A KAM–THEOREM FOR EQUATIONS OF THE KORTEWEG – DE VRIES TYPE

Sergei B. Kuksin

Abstract. We study quasilinear Hamiltonian partial differential equations with one-dimensional space variable in a segment of real line. We assume that the equation has a family of n-frequency time-quasiperiodic solutions, depending on an n-dimensional amplitude vector, and prove that most of these solutions persist under Hamiltonian perturbations of the equation by a nonlinear term which contains less derivatives than the linear part of the unperturbed equation. The result is similar to one proved in [K] for perturbations which contain no derivatives.

Introduction

In this paper we are concerned with quasilinear Hamiltonian partial differential equations with nonlinearities depending on derivatives. We study the equations which are close to a linear equation or to an integrable one. As good examples, let us consider the perturbed Korteweg-de Vries (KdV) equation

$$\dot u(t,x)=\frac{\partial}{\partial x}\,(-u_{xx}+6u^2+\varepsilon^2 f(u,x)),\quad x\in S^1=\mathbb{R}/2\pi\,\mathbb{Z}, \tag{1}$$

and a perturbation of the linear equation, similar to KdV:

$$\dot u(t,x)=\frac{\partial}{\partial x}\,(-u_{xx}+V(x)u+\varepsilon\, f(u,x)),\ x\in S^1. \tag{2}$$

Both (1) and (2) become infinite-dimensional Hamiltonian systems if we consider them as dynamical systems in a space of periodic functions $\{u(x)\}$ with zero mean value:

$$\int_0^{2\pi} u\ dx\equiv 0. \tag{3}$$

This restriction is correct for both equations since for their solutions the mean values in x are time-independent quantities.

Consider the differential operator

$$u(x)\longmapsto \frac{\partial}{\partial x}\,(-u_{xx}+V(x)u) \tag{4}$$

The author was supported by the Russian Fund of Fundamental Researches, grant 95-01-00609a

in the space of square-integrable functions with zero mean value. For $V = 0$ the operator has complete system of eigenfunctions $\exp ikx$ with the eigenvalues $ik^3/4$, $k = \pm 1, \pm 2, \dots$. If the potential $V(x)$ is not too large, i.e. if

$$|V|_{C^1} \le V_0, \tag{5}$$

with an appropriate positive constant V_0, then the operator (4) also has single pure imaginary spectrum $\pm i\lambda_j(V)$ corresponding to complex eigenfunctions $w_j(x)$, $j = \pm 1, \pm 2 \dots$. These eigenfunctions have the form

$$w_j(x) = \varphi^+_{|j|}(x) + i \operatorname{sgn} j \, \varphi^-_{|j|}(x),$$

where $\varphi^\pm_j(x)$ are some real functions (so $w_{-j}(x) = \overline{w}_j(x)$). Besides, the functions $w_j(x)$ and $w_{-j}(x)$ are asymptotically close to the exponents $(j\pi)^{-1/2}\exp ijx$ and $(j\pi)^{-1/2}\exp - ijx$.[1] Accordingly, solutions of $(2)|_{\varepsilon=0}$, (3) can be written as

$$\sum_{k=1}^{\infty} a_k \, e^{i\lambda_k(V)t} \, w_k(x) + \text{c.c.}, \quad a_k \in \mathbb{C}. \tag{6}$$

In particular, all solutions are almost periodic in time.[2] The solutions

$$\sum_{k=1}^{n} a_k \, e^{i\lambda_k t} \, w_k(x) + \text{c.c.} \tag{7}$$

with $n = 1, 2, \dots$ are time-quasiperiodic and jointly are dense in the function space.

It is well-understood now that solutions of the integrable equation $(1)|_{\varepsilon=0}$, (3) look similarly: for each $n \ge 1$ the equation has so-called n-gap solutions which can be written as

$$u_n(t, x) = \Phi_n(q + \omega(p)t; \, p)\,(x). \tag{8}$$

Here the analytic function $\Phi_n(q; p)\,(x)$ is 2π-periodic in the n-dimensional variable q (i.e., $q \in \mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$), and $p \in P \subset \mathbb{R}^n$ is an n-dimensional parameter ((8) is a rough version of the Its-Matveev formula, see [DNM]). To see that solutions (7) are analogous to (8), we write in (7) a_k as

$$a_k = \sqrt{p_k} \, e^{iq_k}, \; p_k \ge 0, \; q_k \in S^1 = \mathbb{R}/2\pi\mathbb{Z}. \tag{9}$$

Now the solution (7) can be written as $2 \operatorname{Re} \sum_{k=1}^{n} p_k \, e^{i(q_k + \lambda_k t)} \, w_k(x)$ and an analogy is obvious.

It is also true that all solutions of $(1)|_{\varepsilon=0}$, (3) are almost periodic in time, but corresponding analogy with the formula (6) is less transparent (see [McT, Ka]).

---

[1] This follows from the classical perturbation theory for discrete spectrum, see e.g. [RS]. The normilizing factor $(j\pi)^{-1/2}$ is convenient here since exactly the functions $(j\pi)^{-1/2}\cos jx$, $(j\pi)^{-1/2}\sin jx$ jointly form a Darboux basis for the symplectic structure corresponding to the KdV equation, which defines the skew-product of two functions u_1, u_2 with zero mean value as $\int_0^{2\pi} u_1(x)(\partial/\partial x)^{-1} u_2(x)\, dx$.

[2] If we reject the assumption (5), then some finite system of eigenvalues λ_j could get nontrivial imaginary parts. Accordingly, some terms in (6) could grow exponentially with time – this is not the phenomenon we are interested in this work.

We can take for the phase space $Z = Z_d$ of the equation (2), (3) some Sobolev space $H_0^d(S^1)$ of 2π-periodic functions with zero mean value, $d \geq 2$, and decompose it as

$$Z = E^{2n} \oplus Y,$$

where

$$E^{2n} = \operatorname{span}(\varphi_1^{\pm}, \dots, \varphi_n^{\pm}), \quad Y = \overline{\operatorname{span}}(\varphi_{n+1}^{\pm}, \varphi_{n+2}^{\pm}, \dots).$$

Let us denote by $y_k^{\pm}$, $k \geq n+1$, the coefficients of decomposition of a vector $y \in Y$ in the basis $\{\varphi_{n+1}^{\pm}, \varphi_{n+2}^{\pm}, \dots\}$. Then

$$(p, q, y), \ p \in \mathbb{R}_+^n, \ q \in \mathbb{T}^n, \ y = (y_{n+1}^{\pm}, y_{n+2}^{\pm}, \dots)$$

with (p, q) as in (9), form a coordinate system in $Z = E^{2n} \oplus Y$. Since the equation (2), (3) is Hamiltonian, then – under the requirement of a proper normalization of the eigenfunctions $w_k^{\pm} = \varphi_k^+ \pm i\varphi_k^-$ – in the coordinates (p, q, y) the equation (2), (3) takes the Hamiltonian form

$$\dot p_j = -\partial \mathcal{H}/\partial q_j, \quad \dot q_j = \partial \mathcal{H}/2p_j, \quad \dot y_m^{\pm} = \mp \partial \mathcal{H}/\partial y_m^{\mp}, \tag{10}$$

where

$$\mathcal{H} = \lambda_1 p_1 + \dots + \lambda_n p_n + \frac{1}{2} \sum_{j=n+1}^{\infty} \lambda_j \, |y_j|^2 + \varepsilon \, H(p, q, y), \ \ |y_j|^2 = y_j^{+^2} + y_j^{-2}, \tag{11}$$

and εH is the Hamiltonian of the perturbation $\varepsilon \frac{\partial}{\partial x} f(x, u)$. On this half-obvious matter see in [K] Part 2.7 and Parts 2.3, 2.6 (where similar examples of nonlinear Schrödinger and wave equations are discussed).

Since the perturbation in (2) is a first-order nonlinear differential operator, then the map

$$y = \sum_{j=n+1}^{\infty} y_m^{\pm} \varphi_m^{\pm} \longmapsto \sum_{j=n+1}^{\infty} \mp \left(\frac{\partial}{\partial y_m^{\mp}} H \right) \varphi_m^{\pm} \tag{12}$$

is smooth as a map $Y_d \longrightarrow Y_{d-1}$ (the spaces Y_r are given norms induced from the Sobolev spaces $H_0^r(S^1)$). This map *is unbounded* as a map in Y_d.

It is remarkable that the equation (1), (3) near the 2n-dimensional submanifold $\mathcal{T}^{2n} \subset Z$,

$$\mathcal{T}^{2n} = \cup \{ \Phi_n(q; p)(\cdot) \,|\, q \in \mathbb{T}^n, \ p \in P \},$$

can be put to a form similar to (10), (11): in [K2] we proved that if for the p-variables in (8) we chose the actions of the integrable system defined by (1), (3) on $\mathcal{T}^{2n}$, then the action-angle coordinates (p, q) in $\mathcal{T}^{2n}$ can be supplemented by an infinite-dimensional coordinate $y = (y_{n+1}^{\pm}, y_{n+2}^{\pm}, \dots)$ in a subspace transversal to $\mathcal{T}^{2n}$ in Z (which can be identified with the space Y as above) in such a way that (p, q, y) form a coordinate system in the vicinity of $\mathcal{T}^{2n}$ in Z. In these coordinates the equations (1), (3) take the form (10) with

$$\mathcal{H} = h(p) + \frac{1}{2} \sum_{j=n+1}^{\infty} \lambda_j(p) \, |y_j|^2 + \varepsilon^2 H_1(p, q, y) + H_3(p, q, y), \tag{13}$$

where $H_3 = O(\|y\|^3)$ and the maps (12) with $H = H_1$ and $H = H_3$ both are smooth as maps $Y_d \longrightarrow Y_{d-1}$. The functional $h(p)$ is the Hamiltonian for the KdV equation $(1)|_{\varepsilon=0}$ restricted to $\mathcal{T}^{2n}$, so $\nabla h(p) = \omega(p)$ (cf. (8)).

In (10), (13) we make the substitution

$$q = \tilde{q}, \ \ p = \xi + \varepsilon \tilde{p}, \ \ y = \sqrt{\varepsilon}\, \tilde{y},$$

where $\xi \in P$ is a parameter of the substitution. In the tilde-variables the equation takes the form (10) with

$$\mathcal{H}(\tilde{p}, \tilde{q}, \tilde{y};\, \xi) = \omega(\xi) \cdot \tilde{p} + \frac{1}{2} \sum_{j=n+1}^{\infty} \lambda_j(\xi)\, |\tilde{y}_j|^2 + \varepsilon\, H(\tilde{p}, \tilde{q}, \tilde{y}; \xi,\, \varepsilon). \tag{14}$$

Due to an elegant result of I. Krichever (see in [BiK1]), the map $\xi \longmapsto \nabla h(\xi) = \omega$ is a local diffeomorphism (i.e., its determinant never vanishes). So in the Hamiltonian (14) we can pass from parameter $\xi \in P$ to $\omega \in \Omega = \nabla h(P)$.

In the tilde-variables the n-gap solutions (8) of $(1)|_{\varepsilon=0}$ with $p = \xi$ take the form

$$\tilde{p} = 0, \ \ \tilde{q} = q_0 + \omega\, t, \ \ \tilde{y} = 0. \tag{15}$$

Similarly with solutions (7) of (10), (11): after the substitution $p = \xi + \tilde{p}$, $q = \tilde{q}$, $y = \tilde{y}$ they take the same form with $\omega = (\lambda_1, \dots, \lambda_n)$. Let us denote by T_0^n the n-torus $\{(0, q, 0) \,|\, q \in \mathbb{T}^n\}$ filled by solutions (15).

The main result of this work – Theorem 1 from Chapter 1 – implies that for most ω the solution (15) persists in equation (10), (14) with sufficiently small $\varepsilon > 0$. Here we give a simplified version of the result.

Theorem A. *Suppose that in (14)*

1) $\lambda_j - K j^{d_1} = o(j^{d_1})$, $|\lambda_j - \lambda_{j-1}| \geq C^{-1} j^{d_1 - 1}$, $|\partial \lambda_j / \partial \omega| \leq C\, j^{d_H}$ *with some* $d_1 > 1$, $d_H < d_1 - 1$;
2) *the map (12) defines a smooth analytic map* $Y_d \longrightarrow Y_{d - d_H}$*, which smoothly depends on the parameter* ω.

Then there exist finite numbers M *and* j_1 *such that if for all* ω *in some subdomain* $\Omega_0 \subset \Omega$ *one has*

$$|s \cdot \omega + \ell_1 \lambda_{n+1}(\omega) + \ldots + \ell_{j_1} \lambda_{n+j_1}(\omega)| \geq C^{-1} > 0 \tag{16}$$

for all integer n*-vectors* s *and* j_1*-vectors* ℓ *such that* $|s| \leq M$, $1 \leq |\ell| \leq 2$, *then there exists a subset* $\Omega_\varepsilon \subset \Omega_0$,

$$mes\,(\Omega_0 \backslash \Omega_\varepsilon) \longrightarrow 0 \quad (\varepsilon \longrightarrow 0),$$

such that for ω *in* Ω_ε *the equation (10), (14) has an invariant* n*-torus* $T_\varepsilon^n \subset Z$,

$$dist\,(T_\varepsilon^n, T_0^n) \leq C\, \varepsilon^{\kappa}, \ \ \kappa > 0,$$

which is filled with time-quasiperiodic solutions with zero Lyapunov exponents.

The numbers j_1, M depend only on the constants which characterize the perturbation H and the asymptotics for λ_j. Thus if the frequencies λ_j are analytic in ω, then the assumption (16) can be replaced by

$$(16') \qquad \begin{gathered} \Lambda_{s,\ell}(\omega) := s\cdot\omega + \ell_1\lambda_{n+1}(\omega) + \ldots + \ell_{j_1}\lambda_{n+j_1}(\omega) \not\equiv 0 \\ \forall\, |s| \le M, \quad 1 \le |\ell| \le 2, \end{gathered}$$

since (16′) implies (16) for all ω outside a small neighborhood of the union of zero-sets of the analytic functions $\Lambda_{s,\ell}$ as in (16′).

Example 1. Let us take equation (2) where the potential $V(x)$ as in (5) analytically (e.g., linearly) depends on an n-dimensional parameter ξ from a ball B_r and f is sufficiently smooth in (u,x),[3] analytic in u. For generic families of potentials $V(x,\xi)$ the map $B_r \ni \xi \longmapsto \omega = (\lambda_1,\ldots,\lambda_n)$ is an analytic diffeomorphism and the functions $\Lambda_{s,\ell}(\omega(\xi))$ are not identically zero (cf. [K]). So generically equation (2), (3) with $V = V(x,\xi)$ and sufficiently small ε is such that for most ξ it has time-quasiperiodic solutions which form solenoids in invariant tori of (2), (3) close to those of the linear equation $(2)|_{\varepsilon=0}$.

Let ω_0 in (8) be the limiting value of the frequency vector ω corresponding to the zero solution of (1).

Example 2. Let us take the perturbed KdV equation (1), where f is sufficiently smooth in (u,x) and analytic in u. Direct evaluations of $\Lambda_{s,\ell}(\omega_0)$ and $\nabla_\omega\Lambda_{s,\ell}(\omega_0)$, carried out in [BoK1], show that (16′) hold for all s and ℓ such that $1 \le |\ell| \le 2$. Therefore most (in the measure sense) solutions (8) persist in equation (1), (3) with sufficiently small ε.

Due to a well-known result of V. Marčenko, the union of all finite-gap manifolds $\mathcal{T}^{2n}$ is dense in each space $Z = Z_d$. Each manifold $\mathcal{T}^{2n}$ "mostly persists" in the perturbed equation (1), (3) when $\varepsilon \longrightarrow 0$. So if QP_ε is the union of all time-quasiperiodic trajectories of (1), (3) with zero Lyapunov exponents (treated as curves in the phase-space Z), then for any fixed $\mathfrak{z} \in Z$,

$$\mathrm{dist}(\mathfrak{z}, QP_\varepsilon) \longrightarrow 0 \ \text{ as } \ \varepsilon \longrightarrow 0.$$

Thus, stable time-quasiperiodic solutions of (1), (3) jointly form in Z a web, asymptotically dense in Z as $\varepsilon \longrightarrow 0$.

Remark. It is not very natural that an ε^2-perturbation of equation (1) implies a deformation of the invariant torus T_0^n of order ε^κ with some possibly small $\kappa > 0$. In fact, the deformation is of order ε^ρ for any $\rho < 2$. To see this, one should apply directly to Hamiltonian (13) Amplification 3 to Theorem 1 from Chapter 1 below (cf. in [K], Part 2.2 and item 3.2.C of the Introduction).

Example 3. Theorem A is also applicable to study perturbations of higher equations from the KdV hierarchy [DMN, McT]. Take, for example, the second equation:

$$(17) \qquad \dot u = \frac{\partial}{\partial x}\,\frac{\delta I_2}{\delta u(x)} = \frac{\partial}{\partial x}\,\Big(u_{xxxx} - 5\,uu_xu_{xx} - \frac{5}{2}\,u^2\,u_{xxx} + 10u^3\Big).$$

[3] H^2-smoothness in x is sufficient

The same functions Φ_n as in (8) define finite-gap solutions of (17), (3) if we replace $\omega(p)$ by a suitable frequency vector $\omega_2(p)$. Moreover, the same coordinate system (p,q,y) in the vicinity of $\mathcal{T}^{2n}$ reduce (17) to the equation (10) with

$$\mathcal{H} = \mathcal{H}_2 := h^2(p) + \frac{1}{2}\sum_{j=n+1}^{\infty} \lambda_j^2(p)\,|y_j|^2 + H_3^2(p,q,y), \quad H_3^2 = O(\|y\|^3).$$

Take any analytic in u, u_x and sufficiently smooth in x function $g(x,u,u_x)$. Then the equation

$$\dot{u} = \frac{\partial}{\partial x}\left(\frac{\delta I_2}{\delta u_x} + \varepsilon\,(g'_u(x,u,x_x) - \frac{\partial}{\partial x}\,g'_{u_x}(x,u,u_x))\right) \tag{18}$$

is a Hamiltonian perturbation of (17) by means of a third order nonlinear operator. In the same way as in Example 2, Theorem A (with $d_1 = 5$, $d_H = 3$) implies that most of finite-gap solutions of (17) persist in (18) with ε sufficiently small.

The assertion of Example 2 is a result of the paper [K2], where the nonresonance relations $(16')$ were taken for granted and it was claimed that the proof of Theorem A given in [K1] for the case $d_H \le 0$ (and $d_1 \ge 1$) also is applicable to prove the result for $d_H > 0$. In this paper we finally pay our delayed debt and present the proof. We use the scheme of the works [K,K1] and profit from the simplifying assumption $d_1 > 1$ (instead of $d_1 \ge 1$ in [K]). The only (but rather nontrivial) complication compare to the case $d_H \le 0$ arises when we solve the homological equations. Somewhat simplifying the problem, we can state it like that: for $j = 1,2,\dots$ we should solve the equations

$$-i(\omega\cdot\nabla)x(q) + \lambda_j\,x(q) + \beta_j(q)x(q) = z(q), \quad q \in \mathbb{T}^n, \tag{19_j}$$

where $\omega\cdot\nabla = \sum \omega_j\,\partial/\partial q_j$, $\beta_j(q) \sim \varepsilon\, j^{d_H}$, ω is a Diophantine n-vector and $z(q)$ is a given analytic function. We can find an analytic function $H_j(q)$ such that $(\omega\cdot\nabla)H_j = \beta_j$ and still $H_j(q) \sim \varepsilon\, j^{d_H}$. The substitution $x = \exp(-iH)y$ reduces (19_j) to the equation

$$-i(\omega\cdot\nabla)y + \lambda_j y = e^{iH_j}z =: \mathfrak{z}(q).$$

If $d_H \le 0$, then $\exp iH_j$ is a factor of order one and we can solve the last equation by decomposing $\mathfrak{z}(q)$ and $y(q)$ to Fourier series, see the Appendix below. But if d_H is positive, then the norm of $\mathfrak{z}(q)$ in a complex neighborhood of the torus grows exponentially with j. This exponential factor appears in Fourier coefficients of the solution $y(q)$ and – in a naive way – also in an estimate for the solution $x(q)$. But for our proof to work we need a *uniform* in j estimate for $x(q)$. We obtain this estimate in Chapter 4. The trick we use there is to approximate the vector ω by vectors $\widetilde{\omega}_\ell$, $\ell = 1,2,\dots$, of the form $\widetilde{\omega}_\ell = r_\ell/\nu_\ell$ where $r_\ell \in \mathbb{Z}^n$ and ν_ℓ is an appropriate real number. For ω replaced by $\tilde{\omega}_\ell$ we find representation for an approximate solution x_ℓ of the equation as a rapidly oscillating (when j grows) one-dimensional integral Fourier with a complex phase function. We show how to shift the contour of integrating to make the phase function real, which implies an estimate for the approximating solutions. This estimate turns out to be uniform in ℓ and implies a desirable j-independent estimate for the exact solution of (19_j).

After the difficulty with the homological equations is overcome, the proof goes like in [K,K1] and even simpler since, first, the complicated boundary case $d_1 = 1$ is now excluded from considerations and, second, for this work we found a simpler proof for the last step of our scheme ("transition to limit", Chapter 2.8).[4]

As we mentioned earlier, our proof follows the "KAM for PDEs" scheme, designed in [K1] and developed in [K]. Independently similar schemes were proposed by C. E. Wayne [W] to study quasiperiodic solutions of some nonlinear wave equations and by L. H. Eliasson [E] to study finite-dimensional hamiltonian systems;[5] also see [P] for some developments of this approach. Lately another KAM-scheme to prove persistence of time-periodic and time-quasiperiodic solutions of *linear* equations (10), (11) under nonlinear perturbations which *contain no derivatives* was proposed by Craig-Wayne [CW] and developed by J. Bourgain [Bou].

The scheme of [K1, K] was initially used to study nonlinear perturbations of linear equations, but it turned out to be a flexible tool to study time-quasiperiodic solutions of nonlinear PDEs: In [K2, BiK2] it was applied to study perturbations of integrable PDEs and in [BoK2, KP] – to study small solutions of nonlinear PDEs. The theorem we prove in this paper essentially extends the domain of applicability of our KAM for PDEs scheme.

Acknowledgement. I sincerely thank the Institute for Advanced Study (Princeton) for their hospitality where this work was done. Discussions with Jean Bourgain during preparation of this paper were very useful to me.

1. The problem and the result

Let Y be a real Hilbert space with the scalar products $\langle\cdot,\cdot\rangle$ and the (Hilbert) basis $\{\varphi_j^\pm \mid j \geq 1\}$ and let Y_s, $s\in\mathbb{R}$, be the Hilbert space with the basis $\{j^{-s}\varphi_j^\pm \mid j=1,2,\dots\}$. So Y_s is formed by vectors $y=\sum y_j^\pm\varphi_j^\pm$ with finite norm $\|y\|_s$, where

$$\|y\|_s^2=\sum j^{2s}\left({y_j^+}^2+{y_j^-}^2\right).$$

In particular, $Y=Y_0$.

Example. If Y is the space of square-summable functions on the segment $[0,2\pi]$ with zero mean value and

$$\varphi_j^+=\pi^{-1/2}\cos jx,\ \ \varphi_j^-=\pi^{-1/2}\sin jx,\ \ j=1,2,\dots,$$

then Y_s is the Sobolev space of 2π-periodic functions with zero mean value. For a natural s the norm in Y_s is given by the formula $\|y(x)\|_s^2=\int_0^{2\pi}|\partial^s y/\partial x^s|^2dx$.

We define the spaces

$$\mathcal{Y}_s=\mathbb{R}^n\times\mathbb{T}^n\times Y_s$$

and consider a neighborhood Q of the torus

$$T_0^n=\{0\}\times\mathbb{T}^n\times\{0\}$$

in $\mathcal{Y}_d$ of the form

$$Q=O(\delta,\mathbb{R}^n)\times\mathbb{T}^n\times O(\delta,Y_d),\quad \delta>0,$$

[4]Due to this simplification, we managed to drop the additional restriction imposed on the equation (10), (14) in [K2] – there we assumed a priori that the equation is correct locally in time.

[5]still, the schemes of [K1], [E] and [W] are sufficiently different; we use exactly the scheme from [K1].

where $d \geq 0$ will be fixed later and for a Banach space B we denote by $O(\delta, B)$ the δ-ball in B centered at the origin.

We denote by J the skew-symmetric operator in Y such that

$$J\varphi_j^{\pm} = \mp\varphi_j^{\mp}, \ \ j \geq 1,$$

and supply the spaces $\mathcal{Y}_s$ with the symplectic structure given by the 2-form

$$\omega_2 = dp \wedge dq + \langle Jdy, dy\rangle \quad (p \in \mathbb{R}^n, \ q \in \mathbb{T}^n, \ y \in Y).$$

Accordingly for two functionals H_1, H_2 on $\mathcal{Y}_s$ we define their Poisson bracket $\{H_1, H_2\}$ as the functional

$$\{H_1, H_2\} = -\nabla_q H_1 \cdot \nabla_p H_2 + \nabla_p H_1 \cdot \nabla_q H_2 + \langle J\nabla_y H_1, \nabla_y H_2\rangle, \tag{1.1}$$

where ∇_y is the gradient in y with respect to the scalar product $\langle\cdot,\cdot\rangle$. For a functional $\mathcal{H}$ we define the Hamiltonian equations with the Hamiltonian $\mathcal{H}$ as

$$\dot{p} = -\nabla_q\mathcal{H}, \quad \dot{q} = \nabla_p\mathcal{H}, \quad \dot{y} = J\nabla_y\mathcal{H}, \tag{1.2}$$

and abbreviate these equations as

$$\dot{\mathfrak{h}} = V_{\mathcal{H}}(\mathfrak{h}), \quad \mathfrak{h} = (p, q, y). \tag{1.2'}$$

If $\mathcal{H}$ is a smooth function on $Q \subset \mathcal{Y}_d$, then the map

$$Q \longrightarrow \mathbb{R}^n \times \mathbb{R}^n \times Y_{-d}, \ \ \mathfrak{h} = (p, q, y) \longmapsto V_{\mathcal{H}}(\mathfrak{h}) = (-\nabla_q\mathcal{H}, \ \nabla_p\mathcal{H}, \ J\nabla_y\mathcal{H})$$

is also smooth. We study strong solutions of (1.2) given by C^1-smooth curves $t \longmapsto \mathfrak{h}(t) \in Q$ such that (1.2′) holds in $\mathbb{R}^n \times \mathbb{R}^n \times Y_{-d}$. See more on equations (1.2) in [K].

We are concerned with Hamiltonians of the form

$$\mathcal{H}_\varepsilon = p \cdot \omega + \frac{1}{2}\ \langle A(\omega)y, y\rangle + \varepsilon H\,(\mathfrak{h};\, \omega, \varepsilon),$$

where $\omega \in \Omega \Subset \mathbb{R}^n$ is a vector-parameter, $A(\omega)$ is an unbounded selfadjoint operator in Y such that

$$A(\omega)\varphi_j^{\pm} = \lambda_j\ (\omega)\ \varphi_j^{\pm}, \quad j \geq 1,$$

and ε is a perturbation parameter, $0 \leq \varepsilon < 1$. All estimates for H will be valid uniformly in ε and dependence of H in ε will be neglected.

For $\varepsilon = 0$ the Hamiltonian vector field $V_{\mathcal{H}_0}$ defines the Hamiltonian equations

$$\dot{p} = 0, \ \ \dot{q} = \omega, \ \ \dot{y} = JAy. \tag{1.3}$$

The torus T_0^n is invariant for (1.3) and is filled with quasiperiodic trajectories $t \longmapsto (0, q_0+\omega t, 0)$. Our goal is to prove that for most values of the parameter $\omega \in \Omega$ the torus T_0^n persists in the equation with Hamiltonian $\mathcal{H}_\varepsilon$ if ε is sufficiently small, provided that the perturbation εH and the spectrum $\{\lambda_j\}$ meet some additional restrictions which we shall now discuss.

1.1 The restrictions.

Let us denote by Y_s^c the complexification of the Hilbert space Y_s and by $\mathcal{Y}_s^c$ the complexification of $\mathcal{Y}_s$,

$$\mathcal{Y}_s^c = \mathbb{C}^n \times (\mathbb{C}^n / 2\pi\mathbb{Z}^n) \times Y_s^c.$$

We also denote by $U(\delta)$ a complex neighbourhood of the torus $\mathbb{T}^n$,

$$U(\delta) = \{\xi \in \mathbb{C}^n / 2\pi\mathbb{Z}^n \bigm| |\operatorname{Im}\xi| < \delta\}$$

and by Q^c – a neighbourhood of the torus T_0^n in $\mathcal{Y}_s^c$:

$$Q^c = O(\delta, \mathbb{C}^n) \times U(\delta) \times O(\delta, Y_d^c).$$

We systematically use Lipschitz maps between metric spaces and denote by Lip their Lipschitz constants. For a map $f : M \to B$ where M is a metric space and B is Banach, we write

$$\|f\|_B^{M,\mathrm{Lip}} = \max\,(\mathrm{Lip}\, f,\ \sup_{m\in M} \|f(m)\|_B).$$

If B_1, B_2 are Banach, $O_1 \subset B_1$ and f maps $O_1 \times M$ to B_2, we write

$$\|f\|_{B_2}^{O_1,M} = \sup_{b_1\in O_1} \|f(b_1,\cdot)\|_{B_2}^{M,\mathrm{Lip}}.$$

Similar for a map $f : O_1 \longrightarrow B_2$ we denote by $\|f\|_{B_2}^{O_1}$ the supremum of its $\|\cdot\|_{B_2}$-norm.

Below we give the assumptions imposed on the spectrum $\{\lambda_j\}$ and on H. By $K_0, K_1, \dots$ we denote different positive constants.

1) The functions $\lambda_j(\omega)$ are Lipschitz and

$$\begin{cases} K_1^{-1} j^{d_1} - K_0 \le \lambda_j(\omega) \le K_1 j^{d_1} \quad \forall\omega, \quad \forall j, \\ |\lambda_j(\omega) - \lambda_k(\omega)| \ge K_1^{-1}\, |j^{d_1} - k^{d_1}| \quad \forall\omega, \quad \forall k, j, \\ \mathrm{Lip}\lambda_j \le K_1\, j^{d_H} \quad \forall j, \end{cases} \tag{1.4}$$

where $d_1 > 1$ and $d_H < d_1 - 1$;

2) $d \ge d_1/2$ and the function $H(\mathfrak{h};\omega)$, $\mathfrak{h} = (p, q, y)$, can be extended to a complex-analytic in $\mathfrak{h} \in Q^c \subset \mathcal{Y}_d^c$ function such that

$$|H|^{Q^c,\Omega} + \|\nabla_y H\|_{d-d_H}^{Q^c,\Omega} \le 1. \tag{1.5}$$

Under the restrictions (1.4), (1.5) we study Hamiltonian system with the Hamiltonian $\mathcal{H}_\varepsilon$:

$$\begin{cases} \dot p = -\varepsilon\nabla_q\ H(\mathfrak{h};\omega), \\ \dot q = \omega + \varepsilon\nabla_p\ H(\mathfrak{h};\omega), \\ \dot y = J(A(\omega)y + \varepsilon\nabla_y\ H(\mathfrak{h};\omega)). \end{cases} \tag{1.6}$$

As above we abbreviate these equations as

$$\dot{\mathfrak{h}} = V_{\mathcal{H}_\varepsilon}(\mathfrak{h}), \quad \mathfrak{h} = (p, q, y). \tag{1.6$'$}$$

1.2 The result.

Let us fix any $\rho < 1/3$.

Theorem 1. *Suppose the assumptions (1.4), (1.5) hold. Then there exist integer* j_1, M_1*, depending on*

$$n,\, d_1,\, d_H,\, K_0,\, K_1,\, K_2, \tag{1.7}$$

such that if

$$|s\cdot\omega + \ell_1\lambda_1(\omega) + \ldots + \ell_{j_1}\lambda_{j_1}(\omega)| \geq K_3 > 0 \tag{1.8}$$

for all $\omega \in \Omega$*, all integer* n*-vectors* s *and* j_1*-vectors* ℓ *such that*

$$|s| \leq M_1, \quad 1 \leq |\ell_1| + \ldots + |\ell_{j_1}| \leq 2,$$

then the n*-torus* T_0^n *persists in (1.6) in the following sense:*

For arbitrary $\gamma > 0$ *and for sufficiently small* $\varepsilon < \bar{\varepsilon} = \bar{\varepsilon}(\gamma)$ *there exists a Borel subset* $\Omega_\varepsilon \subset \Omega$ *and analytic embeddings*

$$\Sigma_\omega : \mathbb{T}^n \longrightarrow \mathcal{Y}_d, \quad \omega \in \Omega_\varepsilon,$$

with the following properties:

a) *mes*$(\Omega \backslash \Omega_\varepsilon) \leq \gamma$,
b) *the torus* $\Sigma_\omega(\mathbb{T}^n)$ *is* ε^ρ*-close to* T_0^n*, is invariant for the flow of (1.6) and is filled with quasiperiodic solutions of the form* $\mathfrak{h}_0(t) = \Sigma_\omega(q + \omega' t)$*, where* $|\omega' - \omega| < C\varepsilon^{1/3}$*.*

Amplification 1. The map $\Sigma : \mathbb{T}^n \times \Omega_\varepsilon \longrightarrow \mathcal{Y}_d$, $(q,\omega) \longmapsto \Sigma_\omega(q)$ is Lipschitz-close to the map $\Sigma^0 : (q,\omega) \longmapsto (0,q,0) \in \mathcal{Y}_d$, i.e. $\mathrm{dist}_d(\Sigma^0,\Sigma) \leq C\varepsilon^\rho$ and $\mathrm{Lip}(\Sigma^0 - \Sigma) \leq C\varepsilon^\rho$. Besides, the map $\omega \longmapsto \omega'$ is Lipschitz and $\mathrm{Lip}(\omega \longmapsto \omega' - \omega) \leq C\varepsilon^{1/3}$.

Amplification 2. Statements of Theorem 1 and Amplification 2 remain true with ρ replaced by one. Also $|\omega - \omega'| + \mathrm{Lip}(\omega' - \omega) \leq C\varepsilon$.

Amplification 3. Assertions of the theorem and of all the amplifications remain true for Hamiltonians $\mathcal{H}_\varepsilon$ of the form

$$\mathcal{H}_\varepsilon = p\cdot\omega + \frac{1}{2}\,\langle A(\omega)y, y\rangle + \varepsilon H(\mathfrak{h};\, \omega, \varepsilon) + H^3(\mathfrak{h};\, \omega, \varepsilon),$$

where A and H are as above and H^3 is an analytic in $\mathfrak{h} \in Q^c$ function such that

$$|H^3|^{\Omega,\mathrm{Lip}} \leq K^1(|p|^2 + |p|\, \|y\|_d + \|y\|_d^3), \quad \|\nabla_y H^3\|_{d-d_H}^{\Omega,\mathrm{Lip}} \leq K^1(|p| + \|y\|_d^2),$$

for anyh $\mathfrak{h}$ in Q^c.

The theorem and Amplification 1 are proven in the next Part 2. We skip the proofs of Amplifications 2 and 3 since they are identical with the proofs given in [K] for the case $d_H \leq 0$.

1.3 Linearized equations.

Let us consider the linearization of the equation (1.6′) about any solution $\mathfrak{h}(t)$:

$$\dot\eta - V_{\mathcal{H}_\varepsilon}(\mathfrak{h}(t))_*\eta = 0,$$

where

$$\eta(t) \in T_{\mathfrak{h}(t)}\mathcal{Y}_d \simeq Z_d = \mathbb{R}^n_p \times \mathbb{R}^n_q \times Y_d.$$

We say that the linearized equation is *well-posed* if for each its solution $\eta(t)$ we have:

$$\|\eta(t+\tau)\|_\theta \le C_1 e^{C_2|\tau|}\|\eta(t)\|_\theta \quad \text{for any} \ \ 0 \le \theta \le d, \tag{1.9}$$

where the constants C_1, C_2 do not depend on $\eta(\cdot)$ (we remark that using usual variation of constant one gets estimates similar to (1.9) for solutions of the equation with a non-zero right hand side, bounded in Z_θ).

We are concerned with linearizations of equation (1.6′) about solutions $\mathfrak{h}_0(t)$ constructed in Theorem 1,

$$\dot\eta = V_{\mathcal{H}_\varepsilon}(\mathfrak{h}_0(t))_*\eta. \tag{1.10}$$

Example 1. If $d_H \le 0$ and $\mathfrak{h}_0(t)$ is bounded in Z_d, then (1.10) is a bounded in Z_θ, $\theta \le d$, perturbation of the linear problem which defines a group of linear isometries of Z_θ. So (1.9) clearly holds.

Example 2. Take for (1.6) the perturbed Korteweg–de Vries equation (1) (being written in the coordinates as in (13) it has the form (1.6) with a (more general) perturbation as in Amplification 3). The linearized equation takes the form

$$\dot v = -v_{xxx} + 12\frac{\partial}{\partial x}u_0 v + \varepsilon^2\frac{\partial}{\partial x}(f'(u_0)v). \tag{1.11}$$

Let u_0 be any time-quasiperiodic solution of (1) in $H^d_0(S^1)$, smooth in t. Then using the equation we can express third space-derivatives of u_0 via its first time-space derivatives. So $u_0(t,\cdot)$ is bounded in $H^{d+2}_0(S^1)$. Multiplying (1.11) by $(-\Delta)^d v(t,x)$ and integrating over S^1 we get that

$$\frac{d}{dt}\|v(t,\cdot)\|_d^2 \le C\|v(t,\cdot)\|_d^2,$$

which implies (1.9).

Example 3. For perturbed higher equations from the KdV hierarchy (e.g., for (18)) everything is the same.

Theorem 2. *Suppose that the assumptions of Theorem 1 hold and $f(t)$ is a constructed in Theorem 1 time-quasiperiodic solution of (1.6) such that the linearized equation (1.10) is well-posed. Then each solution $\eta(t)$ of (1.10) meets the estimate*

$$\|\eta(t)\|_d \le (C_1 + C_2 t)\|\eta(0)\|_d.$$

In particular, all Lyapunov exponents of $\mathfrak{h}_0(t)$ vanish.

Amplification 4. The theorem's assertion remains true for more general perturbations as in Amplification 3.

The theorem is proven in Chapter 5; for the situation described in the amplification the proof remains the same.

2.1 Notations.

We use some additional notations. We introduce increasing sequence $\{e_j\}$, where $e_0 = 0$ and for $m \geq 1$ we set

$$e_m = (1^{-2} + \ldots + m^{-2})/K_*, \quad K_* = 2(1^{-2} + m^{-2} + \ldots)$$

(thus $e_m < 1/2$ for all m), and introduce two decreasing sequences $\{\varepsilon_m\}$ and $\{\delta_m\}$:

$$\varepsilon_m = \varepsilon^{(1+\rho)^m}, \quad \delta_m = \delta_0\,(1 - e(m)).$$

We denote $U_m = U(\delta_m)$ and consider complex neighborhoods O_m of the torus T_0^n of the form

$$O_m = O(\varepsilon_m^{2/3}, \mathbb{C}^n) \times U_m \times O(\varepsilon_m^{1/3}, Y_d^c).$$

Besides we define the intermediate numbers $\delta_m^j = \frac{6-j}{6}\,\delta_m + \frac{j}{6}\,\delta_{m+1}, \quad 0 \leq j \leq 5$, and the intermediate domains

$$O_m^j = O((2^{-j}\varepsilon_m)^{2/3}, \mathbb{C}^n) \times U(\delta_m^j) \times O((2^{-j}\varepsilon_m),^{1/3} Y_d^c).$$

If $\bar{\varepsilon} \ll 1$ (i.e., is sufficiently small) then

$$O_m \supset O_m^1 \supset \ldots \supset O_m^5 \supset O_{m+1} \supset \ldots \supset T_0^n.$$

By C, C_1 etc. we denote different positive constants independent from ε and m; by $C(m)$, $C_1(m)$ etc. – different functions of m of the form $C(m) = C_1 m^{C_2}$; by $C^e(m)$, $C_1^e(m)$ etc. – functions of the form $\exp C(m)$, $\exp C_1(m)$. By $C_*, C_*(m)$, $C_*^e(m)$ etc. we denote fixed constants and functions.

Observe that for any $C^e(m)$ and any $\sigma < 0$ the estimate $C^e(m) < \varepsilon_m^\sigma$ holds for all m provided that $\bar{\varepsilon} \ll 1$. We profit from the assumption that $\varepsilon < \bar{\varepsilon}$ with sufficiently small $\bar{\varepsilon} > 0$ and use inequalities like

$$C^e(m)\varepsilon_m^\rho < 1$$

without extra remark.

The theorem will be proven by the KAM-procedure. That is, for $m = 0, 1, \ldots$ we shall define subsets $\Omega_m \subset \Omega$, analytic functions $\mathcal{H}_m$ on the domains O_m as above, such that $\mathcal{H}_0 = \mathcal{H}_\varepsilon$, and a sequence of symplectic transformations

$$S_m : O_{m+1} \cap \mathcal{Y}_d \longrightarrow O_m \cap \mathcal{Y}_d$$

such that S_0 transforms the initial system $V_{\mathcal{H}_\varepsilon} = V_{\mathcal{H}_0}$ to $V_{\mathcal{H}_1}$, S_1 transforms $V_{\mathcal{H}_1}$ to $V_{\mathcal{H}_2}$ etc. We shall show that the system $V_{\mathcal{H}_m}$ on $O_m \cap \mathcal{Y}_d$ is integrable modulo a term $O(\varepsilon_m^\rho)$ — so the transformation $S_0 \circ \ldots \circ S_{m-1}$ "almost integrates" the initial equations (1.6). Finally, we shall see that the limiting transformation $S_0 \circ S_1 \circ \ldots$ is well-defined and integrates the equations.

We start with an inductive construction the transformation S_m and the Hamiltonian $\mathcal{H}_{m+1}$, given a hamiltonian $\mathcal{H}_m$, and finish with investigating the limiting transformation $S_0 \circ S_1 \circ \ldots$.

2.2 The Hamilton $\mathcal{H}_{\mathbf{m}}$.

On domain O_m we consider Hamiltonian $\mathcal{H}_m(\mathfrak{h};\omega),\ \mathfrak{h}=(p,q,y)$,

$$\mathcal{H}_m = H_{0m}(p,y;\omega) + \varepsilon_m H_m(\mathfrak{h};\omega), \tag{2.1}$$

where

$$H_{0m} = p\cdot\Lambda_m(\omega) + \frac{1}{2}\ \langle A_m(q;\omega)y,y\rangle, \tag{2.2}$$

and $\omega\in\Omega_m,\ \Omega_m$ is a Borel subset of Ω such that

$$\operatorname{mes}(\Omega\setminus\Omega_m)\le\gamma e(m). \tag{2.3}$$

The map $\omega\longmapsto\Lambda_m$ is Lipschitz and

$$|\Lambda_m(\omega)-\omega|^{\Omega_m,\mathrm{Lip}}\le C\varepsilon^\rho e(m). \tag{2.4}$$

The operator A_m is assumed to be diagonal in the basis $\varphi_j^\pm$:

$$A_m\varphi_j^\pm = (\lambda_j^{(m)}(\omega)+\beta_j^{(m)}(q;\omega))\,\varphi_j^\pm,$$

where

$$|\lambda_j^{(m)}-\lambda_j|^{\Omega_m,\mathrm{Lip}}\le j^{d_H}C\varepsilon^\rho e(m) \tag{2.5}$$

and

$$\int\beta_j^{(m)}\,dq=0,\quad |\beta_j^{(m)}|^{U_m,\Omega_m}\le j^{d_H}\,C\varepsilon^\rho e(m). \tag{2.6}$$

In particular, by the Cauchy estimate $|\nabla_q\beta_j^{(m)}|^{U_m,\Omega_m}\le j^{d_2}C(m)\varepsilon^\rho$ and

$$\|\nabla_q A_m\|_{d,d-d_2}^{U_m^1,\Omega_m}\le C(m)\varepsilon^\rho. \tag{2.5$'$}$$

The functional H_m is analytic in O_m and

$$|H_m|^{O_m,\Omega_m}\le C_*(m)\equiv K_4^m, \tag{2.7}$$

$$\|\nabla_y H_m\|_{d_c}^{O_m,\Omega_m}\le\varepsilon_m^{-1/3}C_*(m),\quad d_c:=d-d_H. \tag{2.8}$$

Clearly the initial Hamiltonian $\mathcal{H}_\varepsilon$ has the form $\mathcal{H}_0$. (One should chose $\Lambda_0(\omega)=\omega,\ A_m=A,\ \Omega_m=\Omega$, etc.).

Hamiltonian equations with the Hamiltonian $\mathcal{H}_m$ have the form

$$\dot p = -\frac{1}{2}\ \langle\nabla_q A_m(q;\omega)y,y\rangle\ -\varepsilon_m\nabla_q H_m, \tag{2.9}$$

$$\dot{q} = \Lambda_m(\omega) + \varepsilon_m \nabla_p H_m, \tag{2.10}$$

$$\dot{y} = JA_m(q;\omega)y + \varepsilon_m J\nabla_y H_m. \tag{2.11}$$

Our goal is to construct an analytic map $S_m : O_{m+1} \longrightarrow O_m$ which defines a symplectic transformation $S_m : O_{m+1} \cap \mathcal{Y}_d \longrightarrow O_m \cap \mathcal{Y}_d$ and transforms the Hamiltonian $\mathcal{H}_m$ to $\mathcal{H}_{m+1} = \mathcal{H}_m \circ S_m$ which has the form (2.1) with m replaced by $m+1$. The transformation S_m is constructed in six steps which are essentially identical to the ones described in [K]. The only difference comes during "averaging" when we extract from the perturbation and add to the integrable part H_{0m} the whole diagonal of $\operatorname{Hess} \varepsilon_m H_m$ — not only its averaging in q.[6] Because of this, the operators A_m in (2.2) depend on q (their analogies in [K] are q-independent). Accordingly, homological equations written in terms of these operators become more complicated. Their resolution is based on a new theorem on first-order linear differential equations with variable coefficients which we prove in Chapter 4.

2.3 Step 1. Splitting the perturbation.

We rewrite H_m as

$$H_m = h^q(q;\omega) + p \cdot h^{1p}(q;\omega) + \langle y, h^y(q;\omega)\rangle + \langle h^{yy}(q;\omega)y, y\rangle + H_{3m}(\mathfrak{h};\omega), \tag{2.12}$$

where $H_{3m} = O(|p|^2 + \|y\|_d^3 + |p|\,\|y\|_d)$. Next we change H_m (and so h^q) by an ω-dependent constant to achieve

$$(2\pi)^{-n} \int h^q dq = 0.$$

We denote

$$h^{0p} = (2\pi)^{-n} \int h^{1p} dq, \quad h^p = h^{1p} - h^{0p}, \tag{2.13}$$

and

$$\Lambda_{m+1} = \Lambda_m + \varepsilon_m h^0(\omega). \tag{2.14}$$

Now we rewrite $\mathcal{H}_m$ as

$$\mathcal{H}_m = H'_{0\,m+1}(p, y;\omega) + \varepsilon_m (H_{2m} + H_{3m})\ (\mathfrak{h};\omega),$$

where

$$H'_{0\,m+1} = p \cdot \Lambda_{m+1} + \frac{1}{2}\ \langle A_m y, y\rangle$$

and the function H_{2m} equals to

$$H_{2m} = h^q + p \cdot h^p + \langle y, h^y\rangle + \langle h^{yy} y, y\rangle.$$

[6] We are forced to do so since if $d_H > 0$ (and the perturbing vector field is unbounded), then to kill the diagonal part of $\operatorname{Hess} \varepsilon_m H_m$ the transformation S_m must be unbounded.

Lemma 2.1. *The terms of the decomposition (2.12) may be estimated as follows:*

a)

$$|h_q|^{U_m,\Omega_m} \le 2C_*(m),$$
$$|h^{1p}|^{U_m,\Omega_m} \le 2C_*(m)\varepsilon_m^{-2/3},$$
$$\|h^y\|_{d_c}^{U_m,\Omega_m} \le C_*(m)\varepsilon_m^{-1/3},$$
$$\|h^{yy}\|_{d,d_c}^{U_m,\Omega_m} \le C_*(m)\varepsilon_m^{-2/3}.$$

Besides, the operator h^{yy} is symmetric in Y and is real for real q.

b) *In the domain $O_{m+1} \subset O_m$ the term $\varepsilon_m H_{3m}$ is twice smaller than the admissible disparity of the next step:*

$$\varepsilon_m |H_{3m}|^{O_{m+1},\Omega_m} \le \frac{1}{2}\, C_*(m+1)\varepsilon_{m+1},$$
$$\varepsilon_m \|\nabla_y H_{3m}\|_{d_c}^{O_{m+1},\Omega_m} \le \frac{1}{2}\, C_*(m+1)\varepsilon_{m+1}^{2/3},$$

provided that $\bar{\varepsilon} \ll 1$ and K_4 in (2.7) is sufficiently large.

c) *The functions H_{2m}, H_{3m} are analytic in $\mathfrak{h} \in O_m$ and are real for real arguments.*

The proof is straightforward. See [K, p.59] or [K1].

2.4 Step 2. Formal construction of the transformation $\mathbf{S}_m$ and derivation the homological equations.

We construct S_m as the time-one shift along trajectories of an auxiliary Hamiltonian vector field

$$\dot{q} = \varepsilon_m \nabla_p F, \quad \dot{p} = -\varepsilon_m \nabla_q F, \quad \dot{y} = \varepsilon_m J \nabla_y F, \tag{2.15}$$

where the function F has the same structure as H_{2m}:

$$F = f^q(q;\omega) + p \cdot f^p(q;\omega) + \langle y, f^y(q;\omega)\rangle + \langle f^{yy}(q;\omega)y,\, y\rangle.$$

The flow $\{S^t\}$ of equations (2.15) is formed by canonical transformations (with respect to the symplectic structure ω_2 defined in Chapter 1, see more in [K]) and we set

$$S_m := S^t|_{t=1}.$$

Then formally

$$\mathcal{H}_m(S_m(\mathfrak{h};\omega);\omega) = \mathcal{H}_m(\mathfrak{h};\omega) + \varepsilon_m\{F, \mathcal{H}_m\} + O\,(\varepsilon_m^2),$$

where $\{\cdot,\cdot\}$ is the Poisson bracket defined in (1.1) (see [K]). Taking into account assertion b) of Lemma 2.1, we get that in O_{m+1}

$$\mathcal{H}_m \circ S_m(\mathfrak{h}) = H'_{0\,m+1} + \varepsilon_m(H_{2m} + \nabla_p F \cdot \nabla_q H'_{0\,m+1} - \nabla_q F \cdot \nabla_p H'_{0\,m+1} +$$
$$+ \langle J\nabla_y F, \nabla_y H'_{0\,m+1}\rangle) + O\,(\varepsilon_{m+1}).$$

We observe that

$$\nabla_p H'_{0\,m+1} = \Lambda_{m+1},\ \nabla_q H'_{0\,m+1} = \frac{1}{2}\ \langle \nabla_q A_m\, y, y\rangle,\ \nabla_y H'_{0\,m+1} = A_m\, y$$

and abbreviate

$$\Lambda_{m+1} = \omega', \quad \omega' \cdot \nabla_q = \frac{\partial}{\partial \omega'}, \quad A_m = A.$$

Now we rewrite $\mathcal{H}_m \circ S_m$ as

$$\mathcal{H}_m \circ S_m = H'_{0\,m+1} + \varepsilon_m \left[\frac{1}{2}\ \langle (f^p \cdot \nabla_q A)y, y\rangle - \partial f^q/\partial\omega' - p\cdot\partial f^p/\partial\omega' - \right.$$

$$\langle y, \partial f^y/\partial\omega'\rangle - \langle y, (\partial f^{yy}/\partial\omega')y\rangle + \langle Ay, Jf^y\rangle +$$

$$\left. + 2\langle Ay, Jf^{yy}y\rangle + h^q + p\cdot h^p + \langle y, h^y\rangle + \langle y, h^{yy}y\rangle \right] + O\left(\varepsilon_{m+1}\right). \tag{2.16}$$

(The term in the square brackets equals $H_{2m} + \{F, H_{0\,m+1}\}$).

We wish to find F in such a way that the contents of the square brackets in the r.h.s. of (2.16) vanish up to an admissible disparity we define below. For this end f^q, f^p, f^y and f^{yy} should satisfy the homological equations:

$$\partial f^q/\partial\omega' = h^q(q;\omega), \qquad \partial f^p/\partial\omega' = h^p(q;\omega), \tag{2.17}$$

$$\partial f^y/\partial\omega' - AJf^y = h^y, \tag{2.18}$$

$$\partial f^{yy}/\partial\omega' + f^{yy}JA - AJf^{yy} = h^{yy} + \frac{1}{2}\ f^p\cdot\nabla_q A =: h^{1yy}$$

(the disparity is introduced later). We define the functions b_j,

$$b_j(q;\,\omega) = \frac{1}{2}\ \langle h^{1yy}\varphi_j^+,\, \varphi_j^+\rangle + \frac{1}{2}\ ,\langle h^{1yy}\varphi_j^-,\, \varphi_j^-\rangle \tag{2.19}$$

and the operators B and h^{0yy}, where

$$B(q;\,\omega) = \operatorname{diag}\{b_1, b_1, b_2, b_2 \ldots\}$$

(i.e., $B\varphi_j^\pm = b_j\varphi_j^\pm$), and

$$h^{0yy}(q;\,\omega) = h^{1yy} - B.$$

Both operators h^{0yy} and h^{1yy} depend on the solution f^p of the second equation in (2.17). Using the estimate for f^p we get below in Lemma 2.2 jointly with (2.5′) and Lemma 2.1, we find that

$$\|h^{1yy}\|_{d,d_c}^{U^1_m,\Omega_m\setminus\Omega^1} \le C(m)\varepsilon_m^{-2/3}.$$

Hence,

$$|b_j|^{U^1_m,\Omega_m\setminus\Omega^1} \le j^{d_2}c(m)\varepsilon_m^{-2/3} \quad \forall j,$$

and the operator h^{0yy} meets similar estimate:

$$\|h^{0yy}\|_{d,d_c}^{U^1_m,\Omega_m\setminus\Omega^1} \le C'(m)\varepsilon_m^{-2/3}.$$

We observe that $JA = AJ$ and rewrite the last equation with h^{1yy} replaced by h^{0yy} (i.e., introducing a disparity) as

$$\partial f^{yy}/\partial\omega' + [f^{yy}, JA] = h^{0yy}. \tag{2.20}$$

If $f^q, \ldots, f^{yy}$ solve the equations (2.17) – (2.20) then the contents of the square brackets in (2.16) equals $\langle By, y\rangle$ and

$$\{F, H'_{0\,m+1}\} = -H_{2m} + \langle By, y\rangle. \tag{2.21}$$

2.5 Step 3. Solving the homological equations.

The following lemma which deals with equations (2.17) is classical for the KAM-theory (see e.g. [K, pp. 67–68]):

Lemma 2.2. *Define Ω^1 as*

$$\Omega^1 = \{\omega \in \Omega_m \mid |\omega' \cdot s| \le C^{-1}(m+1)^{-2}|s|^{-n} \text{ for some } s \in \mathbb{Z}^n \backslash \{0\}\}.$$

Then $mes\,\Omega^1 \le \gamma(m+1)^{-2}/3K_*$ if C is chosen sufficiently large and for $\omega \in \widetilde{\Omega} = \Omega_m \backslash \Omega^1$ equations (2.17), (2.18) have analytic solutions real for real arguments and such that

$$|f^q|^{U^1_m,\widetilde{\Omega}} \le C(m), \qquad |f^p|^{U^1_m,\widetilde{\Omega}} \le \varepsilon_m^{-2/3} C(m).$$

Equations (2.18), (2.20) are more complicated than (2.17). We start with the most difficult equation (2.20).

The numbers $\lambda_j^{(m+1)}(\omega)$ were defined for $j \in \mathbb{N}$. Now we define them for all $j \in \mathbb{Z}\backslash\{0\}$ by setting

$$\lambda_{-j}^{(m+1)}(\omega) = -\lambda_j^{(m+1)}(\omega) \quad \forall j \in \mathbb{N}.$$

Lemma 2.3. *There exists a Borel subset $\Omega^2 \subset \Omega_m$ such that*

$$mes\,\Omega^2 \le \gamma(m+1)^{-2}/(3K_*)$$

and

$$\left|\omega' \cdot s + \lambda_j^{(m+1)} - \lambda_k^{(m+1)}\right| \ge \frac{|j^{d_1} - k^{d_1}|}{C_{**}(m)\,\langle s \rangle^{c_1}}$$

for all $\omega \in \widetilde{\Omega} \setminus \Omega^2$, all $j,k \in \mathbb{Z}\backslash\{0\}$ and all $s \in \mathbb{Z}^n$, with some $C_{**}(m)$ and c_1. Here for $j \in \mathbb{Z}$ we write $j^{d_1} = sgn\, j\, |j|^{d_1}$.

The proof follows [K] and is given in Chapter 3 below.

For $j \in \mathbb{N}$ we set

$$w_j = (\varphi_j^+ + i\varphi_j^-)/\sqrt{2}, \quad w_{-j} = (\varphi_j^+ - i\varphi_j^-)/\sqrt{2}.$$

The vectors

$$\{|j|^{-s} w_j \mid j \in \mathbb{Z}\backslash\{0\}\} \tag{2.22}$$

form a Hilbert basis of Y_s^c. The operator JA is diagonal in this basis:

$$JA(q;\omega) w_j = i\ \lambda_j^1(q;\omega) w_j,$$

where $\lambda_{-j}^1 = -\lambda_j^1$ and for $j \in \mathbb{N}$

$$\lambda_j^1(q;\omega) = \lambda_j^{(m+1)}(\omega) + \beta_j^{(m+1)}(q;\omega).$$

Let us denote by $\{f_{kj}(q;\omega)\}$ and $\{h_{kj}(q;\omega)\}$ the matrix elements of the operators f^{yy} and h^{0yy} with respect to the complex basis $\{w_j\}$. Then the equation (2.20) can be rewritten as

$$\frac{\partial}{\partial\omega'}\, f_{kj}(q;\omega) + (\lambda_j^1 - \lambda_k^1)\ (q;\omega)\ f_{kj} = h_{kj}(q;\omega). \tag{2.23}$$

Due to the definition of the operator h^{0yy} its diagonal part vanishes:

$$h_{kk}(q;\omega) \equiv 0 \qquad \forall k.$$

Besides if we supply the spaces Y_d^c, $Y_{d_c}^c$ with the Hilbert bases as in (2.22), then the matrix of the operator $h^{0yy} : Y_d^c \to Y_{d_c}^c$ will be $\{|k|^{d_c} h_{kj} |j|^{-d}\}$. In particular, since $\|h^{0yy}\|_{d,d_c}^{U_m^1,\Omega_m\setminus\Omega^1} \le C(m)\varepsilon_m^{-2/3}$, then

$$|h_{kj}|^{U_m^1} \le C(m)\,\varepsilon_m^{-2/3} |j|^d |k|^{-d_c}.$$

Observe that

$$\lambda_j^1 - \lambda_k^1 = (\lambda_j^{(m+1)} - \lambda_k^{(m+1)})\,(\omega) + (\beta_j^{(m+1)} - \beta_k^{(m+1)})\,(q;\omega)$$

is the sum of a constant which is $\ge \max(|j|,\ |k|)^{d_1-1}/\,C$ (due to (1.4)) and a q-dependent function of order

$$\varepsilon\ \max(|j|,\ |k|)^{d_H}.$$

Since d_H can be positive, then (2.23) is a perturbation of a constant-coefficient equation by a variable-coefficient term which can be arbitrary large. Still since $d_2 < d_1 - 1$, then the "very large" constant-coefficient part of (2.23) suppresses the "large" variable coefficient one: We show in Chapter 4 that for ω from $\Omega_m \setminus (\Omega^1 \cup \Omega^2)$ the equation (2.23) has a unique analytic solution f_{kj} and

$$|f_{kj}|^{U_m^2} \le C^e(m)\ |h_{kj}|^{U_m^1} /\, |j^{d_1} - k^{d_1}|,$$

where for $j \in \mathbb{Z}$ we set $j^{d_1} = \operatorname{sgn} j\ |y|^{d_1}$.

The operator $f^{yy} : Y_d \longrightarrow Y_d$ has the matrix $F = \{F_{kj}\} = \{|k|^d f_{kj}\ |j|^{-d}\}$ (both the spaces are provided with Hilbert bases (2.22)). So for each $q \in U_m^2$

$$|F_{kj}(q)| \le C^2(m)\varepsilon_m^{-2/3} |k|^{d_H} /\, |j^{d_1} - k^{d_1}|.$$

Since $F_{kk} \equiv 0$ and $d_1 > d_H + 1$, then

$$\sum_k |F_{kj}| \le \varepsilon_m^{-2/3} C_1^e(m) \left(\int_{-\infty}^{-1} + \int_1^j + \int_{j+1}^{\infty} \right) \frac{|x|^{d_H}\, dx}{|j^{d_1} - x^{d_1}|} \le$$
$$\le \varepsilon_m^{-2/3} C_2^e(m)\, |j|^{d_H+1-d_1} \log |j| \le C^e(m)\, \varepsilon_m^{-2/3}.$$

Similar estimate holds for ℓ^1-norms of rows of the matrix F. Therefore the norm of the operator $f^{yy}(q) : Y_d \longrightarrow Y_d$ with q in U_m^2 is bounded by $C^e(m)\varepsilon_m^{-2/3}$. (For this classical result see [HLP, Chap. 8]). The same estimate holds for the norm of the operator $f^{yy} : Y_{d_c} \longrightarrow Y_{d_c}$.

So the norm of $f^{yy}(q),\ q \in U_m^2$ is estimated. To estimate the Lipschitz constant, we consider an increment f_Δ^{yy} of the operator $f^{yy},\ f_\Delta^{yy} = f^{yy}(q;\omega_1) - f^{yy}(q;\omega_2)$. For this operator we have the equation

$$\partial f_\Delta^{yy}/\partial\omega' + [f_\Delta^{yy}, JA] = h_\Delta^{0yy} - \nabla_q f^{yy}(q;\omega_2)\cdot(\omega_1-\omega_2) - [f(q;\omega_2), JA_\Delta] =: H_\Delta^{yy},$$

where h_Δ^{0yy} and A_Δ stand for increments of h^{0yy} and A. We immediately see that for $q \in U_m^3$

$$\|H_\Delta^{yy}(q;\omega)\|_{d,d_c} \le C_1^e(m)\varepsilon_m^{-2/3}\ |\omega_1-\omega_2|.$$

So the given above arguments estimate Lipschitz constant in ω for f^{yy} when $q \in U_m^4$. We can use intermediate domains like $U_m^{3/2}$ to get the same estimate for q in U_m^2:

Lemma 2.4. *If $\omega \in \widetilde{\Omega} = \Omega_m \setminus (\Omega^1 \cup \Omega^2)$, then equation (2.20) has an analytic solution f^{yy} which is a symmetric in Y^c operator, real for real q and such that*

$$\|f^{yy}\|_{d,d}^{U_m^2,\widetilde{\Omega}} \le C^e(m)\ \varepsilon_m^{-2/3}. \tag{2.24}$$

Quite similar (but simpler) arguments show solvability of equation (2.18):

Lemma 2.5. *There exists a Borel subset $\Omega^3 \subset \Omega_m$, mes $\Omega^3 \le \gamma(m+1)^{-2}/3K_*$, such that for $\omega \in \Omega_m \setminus (\Omega^1 \cup \Omega^3)$, the equation (2.18) has an analytic solution $f^y(q;\omega)$, real for real q, and such that*

$$\|f^y\|_d^{U_m^2,\Omega_m\setminus(\Omega^1\cup\Omega^3)} \le C^e(m)\,\varepsilon_m^{-2/3}.$$

Now we define the set Ω_{m+1} as

$$\Omega_{m+1} = \Omega_m \setminus (\Omega^1 \cup \Omega^2 \cup \Omega^3). \tag{2.25}$$

Due to the estimates for measures of the sets Ω^1, Ω^2 and Ω^3 we got in Lemmas 2.2, 2.3 and 2.6,

$$\text{mes}\,(\Omega\backslash\Omega_{m+1}) \le \text{mes}\,(\Omega\backslash\Omega_m) + \gamma(m+1)^{-2}/K_* \le \gamma\, e(m+1).$$

So Ω_{m+1} meets estimate (2.3) with m replaced by $m+1$.

2.6 Step 4. Study of the transformation $\mathbf{S_m}$.

We recall that $S_m = S^t|_{t=1}$, where $\{S^t\}$ is the flow of the system (2.15) which we now write as

$$\dot{\mathfrak{h}} = \varepsilon_m V_F(\mathfrak{h}),\quad \mathfrak{h} = \mathfrak{h}(t) = (q,p,y)(t),$$

where $V_F(\mathfrak{h}) = (-\nabla_q F,\ \nabla_p F,\ J\nabla_y F)$.

Lemma 2.6. *If $\bar{\varepsilon} \ll 1$, then for ω in Ω_{m+1} the map S_m is analytic and sends O_m^3 to O_m^2 in such a way that for $\mathfrak{h} \in O_m^3$ and for $-d \le \theta \le d$ we have*

$$\|S_m - \mathfrak{h}\|_d^{O_m,\Omega_{m+1}} \le \varepsilon_m^\rho, \tag{2.26}$$

$$\|S_{m*}(\mathfrak{h}) - id\|_{\theta,\theta} \le \varepsilon_m^\rho \tag{2.27}$$

where $S_{m*}: Z^c_\theta = \mathbb{C}^n \times \mathbb{C}^n \times Y^c_\theta \longrightarrow Z^c_\theta$.

Proof: The estimates of Lemmas 2.2, 2.4, 2.5 and the Cauchy estimate show that for $\mathfrak{h}$ in O^3_m components of the vector field $\varepsilon_m V_F$ meet the estimates

$$|\varepsilon_m \nabla_p F| \le C^e(m)\varepsilon_m^{1/3}, \ \ |\varepsilon_m \nabla_q F| \le C^e(m)\varepsilon_m, \ \ \|\varepsilon_m \nabla_y F\|_d \le C^e(m)\varepsilon_m^{2/3},$$

so S_m sends O^3_m to O^2_m and satisfies (2.26).

To prove (2.27), we denote $\eta(t) = S^t(\mathfrak{h})_*\eta$. Then $\eta(t)$ is the solution of the Cauchy problem

$$\dot\eta(t) = \varepsilon_m \ V_F(\mathfrak{h}(t))_*\eta(t), \ \ \eta(0) = \eta,$$

and

$$\eta(t) = \eta + \varepsilon_m \int_0^t V_F(\mathfrak{h}(\tau))_*\eta(\tau)d\tau.$$

Since the map f^{yy} is symmetric in Y, then by the interpolation theorem (see [RS] and Appendix A in [K]) the estimate (2.24) holds for all $|\theta| \le d$. Hence,

$$\|\eta(t) - \eta\|_\theta \le t\, C^e(m)\varepsilon_m^{1/3}\, \|\eta\|_\theta$$

and (2.27) follows. □

2.7 Step 5. The transformed Hamiltonian.

Now we study the transformed Hamiltonian $\mathcal{H}_m \circ S_m$. Since the functional $\mathcal{H}_m$ is smooth on the space $\mathcal{Y}_d$ and the flow-maps S^t are C^1-smooth in t, then

$$\frac{d}{dt}\ H'_{0\,m+1} \circ S^t = \varepsilon_m \{F,\ H'_{0\,m+1}\} \circ S^t = -\varepsilon_m (H_{2m} - \langle By, y\rangle) \circ S^t,$$

where the second equality follows from (2.21). The first equality is well known for finite-dimensional Hamiltonian systems; for its infinite-dimensional version we use see [K, Part 1]. Now we can calculate the second derivative:

$$\frac{d^2}{dt^2}\ H'_{0\,m+1} \circ S^t = -\varepsilon_m^2 \{F,\ H_{2m} - \langle By, y\rangle\} \circ S^t.$$

Thus,

$$\begin{aligned} H'_{0,m+1} \circ S_m &= H'_{0\,m+1} \circ S^t|_{t=1} = \\ &= H'_{0\,m+1} + \frac{d}{dt}\ H'_{0\,m+1}|_{t=0} + \int_0^1 (1-t)\ \frac{d^2}{dt^2}\ H'_{0\,m+1} \circ S^t dt = \\ &= H'_{0\,m+1} + \varepsilon_m\langle By, y\rangle - \varepsilon_m H_{2m} - \varepsilon_m^2 \int_0^1 (1-t)\, \{F, H_{2m} - \langle By, y\rangle\} \circ S^t dt. \end{aligned}$$

Similar

$$\varepsilon_m(H_{2m}+H_{3m})\circ S_m=\varepsilon_m(H_{2m}+H_{3m})+\varepsilon_m^2\int_0^1\{F,\,H_{2m}+H_{3m}\}\circ S^t dt.$$

Therefore, the transformed Hamiltonian can be written as

$$\begin{aligned}\mathcal{H}_m\circ S_m&=H_{0\,m+1}+\varepsilon_m H_{3m}+\\ &+\varepsilon_m^2\int_0^1(t-1)\{F,H_{2m}-\langle By,y\rangle\}\circ S^t dt+\varepsilon_m^2\int_0^1\{F,H_{2m}+H_{3m}\}\circ S^t dt,\end{aligned}\tag{2.28}$$

where

$$H_{0\,m+1}=H'_{0\,m+1}+\langle By,y\rangle$$

has the form (2.2) with $m:=m+1$ and with

$$A_{m+1}=A_m+2\varepsilon_m B.$$

Since diagonal elements $b_j(q;\omega)$ of the operator B are bounded by $j^{d_2}C(m)\varepsilon_m^{-2/3}$ as well their Lipschitz constants in ω, then diagonal elements $\lambda_j^{(m+1)}+\beta_j^{(m+1)}$ of the operator A_{m+1} satisfy the a priori estimates (2.5), (2.6) with m replaced by $m+1$.

For $j=1,2,3,4$ we denote by $\Delta_j H$ the j-th term in the r.h.s. of (2.28). To prove that the Hamiltonian $\mathcal{H}_{m+1}:=\mathcal{H}_m\circ S_m$ has the form (2.1) we should check that

$$\Delta_2 H+\Delta_3 H+\Delta_4 H=\varepsilon_{m+1}H_{m+1},\tag{2.29}$$

where H_{m+1} is a function satisfying estimates (2.7), (2.8) in the domain O_{m+1}.

The term $\Delta_2 H$ is twice smaller than the r.h.s. of (2.29) by Lemma 2.1. The estimates for $\Delta_3 H,\Delta_4 H$ follow from the following statement:

Lemma 2.7. *If H is a functional such that*

$$|H|^{O^1_m,\Omega_{m+1}}\le C^e(m)\,\varepsilon_m^2,\qquad \|\nabla_y H\|_{d_c}^{O^1_m,\Omega_{m+1}}\le C^e(m)\,\varepsilon_m^{5/3},\tag{2.30}$$

then for $0\le t\le 1$

$$|\{F,H\}\circ S^t|^{O^5_m,\Omega_{m+1}}\le C_1^e(m)\,\varepsilon_m^{4/3},\quad \|\nabla_y(\{F,H\}\circ S^t)\|_{d_c}^{O^5_m,\Omega_{m+1}}\le C_1^e(m)\,\varepsilon_m.\tag{2.31}$$

The lemma is proven in [K, pp. 81–82]. Here we just remark that the first estimate in (2.31) is essentially obvious since $\{F,H\}=-\nabla_q F\cdot\nabla_p H+\nabla_p F\cdot\nabla_q H+\langle J\nabla_y F,\nabla_y H\rangle$, since estimates for $\nabla H=(\nabla_p H,\nabla_q H,\nabla_y H)$ follow from (2.30) (and the Cauchy estimate) and estimates for ∇F result from estimates for its components obtained in Lemmas 2.2, 2.4, 2.5. By the first estimate and the Cauchy one, we get the second estimate with the d_c-norm replaced by the $(-d)$-norm. So to prove the second estimate we just have to control smoothness of the gradient. See in [K] how to do it.

Due to the lemma for $\mathfrak{h}$ in O_{m+1} we have $|\Delta_3 H+\Delta_4 H|\le 2C_1^2(m)\,\varepsilon_m^{4/3}\le\varepsilon_{m+1}$ if $\bar\varepsilon\ll 1$, and similar with gradients of the functionals.

Therefore $\mathcal{H}_{m+1}:=\mathcal{H}_m\circ S_m$ also has the form (2.1).

2.8 Step 6. Transition to limit.

Here we show that the set

$$(S_0 \circ S_1 \circ \ldots)\ \ (T_0^n) \subset \mathcal{Y}_d$$

is a smooth torus, invariant for the equations (1.6).

Let us denote $\Omega_\varepsilon = \cap \Omega_m$. Then Ω_ε is a Borel subset of Ω and by (2.3)

$$\mathrm{mes}(\Omega \backslash \Omega_\varepsilon) \leq \gamma/2.$$

For $\omega \in \Omega_\varepsilon$ and $0 \leq r \leq N$ we denote by Σ_N^r the map

$$\Sigma_N^r(\cdot\,;\omega) : S_r \circ \ldots \circ S_{N-1} : O_N \times \Omega_N \longrightarrow O_r$$

(as usual, Σ_r^r is the projection to O_r). We claim that for all $r, m \geq 0$

$$\|\Sigma_{r+m}^r - \Pi_{\mathcal{Y}}\|_d^{O_{r+m},\Omega_\varepsilon} \leq 3\,\varepsilon_r^\rho, \tag{2.32}$$

where $\Pi_{\mathcal{Y}}(\mathfrak{h};\omega) = \mathfrak{h}$. Indeed, let us denote the l.h.s. in (2.32) by D_{r+m}^r. We rewrite the identity $\Sigma_{r+m}^r(\mathfrak{h};\omega) = S_r(\Sigma_{r+m}^{r+1}(\mathfrak{h};\omega);\omega)$ in the form

$$\Sigma_{r+m}^r - \Pi_{\mathcal{Y}} = (S_r - \Pi_{\mathcal{Y}}) \circ (\Sigma_{r+m}^{r+1} \times \Pi_\Omega) + (\Sigma_{r+m}^{r+1} - \Pi_{\mathcal{Y}}),$$

where $\Pi_\Omega(\mathfrak{h},\omega) = \omega$. By Lemma 2.6 we get

$$D_{r+m}^r \leq \varepsilon_r^\rho\ \left(D_{r+m}^{r+1} + 2\right) + D_{r+m}^{r+1}.$$

As $D_{r+m}^{r+m} = 0$, then (2.32) follows by induction.

Observe that because estimate (2.27), for finite $r \leq N$ and any $\mathfrak{h} \in O_N$ the tangent map $\Sigma_N^r(\mathfrak{h})_*$ is close to the identity:

$$\|\Sigma_N^r(\mathfrak{h})_* - \mathrm{id}\|_{\theta,\theta} \leq 2\ \varepsilon_r^\rho. \tag{2.33}$$

Let us denote by $\mathcal{O}$ the set

$$\mathcal{O} = \{0\} \times U(\delta/2) \times \{0\} \subset \mathcal{Y}_d^c.$$

This set is a complex neighborhood of the torus $T_0^n = \{0\} \times \mathbb{T}^n \times \{0\}$ in $\{0\} \times (\mathbb{C}^n/2\pi\mathbb{Z}^n) \times \{0\}$, which is contained in each O_m since $\delta_m > \delta/2$.

As a consequence of (2.32) we get that for each $m \geq 0$ and each $\omega \in \Omega_\varepsilon$ the maps Σ_{m+N}^m restricted to $\mathcal{O}$ converge to an analytic map

$$\Sigma_\infty^m(\cdot\,;\omega) : \mathcal{O} \longrightarrow O_m \subset \mathcal{Y}_d^c$$

and $\Sigma_p^m \circ \Sigma_\infty^p = \Sigma_\infty^m$ for all $p \leq m$. Because (2.32), (2.33) we have the estimates:

$$\|\Sigma_\infty^r - \Pi_{\mathcal{Y}}\|_d^{\mathcal{O},\Omega_\varepsilon} \leq 3\ \varepsilon_r^\rho \qquad \forall r, \tag{2.34}$$

$$\|\Sigma_\infty^r(\mathfrak{h})_* - \mathrm{id}\|_{\theta,\theta} \leq 2\ \varepsilon_r^\rho \qquad \forall r, \quad \forall \theta \in [-d,d]. \tag{2.35}$$

Due to the recurrent formula (2.14) for the vectors Λ_m and an estimate of Lemma 2.1, the maps Λ_m converge to a Lipschitz map Λ_∞ such that

$$\Lambda_\infty : \Omega_\varepsilon \longrightarrow \mathbb{R}^n, \quad |\Lambda_\infty - \omega|^{\Omega_\varepsilon, \mathrm{Lip}} \leq C\,\varepsilon^{1/3},$$

and

$$|\Lambda_\infty - \Lambda_m| \leq \varepsilon_m^\rho. \tag{2.36}$$

Now we consider a curve

$$\mathfrak{h}_\infty(t) = (0, q_0 + t\Lambda_\infty, 0) \subset T_0^n$$

and the curves $\mathfrak{h}_m(t) = \Sigma_\infty^m \mathfrak{h}_\infty(t) \subset O_m$. We wish to show that $\mathfrak{h}_0(t)$ is a (strong) solution of the equation (1.6). To do so, we use (2.34) (2.36) and the Cauchy inequality to get that

$$\dot{\mathfrak{h}}_m = (0, \Lambda_m(\omega), 0) + O\left(\varepsilon_m^\rho\right). \tag{2.37}$$

Using again the estimate (2.34) and the form of equations (2.9)–(2.11), which we abbreviate as

$$\dot{\mathfrak{h}} = V_{\mathcal{H}_m}(\mathfrak{h}), \quad \mathfrak{h} \in O_m, \tag{2.38}$$

we see that

$$\|V_{\mathcal{H}_m}(\mathfrak{h}_m(t)) - (0, \Lambda_m, 0)\|_{d-d_1} = O(\varepsilon_m^\rho). \tag{2.39}$$

Since $\Sigma_m^0 \mathfrak{h}_m(t) = \mathfrak{h}_0(t)$ and $\Sigma_{m\,*}^0 \dot{\mathfrak{h}}_m = \dot{\mathfrak{h}}_0\,,\ \ \Sigma_{m\,*}^0 V_{\mathcal{H}_m}(\mathfrak{h}_m) = V_{\mathcal{H}_0}(\mathfrak{h}_0)$, then (2.37), (2.39) and (2.33) jointly imply that

$$\dot{\mathfrak{h}}_0 - V_{\mathcal{H}_0}\mathfrak{h}_0 = O\left(\varepsilon_m^\rho\right) \text{ in } \mathbb{R}^n \times \mathbb{R}^n \times Y_{d-d_1}.$$

Since m is arbitrary, we get that the l.h.s. is zero and $\mathfrak{h}_0(t)$ is a solution of the system (1.6) (which coincides with (2.9)–(2.11) when $m = 0$).

Now assertions of Theorem 1 and Amplification 1 follow if we choose $\Sigma_\omega(q) = \Sigma_\infty(q, 0, 0; \omega)$ and $\omega' = \Lambda_\infty(\omega)$.

3. Proof of Lemma 2.3 (estimation of the small divisors)

We denote $\Lambda_{jk}(\omega) = \lambda_j^{(m+1)}(\omega) - \lambda_k^{(m+1)}(\omega)$ and rewrite the assertion of the lemma as

$$|\omega' \cdot s + \Lambda_{jk}(\omega)| \geq \kappa := \frac{|j^{d_1} - k^{d_1}|}{C_{**}(m)\langle s\rangle^{c_1}} \tag{3.1}$$

for all ω in $\widetilde{\Omega}\backslash\Omega^2$ and all $j, k \in \mathbb{Z}\backslash\{0\}$. Here the constants C_{**}, c_1 and the Borel subset Ω^2 such that $\operatorname{mes}\Omega^2 \leq \gamma(m+1)^{-2}/(3K_*)$ have to be found.

If $|s| \leq M_1$ and $j \leq j_1$ then (2.4), (2.5) and the assumption (1.8) of Theorem 1 jointly imply (3.1), so henceforth we may suppose that

$$|s| \geq M_1 \text{ or } j \geq j_1, \tag{3.2}$$

where M_1 and j_1 depending on the numbers listed in (1.7) will be choosen later.

Let us denote for a moment $D(j,k,s) = \omega' \cdot s + \Lambda_{jk}(s)$. Then

$$D(j,k,s) = D(-k,-j,s) = -D(-j,-k,-s)$$

and we may also suppose that

$$j > 0, \quad j \geq |k|, \quad j \neq k \tag{3.3}$$

(for $j = k$ the estimate (3.1) is trivial).

Observe that

$$|\Lambda_{jk}| \geq C_0^{-1} \, |j^{d_1} - k^{d_1}| \tag{3.4}$$

and

$$|j^{d_1} - k^{d_1}| \geq C_1^{-1} j^{d_1 - 1}. \tag{3.5}$$

Indeed, for large j the estimate (3.4) follows from (1.4) and (2.5), for j small it results from the assumption (1.8) with $s = 0$ (j_1 should be sufficiently large).

By (3.4) the estimate (3.1) holds trivially if $j \geq j_1$ and $|s| \leq C^{-1} \, | \, j^{d_1} - k^{d_1}|$, where C is choosen accordingly. So we can suppose below that

$$|s| \geq C^{-1} \, | \, j^{d_1} - k^{d_1}|. \tag{3.6}$$

In particular, $s \neq 0$.

Let us denote by $\mathcal{L}$ the set of all triples (k,j,s) as in (3.2), (3.3), (3.6). For $(k,j,s) \in \mathcal{L}$ we define

$$\Omega(k,j,s) \subset \widetilde{\Omega}$$

as the set of all $\omega \in \widetilde{\Omega}$ violating (3.1) for the choosen triple (k,j,s). Let us take for Ω^2 the union

$$\Omega^2 = \bigcup \{ \Omega(k,j,s) \, | \, (k,j,s) \in \mathcal{L} \}.$$

Clearly, (3.1) holds for ω outside Ω^2. So it remains to estimate the measure of Ω^2. Here the key is the following result:

Lemma 3.1. *For each triple $(k,j,s) \in \mathcal{L}$ we have*

$$\operatorname{mes} \Omega(k,j,s) \leq C\kappa,$$

provided that j_1, M_1 are sufficiently large.

Proof: Let us consider the map

$$\widetilde{\Omega} \ni \omega \longmapsto \omega' = \Lambda_{m+1}(\omega) \in \mathbb{R}^{n+1}.$$

This map is Lipschitz-close the identity, so it is a Lipschitz homeomorphism which changes the diameters of sets and their Lebesgue measure no more than twice (see [K, Appendix C]). So to estimate $\operatorname{mes} \Omega(k,j,s)$ is equivalent to estimate the measure of the set Ω',

$$\Omega' = \Lambda_{m+1}(\Omega(k,j,s)).$$

To make this estimate we express $\lambda_k, \lambda_j, \Lambda_{kj}$ as function of ω' and write Ω' as

$$\Omega' = \{\omega' \in \Lambda_{m+1}(\widetilde{\Omega}) \,\big|\, |\omega' \cdot s - \Lambda_{kj}| \le \kappa\}.$$

Since $|\omega'| \le C$ for each ω', then by the Fubini theorem to estimate $\operatorname{mes}\Omega'$ it is sufficient to estimate the one-dimensional measure of the intersection of Ω' with every line in $\mathbb{R}^n$ parallel to some fixed direction. In particular, parallel to $S = s/|s|$. Take any $\eta \in \mathbb{R}^n$. The intersection of Ω' with the line $L_\eta = \{\eta + tS \,|\, t \in \mathbb{R}\}$ is given by t from the set

$$\{t \,\big|\, |\Gamma(t)| \le \kappa\}, \tag{3.7}$$

where

$$\Gamma(t) := (\omega' \cdot s + \Lambda_{kj}(\omega'))\big|_{\omega' = \eta + tS}.$$

Observe that $(\partial/\partial t)\omega' \cdot s = |s|$, where $\omega' = \eta + tS$. So if we denote $\mathrm{Lip}\Lambda_{kj} = \mathrm{Lip}(\omega' \longmapsto \Lambda_{jk})$, then for $t_1 > t_2$ we have

$$\begin{aligned}\Gamma(t_1) - \Gamma(t_2) &\ge |s|(t_1 - t_2) - (t_1 - t_2)\mathrm{Lip}\,\Lambda_{kj} \ge (t_1 - t_2)\ (|s| - C\, j^{d_H}) \ge \\ &\ge C^{-1}(t_1 - t_2)\ (j^{d_1} - k^{d_1} - c_1\, j^{d_H}) \ge C_2^{-1}(t_1 - t_2)\ (j^{d_1 - 1} - C_3\, j^{d_H})\end{aligned}$$

(we use (3.5) and (3.6)). So if $j \ge j_1$ and j_1 is sufficiently large, then

$$\Gamma(t_1) - \Gamma(t_2) \ge t_1 - t_2.$$

If $j_1 < j_1$, then by (3.2) $|s| \ge M_1$ and again

$$\Gamma(t_1) - \Gamma(t_2) \ge (t_1 - t_2)\ (M_1 - C j_1^{d_H}) \ge t_1 - t_2$$

if we choose $M_1 \ge C j_1^{d_H} + 1$.

Thus, the measure of the set (3.6) is less than 2κ and the assertion of the lemma follows. □

Now an estimate for the measure of Ω^2 is straightforward:

$$\operatorname{mes}\Omega^2 \le \sum_{\mathcal{L}} \operatorname{mes}\Omega(k,j,s) \le \frac{C_2}{C_{**}(m)} \sum_s \langle s\rangle^{-c_1} \sum_{\substack{j,k \\ (j,k,s)\in\mathcal{L}}} (j^{d_1} - k^{d_1}).$$

By (3.5), $j \le C|s|^{d_0}$ where $d_0 = 1/(d_1 - 1)$. So the inner sum in the r.h.s. may be estimated as follows:

$$\sum_{\substack{j,k \\ (j,k,s)\in\mathcal{L}}} (j^{d_1} - k^{d_1}) \le C \sum_{\substack{j,k \\ (j,k,s)\in\mathcal{L}}} |s| \le C_1 \langle s\rangle^{2d_0 + 1}.$$

Therefore, $\operatorname{mes}\Omega^2 \le \gamma(m+1)^{-2}/(3K_*)$ if $c_1 > 2\,d_0 + n + 1$ and $C_{**}(m)$ is sufficiently large.
Lemma 2.3 is proven.

The crux of our resolution the homological equations (2.18), (2.20) in Chapter 2.5 was reduction the equations to infinite systems of non-coupled differential equations on the torus $\mathbb{T}^n$ with large variable coefficients. (Equation (2.23) in Chapter 2). Each equation can be written as

$$-i\,\frac{\partial x}{\partial \omega} + Ex + Bh(q)x = b(q), \quad q \in \mathbb{T}^n, \tag{4.1}$$

where the function h has zero mean value and is of order one, B could be large, but is much smaller than the constant E:

$$E \geq C \ \text{ and } \ E^\theta \geq C_1 B \ \text{ with some } \ 0 < \theta < 1.$$

The frequency vector ω is Diophantine (since, in Chapter 2.5, ω is outside the set Ω^1 as in Lemma 2.2 and ω in equation (4.1) is the ω' from Chapter 2.5),

$$|\omega \cdot s| \geq C^{-1}(m+1)^{-2}|s|^{-n} \quad \forall s \in \mathbb{Z}^n \backslash 0 \tag{4.2}$$

and is incommensurable with E (since in Chapter 2.5 ω is outside the set Ω^2 as in Lemma 2.3),

$$|\omega \cdot s + E| \geq \frac{E}{C(m)\langle s\rangle^{c_1}} \quad \forall s \in \mathbb{Z}^n. \tag{4.3}$$

The functions h, b are analytic:

$$|h|^{U_m}, |b|^{U^1_m} \leq 1.$$

We should prove that equation (4.1) has a unique analytic solution $x(q)$ and

$$|x|^{U^2_m} \leq C^e(m)/E. \tag{4.4}$$

4.1 Uniqueness of the solution.

Since the frequency ω is Diophantine and h is analytic and has zero mean value, then we can find analytic $H(q)$ such that

$$\partial H/\partial \omega = h, \quad |H|^{U^1_m} \leq C(m)$$

(see Appendix). If we substitute in (4.1) $x = e^{-iBH}y$, then for $y(q)$ we get the equation

$$-i\,\frac{\partial y}{\partial \omega} + Ey = e^{iBH}b =: p(q).$$

If p_s, $s \in \mathbb{Z}^n$, are Fourier coefficients of $p(q)$, then Fourier coefficients of the solution $y(q)$ are

$$y_s = \frac{p_s}{s \cdot \omega + E}$$

and by (4.3)

$$|y_s| \leq C(m)\, E^{-1}\langle s\rangle^{c_1}|p_s|.$$

Therefore

$$\textit{analytic solution } x(q) \textit{ of } (4.1) \textit{ is unique}$$

and

$$|y_s| \le C(m)\, E^{-1} \langle s \rangle^{c_1} \, e^{-\delta_m^1 |s|} \, e^{BC(m)}$$

since $|p| \le \exp(BC(m))$ in U_m^1. So for the solution x we have the trivial estimate

$$|x|^{U_m^2} \le e^{BC(m)}/E,$$

which implies (4.4) if $B \le C_1(m)$.

So from now on we can suppose that

$$E \ge C_*(m), \tag{4.5}$$

where E-independent $C_*(m)$ will be choosen later.

To prove estimate (4.4) under the assumption (4.5) we shall approximate the diophantine vector ω by vectors $\tilde{\omega} = \tilde{\omega}_\ell$ with rationally dependent coefficients, which are C/ℓ-close to ω (ℓ=2,3, ...). We shall find integral representation for an approximate solution of equation (4.1) with ω replaces by $\tilde{\omega}$ and prove that the approximate solution satisfies (4.4). Next we send ℓ to infinity to get estimate (4.4) for the exact solution of (4.1).

4.2 Approximations for the frequency vector.

For an integer $\ell \ge 2$ we consider the vector $\ell\omega \in \mathbb{R}^n$ and define $N_\ell \in \mathbb{Z}^n$ as the closest to $\ell\omega$ element of $\mathbb{Z}^n$. Then

$$|\omega - \ell^{-1} N_\ell| \le \frac{\sqrt{n}}{2\ell} =: \rho. \tag{4.6}$$

Lemma 4.1. *There exists real r, satisfying $|r-1| < 1/\ell$, such that $\ell E \notin r\mathbb{Z}$ and for the vector $\widetilde{\omega}$, defined as*

$$\widetilde{\omega} = \widetilde{\omega}_{\ell,r} := r\, \frac{N_\ell}{\ell},$$

and for all $s \in \mathbb{Z}^n$ one has

$$|s \cdot \widetilde{\omega} + E| \ge \frac{E}{\langle s \rangle^{c_2} C_2(m)}, \quad c_2 = c_1 + n + 1, \tag{4.7}$$

with some ℓ, ω-independent $C_2(m)$.

Proof: By (4.3), (4.6),

$$|\widetilde{\omega} \cdot s + E| \ge E \langle s \rangle^{-c_1}/C(m) - C\,|s|\rho \ge \frac{1}{2}\, E\, \langle s \rangle^{-c_1}/C(m)$$

if $\rho \le E\, \langle s \rangle^{-c_1 - 1}/C_1(m)$ or equivalently, if

$$|s| \le \left(\frac{E\,\ell}{C'(m)} \right)^{\frac{1}{c_1+1}} =: N_0.$$

So below we should consider only $|s| > N_0$.

Take any $s_0 \in \mathbb{Z}^n$ which violates (4.7) for some choice of $r \in \Delta := [1-\ell^{-1}, 1+\ell^{-1}]$. Then $|s_0 \cdot \widetilde{\omega}| \geq E/2$ and therefore the set

$$A_{s_0} = \left\{ r \in \Delta \;\Big|\; |s_0 \cdot \widetilde{\omega}_{\ell,r} + E| \leq \frac{E}{\langle s\rangle^{c_2} C_2(m)} \right\}$$

is a segment of the length $\leq 4\langle s\rangle^{-c_2}/C_2(m)$. So

$$\text{mes} \bigcup_{|s|\geq N_0} A_s \leq \frac{C}{C_2(m)}\; N_0^{-c_2+n} = \frac{C}{C_2(m)} \quad \frac{C'(m)}{E\,\ell},$$

which is less than ℓ^{-1} if $C_2(m)$ is choosen sufficiently large.

Therefore, there exists a point $r_0 \in \Delta$ which lies outside all the sets A_s, $|s| \geq N_0$. The corresponding vector $\widetilde{\omega} = \widetilde{\omega}_{\ell,r_0}$ meets all the estimates (4.7). We can choose r to be non equal to the numbers $\ell E/j$, $j = \pm 1, \pm 2, \dots$ and the lemma is proven. $\square$

We denote $\tilde{\ell} = \ell/r$. Then $\widetilde{\omega} = N_\ell/\tilde{\ell}$. We also note that, by (4.6), $|\omega - \widetilde{\omega}| \leq C/\ell$. Therefore

$$|s \cdot \widetilde{\omega}| \geq (2C(m)\,|s|^n)^{-1} \text{ if } 0 < |s| \leq C\ell^{1/(n+1)} =: L \tag{4.8}$$

Denote by h_s Fourier coefficients of $h(q)$ (recall that $h_0 = 0$) and define the resonant and the regular parts of h as

$$h^{res}(q) = \sum_{\substack{s\\ s\cdot\widetilde{\omega}=0}} h_s\, e^{i\, s\cdot q}, \qquad h^{reg}(q) = \sum_{\substack{s\\ s\cdot\widetilde{\omega}\neq 0}} h_s\, e^{i\, s\cdot q}.$$

So $h = h^{res} + h^{reg}$.

Lemma 4.2. *The functions h^{res}, h^{reg} are analytic in U^1_m and*

$$|h^{res}|^{U^1_m} \leq C(m)\,\ell^{-\frac{1}{n+1}}, \quad |h^{reg}|^{U^1_m} \leq C(m).$$

Proof: The estimate for h^{reg} is obvious (see the Appendix). In order to estimate h^{res} we observe that if $s \cdot \widetilde{\omega} = 0$, then by (4.8) $|s| \geq L$ and for q in U^1_m we have

$$|h^{res}| \leq \sum_{|s|\geq L} e^{-|s|\,m^{-2}} \leq C(m) L^{-1}$$

(see the Appendix). Thus, also the estimate for h^{res} is proven. $\square$

Lemma 4.3. *There exists an analytic in U^1_m function $\widetilde{H}$ such that $\partial\widetilde{H}/\partial\widetilde{\omega} = h^{reg}$ and $|\widetilde{H}|^{U^1_m} \leq C(m)$.*

Proof: Let us define $\widetilde{H}$ as Fourier series with the coefficients $\widetilde{H}_s$, where

$$\widetilde{H}_s = \begin{cases} 0, & s\cdot\widetilde{\omega} = 0 \\ h_s/\, s\cdot\widetilde{\omega} & \text{otherwise.} \end{cases}$$

Then, by (4.8), for q in U^1_m we have

$$|\widetilde{H}(q)| \leq 2 \sum_{|s|\leq L} |s|^n C(m)\; e^{-|s|\,m^{-2}} + 2\ell \sum_{|s|>L} e^{-|s|\,m^{-2}}$$

since $|s\cdot\widetilde{\omega}| \geq 1/\tilde{\ell} \geq 1/(2\ell)$ if $s\cdot\widetilde{\omega} \neq 0$. Now the assertion follows. For: the estimate for the first sum by some $C_1(m)$ is obvious and the estimate for the second one follows from (A4) with $k = n+1$ (see the Appendix). $\square$

4.3 Approximating equations.

Let us approximate (4.1) by the equation where ω is replaced by $\widetilde{\omega}_\ell$ and $h(q)$ – by its regular part h^{reg}:

$$(4.9) \qquad -i\,\frac{\partial x}{\partial \widetilde{\omega}} + Ex + Bh^{reg}x = b(q).$$

The substitution $x = e^{-iB\widetilde{H}}y$ with $\tilde{H}$ as as in Lemma 4.3 reduces (4.9) to

$$(4.10) \qquad -i\,\frac{\partial y}{\partial \widetilde{\omega}} + Ey = e^{iB\widetilde{H}}b =: \beta(q).$$

There exists an integral representation for the solution y of (4.10). To get the formula we consider the equation

$$(4.11) \qquad -i\mu\,\frac{\partial z}{\partial t} + Ez = f(t), \quad t \in S^1 = \mathbb{R}/2\pi\mathbb{Z}.$$

If $E \notin \mu\mathbb{Z}$, then the unique reriodic solution of (4.11) can be written as

$$z(t) = \frac{K_{E/\mu}}{\mu}\int_0^{2\pi} e^{-i(E/\mu)\tau} f(t-\tau)\,d\tau,$$

where $K_r = \frac{i}{1-e^{-i2\pi r}}$. Indeed, for $f = e^{ikt}$ we have $z = e^{ikt}/(E+k\mu)$, which is a periodic solution of (4.11). Arbitrary periodic f can be decomposed to Fourier series and the assertion follows.

Next, we take any $R \in \mathbb{T}^n$ and consider the solenoid through R,

$$(4.12) \qquad t \longmapsto R + t\,\tilde{\ell}\,\widetilde{\omega} \in \mathbb{T}^n = \mathbb{R}^n/2\pi\,\mathbb{Z}^n.$$

Since $\tilde{\ell}\,\widetilde{\omega} = N_\ell$ is an integer vector, then the solenoid is a 2π-periodic loop in $\mathbb{T}^n$. On the other hand, since for a function on $\mathbb{T}^n$ and for its restriction to the solenoid one has $\partial/\partial t = \tilde{\ell}\,\partial/\partial\widetilde{\omega}$. Then equation (4.10) restricted to the loop (4.12) takes the form (4.11) with

$$\mu = \tilde{\ell}^{-1}, \quad f(t) = \beta(R + \tilde{\ell}\widetilde{\omega}t).$$

The assumption $E \notin \mu\mathbb{Z}$ is sutisfied since $\ell E \notin r\mathbb{Z}$ by Lemma 4.1. Therefore

$$y(R) = K_{E\tilde{\ell}}\,\tilde{\ell}\int_0^{2\pi} e^{-iE\tilde{\ell}\tau}\beta(R - \tilde{\ell}\widetilde{\omega}\tau)\,d\tau.$$

Finally, we denote $\nu = \widetilde{\omega}/|\widetilde{\omega}|^2$, $z = \tilde{\ell}\tau$ (so $E\tilde{\ell}\tau = E\nu\cdot\widetilde{\omega}z$) and obtain the integral representation for (the unique) solution x of (4.9):

$$(4.13) \qquad x(q) = K_{E\tilde{\ell}}\int_0^{2\pi\tilde{\ell}} e^{-iE(\nu\cdot Q + (B/E)\,(\widetilde{H}(q) - \widetilde{H}(q-Q))}\,b(q-Q)\big|_{Q=\widetilde{\omega}z}dz$$

– now we treat Q as a point in $\mathbb{R}^n$ and H, b as analytic 2π-periodic functions.

So we have represented $x(q)$ as a rapidly oscillating integral Fourier with a complex (for complex q) phase function.

4.4 Study of the oscillating integral (4.13).

We denote $\gamma = B/E$, $S(q,Q) = \widetilde{H}(q) - \widetilde{H}(q-Q)$ and observe that

i) $\gamma \le CE^{\theta-1} \le C_*(m)^{\theta-1} \quad (\theta < 1)$
(see (4.5));

ii) $S(q,0) \equiv 0$;

iii) for q in U^2_m the function S is analytic in Q and

$$|\nabla_Q S(q,\cdot)|^{U(m^{-2}/C)} + |S(q,\cdot)|^{U(m^{-2}/C)} \le C(m)$$

(by Lemma 4.3 and the Cauchy estimate).

Let us consider the substitution

$$Q = R + f(R)\widetilde{\omega} \equiv \Phi(R),$$

where $R \in \mathbb{T}^n$ and f is a complex function. Then

$$\nu \cdot Q + \gamma S(q,Q)\big|_{Q=\Phi(R)} = \nu \cdot R + f(R) + \gamma S(q, R + f(R)\widetilde{\omega}).$$

Consider the equation for $f(R)$:

$$f(R) + \gamma S(q, R + f(R)\widetilde{\omega}) = 0.$$

If $C_*(m)$ in (4.5) is sufficiently large, then by i), iii) and the implicit function theorem the equation has the only solution $f(R) = f(q,R)$ which is complex-analytic in $R \in U(m^{-2}/C)$ and satisfies the estimate

$$|f|^{U(m^{-2}/C)} \le 1/C_{**}(m),$$

where C_{**} goes to infinity with C_*. Besides due to ii) $f(0,q) \equiv 0$.

Now we treat (4.13) as an integral of a holomorphic function along the segment $\Delta = \widetilde{\omega}[0, 2\pi\widetilde{\ell}]$ in the complex plane $\mathbb{C}^1 = \mathbb{C}\widetilde{\omega}$, namely

$$x(q) = K_{E\tilde{\ell}} \int_{\Delta} e^{-iE(\widehat{\omega}\cdot R + \gamma S(q,R))}\, b(q-R)\, dR/|\widetilde{\omega}|.$$

In this integral we can replace the contour $\Delta = \{R\}$ by $\Phi(\Delta) = \{Q\} \subset \mathbb{C}^1$ since both the contours lie in the domain of analiticity and their ends conside. As $f(R) + \gamma S(q, \Phi(R)) = 0$, then

$$\begin{aligned} x(q) &= K_{E\tilde{\ell}} \int_{\Phi(\Delta)} e^{-iE(\widehat{\omega}\cdot Q + \gamma S(q,Q))}\, b(q-Q) \frac{dQ}{|\widetilde{\omega}|} = \\ &= K_{E\tilde{\ell}} \int_{\Delta} e^{-iE\nu\cdot R}\, b(q - Q(R))(1 + |\widetilde{\omega}|\, f'(R)) \frac{dR}{|\widetilde{\omega}|} = K_{E\tilde{\ell}} \int_{\Delta} e^{-iE\nu\cdot R} g(R) \frac{dR}{|\widetilde{\omega}|}, \end{aligned} \tag{4.14}$$

where we use the same notation f for the function f restricted to $\mathbb{C}^1$ and denote

$$g(R) = b(q - Q(R))\ (1 + |\widetilde{\omega}|\ f'(R)), \quad R \in \mathbb{C}^1.$$

This function is analytic in $U(m^{-2}/2C)$ and is bounded by some constant C_1 if $C_*(m)$ is sufficiently large.

In order to estimate the r.h.s. of (4.14) we expand g in Fourier series,

$$g = \sum g_s \, e^{is\cdot R}, \quad |g_s| \le C_1 \, e^{-|s|/(2Cm^2)}, \tag{4.15}$$

(see (A11)). Now we have

$$x(q) = K_{E\tilde{\ell}} \sum_s g_s \int_0^{2\pi\tilde{\ell}} e^{-i(E+\widetilde{\omega}\cdot s)} dt = K_{E\tilde{\ell}} \sum \frac{i g_s}{E - \widetilde{\omega}\cdot s} \left(e^{-iE2\pi\tilde{\ell}} - 1\right),$$

since $\widetilde{\omega}\cdot\tilde{\ell}$ is an integer. Therefore $x(q) = \sum g_s/\left(E + \widetilde{\omega}\cdot s\right)$ and by (4.7), (4.15) and (A2) we have

$$|x(q)| \le C_1(m) \sum_s e^{-|s|/(2Cm^2)} \langle s\rangle^{-c_2} \le C(m)/\,E \quad \text{for} \quad q \in U_m^2. \tag{4.16}$$

We stress that this estimate is l-independent.

4.5 Transition to limit.

Changing the notation, we denote by $x_\ell(q)$ the solution of (4.9) we have constructed, and rewrite (4.9) as

$$-i\widetilde{\omega}_\ell \cdot \nabla x_\ell + E\, x_\ell + Bh(q)x_\ell = b(q) + z_\ell(q),$$

where $z_\ell = B\, h^{res} x_\ell$. Then by (4.16) and Lemma 4.2, $|z_\ell|^{U_m^2} \le C(m)B/\bigl(E\,\ell^{1/(n+1)}\bigr)$ with some ℓ-independent $C(m)$. By (4.16) the sequence $\{x_\ell\}$ contains a subsequence such that $\{x_\ell\}$ and $\{\nabla x_\ell\}$ converge uniformly in U_m^3, namely

$$x_\ell \longrightarrow x, \qquad \nabla x_\ell \longrightarrow \nabla x.$$

As $z_\ell \longrightarrow 0$ and $\widetilde{\omega}_\ell \longrightarrow \omega$, then $x(q)$ is a solution of (4.1) in U_m^3 which meets (4.16) for $q \in U_m^3$. Clearly, we can use intermediate domains like $U_m^{1/2}$ to prove that x solves (4.1) and meets (4.16) for q in U_m^2.

Thus, the proof of the estimate (4.4) is completed. □

5. Proof of Theorem 2

In this chapter we study the linearized equation (1.10). To do it we also consider linearization of the transformed equation (2.38) about the transformed solution $\mathfrak{h}_m = \Sigma_\infty^m(\mathfrak{h}_\infty(t))$:

$$\dot{\eta}_m = V_{\mathcal{H}_m}(\mathfrak{h}_m(t))_*\eta_m. \tag{5.1}$$

This equation coinsides with (1.10) if $m = 0$. The linear transformation

$$\mathcal{L}_m(t) := \Sigma_m^0(\mathfrak{h}_m(t))_*$$

sends solutions of (5.1) to solutions of (1.10). By (2.35) the limiting linear maps $\mathcal{L}_\infty(t)$ exist and define isomorphisms of the spaces $Z_\theta = \mathbb{R}^n \times \mathbb{R}^n \times Y_\theta$, $|\theta| \le d$. Moreover, each linear map $\mathcal{L}_m(t)$ is symplectic since the maps Σ_m^0 are. The limiting maps $\mathcal{L}_\infty(t)$ are symplectic as well.

For $0 \le m \le \infty$ we have the estimates

$$\|\mathcal{L}_m(t)\|_{\theta,\theta} + \|\mathcal{L}_m^{-1}(t)\|_{\theta,\theta} \le 3, \quad |\theta| \le d, \tag{5.3}$$

(see (2.33), (2.35)). Since the linearized equation (1.10) is well-posed by assumptions of the theorem, then due to (5.3) equations (5.1) also are well posed: for any m, the flow-maps $S_{(m)\tau}^{\tau+t}(\mathfrak{h}_m(\tau))_*$ of (5.1) are such that

$$\|S_{(m)\tau}^{\tau+t}(\mathfrak{h}_m(\tau))_*\|_{\theta,\theta} \le Ce^{C_2 t} \text{ for any } t \text{ and any } |\theta| \le d. \tag{5.2}$$

Because (5.3), to estimate solutions $\eta_0(t)$ of (5.1) with $m = 0$ is equivalent to estimate their transformations $\eta_\infty(t) = (\mathcal{L}_\infty(t))^{-1}\eta_0(t)$. We can not directly go to limit in (5.1) to write for $\eta_\infty(t)$ the limiting equation, instead we shall obtain estimates for η_∞ by examining p-, q- and y-components of the curves η_m with large m.

For any $0 \le r \le m \le \infty$ we define the linear transformations $\mathcal{L}_m^r$,

$$\mathcal{L}_m^r(t) = \Sigma_m^r(\mathfrak{h}_m(t))_*, \quad \infty \ge m \ge r.$$

Since $\mathcal{L}_m^r = (\mathcal{L}_r)^{-1} \circ \mathcal{L}_m$, then by (2.35) we have

$$\|\mathcal{L}_m^r - \mathrm{id}\|_{\theta,\theta} \le C\varepsilon_r^\rho. \tag{5.4}$$

Next we write (5.1) as a system of equations for $\eta_m = (\eta_p, \eta_q, \eta_y)$, omitting dependence of solutions on m (and on the parameter ω which is now irrelevant):

$$\begin{cases} \dot\eta_p = -\varepsilon_m \nabla_{q,p} H_m \eta_p - \varepsilon_m \nabla_{q,q} H_m \eta_q - \varepsilon_m \nabla_{q,y} H_m \eta_y, \\ \dot\eta_q = \varepsilon_m \nabla_{p,p} H_m \eta_p + \varepsilon_m \nabla_{p,q} H_m \eta_q + \varepsilon_m \nabla_{p,y} H_m \eta_y, \\ \dot\eta_y = JA_m(q_m(t))\eta_y + \varepsilon_m J\nabla_{y,p} H_m \eta_p + \varepsilon_m J\nabla_{y,q} H_m \eta_q + \varepsilon_m J\nabla_{y,y} H_m \eta_y. \end{cases} \tag{5.1$'$}$$

We need the following refinement of estimates (2.7), (2.8):

Lemma 5.1. *The Hamiltonian $\mathcal{H}_m(\mathfrak{h})$ in O_m meets the following estimates:*

$$\varepsilon_m \left\| \frac{\partial}{\partial p_j} \nabla_y H_m(\mathfrak{h}) \right\|_{d_c} \le C\, e(m), \quad j = 1, \dots, n, \tag{5.5}$$

(the numbers $e(m)$ were defined in Chapter 2.1, C is an m-independent constant), and

$$\varepsilon_m \left| \frac{\partial}{\partial p_j} \frac{\partial}{\partial p_k} H_m(\mathfrak{h}) \right| \le C^e(m), \quad j, k = 1, \dots, n.$$

Proof: For $m = 0$ the estimate (5.5) follows from (1.5) and the Cauchy estimate since the domain of analyticity Q^c is ε-independent. Suppose that the estimate is proven for $m = m$ and show that it also holds for $m = m + 1$. Since $(\partial/\partial p_j)\nabla_y H_{2m} = 0$ (see Chapter 2.3) and $H_m + H_{2m} + H_{3m}$, then $\varepsilon_m H_{3m}$ also meets (5.5). In Chapter 2.7 we constructed $\varepsilon_{m+1} H_{m+1}$ in the form $\varepsilon_{m+1}H_{m+1}=\varepsilon_m H_{3m}+\Delta_3 H+\Delta_4 H$ (see (2.29)). By Lemma 2.7 (the second estimate in (2.31)) and the Cauchy estimate, for $\mathfrak{h}$ in O_{m+1} and $j=1,\dots,n$, $l=3,4$ we have

$$\|\frac{\partial}{\partial p_j}\nabla_y\Delta_l H(\mathfrak{h})\|_{d_c}\le C^e(m)\varepsilon_m^{1/3}\le\frac{1}{2K_*(m+1)^2}.$$

So (5.5) for $m=m+1$ follows.

Proof of the second estimate is analogous. $\square$

By (2.7), (2.8) and the last lemma system (5.1′) can be abbreviated as

$$\begin{cases}\dot\eta_p=O_{p,\eta}(\varepsilon_m^\rho)\eta,\\ \dot\eta_q=O_{q,p}(C^e(m))\eta_p+O_{q,y}(1)\eta_y+O_{q,\eta}(\varepsilon_m^\rho)\eta\\ \dot\eta_y=JA_m(q_m(t))\eta_y+O_{y,p}(1)\eta_p+O_{y,\eta}(\varepsilon_m^\rho)\eta,\end{cases}\tag{5.6}$$

where $O_{p,\eta}(\varepsilon_m^\rho)$ stands for a time-dependent linear operator $Z_d\to\mathbb{R}^n$, $\eta\to p$, of the norm $O(\varepsilon_m^\rho)$ and similar with $O_{q,p}(C^e(m)),\dots,O_{q,\eta}(\varepsilon_m^\rho)$. The linear operators $O_{y,p}(1),O_{y,\eta}(\varepsilon_m^\rho)$ are bounded operator valued in Y_{d_c}.

For $j=1,2,\dots$ let us denote by $\xi_j\in Z$ the unit vector of the form

$$\xi_{j0}=(0,0,y_{j0}),\quad y_{j0}=y^j w_j+\bar{y^j}w_{-j},\quad \|y_{j0}\|_d=1\tag{5.7}$$

(the complex basis w_j was defined in Chapter 2.5) and denote

$$\xi_{j0}^{(m)}:=\mathcal{L}_\infty^m(0)\,\xi_{j0}=\xi_{j0}+O(\varepsilon_m^\rho)$$

(the second equality follows from (5.4)). Let $\xi_j^{(m)}(t)$ be the solution of (5.1) such that

$$\xi_j^{(m)}(0)=\xi_{j0}^{(m)}.$$

For $m=0,1,\dots$ the linear map $\mathcal{L}_m$ sends $\xi_j^{(m)}(t)$ to $\xi_j^{(0)}(t)$.

Since a diagonal element $\beta_j^{(m)}+\lambda_j^{(m)}$ of the operator $A_m(q;\omega)$ equals

$$\langle\text{diagonal element } \lambda_j(\omega) \text{ of the operator } A(\omega)\rangle+2\varepsilon_1 b_j^{(1)}(q;\omega)+\dots 2\varepsilon_m b_j^{(m)}(q;\omega),$$

where $b_j^{(l)}$ stands for the contribution from the l-th step. Since the function $b_j^{(l)}(\cdot;\omega)$ is analytic in U_l^1 and bounded there by $j^{d_2}C(l)\varepsilon_l^{-2/3}$, then for ω in Ω_ε we have the convergences

$$\beta_j^{(m)}(q;\omega)\longrightarrow\beta_j^\infty(q;\omega),\quad \lambda_j^{(m)}(\omega)\longrightarrow\lambda_j^\infty(\omega)\quad\text{as } m\to\infty.$$

The limiting maps are such that

$$|\beta_j^\infty|^{\mathcal{O},\Omega_\varepsilon}+|\lambda_j^\infty-\lambda_j|^{\Omega_\varepsilon,\mathrm{Lip}}\le C\varepsilon_0^\rho j^{d_H}.\tag{5.8}$$

We denote by A_∞ the limiting operator $A_\infty(q;\omega)=\operatorname{diag}\{\lambda_j^\infty+\beta_j^\infty\mid j\ge1\}$ and consider the corresponding limiting equation in the space Y_d:

$$y=JA_\infty(q_0+\omega' t;\omega)y,\quad \omega'=\Lambda_\infty(\omega).\tag{5.9}$$

We recall that $\operatorname{mes}(\Omega\backslash\cap\Omega_m) \leq \gamma/2$. So we can replace $\Omega_\varepsilon = \cap\Omega_m$ by its subset, which we also denote Ω_ε, such that the new one still satisfies assertion a) of Theorem 1 (i.e. $\operatorname{mes}\Omega\backslash\Omega_\varepsilon \leq \gamma$) and

$$
|\omega' \cdot s| \geq C^{-1}|s|^{-n} \quad \forall s \in \mathbb{Z}^n\backslash\{0\}, \quad \forall \omega \in \Omega_\varepsilon \tag{5.10}
$$

with some $C = C(\gamma)$, see the Appendix.

Consider the solution $y(t) = y_j(t)$ of (5.9) with the initial data

$$
y(0) = y_{j0} \quad \text{as in (5.7).} \tag{5.11}
$$

The solution has the form $y_j(t) = y^j(t)w_j + y^{-j}w_{-j}$, where y^j and y^{-j} are complex-conjugated complex functions and for y^j we have the equation

$$
\dot{y}^j = i(\lambda_j^\infty + \beta_j(q_0 + \omega' t))y^j.
$$

Since β_j is analytic with zero mean value and ω' is Diophantine (see (5.8), (5.10)), then there exists an analytic function $\mathfrak{z}_j(q)$, real for real q, such that

$$
\frac{\partial \mathfrak{z}_j}{\partial \omega'}(q) = \beta_j(q)
$$

(see the Appendix). The substitution

$$
y^j(t) = e^{i\mathfrak{z}_j(q_0+\omega' t)} z^j(t)
$$

implies for z^j the equation $\dot{z}^j = i\lambda_j^\infty z^j$. So $|z^j(t)| = \text{const}$ and $|y^j(t)| =$const since $\mathfrak{z}(q)$ is real. – Solution $y_j(t)$ of (5.9), (5.11) is a fast (if $j \gg 1$) rotation in the plane $\mathbb{R}\varphi_j^+ + \mathbb{R}\varphi_j^-$ with a time-dependent velocity and $\|y_j(t)\|_d \equiv 1$.

Now let us consider the curve $\eta_j^{(m)}(t)$ in Z,

$$
\eta_j^{(m)} = (0, \int_0^t O_{q,y}(1) y_j(\tau) d\tau,\, y_j(\tau)),
$$

where $O_{q,y}(1)$ is the linear operator from the second equation in (5.6). Clearly, its Z_d-norm is bounded by $C(t+1)$. Therefore $\eta_j^{(m)}$ satisfies (5.6) with a disparity $\Delta_j^{(m)}(t) \in Z$ such that $\|\Delta_j^{(m)}(t)\|_{d_c} \leq C(t+1)\varepsilon_m^\rho$. Since $\eta_j^{(m)}(0) = (0, 0, y_j(0)) = \xi_{j0}$ and the flow-maps $S^t_{(m)\tau}$ of equation (5.1) satisfies (5.2), then we get the estimate for divergence of $\eta_j^{(m)}$ from the exact solution $\xi_j^{(m)}$:

$$
\|\xi_j^{(m)}(t) - \eta_j^{(m)}(t)\|_{d_c} \leq C\varepsilon_m^\rho e^{C_1 t} \tag{5.12}
$$

with some C, C_1. The operator $\mathcal{L}_m^r$ sends $\xi_j^{(m)}$ to $\xi_j^{(r)}$ and satisfies (4.3). Therefore by (5.12), $\eta_j^{(m)}$ converges (as m grows) to

$$
\xi_j^{(\infty)} = (\mathcal{L}_\infty)^{-1}\xi_j^{(0)}
$$

uniformly for bounded t's. Denoting by Π_p, Π_q, Π_y the natural linear projectors which send Z to $\mathbb{R}^n_p, \mathbb{R}^n_q$ and Y respectively, we get from this convergence that

$$\Pi_p \xi_j^{(\infty)} \equiv 0, \quad \|\Pi_y \xi_j^{(\infty)}(t)\|_d \equiv 1. \tag{5.13}$$

Now let us denote by S^{τ_1,τ_2} the linear operator in Z_d which sends $\eta(\tau_1)$ to $\eta(\tau_2)$, where $\eta(\tau)$ is a solution of (1.10). Let $\tilde{S}^{\tau_1,\tau_2}$ be the conjugated operator:

$$\tilde{S}^{\tau_1,\tau_2} = \mathcal{L}_\infty(\tau_2)^{-1} \circ S^{\tau_1,\tau_2} \circ \mathcal{L}_\infty(\tau_1)$$

(it sends $\eta_\infty(\tau_1)$ to $\eta_\infty(\tau_2)$). We write Z_d as $\mathbb{R}^n_p \times \mathbb{R}^n_q \times Y_d$ and accordingly write $\tilde{S}^{\tau_1,\tau_2}$ in the matrix form,

$$\tilde{S}^{\tau_1,\tau_2} = \begin{pmatrix} s_{pp} & s_{pq} & s_{py} \\ s_{qp} & s_{qq} & s_{qy} \\ s_{yp} & s_{yq} & s_{yy} \end{pmatrix}.$$

As $\xi_j^{(0)} = \mathcal{L}_\infty(0)\xi_j$, then $\tilde{S}^{0,t}(\xi_j) = \xi_j^{(\infty)}(t)$ and we get from (5.13) that

$$s_{py} = 0, \quad \|s_{yy}\|_{d,d} \equiv 1.$$

For each $q \in \mathbb{T}^n$, the map Σ_ω sends the curve $q + \omega' t \in \mathbb{T}^n$ to a solution of the initial equation (1.6). So Σ_ω conjugates translation of $\mathbb{T}^n$ along ω' with the flow of (1.6) and its linearization $\Sigma_{\omega *} = \mathcal{L}_\infty\big|_{\{0\}\times\mathbb{R}^n_q\times\{0\}}$ conjugates linearization of the translation with the corresponding operator $\tilde{S}$. It means that

$$s_{pq} = 0, \quad s_{qq} = \mathrm{id}, \quad s_{yq} = 0. \tag{5.14}$$

Each map $\tilde{S}^{\tau_1,\tau_2}$ is symplectic as a composition of symplectic maps. Hence,

$$\alpha_2[\tilde{S}^{\tau_1,\tau_2}(\delta p_1, 0, 0), (0, \delta q_2, 0)] = \langle \delta p_1, \delta q_2 \rangle_{\mathbb{R}^n} \quad \forall\, \delta p_1, \delta q_2 \in \mathbb{R}^n.$$

Because (5.14) this implies that $\langle s_{pp}\delta p_1, \delta q_2 \rangle_{\mathbb{R}^n}$. Hence,

$$s_{pp} = \mathrm{id}. \tag{5.15}$$

By the estimate (1.9),

$$\|\tilde{S}^{\tau_1,\tau_2}\|_{d,d} \le C \|S^{\tau_1,\tau_2}\|_{d,d} \le C e^{C_1|\tau_1-\tau_2|}. \tag{5.16}$$

Now we can estimate the norm of the operator $\tilde{S}^{0,T}$ with large T. To do it let us write Z_d as

$$Z_d = \mathbb{R}^n_p \times E, \quad E = \mathbb{R}^n_q \times Y_d = \{\mu = (q, y)\}.$$

Accordingly we enlarge blocs of $\tilde{S}^{\tau_1,\tau_2}$ and write this operator as

$$\tilde{S}^{\tau_1,\tau_2} = \begin{pmatrix} \mathfrak{s}_{pp} & \mathfrak{s}_{p\mu} \\ \mathfrak{s}_{\mu p} & \mathfrak{s}_{\mu\mu} \end{pmatrix}.$$

By (5.14), (5.15), (5.16) we have:

$$\mathfrak{s}_{p\mu}=0,\ \ \mathfrak{s}_{pp}=\mathrm{id},\ \ \|\mathfrak{s}_{\mu\mu}\|=1,\ \ \|\mathfrak{s}_{\mu p}\|\le C'e^{C_1|\tau_1-\tau_2|}. \tag{5.17}$$

For any $(p_0,\mu_0)\in Z_d$ and $T\in\mathbb{N}$ we can write $\tilde S^{0,T}(p_0,\mu_0)$ as

$$\tilde S^{0,T}(p_0,\mu_0)=\tilde S^{T-1,T}\circ\cdots\circ\tilde S^{0,1}(p_0,\mu_0).$$

Denoting $(p_j,\mu_j)=\tilde S^{j-1,j}\circ\cdots\circ\tilde S^{0,1}(p_0,\mu_0)$ and using (5.17) we see that

$$|p_j|=|p_{j-1}|,\quad \|\mu_j\|_d\le\|\mu_{j-1}\|_d+C_2|p_{j-1}|\quad\text{where}\quad C_2=C'e^{C_1}.$$

Therefore we get the component-wise inequality:

$$\begin{pmatrix}|p_T|\\ \|\mu_T\|_d\end{pmatrix}\le\begin{pmatrix}1&0\\ C_2&1\end{pmatrix}^T\begin{pmatrix}|p_0|\\ \|\mu_0\|_d\end{pmatrix}=\begin{pmatrix}|p_0|\\ \|\mu_0\|+(C_2+T)|p_0|\end{pmatrix}.$$

We have seen that any solution $\eta(t)$ of (1.10) meets the estimate

$$\|\eta(t)\|_d\le 3\|\eta_\infty(t)\|_d\le(C_1+tC_2)\|\eta(0)\|_d$$

and Theorem 2 is proven.

Remark. If $d_2\le 0$ then the functions b_j in (2.19) can be replaced by their mean-values in q since the variable part of the diagonal of the opeartor h^{1yy} can be killed by the transformation S_m (see [K]). So the functions β_j^∞ (see (5.8)) are zero and the thansformation $\mathcal{L}_\infty$ reduces (1.10) to the equation in $\mathbb{R}^n\times\mathbb{R}^n\times Y_d$ with a matrix

$$\begin{pmatrix}0&0&0\\ m_{qp}&0&m_{qy}\\ m_{yp}&0&JA_\infty\end{pmatrix},$$

where A_∞ is a diagonal constant-coefficient matrix. The Jordan part m_{qp} and "half" of the matrices m_{qy},m_{yp} can be nontrivial.[7] In [K] we claimed (without giving datailed arguments) that the blocks m_{qp},m_{qy},m_{yp} vanish, which is not true. Still, the main assertion in [K] concerning the linearized equation – that its Lyapunov exponents vanish – is correct as shows the detailed proof given above.

Appendix. Some inequalities for Fourier series

Let B be a complex Banach space and $f:U(\delta)\longrightarrow B$ be a complex-analytic map. We can write f as Fourier series,

$$f(q)=\sum_{s\in\mathbb{Z}^n}f_s\,e^{is\cdot q},\ \ f_s=\int_{\mathbb{T}^n}f(q)\,e^{-iq\cdot s}dq/(2\pi)^n\in B.$$

[7]Example: Take in (1.6) $n=1$ and $H(p,q,y)=p^2$. Then $\Sigma_\omega(q)=(0,q,0)$ and $m_{qp}=\varepsilon/2$.

If $\|f\| \leq 1$ in $U(\delta)$, then we can replace the integrating over $\mathbb{T}^n$ by integrating over the set $\mathbb{T}^n - i(\delta-\varepsilon)\,\frac{s}{|s|} \subset U(\delta)$ to get that $\|f_s\| \leq \exp(-|s|\,(\delta-\varepsilon))$ for each positive ε. Thus,

$$\|f_s\| \leq e^{-|s|\,\delta}. \tag{A1}$$

Conversely, if for some $d \geq 0$ we have $\|f_s\| \leq \langle s\rangle^d\ e^{-|s|\,\delta}$ for all s, and if $0<\gamma<\delta$, then

$$\|f\|^{U(\delta-\gamma)} \leq \sum_{s\in\mathbb{Z}^n} \langle s\rangle^d\ e^{-\gamma\,|s|} \leq C \int\limits_{\mathbb{R}^n} |x|^d\ e^{-\gamma\,|x|} dx = C_d \gamma^{1-n-d}. \tag{A2}$$

In particular, if $f_0 = 0$ and

$$\omega \cdot \nabla g(q) = f(q)$$

where ω is an n-vector such that

$$|\omega \cdot s| \geq \widetilde{C}\,|s|^{-n} \quad \forall s \in \mathbb{Z}^n \backslash \{0\},$$

then $g_s = f_s/(is\cdot\omega)$. So if $\|f\|^{U(\delta)} \leq 1$, then by (A1), (A2)

$$\|g\|^{U(\delta-\gamma)} \leq (C_n/\widetilde{C})\gamma^{-2n}.$$

If for any analytic function $f(q)$ such that $\|f\|^{U(\delta)} \leq 1$, we cut off its low-frequency part and for any $R>0$ define f^R as

$$f^R(q) = \sum_{|s|\geq R} f_s\ e^{is\cdot q},$$

then for any $0<\gamma<\delta$ we have:

$$\begin{aligned} \|f^R\|^{U(\delta-\gamma)} &\leq \sum_{|s|\geq R} e^{-|s|\,\gamma} \leq C \int\limits_R^\infty e^{-t\gamma} t^{n-1} dt = \\ &= C \sum_{m=1}^n \gamma^{-m}\ \frac{(n-1)!}{(n-m)!}\ e^{-\gamma\,R}\, R^{n-m}. \end{aligned} \tag{A3}$$

Take any $k \in \mathbb{N}$. Then by (A3)

$$\|f^R\|^{U(\delta-\gamma)} \leq C_{n,k}\, R^{-k} \gamma^{-n-k}. \tag{A4}$$

References

[BiK1] R.F. Bikbaev, S.B. Kuksin, *"On the parametrization of finite-gap solutions by frequency vector and wave-number vector and a theorem of I. Krichever"*. Lett. Math. Phys. **28** (1993), pp. 115–122.

[BiK2] R.F. Bikbaev, S.B. Kuksin, *"A periodic boundary value problem for the Sine-Gordon equation, small Hamiltonian perturbations of it, and KAM-stability of finite-gap tori"*. St.-Petersburg Math. J. **4** (1993), pp. 439–468.

[BoK1] A.I. Bobenko, S.B. Kuksin, *"Finite-gap periodic solutions of the KdV equation are nondegenerate"*. Phis. Lett. **A 161:3** (1991), pp. 274–276.

[BoK2] A.I. Bobenko, S.B. Kuksin, *"The nonlinear Klein-Gordon equation on an interval as a perturbed Sine-Gordon equation"*. Comment. Math. Helv. **70** (1995), pp. 63–112.

[Bou] J. Bourgain, *"Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE"*. International Mathematics Research Notes, No. 11 (1994), pp. 1–23.

[CW] W. Craig, C.E. Wayne, *"Newton's method and periodic solutions of nonlinear wave equation"*. Comm. Pure. Appl. Math. **46** (1993), pp. 1409–150l.

[DMN] B.A. Dubrovin, V.F. Matveev, S.P. Novikov, *"Nonlinear equations of Korteweg – de Vries type, finite zone linear operators, and abelian varieties"*. Russ. Math. Surv. **31:1** (1976).

[E] L.H. Eliasson *"Perturbations of stable invariant tori"*. Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV Ser.15 (1988), pp. 115–147.

[HLP] G.H. Hardy, J.E. Littlewood, and G. Polya, *"Inequalities"*. Cambridge: Cambridge Univ. Press, (1967).

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

[K] S.B. Kuksin, *"Nearly integrable infinite-dimensional Hamiltonian systems"*. Lecture Notes Math. **1556**, Springer-Verlag (1993).

[K1] S.B. Kuksin, *"Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum"*. Funct. Anal. Appl. **21** (1987), pp. 192–205.

[K2] S.B. Kuksin, *"Perturbation theory for quasiperiodic solutions of infinite-dimensional Hamiltonian systems and its applications to the Korteweg – de Vries equation"*. Math. USSR Sbornik **64** (1989), pp. 397–413.

[KP] S.B. Kuksin, J. Pöschel *" Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation"*. Annals of Mathematics **143** (1996), pp. 149–179.

[McK] H.P. McKean, E. Trubowitz, *"Hill's operator and hyperelliptic function theory in the presence of infinitely many branch points"*. Comm. Pure Appl. Math. **29** (1976), pp. 143–226.

[P] J. Pöschel *" A KAM-theorem for some nonlinear PDEs"*. Ann. Scuola Norm. Sup. Pisa, Cl. Sci., IV Ser. 15 **23** (1996), pp. 119-148.

[RS] M. Reed, B. Simon, *"Methods of modern mathematical physics"*. Vols. 1 2, Academic Press (1980).

[W] C.E. Wayne *"Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory"*. Commun. Math. Phys. **127** (1990), pp. 479–528.

Steklov Mathematical Institute, Vavilova St., 42, 117966 Moscow
E-mail address: kuksin@sci.lpi.ac.ru