Continuous averaging in multi-frequency slow-fast systems

A.Pronin, D.Treschev

Abstract

It is well-known that in real-analytic multi-frequency slow-fast ODE systems the dependence of the right-hand sides on fast angular variables can be reduced to an exponentially small order by a near-identical change of the variables. Realistic constructive estimates for the corresponding exponentially small terms are obtained.

1 Motivations

Consider a slow-fast system

$$\dot X = \Omega(Y) + \varepsilon U(X,Y,\varepsilon), \quad \dot Y = \varepsilon V(X,Y,\varepsilon). \tag{1.1}$$

The variables $X \in \mathbf{T}^N$ and $Y \in D \subset \mathbf{R}^M$ are called respectively fast and slow. The parameter $\varepsilon \geq 0$ is assumed to be small.

There is a classical idea to describe the behavior of Y by the averaged system

$$\dot Y = \varepsilon \overline{V}(Y), \quad \overline{V} = \frac{1}{(2\pi)^N} \int_{\mathbf{T}^N} V(X,Y,0)\, dX.$$

It is well known that for $N \geq 2$ such averaging can be unsatisfactory because of the presence of resonances.

Let us fix the frequencies $\Omega^0 = \Omega(Y^0)$ and consider a small neighborhood of the point Y^0 in the domain D.

The vector Ω^0 can be resonant. In this case there exists an integer matrix K, $\det K = 1$ such that $K\Omega^0 = \begin{pmatrix} \omega \\ 0 \end{pmatrix}$, where the vector $\omega \in \mathbf{R}^n, n \leq N$ is nonresonant.

Introducing the notation

$$Y = Y_0 + \sqrt{\varepsilon} y'', \quad X = K^{-1} \begin{pmatrix} x \\ y' \end{pmatrix},$$
$$x \in \mathbf{T}^n, \quad y' \in \mathbf{R}^{N-n}, \quad y'' \in \mathbf{R}^M,$$
$$y = \begin{pmatrix} y' \\ y'' \end{pmatrix} \in \mathbf{R}^m, \quad m = N + M - n,$$

we transform the system (1.1):

$$\dot x = \omega + \sqrt{\varepsilon} u(x,y,\sqrt{\varepsilon}), \quad \dot y = \sqrt{\varepsilon} v(x,y,\sqrt{\varepsilon}). \tag{1.2}$$

2 Set up of the problem

Redefining for convenience $\sqrt{\varepsilon} \to \varepsilon$ and putting $u = \overline{u} + \hat u$, $v = \overline{v} + \hat v$,

$$\overline{u}(y) = \frac{1}{(2\pi)^n} \int_{\mathbf{T}^n} u(x,y,0)\, dx, \quad \overline{v}(y) = \frac{1}{(2\pi)^n} \int_{\mathbf{T}^n} v(x,y,0)\, dx,$$

we transform (1.2) as follows:

$$\dot x = \omega + \varepsilon(\overline{u}(y) + \hat u(x,y,\varepsilon)), \quad \dot y = \varepsilon(\overline{v}(y) + \hat v(x,y,\varepsilon)). \tag{2.1}$$

The functions $\overline{u}, \overline{v}, \hat{u}, \hat{v}$ are assumed to be real-analytic. Obviously, average in x of the functions $\hat{u}, \hat{v}$ equals $O(\varepsilon)$.

In the present paper we solve the following problem. To weaken the dependence of the right-hand side of the system (2.1) from fast variables by a close to the identity change of the variables $(x \bmod 2\pi, y) \mapsto (x_\bullet \bmod 2\pi, y_\bullet)$.

We put

$$z = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \overline{\omega} = \begin{pmatrix} \omega \\ 0 \end{pmatrix}, \quad \overline{w} = \begin{pmatrix} \overline{u} \\ \overline{v} \end{pmatrix}, \quad \hat{w} = \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix}.$$

Note that the vector fields $\overline{\omega}$ and $\overline{w}$ commute.

The system (2.1) takes the form

$$\dot{z} = \overline{\omega} + \varepsilon(\overline{w} + \hat{w}). \tag{2.2}$$

We have the following assumptions.

A1. Diophantine condition.

We assume that $n \geq 2$ and the frequency vector ω is Diophantine:

$$|\langle k, \omega\rangle| \geq \gamma_0 ||k||^{-\gamma}, \qquad k \in \mathbf{Z}^n \setminus \{0\}. \tag{2.3}$$

Recall that $\gamma \geq n-1$. Otherwise the set of vectors ω, satisfying (2.3), is empty.

To formulate the next assumption, we need some definition. Let g^t be the phase flow of the system

$$\dot{z} = \overline{w}(y). \tag{2.4}$$

The maps g^t are such that for some functions $\mathcal{X}$ and $\mathcal{Y}$

$$g^t \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + \mathcal{X}(y,t) \\ y + \mathcal{Y}(y,t) \end{pmatrix},$$

where $\mathcal{X}(y,0) = 0$, $\mathcal{Y}(y,0) = 0$. In particular, for any real-analytic function $f(y)$ with values in $\mathbf{C}^{n+m}$ and the vector $k \in \mathbf{Z}^n$ we put $f_k = f(y)e^{i\langle k,x\rangle}$. Then the function

$$\mathbf{g}_k^s f = e^{-i\langle k,x\rangle} g_*^{-is}(f_k \circ g^{is}), \qquad s \in \mathbf{C}$$

does not depend on x. Here g_*^t is the differential of the map g^t. Below the operators $\mathbf{g}_k^s$ with real s are used multiply. Note that $\mathbf{g}_k^s$, $s \in \mathbf{R}$ include shifts along g^t in purely imaginary direction.

We put

$$\begin{aligned} \Sigma_q &= \{x \in \mathbf{C}^n/(2\pi\mathbf{Z})^n : |\mathrm{Im}\, x_j| \leq q, \quad j = 1,\dots,n\}, \\ V_\nu &= \{y \in \mathbf{C}^m : \mathrm{Re}\, y \in (\overline{\mathcal{D}} + \nu) \subset \mathbf{R}^m, \quad |\mathrm{Im}\, y_l| \leq \nu, \quad l = 1,\dots,m\}, \end{aligned}$$

where $\overline{\mathcal{D}}$ is a compact domain and $\overline{\mathcal{D}} + \nu$ is the ν-neighborhood of. $\overline{\mathcal{D}}$ in the l_1-norm.

Expand the function $\hat{w}$ into the Fourier series:

$$\hat{w}(z,\varepsilon) = \sum_{k \in \mathbf{Z}^n} \hat{w}^k(y,\varepsilon) e^{i\langle k,x\rangle}.$$

A2. Analyticity.

Let the constant $\alpha > 0$ be such that for any real $s \in [-\alpha, \alpha]$, for any $k \in \mathbf{Z}$ and for any $(x,y) \in \overline{\Sigma}_q \times \overline{V}_\nu =$ closure$(\Sigma_q) \times$ closure(V_ν) the map $g^{is}(x,y)$ is analytic in x, y and the map $\mathbf{g}_k^s \hat{w}^k$ is analytic in y.

Moreover, suppose that for some real ρ

$$|\mathbf{g}_k^s \hat{w}^k|_{\overline{V}_\nu \times [0,\varepsilon_0)} \leq \mu ||k||^{-\rho} e^{-q||k||}, \qquad k \neq 0. \tag{2.5}$$

The function $\hat{w}^0$ is of order ε. So, we assume that

$$|\mathbf{g}_0^s \hat{w}^0|_{\overline{V}_\nu \times [0,\varepsilon_0)} \leq \varepsilon \mu_0. \tag{2.6}$$

3 Main result

Below we denote by c with various sub- and superscripts positive constants independent of ε and μ. The main constants have unique names, but other ones can be denoted in the same way, say, by c. Hope that this will not confuse the reader.

Define the following function $a(\delta)$:

$$
\begin{aligned}
a(\delta) &= c_a \varepsilon \mu P(\varepsilon B + M(\delta)) + c_a \mu B^{-1} e^{-q_0 \delta^{-1/(1+\gamma)}}, && (3.1)\\
P &= c_P \varepsilon^{-1/(\gamma+1)}, \quad c_P \text{ is sufficiently large}, && (3.2)\\
B &= \max\left\{\varepsilon^{-\gamma/(2\gamma+2)}, \mu_0 \mu^{-1}\right\}, && (3.3)\\
M(\delta) &= c \min\{M_1, M_2\}, && (3.4)\\
M_1 &= \begin{cases} 1 + P^{n-\rho} & \text{if } \rho \neq n, \\ \log P & \text{if } \rho = n, \end{cases} && (3.5)\\
M_2 &= \begin{cases} \delta^{-\rho/\gamma} + \delta^{-1} P^{\gamma-\rho} & \text{if } \rho \neq 0,\ \rho \neq \gamma, \\ \log P + \delta^{-1} P^{\gamma} & \text{if } \rho = 0, \\ \delta^{-\rho/\gamma} + \delta^{-1} \log P & \text{if } \rho = \gamma. \end{cases} && (3.6)
\end{aligned}
$$

The constant q_0 is positive.

We put

$$A(\delta) = \int_0^\delta a(s)\, ds, \qquad A_\bullet = A(\alpha/\varepsilon). \tag{3.7}$$

Below it is important to have $A_\bullet$ small. Simple sufficient condition which implies that $A_\bullet \to 0$ as $\varepsilon \to 0$ is as follows:

$$\mu, \mu_0 \sim \text{const}, \quad \rho > \gamma/(\gamma+1). \tag{3.8}$$

Although $A_\bullet$ is expressed in terms of elementary functions of ε, it is rather cumbersome. We just note that

$$A_\bullet \leq c\Big(\varepsilon^{\frac{\gamma}{2(\gamma+1)}} + \varepsilon^{\rho - \frac{\gamma}{\gamma+1}} + \varepsilon^{\frac{\rho}{\gamma+1}} \log \varepsilon\Big) \tag{3.9}$$

provided conditions (3.8) hold and $\rho < \gamma$.

Theorem 1 *Suppose that assumptions* **A1**–**A2** *hold,* $2\nu < Pq$*, and* $A_\bullet < \nu^2/4$*. Then there exists a change of variables*

$$z \mapsto z_\bullet = f(z, \varepsilon), \quad f : \Sigma_{2q/3} \times V_{2\nu/3} \times [0, \varepsilon_0) \to \mathbf{C}^n/(2\pi \mathbf{Z}^n) \times \mathbf{C}^m \tag{3.10}$$

such that

(1) f is analytic in z, smooth in ε,

(2) $f(z, \varepsilon) = z + O(\varepsilon)$,

(3) the system (2.2) *takes the form*

$$\dot z_\bullet = \overline{\omega} + \varepsilon(\overline{w}(y_\bullet) + w_\bullet(z_\bullet, \varepsilon)). \tag{3.11}$$

Let $w^0_\bullet(y_\bullet, \varepsilon)$ be the average in $x_\bullet$ of $w_\bullet$. Then $w^0_\bullet(y_\bullet, 0) = 0$. Moreover,

$$|w_\bullet - w^0_\bullet| \leq C \mu \varepsilon^{\rho/(\gamma+1)} e^{-\overline{q}\varepsilon^{-1/(\gamma+1)}}, \qquad z_\bullet \in \Sigma_{q/2} \times V_{\nu/2}, \tag{3.12}$$

where C is constant, not depending on ε and μ,

$$\overline{q} = (1 + \gamma^{-1})(\gamma \gamma_0 \alpha q^\gamma)^{1/(\gamma+1)}. \tag{3.13}$$

Remark 3.1 *If the system* (2.2) *is Hamiltonian with respect to a certain symplectic structure Ω then the system* (3.11) *is also Hamiltonian with respect to Ω, and the change* (3.10) *is Ω-symplectic.*

Let us put $w_\bullet - w_\bullet^0 = w'(z_\bullet, \varepsilon) + w''(z_\bullet, \varepsilon)$. Both w' and w'' are presented by $\sum w_\bullet^k(y_\bullet, \varepsilon) e^{i\langle k, x_\bullet\rangle}$ where the summation for w' is performed in $k \in \mathbf{Z}^n$ such that

$$q||k|| + |\langle\omega, k\rangle|\alpha\varepsilon^{-1} \le qP, \quad 0 < ||k|| \le P, \tag{3.14}$$

and the sum w'' contains all other nonzero k. In a sense, w' is formed by small denominators with not too big $||k||$. Actually, there are a few of them.

Theorem 2 *Suppose that conditions of Theorem 1 hold. Then there exists a constant $\overline{q}_0 > \overline{q}$ such that for any $z_\bullet \in \Sigma_{q/2} \times V_{\nu/2}$*

$$|w''(z_\bullet, \varepsilon)| < C\mu e^{-\overline{q}_0 \varepsilon^{-1/(\gamma+1)}}. \tag{3.15}$$

For any k, satisfying (3.14), *we put $\sigma = \mathrm{sign}\langle k, \omega\rangle$. Then*

$$\left|w_\bullet^k(y, \varepsilon) - e^{-|\langle k,\omega\rangle|\alpha/\varepsilon} \mathbf{g}_k^{\sigma\alpha} \hat{w}^k(y, \varepsilon)\right| \le c' A_\bullet \mu ||k||^{-\rho} e^{-|\langle k,\omega\rangle|\alpha/\varepsilon - q||k||}. \tag{3.16}$$

Theorem 2 means that the series $w_\bullet - w_\bullet^0$ can be divided into 2 parts: w'' which is small (compare (3.12) with (3.15)), and the main one w' for which we have a sort of control (3.16) (recall that $A_\bullet$ is small: see for example, (3.9)).

Recall that we do not consider the one-frequency ($n = 1$) case in this paper. However, in this case Theorem 1 remains true with $\gamma = 0$ and $\overline{q} = \alpha\omega$. Estimates analogous to (3.12) for $n = 1$ were obtained by Neishtadt [10] and later by Ramis and Schäfke [11]. However, no constructive estimates for $\overline{q}$ were given. Analogous result, containing the equality $\overline{q} = \alpha\omega$, (provided $\overline{u} = \hat{u} = 0$) is presented in [16].

Results analogous to Theorem 1 in multi-frequency case (again without estimates for $\overline{q}$) are contained in [13, 2].

The main motivation for us to study this problem is the problem of exponentially small splitting of asymptotic surfaces in near-integrable Hamiltonian systems. We regard Theorem 2 as a first step to the evaluation of the rate of the splitting in problems not containing any additional small parameter. We used this idea in the one-frequency case (see [14, 15]). Other methods, working good in such problems are given in [9, 5, 7, 8].

The problem of splitting of asymptotic surfaces in multi-frequency case is studied much weaker. All known results deal with problems containing an additional small parameter exponentially [1] or polynomially [3, 6, 12, 4] small with respect to the main one.

4 Continuous averaging

The change f is constructed as a shift $z = z(0) \mapsto z(\delta_\bullet) = z_\bullet$ along trajectories of the phase flow of the ODE system

$$z_\delta = F(z, \varepsilon, \delta), \quad 0 \le \delta \le \delta_\bullet.$$

Here and below the subscript δ denotes the derivative.

In the variables $z = z(\delta)$ the equation (2.2) has the form

$$\dot{z} = \overline{\omega} + \varepsilon(\overline{w}(y) + w(z, \varepsilon, \delta)).$$

We will search for the vector field F in the form $F = \varepsilon\xi w$, where ξ is some linear operator on the set of functions periodic in x. Then it is easy to see that

$$w_\delta = -[\xi w, \overline{\omega} + \varepsilon(\overline{w} + w)], \quad w(z, \varepsilon, 0) = \hat{w}(z, \varepsilon). \tag{4.1}$$

Here $[u_1, u_2] = \partial_{u_1} u_2 - \partial_{u_2} u_1$ is ordinary vector commutator.

The operator ξ is constructed as follows. Expand $w(x, y, \varepsilon, \delta)$ into the Fourier series x:

$$w = \sum_{k \in \mathbf{Z}^n} w^k(y, \varepsilon, \delta) e^{i\langle k, x\rangle}.$$

Consider the sets

$$\begin{aligned} D^{\pm}_{q,P}(\delta) &= \{k\in\mathbf{Z}^n : \pm\langle\omega,k\rangle>0,\ q||k||+|\langle\omega,k\rangle|\delta\le qP\},\\ D^{*}_{q,P}(\delta) &= \mathbf{Z}^n\setminus\Big(\{0\}\cup D^{+}_{q,P}(\delta)\cup D^{-}_{q,P}(\delta)\Big). \end{aligned}$$

Here $P=P(\varepsilon)>0$ is a large fixed constant. Below for briefness instead of $D^{\pm}_{q,P}(\delta)$ we write $D^{\pm}(\delta)$.

We put

$$\sigma_k(\delta)=\begin{cases} 1 & k\in D^+(\delta),\\ -1 & k\in D^-(\delta),\\ 0 & k\notin D^+(\delta)\cup D^-(\delta).\end{cases}$$

Define

$$\xi w = i\sum_{k\in\mathbf{Z}^n}\sigma_k(\delta)w^k e^{i\langle k,x\rangle}.$$

If $P=+\infty$ and $n=1$, ξw is called the Hilbert transform of w.

Define the following functions

$$w^{\pm}=\sum_{k\in D^{\pm}(\delta)} w^k e^{i\langle k,x\rangle},\quad w^*=\sum_{k\in D^*(\delta)} w^k e^{i\langle k,x\rangle}.$$

Then $w=w^0+w^++w^-+w^*$ and the equality $\xi w=i(w^+-w^-)$ holds.

The equation (4.1) takes the form

$$w_\delta=-i[w^+-w^-,\overline{\omega}+\varepsilon(\overline{w}+w^0+w^++w^-+w^*)],\quad w(z,\varepsilon,0)=\hat{w}(z,\varepsilon). \quad (4.2)$$

Proof of Remark 3.1.

If the vector field w is Hamiltonian with Hamiltonian H with respect to the symplectic structure Ω, so is the vector field ξw. The corresponding Hamiltonian equals $i(H^+-H^-)$, where the functions $H^{\pm}$ are defined in the same way as $w^{\pm}$. Here we have used the linearity of the operator which assigns to a Hamiltonian the corresponding vector field.

In this case the equation (4.2) can be rewritten in the following Hamiltonian form:

$$H_\delta=-i\{H^+-H^-,H^{\overline{\omega}}+\varepsilon(H^{\overline{w}}+H^0+H^++H^-+H^*)\},\quad H(z,\varepsilon,0)=\hat{H}(z,\varepsilon), \quad (4.3)$$

where $\{\cdot,\cdot\}$ is the Poisson bracket, $H^{\overline{\omega}}$, $H^{\overline{w}}$, etc are Hamiltonians generating the vector fields $\overline{\omega}$, $\overline{w}$, etc.

Hence, if (4.2) has a unique solution, so is (4.3). Moreover, H is the Hamiltonian for w, and ξH is the Hamiltonian for $\xi w=F$.

5 Averaging equations

The system (4.2) can be rewritten in the form of the following infinite system of partial differential equations for the Fourier coefficients w^k:

$$\begin{aligned} w^k_\delta &= -\sigma_k\langle k,\omega\rangle w^k+i\varepsilon\sigma_k e^{-i\langle k,x\rangle}[\overline{w},w^k e^{i\langle k,x\rangle}]+e^{-i\langle k,x\rangle}\eta^k,\\ w^k|_{\delta=0} &= \hat{w}^k, \end{aligned} \quad (5.1)$$

where the functions $\eta^k=\eta^k(z,\varepsilon,\delta)$ are defined as follows:

$$\begin{aligned} \eta^k = {} & i\varepsilon\sigma_k(\delta)[w^0,w^k e^{i\langle k,x\rangle}]\\ & -2i\varepsilon\sum_{l^\pm\in D^\pm(\delta),l^++l^-=k}\Big[w^{l^+}e^{i\langle l^+,x\rangle},w^{l^-}e^{i\langle l^-,x\rangle}\Big]\\ & -i\varepsilon\sum_{\sigma\in\{+,-\}}\sigma\sum_{l\in D^\sigma(\delta),l^*\in D^*(\delta),l+l^*=k}\Big[w^l e^{i\langle l,x\rangle},w^{l^*}e^{i\langle l^*,x\rangle}\Big]. \end{aligned} \quad (5.2)$$

We complete the section by an informal studying of the system (5.1). Putting $P=+\infty$ and $\varepsilon=0$, we have:

$$w^k_\delta=-|\langle k,\omega\rangle|w^k,\quad w^k|_{\delta=0}=\hat{w}^k. \quad (5.3)$$

Solution of this system has the form $w^k(y,\varepsilon,\delta) = e^{-|\langle k,\omega\rangle|\delta}\hat{w}^k(y,\varepsilon)$. Since the frequency vector ω is nonresonant, the Fourier coefficients w^k with $k \neq 0$ tend to zero as $\delta \to +\infty$. Note that the coefficients corresponding to small denominators, ($\langle k,\omega\rangle \approx 0$), decrease slowly. Unfortunately, the system (5.1) with $P = +\infty$, apparently, has no solution. On the other hand, coefficients $\hat{w}^k_\delta$ with large $||k||$ are small and do not need to be averaged. Below we take P satisfying (3.2).

For finite P instead of (5.3) we have:

$$w^k_\delta = -\sigma_k(\delta)\langle k,\omega\rangle w^k, \quad w^k|_{\delta=0} = \hat{w}^k. \tag{5.4}$$

As a solution we get: $w^k(y,\varepsilon,\delta) = e^{-\langle k,\omega\rangle S_k(\delta)}\hat{w}^k(y,\varepsilon)$, where

$$S_k(\delta) = \int_0^\delta \sigma_k(s)\,ds.$$

When $\delta \to \infty$ we have:

$$w^k(y,\delta) = \begin{cases} e^{-qP+q||k||}\hat{w}^k & \text{if } k \in D^+(0) \cup D^-(0), \\ \hat{w}^k & \text{if } k \notin D^+(0) \cup D^-(0). \end{cases}$$

Since $\hat{w}^k$ is of the order $||k||^{-\rho}e^{-q||k||}$, we see that $w - w^0$ is roughly of the order $P^{n-\rho}e^{-qP}$. The last quantity is small.

Now consider a more precise approximation for (5.1). Putting formally $\eta^k = 0$, we get:

$$w^k_\delta = -\sigma_k\langle k,\omega\rangle w^k + i\varepsilon\sigma_k e^{-i\langle k,x\rangle}[\overline{w}, w^k e^{i\langle k,x\rangle}], \quad w^k|_{\delta=0} = \hat{w}^k.$$

The solution again can be written down explicitly:

$$w^k(y,\varepsilon,\delta) = e^{-\langle k,\omega\rangle S_k(\delta)}\mathbf{g}_k^{\varepsilon S_k(\delta)}\hat{w}^k(y,\varepsilon).$$

We see that singularities of the functions $\mathbf{g}_k^{\varepsilon S_k(\delta)}\hat{w}^k(y,\varepsilon)$ hinder to take δ arbitrarily large (recall that $\mathbf{g}_k^s$ include shifts in imaginary direction along the phase flow g^t). However, according to our analyticity assumptions (see (2.5)–(2.6)), δ can be taken of order ε^{-1}. For such δ the multipliers $e^{-\langle k,\omega\rangle S_k(\delta)}$ provide a strong averaging.

6 Majorants

The main tool we use for the analysis of system (5.1) is the majorant method. Below we denote by U^m a neighborhood of zero in $\mathbf{C}^m$, and by U^1 a neighborhood of zero in $\mathbf{C}$.

Let $f = \sum f_j y^j$ and $g = \sum g_j y^j$ be two complex-valued functions on U^m:

$$j = (j_1,\ldots,j_m) \in \mathbf{Z}^m, \quad y^j = y_1^{j_1}\cdots y_m^{j_m}, \quad j_s \geq 0.$$

Then we say that $f \ll g$ if for all multi-indices j the inequalities $|f_j| \leq g_j$ hold. In this case g is called a majorant for f. If f takes values in C^l, $l \in \mathbf{N}$, we say that $f \ll g$ if g majorates each component of the function f.

The following properties of majorants are obvious.

(1) If $f_1 \ll F_1$, $f_2 \ll F_2$, and $c_1, c_2 \in \mathbf{C}$ then

$$c_1 f_1 + c_2 f_2 \ll |c_1|F_1 + |c_2|F_2, \quad f_1 f_2 \ll F_1 F_2.$$

(2) If $f \ll F$ and ∂ is any differential operator in y with constant positive coefficients then $\partial f \ll \partial F$.

(3) If f and F depend on some real parameter s and for any $s \in [a,b]$ we have $f(y,s) \ll F(y,s)$ then $\int_a^b f(y,s)\,ds \ll \int_a^b F(y,s)\,ds$.

Since $\overline{\mathcal{D}}$ is compact, it is sufficient to perform an analysis of the averaging system in a neighborhood of an arbitrary point $y_0 \in \overline{\mathcal{D}}$. For convenience we assume that y_0 has zero coordinates. Below we always denote

$$Y = y_1 + \ldots y_m.$$

For $f : U^m \to \mathbf{C}^l$ and $F : U^1 \to \mathbf{C}$ we say that $f(y) \ll F(Y)$ if $F(y_1 + \ldots + y_m)$ majorates each component of f.

Proposition 6.1 *For any $k \in \mathbf{Z}^n$, $\varepsilon \in [0, \varepsilon)$, and any $s \in [-\alpha, \alpha]$*

$$\begin{aligned} \mathbf{g}_0^s \hat{w}^0(y, \varepsilon) &\ll \frac{\varepsilon \mu_0}{\nu - Y}, \\ \mathbf{g}_k^s \hat{w}^k(y, \varepsilon) &\ll \frac{\mu e^{-q||k||}}{||k||^\rho (\nu - Y)}, \qquad k \neq 0. \end{aligned}$$

Proof of Proposition 6.1 easily follows from assumption **A2** and the Cauchy integral formula.

7 The majorant commutator

Definition 7.1 *For any two functions $F, G : U^1 \to \mathbf{C}$, and any two vectors $l, k \in \mathbf{Z}^n$ we define the majorant commutator:*

$$[[F, G]]^{l,k} = (||l||_1 + ||k||_1) FG + m(FG)_Y. \tag{7.1}$$

Here $||\cdot||_1$ is the following norm on $\mathbf{R}^n$: $||j||_1 = |j_1| + \ldots + |j_n|$.

The following two propositions can be regarded as a motivation for this definition.

Proposition 7.1 *Suppose that $f_1, f_2 : U^m \to \mathbf{C}^{n+m}$, $F_1, F_2 : U^1 \to \mathbf{C}$, and $f_1 \ll F_1$, $f_2 \ll F_2$.*

Then for any $k_1, k_2 \in \mathbf{Z}^n$

$$e^{-i\langle k_1 + k_2, x\rangle} \Big[f_1 e^{i\langle k_1, x\rangle}, f_2 e^{i\langle k_2, x\rangle} \Big] \ll [[F_1, F_2]]^{k_1, k_2}. \tag{7.2}$$

The proof is straightforward.

Proposition 7.2 *Suppose that $\mathbf{g}_{k_1}^h f_1(y) \ll F_1(Y)$, $\mathbf{g}_{k_1}^h f_2(y) \ll F_2(Y)$. Then*

$$\mathbf{g}_{k_1 + k_2}^h \Big(e^{-i\langle k_1 + k_2, x\rangle} \big[f_1 e^{i\langle k_1, x\rangle}, f_2 e^{i\langle k_2, x\rangle} \big] \Big) \ll [[F_1, F_2]]^{k_1, k_2}. \tag{7.3}$$

Proof of Proposition 7.2. The left-hand side of (7.3) equals

$$e^{-i\langle k_1 + k_2, x\rangle} \big[e^{i\langle k_1, x\rangle} \mathbf{g}_{k_1}^h f_1, e^{i\langle k_2, x\rangle} \mathbf{g}_{k_2}^h f_2 \big].$$

Hence, (7.3) follows from (7.2).

8 The functions ψ^β.

For briefness we do not write ε among arguments of functions.

Consider the functions ψ, ψ^β, $\beta \geq 0$ defined as follows

$$\psi_\delta = a(\delta)\psi\psi_y, \quad \psi|_{\delta=0} = 1/(\nu - Y), \tag{8.1}$$

$$\psi^\beta = \psi \qquad \text{if } \beta \leq P, \tag{8.2}$$

$$\psi_\delta^\beta = a(\delta)\Big(\psi\psi_Y^\beta + \frac{\beta}{P}\psi\psi^\beta\Big), \quad \psi^\beta|_{\delta=0} = \frac{1}{\nu - Y}, \quad \text{if } \beta > P. \tag{8.3}$$

Proposition 8.1

$$\psi = \frac{2}{(\nu - Y) + \sqrt{(\nu - Y)^2 - 4A(\delta)}}, \quad \psi^\beta = \psi e^{A(\delta)\beta\psi/P}, \quad \beta > P.$$

The straightforward proof is based on (8.1)–(8.3).

The functions ψ^β can be regarded as majorants (i.e., $\psi^\beta \gg 0$) provided $A(\delta) < \nu^2/4$.

9 Main lemma

To analyze solution of the averaging system, we use local argument: y takes values near some value y_0. Since $y_0 \in D$ can be chosen arbitrarily, this way we obtain the required estimates for the function w. Without loss of generality we can assume that $y_0 = 0$.

Lemma 9.1 *For any real s and $\delta \geq 0$ such that $|s| + \varepsilon\delta \leq \alpha$, and for any $k \in \mathbf{Z}^n$*

$$\mathbf{g}_k^s w^k(y,\delta) \ll ||k||^{-\rho} e^{-\langle k,\omega\rangle S_k(\delta) - q||k||} W^k(Y,\delta), \quad k \neq 0, \tag{9.1}$$

$$\mathbf{g}_0^s w^0(y,\delta) \ll W^0(Y,\delta) \tag{9.2}$$

where

$$W^0(Y,\delta) = \varepsilon\mu B\psi, \quad W^k(Y,\delta) = \mu\psi^{||k||}.$$

10 Proof of the main lemma

Our method of the proof can be illustrated by the following trivial example. Consider the Cauchy problem for real scalar function $u(t)$:

$$\dot{u} = u^2 - 1, \qquad u(0) = 2. \tag{10.1}$$

It can be easily solved explicitly. However, suppose that we want to prove the following

Proposition 10.1 *For any $t \in [0, 1/9)$*

$$|u(t)| < \mathbf{u}(t) \equiv 3 + 3t. \tag{10.2}$$

Proof. The Cauchy problem (10.1) is equivalent to the following integral equation: $u(t) = 2 + \int_0^t (u^2(s) - 1)\, ds$. Hence, for $t \geq 0$ we have:

$$|u(t)| \geq 2 + \int_0^t (u^2(s) + 1)\, ds.$$

The inequality (10.2) holds for $t = 0$. Suppose that it holds for $t \in [0, t_0)$ and let us check that

$$|\mathbf{u}(t_0)| \geq 2 + \int_0^{t_0} (\mathbf{u}^2(s) + 1)\, ds. \tag{10.3}$$

Straightforward estimates show that (10.3) is satisfied provided $t_0 < 1/9$. This proves the proposition. Indeed, consider the set

$$\mathbf{S} = \{t_* \geq 0 : \text{ (10.2) holds for any } t \in [0, t_*]\}.$$

This set is connected and open in $\mathbf{R}^+ = \{t_* \in \mathbf{R} : t_* \geq 0\}$. Our check shows that $\mathbf{S}$ is closed in $[0, 1/9]$. Therefore, $\mathbf{S}$ contains the interval $[0, 1/9]$.

To apply the proposed method to our problem, we just need to replace in estimates the order relation $<$ by the majorant order $\ll$. Here we should confess that we do not have a complete rigorous justification of this method for the analysis of the averaging equations, although, the procedure looks natural and trivial. Once again: first, we replace the Cauchy problem by an integral equation, secondly, we replace the integral equation by its majorant version, thirdly we present a conjecture about the behaviour of the solution of the majorant equation, and finally we "check" the conjecture by the direct substitution into the equation.

Let us define

$$L(l', l'', s) = -\langle l', \omega\rangle - \langle l'', \omega\rangle + \langle l' + l'', \omega\rangle - q(||l'|| + ||l''|| - ||l' + l''||).$$

We will use the integral version of system (5.1).

$$e^{\langle k,\omega\rangle S_k(\delta)} w^k(y,\delta) = \mathbf{g}_k^{\varepsilon S_k(\delta)} \hat{w}^k(y) + ||k||^{-\rho} e^{-q||k||} \int_0^\delta \zeta_k(y,s)\, ds, \tag{10.4}$$

$$\zeta_k(y,s) = ||k||^\rho e^{\langle k,\omega\rangle S_k(s) + q||k||} \mathbf{g}_k^{\varepsilon(S_k(\delta) - S_k(s))} \big(e^{-i\langle k,x\rangle} \eta^k(z,s)\big). \tag{10.5}$$

Now we prove the main lemma. The plan of the proof is as follows. According to Proposition 6.1, Lemma 9.1 holds for $\delta = 0$. Assuming that Lemma 9.1 holds for $\delta \in [0, \delta_0)$ we check that the majorant for the left-hand side of (10.4) taken from (9.1),(9.2) is greater than the corresponding majorant of the right-hand side of (10.4) for $\delta = \delta_0$.

Proposition 10.2 *Suppose that Lemma 9.1 holds. Then*

$$\zeta^k(y,s) \ll \mu\psi_s^{||k||}(Y,s), \quad \zeta^0(y,s) \ll \varepsilon\mu B\psi_s(Y,s)$$

Proof. By using (10.5), (5.2), and Proposition 7.2, we get

$$\begin{aligned} \zeta^k &\ll \varepsilon|\sigma_k(s)|\,[[\varepsilon\mu B\psi(Y,s), \mu\psi(Y,s)]]^{0,k} \\ &+ 2\varepsilon \sum_{l^\pm \in D^\pm(\delta), l^+ + l^- = k} \Big(\frac{||k||}{||l^+||\,||l^-||}\Big)^\rho e^{L(l^+,l^-,s)}[[\mu\psi(Y,s), \mu\psi(Y,s)]]^{l^+,l^-} \\ &+ \varepsilon \sum_{l\in D^\pm(\delta), l^* \in D^*(\delta), l+l^*=k} \Big(\frac{||k||}{||l||\,||l^*||}\Big)^\rho e^{L(l,l^*,s)}[[\mu\psi(Y,s), \mu\psi^{||l^+||}(Y,s)]]^{l,l^*}. \end{aligned}$$

By using (7.1), we get:

$$\begin{aligned} \zeta^k \ll c\varepsilon^2\mu^2 B(P\psi^2 + (\psi^2)_Y) \quad &+ \quad c\varepsilon\mu^2(2||k||\psi^2 + (\psi^2)_Y)\Sigma_1 \\ &+ \quad c\varepsilon\mu^2(2||k||\psi\psi^{||k||+P} + 2\psi\psi_Y^{||k||+P})\Sigma_2, \end{aligned}$$

where

$$\Sigma_1 = \sum_{l^\pm \in D^\pm(\delta), l^+ + l^- = k} \Big(\frac{||k||}{||l^+||\,||l^-||}\Big)^\rho e^{L(l^+,l^-,\delta)}, \tag{10.6}$$

$$\Sigma_2 = \sum_{l\in D^\pm(\delta), l^*\in D^*(\delta) l+l^*=k} \Big(\frac{||k||}{||l||\,||l^*||}\Big)^\rho e^{L(l,l^*,\delta)}. \tag{10.7}$$

We proceed with the help of Propositions A.1 and B.2:

$$\begin{aligned} \zeta^k(Y,s) &\ll 2\big(c\varepsilon^2\mu^2 B + 2c\varepsilon\mu^2 M(s)\big)(2||k||\psi\psi^{||k||+P} + 2\psi\psi_Y^{||k||+P}) \\ &\ll 4c\varepsilon P\mu^2 e^q\big(\varepsilon B + 2M(s)\big)(\psi\psi_Y^{||k||} + ||k||P^{-1}\psi\psi^{||k||}) \\ &\ll 8c\varepsilon P\mu^2 e^q\big(\varepsilon B + 2M(s)\big)a^{-1}(s)\psi_s^{||k||}(Y,s) \\ &\ll \mu\psi_s^{||k||}(Y,s). \end{aligned}$$

For $k = 0$ we get:

$$\begin{aligned} \zeta_0(y,s) &\ll c\varepsilon\mu^2 e^{-q_0 s^{-1/(1+\gamma)}}\psi(Y,s)\psi_Y(Y,s) \\ &\ll c\varepsilon\mu^2 e^{-q_0 s^{-1/(1+\gamma)}}a^{-1}(s)\psi_s(Y,s) \\ &\ll \varepsilon\mu B\psi_s(Y,s). \end{aligned}$$

We should check that

$$\begin{aligned} \varepsilon\mu B\psi(Y,\delta) &\gg \frac{\varepsilon\mu_0}{\nu - Y} + \int_0^\delta \varepsilon\mu B\psi_s(Y,s)\,ds, \\ \mu\psi^{||k||}(Y,\delta) &\gg \frac{\mu}{\nu - Y} + \int_0^\delta \mu\psi_s^{||k||}(Y,s)\,ds \end{aligned}$$

These relations are obvious, because according to the definition of B, we have: $\mu B \geq \mu_0$.

11 An estimate for $w_\bullet$

The vector field $w_\bullet$ can be presented as a Fourier series. We break this series into 4 parts:

$$w_\bullet = \sum_{k\in\mathbf{Z}} w_\bullet^k e^{i\langle k,x\rangle} = w_\bullet^0 + \Sigma_1^\bullet + \Sigma_2^\bullet + \Sigma_3^\bullet,$$

where the sum $\Sigma_1^\bullet$ is taken for $||k|| > P$,
the sum $\Sigma_2^\bullet$ is taken for $q||k|| + |\langle\omega, k\rangle|\alpha\varepsilon^{-1} \geq qP$, $||k|| \leq P$,
the sum $\Sigma_3^\bullet$ is taken for $q||k|| + |\langle\omega, k\rangle|\alpha\varepsilon^{-1} \leq qP$, $0 < ||k|| \leq P$.
Let us estimate these sums separately:

$$\begin{aligned}
\Sigma_1^\bullet &\ll \sum \frac{\mu\psi^{||k||}}{||k||^\rho} e^{-q||k||} \ll \sum \frac{\mu\psi}{||k||^\rho} e^{\nu(||k||-P)/P - q||k||} \\
&\ll \sum \frac{\mu\psi}{||k||^\rho} e^{(\nu/P - q)(||k||-P) - qP}
\end{aligned}$$

Recalling that $q > \frac{\nu}{P}$, we have

$$\Sigma_1^\bullet \ll c\mu P^{n-1+|\rho|} e^{-qP}$$

For the other two sums we have:

$$\begin{aligned}
\Sigma_2^\bullet &\ll \sum \frac{\mu\psi}{||k||^\rho} e^{-q||k|| - |\langle\omega,k\rangle|\Delta(k,\alpha/\varepsilon)} = \sum \frac{\mu\psi}{||k||^\rho} e^{-qP} \ll c\mu P^{n+|\rho|} e^{-qP}, \\
\Sigma_3^\bullet &\ll \sum_{k\in\mathbf{Z}^n} \frac{\mu\psi}{||k||^\rho} e^{-q||k|| - |\langle\omega,k\rangle|\alpha/\varepsilon} \ll \sum_{k\in\mathbf{Z}^n} \frac{\mu\psi}{||k||^\rho} e^{-q||k|| - \gamma_0 ||k||^{-\gamma}\alpha/\varepsilon}.
\end{aligned}$$

By straightforward calculations we get: $\Sigma_3^\bullet \ll c\mu\varepsilon^{\frac{\rho}{\gamma+1}} e^{-\overline{q}\varepsilon^{-1/(\gamma+1)}}$. Applying (3.2), we have $\Sigma_3^\bullet \ll c\mu c_P^{|\rho|} P^{|\rho|} e^{-\overline{q}\varepsilon^{-1/(\gamma+1)}}$.
Hence,

$$w_\bullet - w_\bullet^0 \ll c\mu P^{|\rho|}\left(c_P^{|\rho|} + P^{n-1} + P^n\right) e^{-\overline{q}\varepsilon^{-1/(\gamma+1)}}.$$

12 An estimate for $w_\bullet^k$

The estimate (3.16) of Theorem 2 is derived from the following lemma.

Lemma 12.1 *For any $k \in \mathbf{Z}^n$, $0 < ||k|| \leq P$, we put $\sigma = \operatorname{sign}\langle k, \omega\rangle$. Then*

$$w_\bullet^k - e^{-\langle k,\omega\rangle S_k(\alpha/\varepsilon)} \mathbf{g}_k^{\sigma\alpha} \hat{w}^k \ll ||k||^{-\rho} e^{-\langle k,\omega\rangle S_k(\alpha/\varepsilon) - q||k||} \frac{A_\bullet}{(\nu - Y)^2 - 4A_\bullet}.$$

Proof. According to (10.4),

$$w_\bullet^k - e^{-\langle k,\omega\rangle S_k(\alpha/\varepsilon)} \mathbf{g}_k^{\sigma\alpha} \hat{w}^k = ||k||^{-\rho} e^{-\langle k,\omega\rangle S_k(\alpha/\varepsilon) - q||k||} \int_0^{\alpha/\varepsilon} \zeta_k(y, s)\, ds.$$

Due to Proposition 10.2,

$$\int_0^{\alpha/\varepsilon} \zeta_k(y, s)\, ds \ll \mu(\psi(Y, \alpha/\varepsilon) - \psi(Y, 0)).$$

It remains to use assertion (f) of Proposition A.1.

A Some properties of the functions $\psi, \psi^{||k||}$

Proposition A.1 *Suppose that $\delta \geq 0$. Then*
(a) $1/(\nu - Y) \ll \psi \ll \psi_Y$,
(b) $\psi \ll \psi^\nu$ *for any* $\nu \geq 0$,
(c) $\psi_Y \psi^\nu \ll \psi\psi_Y^\nu$,
(d) $\psi^\beta \ll \psi e^{\nu\beta/P}$,
(e) $\psi^{\beta''} \ll \psi^{\beta'} e^{\nu(\beta''-\beta')/P}$, $\beta' < \beta''$.
(f) $\psi(Y, \delta) - \psi(Y, 0) \ll \frac{A(\delta)}{(\nu - Y)^2 - 4A(\delta)}$.

Proof. Assertions (a)–(c) are obvious. Let us check (d). First, let us show that

$$\psi^2 \ll \nu\psi/A. \tag{A.1}$$

Indeed,

$$\begin{aligned}\psi^2 &= \frac{4}{(\nu - Y + \sqrt{(\nu-Y)^2 - 4A})^2} = \frac{\nu - Y - \sqrt{(\nu-Y)^2-4A}}{A(\nu - Y + \sqrt{(\nu-Y)^2-4A})} \\ &= \frac{2\nu - 2Y - (\nu - Y + \sqrt{(\nu-Y)^2-4A})}{A(\nu - Y + \sqrt{(\nu-Y)^2-4A})} \ll \frac{\nu}{A}\psi.\end{aligned}$$

By using (A.1), we have:

$$\begin{aligned}\psi^\beta &= \psi e^{A\psi\beta/P} = \psi + \frac{A\beta}{P1!}\psi^2 + \frac{A^2\beta^2}{P^2 2!}\psi^3 + \dots \\ &\ll \psi\Big(1 + \frac{\nu\beta}{P1!} + \frac{\nu^2\beta^2}{P^2 2!} + \dots\Big) = \psi e^{\nu\beta/P}.\end{aligned}$$

So, we get the required estimate.

Assertion (e) can be proven analogously.

Let us prove (f). We have:

$$\psi(Y,\delta) - \psi(Y,0) = \frac{4A}{\big(\nu - Y + \sqrt{(\nu-Y)^2-4A}\big)^2}.$$

So, we just should check that

$$\frac{2}{\nu - Y + \sqrt{(\nu-Y)^2-4A}} \ll \frac{1}{\sqrt{(\nu-Y)^2-4A}}.$$

Subtracting from the right-hand side of this inequality the left-hand side, we get:

$$\frac{4A}{\sqrt{(\nu-Y)^2-4A}\,\big(\nu - Y + \sqrt{(\nu-Y)^2-4A}\big)^2}.$$

This expression is obviously greater ($\gg$) than zero.

B Estimates for sums

Proposition B.1 *For some positive constants c_0 and q_0*

$$\Sigma_0 = \sum_{l\in\mathbf{Z}^n} e^{-|\langle l,\omega\rangle|\delta - q||l||} \le c_0 e^{-q_0\delta^{1/(\gamma+1)}}.$$

Proof. According to the Diophantine conditions

$$|\langle l,\omega\rangle|\delta \ge \gamma_0\delta||l||^{-\gamma}$$

Therefore, the sum Σ_0 can be estimated by the integral

$$\int_{\mathbf{R}^n} e^{-\gamma_0\delta||l||^{-\gamma} - q||l||}dl,$$

and we arrive at the required inequality

Proposition B.2 *Let $M(\delta)$ be defined by (3.4). Then the sums (10.6)–(10.7) satisfy the following estimates: $\Sigma_1 \le M(\delta)$, $\Sigma_2 \le M(\delta)$.*

Proof. We prove Proposition B.2 in 2 steps. First, we show that Σ_1 and Σ_2 do not exceed $\mathcal{S}' + \mathcal{S}''$, where

$$\mathcal{S}' = 4^\rho \sum_{\Lambda(\delta)\le||l||\le P} ||l||^{-\rho} e^{-|\langle l,\omega\rangle|\delta}, \quad \mathcal{S}'' = 4^\rho \sum_{\Lambda(\delta)\le||k-l||\le P} ||k-l||^{-\rho} e^{-|\langle l,\omega\rangle|\delta}.$$

And secondly, we estimate $\mathcal{S}'$ and $\mathcal{S}''$.

1. Consider the sum Σ_1 for positive $\langle k,\omega\rangle$ (the case of negative $\langle k,\omega\rangle$ is analogous). We have $L(l^+,l^-,\delta) \le -2|\langle l^-,\omega\rangle|\delta$. On the other hand,

$$\frac{||k||}{||l^+||\,||l^-||} \le \frac{2}{\min\{||l^+||,||l^-||\}} \le \frac{2}{||l^-||} + \frac{2}{||l^+||}.$$

Hence, Σ_1 does not exceed $\mathcal{S}' + \mathcal{S}''$.

In the sum Σ_2 first, assume that $||k|| > P$. Then $S_k(\delta) = 0$, and therefore,

$$e^{L(l^+,l^-,\delta)} \le e^{-\langle l^*,\omega\rangle S(\delta) - |\langle l,\omega\rangle|\delta}. \tag{B.1}$$

Let us show that

$$\Big(\frac{||k||}{||l^*||}\Big)^{\rho} e^{-\langle l^*,\omega\rangle S(\delta)} \le 4^{\rho}. \tag{B.2}$$

Indeed, if $||l^*|| \ge P/2$, we have:

$$\frac{||k||}{||l^*||} \le \frac{||l|| + ||l^*||}{||l^*||} \le \frac{P + ||l^*||}{||l^*||} \le 3.$$

If $||l^*|| < P/2$, we use the inclusion $l \in D^*(\delta) \cap D^{\pm}(0)$. This means that $\langle l^*,\omega\rangle S(\delta) + q||l^*|| = qP$. Hence, $\langle l^*,\omega\rangle S(\delta) \ge qP/2$. On the other hand, $||k|| \le ||l|| + ||l^*|| \le 3P/2$. Since P is large, we again obtain (B.2).

By using (B.2) and (B.1), we get: $\Sigma_2 \le \mathcal{S}'$.

If $||k|| \le P$, we have:

$$\langle k,\omega\rangle S_k(\delta) + \sigma||k|| \le \sigma P \le \langle l^*,\omega\rangle S_{l^*}(\delta) + \sigma||l^*||.$$

In this case $\Sigma_2 \le \sum_{||l||\le P} e^{-|\langle l,\omega\rangle|\delta - q||l||}$. So, $\Sigma_2 \le M(\delta)$ by Proposition B.1.

2. Hence, we need to estimate $\mathcal{S}'$ and $\mathcal{S}''$. For any $\delta \ge 0$ we have: $\mathcal{S}' \le \sum_{||l||\le P} ||l||^{-\rho} \le c(1 + P^{n-\rho})$ if $n \ne \rho$, or $\mathcal{S}' \le c\log P$ if $n = \rho$. The same estimate holds for $\mathcal{S}''$.

Now suppose that $\delta > 0$. We have:

$$\begin{aligned}
\mathcal{S}' &\le \sum_{j\ge 0, 1\le e^r \le c_1 P} \mathcal{S}_{j,r}, \qquad j,r,\in \mathbf{Z},\\
\mathcal{S}'_{j,r} &= \sum_{c_2 e^r \le ||l|| \le c_2 e^{r+1}, j \le |\langle l,\omega\rangle|\delta \le j+1} ||l||^{-\rho} e^{-j}\\
&\le \sum_{||l|| \le c_2 e^{r+1}, j \le |\langle l,\omega\rangle|\delta \le j+1} c_2^{-\rho} e^{-r\rho - j}.
\end{aligned}$$

According to the Diophantine conditions, in the sum $\mathcal{S}'_{j,r}$ we have

$$\frac{\gamma_0}{c_2^{\gamma} e^{\gamma(r+1)}} \le \frac{\gamma_0}{||l||^{\gamma}} \le |\langle l,\omega\rangle| \le \frac{j+1}{\delta}.$$

Hence,

$$e^r \ge R = c_R\big(\delta/(j+1)\big)^{1/\gamma}.$$

According to Proposition C.1, the number of terms in the sum $\mathcal{S}'_{j,r}$ does not exceed $2 + 2(\delta\gamma_0)^{-1}(2c_2 e^{r+1})^{\gamma}$. Hence,

$$\mathcal{S}' \le \sum_{j\ge 0, R\le e^r \le c_1^{-1} P} c_3(1 + \delta^{-1} e^{r\gamma}) e^{-r\rho - j}, \qquad j,r \in \mathbf{Z}.$$

First, assume that $\rho > 0$ and $\rho \ne \gamma$. Making summation with respect to r, we get:

$$\mathcal{S}' \le \sum_{j\ge 0} c_4\Big(\Big(\frac{j+1}{\delta}\Big)^{\rho/\gamma} + \frac{(j+1)^{\rho/\gamma - 1}}{\delta^{\rho/\gamma}} + \frac{P^{\gamma-\rho}}{\delta}\Big) e^{-j}.$$

Summing up with respect to j, we obtain the estimate

$$\mathcal{S}' \le c(\delta^{-\rho/\gamma} + \delta^{-1} P^{\gamma-\rho})$$

Analogously we get:

$$\begin{aligned}\mathcal{S}' &\leq c(\log P + \delta^{-1} P^{\gamma}) \quad \text{if } \rho = 0,\\ \mathcal{S}' &\leq (\delta^{-\rho/\gamma} + \delta^{-1}\log P) \quad \text{if } \rho = \gamma\end{aligned}$$

The sum $\mathcal{S}''$ can be estimated analogously. In this case it is convenient to put

$$\mathcal{S}''_{j,r} = \sum_{ce^r \leq ||k-l|| \leq ce^{r+1}, j \leq |\langle l,\omega\rangle|\delta \leq j+1} ||k-l||^{-\rho} e^{-j}.$$

C Auxiliary statements

Proposition C.1 *Let $\mathcal{N}$ be the number of elements in the set $\mathcal{A}(a,b) \cap B(l_0, R)$, where*

$$\mathcal{A}(a,b) = \{l \in \mathbf{Z}^n : a \leq |\langle l,\omega\rangle| \leq b\}, \quad B(l_0,R) = \{l \in \mathbf{Z}^n : ||l - l_0|| \leq R\}.$$

Then $\mathcal{N} \leq 2 + 2(b-a)\gamma_0^{-1}(2R)^{\gamma}$.

Proof. We have the equality:

$$\mathcal{A} = \mathcal{A}^+ \cup \mathcal{A}^-, \quad \mathcal{A}^{\pm} = \{l \in \mathcal{A} : \pm\langle l,\omega\rangle \geq 0\}.$$

Consider the set $\mathcal{A}^+$ (analogous argument obviously hold for $\mathcal{A}^-$). Let $\mathcal{N}^+$ be the number of its elements. There exist two vectors $l', l'' \in \mathcal{A}^+ \cap B(l_0, R)$ such that

$$|\langle l',\omega\rangle - \langle l'',\omega\rangle| \leq (b-a)/(\mathcal{N}-1).$$

We put $l = l' - l''$. Since $||l|| \leq 2R$, we have

$$\gamma_0(2R)^{-\gamma} \leq |\langle l,\omega\rangle| \leq (b-a)/(\mathcal{N}-1).$$

This implies the required estimate.

Acknowledgement. The work was partially supported by Russian Foundation of Basic Research grants No 96-15-96855 and 99-01-00953, and by INTAS grant 97-10771.

References

[1] Arnold V., Instability of dynamical systems with several degrees of freedom, Dokl. Akad. Nauk SSSR, 1964, V. 156, No. 1, P. 9–12, English transl. in *Sov. Mat. Dokl.*, 1964, V. 5, P. 581–585.

[2] Bambusi D., Long time stability of some small amplitude solutions in nonlinear Schrödinger equations. Comm. Math. Phys. 1997, V. 189, 205–226.

[3] Delshams A., Gelfreich V., Jorba A., Seara T., Exponentially small splitting of separatrices under fast quasi-periodic forcing, Comm. Math. Phys. **189**, 1997, 35–71.

[4] Delshams A., Gutiérres D., Splitting potential and Poincaré-Melnikov theory for whiskered tori in Hamiltonian systems, Preprint, 1998, http://www-ma1.upc.es

[5] Delshams A., Seara T., An asymptotic expression for the splitting of separatrices of the rapidly forced pendulum, Comm. Math. Phys. **150**, 1992, 433–463.

[6] Gallavotti G., Twistless KAM tori, quasi-flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable systems; a review, Reviews in Mathematical Physics, **6**, 1994, 343–411.

[7] Gelfreich V., Reference systems for splitting of separatrices. *Nonlinearity* V. 10 (1997) P. 175–193.

[8] V.Gelfreich., A proof of the exponentially small transversality of the separatrices for the standard map, Comm.Math.Phys. vol. 201, 155-216, 1999.

[9] Lazutkin V., Splitting of separatrices for standard Chirikov's mapping, VINITI no. 6372-84, 24 Sept. 1984 (in Russian).

[10] Neishtadt A., The separation of motions in systems with rapidly rotating phase, *Prikl. Mat. Mekh.*, 1984, V. 48, No. 2, P. 197–204. (in Russian) English transl. in *J. Appl. Math. Mech.*, 1984, V. 48, No. 2, P. 133–139.

[11] Ramis J. P., Schäfke R., Gevrey separation of fast and slow variables. *Nonlinearity*, V. 9, 1996, P. 353–384.

[12] Sauzin D., A new method for measuring the splitting of invariant manifolds, 1999, Notes scientifiques et techniques du Bureau des Longitudes.

[13] Simó C., Averaging under fast quasi-periodic forcing in Hamiltonian mechanics: Integrability and chaotic behavior, J. Seimenis ed., NATO Adv. Sci. Inst. Ser. B Phys. **331**, Plenum Press, New York 1994, 13–34.

[14] D.Treschev, Separatrix splitting for a pendulum with rapidly oscillating suspension point. Russian J. of Math. Phys., 1997, V. 5, No. 1, P. 63-98.

[15] D.Treschev, An averaging method for Hamiltonian systems, exponentially close to integrable ones. Chaos, V. 6, No. 1, 1996, P. 6-14.

[16] D.Treschev, The continuous averaging method in the problem of separation of fast and slow motions. Regular and Chaotic Dynamics, V.2, No.3/4, 1997, P. 9-20. (In Russian)