# Continuous averaging method in the problem of separation of fast and slow motions

D. Treschev

## 1 Fast phase averaging

Mathematical models of various physical processes use systems of ODE which contain an angular variable changing much faster than other variables in the system. Taking the fast phase as a new time, we can rewrite the equations in the form

$$\dot{z} = \varepsilon \hat{v}(z, t, \varepsilon), \qquad z \in M, \tag{1.1}$$

where $M$ is the $m$-dimensional phase space of the system, and $\varepsilon$ is a small parameter. The last one equals to the ratio of a typical velocity of slow variable change to a typical velocity of phase rotation. The vector field $\hat{v}$ is assumed to be smooth[1] and to depend on time $2\pi$-periodically.

It is well known that by a change of the variables it is possible to weaken considerably the dependence of the system (1.1) on time. In particular, by using the standard averaging method (see for example, [2]), it is easy to construct a $2\pi$-periodic in $t$ change of the variables $z \mapsto z_*$ such that the equations (1.1) take the form

$$\dot{z}_* = \varepsilon \hat{v}^0(z_*) + \varepsilon^2 \hat{v}_*(z_*, \varepsilon) + \varepsilon \tilde{v}(z_*, t, \varepsilon). \tag{1.2}$$

Here the only term in the right-hand side depending explicitly on time is $\varepsilon\tilde{v} = O(\varepsilon^K)$. The natural $K$ is arbitrary and

$$\hat{v}^0(z) = \frac{1}{2\pi} \int_0^{2\pi} \hat{v}(z, t, 0)\, dt$$

is the time average of $\hat{v}(z, t, 0)$.

Now suppose that the functions $\hat{v}$ depend analytically on the phase variables. Poincaré noted that in this case power series in the small parameter presenting a change of variables eliminating time from the equations, exist but diverge: terms at $\varepsilon^k$ in these series have the order $k!$. In a general situation this statement has been proved by Sausin [8].

[1] Actually, in the averaging problem only smoothness in $z$ (and in a smaller degree in $\varepsilon$) is important. It is anough to have continuity in $t$ and sometimes even this requirement can be weakened.

Neishtadt [6, 1] noted that in the case of analytic in phase variables functions $\hat{v}$ it is possible to obtain in the equations (1.2)

$$\tilde{v} = O(e^{-\alpha/\varepsilon}), \tag{1.3}$$

where $\alpha > 0$ is some constant and the parameter $\varepsilon$ is assumed to be nonnegative. Hence, the explicit dependence of the equations on time can be made exponentially small. The method Neishtadt used to prove this assertion is based on a large (of order $1/\varepsilon$) number of successive changes of variables. These changes weaken gradually explicit dependence of the equations on time. Ramis and Schafke [7] obtained analogous results analyzing diverging series, produced by the standard averaging method.

It is known also that in general a constant $A > \alpha$ exists such that it is impossible to construct $2\pi$-periodic in $t$ change of the variables $z \to z_*$, such that $\tilde{v} = O(e^{-A/\varepsilon})$. This statement follows, for example, from an estimate of the separatrix splitting rate in Hamiltonian systems of type (1.1) with one and a half degrees of freedom (see., [5], [3], [13], [4]).

Below we evaluate a "maximal" $\alpha$ for which the estimate (1.3) is possible (or at least, we obtain realistic estimates for $\alpha$).

Suppose that the manifold $M$ is real-analytic. We fix its complex neighborhood $M_{\mathbf{C}}$. We denote by $g^t$ phase flow of the averaged system

$$\dot{z} = \hat{v}^0(z). \tag{1.4}$$

Let $Q$ be a compact in $M$, $V_Q$ its (small) neighborhood in $M_{\mathbf{C}}$, and

$$\Sigma_\rho = \{t \in \mathbf{C} : |\operatorname{Im} t| \le \rho\}.$$

Suppose that for any real $s$ such that $|s| < \alpha$ and for any point $z \in V_Q$ the map $g^{is}$ is analytic at the point $z$ and moreover, $g^{is}(z) \in M_{\mathbf{C}}$. We define the set

$$U_{Q,\alpha} = \bigcup_{-\alpha < s < \alpha} g^{is}(V_Q).$$

**Theorem 1** *Let the following assumptions hold.*

*(1) The vector field $\hat{v}^0$ has no singular points in $Q$.*

*The positive constants $\alpha, \rho, \varepsilon_0$ are such that*

*(2) The set $U_{Q,\alpha}$ is defined.*

*(3) The vector field $\hat{v}$ is analytic on the set $U_{Q,\alpha} \times \Sigma_\rho \times [0, \varepsilon_0]$.*

*Then if $\varepsilon_0$ is sufficiently small, there exists a $2\pi$-periodic in $t$ real-analytic map*

$$F : V'_Q \times \Sigma_\rho \times (0, \varepsilon_0) \to M_{\mathbf{C}}, \qquad Q \subset V'_Q \subset V_Q,$$

*which differs from projection to the first multiplier only in terms of order $O(\varepsilon)$ and transforms the vector field $\varepsilon\hat{v}$ to $\varepsilon v_*$ where*

$$v_*(z, t, \varepsilon) = \hat{v}^0(z) + \varepsilon \hat{v}_*(z, \varepsilon) + \tilde{v}(z, t, \varepsilon). \tag{1.5}$$

*The set $V'_Q$ is open and the following estimate holds:*

$$|\tilde{v}(z,t,\varepsilon)| \le Ce^{-\alpha/\varepsilon}, \qquad z \in V'_Q, \quad t \in \Sigma_\rho, \quad \varepsilon \in [0,\varepsilon_0). \tag{1.6}$$

**Remark 1.1** *Actually, the function $\tilde{v}$ is analytic in $t$ in the strip $\Sigma_{\rho+\alpha/\varepsilon}$, but the estimate (1.7) holds only in $\Sigma_{\mathrm{const}}$, where the constant does not depend on $\varepsilon$. In Theorem 2 we take* $\mathrm{const} = \rho$ *for definitness.*

**Remark 1.2** *The same result holds without the assumption of analyticity of $\hat{v}$ with respect to time, but the proof is more complicated.*

Theorem 1 means in particular that in the case when components of the field $\hat{v}$ are entire functions of the variables $z$, the quantity $\alpha$ in relation (1.6) can be arbitrary positive number such that for all $s \in [-\alpha, \alpha]$ the maps $z \mapsto g^{is}(z)$ are holomorphic at the point $z^0$.

Theorem 1 will be obtained below as a corollary of its local version (Theorem 2). To formulate the last one we consider instead of the compact $Q$ the point $z^0 \in M$. Let $V \subset M_{\mathbf{C}}$ be its neighborhood.

Suppose that for any real $s$ such that $|s| < \alpha$ and for any point $z \in V$ the map $g^{is}$ is analytic at the point $z$ and moreover, $g^{is}(z) \in M_{\mathbf{C}}$. Define the set $U_\alpha \subset M_{\mathbf{C}}$ by the equality

$$U_\alpha = \bigcup_{-\alpha < s < \alpha} g^{is}(V).$$

**Theorem 2** *Let the following assumptions hold.*

*(1) The point $z^0$ is not singular for the vector field $\hat{v}^0$.*

*The positive constants $\alpha, \rho, \varepsilon_0$ are such that*

*(2) The set $U_\alpha$ is defined.*

*(3) The vector field $\hat{v}$ is analytic on the set $U_\alpha \times \Sigma_\rho \times [0, \varepsilon_0]$.*

*Then if $\varepsilon_0$ is sufficiently small, there exist $2\pi$-periodic in $t$ real-analytic map $f : V' \times \Sigma_\rho \times (0, \varepsilon_0) \to V$, where $V' \subset V$ is a neighborhood containing the point $z^0$, and the following assertions hold.*

*(a) The map $f$ is smooth in $\varepsilon$ and differs from the projection to the first multiplier only in terms of the order $O(\varepsilon)$.*

*(b) The vector field $\varepsilon\hat{v}$ is transformed by the map $f$ to $\varepsilon v_*$ (see (1.5)), where for some constant $C_0$ the following estimate holds*

$$|\tilde{v}(z,t,\varepsilon)| \le C_0 e^{-\alpha/\varepsilon}, \qquad z \in V', \quad t \in \Sigma_\rho, \quad \varepsilon \in [0,\varepsilon_0). \tag{1.7}$$

The map $f$ can be regarded as change of coordinates in the system (1.1) which reduces explicit dependence of the equations on time to exponentially small terms.

**Remark 1.3** *The assumption that the point $z^0$ is not singular for $\hat{v}^0$ is of purely technical nature. An analogous theorem holds also in the case of singular point $z^0$. Furthermore, the averaging system (see below) in this case can be chosen the same. Proof of the fact that the averaging procedure leads to the required result, differs from proof of Theorem 2 only in technical part.*

Theorem 1 can be deduced from Theorem 2 if we take into account that the averaging system we use in the proof of Theorem 2, is defined independently of the neighborhood $V$ and of the point $z^0$. Hence, the maps $f$ we construct in Theorem 2, pasted together, produce a real-analytic map $F$. We can put $Q' = \cup_z V'(z)$, where $z \in Q$ and the neighborhoods $V'(z)$ are defined in Theorem 2. Due to compactness of $Q$ we can assume that the union is taken over a finite system of sets. The constant $C$ in (1.6) can be taken as maximal one among the constants $C_0$ corresponding to the sets $V'(z)$ from this system.

As an example consider a pendulum with periodically vertically oscillating suspension point. Hamiltonian of the system has the form

$$H(\hat{x}, \hat{y}, t, \varepsilon) = \hat{y}^2/2 + (\Omega^2 + \theta(\omega t)) \cos \hat{x}. \tag{1.8}$$

Here $\hat{x} = \hat{x} \bmod 2\pi$ is the angle between the pendulum and the vertical, $\hat{y}$ is the corresponding momentum, $\Omega > 0$ is the "internal frequency" of the system ($\Omega^2$ equals gravity acceleration divided by length of the pendulum), $\omega$ is the frequency of the suspension point oscillations. The $2\pi$-periodic function $\theta$ determines form of the oscillations. We assume that average of $\theta$ vanishes.

Let the period of the suspension point oscillations is small. More precisely, the equalities

$$\Omega = 1, \quad \omega = 1/\varepsilon$$

hold. Then by using the change $t = \varepsilon\tau$, $H = \varepsilon^{-1}\mathcal{H}$, we get:

$$\mathcal{H} = \varepsilon(\mathcal{H}_0 + \mathcal{H}_1), \qquad \mathcal{H}_0 = y^2/2 + \cos x, \quad \mathcal{H}_1 = \theta(\tau)\cos x. \tag{1.9}$$

System with the Hamiltonian $\mathcal{H}$ obviously has the form (1.1), where the dot denotes the derivative with respect to the new time $\tau$. The corresponding first approximation averaged system (1.4) is also Hamiltonian: the Hamiltonian function equals $\mathcal{H}_0$.

Solutions of the averaged system are elliptic functions of the time $\tau$. Distances to their complex singularities can be expressed in terms of elliptic integrals. Consider the energy level $L_1 = \{(x, y) \in \mathbf{R}^2 : \mathcal{H}_0(x, y) = 1\}$, corresponding to separatrix of the unperturbed pendulum. Distance from the real axis $\operatorname{Re}\tau$ to a singularity of a solution with initial conditions lying on $L_1$ can be evaluated explicitly. It equals $\pi/2$. Hence, according to Theorem 1, if $Q$ lies in a small neighborhood of the separatrices, the constant $\alpha$ in the exponential estimate of the perturbation $\tilde{v}$ can be chosen close to $\pi/2$.

These arguments were used to solve the problem on separatrix splitting in the system with Hamiltonian (1.9) and $\theta(\tau) = \cos\tau$ (see [11, 12, 13]). Note however, that in the indicated papers an asymptotics of the perturbation $\tilde{v}$ is constructed. This naturally requires considerably more delicate assertions in comparison with Theorem 1.

## 2 The continuous averaging method

Proof of Theorem 2 is based on the continuous averaging method.

This method appeared as developement of the Neishtadt averaging procedure [6]. First it was used in papers [9, 10] in the reducibility problem for a linear system with quasi-periodic coefficients. Now let us turn to description of the method.

Let the system of ODE

$$\dot{z} = \hat{a}(z), \tag{2.1}$$

is transformed by the change of the variables

$$z \mapsto Z(z, \Delta). \tag{2.2}$$

Here $z$ is a point of the manifold $M$, $\hat{a}$ is a smooth vector field on $M$, $\Delta$ is a nonnegative parameter and the change (2.2) is defined as a shift along solutions of the equation

$$Z' = f(Z, \delta), \qquad Z(z, 0) = z, \quad 0 \le \delta \le \Delta. \tag{2.3}$$

Prime denotes here derivative with respect to $\delta$.

Let the change $z \mapsto Z$ transforms (2.1) into the following system:

$$\dot{Z} = a(Z, \delta). \tag{2.4}$$

Differentiating the equality (2.4) with respect to $\delta$, we have:

$$\dot{f}(Z, \delta) = a_\delta(Z, \delta) + \partial_f a(Z, \delta) \quad \text{or} \quad a_\delta = [a, f].$$

Here $\partial_f$ is the differential operator on $M$, corresponding to the vector field $f$, the subscript $\delta$ denotes partial derivative and $[\cdot, \cdot]$ is the ordinary vector commutator: $[u_1, u_2] = \partial_{u_1} u_2 - \partial_{u_2} u_1$. Putting $f = \xi a$, where $\xi$ is a certain fixed linear operator, we obtain the Cauchy problem

$$a' = -[\xi a, a], \qquad a|_{\delta=0} = \hat{a}. \tag{2.5}$$

We call the system (2.5) averaging.

Below we will need a non-autonomous analog of the equation (2.5). Suppose that the vector fields $\hat{a}$ and $f$ depend on time explicitly. Then the vector field $a$ also explicitly depends on $t$. To obtain an equation for $a$ we use reduction to the autonomous case. Let $\tau$ be the new variable such that $\dot{\tau} = 1$. The system

$$\dot{z} = \hat{a}(z, \tau), \quad \dot{\tau} = 1$$

is autonomous in the extended phase space $M \times \mathbf{R}_\tau$. Suppose that during the averaging process the coordinate $\tau$ does not change. The corresponding equation (2.5) can be easily transformed to the following one:

$$a_\delta = (\xi a)_t - [\xi a, a], \qquad a|_{\delta=0} = \hat{a}(z,t). \tag{2.6}$$

This is the non-autonomous analog of the system (2.5).

When proving Theorem 2 we solve the Cauchy problem (2.6), where instead of $a$ and $\hat{a}$ we write $\varepsilon v$ and $\varepsilon \hat{v}$:

$$v_\delta = (\xi v)_t - \varepsilon[\xi v, v], \qquad v|_{\delta=0} = \hat{v}(z,t,\varepsilon). \tag{2.7}$$

We define the operator $\xi$ as follows. Let

$$v(z,t,\varepsilon,\delta) = \sum_{k\in\mathbf{Z}} v^k(z,\varepsilon,\delta)e^{ikt}, \quad v^{\pm} = \sum_{\pm k>0} v^k(z,\varepsilon,\delta)e^{ikt}.$$

Then $v = v^0 + v^+ + v^-$. We put

$$\xi v(z,t,\varepsilon,\delta) = i(v^+ - v^-). \tag{2.8}$$

The required change of the variables corresponds to the value $\delta = \alpha/\varepsilon$. Due to the following informal arguments one can hope that the system (2.8) really performs an averaging. Suppose that it is possible to neglect in (2.7) by the terms $\varepsilon[\xi v, v^{\pm}]$, and to change the initial conditions by $\hat{v}(z,t,0)$. In the simplified system we obtain equations for the functions $v^k$ separate:

$$v^k_\delta = -|k|v^k + i\varepsilon \mathrm{sign}\, k\, [v^0, v^k], \qquad v^k|_{\delta=0} = \hat{v}^k(z).$$

Solution of this Cauchy problem can be obtined explicitly:

$$v^k(z,\delta) = e^{-|k|\delta}\hat{v}^k \circ g^{i\varepsilon\delta \mathrm{sign}\, k}(z). \tag{2.9}$$

Here $g^t$ is the phase flow defined in section 1. According to conditions of Theorem 2 the functions (2.9) are defined for $z \in V$ and $\delta \in [0, \alpha/\varepsilon]$. For $\delta = \alpha/\varepsilon$ and $k \neq 0$ the functions (2.9) are exponentially small. Proof of the fact that the solution of the precise system (2.7)–(2.8) behaves analogously is contained in section 4.

Note that according to definition the averaging procedure has the following important property. Suppose that for any fixed $t$ and $\varepsilon$ the vector field $\hat{v}$ belongs to a certain subalgebra $\chi$ in the Lie algebra of vector fields on $M$. Then for fixed $t$, $\varepsilon$ and $\delta$ the vector field $v(z,t,\varepsilon,\delta)$ also belongs to $\chi$. Hence, the diffeomorphism $f$ belongs to the correspondent Lie group. Moreover, if $\hat{v}$ is reversible with respect to some involution $I : M \to M$, the vector field $v(z,t,\varepsilon,\delta)$ is also $I$-reversible. In particular, if the initial vector field $\hat{v}$ is Hamiltonian then $v_* = v(z,t,\varepsilon,\alpha/\varepsilon)$ is also Hamiltonian and the corresponding change $F$ is symplectic.

# 3 Majorants and their properties

Equations of the type (2.5)–(2.6) are not simple to analyse. Even existence of their solution is not trivial. The main tool we use in the analysis of these equations is majorant estimates. The present section is auxiliary. It contains some simple properties of majorants.

Recall that for any two functions $f(z), g(z)$, $z = (z_1, \dots, z_m)$ analytic at the point $z = 0$:

$$f(z) = \textstyle\sum_\beta f_\beta z^\beta, \qquad g(z) = \textstyle\sum_\beta g_\beta z^\beta,$$
$$\beta = (\beta_1, \dots, \beta_m), \qquad \beta_j \geq 0, \qquad z^\beta = z_1^{\beta_1} \cdots z_m^{\beta_m}$$

$g$ is said to be a majorant for $f$ ($f \ll g$) if for any multi-index $\beta$ we have $g_\beta \geq |f_\beta|$. In the case when $f$ and $g$ are vector-valued functions we say that $f \ll g$ if any component of the vector $g$ majorates the corresponding component of the vector $f$.

**Lemma 3.1** *The relation $\ll$ satisfies the following properties:*

*(1) If $f_1 \ll g_1$ and $f_2 \ll g_2$ then $f_1 + f_2 \ll g_1 + g_2$ and $f_1 f_2 \ll g_1 g_2$.*

*(2) If $f \ll g$ then for any $j$ the estimate $\partial f/\partial z_j \ll \partial g/\partial z_j$ holds.*

*(3) If $f(z,\lambda) \ll g(z,\lambda)$ for any value of the parameter $\lambda \in [a,b]$ then*

$$\int_a^b f(z,\lambda)\, d\lambda \ll \int_a^b g(z,\lambda)\, d\lambda.$$

*(4) Let the variable $z$ be scalar. Then for any positive $b < a$ the following estimates hold*

$$1 \ll \frac{b}{b-z}, \qquad \frac{1}{(a-z)(b-z)} \ll \frac{1}{(a-b)(b-z)}.$$

*(5) Let $|f(z)| \leq c$ in the domain*

$$\{z = (z_1, \dots, z_m) : |z_j| \leq b,\ j = 1, \dots, m\}.$$

*Then $f(z) \ll c/w$, where $w = b^{-m}(b - z_1) \cdots (b - z_m)$.*

Assertions (1)–(3) of Lemma 3.1 are obvious. Assertion (4) follows from the formula

$$\frac{1}{(a-z)(b-z)} + \frac{1}{(a-b)(a-z)} = \frac{1}{(a-b)(b-z)}.$$

Assertion (5) follows from the Cauchy formula

$$f(z) = \frac{1}{(2\pi i)^m} \oint d\zeta_1 \dots \oint d\zeta_{m-1} \oint \frac{f(\zeta) d\zeta_m}{(\zeta_1 - z_1) \dots (\zeta_m - z_m)},$$

where the integration is performed along the circles $\{|z_j| = b\}$. Indeed, we have the estimates:

$$\begin{aligned}\Big|\frac{\partial^{\beta_1+\ldots+\beta_m} f}{\partial z_1^{\beta_1}\ldots\partial z_m^{\beta_m}}(0)\Big| &= \Big|\frac{1}{(2\pi i)^m}\oint d\zeta_1\ldots\oint d\zeta_{m-1}\oint\frac{\beta_1!\ldots\beta_m! f(\zeta)d\zeta_m}{\zeta_1^{\beta_1+1}\ldots\zeta_m^{\beta_m+1}}\Big| \\ &\leq \frac{\beta_1!\ldots\beta_m! c}{b^{\beta_1+\ldots\beta_m}} = \frac{\partial^{\beta_1+\ldots+\beta_m}}{\partial z_1^{\beta_1}\ldots\partial z_m^{\beta_m}}\Big|_{z=0}\frac{c}{w}.\end{aligned}$$

By using the following lemma one can estimate Fourier coefficients of an analytic $2\pi$-periodic function in terms of norm of the function in $\Sigma_\rho$.

**Lemma 3.2** *Suppose that the function $\varphi(x)$ is analytic and $2\pi$-periodic in $x$, and $|\varphi| \leq N$ for $x \in \Sigma_\rho$. Let its Fourier expansion have the form $\varphi = \sum_{k\in\mathbf{Z}}\varphi^k e^{kx}$. Then the Fourier coefficients $\phi^k$ satisfy the following estimates:*

$$|\varphi^k| \leq e^{-|k|\rho}N.$$

*Proof of the lemma.* Let us put $\sigma = \operatorname{sign} k$. We have the equalities:

$$\begin{aligned}\varphi^k &= \frac{1}{2\pi}\int_0^{2\pi} e^{-ikx}\varphi(x)dx \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{-ik(x-i\sigma\rho)}\varphi(x-i\sigma\rho)\,dx.\end{aligned}$$

Since the function $\varphi$ in absolute value does not exceed $N$, we obtain the estimate:

$$|\varphi^k| \leq \frac{1}{2\pi}\int_0^{2\pi} e^{-|k|\rho}N dx = e^{-|k|\rho}N.$$

Lemma 3.2 is proved.

# 4 Analytic properties of the averaging procedure

Below for brevity we do not write $\varepsilon$ among arguments of the functions we study.

It is convenient to analyze the averaging system in certain special coordinates $(\tau, \theta)$, $\theta = (\theta_1, \ldots, \theta_{m-1})$, which satisfy the following properties

(1) coordinates of the point $z^0$ vanish;

(2) the maps $z \mapsto \tau(z)$ and $z \mapsto \theta(z)$ are real-analytic;

(3) the phase flow $g^t$ of the system (1.4) in coordinates $(\tau, \theta)$ acts as follows: $(\tau, \theta) \mapsto (\tau + t, \theta)$.

In other words, the functions $\theta$ are real-analytic first integrals of the system (1.4) and $\tau$ is time on solutions of this system.

The indicated coordinates can be constracted as follows. By the theorem on straightening of trajectories, in a neighborhood of the point $z^0$ on $M$ there exist real-analytic coordinates $(\tau, \theta)$ such that the system (1.4) has the form

$$\dot{\tau} = 1, \quad \dot{\theta}_1 = \ldots = \dot{\theta}_{m-1} = 0.$$

It is possible to assume that coordinates of $z^0$ vanish. According to conditions of Theorem 2 the variables $\tau, \theta$ are defined in the domain

$$W_\kappa = \{(\tau, \theta) : |\mathrm{Im}\, \tau| \le \alpha + \kappa,\ |\mathrm{Re}\, \tau| \le \kappa,\ |\theta_j| \le \kappa,\ j = 1, \ldots, m-1\}.$$

for sufficiently small constant $\kappa$. Moreover, the vector field $\hat{v}(\tau, \theta, t)$ is analytic in the domain $W_k \times \Sigma_\rho$.

**Lemma 4.1** *For any real $s \in [-\alpha, \alpha]$ the Fourier coefficients $\hat{v}^k$, $k \ne 0$ in the coordinates $(\tau, \theta)$ satisfy the majorant estimates*

$$\hat{v}^k(\tau - is, \theta, \varepsilon) \ll \frac{\beta e^{-\rho|k|}}{w_*(\kappa)}\mathbf{1}, \qquad w_*(\kappa) = \frac{\kappa - \tau - \theta_1 - \ldots - \theta_{m-1}}{\kappa},$$

*where $\mathbf{1} = (1, \ldots, 1)^T \in \mathbf{R}^m$ and $\beta$ is some constant.*

*Proof.* Let $\beta = \max_{W_k \times \Sigma_\rho} |\hat{v}|$. Then according to Lemma 3.2 we have: $\max_{W_k} |\hat{v}^k| \le \beta e^{-\rho|k|}$. By using assertion (5) of Lemma 3.1, for any $s \in [-\alpha, \alpha]$ we obtain the estimate

$$\hat{v}^k(\tau - is, \theta, \varepsilon) \ll \frac{\beta e^{-\rho|k|}}{w(\kappa)}\mathbf{1}, \qquad w(\kappa) = \frac{(\kappa - \tau)(\kappa - \theta_1)\ldots(\kappa - \theta_{m-1})}{\kappa^m}.$$

It remains to use the simple inequality $1/w \ll 1/w_*$ which can be easily checked by explicit differentiation.

The system (2.7), (2.8) can be rewritten in the following form ($n \in \mathbf{N}$):

$$\begin{aligned} v^0_\delta &= -2i\varepsilon \sum_{k>0} [v^k, v^{-k}], \\ v^n_\delta &= -nv^n + i\varepsilon[v^0, v^n] - 2i\varepsilon \sum_{k>0} [v^{n+k}, v^{-k}], \\ v^{-n}_\delta &= -nv^{-n} - i\varepsilon[v^0, v^{-n}] - 2i\varepsilon \sum_{k>0} [v^k, v^{-n-k}]. \end{aligned}$$

Let us put

$$\begin{aligned} v^0(\tau, \theta, \delta) &= (1, 0, \ldots, 0) + u^0(\tau, \theta, \delta), \\ v^l(\tau, \theta, \delta) &= u^l(\tau + i\varepsilon\delta\, \mathrm{sign}\, l, \theta, \delta) e^{-(\rho+\delta)|l|}, \qquad l \ne 0. \end{aligned} \tag{4.1}$$

The functions $u^l$ satisfy the equations

$$\begin{aligned}
u^0_\delta(\tau,\theta,\delta) &= -2i\varepsilon\sum_{k>0}[u^k(\tau+i\varepsilon\delta,\theta,\delta),u^{-k}(\tau-i\varepsilon\delta,\theta,\delta)]e^{-2(\rho+\delta)k},\\
u^n_\delta(\tau+i\varepsilon\delta,\theta,\delta) &= i\varepsilon[u^0(\tau,\theta,\delta),u^n(\tau+i\varepsilon\delta,\theta,\delta)]\\
&\quad -2i\varepsilon\sum_{k>0}[u^{n+k}(\tau+i\varepsilon\delta,\theta,\delta),u^{-k}(\tau-i\varepsilon\delta,\theta,\delta)]e^{-2(\rho+\delta)k},\\
u^{-n}_\delta(\tau-i\varepsilon\delta,\theta,\delta) &= -i\varepsilon[u^0(\tau,\theta,\delta),u^{-n}(\tau-i\varepsilon\delta,\theta,\delta)]\\
&\quad -2i\varepsilon\sum_{k>0}[u^{k}(\tau+i\varepsilon\delta,\theta,\delta),u^{-n-k}(\tau-i\varepsilon\delta,\theta,\delta)]e^{-2(\rho+\delta)k},\\
u^0(\tau,\theta,0) &= 0,\quad u^l(\tau,\theta,0)=\hat v^l(\tau,\theta),\qquad l\in\mathbf{Z}.
\end{aligned}$$

Here we stress that in the left-hand sides of these differential equations the derivative in $\delta$ is partial, i.e. differentiation in $\delta$ that enter expressions of the form $\tau\pm i\varepsilon\delta$ is not performed.

For any two vector fields

$$W'(z)=(W_1'(z),\dots,W_m'(z)),\quad W''(z)=(W_1''(z),\dots,W_m''(z)),$$

where $z=(z_1,\dots,z_m)$ denotes the coordinates $(\tau,\theta)$, we put

$$[[W',W'']]=\sum_{s=1}^m\big(W_s'\,\partial W''/\partial z_s+W_s''\,\partial W'/\partial z_s\big).$$

The bi-linear operation $[[\,,]]$ is not defined in an invariant way: it depends on the coordinates. However, it majorates the commutator in the following sense. Suppose that the estimates $W_1\ll W'$, $W_2\ll W''$ hold. Then $[W_1,W_2]\ll[[W',W'']]$.

Consider the so called majorant system ($n\in\mathbf{N}$)

$$\begin{aligned}
U^0_\delta &= 2\varepsilon\sum_{k>0}[[U^k,U^{-k}]]e^{-2(\rho+\delta)k},\\
U^n_\delta &= \varepsilon[[U^0,U^n]]+2\varepsilon\sum_{k>0}[[U^{n+k},U^{-k}]]e^{-2(\rho+\delta)k},\\
U^{-n}_\delta &= \varepsilon[[U^0,U^{-n}]]+2\varepsilon\sum_{k>0}[[U^{k},U^{-n-k}]]e^{-2(\rho+\delta)k},\\
U^0(\tau,\theta,0) &= 0,\quad U^l(\tau,\theta,0)=\beta w_*^{-1}(\kappa)\mathbf{1},\qquad l\neq 0.
\end{aligned}$$

The following proposition plays a key role in the proof of Theorem 2.

**Proposition 4.1** *Let the solution $U^l$ of the majorant system is defined for $0\le\delta\le\Delta$, $\varepsilon\Delta\le\alpha$. Then solution of the Cauchy problem for the functions $u^l$ is defined for $0\le\delta\le\Delta$ and for any real $s$ such that $|s|\le\alpha-\varepsilon\delta$ the estimates*

$$u^l(\tau+i\varepsilon\delta\,\mathrm{sign}\,l+is,\theta,\delta)\ll U^l(\tau,\theta,\delta)$$

*hold, where the majorant inequality is considered with respect to the variables $\tau$ and $\theta$.*

Proof of the proposition is obvious. Indeed, the right-hand sides of the equations for the functions $U^l$ majorate the right-hand sides of the equations for the functions $u^l$. According to Lemma 4.1 the initial conditions for the corresponding Cauchy problems are in the same relation. This implies the required assertion.

It is possible to search for a solution of the majorant system in the form $U^l = U$, $l \neq 0$, where the vector-function $U$ does not depend on $l$. One can assume that the functions $U^0$ and $U$ depend on the variables $(\tau, \theta)$ only through the combination $x = \tau + \theta_1 + \ldots + \theta_{m-1}$. Moreover, the vectors $U^0$ and $U$ can be regarded parallel to the vector $\mathbf{1}$. We denote components of the vectors $U^0$ and $U$ by the symbols $y^0$ and $y$ respectively, where $y^0 = y^0(x,\delta)$ and $y = y(x,\delta)$. The majorant system gives us the following Cauchy problem:

$$\begin{aligned} y^0_\delta &= \frac{4m\varepsilon e^{-2(\rho+\delta)}}{1-e^{-2(\rho+\delta)}} y y_x, \\ y_\delta &= m\varepsilon(y^0 y)_x + \frac{4m\varepsilon e^{-2(\rho+\delta)}}{1-e^{-2(\rho+\delta)}} y y_x, \\ y^0|_{\delta=0} &= 0, \qquad y|_{\delta=0} = \frac{\beta\kappa}{\kappa - x}. \end{aligned} \tag{4.2}$$

We simplify the equations making right-hand sides of these equations slightly greater. Indeed, by using the inequality

$$\frac{4me^{-2\rho}}{1-e^{-2(\rho+\delta)}} \leq 2a, \qquad a = \frac{2me^{-2\rho}}{1-e^{-2\rho}},$$

we obtain the following Cauchy problem solutions of which for $\delta \geq 0$ majorate solutions of the problem (4.2):

$$\begin{aligned} y^0_\delta &= a\varepsilon e^{-2\delta}(y^2)_x, \\ y_\delta &= m\varepsilon(y^0 y)_x + a\varepsilon e^{-2\delta}(y^2)_x, \\ y^0|_{\delta=0} &= 0, \qquad y|_{\delta=0} = \frac{\beta\kappa}{\kappa - x}. \end{aligned} \tag{4.3}$$

We perform analysis of the solution for the system (4.3) in two steps. The first step corresponds to the magnitudes $\delta \in [0, \log\varepsilon^{-1}]$, and the second one to $\delta \in [\log\varepsilon^{-1}, \alpha/\varepsilon)$.

**Lemma 4.2** *For $\delta \in [0, \log\varepsilon^{-1}]$ solution of the system (4.3) satisfies the following estimates:*

$$\begin{aligned} y^0(x,\delta) \ll y(x,\delta) &\ll Y(\kappa, \beta, x, 2(m+a)\varepsilon\delta), \\ Y(\kappa,\beta,x,\lambda) &= \frac{2\beta\kappa}{\kappa - x + \sqrt{(\kappa-x)^2 - 4\beta\kappa\lambda}}. \end{aligned}$$

*Proof of Lemma 4.2.* First, note that solution of the system (4.3) satisfies the majorant inequality $y^0(x,\delta) \ll y(x,\delta)$ for all magnitudes $\delta \geq 0$, such that both functions are defined. The reason for this is that the indicated estimate holds for $\delta = 0$ and the right-hand side of the differential equation (4.3) for $y$ majorates the right-hand side of the equation (4.3) for $y^0$.

This remark implies that the solution $\varphi(x,\delta)$ of the Cauchy problem

$$\varphi_\delta = (m+a)\varepsilon(\varphi^2)_x, \qquad \varphi|_{\delta=0} = \frac{\beta\kappa}{\kappa - x}$$

majorates the function $y(x,\delta)$ for any $\delta \geq 0$ such that $\varphi(x,\delta)$ is defined.

The function $\varphi$ can be evaluated explicitly: $\varphi(x,\delta) = Y(\kappa,\beta,x,2(m+a)\varepsilon\delta)$. Thus, the lemma is proved.

The function $\varphi$ is defined on the interval $\delta \in [0,\alpha/\varepsilon]$ only for sufficiently small $\alpha$. We do not want to assume that $\alpha$ is small. So, for $\delta \geq \log\varepsilon^{-1}$ we use more delicate estimates.

**Lemma 4.3** *For small $\varepsilon > 0$ and $\delta \in [\log\varepsilon^{-1},\alpha/\varepsilon]$ solution of the system (4.3) satisfies the following estimates:*

$$\begin{aligned} y^0(x,\delta) &\ll y(x,\delta) \ll Y(\kappa/2,\beta_*,x,4m\varepsilon^2\log\varepsilon^{-1}\,\delta), \\ \beta_* &= \max\{4\beta, 256\beta^2 a/\kappa\}. \end{aligned} \tag{4.4}$$

Note that Lemmas 4.2–4.3 imply Theorem 2. Indeed, the functions $y^0, y$ turn out to be analytic at zero for all $\delta \in [0,\alpha/\varepsilon]$. Then solution of the majorant system satisfies the estimates

$$U^l(\tau,\theta,\delta) \ll y(x,\delta), \qquad 0 \leq \delta \leq \alpha/\varepsilon.$$

According to Proposition 4.1 we have:

$$u^l(\tau + i\varepsilon\delta\,\mathrm{sign}\,l,\theta,\alpha/\varepsilon) \ll U^l(\tau,\theta,\alpha/\varepsilon).$$

Hence, due to the equalities (4.1)

$$\begin{aligned} v^0(\tau,\theta,\alpha/\varepsilon) - (1,0,\ldots,0) &\ll y(x,\alpha/\varepsilon), \\ v^l(\tau,\theta,\alpha/\varepsilon) &\ll y(x,\alpha/\varepsilon)e^{-(\rho+\alpha/\varepsilon)|l|}, \qquad l \neq 0, \end{aligned}$$

what implies the estimate (1.7).

In the rest part of this section we prove Lemma 4.3.

We need more precise than in Lemma 4.2 estimate for the function $y^0$:

$$\begin{aligned} y^0(x,\log\varepsilon^{-1}) &= \int_0^{\log\varepsilon^{-1}} a\varepsilon e^{-2\mu}(y^2(x,\mu))_x\,d\mu \\ &\ll \int_0^{\log\varepsilon^{-1}} a\varepsilon(Y^2(\kappa,\beta,x,2(m+a)\varepsilon\mu))_x\,d\mu \\ &\ll a\varepsilon\log\varepsilon^{-1}(Y^2(\kappa,\beta,x,2(m+a)\varepsilon\log\varepsilon^{-1}))_x. \end{aligned}$$

**Proposition 4.2** *For small $\varepsilon > 0$ the following relations hold:*

$$Y\big(\kappa,\beta,x,2(m+a)\varepsilon\log\varepsilon^{-1}\big) \quad \ll \quad \frac{2\beta\kappa}{\kappa/2-x} \tag{4.5}$$

$$\Big(Y^2\big(\kappa,\beta,x,2(m+a)\varepsilon\log\varepsilon^{-1}\big)\Big)_x \quad \ll \quad \frac{128\beta^2}{\kappa/2-x}. \tag{4.6}$$

Proof of Proposition 4.2 is based on direct estimates. Indeed,

$$Y\big(\kappa,\beta,x,2(m+a)\varepsilon\log\varepsilon^{-1}\big) = \tfrac{2\beta\kappa}{\kappa-x+Q},$$
$$Q = \sqrt{(\kappa-x)^2 - 8\beta\kappa(m+a)\varepsilon\log\varepsilon^{-1}}.$$

Since $0 \ll (\kappa - x + Q)^{-1} \ll Q^{-1}$, to prove the relation (4.5), it is suffisient to sheck that $1/Q \ll 1/(\kappa/2-x)$. Actually, the stronger estimate holds:

$$\frac{1}{Q} \ll \frac{1}{\kappa_* - x}, \qquad \kappa_* = \kappa - \sqrt{\lambda_*}, \quad \lambda_* = 8\beta\kappa(m+a)\varepsilon\log\varepsilon^{-1}.$$

Indeed,

$$\frac{1}{Q} = \frac{1}{\sqrt{(\kappa_* - \sqrt{\lambda_*} - x)(\kappa_* + \sqrt{\lambda_*} - x)}} \ll \frac{1}{(\kappa_* - \sqrt{\lambda_*} - x)}.$$

The relation (4.5) is proved since $\kappa_* > \kappa/2$ for small $\varepsilon > 0$ (below we use the inequality $\kappa_* > 3\kappa/4$).

Now let us check the estimate (4.6). We have:

$$\begin{aligned}(Y^2(\kappa,\beta,x,2(m+a)\varepsilon\log\varepsilon^{-1}))_x \quad &= \quad \frac{8\beta^2\kappa^2}{(\kappa-x+Q)^2Q}\\ &\ll \quad \frac{8\beta^2\kappa^2}{Q^3} \ll \frac{8\beta^2\kappa^2}{(\kappa_*-x)^3}.\end{aligned}$$

Hence, it is sufficient to show that

$$\frac{8\beta^2\kappa^2}{(\kappa_*-x)^3} \ll \frac{128\beta^2}{\kappa/2-x}.$$

By using assertion (4) of Lemma 3.1, we get:

$$\frac{128\beta^2}{\kappa/2-x} \ll \frac{32\beta^2\kappa}{(\kappa/2-x)(3\kappa/4-x)} \ll \frac{8\beta^2\kappa^2}{(\kappa/2-x)(3\kappa/4-x)^2} \ll \frac{8\beta^2\kappa^2}{(\kappa_*-x)^3}.$$

The proposition is proved.

According to Proposition 4.2, the following estimates are valid:

$$y^0(x,\log\varepsilon^{-1}) \ll \frac{128\beta^2 a\varepsilon\log\varepsilon^{-1}}{\kappa/2-x}, \qquad y(x,\log\varepsilon^{-1}) \ll \frac{2\beta\kappa}{\kappa/2-x}.$$

Hence, for $\delta \geq \log \varepsilon^{-1}$ the functions $y^0, y$ are majorated respectively by the functions $\psi^0, \psi$. The last ones satisfy the system

$$
\begin{aligned}
\psi^0_\delta &= a\varepsilon^3(\psi^2)_x, \\
\psi_\delta &= m\varepsilon(\psi^0\psi)_x + a\varepsilon^3(\psi^2)_x, \qquad (4.7)\\
\psi^0|_{\delta=\log\varepsilon^{-1}} &= \varepsilon\log\varepsilon^{-1} f(x), \qquad \psi|_{\delta=\log\varepsilon^{-1}} = f(x), \\
f(x) &= \frac{\beta_*\kappa/2}{\kappa/2 - x}, \qquad (4.8)
\end{aligned}
$$

where the quantity $\beta_*$ is defined by the equality (4.4). We obtain the system (4.7) from (4.3) replacing $e^{-2\delta}$ by the greater magnitude $\varepsilon^2$. To estimate solution of this system, we put

$$\psi^0 = \varepsilon\log\varepsilon^{-1} f(x) + \varepsilon\sqrt{a/m}\,\hat\psi^0,$$

where $\hat\psi^0$ is new unknown function. The system (4.7) takes the form

$$
\begin{aligned}
\hat\psi^0_\delta &= \sqrt{ma}\,\varepsilon^2(\psi^2)_x, \\
\psi_\delta &= m\varepsilon^2\log\varepsilon^{-1}(f\psi)_x + \sqrt{ma}\,\varepsilon^2(\hat\psi^0\psi)_x + a\varepsilon^3(\psi^2)_x, \\
\hat\psi^0|_{\delta=\log\varepsilon^{-1}} &= 0, \qquad \psi|_{\delta=\log\varepsilon^{-1}} = f(x).
\end{aligned}
$$

For $\delta \geq \log\varepsilon^{-1}$ the obvious estimates $\psi \gg \hat\psi^0$, $\psi \gg f$ hold. Hence, the relation

$$\psi(x,\delta) \ll \chi(x,\delta), \qquad \delta \geq \log\varepsilon^{-1}$$

is valid, where $\chi$ is the solution of the following Cauchy problem:

$$
\begin{aligned}
\chi_\delta &= (m\log\varepsilon^{-1} + \sqrt{ma} + a\varepsilon)\varepsilon^2(\chi^2)_x, \\
\chi|_{\delta=\log\varepsilon^{-1}} &= f(x).
\end{aligned}
$$

The function $\chi$ can be found explicitly:

$$\chi(x,\delta) = Y(\kappa/2, \beta_*, x, 2(m\log\varepsilon^{-1} + \sqrt{ma} + a\varepsilon)\varepsilon^2\delta) \ll Y(\kappa/2, \beta_*, x, 4m\log\varepsilon^{-1}\varepsilon^2\delta).$$

Lemma 4.3 is proved.

# References


[1] Arnold V. I., Kozlov V. V., and Neishtadt A. I. Mathematical aspects of classical and celestial mechanics. Enciclopedia of Mathematical Sciences, Vol.3, V. I. Arnold Ed., Springer Verlag, 1988.

[2] Bogolyubov N. N. and Mitropolsky Yu. A. Asymptotic methods in the theory of nonlinear oscillations. ??????

[3] Delshams A., Seara T. M., An asymptotic Expressions for the Splitting of Separatrices of the Rapidly Forced Pendulum. *Comm. Math. Phys.* 1992, V. 150, P. 433–463.

[4] Gelfreich V., Reference systems ... Preprint.

[5] Holmes P., Marsden J., Scheurle J., Exponentially small splittings of separatrices with applications to KAM theory and degenerate bifurcations *Contemp. Math.*, 1988, V. 81, P. 213–244.

[6] Neishtadt A. I., 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.

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

[8] Sauzin D., Caractère Gevrey des solutions formelles dún problème de moyennisation. *C. R. Acad. Sci. Paris*, T. 315, Serie I, 1992, P. 991–995.

[9] Treschev D., An Estimate of Irremovable Nonconstant Terms in the Reducibility Problem, *Amer. Math. Soc. Transl.* (2), 1995, V. 168, P. 91–128.

[10] Treschev D., On the Reducibility of the One-Dimensional Schrödinger Equation with Quasi-Periodic Potential *Amer. Math. Soc. Transl.* (2), 1995, V. 168, P. 129–140.

[11] Treschev D., Separatrix Splitting for a Pendulum with Rapidly Oscillating Suspension Point. *Preprint* no. 273, Centre de Recerca Matematica, Institute D'Estudis Catalans, Barcelona, Spain, 1994.

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

[13] Treschev D., Separatrix Splitting for a Pendulum with Rapidly Oscillating Suspension Point. *Russian J. of Math. Phys.*, 1997.