Lecture Notes in Mathematics 1556

Sergej B. Kuksin

Nearly Integrable Infinite-Dimensional Hamiltonian Systems

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Author

Sergej B. Kuksin
Institute for Information
Transmission Problems
Ermolovoy St. 19
101 447 Moscow, Russia

Mathematics Subject Classification (1991): 34G20, 35Q55, 58F07, 58F27, 58F39

ISBN 3-540-57161-2 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-57161-2 Springer-Verlag New York Berlin Heidelberg

Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2146/3140-543210 - Printed on acid-free paper

Contents

Introduction

The book is devoted to nonlinear Hamiltonian perturbations of integrable (linear and nonlinear) Hamiltonian systems of large and infinite dimension. Such systems arise in physics in many different ways. As a working hypothesis for theirs study it was postulated in the physical literature after the works of Boltzmann that in a "typical situation" their solutions are stochastic. This postulate ("ergodic hypothesis") was successfully used to explain many properties of matter. On the other hand, a lot of numerical experiments starting from the ones of Fermi–Pasta–Ulam (see [FPU], [U]) have shown quite regular recurrent behavior of many solutions of the systems under consideration (see e.g. [ZIS]). This effect cannot be explained by means of the Poincaré recurrence theorem [AKN] because the Poincaré recurrence time is much larger than the one obtained in the experiments. It seems that the investigated systems have in abundance quasiperiodic trajectories or trajectories abnormally close to the quasiperiodic ones (see [LL], [DEGM], [Mo]). These trajectories correspond to low-frequency oscillations of the underlying physical object. In these oscillations the energy is frozen in low frequencies for a very long time. So the recurrence effect causes a low rate of stochasticity (the ergodic hypothesis works now in a slow way). This effect seemed rather strange to the physicists who observed it.

Our goal in this book is to obtain some general theorem to prove that "many" quasiperiodic solutions of the unperturbed integrable system, which describes a conservative physical system with one spatial dimension, persist under perturbations. The theorem gives some explanation to the recurrence effect in spatially one-dimensional systems. It proves that in some strict sense the *one-dimensional world* "is not very ergodic".

We consider *discrete*-spectrum systems only. For Hamiltonian systems with *continuous* spectrum time-quasiperiodic solutions play rather unessential role. To study nearintegrable continuous-spectrum systems various types of averaging theorems in time- and space-variables have been developed. We avoid discussing of this expanded subject.

The main part of the book deals with perturbations of *linear* Hamiltonian equations, depending on a finite-dimensional parameter. However, it turns out that the problem of persistence quasiperiodic solutions of a *nonlinear* integrable system can be reduced to the same problem for a parameter-depending linear equation (after the reduction the frequency vector of the unperturbed nonlinear quasiperiodic oscillation plays a role of the parameter we need). Similar finite-dimensional reduction is well-known; see Remark in item 1.3 below. We discuss the infinite-dimensional case in item 3.2 of the introduction and refer the reader for details to original papers (and, hopefully, to the next book of the author). We formulate the main theorem of the book in a way to simplify its nonlinear applications. Therefore the title of the book refer to *integrable* (linear and nonlinear) systems rather than to *linear* systems only.

The introduction is devoted to a rather expanded discussion of the theorem and its applications. Sometimes the discussion supplements the results from the main text. We preface the survey of our results with a survey of the finite dimensional situation.

1. Finite dimensional situation

"Regular" (periodic and quasiperiodic) solutions of $2n$-dimensional Hamiltonian systems are important for classical and celestial mechanics. Some quite general existence theorems for this class of solutions have been obtained. Here we are interested in perturbation-type results only.

1.1 Lyapunov and Poincaré theorems

The first classical results in this direction where the Lyapunov and Poincaré theorems (see [AKN], [SM]), stating that nonresonant periodic solutions of a Hamiltonian system survive under Hamiltonian perturbation. More exactly, the Lyapunov theorem states that if the unperturbed system is a linear Hamiltonian system with the spectrum

$$\{\pm i\lambda\} \cup \{\pm\mu_1, \dots, \pm\mu_{n-1}\} ,$$

where $\lambda \in \mathbb{R}\backslash\{0\}$, $\mu_1, \dots, \mu_{n-1}$ are complex numbers and

$$ik\lambda \neq \mu_j \quad \forall j = 1, \dots, n-1 , \quad \forall k \in \mathbb{Z} ,$$

then the perturbed system has a two-dimensional invariant manifold filled with periodic solutions of frequencies close to λ (i.e., of periods close to $2\pi/\lambda$).

The Poincaré theorem states that if a Hamiltonian system has a periodic solution such that the linearization of the corresponding isoenergetic Poincaré map at the fixed point does not have an eigenvalue equal to one, then this solution lies in two-dimensional invariant manifold filled with periodic solutions. The periodic solutions and the manifold they fill persist under Hamiltonian perturbations of the equation.

1.2 Kolmogorov theorem

The second classical result concerning the subject is the Kolmogorov theorem [Kol] (stated in [Kol] with a scheme of a proof given, and proven in details by Arnold and Moser), which inspired Arnold and Moser to create a powerful technique to handle nonlinear problems, well known nowadays as KAM (Kolmogorov–Arnold–Moser) theory; see [A2], [A3], [AA], [Mo], [SM] and bibliographies of the last three books. Kolmogorov's theorem states that most of the quasiperiodic n-frequency solutions of a nondegenerate integrable analytical system with n degrees of freedom persist under analytic Hamiltonian perturbations or, equivalently, Hamiltonian perturbations preserve most of invariant n-tori of a nondegenerate integrable system. Here integrability means that in a phase space $\mathbb{T}^n \times P$ (P is a bounded n-dimensional domain) the system has the form:

$$\dot{q} = \nabla h(p) , \quad \dot{p} = 0 , \qquad (2)$$

(i.e., it has a hamiltonian h depending on the actions $p \in P$ only) and the nondegeneracy means that

$$\operatorname{Hess} h(p) := \det\{\partial^2 h(p)/\partial p_i \partial p_j\} \not\equiv 0 \,. \tag{3}$$

Invariant tori of the system (2) are of the form

$$T^n(p) = \mathrm{T}^n \times \{p\} \,, \quad p \in P \,, \tag{4}$$

and most of them survive in the perturbed system with the hamiltonian $h(p) + \varepsilon H(q,p)$,

$$\dot{q} = \nabla_p\big(h(p) + \varepsilon H(q,p)\big) \,, \quad \dot{p} = -\varepsilon \nabla_q H(q,p) \,, \tag{5}$$

if positive ε is small enough. That means that for $\rho < 1$ there exists a subset $P_\varepsilon \subset P$ such that $\operatorname{mes}(P \backslash P_\varepsilon) \to 0$ as $\varepsilon \to 0$, and for $p \in P_\varepsilon$ there exists a map $\Sigma_p : \mathrm{T}^n \to \mathrm{T}^n \times P$ and an n-vector $\omega(p)$ such that $|\omega(p) - \nabla h(p)| \leq C\varepsilon$, for all $q \in \mathrm{T}^n \operatorname{dist}\big(\Sigma_p(q),(q,p)\big) < \varepsilon^\rho$ and the curve

$$t \longmapsto \Sigma_p\big(q + t\omega(p)\big) \tag{6}$$

is a solution of (5).

For other versions and important improvements of the theorem se [AKN], [Bru1], [Bru2], [Her], [Laz], [Mo], [Mo1], [P3], [Ru], [Sev], [SZ], [Z1].

1.3 Melnikov theorem

The Lyapunov and Poincaré theorems state the persistence of nondegenerate one-dimensional invariant tori (= periodic solutions) under Hamiltonian perturbations, and the Kolmogorov theorem states the persistence most of the invariant N-tori of integrable system with N degrees of freedom. The natural question is if most of invariant tori of an intermediate dimension n, $1 < n < N$, survive under perturbations. For perturbations of a linear Hamiltonian system with $N = n + m$ degrees of freedom the question means the following. In the phase space

$$\mathrm{T}^n \times P \times \mathrm{R}^{2m} = \big\{(q,p,z)\big\} \,, \quad z = (z_+, z_-) \in \mathrm{R}^{2m} \,, \tag{7}$$

$(n \geq 2, m \geq 1)$ the Hamiltonian equations

$$\dot{q} = \lambda + \varepsilon \nabla_p H \,, \quad \dot{p} = -\varepsilon \nabla_q H \,, \quad \dot{z} = J(Az + \varepsilon \nabla_z H) \,, \tag{8}$$

are considered. Here $J(z_+, z_-) = (-z_-, z_+)$, A is a symmetric linear operator in R^{2m}, $\varepsilon H = \varepsilon H(q,p,z)$ is an analytic perturbation and $\lambda \in \Lambda \subset \mathrm{R}^n$ is a parameter. For $\varepsilon = 0$ the system (8) has invariant n-tori $T^{n,m}(p) = \mathrm{T}^n \times \{p\} \times \{0\}$, $p \in P$. The question is if these tori persist in the system (8) for $\varepsilon > 0$.

Let us denote the spectrum of the operator JA by $M = \{\mu_1, \ldots, \mu_{2m}\}$. We should distinguish three cases:

a) (*nondegenerate hyperbolic tori*) $M \subset C \backslash i\mathrm{R}$, $\mu_j \neq \mu_k \quad \forall_j \neq k$. In this situation a hyperbolic torus $T^{n,m}(p)$ persist for most λ. That is, for positive ε small enough and for $\lambda \in \Lambda(\varepsilon,p)$, where $\operatorname{mes} \Lambda\backslash\Lambda(\varepsilon,p) \to 0$ $(\varepsilon \to 0)$, the equation (8) has an invariant torus at a distance $< \varepsilon^\rho$ from $T^{n,m}(p)$. See [Gr], [Mo], [Z1].

b$_0$) (*nondegenerate elliptic tori*) $M \subset i\mathbb{R}\backslash\{0\}$, $\mu_j \neq \mu_k$ $\forall j \neq k$. This situation is more complicated. The preservation theorem for the elliptic torus $T^{n,m}(p)$ for most λ was formulated by Melnikov [Me1], [Me2]. The complete proof of the theorem was published only 15 years later by Eliasson [El], Pöschel [P1] and the author [K1], [K2] (the infinite-dimensional theorems of the last two works are applicable to equations (8) as well). The proofs given in the papers just mentioned are also valid in the more general situation:

b) (*nondegenerate tori*) $0 \notin M$, $\mu_j \neq \mu_k$ $\forall j \neq k$.

In the degenerate case

c) $0 \in M$ or $\mu_j = \mu_k$ for some $j \neq k$

no preservation theorem for the tori $T^{n,m}(p)$, formulated in terms of the unperturbed equation (8) with $\varepsilon = 0$ only, is known yet.

Remark. Melnikov theorem (case b$_0$)) remains true for $m = 0$, too. In such a case $Y = \{0\}$ and the theorem asserts the preservation of the n-dimensional invariant torus $\mathbb{T}^n \times \{0\}$ of the system with the linear hamiltonian $h(p) = \lambda \cdot p$,

$$\dot{q} = \lambda\,, \quad \dot{p} = 0\,,$$

under small analytic Hamiltonian perturbations for most parameters $\lambda \in \Lambda$. This result implies Kolmogorov's theorem as it was formulated above via some simple substitution; see [Mo1], p. 171.[1)] Conversely, one can easily extract somewhat different version of Melnikov's theorem with $m = 0$ from Kolmogorov's theorem.[2)] So these two statement are essentially equivalent. This equivalence (we had found it in the paper [Mo1]) was important for our insight into infinite-dimensional problems.

Remark. The equation (8) arises in studies of nearintegrable systems (5) with $n := N$ near the tori (4) "with some cycles shrinked to zero". It means that we suppose the system (5) be in the Birkhoff normal form (i.e., in the phase-space $\mathbb{R}^N_x \times \mathbb{R}^N_y$ with the usual symplectic structure it has an analytic hamiltonian $h(x_1^2 + y_1^2, \dots, x_N^2 + y_N^2)$) and study its perturbations near an invariant n-torus $\{x_j^2 + y_j^2 = 2I_j\}$, where $I_1, \dots, I_n > 0 = I_{n+1} = \cdots = I_N$.

Lower-dimensional invariant tori also fill resonant Lagrangian (= half-dimensional) tori (4), but they always lead to the degenerate case c). In particular, if for

---

1) In (5) substitute $p = a + \sqrt{\varepsilon}\tilde{p}$, $q = \tilde{q}$, regarding $a \in P$ as a parameter of the substitution. In the tilde-variables the hamiltonian $h + \varepsilon H$ equals $\mathrm{const} + \nabla h(a)\cdot\tilde{p} + O(\sqrt{\varepsilon})$, and we get the system (8) with $\varepsilon := \sqrt{\varepsilon}$, $m = 0$, $\lambda = \nabla h(a)$. As $\operatorname{Hess} h \not\equiv 0$, then we can treat $\lambda \in \nabla h(P)$ as a new parameter and apply Melnikov theorem.

2) Given a system (8) with $m = 0$, consider the extended phase-space $(\mathbb{T}^n \times P) \times (\mathbb{T}^n \times \Lambda) = \{(q,p,\tilde{q},\tilde{p})\}$ and the hamiltonian $\tilde{H}_\varepsilon = p \cdot \tilde{p} + \varepsilon H(q,p)$. The nondegeneracy assumption (3) holds for the function $(p,\tilde{p}) \longmapsto p\cdot\tilde{p}$ and Kolmogorov theorem (with $n := 2n$) can be aplied. The invariant tori of the hamiltonian $\tilde{H}_\varepsilon$ have the form $T^n_\varepsilon \times (\mathbb{T}^n \times \{\tilde{p}\})$, where $\tilde{p} \in \Lambda$ and $T^n_\varepsilon \subset \mathbb{T}^n \times P$ is an invariant torus of (8).

$E(p) := \mathbb{Q}\partial h/\partial p_1 + \cdots + \mathbb{Q}\partial h/\partial p_N$ we have $\dim_{\mathbb{Q}} E(p) = N-1$, then the torus $T^N(p)$ is filled with invariant $(N-1)$-tori. Near each $(N-1)$-torus the perturbed system (5) may be reduced to a system (8) with JA equal to the Jordan 2×2-cell with zero eigenvalue. For more information see [Lo] and references therein.

Famous finite-gap time-quasiperiodic solutions of integrable PDE's form finite-dimensional invariant tori of the corresponding infinite-dimensional Hamiltonian integrable systems, obtained by "shrinking" (not "degenerating" !) of half-dimensional invariant tori. See below and [McT], [K7].

2. Infinite dimensional systems

2.1 The Problem

In a Hilbert space Z with inner product $\langle\cdot,\cdot\rangle$ we consider the equation

$$\dot u(t) = J\nabla\mathcal{K}\big(u(t)\big)\ , \quad u(t)\in Z\ . \tag{9}$$

Here J is an antiselfadjoint operator in Z and $\nabla\mathcal{K}$ is the gradient of a functional $\mathcal{K}$ relative to the inner product $\langle\cdot,\cdot\rangle$. In the most interesting situations the linear operator J, or the nonlinear operator $\nabla\mathcal{K}$, or both of them are unbounded. So one has to be careful with the equation and its solutions. For the exact definition of solutions of (9) and for some their properties see [Bre], [Lio] and Part 1 of the main text. Equation (9) is Hamiltonian if the phase space Z is provided with a symplectic structure by means of 2-form $-\langle J^{-1}du, du\rangle$ (by definition, $-\langle J^{-1}du,du\rangle[\xi,\eta] = -\langle J^{-1}\xi,\eta\rangle$ for ξ,η in Z).

In this book we are most interested in equations of the form

$$\dot u(t) = J\Big(Au(t) + \varepsilon\nabla H\big(u(t)\big)\Big)\ . \tag{10}$$

This equation is Hamiltonian with the hamiltonian

$$\mathcal{K}_\varepsilon = \frac{1}{2}\langle Au,u\rangle + \varepsilon H(u)\ .$$

Here A is a selfadjoint linear operator in Z and H is an analytic functional. The linear operators J, A and the nonlinear operator ∇H are assumed to be characterized by their *orders* d^J, d^A and d^H. We suppose that

$$d^J \ge 0\ ,\ d^A \ge 0\ ,\ d^J + d^A \ge 1\ ,\ d^J + d^H \le 0\ . \tag{11}$$

In the most important examples Z is the L_2-space of square-summable functions on a segment, and J and A are differential operators. In such a case d^J, d^A are the orders of the differential operators and $\nabla H(u)$ is a variational derivative $\delta H/\delta u(x)$. In particular, if

$$H(u) = \int h\big(x,u(x)\big)\,dx\ ,$$

then $\delta H/\delta u(x) = h_u\big(x,u(x)\big)$ and $d^H = 0$; if the density h depends on integral of $u(x)$ instead of $u(x)$ itself, then $d^H < 0$. To define the orders d^J, d^A, d^H in a general case, we include the space Z into a scale of Hilbert spaces. See Part 1 below.

The assumption (11) implies that equation (10) is quasilinear. This assumption is rather natural for the study of long-time behavior of solutions because for some strongly nonlinear Hamiltonian equations (i.e. ones of the form (10) with $d^H = d^A$) it is known that the equations have no nontrivial solutions existing for all time; see [Lax].

We suppose that J and A commute and that Z admits an orthonormal basis $\{\varphi_j^\pm \mid j \geq 1\}$ such that

$$A\varphi_j^\pm = \lambda_j^A \varphi_j^\pm \ , \quad J\varphi_j^\pm = \mp\lambda_j^J \varphi_j^\mp \ , \quad \forall j \geq 1 \ . \tag{12}$$

So, in particular, the spectrum of the operator JA is equal to

$$\{\pm i\lambda_j \mid j \geq 1, \lambda_j = \lambda_j^A \lambda_j^J\} \ .$$ [3]

Let us fix some $n \geq 1$. The $2n$-dimensional linear space

$$Z^0 = \operatorname{span}\{\varphi_j^\pm \mid 1 \leq j \leq n\}$$

is invariant for the flow of equation (10) with $\varepsilon = 0$, is foliated into invariant n-tori

$$T^n(I) = \{\sum_{j=1}^n x_j^\pm \varphi_j^\pm \mid x_j^{+^2} + x_j^{-^2} = 2I_j \ \forall j\} \ ,$$

$I = (I_1, \dots, I_n) \in \mathbb{R}_+^n$, and every torus $T^n(I)$ is filled with quasiperiodic solutions of the equation.[4]

One can treat (10) with $\varepsilon = 0$ as an infinite chain of free harmonic oscillators with the frequencies $\lambda_1, \lambda_2, \dots$. The solutions lying on the tori $T^n(I)$ correspond to oscillations with only the first n oscillators being excited. One can treat these solutions as low-frequency oscillations.

We study the question: *under what assumptions do the tori $T^n(I)$ and the corresponding low-frequency quasiperiodic solutions persist in equation (10) for $\varepsilon > 0$?*

It is convenient to introduce the angle-action variables $(q_1, \dots, q_n, p_1, \dots, p_n)$ in the space Z^0,

$$x_j^+ + ix_j^- = \sqrt{2p_j}\exp(iq_j) \ , \quad j = 1, \dots, n$$

($x_j^\pm$ are the coordinates with respect to the basis $\{\varphi_j^\pm \mid 1 \leq j \leq n\}$); to denote by $Y = Z \ominus Z^0$ the closure of $\operatorname{span}\{\varphi_j^\pm \mid j \geq n+1\}$ and to pass to the variables (q, p, y),

$$q = (q_1, \dots, q_n) \in \mathbb{T}^n \ , \quad p = (p_1, \dots, p_n) \in \mathbb{R}_+^n \ , \quad y \in Y \ . \tag{13}$$

---

[3] The assumption (12) may be essentially weakened. See Part 2.7.

[4] We remain that a solution $u(t)$ is called *quasiperiodic* with n frequencies if there exist a continuous map $U : \mathbb{T}^n \to Z$ and n-vector ω (called the *frequency vector* of the solution) such that $u(t) \equiv U(\omega t)$. A quasiperiodic solution with one frequency is periodic, so quasiperiodic solutions represent a natural extension of the class of periodic solutions.

Let us denote by Σ^0 the imbedding

$$\Sigma^0 : \mathrm{T}^n \times \mathrm{R}^N_+ \longrightarrow Z\ ,\quad (q,p) \longmapsto (q,p,0)\ .$$

(we use in Z the coordinates (13)). The invariant space Z^0 is the image of this map.

In the new variables (13) equation (10) takes the form:

$$\dot q = \nabla_p \mathcal{H}\ ,\quad \dot p = -\nabla_q \mathcal{H}\ ,\quad \dot y = J^Y \nabla_y \mathcal{H} \tag{14}$$

with

$$\mathcal{H} = \mathcal{H}_\varepsilon = \omega \cdot p + \frac{1}{2}\langle A^Y y, y\rangle + \varepsilon H(q,p,y)\ .$$

Here $\omega = (\lambda_1, \ldots, \lambda_n)$, $J^Y = J_{|Y}$, $A^Y = A_{|Y}$. So the operator $J^Y A^Y$ has pure imaginary spectrum $\{\pm i\lambda_j \mid j \geq n+1\}$ and one can easily recognize in the last equations an infinite-dimensional analogy to the elliptic case of the system (8). The form of Melnikov's theorem we gave above in Section 1.3 has a natural infinite-dimensional reformulation. It is remarkable that this reformulation becomes a true statement after adding essentially just two infinite-dimensional conditions.

2.2 The result

Keeping in mind the applications, we suppose that equation (10) analytically depends on n outer parameters $(a_1, \ldots, a_n) = a \in \mathfrak{A}$, where $\mathfrak{A}$ is a connected bounded open domain in R^n. So $A = A_a$, $H = H_a$ and $\lambda_j = \lambda_j(a)$. Let us assume that

$$\det\bigl\{\partial\lambda_j(a)/\partial a_k \bigm| 1 \leq j,k \leq n\bigr\} \not\equiv 0\ . \tag{15}$$

This assumption means that we can replace the parameter a by $\omega = (\lambda_1, \ldots, \lambda_n)(a)$. Later we refer to ω as to the *natural parameter* of the equation.

We consider a torus $T^n(I_1, \ldots, I_n)$ such that $I_k > 0\ \forall k$.

Theorem 1. Let us suppose that the assumptions (12), (15) hold together with

1) *(quasilinearity)*

$$d^J \geq 0\ ,\quad d^A \geq 0\ ,\quad d_1 := d^J + d^A \geq 1\ ,\quad d^J + d^H \leq 0\ ,\quad d^J + d^H < d_1 - 1\ ,$$

2) *(spectral asymptotics)*

$$\lambda_j(a) = K_1 j^{d_1} + K_2 + \mu_j(a)\ ,$$

where

$$|\mu_j(a)| + |\nabla \mu_j(a)| \leq K_3 j^{d_1 - \kappa}$$

for some $\kappa > 1$;

3) for some $N \geq n$ and $M \geq 1$ depending on the problem (10) the *nonresonance relations*

$$s_1\lambda_1(a) + s_2\lambda_2(a) + \ldots + s_N\lambda_N(a) \not\equiv 0 \tag{16}$$

hold for all $s \in Z^N$ such that $1 \le |s| \le M$ and $|s_{n+1}| + \cdots + |s_N| \le 2$.

Then for arbitrary $\rho < 1$ and for positive ε small enough there exist a Borel subset $\mathfrak{A}_\varepsilon(I) \subset \mathfrak{A}$ and analytic embeddings

$$\Sigma^\varepsilon_{a,I} : \mathrm{T}^n \longrightarrow Z \ , \quad a \in \mathfrak{A}_\varepsilon(I) \ , \tag{17}$$

such that

a) $\operatorname{mes}\big(\mathfrak{A} \backslash \mathfrak{A}_\varepsilon(I)\big)/\operatorname{mes}\mathfrak{A} \to 0$ $(\varepsilon \to 0)$;

b) the map $(q, I, a) \longmapsto \sum^\varepsilon_{a,I}(q)$ is Lipschitz and is ε^ρ-close to the map $(q, I, a) \longmapsto \Sigma^0(q, I)$;

c) for $a \in \mathfrak{A}_\varepsilon(I)$ the torus $\Sigma^\varepsilon_{a,I}(\mathrm{T}^n)$ is invariant for the equation (10) and is filled with quasiperiodic solutions of the form $u_\varepsilon(t) = \Sigma^\varepsilon_{a,I}(q + \omega_\varepsilon t)$ with a frequency vector $\omega_\varepsilon \in \mathrm{R}^n$ which is $C\varepsilon$-close to $\omega = (\lambda_1, \ldots, \lambda_n)$. All Lyapunov exponents of these solutions are equal to zero.

Refinement (see Part 3, Theorem 1.1). In the variables (13) the unperturbed hamiltonian is equal to $\omega \cdot p + \frac{1}{2}\langle A^Y y, y\rangle$ and the perturbation is $\varepsilon H_a(q, p, y)$. The statements of the theorem remain true for perturbations of the more general form

$$\varepsilon H_a = \varepsilon H_{1a}(q, p, y) + H^3_a(q, p, y) \ , \quad H^3_a = O\big(|p - I|^2 + \|y\|^3 + \|y\|\,|p - I|\big) \ .$$

This form of the result is suitable for applications to perturbations of nonlinear problems (see below).

The formulations of our results given above are "almost exact". For the exact statements see the main text. In Part 2 of the text we state local and global in a versions of Theorem 1 (Theorem 2.1.1 and 2.2.2 respectively, where the latter is a rather simple consequence of the former); we give various applications of the theorems to nonlinear perturbations of linear PDE's and postpone the proof of Theorem 2.1.1 till Part 3. There we reformulate the theorem in a more general form to facilitate its applications to nonlinear problems (the reformulated theorem also includes Refinement given above) and prove the result. In Part 3 we also give a version of Theorem 1, applicable to the problems with the natural parameter $\omega = (\lambda_1, \ldots, \lambda_n)$ varying in a domain of small diameter of order ε^κ, $0 < \kappa < 1$. This result is useful to study small-amplitude oscillations in nonlinear PDE's (see item 3.2.B below).

Remark. If the natural parameter ω is chosen for the parameter of the equation and λ_j does not depend on ω for $j \ge n + 1$, then the assumption 3) is fulfilled trivially. If in addition $\dim Z < \infty$, then the assumptions 1), 2) hold trivially, too. So for finite-dimensional systems (written in the form (14)) Theorem 1 coincides with Melnikov's theorem.

As another infinite-dimensional Melnikov-type theorem we mention the result of Wayne's paper [W1], devoted to the nonlinear string equation with a random potential. We discuss the approach, the work [W1] is based on, below.

Remark. If the hamiltonian of the perturbation is quadratic, then the theorem's statements for all parameters a (and ε small enough) immediately result from the classical perturbation theory for the discrete spectrum of a linear operator in Hilbert space (see, e.g. [RS]). In the nonlinear case the theorem's statements certainly does not hold for *all* parameters because of resonances between the frequencies $\{\lambda_j\}$, which occur for some a and give rise to much more complicated phenomenous. For discussions some of them in the finite-dimensional situation see [AKN] and [Mo].

Remark. As the map (17) is ε^ρ-close to the map $q \longmapsto \Sigma^0(q,I)$, then the solutions $u_\varepsilon(t)$ are ε^ρ-close to the curves $t \longmapsto \Sigma^0(q+\omega_\varepsilon t, I)$ for all t. The vector ω_ε is equal to $\omega+\varepsilon\omega_1+\varepsilon^2\omega_2+\cdots$, where the vector ω_1 may be obtained via some natural averaging (see [K4]). So Theorem 1 gives an averaging procedure for low-frequency solutions of equation (10) as a simple consequence.

Under the assumptions of the theorem an unperturbed torus $T^n(I)$ with

$$I \in \mathcal{I} = \{x \in \mathbb{R}^n \mid K^{-1} \le x_j \le K \ \forall j\}$$

survives in the equation (10) if $\varepsilon \le \varepsilon_0$ and a belongs to a set $\mathfrak{A}_\varepsilon(I)$ such that

$$\operatorname{mes}(\mathfrak{A}\backslash\mathfrak{A}_\varepsilon(I)) \le \nu(\varepsilon)\operatorname{mes}\mathfrak{A}\ ,$$

where $\nu(\varepsilon) \to 0$ as $\varepsilon \to 0$. The number ε_0 and the function $\nu(\varepsilon)$ do not depend on I (but depend on K). Let us denote

$$\mathcal{I}_\varepsilon(a) = \{I \in \mathcal{I} \mid a \in \mathfrak{A}_\varepsilon(I)\}\ .$$

The torus $T^n(I)$ persists if $I \in \mathcal{I}_\varepsilon(a)$. By Fubini theorem,

$$(\operatorname{mes}\mathfrak{A})^{-1}\int_{\mathfrak{A}} \operatorname{mes}(\mathcal{I}\backslash\mathcal{I}_\varepsilon(a))\,da = (\operatorname{mes}\mathfrak{A})^{-1}\int_{\mathcal{I}} \operatorname{mes}(\mathfrak{A}\backslash\mathfrak{A}_\varepsilon(I))\,dI \le \operatorname{mes}\mathcal{I}\,\nu(\varepsilon)\ . \tag{18}$$

Let us consider the sets

$$Z_K^0 = \{(q,I) \in Z^0 \mid I \in \mathcal{I}(a)\}\ ,\quad Z_K^\varepsilon = \{(q,I) \in Z^0 \mid I \in \mathcal{I}_\varepsilon(a)\}\ .$$

By (18) for a typical a the relative measure of Z_K^ε in Z_K^0 is no less then $1-\nu(\varepsilon)$. The image of the set Z_K^ε under the map

$$(q,I) \longmapsto \Sigma_{a,I}^\varepsilon(q) \tag{19}$$

is invariant for the flow of equation (10) and is filled with quasiperiodic solutions. The mapping (19) is Lipschitz and ε^ρ-close to the embedding Σ^0. So the Hausdorf measure $\mathcal{H}^{2n}$ (see [Fe]) of the invariant set as above is no less then

$$(1-\nu_1(\varepsilon))\operatorname{mes}_{2n}Z_K^0\ , \tag{20}$$

with some $\nu_1(\varepsilon) \to 0$ as $\varepsilon \to 0$. Taking K large enough and ε sufficiently small one can make (20) as large as desired. So we have seen that under the assumptions of Theorem 1 for typical a and for ε small enough the equation (10) has invariant sets of the Hausdorf measure $\mathcal{H}^{2n}$ as large as desired. These sets are filled with quasiperiodic trajectories with zero Lyapunov exponents. They form obstacles to the fast stochastisation of solutions of a typical system of form (10). Our guess is that the recurrence effect "of FPU type" is caused by such sets.

Our results leave without any answer the natural question: do the *infinite-dimensional* invariant tori of the system (10) with $\varepsilon = 0$ persist under Hamiltonian perturbations? The answer is affirmative if the following three assumptions are satisfied:

a) the perturbation H has short range interactions, i.e. for $u(t)$ written as $\Sigma x_k^{\pm}(t)\varphi_k^{\pm}$, and for some finite N the equation for $x_k^{\pm}$ does not depend on $x_m^{\pm}$ with $|k-m| \geq N$ (or depends on $x_m^{\pm}$ in an exponentially small with respect to $|k-m|$ way);

b) $|H(u)| = O(\|u\|^d)$ for some $d > 2$;

c) the coefficients $x_k^{\pm}$ decrease, for example, exponentially when k is growing.

The assumptions a), b) never hold for nonlinear partial differential equations (but they are fulfilled for some equations from the physics of crystals). For the exact statements see [FSW], [VB] and [P2], [W2], [AlFS], [ChP]. We remark that the works [FSW], [VB] were the first ones where KAM theory was applied to infinite-dimensional Hamiltonian systems.

Without the assumptions a)–b) the maximal magnitude of the perturbation which allows one to prove Kolmogorov's theorem (=to prove preservation most of half-dimensional tori) exponentially decrease with the dimension of the phase-space (see e.g. [P2, p.364]). We suppose that the exponential estimate is the best possible one. In particular, infinite-dimensional tori "in general" do not survive under the system's perturbations.

We end this part with the remark that some results concerning the preservation of infinite-dimensional tori in equation (10) with the sectrum $\{\pm i\lambda_j\}$ of a special type may be obtained via infinite-dimensional versions of Siegel's theorem. See [War], [Z2] and especially [Nik].

3. Applications

In this item we show that Theorem 1 gives a flexible tool to study nonlinear Hamiltonian PDE's with one-dimensional spatial variable. We discuss applications to nonlinear perturbations of linear equations, to small-amplitude oscillations in nonlinear equations and to perturbations of the integrable PDE's. We end the item with some arguing why the theorem can not be applied to multi-dimensional PDE's.

3.1 Perturbations of linear differential equations

As a rule, the assumption 1) of Theorem 1 is fulfilled if JA is a differential operator on a segment with some self-adjoint boundary conditions. So the theorem is applicable to spatially one-dimensional quasilinear Hamiltonian partial differential equations, depending on a vector parameter.

Example 1 (see [K1] and Part 2.3). Let us consider nonlinear Schrödinger equation with a bounded real potential $V(x;a)$, depending on an n-dimensional parameter a:

$$\begin{aligned} \dot{u} &= i\big(-u_{xx} + V(x;a)u + \varepsilon\varphi'(x,|u|^2;a)u\big) , \\ u &= u(t,x) , \ t \in \mathbb{R} , \ x \in (0,\pi) ; \quad u(t,0) \equiv u(t,\pi) \equiv 0 . \end{aligned} \tag{21}$$

Here φ is a real function analytic in $|u|^2$ and $\varphi' = \partial\varphi/\partial|u|^2$. To apply the theorem one has to set Z equal to the space of square-summable complex-valued functions on $(0,\pi)$ (and consider it as a real Hilbert space), to set A_a equal to the differential operator $-\partial^2/\partial x^2 + V(x;a)$ under the Dirichlet boundary condition, to set $Ju(x) = iu(x)$ and

$$H_a\big(u(x)\big) = \frac{1}{2}\int_0^\pi \varphi\big(x,|u(x)|^2;a\big)\,dx .$$

Let us denote by $\{\varphi_j(x;a)\}$, $\{\lambda_j(a)\}$ complete systems of real eigenfunctions and eigenvalues of the operator A_a. The invariant n-tori of the unperturbed problem are of the form

$$T(I) = \Big\{\sum_{j=1}^{n}(\alpha_j^+ + i\alpha_j^-)\varphi_j(x;a) \mid {\alpha_j^+}^2 + {\alpha_j^-}^2 = 2I_j > 0 \ \forall j\Big\} .$$

By the well-known asymptotics of the spectrum of the Sturm–Liouville problem ([Ma], [PT]), $\lambda_j(a) = j^2 + O(1)$ and the assumption 1) of Theorem 1 is fulfilled with $d_1 = 2$, $\kappa = 3/2$. The theorem is applicable to the problem (21); therefore the torus $T(I)$ persists in the probelm (21) for most of a and ε small enough, if the potential V depends on a in a nondegenerate way. So for nondegenerate families of potentials $\{V(\cdot;a)\}$ and for typical parameters a equation (21) has a lot of quasiperiodic in t solutions, localized in the phase-space Z in a ε^ρ-neighborhood of the low-frequency tori $T(I)$.

Example 2 (see Part 2.4). We consider nonlinear Schrödinger equation with real random potential $V_\nu(x)$ under the Dirichlet boundary conditions:

$$\dot{u} = i\big(-u_{xx} + V_\nu(x)u + \varepsilon\varphi'(x,|u|^2)u\big) , \quad u(t,-\pi) \equiv u(t,\pi) \equiv 0 , \tag{22}$$

where ν is a random parameter. We denote by $QP_\varepsilon = QP_\varepsilon(\nu) \subset Z$ the random subset of the phase space $Z = L_2(-\pi,\pi;\mathbb{C})$, equal to the union of all time quasiperiodic solutions with zero Lyapunov exponents (we treat the solutions as curves in Z). It occurs that if the potential V is x-periodic "with good randomness properties", then the set QP_ε is asymptotically dense in the phase space as $\varepsilon \to 0$: for any complex function $\mathfrak{z}(x)$

$$\operatorname{dist}\big(\mathfrak{z}(\cdot), QP_\varepsilon\big) \longrightarrow 0 \quad (\varepsilon \to 0)$$

in probability.

To prove this statement we treat (22) as an equation (21) with an infinite-dimensional parameter a. We 1) apply the results of Example 1 to construct invariant tori of dimensions 1,2,... ; 2) prove that the union of these finite-dimensional invariant tori is asymptotically dense in Z when $\varepsilon \to 0$.

This result explains (and predicts) long-time regular behaviour of "typical" solutions of (22), trapped by linearly-stable regular solutions from QP_ε. The solutions in QP_ε can be eternally approximated with accuracy $C\varepsilon$ by the quasiperiodic solutions of linear equation $(22)|_{\varepsilon=0}$ with the frequency vector ω replaced by some averaged vector ω_ε (see Corollary 2.1.1 in Part 2). We think that this result, which is also true for other hamiltonian PDE's with random coefficients, can be treated as a kind of averaging theorem for nonlinear PDE's.

Example 3 (see Part 2.5). Theorem 1 can be applied to study nonlinear perturbations of the quantized harmonic oscillator

$$\begin{aligned} \dot u &= i\Big(-u_{xx} + \big(x^2 + V_0(x;a)\big)u + \varepsilon\nabla H_a(u)\Big) , \\ u &= u(t,x) , \ x \in \mathbb{R} , \ u(t,\cdot) \in L_2(\mathbb{R}) , \end{aligned} \tag{23}$$

where the function V_0 vanishes at $x = \pm\infty$. The operator $A_a = -\partial^2/\partial x^2 + x^2 + V_0$ has a discrete spectrum $\{\lambda_j(a)\}$, which obeys Bohr's quantization law: $\lambda_j \sim C(j+1/2)$. Moreover, $|\lambda_j - C(j+1/2)| \le C_1 j^{-1/2}$. So the spectral asymptotic assumption holds with $d_1 = 1$ and Theorem 1 can be applied to (23), provided that the gradiental map $\nabla H_a(u)$ is of a negative order. In particular, if

$$H_a = \frac{1}{2}\int \varphi\big(|u * \xi(x)|^2 ; a\big)\, dx$$

($u * \xi$ is the convolution with a smooth real-valued function ξ, vanishing at infinity).

We can also consider perturbed unharmonic oscillator

$$\dot u = i\Big(-u_{xx} + \big(x^2 + \mu x^4 + V_0(x;a)\big)u + \varepsilon\varphi'(x, |u|^2 ; a)u\Big) , \tag{24}$$

where $\mu > 0$. Now

$$\lambda_j = C_1\big(n + \frac{1}{2}\big)^{4/3} + C_2\big(n + \frac{1}{2}\big)^{2/3} + O(1) ,$$

so the assumption 2) of Theorem 1 holds in a slightly generalized form with $d_1 = \kappa = 4/3$ (below in Part 3 the theorem is stated and proven with the spectral assumption exactly in this form). The first assumption of Theorem 1 holds with $d_H = 0$, $d_J = 0$, $d_1 = 4/3$. So typically equation (24) has many time-quasiperiodic solutions, if ε is small enough.

Example 4 (see [K2] and Part 2.6). Let us consider the equation of oscillations of a string with fixed ends in nonlinear-elastic media depending on n-dimensional parameter:

$$\begin{aligned} \ddot w &= \big(\partial^2/\partial x^2 - V(x;a)\big)w - \varepsilon\varphi_w(x, w; a) , \\ w &= w(t,x) , \ \ 0 \le x \le \pi , \ \ t \in \mathbb{R} ; \ \ w(t,0) \equiv w(t,\pi) \equiv 0 . \end{aligned} \tag{25}$$

After some reduction (see Part 2.6) Theorem 1 is applicable to this problem with the choice $d_1 = 1$, $\kappa = 3/2$, $d_H = -1$. So in a nondegenerate case quasiperiodic in

t solutions of the unperturbed problem (25) with $\varepsilon = 0$ persist in the problem (25) for most of a and for ε small enough.

Concrete examples of nondegenerate potentials are given in Part 2.6 (in particular if $n = 1$, then one can take $V(x; a) = a$).

If $n = 1$, then the theorem deals with time-periodic solutions

$$w(t,x) = I\varphi_j(x;a)\sin\left(\sqrt{\lambda_j(a)}\,(t+q)\right)$$

of linear string equation $(25)|_{\varepsilon=0}$, depending on a one-dimensional parameter a (as above, $\{\varphi_j(x;a)\}$ and $\{\lambda_j(a)\}$ are eigenfunctions and eigenvalues of the operator $-\partial^2/\partial x^2 + V(x;a)$). These solutions persist in the perturbed equation (25) for most of a, if $\lambda_j'(a) \not\equiv 0$ and

$$m \cdot \sqrt{\lambda_j(a)} \not\equiv \sqrt{\lambda_N(a)}\ , \quad m \cdot \sqrt{\lambda_j(a)} \not\equiv \sqrt{\lambda_N(a)} \pm \sqrt{\lambda_M(a)}\ ,$$

where m is an arbitrary integer and the numbers j, N, M are pairwise different. This statement is an infinite-dimensional analog of the Lyapunov theorem (see item 1.1 above) with an additional second-order nonresonance condition. Recently Craig and Wayne [CW] proved that the extra condition may be omitted provided that the functions V and φ are analytic in x.

Time-periodic solutions of nonlinear string equation under the Dirichlet boundary conditions have been studied by many authors (see the survey [Bre1]). Under different restrictions on the nonlinear term of the equation it was proven that the equation has a *countable* family of time-periodic solutions. Our tools enable us to prove that for typical potentials the equation has time-periodic solutions, parametrized by the points of some *one-dimensional* sets (see (18)). In [BoK] similar result is proven for parameter-independent equation (25) with $V = 1$, provided that $\varphi = \kappa\varphi^4 + O(|\varphi|^5)$, $\kappa \neq O$ (the proof is based on an application of Theorem 1 to perturbations of the Sine-Gordon and Sinh-Gordon equations).

3.2 Perturbations of nonlinear systems

A) *Perturbations of Birkhoff-integrable systems* (see [K5], Example 1).

We call a Hamiltonian system *Birkhoff integrable* if it may be analytically reduced to an infinite sequence of Hamiltonian equations of the form

$$\dot{x}_j^+ = \partial H_0/\partial x_j^-\ , \quad \dot{x}_j^- = -\partial H_0/\partial x_j^+\ , \quad j = 1, 2, \dots$$

with

$$H_0 = H_0(p_1, p_2, \dots)\ , \quad p_j = \frac{1}{2}\left(x_j^{+^2} + x_j^{-^2}\right)$$

(i.e., it may be analytically reduced to the Birkhoff normal form, see [Mo], [SM]). The n-tori

$$T(p) = \left\{x | x_j^{+^2} + x_j^{-^2} = 2p_j\ , \quad j = 1, \dots, n\ ; \quad 0 = x_{n+1}^{\pm} = x_{n+2}^{\pm} = \cdots\right\}$$

are invariant for the system. It is convenient to pass to the variables (q,p,y) as in (13) with $p=(p_1,\dots,p_n)$, $y=(y_1^+,y_1^-,y_2^+,\dots)$, $y_j^\pm = x_{n+j}^\pm$ $(j=1,2,\dots)$. In these variables the equations have the form (14) with

$$\mathcal{H}_0(q,p,y) = h(p) + \frac{1}{2}\langle A(p)y,y\rangle + O(|y|^3)\,,$$

where

$$h(p) = H_0(p_1,\dots,p_n,0,\dots)\,,$$

and

$$\langle A(p)y,y\rangle = \sum_{j=1}^{\infty}\left(y_j^{+^2}+y_j^{-^2}\right)\frac{\partial}{\partial p_{n+j}}h(p_1,\dots,p_n,0,\dots)\,.$$

So the ε-perturbed hamiltonian in the new variables is equal to

$$\mathcal{H}_\varepsilon = \mathcal{H}_0 + \varepsilon H_1 = h(p) + \frac{1}{2}\langle A(p)y,y\rangle + O(\|y\|^3) + \varepsilon H_1\,. \tag{26}$$

Let us fix for a moment some $a\in \mathbb{R}_+^n$ and rewrite $\mathcal{H}_\varepsilon$ as follows:

$$\mathcal{H}_\varepsilon = \left[h(a)-\omega(a)\cdot a\right] + \omega(a)\cdot p + \frac{1}{2}\langle A(a)y,y\rangle + \varepsilon H_1 + O(\|y\|^3 + |p-a|^2 + |p-a|\,\|y\|^2)\,,$$

where $\omega(a) = \nabla h(a)$. The term in the square brackets does not affect the dynamics and may be neglected. Let us suppose that the system possesses nondegenerate *amplitude-frequency modulation*:

$$\operatorname{Hess} h(a) = \det\{\partial\omega_j(a)/\partial a_k\} \not\equiv 0\,. \tag{27}$$

Then one can treat the vector a as a parameter of the problem and apply to the perturbed problem Theorem 1, taking into account Refinement. So if spectral asymptotics and nondegeneracy assumptions are fulfilled, then most of invariant tori $T(a)$ survive under perturbations.

The trick we have just discussed is well suited to study perturbations of finite-dimensional integrable systems but not perturbations of integrable partial differential equations of Hamiltonian form. The reason is that in the last case the transition to the Birkhoff coordinates (or to the action-angle ones) is not regular.[5] To handle the integrable PDE's one needs more sophisticated approach; see item C) below.

B) *Use of the partial Birkhoff normal form*

One can treat the unperturbed linear Hamiltonian system (10) as a Birkhoff integrable system with the quadratic hamiltonian $\mathcal{H}_0 = h(p) + \frac{1}{2}\langle Ay,y\rangle$, where $\frac{1}{2}\langle Ay,y\rangle = \frac{1}{2}\Sigma_{j=n+1}^{\infty}\lambda_j\left(x_j^{+^2}+x_j^{-^2}\right)$ and $h(p)=\lambda_1 p_1+\dots+\lambda_n p_n$, $\omega(p)=(\lambda_1,\dots,\lambda_n)$.

[5] At least, the smoothness or analyticity of the action-angle variables is not proven yet. See in [McT] *continuous* action-angle variables for the KdV equation; see [Kap] for the fact that the constructed in [McT] foliation of the phase-space to invariant tori — not the action-angle variables themselves! — is analytic in L_2-norm.

Now the condition (27) is broken and one can not use an amplitude-frequency modulation to avoid outer parameters a. Nevertheless sometimes one can extract the modulation from the perturbation. This trick was successfully used in a number of works, starting (as far as we know) with Arnold's paper [A3] devoted to Hamiltonian systems with proper degeneration (see also [AKN]); Pöschel [P1] used the trick in his investigations of lower-dimensional tori, Wayne [W1] used similar approach to prove the existence of quasiperiodic in time solutions of nonlinear string equation with a random potential. Now we turn to its discussion.

For the sake of symplicity we restrict ourselves to the perturbations of the form $H = H^3 + H^4$ with homogenous of order j functions H^j, $j = 3, 4$. Let us pass to the variables (13). Then the perturbed hamiltonian is $\mathcal{H}_\varepsilon = \mathcal{H}_0 + \varepsilon H_1$ with

$$H_1 = H^0(q,p) + \langle H^1(q,p), y\rangle + \frac{1}{2}\langle H^2(q,p)y, y\rangle + O(\|y\|^3)\,.$$

Here H^1 is a vector in Y and H^2 is a selfadjoint operator. So

$$\mathcal{H}_\varepsilon = h(p) + \frac{1}{2}\langle Ay, y\rangle + \varepsilon\big[H^0(q,p) + \langle H^1(q,p), y\rangle + \frac{1}{2}\langle H^2(q,p)y, y\rangle + O(\|y\|^3)\big]\,.$$

It is known since Birkhoff that with the help of a formally-analytic symplectic change of variables $\mathcal{H}_\varepsilon$ may be put into a partial normal form as follows:

$$\mathcal{H}_\varepsilon = h^1(p) + \frac{1}{2}\langle A^1(p)y, y\rangle + \varepsilon^2 H_\Delta(q,p,y) + \varepsilon O(\|y\|^3)\,.^{6)} \tag{28}$$

Here $h^1(p) = h(p) + \varepsilon \bar{H}^0(p)$ (the bar means the averaging over $q \in \mathrm{T}^n$) and $A^1 = A + \varepsilon A_\Delta(p)$ with some operator $A_\Delta(p)$ constructed in terms of the operator $\bar{H}^2(p)$. The function (28) is of the same form as (26) and in general the assumption (27) is fulfilled for the function $h^1(p)$.

The natural parameter $\omega = \nabla h^1(p)$ varies now in a domain of a small diameter δ_a, $\delta_a \sim \varepsilon$ (the perturbation is much smaller — of order ε^2). So Theorem 1 can not be directly applied to the equation (28). To handle this class of problems we state in Part 3 of the book Theorem 3.1.2, devoted to the equation (10) written in the variables (13), with the set $\mathfrak{A}$ equal to the ball of a small radius δ_a. The theorem states that the assertions of Theorem 1 remain true if

$$\varepsilon H_a = \varepsilon^{1+\mu} H_{1a}(q,p,y) + \delta_a H_a^3(q,p,y)\,,$$

$$H_a^3 = O\big(|p - I|^2 + \|y\|^3 + \|y\|^2\,|p - I|\big)\,,$$

where $\mu > 0$ and $1 \geq \delta_a \geq C\varepsilon$. (For the exact statement of the result see Part 3.1.)

The application of this result with $\delta_a = \varepsilon$, $\mu = 1$ to the equation with the hamiltonian (28) proves persistence most of the invariant n-tori $T^n(p)$ of the initial

---

6) One can achieve this normal form by formal applying to H_ε the transformation S_0 from Part 3.2 (the "KAM-step" of our proof), taking for granted that all the involved series converge.

linear system, provided that the transformation to the partial normal form converges and the nondegeneracy assumption

$$\operatorname{Hess} h_1(p) \not\equiv 0$$

holds together with the nonresonance one (the assumption 3) of Theorem 1, where $\lambda = \tilde{\lambda}$, $(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n)(p) = \nabla h_1(p)$ and $\{\tilde{\lambda}_{n+1}(p), \tilde{\lambda}_{n+2}(p), \ldots\}$ is the spectrum of the operator $A^1(p)$). Below we call these assumptions *nonlinear*, because they reflect nonlinear nature of the equation with the hamiltonian (28) (the frequencies $\{\tilde{\lambda}_j\}$ of oscillations depend on their amplitude-vector p).

The exact formulae, which can be constructed as in [P1], or, due to the last footnote, can be extracted from the proof of Theorem 1 (see below Part 3.2 with $m = 0$ and Part 3.8), show that the normal-form transformation is defined as a series with some regular numerators and with denominators of the form $D(s) = s_1\lambda_1 + \cdots + s_N\lambda_N$. Here N is an arbitrary natural number $\geq n+1$ and

$$3 \leq |s_1| + \cdots + |s_N| \leq 4\,, \quad s_N \neq 0\,, \quad |s_{n+1}| + \cdots + |s_N| \leq 2\,. \tag{29}$$

So if

$$|D(s)| \geq C^{-1} \tag{30}$$

for all s as above, then the normal-form transformation converges.

The condition (30) is not very restrictive because it holds for typical sequences $\{\lambda_j\}$ satisfying assumption 2) of Theorem 1. Now we check this statement for $d_1 > 1$. To do it let us take some $N \geq n+1$. Then

$$\begin{aligned} |D(s)| &\geq |\lambda_{n+1}s_{n+1} + \cdots + \lambda_N s_N| - |\lambda_1 s_1 + \cdots + \lambda_n s_n| \\ &\geq |\lambda_N - \lambda_{N-1}| - 3\max\{|\lambda_j| \mid 1 \leq j \leq n\} \geq C_1 N^{d_1 - 1} - C_2\,. \end{aligned}$$

So (30) holds with $C = 1$ if N is greater than some N_0. Therefore the inequality (30) holds if $D(s) \neq 0$ for the finite set of resonance relations consisting of all admissible relations with $N \leq N_0$ (one can choose C^{-1} equal to $\min\{1, \min\{|D(s)| \mid N \leq N_0\}\}$).

Thus, one can guarantee the convergence of the normal-form transformation for a parameter-dependent system for most values of the parameter. Due to simplicity of the involved resonance relations, often it is sufficient to have a *one-dimensional* parameter to obtain the convergence for *any fixed* $n \geq 1$.

The scheme we have just explained is applicable to study parameter-depending perturbed equation (10), if we

1) take away a set $\mathfrak{A}_c$ of parameters $a \in \mathfrak{A}$, violating the estimate (30) for some s as in (29) (this set is small, if C is large enough) and transform the equation to the normal form (28);

2) check the nonlinear nondegeneracy and nonresonance assumptions for $\{\tilde{\lambda}_j(p)\}$ and apply Theorem 3.1.2, treating $\omega = \nabla h^1(p)$ as the parameter.

The advantage of this approach to prove persistence of invariant n-tori of the linear system $(10)|_{\varepsilon=0}$ is that the set $\mathfrak{A}_c$ of "bad" parameters a is now an I- and

ε-independent set, which is relatively under the control. The disadvantage is the necessity to check the nonlinear nondegeneracy and the nonresonance assumptions for the "averaged" spectrum $\{\tilde{\lambda}_j(p)\}$, in addition to the ones for the initial spectrum $\{\lambda_j(a)\}$ which we need to make the first step.

This approach is applicable to study nonlinear Schrödinger and nonlinear string equations we discussed in Examples 1, 4 above. It is not difficult to check that the nonlinear nondegeneracy and nonresonance assumptions hold e.g., if in (25) the nonlinear term $-\varepsilon\varphi_w$ is equal to $-\varepsilon w^3$. The existence of time-quasiperiodic solutions of the equation

$$\ddot{w} = w_{xx} - V(x)w - \varepsilon w^3 , \quad w(t,0) \equiv w(t,\pi) \equiv 0 , \tag{31}$$

with the potential $V(x)$ lying outside a small set of "bad" potentials was obtained in a similar way by C.E. Wayne. In his paper [W1] the set of all potentials is given some Gaussian measure and the set of "bad" potentials is constructed as its small-measure subset.

In fact, the infinite-dimensional parameter $V(x)$, used in [W1], is much excessive: the scheme given above is applicable to (31) with $V(x) \equiv m \in \mathbb{R}^+$. It allows to prove existence of time-quasiperiodic solutions for most "masses" $m > 0$.

Remark. If the quadratic hamiltonian $\mathcal{H}_0$ is perturbed by higher-order terms starting from fourth order, then $\mathcal{H}_\varepsilon(z) = \mathcal{H}_0(z) + H_4(z) + H_5(z) + \cdots$. To study small-amplitude solutions of the corresponding Hamiltonian equation one can rescale $z = \mu u$, $\mu << 1$, obtain for u the equation with the hamiltonian $\tilde{\mathcal{H}}_\varepsilon(u) = \mathcal{H}_0(u) + \mu^2 H_4(u) + \mu^3 H_5(u) + \cdots$ and proceed exactly as above (with $\varepsilon = \mu^2$).

If the perturbation includes cubic terms, then the rescaled hamiltonian $\tilde{\mathcal{H}}_\varepsilon(u)$ contains the term $\mu H_3(u)$ which does not contribute to the function $h^1(p)$. Now the perturbation in (28) is larger than $\operatorname{Hess} h^1$ (the perturbation is of order μ and the Hessian — μ^2); so we can not use $\omega = \nabla h^1$ as a parameter to apply the theorem.

C) *On the integrable equations of mathematical physics*

One of the main achievements of mathematical physics during the last decades was the discovery of theta-integrable nonlinear partial differential equations (see e.g. [DEGN], [NMPZ]). Such equations are quasilinear Hamiltonian equations of the form (9). They possess invariant symplectic $2n$-dimensional manifolds $\mathcal{T}^{2n}$ such that the restriction of the system (9) on $\mathcal{T}^{2n}$ is integrable. So $\mathcal{T}^{2n}$ is symplectomorphic to $\mathbb{T}^n_q \times P_p$, $P \subset \mathbb{R}^n$, and in coordinates (q,p) the restriction of the system onto $\mathcal{T}^{2n}$ has the form

$$\dot{q} = \nabla h(p) , \quad \dot{p} = 0 .$$

Therefore $\mathcal{T}^{2n}$ is foliated into invariant n-tori $T^n(p) = \{(q,p) | p = \text{const}\}$ filled with quasiperiodic solutions $u_0(t) = \big(q + t\nabla h(p), p\big)$. The question is if the tori $T^n(p)$ survive under Hamiltonian perturbations of the equation. To formulate the corresponding result we have to consider variational equations about the solutions $u_0(t)$:

$$\dot{v} = J\Big(\nabla \mathcal{K}\big(u_0(t)\big)\Big)_* v$$

and to suppose that these equations are reducible to constant coefficient linear equations by means of a quasi-periodic substitutions $v = B(t,p)V$ (B is linear operator in Z quasiperiodically depending on t). It is proved (see [K3], [K5], [K8]) that under the reducibility assumption the quasilinear equation (9) near the manifold $\mathcal{T}^{2n}$ may be written in the form (14) with

$$\mathcal{H} = h(p) + \frac{1}{2}\langle A(p)y, y\rangle + O(\|y\|^3)\,.$$

A perturbed equation under this reduction takes exactly the form (26). So as in item A) one can prove that in a nondegenerate situation most of the tori $T^n(p)$ persist under perturbations.

The integrable nonlinear PDE's, linearized about their time-quasiperiodic solutions, *are* reducible to constant-coefficient equations. So Theorem 1 is applicable to study their perturbations. For an exact realization of this scheme for a perturbed Korteweg–de Vries and Sine–Gordon equations see [K5] and [BiK], [BoK].

See [K7] for a more detailed discussion of this group of applications of Theorem 1.

D) *Some remarks on multidimensional problems*

The most restrictive for applications among the assumptions of Theorem 1 is the assumption 2) (spectral asymptotics). As we have seen, this assumption holds for the differential equations with x-variable in a finite segment (or in the whole real line if the potential of the equation grows at infinity fast enough). The assumption 2) may be somewhat weakened with the same proof being applicable (see Remark 7 in Part 1.2). However, to carry out the proof the "separation condition"

$$\inf_{j\neq k} |\lambda_j(a) - \lambda_k(a)| \geq \delta > 0 \qquad (*)$$

must be fulfilled (possibly, under some additional restriction one could replace $(*)$ by the somewhat weaker assumption

$$|\lambda_j(a) - \lambda_k(a)| \geq \delta \max(j,k)^{-m} \quad \forall j \neq k\,, \qquad (**)$$

with some "not too large"positive m).

We do not know any example where these conditions hold for a differential equation with multidimensional x. Conversely, $(*)$ and $(**)$ do not hold if the quantization arguments can be applied to construct quasimodes of the equation ([Laz], [GW]). In particular, $(**)$ does not hold for the spectrum of the Dirichlet problem for the Laplace operator in a bounded convex two-dimensional domain with an analytic boundary [Laz].

Thus the nonlinear hamiltonian PDE's with x-variable in a segment form the distinguished class of equations with regular behavior of typical small-amplitude solutions.

4. Remarks on averaging theorems

Above we have proposed as an explanation for the recurrence effect of the FPU-type in partial differential equations of Hamiltonian form the theorem on persistence most of quasiperiodic solutions under Hamiltonian perturbations. It is well understood however that the long-time regular behavior of solutions may be explained by means of averaging theorems as well. In a finite-dimensional situation *Nekhoroshev's theorem* (see [N], [BGG], [Lo], [P4]) suites this purpose very well. For infinite-dimensional systems with discrete spectrum versions of this result are known only for systems with short range interactions ([W3], [BFG]). We are rather sceptical that there exists a version of Nekhoroshev's theorem applicable to nearly-integrable nonlinear partial differential equation.

First-order averaging theorems of Krylov–Bogolyubov type hold for a wide class of finite-dimensional systems. In the infinite-dimensional situation similar results are proven for lower-frequency initial data only (but for multidimensional in x equations also, see [Kri], [K3], [K4]). It is an open question if a first order averaging theorem for solutions of nearly integrable PDE's can be proven without this restriction.

5. Remarks on nearly integrable symplectomorphisms

Instead of differential equations (9) one can consider a "discrete-time equation" in the same infinite-dimensional phase-space $\big(Z, \alpha = -\langle J^{-1}dz, dz\rangle\big)$, i.e., a symplectic map

$$S : Z \longrightarrow Z\ , \quad S^{*}\alpha = \alpha\ . \tag{32}$$

The same phenomenon of pathologically regular behavior trajectories of nearly integrable sustem (32) (=iterations of the map S) can be observed; and the same question whether this phenomenon can be explained by existence many of finite-dimensional invariant tori of the map S appears.

A discrete-time theory parallel to the one for continuous-time systems we have discussed, can be developed. Fortunately, this work should not be done anew because of the following

Interpolation theorem. In the extended phase-space

$$\Big(Z \times \big\{(x,y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 2\big\}\ ,\ \alpha \oplus dx \wedge dy\Big)$$

one can find a nearly integrable analytic Hamiltonian vector field with a hamiltonian $H(z,x,y)$ such that its isoenergetic Poincaré map with respect to the manifold $\{y = 0\} \cap \{H = \text{const}\}$ is conjugated with S. In particular, if S is closed to the linear symplectic map $\exp JA$, then H is close to $\frac{1}{2}\langle Az, z\rangle + \pi\big(|x|^2 + |y|^2\big)$.

A possible reformulation of the result is that *the nearly integrable analytic symplectic map S is conjugated with time-one shift along trajectories of analytic 1-periodic time-dependent Hamiltonian vector-field, close to an autonomous integrable one.*

So S inherits invariant finite-dimensional tori of the interpolating nearly integrable vector-field.

As far as we see, the constructive proof of a finite-dimensional version of this result, given in [K9] (see also [KP]), is also applicable in the infinite-dimensional setting. We did not include into [K9] an infinite-dimensional interpolation theorem mostly because we are not aware of any concrete infinite-dimensional symplectomorphism S of physical interest.

6. Notations

The list of notations we use is given at the end of the book. As usual, we refer to formula (2.3) form Part 1 as (1.2.3), if we are outside Part 1; we refer to Chapter 3.2 of Part 3 as to §2, if we are inside Part 3.

Acknowledgements

I am grateful to V.I. Arnold, V.F. Lazutkin and M.I. Vishik for useful discussions during the writing of the papers on which this book is based. I am indebted to Jürgen Pöschel for helpful remarks and friendly discussions during the preparation of the book; to M.B. Sevrjuk and Gene Wayne for pointing out a lot of misprints and mistakes; to Jürgen Moser and Eduard Zehnder for encouragement. I am much obliged to my wife Julia for encouragement and for typing the original Russian version of *all* my papers.

The draft of this text was written while I was visiting the Max–Planck–Institut für Mathematik in Bonn and was originally printed as the preprint of the MPI [K6]. The text was essentially revised and extended during my staying at the Forschungsinstitut für Mathematik (ETH, Zürich). I am thankful to both institutes for their warm hospitality, and to Rahel Boller (ETH) for her quality typing the text and her patience in carrying out countless corrections and improvements.