is non-degenerate as it was explained above. In the first three examples the potential $V(x; a)$ is real, smooth in x and analytic in a. The function $f(x, v; a)$ is real, smooth in x and analytic in v and a. Details can be found in [52,53,59,57].

EXAMPLE 5.2. Non-linear Schrödinger equation (NLS), cf. Example 2.8:

$$\dot{u} = i\big(-u_{xx} + V(x; a)u + \varepsilon f\big(x, |u|^2; a\big)u\big), \quad u = u(t, x), x \in [0, \pi]; \tag{5.9}$$

$$u(t, 0) \equiv u(t,\ \pi) \equiv 0. \tag{5.10}$$

Now $d_J = 0$, $d_A = 2$, $\tilde{d} = d_H = 0$ and we view the Dirichlet boundary conditions as the odd periodic ones (cf. Example 2.9). The theorem applies in the scale of odd periodic functions with $d_0 = 1$ or 2. If f is even and 2π-periodic in x, then the theorem applies with any $d_0 \geqslant 1$ and the constructed quasiperiodic solutions are smooth.

EXAMPLE 5.3. Non-linear string equation: $w(t,x)$ satisfies (5.10) and

$$\ddot{w} = w_{xx} - V(x;a)w + \varepsilon f(x,w;a),$$

where now $V > 0$ and $f(x,w) = 0$ if $w = 0$ or $x = 0$. Let us denote $U = (u, -(-\Delta)^{-1/2}\dot{u})$. It is a matter of direct verification that U satisfies a semilinear Hamiltonian equation (5.1) in a suitable symplectic Hilbert scale, formed by odd periodic Sobolev vector-functions (cf. Equation (4.6)). Now $d_A = 1$, $d_J = 0$, $\tilde{d} = d_H = -1$. Cf. [79] and [16,4].

EXAMPLE 5.4 (*KdV-type equations*). KdV-type equation

$$\dot{u} = \frac{\partial}{\partial x}\big(-u_{xx} + V(x;a)u + \varepsilon f(x,u;a)\big); \quad x \in S^1,\ \int_{S^1} u\,dx \equiv 0, \tag{5.11}$$

cf. Example 2.7. Now $d_J = 1$, $d_A = 2$, $\tilde{d} = d_H = 0$.

Theorem 5.1 also applies if $x \in \mathbb{R}^1$ and the potential $V(x;\ a)$ grows sufficiently fast when $x \to \infty$.

EXAMPLE 5.5. Non-linear Schrödinger equation on the line:

$$\dot{u} = i\big(-u_{xx} + \big(x^2 + \mu x^4 + V(x;a)\big)u + \varepsilon f\big(|u|^2;a\big)u\big), \quad \mu > 0,$$
$$u = u(t,\ x),\ x \in \mathbb{R},\ u \to 0 \text{ as } |x| \to \infty.$$

Here the potential V is smooth, analytic in a and vanishes as $|x| \to \infty$. The real-valued function f is analytic. Now $d_J = 0$, $d_A = 4/3$, $d_H = 0$. Another example of this sort see in [52], Section 2.5.

The time-quasiperiodic solutions, constructed in Examples 5.2–5.5, are linearly stable. Therefore they should be observable in numerical models for the corresponding equations. Indeed, quasiperiodic behaviour of solutions for 1D HPDEs with small non-linearity was observed in many experiments, starting from the famous numerics of Fermi, Pasta and Ulam [36]; e.g., see [82].

5.3. *Multiple spectrum*

In Examples 5.2, 5.3 the equations are considered under the Dirichlet boundary conditions. If we replace them by the periodic ones

$$u(t,x) \equiv u(t, x + 2\pi),$$

then Theorem 5.1 would not apply since now the frequencies of the corresponding linear equations are asymptotically double: they have the form $\{\lambda_j^{\pm}, j \geqslant 1\}$, where $|\lambda_j^+ - \lambda_j^-| \to 0$ as $j \to \infty$. It is clear that the numbers $\{\lambda_j^{\pm}\}$ cannot be re-ordered to meet the spectral asymptotic condition (H2). Still, for some semilinear equations (5.1) assertions of the theorem remain true if the frequencies λ_j are not single, but asymptotically they have the same multiplicity $m \geqslant 2$ and behave regularly. A corresponding result is proved by Chierchia and You in [27], using the scheme, explained in Section 5.1. We do not give precise statement of their theorem, but note that it applies to the non-linear string equation in Examples 5.3 under the periodic boundary conditions. The result is the same: if the non-degeneracy condition holds, then for ε small enough and for most (in the sense of measure) values of the n-dimensional parameter a, solutions of the linear equation (5.3) which fill in a torus $T^n(I)$, $I \in \mathbb{R}_+^n$, persist as linearly stable time-quasiperiodic solutions of the corresponding non-linear equation (5.1).

We note that this persistence result was proved earlier by Bourgain [16], who used another KAM-scheme, discussed in the next section.

5.4. *Space-multidimensional problems*

The abstract Theorem 5.1 is a flexible tool to study 1D HPDEs, but it *never* applies to space-multidimensional equations since the spectral assumption (H2) never holds in dimensions > 1. The first KAM-theorem which applies to higher-dimensional HPDEs, is due to Bourgain [19]. In that work the 2D NLS equation as in Example 2.8 is considered. For technical reasons the potential term Vu is replaced there by the convolution $V * u$:

$$\dot{u} = i\left(-\Delta u + V(x; a) * u + \varepsilon \frac{\partial}{\partial \bar{u}} g(u, \bar{u})\right), \quad u = u(t, x),\ x \in \mathbb{T}^2. \tag{5.12}$$

The potential $V(x; a)$ is real analytic and $g(u, \bar{u})$ is a real-valued polynomial of u and $\bar{u}$. This equation has the form (5.1), where $Au = -\Delta u + V * u$ and $Ju = iu$. The basis $\{\varphi_k\}$ as in (5.2) is formed by normalised exponents $\{e^{is\cdot x}$ and $ie^{is\cdot x}, s \in \mathbb{Z}^2\}$, re-numerated properly, and

$$\lambda_s^J \equiv 1, \qquad \lambda_s^A = |s|^2 + \widehat{V}(s; a),$$

where $\{\widehat{V}(s; a)\}$ are the Fourier coefficients of V. For any n, the linear equation $(5.12)|_{\varepsilon=0}$ has quasiperiodic solutions

$$u = \sum_{j=1}^{n} z_{s_j} e^{i\lambda_{s_j}^A t} \varphi_{s_j}(x) \tag{5.13}$$

(these are trajectories of Equation (5.6) on the n-torus (5.5), where $I_j = \frac{1}{2}|z_{s_j}|^2$ and $I_s = 0$ if s differs from all s_j). For simplicity let us assume that $a_j = \widehat{V}(s_j;\ a)$, $j = 1, \ldots, n$. Then the result of [19] is that for most values of the parameter a (in the same sense as in Theorem 5.1), the solution (5.13) persists as a time-quasiperiodic solution of Equation (5.12). In contrast to the 1D case it is unknown if the new solutions are linearly stable.

The proof in [19] is based on a KAM-scheme, different from that described in Section 5.1. Originally this scheme is due to Craig and Wayne [29] who used it to construct periodic solutions of non-linear wave equations, using certain techniques due to Fröhlich–Spencer [38]. Also see [16].

Now we briefly describe the scheme, using the notations from Section 5.1. When the perturbation εH_1 is decomposed as in (5.8), we extract the term $\varepsilon\langle h^{yy}(q)y,\ y\rangle$ from εH_1^1 and add it to the integrable part H_0. After this the Hamiltonian to be killed is the sum of the three terms $h(q) + h^p(q) + \langle h^y(q), y\rangle$; accordingly the Hamiltonian F is a sum of three terms as well. We have to find them from the first three homological equations. The first two are not difficult, but the third one is a real problem since the operator A no longer has constant coefficients but equals $A_0 + \hat{A}(q)$, where $\hat{A}$ is a bounded operator of order ε (it changes from one KAM-step to another). The resolution of this equation for high KAM steps is the most difficult part of implementation the Craig–Wayne–Bourgain KAM-scheme.

Recently Bourgain managed to develop this scheme father and applied it to high-dimensional equations. We are not ready to discuss this and related results, and instead refer the reader to the original publications [23]. Also see [34].

5.5. *Perturbations of integrable equations*

Let us consider a quasilinear HPDE on a finite space-interval, which is an integrable Hamiltonian equation (4.1) in some symplectic Hilbert scale ($\{X_s\}$, $\alpha_2 = \bar{J}\,dx \wedge dx$). As we explained in Section 4.1, this equation has invariant finite-gap symplectic manifolds $\mathcal{T}^{2n}$ such that restriction of (4.1) to any of them is integrable. In this section we discuss the results on persistence of quasiperiodic solutions that fill in these manifolds, provided by the KAM for PDEs theory. We shall see that they are very similar to the celebrated Kolmogorov theorem, which states that *most of quasiperiodic solutions of a non-degenerate analytic integrable* (*finite-dimensional*) *Hamiltonian system persist under small perturbations of the Hamiltonian*; see [1,65,78] and Addendum in [59]. We state the main result as a

THEOREM 5.6 (Metatheorem). *Most of quasiperiodic solutions that fill in any finite-gap manifold $\mathcal{T}^{2n}$ as above persist under small Hamiltonian quasilinear analytic perturbations of the integrable equation. If the finite-gap solutions in $\mathcal{T}^{2n}$ are linearly stable, then the new solutions are linearly stable as well.*

In the assertion above the statement 'most of quasiperiodic solutions persist' means the following. Due to the Liouville–Arnold theorem [2,43], the manifold $\mathcal{T}^{2n}$ can be covered by charts, diffeomorphic to $B \times \mathbb{T}^n = \{p, q\}$ (B is a ball in $\mathbb{R}^n$), with chart-maps $\Phi_0 : B \times \mathbb{T}^n \to \mathcal{T}^{2n}$ such that $\Phi_0^*\alpha_2 = dp \wedge dq$, and the curves $\Phi_0(p, q + t\nabla h(p))$ are solutions of the integrable equation, where $h(p) = H \circ \Phi_0(p, q)$. Let us denote by ε the small coefficient in front of the perturbation. Then for every chart there exists a Borel subset $B_\varepsilon \subset B$ and a map $\Phi_\varepsilon : B_\varepsilon \times \mathbb{T}^n \to X_d$ (d is fixed), with the following properties:

(i) $\operatorname{mes}(B \setminus B_\varepsilon) \to 0$ as $\varepsilon \to 0$;
(ii) the map $\Phi_\varepsilon : B_\varepsilon \times \mathbb{T}^n \to X_d$ is $C\sqrt{\varepsilon}$-close to Φ_0 in the Lipschitz norm and is analytic in $q \in \mathbb{T}^n$;
(iii) there exists a map $\omega_\varepsilon : B_\varepsilon \to \mathbb{R}^n$, $C\varepsilon$-close to the gradient map ∇h in the Lipschitz norm, such that the curves $t \mapsto \Phi_\varepsilon(p, q + t\omega_\varepsilon(p))$, $p \in B_\varepsilon$, $q \in \mathbb{T}^n$, are solutions for the perturbed equation.

The statement of Theorem 5.6 is proven under a number of assumptions (see [59,35]). These assumptions are checked for such basic integrable HPDEs as KdV, Sine- and Sinh-Gordon equations. There are no doubts that they also hold for the Zakharov–Shabat equations[7] (but the theorem in [59,35] does not apply to the Kadomtsev–Petviashvili equation). Below we present a scheme of the proof and discuss the restrictions on the integrable HPDE which allow to implement it.

We view (4.1) as an equation in the Hilbert space X_d, and denote the quasilinear Hamiltonian of the perturbed equation as

$$H_\varepsilon = \frac{1}{2}\langle Ax,\, x\rangle + h_0(x) + \varepsilon h_1(x).$$

Accordingly, $H_0 = \frac{1}{2}\langle Ax,\, x\rangle + h_0$ is the Hamiltonian H of the unperturbed equation (4.1).

Step 1. Let us consider any finite-gap solution $u_0(t) = \Phi_0(p_0,\, q_0 + t\nabla h(p_0))$ and linearise (4.1) about it:

$$\dot{v} = J\big(\nabla H\big(u_0(t)\big)\big)_* v. \tag{5.14}$$

The theory of integrable equations provides tools to reduce this equation to constant coefficients by means of a time-quasiperiodic substitution $v(t) = G(p_0, q_0 + t\nabla h(p_0))\tilde{v}(t)$, where $G(p,q)$, $(p,q) \in B \times \mathbb{T}^n$, is a symplectic linear map $G(p,q) : Y_d \to Z_d$ (see [59, Sections 5, 6]). Here Y_d is a fixed symplectic subspace of Z_d of codimension $2n$. The restriction, which we impose at this step, is that the operator $G(p,q)$ is a compact perturbation of the embedding $Y_d \to Z_d$, which analytically depends on (p,q).

Step 2. The map G from the Step 1 defines an analytic map

$$B \times \mathbb{T}^n \times Y_d \to X_d,$$

linear and symplectic in $y \in Y_d$. This map defines a symplectomorphism

$$B \times \mathbb{T}^n \times B_\delta(Y_d) \to X_d, \quad B_\delta(Y_d) = \big\{\|y\|_d < \delta\big\}, \tag{5.15}$$

such that linearisation in y at $y = 0$ of the latter equals the former ([59, Section 7]).

Step 3. We use the map (5.15) to pass in the Hamiltonian H_ε to the variables (p,q,y). Retaining linear and quadratic in y terms we get

$$H_\varepsilon(p,\, q,\, y) = h(p) + \frac{1}{2}\big\langle \mathcal{A}(p)y, y\big\rangle + h_3(p,q,y) + \varepsilon h_1(p,q,y), \tag{5.16}$$

[7] See [41] for an *ad hoc* KAM-theorem for the defocusing equation.

where $h_3 = O(\|y\|_d^3)$. Calculations show that $h_3(p,q,y)$ contains terms such that their gradient maps have the same order as the operator $\mathcal{A}(p)$. If this really was the case, then the Hamiltonian equation would not be quasilinear, which would complicate its study a lot. Fortunately, this does not happen due to a cancellation of a very general nature (see Lemma 7.5 in [59]), and we have

$$\operatorname{ord}\nabla h_3 < \operatorname{ord}\mathcal{A}(p) - 1. \tag{5.17}$$

Step 4. Invariant tori of the unperturbed system with the Hamiltonian $H_0(p,q,y)$ have the form $\{p = \text{const}, y = 0\}$. Let us scale the variables near a torus $\{p = a, y = 0\}$: $p = a + \varepsilon^{2/3}\tilde{p}$, $q = \tilde{q}$, $y = \varepsilon^{1/3}\tilde{y}$. In the scaled variables the perturbed equation has the Hamiltonian

$$\text{const} + \omega(a)\cdot\tilde{p} + \frac{1}{2}\big\langle\mathcal{A}(a)\tilde{y},\tilde{y}\big\rangle + O\big(\varepsilon^{1/3}\big), \quad \omega(a) = \nabla h(a). \tag{5.18}$$

So we have got the system (5.1), written in the form (5.8), with ε replaced by $\varepsilon^{1/3}$. If Theorem 5.1 applies, then most of the finite-gap tori $\{p = \text{const}\}$ persist in the perturbed equation, as states the Metatheorem. To be able to use the theorem we have to check the assumptions (H1)–(H4).

The condition (H2) holds if the integrable equation is 1D (if the spectrum is asymptotically double, e.g., if the unperturbed equation is the Sine-Gordon equation under the periodic boundary conditions, then one should use a version of the Metatheorem, based on the Chierchia–You result). The quasilinearity condition (H3) holds due to (5.17). The assumption (H1) now takes the form

$$\operatorname{Hess} h(p) \not\equiv 0. \tag{5.19}$$

This is exactly Kolmogorov's non-degeneracy condition for the integrable system on $\mathcal{T}^{2n}$. The assumption (H4) with $\omega = \nabla h(a)$ is the second non-degeneracy condition, which needs verification.

Summing up what was said above, we see that Theorem 5.1 implies the Metatheorem if the unperturbed integrable equation is 1D quasilinear, the linear operator $G(p,q)$ from Step 1 possesses the required regularity properly and the non-degeneracy assumptions (5.19) and (5.7) hold true.

The scheme we have just explained was suggested in [51], where it was used to prove an abstract KAM-theorem, which next was applied to Birkhoff-integrable infinite-dimensional systems and to perturbed KdV equations. See [59,35] for a more general abstract theorem, based on the same scheme.

Steps 1–2 are not the only way to reduce an integrable equation to the normal form (5.16). Another approach to get it had been initiated by Kappeler [44]. It was developed further in a number of publications and finally in [45] it was proved that the KdV equation is Birkhoff-integrable. It means the following. Let us take the Darboux scale $(\{X_s\},\alpha_2)$ with the index-set $\mathcal{Z} = \mathbb{Z}_0$, and $\theta_k = |k|$ (see Example 2.6). Then there exists a map $\Phi : X_\infty \to H^\infty(S^1)_0$ which extends to analytic maps $X_s \to H^s(S^1)_0$, $s \geqslant 0$, such that

$$h \circ \Phi(u) = \sum_{j=1}^{\infty} j^3\big(u_j^2 + u_{-j}^2\big) + \big\langle \text{a function of } u_l^2 + u_{-l}^2,\ l = 1, 2, \dots\big\rangle. \quad (5.20)$$

Here $\{u_k, k \in \mathbb{Z}_0\}$ are coefficients of decomposition of $u \in X_s$ in the basis $\{\varphi_k\}$ and h is the KdV-Hamiltonian (see Example 2.7). Moreover, the Hamiltonian (5.20) defines an analytic Hamiltonian vector field of order three in each space X_d, $d \geqslant 1$. In the transformed variables the N-gap tori of the KdV equation take the form (5.5), where $n \geqslant N$ and exactly N numbers I_j are non-zero. Now let us take a torus (5.5), where $I \in \mathbb{R}^n_+$. Making a change of variables as in Section 5.1, we arrive at the Hamiltonian (5.18). Detailed and readable derivation of the normal form (5.20) see in [46].

Reduction to the Birkhoff normal form (5.20) uses essentially specifics of the KdV's L-operator. Still, similar arguments apply as well to the defocusing Zakharov–Shabat equation, see [41]. Presumably, the Birkhoff normal forms exist for some other integrable equations with selfadjoint L-operators, but not for equations with non-selfadjoint operators. In particular, the focusing Zakharov–Shabat equation cannot be reduced to the form (5.20) since for this equation some finite-gap tori are linearly unstable [26], while all invariant tori of the form (5.5) for the Hamiltonian (5.20) are linearly stable.

EXAMPLE 5.7 (*Perturbed KdV equation*). Consider the equation

$$\dot u(t,x) = \frac{1}{4}\frac{\partial}{\partial x}\big(u'' + 3u^2 + \varepsilon f(x,u)\big), \quad x \in S^1; \ \int_{S^1} u\,dx \equiv 0, \quad (5.21)$$

where f is smooth in x, u and analytic in u. The Metatheorem applies and implies that most of finite-gap KdV-solutions persist as time-quasiperiodic solutions of (5.21). Moreover, these solutions are smooth and linearly stable.

This result was first stated in [51]. The proof contains some gaps. Two the most serious of them are that Theorem 5.1, proved then only for semilinear equations, was used in a quasilinear case, and that the non-degeneracy assumptions (5.19) and (5.7) were taken for granted. These gaps were filled in later. The quasilinear version of Theorem 5.1 was proved in [57] (preprint of this paper appeared in 1995), and the non-degeneracy conditions were verified in [12]. Also see [59, Section 6.2.1]. The arguments in [12,59] are general and applies to other equations.

For a complete proof of 'KAM for KdV' see [59,35] and [46].

The Metatheorem (in its rigorous form as in [59,35] and [46]), applies to quasilinear Hamiltonian perturbations of any higher equation from the KdV-hierarchy, provided that the non-degeneracy relations are checked for this equation. It can be done in the same way as in Example 5.7. See [46], where the non-degeneracy of the second KdV equation is verified.

EXAMPLE 5.8 (*Perturbed SG equation*). Consider the equation

$$\ddot{u} = u_{xx} - \sin u + \varepsilon f(u, x), \qquad u(t,\, 0) = u(t, \pi) = 0, \tag{5.22}$$

where $f(0, x) \equiv 0$ (and $f \in C^\infty$ is analytic in u). The Metatheorem applies to prove persistence most of finite-gap solutions of the SG-equation, see [11,59,35]. In general, due to the phenomenon explained in Example 2.9, the persisted solutions are only H^2-smooth in x. But if f is x-independent and odd in u, then they are smooth.

In difference with the KdV-case, large amplitude finite-gap SG-solutions, as well as the corresponding persisted solutions of (5.22), in general are not linearly stable.

To end this section we note that since the persisted solutions $u_\varepsilon(t)$ have the form

$$u_\varepsilon(t) = \Phi_\varepsilon\big(p, q + t\omega_\varepsilon(p)\big) = \Phi_0\big(p, q + t\omega_\varepsilon(p)\big) + \mathrm{O}\big(\sqrt{\varepsilon}\big),$$

then to calculate them with the accuracy $\sqrt{\varepsilon}$ for all values of time t, we can use the "finite gap map" Φ_0 with the corrected frequency vector. Moreover, $\omega_\varepsilon(p) = \nabla h(p) + \varepsilon W_1(p) + \mathrm{O}(\varepsilon^2)$, where the vector $W_1(p)$ can be obtained by averaging over the corresponding finite-gap torus of some explicit quantity, see [59, p. 147].

5.6. *Small amplitude solutions of HPDEs*

Let us consider the non-linear string equation

$$u_{tt} = u_{xx} - mu + f(u), \quad u = u(t, x),\ 0 \leqslant x \leqslant \pi; \qquad u(t, 0) = u(t, \pi) = 0. \tag{5.23}$$

Here $m > 0$ and f is an odd analytic function of the form

$$f(u) = \kappa u^3 + \mathrm{O}\big(u^5\big), \quad \kappa > 0.$$

Since $m,\ \kappa > 0$, then constants $a,\ b > 0$ can be found such that $-mu + f(u) = -a \sin bu$. Hence, Equation (5.23) can be written as

$$u_{tt} = u_{xx} - a \sin bu + \mathrm{O}\big(u^5\big).$$

After the scaling $u = \varepsilon w$, $\varepsilon \ll 1$, the higher-order perturbation transforms to a small one, and we can apply the Metatheorem (cf. Example 5.8) to prove that small-amplitude parts of the finite-gap manifolds $\mathcal{T}^{2n}$, $n = 1, 2, \ldots$, for the SG equation $u_{tt} = u_{xx} - a \sin bu$ with the Dirichlet boundary conditions mostly persist in (5.23). To put this scheme through, the small-amplitude parts

$$\mathcal{T}_\delta^{2n} = \big\{(u, \dot{u}) \in \mathcal{T}^{2n} \mid \|u\| + \|\dot{u}\| < \delta\big\}, \quad 0 < \delta \ll 1,$$

of the manifolds $\mathcal{T}^{2n}$ have to be studied in details. This task was accomplished in [14], where the following results were proved:

(i) the sets $\overline{\mathcal{T}_\delta^{2n}}$ are smooth manifolds which contain the origin,
(ii) they are in one-to-one correspondence with their tangent spaces at the origin,
(iii) these tangent spaces are invariant spaces for the Klein–Gordon equation $u_{tt} = u_{xx} - (ab)u$.

Another proof of (i)–(iii) was suggested in [59]. It is based on some ideas from [44] and applies to other integrable equations. After (i)–(iii) are obtained, a version of the Metatheorem (or a version of Theorem 5.1) applies to prove that most of finite-gap solutions from a manifold $\mathcal{T}_\delta^{2n}$ persist in (5.23) in the following sense: the $2n$-dimensional Hausdorff measure of the persisted part of the manifold, divided by a similar measure of $\mathcal{T}_\delta^{2n}$, converges to one as $\delta \to 0$. See [13] for a proof and [53] for discussion.

Similar results hold for the NLS equation

$$i\dot{u} = u_{xx} + mu + f\big(|u|^2\big)u, \qquad f(0) = 0, \qquad f'(0) = \gamma \neq 0, \tag{5.24}$$

where f is analytic, since it is a higher-order perturbation of the Zakharov–Shabat equation (4.7). But it turns out that it is easier to approximate (5.24) near the origin by its partial Birkhoff normal form. The latter is an integrable infinite-dimensional Hamiltonian system (which is not an HPDE), and a sibling of the Metatheorem applies to prove that most of its time-quasiperiodic solutions persist in (5.24), see [60]. More on the techniques of Birkhoff normal forms in HPDE see in [74] and [46]. The classical reference for finite-dimensional Birkhoff normal forms is the book [65].

6. Around the Nekhoroshev theorem

The classical Nekhoroshev theorem [66] deals with nearly-integrable Hamiltonian systems with analytic Hamiltonians $H_\varepsilon(p, q) = h(p) + \varepsilon H(p, q)$ on the phase-space $P \times \mathbb{T}^n$, $P \subset \mathbb{R}^n$, given the usual symplectic structure $dp \wedge dq$. Under the assumption that the Hamiltonian $h(p)$ satisfies a mild non-degeneracy condition called *the steepness*, the theorem states that the action variables change exponentially slow along trajectories of the system. Namely, there exist constants $a, b \in (0, 1)$ such that for any trajectory $(p(t), q(t))$ of the system we have

$$\big|p(t) - p(0)\big| \leqslant C\varepsilon^a \quad \text{if } |t| \leqslant \exp\big(\varepsilon^{-b}\big). \tag{6.1}$$

Strictly convex functions $h(p)$ form an important class of the steep Hamiltonians. An alternative proof of the theorem which applies in the convex case was suggested by Lochak [63]. It is based on clever approximation of a trajectory $(p(t), q(t))$ by a time-periodic solution of the equation which is a high-order normal form for H_ε. So rational frequency-vectors play for the Lochak approach very important role.

Original Nekhoroshev's proof contains two parts, analytical and geometrical. The techniques, developed in the analytical part of the proof, allow to get the following result, which we call below the quasi-Nekhoroshev theorem: Let us consider the Hamiltonian H_ε, depending on an additional vector-parameter $\omega \in \Omega \Subset \mathbb{R}^n$, $H_\varepsilon = p \cdot \omega + \varepsilon H(p, q)$. Then for any $\gamma > 0$ there exists a Borel subset $\Omega_\gamma \subset \Omega$ ('the Diophantine subset') such that $\mathrm{mes}(\Omega \setminus \Omega_\gamma) < \gamma$, and (6.1) with $C = C_\gamma$ holds if $\omega \in \Omega_\gamma$. Note that in the Cartesian coordinates (x, y), corresponding to the action-angle variables (p, q) (i.e., $x_j = \sqrt{2p_j}\cos q_j$, $y_j = \sqrt{2p_j}\sin q_j$), the Hamiltonian H_ε reeds as

$$H_\varepsilon = \frac{1}{2}\sum \omega_j\big(x_j^2 + y_j^2\big) + \varepsilon H(x, y).$$

That is, H_ε is a perturbation of the quadratic Hamiltonian H_0. So the quasi-Nekhoroshev theorem implies long-time stability of the zero equilibrium for an analytical Hamiltonian

$$H(x, y) = H_0 + h, \quad h = \mathrm{O}\big(\big|(x, y)\big|^3\big), \tag{6.2}$$

provided that the vector ω belongs to the Diophantine set. In [67] Niederman used the Lochak approach to get a stronger theorem on stability for (6.2). Namely, he proved that the equilibrium is stable during the exponentially long time if the vector ω does not satisfies resonant relations up to order four, and h is convex in a certain sense.[8]

To get a corresponding theorem which applies to all small initial data is a non-trivial task, resolved by Niederman [67] by means of the Lochak approach.

No analogy of the Nekhoroshev theorem for HPDEs is known yet, but a number of *ad hoc* quasi-Nekhoroshev theorems for HPDEs were proved, mostly by Bourgain and Bambusi, see [3,4,22] and references therein. These works discuss stability of the equilibrium for HPDEs (mostly 1D) with Hamiltonians of the form (6.2). Under some restrictions on the quadratic part H_0 and on the higher-order part h, it is proved that if the initial data u_0 is an ε-small and 'very' smooth function, then a solution stays very close to the corresponding invariant torus of the linear system with the Hamiltonian H_0, during the time which is polynomially large in ε^{-1}, or even exponentially large. This result is obtained either under the 'quasi-Nekhoroshev' condition that the spectrum of the operator A is 'highly non-resonant', or under the opposite assumption (needed to apply the Lochak–Niederman technique) that the spectrum is 'very resonant'. In particular, the following result is proved in [3] (also see [75,22]): Let us consider the NLS equation (5.24) in the scale $\{H_0^s(0, \pi)\}$ of odd 2π-periodic functions. Assume that $u_0(x) = \sum_{k=1}^N u_{k0}\sin kx$, denote $\varepsilon = |u_0(x)|_{L_2} \ll 1$ and write the solution $u(t, x)$ of (5.24) as $u = \sum u_k(t)\sin kx$. Then there exist $\varepsilon_* > 0$ and constants $C_1, C_2 > 0$ such that for $\varepsilon < \varepsilon_*$ and $|t| \leqslant C_1\exp(\varepsilon_*/\varepsilon)^{1/N} =: T_\varepsilon$ we have

$$\sum_{k=1}^{\infty}\big(\big|u_k(t)\big|^2 - |u_{k0}|^2\big)^2 \leqslant C_2\varepsilon^{4+1/N}. \tag{6.3}$$

Let us set $T^N = \{u(x) = \sum_{k=1}^N u_k \sin kx \mid |u_k| = |u_{k_0}|\}$. This is an n-torus of diameter $\sim \varepsilon$ and (6.3) implies that

$$\mathrm{dist}_{H_0^s}\big(u(t), T^n\big) \leqslant C_s\varepsilon^{1+1/N} \quad \forall |t| \leqslant T_\varepsilon,$$

[8] Independently this result was obtained in [9] by means of the Nekhoroshev's techniques.

if $s < -1/4$. Thus, during the time T_ε the trajectory $u(t)$ remains very close to its projection to T^N. The latter is a trajectory of an N-dimensional dynamical system, so the time of its return to a $\rho\varepsilon$-neighbourhood ($\rho \ll 1$) of the initial point 'should' be of order ρ^{-N}. Same is true for the trajectory $u(t)$, if ε is small in terms of ρ. The phenomenon of the pathologically good recurrence properties of small-amplitude trajectories of some non-integrable 1D HPDEs is well known from numerics (e.g., see [82]). We have seen that the quasi-Nekhoroshev theorems as above explain it up to some extend.

7. Invariant Gibbs measures

If Equation (4.1) is a finite-dimensional Hamiltonian system with $u = (p, q) \in (\mathbb{R}^{2n}, dp \wedge dq)$, then any measure $f(H(p, q))\, dp\, dq$ such that the function $f \circ H$ is Lebesgue-integrable, is invariant for the equation. The most important among these measures is the Gibbs measure $e^{-H}\, dp\, dq$ (the Hamiltonian H is assumed to grow to infinity with $|(p, q)|$). Now let us consider an HPDE (4.1). Say, the zero-mass ϕ^4-equation

$$\ddot{u} = u_{xx} - u^3, \quad u = u(t, x),\ x \in S^1.$$

This equation is equivalent to the system

$$\begin{aligned} \dot{u} &= -Bv, \\ \dot{v} &= Bu + B^{-1}\big(u^3 - u\big), \end{aligned} \tag{7.1}$$

where $B = \sqrt{1 - \Delta}$. Denoting $\xi = (u, v)$ we can see that this is a Hamiltonian system in the symplectic scale $(\{Z_s = H^{s+1/2}(\mathbb{T}^2; \mathbb{R}^2)\}, \alpha_2 = \bar{J}\, d\xi \wedge d\xi)$, where $J(u, v) = (-v, u)$, with the Hamiltonian

$$H(\xi) = \frac{1}{2}\|\xi\|_0^2 + \int \left(\frac{1}{4}|u|^4 - \frac{1}{2}|u|^2\right) dx, \quad \xi = (u, v).$$

Here $\|\cdot\|_0$ is the norm in the space $H^{1/2}(S^1; \mathbb{R}^2)$ (cf. Section 8.3). The natural question is if the formal expression

$$\mu = e^{-H(\xi)}\, d\xi \tag{7.2}$$

defines a measure in a suitable function space $\Xi = \{\xi(x)\}$, invariant for flow-maps of Equation (7.1). Since the Lebesgue measure $d\xi$ does not exist in an infinite-dimensional function space, then to make the right-hand side of (7.2) meaningful we write it as

$$\mu = e^{-\int (\frac{1}{4}|u|^4 - \frac{1}{2}|u|^2)\, dx} e^{-\frac{1}{2}\|\xi\|_0^2}\, d\xi.$$

Now $\exp -\frac{1}{2}\|\xi\|_0^2\, d\xi$ is a well-defined Gaussian measure, supported by a suitable space Ξ, formed by functions of low smoothness, and $0 < p(\xi) \leqslant C$, where $p(\xi) = e^{-\int (\frac{1}{4}|u|^4 - \frac{1}{2}|u|^2)\, dx}$. Therefore if

(i) $p(\xi)$ is a Borel function on Ξ, then μ is a well-defined Borel measure on Ξ.

To check that it is invariant for Equation (7.1) we have to verify that

(ii) the flow-maps of (7.1) are well-defined on $\operatorname{supp}\mu$ and preserve the measure.

The corresponding result was first stated by Friedlander [37]. Unfortunately, his arguments contain serious flaws. Complete proofs appeared later in works of Zhidkov, McKean and Vaninsky and Bourgain, see the books [20,84] and references therein. Similar arguments apply to the 1D NLS equation (2.4), where the non-quadratic term q satisfies certain restrictions.

For higher-dimensional HPDEs the task of constructing the Gibbs measures becomes much more difficult. The only known result is due to Bourgain who proved that for the defocusing 2D NLS equation

$$i\dot{u} = \Delta u - |u|^2 u, \quad x \in \mathbb{T}^2,$$

the Gibbs measure (7.2) exists and is invariant. The main difficulty here is the step (ii) which is now based on highly non-trivial results on regularity of corresponding flow-maps in Sobolev spaces of low smoothness; see in [20].

8. The non-squeezing phenomenon and symplectic capacity

8.1. *The Gromov theorem*

Let $(\mathbb{R}^{2n}, \beta_2)$ be the space $\mathbb{R}^{2n} = \{x_1, x_{-1}, \dots, x_{-n}\}$ with the Darboux symplectic form $\beta_2 = \sum dx_j \wedge dx_{-j}$. By $B_r(x) = B_r(x; \mathbb{R}^{2n})$ and $C^j_\rho = C^j_\rho(\mathbb{R}^{2n})$, $1 \leqslant j \leqslant n$, we denote the following balls and cylinders in $\mathbb{R}^{2n}$:

$$B_r(x) = \{y \mid |y - x| < r\}, \qquad C^j_\rho = \{y = (y_1, \dots, y_{-n}) \mid y_j^2 + y_{-j}^2 < \rho^2\}.$$

The famous (*non-*)*squeezing theorem* by M. Gromov [42] states that if f is a symplectomorphism $f : B_r(x) \to \mathbb{R}^{2n}$ such that its range belongs to some cylinder $x_1 + C^j_\rho$, $x_1 \in \mathbb{R}^{2n}$, then $\rho \geqslant r$. For an alternative proof, references and discussions see [43].

8.2. *Infinite-dimensional case*

Let us consider a symplectic Hilbert scale $(\{Z_s\}, \alpha_2)$ with a basis $\{\varphi_j \mid j \in \mathbb{Z}_0\}$. We assume that this is a shifted Darboux scale (cf. Example 2.4 in Section 2.2). It means that the basis can be renormalised to a basis $\{\tilde\varphi_j \mid j \in \mathbb{Z}_0\}$ (each $\tilde\varphi_j$ is proportional to φ_j) which is a Darboux basis for the form α_2 and a Hilbert basis of some space Z_d:

$$\langle \tilde\varphi_j, \tilde\varphi_k \rangle_d = \delta_{j,k}, \qquad \alpha_2[\tilde\varphi_j, \tilde\varphi_{-k}] = \operatorname{sgn} j\, \delta_{j,k} \quad \forall j, k. \tag{8.1}$$

These relations imply that

$$\alpha_2[\xi, \eta] = \langle \bar J \xi, \eta \rangle_d, \qquad \bar J \tilde\varphi_j = \operatorname{sgn} j\, \tilde\varphi_{-j} \quad \forall j. \tag{8.2}$$

In particular, $\bar{J} = J$.

Below we skip the tildes and re-denote the new basis back to $\{\varphi_j\}$.

In this scale we consider a semilinear Hamiltonian equation with the Hamiltonian $H(u) = \frac{1}{2}\langle Au, u\rangle_d + h(u, t)$. Due to (8.2) it can be written as

$$\dot{u} = JAu + J\nabla^d h(u, t), \tag{8.3}$$

where ∇^d signifies the gradient in u with respect to the scalar product of Z_d.

If a Hamiltonian PDE is written in the form (8.3), then the symplectic space (Z_d, α_2) is called the (*Hilbert*) *Darboux phase space* for this PDE. Below we study properties of flow-maps of Equation (8.3) in its Darboux phase space.

Let us assume that the operator A has the form

(H1) $Au = \sum_{j=1}^{\infty} \lambda_j (u_j \varphi_j + u_{-j}\varphi_{-j})\ \forall u = \sum u_j \varphi_j$, where λ_j's are some real numbers.

Then $JAu = \sum_{j=1}^{\infty} \lambda_j (u_{-j}\varphi_{-j} - u_j \varphi_j)$, so the linear operators e^{tJA} are direct sums of rotations in the planes $\mathbb{R}\varphi_j + \mathbb{R}\varphi_{-j} \subset Z_d$, $j = 1, 2, \ldots$.

We also assume that the gradient map $\nabla^d h$ is smoothing:

(H2) there exists $\gamma > 0$ such that $\operatorname{ord} \nabla^d h = -\gamma$ for $s \in [d - \gamma, d + \gamma]$. Moreover, the maps

$$\nabla^d h : Z_s \times \mathbb{R} \to Z_{s+\gamma}, \quad s \in [d - \gamma, d + \gamma],$$

are C^1-smooth and bounded.[9]

For any t and T we denote by O_t^T any open subset of the domain of definition of the flow-map S_t^T in Z_d, such that for each bounded subset $Q \subset O_t^T$ the set $\bigcup_{\tau \in [t,T]} S_t^\tau(Q)$ is bounded in Z_d.[10]

In the theorem below the balls B_r and the cylinders C_ρ^j, $j \geqslant 1$, are defined in the same way as in Section 8.1.

THEOREM 8.1. *Assume that* (H1) *and* (H2) *hold and that a ball* $B_r = B_r(u_0; Z_d) := \{\|y - u_0\|_d < r\}$ *belongs to* O_t^T *together with some* ε*-neighbourhood,* $\varepsilon > 0$. *Then the relation*

$$S_t^T(B_r) \subset v_0 + C_\rho^j(Z_d) \tag{8.4}$$

with some $v_0 \in Z_d$ *and* $j \geqslant 1$ *implies that* $\rho \geqslant r$.

PROOF. Without lost of generality we may assume that

$$v_0 = 0, \qquad j = 1.$$

Arguing by contradiction we assume that (8.4) holds with $\rho < r$ and choose any $\rho_1 \in (\rho, r)$.

[9] I.e., they send bounded sets to bounded.

[10] This set should be treated as a ‘regular part of the domain of definition’.

For $n \geqslant 1$ we denote by E^{2n} the subspace of Z_d, spanned by the vectors $\{\varphi_j, |j| \leqslant n\}$, and provide it with the usual Darboux symplectic structure (it is given by the form $\alpha_2|_{E^{2n}}$). By Π_n we denote the orthogonal projection $\Pi_n : Z_d \to E^{2n}$. We set

$$H^n = \frac{1}{2}\langle Au, u\rangle_d + h\big(\Pi_n(u), t\big)$$

and denote by $S^T_{(n)t}$ flow-maps of the Hamiltonian vector filed V_{H^n}. Any map $S^T_{(n)t}$ decomposes to the direct sum of a symplectomorphism of E^{2n} and of a linear symplectomorphism of $Z_d \ominus E^{2n}$. So the theorem's assertion with the map S^T_t replaced by $S^T_{(n)t}$ follows from the Gromov theorem, applied to the symplectomorphism

$$E^{2n} \to E^{2n}, \qquad x \mapsto \Pi_n S^T_{(n)t}\big(i(x) + u_0\big),$$

where i stands for the embedding of E^{2n} to Z_d.

Proofs of the two easy lemmas below can be found in [54].

LEMMA 8.2. *Under the theorem's assumptions the maps $S^T_{(n)t}$ are defined on B_r for $n \geqslant n'$ with some sufficiently large n', and there exists a sequence $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$ such that*

$$\big\| S^T_t(u) - S^T_{(n)t}(u)\big\| \leqslant \varepsilon_n \tag{8.5}$$

for $n \geqslant n'$ and for every $u \in B_r$.

LEMMA 8.3. *For any $u \in B_r$ we have $S^T_t(u) = e^{(T-t)JA}u + \widetilde{S}^T_t(u)$, where $\widetilde{S}^T_t$ is a C^1-smooth map in the scale $\{Z_s\}$ and* $\operatorname{ord} \widetilde{S}^T_t = -\gamma$ *for $s \in [d-\gamma, d+\gamma]$.*

Now we continue the proof of the theorem. Since its assertion holds for any map $S^T_{(n)t}$ $(n \geqslant n')$ and since the ball B_r belongs to this map's domain of definition (see Lemma 8.2), then for each $n \geqslant n'$ there exists a point $u_n \in B_r$ such that $S^T_{(n)t}(u_n) \notin C^1_{\rho_1}(0)$. That is,

$$\big|\Pi_1 S^T_{(n)t}(u_n)\big| \geqslant \rho_1. \tag{8.6}$$

By the weak compactness of a Hilbert ball, we can find a weakly converging subsequence

$$u_{n_j} \rightharpoonup u \in B_r, \tag{8.7}$$

so

$$u_{n_j} \to u \quad \text{strongly in } Z_{d-\gamma}.$$

Due to Lemma 8.3 this implies that $\widetilde{S}^T_t(u_{n_j}) \to \widetilde{S}^T_t(u)$ in Z_d, and using (8.7) we obtain the convergence:

$$S^T_t(u_{n_j}) \rightharpoonup S^T_t(u). \tag{8.8}$$

Noting that $|\Pi_1 S_t^T(u_n)| = |\Pi_1 S_{(n)t}^T u_n + \Pi_1(S_t^T - S_{(n)t}^T)u_n|$ and using (8.6), (8.5) we get:

$$\left|\Pi_1 S_t^T(u_n)\right| \geqslant \rho_1 - \varepsilon_n, \quad n \geqslant n'. \tag{8.9}$$

Since by (8.8) $\Pi_1 S_t^T(u_{n_j}) \to \Pi_1 S_t^T(u)$ in E^2, then due to (8.9) we have $|\Pi_1 S_t^T(u)| \geqslant \rho_1$. This contradicts (8.4) because $\rho_1 > \rho$. The obtained contradiction proves the theorem. □

8.3. *Examples*

EXAMPLE 8.4. Let us consider the non-linear wave equation

$$\ddot{u} = \Delta u - \tilde{f}(u; t, x), \tag{8.10}$$

where $u = u(t, x)$, $x \in \mathbb{T}^n$. The function $\tilde{f}$ is a polynomial in u of a degree D such that its coefficients are smooth functions of t and x. We set $f = \tilde{f} - u$, denote by B the linear operator $B = \sqrt{1 - \Delta}$ and write (8.10) as the system of two equations:

$$\begin{aligned} \dot{u} &= -Bv, \\ \dot{v} &= Bu + B^{-1} f(u; t, x). \end{aligned} \tag{8.11}$$

Let us take for $\{Z_s\}$ the shifted Sobolev scale $Z_s = H^{s+1/2}(\mathbb{T}^n; \mathbb{R}^2)$, where $\langle \xi, \eta \rangle_s = \int_{\mathbb{T}^n} B^{2s+1} \xi \cdot \eta \, dx$ (its basic scalar product is the scalar product in $H^{1/2}$). We set $\alpha_2 = \bar{J}\, d\xi \wedge d\xi$, where $J\xi = (-v, u)$ for $\xi = (u, v)$. Choosing for $\{\psi_j, j \in \mathbb{N}\}$ a Hilbert basis of the space $H^{1/2}(\mathbb{T}^n)$, formed by properly normalised and enumerated non-zero functions $\sin s \cdot x$ and $\cos s \cdot x$ ($s \in \mathbb{Z}^n$), we set

$$\tilde{\varphi}_j = (\psi_j, 0), \qquad \tilde{\varphi}_{-j} = (0, \psi_j), \quad j \in \mathbb{N}.$$

The obtained symplectic scale $(\{Z_s\}, \alpha_2)$ is a Darboux scale. It is easy to see that (8.11) is a Hamiltonian equation with the Hamiltonian

$$H(u, v) = \frac{1}{2}\big\langle B(u, v), (u, v)\big\rangle_0 + \int F(u; t, x)\, dx,$$

where $F'_u = f$. So $Z_0 = H^{1/2}(\mathbb{T}^n, \mathbb{R}^2)$ is the Darboux phase space for the non-linear wave equation, written in the form (8.11).

To apply Theorem 8.1 we have to check the conditions (H1) and (H2). The first one (with $A = B$) holds trivially since $\tilde{\varphi}_j$'s are eigenfunctions of the Laplacian. The condition (H2) holds in the following three cases:

(a) $n = 1$,
(b) $n = 2$, $D \leqslant 4$,
(c) $n = 3$, $D \leqslant 2$.

The case (a) and the case (b) with $D \leqslant 2$ can be checked using elementary tools, see [54]. Arguments in the case (b) with $3 \leqslant D \leqslant 4$ and in the case (c) are based on a Strichartz-type inequality, see [17].

In the cases (a)–(c), Theorem 8.1 applies to Equation (8.10) in the form (8.11) and shows that the flow maps cannot squeeze $H^{1/2}$-balls to narrow cylinders. This result can be interpreted as impossibility of 'locally uniform' energy transition to high modes, see in [54].

EXAMPLE 8.5. For a non-linear Schrödinger equation

$$\dot{u} = i\Delta u + i f'_u\big(|u|^2\big)u, \quad x \in \mathbb{T}^n \tag{8.12}$$

(cf. Example 2.7), the Darboux phase space is the L_2-space $L_2(\mathbb{T}^n; \mathbb{C})$ with the basis, formed by normalised exponents $\{e^{is\cdot x}, ie^{is\cdot x}\}$. Now the assumption (H2) fails (and it is very unlikely that the flow-maps of (8.12) satisfy the assertions of Lemmas 8.2 and 8.3). So we smooth out the Hamiltonian of (8.12) and replace it by

$$H_\xi = \frac{1}{2}\int \big(|\nabla u|^2 + f\big(|U|^2\big)\big)\,dx, \quad U = u * \xi,$$

where $u * \xi$ is the convolution of u with a function $\xi \in C^\infty(\mathbb{T}^n, \mathbb{R})$. The corresponding Hamiltonian equation is

$$\dot{u} = i\Delta u + i\big(f'\big(|U|^2\big)U\big) * \xi. \tag{8.13}$$

This smoothed equation satisfies (H1), (H2), and Theorem 8.1 applies to its flow-maps.

8.4. *Symplectic capacity*

Another way to prove Theorem 8.1 uses a new object—symplectic capacity —which is interesting on its own.

Symplectic capacity in a Hilbert Darboux space (Z_d, α_2) as in Section 8.2 (below we abbreviate Z_d to Z), is a map c which associates to any open subset $O \subset Z$ a number $c(O) \in [0, \infty]$ and satisfies the following properties:

(1) *Translational invariance*: $c(O) = c(O + \xi)$ for any $\xi \in Z$;
(2) *Monotonicity*: if $O_1 \supset O_2$, then $c(O_1) \geqslant c(O_2)$;
(3) 2-*homogeneity*: $c(\tau O) = \tau^2 c(O)$;
(4) *Normalisation*: for any ball $B_r = B_r(x; Z)$ and any cylinder $C_r^j = C_r^j(Z)$ we have $c(B_r) = c(C_r^j) = \pi r^2$.

(We note that for $x = 0$ the cylinder contains the ball and is 'much bigger', but both sets have the same capacity.)

(5) *Symplectic invariance*: for any symplectomorphism $\Phi : Z \to Z$ and any domain O, $c(\Phi(O)) = c(O)$.

If (Z, α_2) is a finite-dimensional Darboux space, then existence of a capacity with properties (1)–(5) is equivalent to the Gromov theorem. Indeed, if a capacity exists, then the squeezing (8.4) with $\rho < r$ is impossible due to (2), (4) and (5). On the opposite, the quantity

$$\tilde{c}(O) = \sup\left\{\pi r^2 \mid \text{there exists a symplectomorphism which sends } B_r \text{ in } O\right\}$$

obviously satisfies (1)–(3) and (5). Using the Gromov theorem we see that $\tilde{c}$ satisfies (4) as well.

If (Z, α_2) is a Hilbert Darboux space, then the finite-dimensional symplectic capacity, obtained in [43], can be used to construct a capacity c which meets (1)–(4). This capacity turns out to be invariant under symplectomorphisms, which are flow-maps S_t^T as in Theorem 8.1, see [54]. This result also implies Theorem 8.1.

9. The squeezing phenomenon and the essential part of the phase-space

Example 8.4 shows that flow-maps of the non-linear wave equation (8.11) satisfy the Gromov property. This means (more or less) that *flow of generalised solutions for a non-linear wave equation cannot squeeze a ball in a narrow cylinder*. On the contrary, behaviour of the flow formed by *classical* solutions for the non-linear wave equation in sufficiently smooth Sobolev spaces exhibits ‘a lot of squeezing’, at least if we put a small parameter δ in front of the Laplacian. Corresponding results apply to a bigger class of equations. Below we discuss them for non-linear Schrödinger equations; concerning the non-linear wave equation (8.10) see the author’s paper in GAFA 5:4.

Let us consider the non-linear Schrödinger equation:

$$\dot{u} = -i\delta\Delta u + i|u|^{2p}u, \tag{9.1}$$

where $\delta > 0$ and $p \in \mathbb{N}$, supplemented by the odd periodic boundary conditions:

$$\begin{aligned} u(t,x) &= u(t, x_1, \ldots, x_j + 2\pi, \ldots, x_n) \\ &= -u(t, x_1, \ldots, -x_j, \ldots, x_n), \quad j = 1, \ldots, n, \end{aligned} \tag{9.2}$$

where $n \leqslant 3$. Clearly, any function which satisfies (9.2) vanishes at the boundary of the cube K^n of half-periods, $K^n = \{0 \leqslant x_j \leqslant \pi\}$. The problem (9.1), (9.2) can be written in the Hamiltonian form (2.2) if for the symplectic Hilbert scale $(\{X_s\}, \alpha_2)$ one takes the scale formed by odd periodic complex Sobolev functions, $X_s = H^s_{\text{odd}}(\mathbb{R}^n/2\pi\mathbb{Z}^n; \mathbb{C})$, and $\alpha_2 = i\, du \wedge du$ (cf. Example 2.8).

Due to a non-trivial result of Bourgain (which can be extracted from [15]), flow-maps S^t for (9.1), (9.2) are well defined in the spaces X_s, $s \geqslant 1$. In particular, they are well defined in the space C^∞ of smooth odd periodic functions. Denoting by $|\cdot|_m$ the

C^m-norm, $|u|_m = \sup_{|\alpha|=m} \sup_x |\partial_x^\alpha u(x)|$, we define below the set $\mathfrak{A}_m \subset C^\infty$ which we call the essential part of the smooth phase-space for the problem (9.1), (9.2) with respect to the C^m-norm, or just the *essential part of the phase-space*:

$$\mathfrak{A}_m = \big\{u \in C^\infty \mid u \text{ satisfies (9.2) and the condition (9.3)}\big\},$$

where

$$|u|_0 \leqslant K_m \delta^\mu |u|_m^{1/(2pm\varkappa+1)}, \tag{9.3}$$

with a suitable $K_m = K_m(\varkappa)$ and $\mu = m\varkappa/(2pm\varkappa + 1)$. Here $\varkappa$ is any fixed constant $\varkappa \in (0, 1/3)$.

Intersection of the set $\mathfrak{A}_m$ with the R-sphere in the C^m-norm (i.e., with the set $\{|u|_m = R\}$) has the C^0-diameter $\leqslant 2K_m\delta^\mu R^{1/(2pm\varkappa+1)}$. Asymptotically (as $\delta \to 0$ or $R \to \infty$) this is much smaller than the C^0-diameter of the sphere, which equals $C_m R$. Thus, $\mathfrak{A}_m$ is an 'asymptotically narrow' subset of the smooth phase space.

The theorem below states that for any $m \geqslant 2$ the set $\mathfrak{A}_m$ is a recursion subset for the dynamical system, and gives a control for the recursion time:

THEOREM 9.1. *Let $u(t) = u(t, \cdot)$ be a smooth solution for* (9.1), (9.2) *and $|u(t_0)|_0 = U$. Then there exists $T \leqslant t_0 + \delta^{-1/3}U^{-4p/3}$ such that $u(T) \in \mathfrak{A}_m$ and $\frac{1}{2}U \leqslant |u(T)|_0 \leqslant \frac{3}{2}U$.*

Since L_2-norm of a solution is an integral of motion (see Example 3.5) and $|u(t)|_0 \geqslant |u(t)|_{L_2(K^n)}$, then we obtain the following

COROLLARY 9.2. *Let $u(t)$ be a smooth solution for* (9.1), (9.2) *and $|u(t)|_{L_2(K^n)} \equiv W$. Then for any $m \geqslant 2$ this solution cannot stay outside $\mathfrak{A}_m$ longer than the time $\delta^{-1/3}W^{-4p/3}$.*

For the theorem's proof we refer the reader to Appendix 3 in [58]. Here we explain why 'something like this result' should be true. Presenting the arguments it is more convenient to operate with the Sobolev norms $\|\cdot\|_m$. Let us denote $\|u(t_0)\|_0 = A$. Arguing by contradiction, we assume that for all $t \in [t_0, t_1] = L$, where $t_1 = t_0 + \delta^{-1/3}U^{-4p/3}$, we have

$$C\delta^a \|u\|_m^b < \|u\|_0, \tag{9.4}$$

where $m \geqslant 3$ is a fixed number. Since $\|u(t)\|_0 \equiv A$, then (9.4) and the interpolation inequality imply the upper bounds

$$\big\|u(t)\big\|_l \leqslant C_l A^{1-\frac{l}{m}+\frac{l}{mb}} \delta^{-\frac{la}{mb}}, \quad 0 \leqslant l \leqslant m,\ t \in L. \tag{9.5}$$

In particular, $\delta\|\Delta u\|_1 \leqslant C_3 A^{1-\frac{3}{m}+\frac{3}{mb}} \delta^{1-\frac{3a}{mb}}$. Therefore if $mb > 3a$, then for $t \in L$ Equation (9.1), treated as a dynamical system in H^1_{odd}, is a perturbation of the trivial equation

$$\dot{u} = i|u|^{2p}u. \tag{9.6}$$

Elementary arguments show that the H^1-norm of each non-zero solution for (9.6) grows linearly with time. This implies a lower bound for $\sup_{t\in L}\|u(t)\|_1$, where $u(t)$ is the solution for (9.1), (9.2) which we discuss. It turns out that one can choose a and b in such a way that $mb > 3a$ and the lower bound we have just obtained contradicts (9.5) with $l = 1$. This contradiction shows that (9.4) cannot be true for all $t \in L$. In other words, $\|u(\tau)\|_0 \leqslant C\delta^a\|u(\tau)\|_m^b$ for some $\tau \in L$. At this moment τ the solution enters a domain, similar to the essential part $\mathfrak{A}_m$.

Let us consider any trajectory $u(t)$ for (9.1), (9.2) such that $|u(t)|_{L_2(K^n)} \equiv W \sim 1$, and discuss the time-averages $\langle |u|_m\rangle$ and $\langle \|u\|_m^2\rangle^{1/2}$ of its C^m-norm $|u|_m$ and its Sobolev norm $\|u\|_m$, where we set

$$\big\langle |u|_m\big\rangle = \frac{1}{T}\int_0^T |u|_m\,dt, \qquad \big\langle \|u\|_m^2\big\rangle^{1/2} = \left(\frac{1}{T}\int_0^T \|u\|^2\,dt\right)^{1/2},$$

and the time T of averaging is specified below. While the trajectory stays in $\mathfrak{A}_m$, we have

$$|u|_m \geqslant \big(WK_m^{-1}\delta^{-\mu}\big)^{1/(1-2p\mu)}.$$

One can show that this inequality implies that each visit to $\mathfrak{A}_m$ increases the integral $\int |u|_m\,dt$ by a term bigger than δ to a negative degree. Since these visits are sufficiently frequent by the corollary, then we obtain a lower estimate for the quantity $\langle |u|_m\rangle$. Details can be found in [55]. Here we present a better result which estimates the time-averaged Sobolev norms. For a proof see Section 4.1 of [58].

THEOREM 9.3. *Let $u(t)$ be a smooth solution for Equation* (9.1), (9.2) *such that $|u(t)|_{L_2(K^n)} \geqslant 1$. Then there exists a sequence $k_m \nearrow 1/3$ and constants $C_m > 0$, $\delta_m > 0$ such that $\langle \|u\|_m^2\rangle^{1/2} \geqslant C_m\delta^{-2mk_m}$, provided that $m \geqslant 4$, $\delta \leqslant \delta_m$ and $T \geqslant \delta^{-1/3}$.*

The results stated in Theorems 9.1, 9.3 remain true for Equations (9.1) with dissipation. I.e., for the equations with δ replaced by $\delta\nu$, where ν is a unit complex number such that $\operatorname{Re}\nu \geqslant 0$ and $\operatorname{Im}\nu \geqslant 0$.[11] If $\operatorname{Im}\nu > 0$, then smooth solutions for (9.1), (9.2) converge to zero in any C^m-norm. Since the essential part $\mathfrak{A}_m$ clearly contains a sufficiently small C^m-neighbourhood of zero, then eventually any smooth solution enter $\mathfrak{A}_m$ and stays there forever. Theorem 9.3 states that the solution will visit the essential part much earlier, before its norm decays. Moreover, results, similar to Theorem 9.3, are true for solutions of the damped-driven equation $\dot u + \delta\Delta u - i|u|^2u = \eta(t,x)$, where the force η is a random field, smooth in x, and stationary mixing in t. See [56] and [58].

Acknowledgements

I thank for the hospitality FIM (ETH, Zürich), where this paper was completed. The research was supported by EPSRC, grants GR/N63055/01 and GR/S68712/01.

[11] The only correction is that if $\operatorname{Im}\nu > 0$, then in Theorem 9.3 one should take $T = \delta^{-1/3}$.

Appendix. Families of periodic orbits in reversible PDEs, by D. Bambusi

A.1. *Introduction*

Some families of periodic solutions of PDEs can be constructed using KAM theory; however a different approach leading to stronger results and simpler proofs is available. It is based on the Lyapunov–Schmidt decomposition combined with a suitable analysis of small denominators. The main advantage of this approach is elimination of the second Melnikov condition (see (5.7)). As a consequence it is applicable to problems with periodic boundary conditions and to some equations in more than one space dimension. Most of the general theory has been developed for equations that are of second order in time and we will mainly deal with this case. Moreover, we will concentrate on problems involving small denominators and only briefly report on results of a different kind.

A.2. *An abstract theorem for non-resonant PDEs*

Let $\{X_s\}$ be a scale of Hilbert spaces with norms $\|\cdot\|_s$ and scalar product $\langle\cdot;\cdot\rangle_s$. Let A be a (linear) morphism of the scale, and assume that there exists a Hilbert basis $\{\varphi_j\}_{j=1}^{\infty}$ of X_0 such that

$$A\varphi_j = \omega_j^2\varphi_j, \quad \omega_j > 0.$$

Let us fix s, consider a neighbourhood $\mathcal{U}$ of the origin in X_s and a smooth map $g:\mathcal{U}\to X_s$, having at the origin a zero of second order. We are interested in families of small amplitude periodic solutions of the equation

$$\ddot{x} + Ax = g(x). \tag{A.1}$$

EXAMPLE A.1. The non-linear wave equation with periodic boundary conditions:

$$w_{tt} - w_{xx} + V(x)w = f(x,w), \tag{A.2}$$

$$w(x,t) = w(x+2\pi,t), \qquad w_x(x,t) = w_x(x+2\pi,t), \tag{A.3}$$

where the potential V and the non-linearity f are smooth periodic of period 2π in x, and $f(x,w) = \mathrm{O}(|w|^2)$. Let λ_j be the periodic eigenvalues of the Sturm–Liouville operator $-\partial_{xx} + V(x)$ and assume $\lambda_j > 0\ \forall j$. Then the frequencies are $\omega_j := \sqrt{\lambda_j}$. In this case $X_s = H^s(\mathbb{T})$, and f induces a smooth operator from X_s to itself, provided that $s > 1/2$.

EXAMPLE A.2. The non-linear plate equation in the d-dimensional cube:

$$w_{tt} + \Delta\Delta w + aw = f(w), \quad x\in\mathcal{Q}, \tag{A.4}$$

$$w|_{\partial\mathcal{Q}} = \Delta w|_{\partial\mathcal{Q}} = 0, \tag{A.5}$$

where $a > 0$, $f(w) = O(|w|^3)$ and

$$\mathcal{Q} := \left\{ x = (x_1, \dots, x_d) \in \mathbb{R}^d \colon 0 < x_i < \pi \right\}.$$

Then the eigenfunctions of the linearised system are given by

$$\varphi_n = \sin(n_1 x_1) \sin(n_2 x_2) \cdots \sin(n_d x_d)$$

and the corresponding frequencies are $\omega_n = \sqrt{(n_1^2 + \cdots + n_d^2)^2 + a}$, where $n \in \mathbb{Z}^d$ and $n_i \geqslant 1$. To fit the abstract scheme we order the basis in such a way that the frequencies are in non-decreasing order. Now $X_0 = L^2(\mathcal{Q})$, and $X_s = D((\Delta\Delta)^s) \subset H^{4s}$ endowed with the graph norm. If the non-linearity f is smooth and odd (i.e. $f(-w) = -f(w)$), then it defines a smooth map from X_s to itself for any $s > [d/2]/4$ (see Example 2.5).

In the linear approximation ($g \equiv 0$) the general solution of (A.1) is the superposition of the linear normal modes, i.e. of the families of periodic solutions

$$x^{(j)}(t) = \big(a_j \cos(\omega_j t) + b_j \sin(\omega_j t)\big)\varphi_j. \tag{A.6}$$

Fix one of the families, say $x^{(1)}$. To ensure its persistence in the non-linear problem we make the following assumptions:

(H1) (Non-resonance) For small enough $\gamma > 0$ there exists a closed set $W_\gamma \subset \mathbb{R}^+$ having ω_1 as an accumulation point both from the right and from the left, and such that for any $\omega \in W_\gamma$ one has

$$|\omega l - \omega_j| \geqslant \frac{\gamma}{l}, \quad \forall l \geqslant 1,\ \forall j \geqslant 2. \tag{A.7}$$

(H2) (Non-degeneracy) Let $g_r(x)$ be the first non-vanishing (homogeneous) Taylor polynomial of g. Assume that $r \geqslant 3$ and $\beta_0 \neq 0$, where

$$\beta_0 := \begin{cases} \langle g_r(\varphi_1), \varphi_1 \rangle_0 & \text{if } r \text{ is odd,} \\ \langle g_{r+1}(\varphi_1), \varphi_1 \rangle_0 & \text{if } r \text{ is even.} \end{cases} \tag{A.8}$$

Denoting $\xi_1(\omega_1 t) = \cos(\omega_1 t)\varphi_1$ one has

THEOREM A.3. *Suppose that assumptions* (H1), (H2) *hold. Then there exist a set $\mathcal{E} \subset \mathbb{R}$ having zero as an accumulation point, a positive ω_*, and a family of periodic solutions $\{x_\varepsilon(t)\}_{\varepsilon \in \mathcal{E}}$ of* (A.1) *with frequencies $\{\omega^\varepsilon\}_{\varepsilon \in \mathcal{E}}$ fulfilling*

$$\sup_t \left\| x_\varepsilon(t) - \varepsilon \xi_1(t\omega^\varepsilon) \right\|_s \leqslant C\varepsilon^r, \quad \left|\omega^\varepsilon - \omega_1\right| \leqslant C\varepsilon^{r-1}. \tag{A.9}$$

Moreover, the set $\mathcal{E}$ is in one to one correspondence either with $W_\gamma \cap [\omega_1, \omega_1 + \omega_*)$ if $\beta_0 < 0$, or with $W_\gamma \cap (\omega_1 - \omega_*, \omega_1]$ if $\beta_0 > 0$.

PROOF. We consider only the case of odd r, the general case can be obtained by a slightly different treatment of the forthcoming equation ω. We are looking for an X_s-valued function $q(t)$ which is 2π-periodic and reversible (i.e. $q(t) = q(-t)$), and for a positive ω, close to ω_1, such that $q(\omega t)$ is a solution of (A.1). They must satisfy the equation

$$L_\omega q = g(q), \quad L_\omega := \omega^2 \frac{d^2}{dt^2} + A, \tag{A.10}$$

which will be considered as an ω-dependent functional equation in the space $\mathcal{H} \subset H^1(\mathbb{T}, X_s)$, formed by the reversible periodic functions. Equation (A.10) is studied using the Lyapunov–Schmidt decomposition, namely by decomposing it into an equation on $\operatorname{Ker} L_{\omega_1} \equiv \operatorname{span}(\xi_1)$ and an equation on its orthogonal complement R. Precisely, denote by Q the projector on ξ_1 and by P the projector on R and make the Ansatz $q = \varepsilon\xi_1 + \varepsilon^r u$, where $u \in R$. Then (A.10) is equivalent to the system

$$\omega^2 = \omega_1^2 + \beta\varepsilon^{r-1}, \tag{A.11}$$

$$L_\omega u = P g_r(\xi_1) + PG(\varepsilon, u), \tag{A.12}$$

$$-\beta\xi_1 = Q g_r(\xi_1) + QG(\varepsilon, u) \tag{A.13}$$

for the unknowns (ε, u, β). Here G contains all higher-order corrections and $\omega \in W_\gamma$ is a parameter. Equations (A.11), (A.12) and (A.13) are called the ω, the P and the Q equation, respectively.

First one solves the P equation (A.12). To this end one has to invert the linear operator $L_\omega|_R$. Its eigenfunctions are $\cos(lt)\varphi_j$, and the corresponding eigenvalues are

$$\lambda_{jl} = -l^2\omega^2 + \omega_j^2 = (l\omega + \omega_j)(\omega_j - l\omega), \quad j \geqslant 2,\ l \geqslant 1.$$

By (A.7), $|\lambda_{jl}| > C\gamma$. So $(L_\omega|_R)^{-1}$ exists and is bounded. Applying this operator to the P equation and using the implicit function theorem one obtains a smooth function $u(\varepsilon)$ that depends parametrically on $\omega \in W_\gamma$ and solves the P equation.

Inserting $u(\varepsilon)$ in the Q equation one determines the parameter β as a function of ε. In particular one has $\beta(\varepsilon) = C\beta_0 +$ higher-order corrections, where $C > 0$. Inserting $\beta(\varepsilon)$ in the ω equation one gets an equation for ε (remember that ω is fixed), which is a perturbation of the equation $\omega^2 - \omega_1^2 = C\beta_0\varepsilon^{r-1}$. By the non-degeneracy this can be reduced to a fixed point equation for ε^{r-1} which is solvable by the contraction mapping principle. □

REMARK A.4. The theorem holds also in the case $r = 2$, but in this case the non-degeneracy condition takes a more complicated form.

Theorem A.3 was proved in [5]. The technique of the Lyapunov–Schmidt decomposition was used for the first time to construct families of periodic solutions in PDEs by Craig and Wayne [29] who considered the model problem of the wave equation with periodic boundary conditions (see Example A.1); we will report on this work in Section A.4.

EXAMPLE A.5. Consider the non-linear wave equation with periodic boundary conditions (see Example A.1). Let ω_1 be such that $\omega_1 \neq \omega_j$ for each $j \neq 1$. Decompose V into its average a and a part $\widetilde{V}$ of zero average, then condition (H1) is satisfied if a belongs to an uncountable set which is dense in a neighbourhood of the origin (for the proof see Lemma 3.1 of [7]). Condition (H2) can be expressed in terms of the eigenfunctions of the Sturm–Liouville operator. If it holds, then Theorem A.3 applies and ensures persistence of the corresponding family of periodic orbits. Note that, in a difference with the case of Dirichlet boundary conditions (see Example 5.3), the non-linearity does not need to have some particular parity.

EXAMPLE A.6. Consider the non-linear plate equation (see Example A.2). In the case $d = 1$ all the frequencies are simple and the assumption (H1) is satisfied if a is chosen in a subset of $\mathbb{R}^+$ having full measure. In the case $d > 1$, all the frequencies are multiple except the smallest one. Taking for ω_1 the smallest frequency, (H1) is fulfilled if a belongs to a dense uncountable subset of $[0, 1/4]$. (H2) holds trivially provided the Taylor expansion of f at zero does not vanish identically (remember that $f(-w) = f(w)$). Then Theorem A.3 ensures persistence of the corresponding family of periodic orbits (for details see [7]).

A.3. *The resonant case*

It is possible to generalise the above theorem to the case when the frequencies satisfy some resonance relations. We will consider only the Lagrangian case, when $g = -\nabla H$.

Fix a frequency ω_1 of the linearised system. We replace the assumption (H1) by the following one:

(H1R) For any small enough γ there exists a closed set $W_\gamma \subset \mathbb{R}^+$ having ω_1 as an accumulation point both from the right and from the left, and such that for any $\omega \in W_\gamma$ one has

$$\text{either} \quad |\omega l - \omega_j| \geqslant \frac{\gamma}{l}, \quad \text{or} \quad l\omega_1 - \omega_j = 0. \tag{A.14}$$

To pass to the non-degeneracy assumption, we define the resonant set as

$$\mathcal{I}_R := \{k \geqslant 1 \colon \exists l \geqslant 1 \colon l\omega_1 - \omega_k = 0\}, \tag{A.15}$$

consider the linear space generated by $\{\varphi_k\}_{k \in \mathcal{I}_R}$, and denote by $\mathcal{N}$ its closure in the graph norm of $D(A)$. Note that all solutions of the linearised system with initial datum in $\mathcal{N}$ and vanishing initial velocity are periodic of period $2\pi/\omega_1$. Let H_r be the first non-vanishing Taylor coefficient of H. For $x \in \mathcal{N}$ define the average of H_r by

$$\langle H_r \rangle(x) := \frac{\omega_1}{2\pi} \int_0^{2\pi/\omega_1} H_r\big(\cos(At)x\big)\, dt.$$

Consider the hypersurface $\mathcal{S} \subset \mathcal{N}$ of the points $x \in \mathcal{N}$ such that $\langle x; Ax \rangle_0 = 1$.

(H2R) There exists a non-degenerate critical point x_0 of the functional $\langle H_r \rangle|_{\mathcal{S}}$. The corresponding Lagrange multiplier β_0 does not vanish.

Denote by $\xi_0(\omega_1 t)$ the solution of the linearised system with initial datum x_0 and vanishing initial velocity.

THEOREM A.7 [6]. *Suppose the assumptions* (H1R), (H2R) *hold. Then there exists a family of periodic solutions* $\{x_\varepsilon(t)\}_{\varepsilon\in\mathcal{E}}$ *of* (A.1) *with frequencies* ω^ε, *satisfying*

$$\sup_t \left\| x_\varepsilon(t) - \varepsilon\xi_0(t\omega^\varepsilon)\right\|_s \leqslant C\varepsilon^r, \quad \left|\omega^\varepsilon - \omega_1\right| \leqslant C\varepsilon^{r-1}. \tag{A.16}$$

The set $\mathcal{E}$ *has the same properties as in the non-resonant case.*

The proof is obtained by proceeding as in the non-resonant case. The only difference is that in this case the kernel of L_{ω_1} is no longer one-dimensional, but is isomorphic to $\mathcal{N}$ (the isomorphism being given by the map $x \mapsto \cos(At/\omega_1)x$). So the Q equation can be transformed into an equation in $\mathcal{N}$. The latter turns out to be a perturbation of the equation for the critical points of $\langle H_r \rangle|_{\mathcal{S}}$, and the non-degeneracy condition (H2R) allows to solve it by the implicit function theorem.

Applying the above theorem, one can construct countably many families of periodic solutions of the ϕ^4-model

$$w_{tt} - w_{xx} = \pm w^3 + \text{higher-order terms}$$

with Dirichlet boundary conditions, and also higher frequency periodic solutions of the non-linear plate equation of Example A.2 (see [6,7], see also [62,21]).

In general it is difficult to check condition (H2R). In the case of Hamiltonian systems with $n < \infty$ degrees of freedom, topological arguments allow to avoid it. Indeed, the Weinstein–Moser theorem (see [80,64]) ensures that close to a minimum of the energy, on each surface of a constant energy there exist at least n periodic orbit. In general they do not form regular families. A corresponding result for PDEs is not available at present. However there exists an *ad hoc* variational result for the wave equation

$$w_{tt} - w_{xx} = \pm w^p + \text{higher-order terms}, \quad p \geqslant 2, \tag{A.17}$$

which ensures that, having fixed $j \geqslant 1$, there exists a sequence of periodic orbits accumulating at zero, whose frequencies accumulate at j (which plays here the role of the jth linear frequency). The corresponding theorem is due to Berti and Bolle [10].

Periodic solutions in the non-linear wave equation

$$w_{tt} - w_{xx} + f(x, w) = 0, \qquad u(0,t) = u(\pi,t) = 0, \tag{A.18}$$

where constructed for the first time by Rabinowitz [76] using global variational methods and a Lyapunov–Schmidt decomposition. Rabinowitz proved that, under suitable assumptions on f, Equation (A.18) has at least one periodic solution with period $T = 2\pi p/q$, for any choice of the integers p and q. Note that, when the period T is commensurable with

2π, the operator $L_\omega|_R$ has a compact inverse, i.e. there are no small denominators. The work [76] was followed by a series of papers, simplifying the proof and sharpening the result (see [24] and references therein). In particular, we mention the paper [25] by Brezis, Coron and Nirenberg, where existence of periodic orbits is proved by a particularly simple method: the authors write a variational principle, dual to the usual one, and look for its critical points, using the mountain pass lemma. It is remarkable that in this approach the Q equation becomes trivial.

A.4. *Weakening the non-resonance condition*

The main limitation of the results presented in Sections A.2 and A.3 rests in the non-resonance conditions (H1) and (H1R). Indeed, such conditions are fulfilled with large probability (in a suitable parameter space) when $\omega_j \sim j^\nu$ with $\nu > 1$; when $\nu = 1$ the non-resonance conditions are satisfied typically on uncountable sets of zero measure, but when $\nu < 1$ they are satisfied only exceptionally (as in the plate equation). As a consequence the results of Sections A.2 and A.3 are not applicable to general equations in more than one space dimensions. Furthermore, the method of Lyapunov–Schmidt decomposition can be extended to the case of reversible systems of first order in time, but the approach of Section A.2 is no more applicable.

In order to avoid such limitations one would like to be able to work with the weaker non-resonance condition "there exists a $\tau > 0$ such that $|l\omega - \omega_j| \geqslant \gamma/l^\tau$". This was done by Craig and Wayne [29] who used the Nash–Moser theorem to solve the P equation. The application of the Nash–Moser theorem requires to construct and estimate the inverse of the linear operator describing the linearisation of the P equation at an approximate solution. This is the main difficulty of Craig–Wayne's approach. To overcome it they use the techniques by Fröhlich and Spencer [38], performing a careful analysis of small denominators (cf. Section 5.3). The method by Craig and Wayne was extended by Bourgain in order to construct periodic (and also quasiperiodic) solutions in higher-dimensional equations. The resulting method seems very general, but at present a theorem "ready for application" is not available. We present here the result obtained by Bourgain by applying this method to the non-linear wave equation

$$w_{tt} - \Delta w + aw + w^3 = 0 \tag{A.19}$$

on $\mathbb{T}^d$. Fix a multiindex $n \in \mathbb{Z}^d$ different from zero, and let

$$\xi_n(\omega_n t, x) := \cos(n \cdot x + \omega_n t), \quad \omega_n := \sqrt{n_1^2 + \cdots + n_d^2 + a},$$

be the corresponding symmetric reversible solution.

THEOREM A.8 [18]. *If a belongs to a certain subset of $\mathbb{R}^+$ of full measure, then there exists a Cantor set $\mathcal{E}$ of positive measure, accumulating at zero, and a family of periodic solutions $\{w_\varepsilon(t,x)\}_{\varepsilon\in\mathcal{E}}$ of* (A.19) *with frequencies ω^ε, satisfying*

$$\left|\varepsilon\xi_n(\omega^\varepsilon t, x) - w_\varepsilon(t,x)\right| \leqslant C\varepsilon^3, \quad \left|\omega_n - \omega^\varepsilon\right| \leqslant C\varepsilon^2.$$

In the case $d = 1$, the result was proved in [29]; subsequently, still in the case $d = 1$, Kuksin introduced a simpler technique to find the “large measure result” of Theorem A.8 (see in [20, pp. 90–94]).

The Craig–Wayne–Bourgain method also allows to deal with first order in time equations. For example, it was applied to the Schrödinger equation in one [30] or two space dimensions [19] (see Section 5.4).

A.5. *The water wave problem*

A particular problem that has attracted the attention of many researchers since the very beginning of the theory of PDEs is that of existence of standing water waves. The first rigorous proof of their existence was obtained only recently by Plotnikov and Toland [70]; we present here their result.

Consider a perfect fluid lying above a horizontal bottom, and confined between two parallel vertical walls. The fluid is subject to gravity, and atmospheric pressure acts at the free surface. This is a dynamical system governed by the Euler equations supplemented by appropriate boundary conditions. It was pointed out by Zakharov that this system is Hamiltonian (see [81]). The corresponding Hamiltonian function is the energy of the fluid, and conjugated variables are given by the wave profile and the velocity potential at the free surface.

In the linear approximation the general solution is given by the superposition of the normal modes. The problem is to continue the normal modes to families of periodic solutions of the non-linear system (the standing waves). Fix one of the normal modes, and denote by $\eta(t, x_1)$ the corresponding profile of the free surface (x_1 being the horizontal variable). Then it is possible to choose the depth h, the width l of the region occupied by the fluid and the gravitational constant g in such a way that the period of the solution is normalised to 2π and the linear frequencies fulfil a suitable non-resonance condition. Denote by (g_0, l_0, h_0) a choice of the parameters realising such conditions, then one has

THEOREM A.9 [70]. *There exists an infinite set $\mathcal{E} \subset \mathbb{R}$ having zero as an accumulation point and, for any $\varepsilon \in \mathcal{E}$, there exist g_ε, l_ε and a standing wave solution of the water wave problem with gravity g_ε in a box of width l_ε. Moreover, denoting by η_ε the corresponding profile of the free surface, one has*

$$\left|\eta_\varepsilon(t, x_1) - \varepsilon^2 \eta(t, x_1)\right| < C\varepsilon^3, \qquad |g_\varepsilon - g_0| + |l_\varepsilon - l_0| \leqslant C\varepsilon.$$

The main difficulties in proving this result are as follows: firstly, the linear frequencies behave as $\omega_n \sim n^{1/2}$, so the non-resonance conditions that can be satisfied are quite weak. Secondly, the mathematical formulation of the problem involves an unbounded non-linear and non-local operator. To overcome these difficulties, Plotnikov and Toland use the Lagrangian description of the fluid motion and apply the Lyapunov–Schmidt approach to handle the resulting non-linear problem. The P equation now is solved by means of the Nash–Moser theorem. The required invertibility of the linearised operator is obtained in two steps: first it is reduced to a suitable canonical form, and next this canonical form (which is essentially a perturbation of an operator involving derivatives and Hilbert transform) is studied in detail.

References

[1] V.I. Arnold, *Proof of a theorem of A.N. Kolmogorov on the conservation of quasiperiodic motions under a small change of the Hamiltonian function*, Russian Math. Surveys **18** (5) (1963), 9–36.

[2] V.I. Arnold, *Mathematical Methods in Classical Mechanics*, 3rd edn, Springer, Berlin (1989).

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

[4] D. Bambusi, *On long time stability in Hamiltonian perturbations of non-resonant linear PDEs*, Nonlinearity **12** (1999), 823–850.

[5] D. Bambusi, *Lyapunov center theorem for some nonlinear PDEs: A simple proof*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **29** (2000), 823–837.

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

[7] D. Bambusi and S. Paleari, *Families of periodic orbits for some PDE's in higher dimensions*, Comm. Pure Appl. Anal. **1** (2002), 269–279.

[8] E.D. Belokolos, A.I. Bobenko, V.Z. Enolskii, A.R. Its and V.B. Matveev, *Algebro-Geometric Approach to Nonlinear Integrable Equations*, Springer, Berlin (1994).

[9] G. Benettin, F. Fasso and M. Guzzo, *Nekhoroshev stability of elliptic equilibra of Hamiltonian systems*, Comm. Math. Phys. **197** (1998), 347–360.

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

[11] R.F. Bikbaev and S.B. Kuksin, *A periodic boundary-value problem for the Sine-Gordon equation, small Hamiltonian perturbations of it, and KAM-deformations of finite-gap tori*, St.-Petersburg Math. J. **4** (1993), 439–468.

[12] A.I. Bobenko and S.B. Kuksin, *Finite-gap periodic solutions of the KdV equation are nondegenerate*, Phys. Lett. A **161** (3) (1991), 274–276.

[13] A.I. Bobenko and S.B. Kuksin, *The nonlinear Klein–Gordon equation on an interval as a perturbed Sine-Gordon equation*, Comment. Math. Helv. **70** (1995), 63–112.

[14] A.I. Bobenko and S.B. Kuksin, *Small-amplitude solutions of the Sine-Gordon equation on an interval under Dirichlet or Neumann boundary conditions*, J. Nonlinear Sci. **5** (1995), 207–232.

[15] J. Bourgain, *Fourier transform restriction phenomenon for certain lattice subsets and applications to nonlinear evolution equations*, Geom. Funct. Anal. **3** (1993), 107–156 and 209–262.

[16] J. Bourgain, *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices (1994), 475–497.

[17] J. Bourgain, *Aspects of long time behaviour of solutions of nonlinear Hamiltonian evolution equations*, Geom. Funct. Anal. **5** (1995), 105–140.

[18] J. Bourgain, *Construction of periodic solutions of nonlinear wave equations in higher dimension*, Geom. Funct. Anal. **5** (1995), 629–639.

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

[20] J. Bourgain, *Nonlinear Schrödinger Equations*, Hyperbolic Equations and Frequency Interactions, Amer. Math. Soc. (1999).

[21] J. Bourgain, *Periodic solutions of nonlinear wave equations*, Harmonic Analysis and Partial Differential Equations, Chicago University Press (1999), 69–97.

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

[23] J. Bourgain, *Green's Function Estimates for Lattice Schrödinger Operators and Applications*, Ann. Math. Studies, Princeton University Press, Princeton (2004).

[24] H. Brezis, *Periodic solutions of nonlinear vibrating string and duality principle*, Bull. Amer. Math. Soc. **8** (1983), 409–426.

[25] H. Brezis, J. Coron and L. Nirenberg, *Free vibrations for a nonlinear wave equation and a theorem by P. Rabinowitz*, Comm. Pure Appl. Math. **33** (1980), 667–689.

[26] D. Cai, D.W. McLaughlin and K.T.R. McLaughlin, *The nonlinear Schrödinger equation as both a PDE and a dynamical system*, Handbook of Dynamical Systems, Vol. 2, B. Fiedler, ed., Elsevier, Amsterdam (2002).

[27] L. Chierchia and J. You, *KAM tori for 1D nonlinear wave equations with periodic boundary conditions*, Comm. Math. Phys. **211** (2000), 497–525.

[28] W. Craig, *Problèmes de petitts diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthéses, no. 9, Société Mathématique de France (2000).

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

[30] W. Craig and C.E. Wayne, *Periodic solutions of nonlinear Schrödinger equations and the Nash–Moser method*, Hamiltonian Mechanics. Integrability and Chaotic Behavior, NATO ASI, Vol. B331, Plenum Press (1994), 103–122.

[31] B.A. Dubrovin, *Theta-functions and nonlinear equations*, Russian Math. Surveys **36** (2) (1981), 11–80.

[32] B.A. Dubrovin, V.B. Matveev and S.P. Novikov, *Nonlinear equations of Korteweg–de Vries type, finite zone linear operators, and Abelian varieties*, Russian Math. Surveys **31** (1) (1976), 55–135.

[33] L.H. Eliasson, *Perturbations of stable invariant tori*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. IV Ser. **15** (1988), 115–147.

[34] L.H. Eliasson and S.B. Kuksin, *KAM for NLS* (2005), Preprint.

[35] L.H. Eliasson, S.B. Kuksin, S.Marmi and J.-C. Yoccoz, *Dynamical systems and small divisors*, Lecture Notes in Math., Vol. 1784, ch. KAM-persistence of finite-gap solutions, Springer, Berlin (2002).

[36] E. Fermi, J.R. Pasta and S.M. Ulam, *Studies of nonlinear problems*, Collected works of E. Fermi, Vol. 2, Chicago University Press, Chicago (1965).

[37] L. Friedlander, *An invariant measure for the equation* $u_{tt} - u_{xx} + u^3 = 0$, Comm. Math. Phys. **98** (1985), 1–16.

[38] J. Fröhlich and T. Spencer, *Absence of diffusion in Anderson tight binding model for large disorder or low energy*, Comm. Math. Phys. **88** (1983), 151–184.

[39] J. Fröhlich, T. Spencer and C.E. Wayne, *Localization in disordered nonlinear dynamical systems*, J. Statist. Phys. **42** (1986), 247–274.

[40] G.E. Giacaglia, *Perturbation Methods in Non-Linear Systems*, Springer, Berlin (1972).

[41] B. Grébert and T. Kappeler, *Perturbation of the defocusing nonlinear Schrödinger equation*, Milan J. Math. **71** (2003), 141–174.

[42] M. Gromov, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), 307–347.

[43] H. Hofer and E. Zehnder, *Symplectic Invariants and Hamiltonian Dynamics*, Birkhäuser, Basel (1994).

[44] T. Kappeler, *Fibration of the phase-space for the Korteweg–de Vries equation*, Ann. Inst. Fourier **41** (1991), 539–575.

[45] T. Kappeler and M. Makarov, *On Birkhoff coordinates for KdV*, Ann. H. Poincaré **2** (2001), 807–856.

[46] T. Kappeler and J. Pöschel, *KAM & KdV*, Springer (2003).

[47] T. Kato, *Quasi-linear equations of evolutions, with applications to partial differential equations*, Lecture Notes in Math., Vol. 448, Springer, Berlin (1975), 25–70.

[48] I.M. Krichever and D.H. Phong, *Symplectic forms in the theory of solitons*, Surv. Differ. Geom., Vol. IV, Int. Press, Boston (1998), 239–313.

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

[50] S.B. Kuksin, *Perturbations of quasiperiodic solutions of infinite-dimensional Hamiltonian Systems*, Izv. Akad. Nauk SSSR Ser. Mat. **52** (1988), 41–63: English transl. in Math. USSR-Izv. **32** (1) (1989).

[51] S.B. Kuksin, *The perturbation theory for the quasiperiodic solutions of infinite-dimensional Hamiltonian systems and its applications to the Korteweg–de Vries equation*, Math. USSR-Sb. **64** (1989), 397–413.

[52] S.B. Kuksin, *Nearly Integrable Infinite-Dimensional Hamiltonian Systems*, Springer, Berlin (1993).

[53] S.B. Kuksin, *KAM-theory for partial differential equations*, Proceedings of the First European Congress of Mathematics, Vol. 2, Birkhäuser (1994), 123–157.

[54] S.B. Kuksin, *Infinite-dimensional symplectic capacities and a squeezing theorem for Hamiltonian PDEs*, Comm. Math. Phys. **167** (1995), 531–552.

[55] S.B. Kuksin, *Growth and oscillations of solutions of nonlinear Schrödinger equation*, Comm. Math. Phys. **178** (1996), 265–280.

[56] S.B. Kuksin, *Oscillations in space-periodic nonlinear Schrödinger equations*, Geom. Funct. Anal. **7** (1997), 338–363.

[57] S.B. Kuksin, *A KAM-theorem for equations of the Korteweg–de Vries type*, Rev. Math. Math. Phys. **10** (3) (1998), 1–64.
[58] S.B. Kuksin, *Spectral properties of solutions for nonlinear PDEs in the turbulent regime*, Geom. Funct. Anal. **9** (1999), 141–184.
[59] S.B. Kuksin, *Analysis of Hamiltonian PDEs*, Oxford University Press, Oxford (2000).
[60] S.B. Kuksin and J. Pöschel, *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Ann. Math. **143** (1996), 149–179.
[61] P.D. Lax, *Periodic solutions of the KdV equations*, Comm. Pure Appl. Math. **28** (1975), 141–188.
[62] 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.
[63] P. Lochak, *Canonical perturbation theory via simultaneous approximation*, Russian Math. Surveys **47** (6) (1992), 57–133.
[64] J. Moser, *Periodic orbits near an equilibrium and a theorem by Alan Weinstein*, Comm. Pure Appl. Math. **29** (1976), 724–747.
[65] J. Moser and C.L. Siegel, *Lectures on Celestial Mechanics*, Springer, Berlin (1971).
[66] N.N. Nekhoroshev, *Exponential estimate of the stability of near integrable Hamiltonian systems*, Russian Math. Surveys **32** (6) (1977), 1–65.
[67] L. Niederman, *Nonlinear stability around an elliptic equilibrium point in an Hamiltonian system*, Nonlinearity **11** (1998), 1465–1479.
[68] S.P. Novikov, *A periodic problem for the Korteweg–de Vries equation, I*, Funct. Anal. Appl. **8** (1974), 236–246.
[69] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, Berlin (1983).
[70] P. Plotnikov and J. Toland, *Nash–Moser theory for standing water waves*, Arch. Rational Mech. Anal. **159** (2001), 1–83.
[71] J. Pöschel, *On elliptic lower dimensional tori in Hamiltonian systems*, Math. Z. **202** (1989), 559–608.
[72] J. Pöschel, *Small divisors with spatial structure in infinite dimensional Hamiltonian systems*, Comm. Math. Phys. **127** (1990), 351–393.
[73] J. Pöschel, *A KAM-theorem for some nonlinear PDEs*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. IV Ser. 15 **23** (1996), 119–148.
[74] J. Pöschel, *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv. **71** (1996), 269–296.
[75] J. Pöschel, *On Nekhoroshev estimates for a nonlinear Schrödinger equation and a theorem by Bambusi*, Nonlinearity **12** (1999), 1587–1600.
[76] P. Rabinowitz, *Free vibrations for a semilinear wave equation*, Comm. Pure Appl. Math. **31** (1978), 31–68.
[77] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. 2, Academic Press, New York (1975).
[78] M.B. Sevryuk, *The classical KAM theory at the dawn of the twenty-first century*, Moscow Math. J. **3** (3) (2003), 1113–1144.
[79] C.E. Wayne, *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. **127** (1990), 479–528.
[80] A. Weinstein, *Normal modes for nonlinear Hamiltonian systems*, Invent. Math. **20** (1973), 47–57.
[81] V.E. Zakharov, *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, Appl. Mech. Tech. Phys. **2** (1968), 190–194.
[82] V.E. Zakharov, M.F. Ivanov and L.N. Shur, *On the abnormally slow stochastisation in some two-dimensional field theory models*, JETP Lett. **30** (1) (1979), 39–44.
[83] V.E. Zakharov, S.V. Manakov, S.P. Novikov and L.P. Pitaevskij, *Theory of Solitons*, Plenum Press, New York (1984).
[84] P.E. Zhidkov, *Korteweg–de Vries and Nonlinear Schrödinger Equations: Qualitative Theory*, Springer, Berlin (2001).