HAL
open science

# KRYLOV- BOGOLYUBOV AVERAGING

Wenwen Jian, Sergei Kuksin, Yuan Wu

▶ **To cite this version:**

Wenwen Jian, Sergei Kuksin, Yuan Wu. KRYLOV- BOGOLYUBOV AVERAGING. Russian Mathematical Surveys, 2020, 75, pp.427-444. hal-02386046

**HAL Id: hal-02386046**

**https://hal.science/hal-02386046**

Submitted on 29 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THE BOGOLYUBOV–KRYLOV AVERAGING

WENWEN JIAN, SERGEI KUKSIN, AND YUAN WU

Abstract. We present the modified approach to the classical Bogolyubov-Krylov averaging, developed recently for the purpose of PDEs. It allows to treat Lipschitz perturbations of linear systems with pure imaginary spectrum and may be generalized to treat PDEs with small nonlinearities.

## 1. Introduction

The classical Bogolyubov-Krylov averaging method is a method for approximated analysis of nonlinear oscillating process. Among a number of its equivalent or closely related formulations, we choose the following. In the space $\mathbb{R}^N$, let us consider the differential equation of the form

$$\frac{\partial v}{\partial t} + Av = \epsilon P(v), \qquad v(0) = v_0, \qquad 0 < \epsilon \leq 1, \tag{1.1}$$

where $A$ is a linear operator with pure imaginary eigenvalues without Jordan cells, and $P(v)$ is a nonlinearity. The task is to study the behavior of solutions for (1.1) on time-interval of order $\epsilon^{-1}$. Let us firstly pass to the slow time $\tau = \epsilon t$ and rewrite the equation as

$$\frac{\partial v}{\partial \tau} + \epsilon^{-1} Av = P(v), \qquad v(0) = v_0, \tag{1.2}$$

where now $\tau$ is a time-variable of order one. Secondly, in equation (1.2), let us pass to the interaction representation variables [1]

$$a(\tau) = e^{\epsilon^{-1}\tau A} v(\tau), \tag{1.3}$$

and rewrite the equation as

$$\frac{\partial a}{\partial \tau} = e^{\epsilon^{-1}\tau A} P(e^{-\epsilon^{-1}\tau A} a), \qquad a(0) = v_0. \tag{1.4}$$

The Bogolyubov-Krylov averaging theorem is the following result:

**Theorem 1.1.** *Assume that the vector-field $P$ is locally Lipschitz continuous. Then*
*1) the limit*

$$\langle\langle P \rangle\rangle(a) = \lim_{T \to \pm\infty} \frac{1}{|T|} \int_0^T e^{\epsilon^{-1} sA} P(e^{-\epsilon^{-1} sA} a) \mathrm{d}s \tag{1.5}$$

*exists for all $a \in \mathbb{R}^n$.*
*2) There exists $\theta = \theta(|v_0|) > 0$, such that for $|\tau| \leq \theta$, a solution $a^\epsilon(\tau)$ of equation (1.4) is $o(1)$-close, as $\epsilon \to 0$, to a solution of the equation*

$$\frac{\partial a}{\partial \tau} = \langle\langle P \rangle\rangle(a), \qquad a(0) = v_0.$$

[1] under this name the change of variable (2.3) is known in physics.

We stress that the only restriction imposed on the spectrum of the operator $A$ is that it is pure imaginary. Theorem 1.1 and related results were proved by Bogolyubov-Krylov in a number of works in 1930's. The research was summarized in the book [4], also see in [2]. In our work we present a proof of Theorem 1.1, based on a variation of the Bogolyubov-Krylov argument, developed recently for the purposes of partial differential equations in [7, 8]. It allows to prove the averaging theorem above under minimal restrictions on the smoothness of the nonlinearity $P$ – only its Lipschitz continuity is required – and it generalizes to a class of perturbative problems in PDEs. Theorem 1.1 is proved in Sections 3-4. In Section 5, we discuss its applications to the case when (1.2) is a Hamiltonian system. We remind that the Bogolyubov-Krylov averaging method was the first rigorously justified averaging theory. Before the work of Bogolyubov-Krylov the method of averaging existed as a heuristic theory, after that other rigorous averaging theories were created, see in [2]. In particular, now the method of averaging applies to equations with added stochasticity. The approach of our work, enriched with the ideas of the seminal work [9], suits well to the situation when the stochasticity is added to the problem in the form a stochastic force; both in the ODE and PDE settings. See the second half of the paper [8] and references in that work.

Our paper is based on the lecture notes for a course that SK was teaching at a CIMPA School on Dynamical Systems in Kathmandu (October 25 – November 5, 2018), which was attended by WJ and YW.

**Notation.** Abbreviation l.h.s. (r.h.s.) stands for "left hand side" ("right hand side"). By $\mathbb{R}_+$ (by $\mathbb{Z}_+$) we denote the set of non-negative real numbers (non-negative integers), denote by $B_R$ the open ball $B_R = \{v| : |v| < R\}$, $R > 0$, and by $\bar{B}_R$ – its closure.

**Acknowledgements.** SK was supported by the grant 18-11-00032 of Russian Science Foundation. WJ and YW were supported by the National Natural Science Foundation of China (No. 11790272 and No. 11421061).

## 2. PRELIMINARIES

Consider again equation (1.2), and assume that the linear operator $A$ has $N$ eigenvalues, counted with their geometrical multiplicities. Assume also that these eigenvalues are pure imaginary. Then they go in pairs $\pm i\lambda_j$, where $0 \neq \lambda_j \in \mathbb{R}$ (see [5]). So $N$ is an even number, $N = 2n$. The imposed restrictions on $A$ are equivalent to the following conditions (see [5]): $KerA = \{0\}$ and in $\mathbb{R}^{2n}$ exists a basis $\{e_1^+, e_1^-, e_2^+, e_2^-, ..., e_n^+, e_n^-\}$ such that in the corresponding coordinates $\{x_1, y_1, x_2, y_2, ..., x_n, y_n\}$, the matrix of the linear operator $A$ has the form

$$\left(\begin{array}{c|c|c|c} \begin{matrix} 0 & -\lambda_1 \\ \lambda_1 & 0 \end{matrix} & \mathbf{0} & \dots & \mathbf{0} \\ \hline \mathbf{0} & \begin{matrix} 0 & -\lambda_2 \\ \lambda_2 & 0 \end{matrix} & \dots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \dots & \begin{matrix} 0 & -\lambda_n \\ \lambda_n & 0 \end{matrix} \end{array}\right). \tag{2.1}$$

Thus, the original unperturbed linear system $(1.1)|_{\epsilon=0}$ reads:

$$\begin{cases} \dot{x}_1 - \lambda_1 y_1 = 0, \\ \dot{y}_1 + \lambda_1 x_1 = 0, \\ \dots \\ \dot{x}_n - \lambda_n y_n = 0, \\ \dot{y}_n + \lambda_n x_n = 0. \end{cases} \tag{2.2}$$

Note that this linear system can be written in the Hamiltonian form

$$\dot{x} = -\frac{\partial h}{\partial y}, \qquad \dot{y} = \frac{\partial h}{\partial x},$$

where $h = -\frac{1}{2}\sum_{j=1}^{n} \lambda_j (x_j^2 + y_j^2)$.

2.1. **Complex structures in $\mathbb{R}^{2n}$ and real analysis in $\mathbb{C}^n$. Systems (2.2) and (1.2) in complex notation.** The systems (2.2) and (1.2) can be written more compactly if we introduce in the space $\mathbb{R}^{2n}$ a complex structure and write $A$ and the perturbation $P$ in its terms. Corresponding construction is performed in this section and is used below to prove Theorem 1.1; the complex language allows to shorten the proof significantly.

Vectors in the space $\mathbb{R}^{2n}$ are caracterised by the coordinates $(x_1, y_1, x_2, y_2, ..., x_n, y_n)$. Let us introduce in $\mathbb{R}^{2n}$ a complex structure by denoting

$$\begin{cases} z_1 = x_1 + \mathbf{i} y_1, \\ z_2 = x_2 + \mathbf{i} y_2, \\ \dots \\ z_n = x_n + \mathbf{i} y_n. \end{cases} \tag{2.3}$$

Then the real space $\mathbb{R}^{2n}$ becomes a space of complex sequences $z = (z_1, z_2, ..., z_n)$ with $z_j \in \mathbb{C}$. That is, we have achieved that

$$\mathbb{R}^{2n} \simeq \mathbb{C}^n.$$

In the complex notation, the Euclidean scalar product $\langle \cdot, \cdot \rangle$ in $\mathbb{R}^{2n} \simeq \mathbb{C}^n$ reads

$$\langle z, z' \rangle = \Re(\sum_j z_j \bar{z}'_j) =: \Re(z \cdot \bar{z}'). \tag{2.4}$$

For the real numbers $\lambda_1, \lambda_2, ..., \lambda_n$ as in (2.2) let us consider the linear operator

$$\mathrm{diag}\{\mathbf{i}\lambda_j\} : \ \mathbb{C}^n \to \mathbb{C}^n, \quad (z_1, z_2, ..., z_n) \mapsto (\mathbf{i}\lambda_1 z_1, \mathbf{i}\lambda_2 z_2, ..., \mathbf{i}\lambda_n z_n).$$

In the real coordinates $(x_1, y_1, x_2, y_2, ..., x_n, y_n)$ it reads

$$(x_1, y_1, x_2, y_2, ..., x_n, y_n) \mapsto (-\lambda_1 y_1, \lambda_1 x_1, -\lambda_2 y_2, \lambda_2 x_2, ..., -\lambda_n y_n, \lambda_n x_n).$$

That is, in the complex coordinates the operator $A$ with the matrix (2.1) is the operator $\mathrm{diag}\{\mathbf{i}\lambda_j\}$, so the system of linear equations $(1.1)|_{\epsilon=0} = (2.2)$ reduces to the diagonal complex system

$$\dot{v}_j + \mathbf{i}\lambda_j v_j = 0, \quad 1 \le j \le n.$$

In the complex notation the perturbed system (1.1) reads

$$\dot{v}_j + \mathbf{i}\lambda_j v_j = \epsilon P_j(v), \quad v(0) = v_0, \quad v = (v_1, v_2, ..., v_n) \in \mathbb{C}^n. \tag{2.5}$$

Below we assume that the vector-field $P$ is locally Lipschitz, i.e. its restrictions to bounded balls $B_R$, $R > 0$, are Lipschitz-continuous. The case of polynomial vector-field $P$ will be for us of special interest, and we start with its brief discussion.

**Definition 2.1.** A complex function $F:\mathbb{C}^n\to\mathbb{C}$ is a polynomial if it can be written as

$$F(z)=\sum_{0\le|\alpha|,|\beta|\le M} C_{\alpha\beta}z^\alpha\bar z^\beta,$$

where $\alpha=(\alpha_1,\alpha_2,...\alpha_n),\beta=(\beta_1,\beta_2,...\beta_n)\in\mathbb{Z}_+^n$ are multi-indices with the norm $|\alpha|=|\alpha_1|+|\alpha_2|+\cdots+|\alpha_n|$, $C_{\alpha\beta}$ are some complex numbers, and

$$z^\alpha=\prod_{j=1}^n z_j^{\alpha_j},\qquad \bar z^\beta=\prod_{j=1}^n \bar z_j^{\beta_j}.$$

**Definition 2.2.** A vector-field $P(z)$ is polynomial if every its component $P_j$ is a polynomial function.

We recall that for a function $f(z)$ (real or complex) of a complex variable $z=x+\mathbf{i}y$, the derivatives $\partial f/\partial z$ and $\partial f/\partial\bar z$ are defined as $\frac{\partial f}{\partial z}=\frac12(\frac{\partial f}{\partial x}-\mathbf{i}\frac{\partial f}{\partial y})$ and $\frac{\partial f}{\partial\bar z}=\frac12(\frac{\partial f}{\partial x}+\mathbf{i}\frac{\partial f}{\partial y})$. The lemma below follows by elementary calculation:

**Lemma 2.3.** *Let $z\in\mathbb{C}$ and consider a complex polynomial $F(z)=\sum\limits_{m,n=1}^M C_{mn}z^m\bar z^n$. Then $\frac{\partial F}{\partial z}=\sum\limits_{m,n=1}^M mC_{mn}z^{m-1}\bar z^n$, and $\frac{\partial F}{\partial\bar z}=\sum\limits_{m,n=1}^M nC_{mn}z^m\bar z^{n-1}$. If $F:\mathbb{C}^n\to\mathbb{R}$ is a real-valued $C^1$-smooth function, then $\overline{\frac{\partial F}{\partial z_j}}=\frac{\partial F}{\partial\bar z_j}$, for any $1\le j\le n$.*

**Example 2.4.** For $n=1$, consider the system of equations

$$\dot x-\omega y=2\epsilon x(x^2-y^2),\quad \dot y+\omega x=4\epsilon yx^2. \tag{2.6}$$

Now $\lambda=\omega$ and in the complex coordinate $v=x+iy$ the system reads

$$\dot v+i\lambda v=\epsilon(v^2\bar v+v^3). \tag{2.7}$$

Let us come back to the general case of locally Lipschitz vector-fields $P$.

**Definition 2.5.** Let $\mathcal{X}:\mathbb{R}_+\to\mathbb{R}_+$ be a non-decreasing continuous function and $f:\mathbb{C}^n\to\mathbb{C}^n$ be a continuous vector-field. We say that $f\in\mathrm{Lip}_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$, if for any $R\ge0$, $|f|_{B_R}\le\mathcal{X}(R)$ and $\mathrm{Lip}f|_{B_R}\le\mathcal{X}(R)$.

**Example 2.6.** Let $P:\mathbb{C}^n\to\mathbb{C}^n$ be a $C^1$-smooth vector-field. For $v\in\mathbb{C}^n$ we denote by $dP(v)$ the differential of $P$ at $v$ (this is a linear over real numbers map from $\mathbb{C}^n$ to $\mathbb{C}^n$). Denote $\widetilde{\mathcal{X}}(R;P)=\max\left\{\sup_{B_R}\|dP(v)\|,\sup_{B_R}|P(v)|\right\}$. Then $\widetilde{\mathcal{X}}$ defines a continuous function of $R\ge0$, and $P\in\mathrm{Lip}_{\widetilde{\mathcal{X}}}(\mathbb{C}^n,\mathbb{C}^n)$. Indeed, the continuity of $\mathcal{X}$ is obvious, while the second property follows from the mean-value theorem which implies that

$$|P(u_2)-P(u_1)|\le\widetilde{\mathcal{X}}(R;P)|u_2-u_1|\quad\text{if}\ \ u_1,u_2\in B_R.$$

**Lemma 2.7.** *Let $P\in Lip_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$ and $v_0\in\bar B_R$. Denote $\theta=\frac{R}{\mathcal{X}(2R)}$. Then a solution $v(t)$ of (2.5) exists for $|t|\le\epsilon^{-1}\theta$ and stays in the ball $\bar B_{2R}$.*

*Proof.* Since $P$ is a locally Lipschitz vector-field, then a solution $v(t)$ of equation (2.5) exists till the blow-up time. Taking the scalar product of equation (2.5) with $v(t)$ (see (2.4)), we get

$$\frac12\frac{\mathrm{d}}{\mathrm{d}t}|v(t)|^2=-\mathbf{i}\langle\mathrm{diag}(\lambda_j)v,v\rangle+\epsilon\langle P(v),v\rangle=\epsilon\langle P(v),v\rangle.$$

Denote $T = \inf\left\{t \in [0, \epsilon^{-1}\theta] : |v(t)| \geq 2R\right\}$, where $T$ equals $\epsilon^{-1}\theta$ if the set under the inf-sign is empty. Then for $0 < t \leq T$ we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|v(t)|^2 \leq \epsilon|v||P(v)| \leq \epsilon\mathcal{X}(2R)|v|.$$

Thus, $\frac{\mathrm{d}}{\mathrm{d}t}|v(t)| \leq \epsilon\mathcal{X}(2R)$ and $|v(t)| \leq R + \epsilon\mathcal{X}(2R)t < 2R$ for all $0 < t < \epsilon^{-1}\theta$. So $T = \epsilon^{-1}\theta$ and the result follows. □

2.2. **Slow time and interaction representation.** Denote $\tau = \epsilon t$. Then $\frac{\partial v}{\partial t} = \epsilon\frac{\partial v}{\partial \tau}$, so equations (2.5) reduce to

$$\frac{\partial v_j}{\partial \tau} + \mathbf{i}\epsilon^{-1}\lambda_j v_j = P_j(v), \quad 1 \leq j \leq n. \tag{2.8}$$

Let us substitute in (2.8)

$$v_j(\tau) = e^{-\mathbf{i}\epsilon^{-1}\lambda_j\tau}a_j(\tau), \quad 1 \leq j \leq n. \tag{2.9}$$

Then (2.8) becomes

$$\dot{a}_j = e^{\mathbf{i}\epsilon^{-1}\lambda_j\tau}P_j(v), \quad 1 \leq j \leq n. \tag{2.10}$$

Denote $a(\tau) = (a_1(\tau), a_2(\tau), ..., a_n(\tau)) \in \mathbb{C}^n$. For a real vector $w = (w_1, w_2, ..., w_n) \in \mathbb{R}^n$ let $\Phi_w$ be the rotation operator

$$\Phi_w : \mathbb{C}^n \to \mathbb{C}^n, \quad \Phi_w = \mathrm{diag}\{e^{\mathbf{i}w_1}, e^{\mathbf{i}w_2}, ..., e^{\mathbf{i}w_n}\}.$$

It is easy to see that

$$(\Phi_w)^{-1} = \Phi_{-w}, \quad \Phi_{w_1} \circ \Phi_{w_2} = \Phi_{w_1+w_2}, \quad \Phi_0 = \mathrm{id},$$

and that each $\Phi_w$ is a unitary transformation.
Denote by $\Lambda$ the vector $(\lambda_1, \lambda_2, ..., \lambda_n) \in \mathbb{R}^n$. Then (2.9) can be written as $v(\tau) = \Phi_{-\tau\epsilon^{-1}\Lambda}a(\tau)$, or $a(\tau) = \Phi_{\tau\epsilon^{-1}\Lambda}v(\tau)$. Thus, the system (2.10) reads as

$$\frac{\partial a}{\partial \tau} = \Phi_{\tau\epsilon^{-1}\Lambda} \circ P(\Phi_{-\tau\epsilon^{-1}\Lambda}a(\tau)), \quad |\tau| \leq \theta, \tag{2.11}$$

with the initial condition

$$a(0) = v_0, \quad |v_0| =: R. \tag{2.12}$$

Note that

$$|a_j(\tau)| = |v_j(\tau)|, \quad \forall\tau, \quad 1 \leq j \leq n. \tag{2.13}$$

## 3. Averaging of vector-fields

We recall that a diffeomorphism $G : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ transforms a vector-field $W$ on $\mathbb{R}^{2n}$ to the vector-field $(G_*W)(v) = dG(u)(W(u))$, $u = G^{-1}(v)$, see in [1]. Accordingly, a linear isomorphism $\Phi_{\Lambda t}$, $t \in R$, transforms the vector-field $P$ to

$$((\Phi_{\Lambda t})_*P)(v) = \Phi_{\Lambda t} \circ P(\Phi_{-\Lambda t}v).$$

Our goal in this section is to study the averaging in $t$ of the vector-field above.

For a continuous vector-field $P$ on $\mathbb{C}^n$ and a vector $\Lambda \in (\mathbb{R} \setminus \{0\})^n$, we denote

$$\langle\langle P\rangle\rangle(a) = \lim_{T\to\pm\infty}\frac{1}{|T|}\int_0^T \Phi_{\Lambda t} \circ P(\Phi_{-\Lambda t}a)\mathrm{d}t, \tag{3.1}$$

if the limit exists (for $T < 0$ we understand $\int_0^T \ldots \mathrm{d}t$ as the integral $\int_T^0 \ldots \mathrm{d}t$). Denote

$$y^t(a) = \Phi_{\Lambda t} \circ P(\Phi_{-\Lambda t}a), \tag{3.2}$$

and for $T \neq 0$ set

$$\langle\langle P\rangle\rangle^T(a) = \frac{1}{|T|}\int_0^T y^t(a)\mathrm{d}t. \tag{3.3}$$

Then

$$\langle\langle P\rangle\rangle(a) = \lim_{T\to\pm\infty}\langle\langle P\rangle\rangle^T(a).$$

The vector-field $\langle\langle P\rangle\rangle$ is called the averaging of a field $P$ in the direction of a vector $\Lambda$, and $\langle\langle P\rangle\rangle^T(a)$ is called the partial averaging. The latter always exists. Our goal in this section is to prove that the former also exists, if the mapping $P$ is locally Lipschitz. To indicate the dependence of the two introduced objects on $\Lambda$, sometimes we will write them as $\langle\langle\cdot\rangle\rangle_\Lambda$ and $\langle\langle\cdot\rangle\rangle_\Lambda^T$. We recall that being written in the special basis the matrix of operator $A$ takes the form (2.1), and that after introducing in $\mathbb{R}^{2n}$ the complex structure (2.3) the matrix becomes $\mathrm{diag}\{\mathbf{i}\lambda_j\}$. Since $\Phi_{\Lambda t} = \exp(\mathrm{diag}\{\mathbf{i}\lambda_j\}t)$, then the definition of $\langle\langle P\rangle\rangle$ agrees with that in (1.5).

**Lemma 3.1.** *If $P \in Lip_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$, then $\langle\langle P\rangle\rangle^T \in Lip_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$ for any $T \neq 0$.*

*Proof.* If $a \in \bar{B}_R$, then $\Phi_{-\Lambda t}a \in \bar{B}_R$. Since $P \in \mathrm{Lip}_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$, one obtains $|P(\Phi_{-\Lambda t}a)| \leq \mathcal{X}(R)$, for each $t$, and then

$$|y^t(a)| \leq \mathcal{X}(R), \quad \forall t. \tag{3.4}$$

Similarly, for any $a_1, a_2 \in \bar{B}_R$, we have

$$\begin{aligned} |y^t(a_2) - y^t(a_1)| &= |P(\Phi_{-\Lambda t}a_2) - P(\Phi_{-\Lambda t}a_1)| \\ &\leq \mathcal{X}(R)|\Phi_{-\Lambda t}a_2 - \Phi_{-\Lambda t}a_1| \\ &\leq \mathcal{X}(R)|a_2 - a_1|, \quad \forall t. \end{aligned} \tag{3.5}$$

From (3.4) and (3.5), one obtains

$$|\langle\langle P\rangle\rangle^T(a)| \leq \sup_{|t|\leq|T|}|y^t(a)| \leq \mathcal{X}(R),$$

and

$$|\langle\langle P\rangle\rangle^T(a_2) - \langle\langle P\rangle\rangle^T(a_1)| \leq \sup_{|t|\leq|T|}|y^t(a_2) - y^t(a_1)| \leq \mathcal{X}(R)|a_2 - a_1|.$$

Thus, $\langle\langle P\rangle\rangle^T \in \mathrm{Lip}_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$ for $T \neq 0$. □

**Lemma 3.2** (The main lemma of averaging)**.** *For any $\Lambda \in (\mathbb{R}\setminus\{0\})^n$ and $P \in Lip_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$, the limit of (3.1) exists for any $a \in \mathbb{C}^n$, and $\langle\langle P\rangle\rangle \in Lip_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$. If $a \in \bar{B}_R$, then the rate of convergence in (3.1) depends only on $R, \Lambda$ and $P$.*

Before proving the general case of Lemma 3.2, we firstly consider the case when $P$ is a polynomial vector-field. Then,

$$P_j(v) = \sum_{0\leq|\alpha|,|\beta|\leq N} C_j^{\alpha\beta} v^\alpha \bar{v}^\beta, \quad 1 \leq j \leq n, \tag{3.6}$$

and one has

$$\begin{aligned} y_j^t(v) &= e^{\mathbf{i}\lambda_j t}\sum_{0\leq|\alpha|,|\beta|\leq N} C_j^{\alpha\beta}\prod_i(e^{-\mathbf{i}\lambda_i t}v_i)^{\alpha_i}\prod_i(e^{\mathbf{i}\lambda_i t}\bar{v}_i)^{\beta_i} \\ &= \sum_{0\leq|\alpha|,|\beta|\leq N} C_j^{\alpha\beta} e^{\mathbf{i}t(\lambda_j - \Lambda\cdot\alpha + \Lambda\cdot\beta)} v^\alpha \bar{v}^\beta. \end{aligned}$$

It follows that

$$\langle\langle P\rangle\rangle_j^T(v) = \sum_{0\le|\alpha|,|\beta|\le N} C_j^{\alpha\beta}\left(\frac{1}{|T|}\int_0^T e^{\mathbf{i}t(\lambda_j-\Lambda\cdot\alpha+\Lambda\cdot\beta)}\mathrm{d}t\right)v^\alpha\bar{v}^\beta.$$

**Definition 3.3.** A pair $(\alpha,\beta)$ is called $(\Lambda,j)$-resonant if $\lambda_j-\Lambda\cdot\alpha+\Lambda\cdot\beta=0$. The resonant part of the polynomial vector-field (3.6) is another polynomial vector-field $P^{\mathrm{res}}(v)$ such that

$$P_j^{\mathrm{res}}(v) = \sum_{\substack{\text{pair } (\alpha,\beta) \text{ is } (\Lambda,j)\text{-resonant},\\ 0\le|\alpha|,|\beta|\le N}} C_j^{\alpha\beta}v^\alpha\bar{v}^\beta \quad \text{for} \quad 1\le j\le n.$$

**Example 3.4.** For equation (2.7) in Example 2.4, we have $P(v)=v^2\bar{v}+v^3=:P_1(v)+P_2(v)$. The monomial $P_1$ is resonant since $\omega-\omega\cdot 2+\omega=0$. So now $P^{\mathrm{res}}(v)=P_1(v)$.

Note that

$$\begin{aligned}&\frac{1}{|T|}\int_0^T e^{-\mathbf{i}t(\lambda_j-\Lambda\cdot\alpha+\Lambda\cdot\beta)}\mathrm{d}t\\ &\quad=\begin{cases}1, & \text{if } (\alpha,\beta) \text{ is } (\Lambda,j)\text{-resonant},\\ \frac{1}{-\mathbf{i}T(\lambda_j-\Lambda\cdot\alpha+\Lambda\cdot\beta)}(e^{-\mathbf{i}T(\lambda_j-\Lambda\cdot\alpha+\Lambda\cdot\beta)}-1), & \text{otherwise.}\end{cases}\end{aligned}$$

So,

$$\lim_{T\to\pm\infty}\frac{1}{|T|}\int_0^T e^{-\mathbf{i}t(\lambda_j-\Lambda\cdot\alpha+\Lambda\cdot\beta)}\mathrm{d}t=\begin{cases}1, & \text{if } (\alpha,\beta) \text{ is } (\Lambda,j)\text{-resonant},\\ 0, & \text{otherwise.}\end{cases} \tag{3.7}$$

Thus, we have

$$\langle\langle P\rangle\rangle_j^T(v)\xrightarrow[T\to\pm\infty]{} P_j^{\mathrm{res}}(v). \tag{3.8}$$

Therefore, in the polynomial case the limit in (3.1) exists.

**Lemma 3.5.** *If $P(v)\in Lip_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$ is a polynomial vector-field of the form (3.6), then the limit $\langle\langle P\rangle\rangle$ in (3.1) exists for all $a\in\mathbb{C}^n$, equals to the resonant part $P^{res}$ of $P$ and satisfies $\langle\langle P\rangle\rangle\in Lip_{\mathcal{X}}(\mathbb{C}^n,\mathbb{C}^n)$. Moreover, if $a\in\bar{B}_R$, then the rate of convergence (3.1) depends only on $R,\Lambda$ and $P$.*

*Proof.* The existence of the limit $\langle\langle P\rangle\rangle$ already is proved, and its Lischitz continuity easily follows Lemma 3.1. The second assertion holds since the rate of convergence in (3.7) and (3.8) depends only on the indicated quantities. We omit the details. □

Now we begin to prove Lemma 3.2.

*Proof.* To show that the limit exists we have to verify that $\frac{1}{|T|}\int_0^T e^{\mathbf{i}\lambda_j t}\circ P_j(\Phi_{-\Lambda t}a)\mathrm{d}t$ converges to a limit as $T\to\pm\infty$. It suffices to show that for any $\xi>0$, there exists $T'=T'(\xi,R,\Lambda,P)>0$, such that

$$|\langle\langle P\rangle\rangle_j^{T_1}(a)-\langle\langle P\rangle\rangle_j^{T_2}(a)|\le\xi, \quad \text{if } |T_1|,|T_2|\ge T'. \tag{3.9}$$

For any $j$ consider the restriction of $P_j$ to the closed ball $\bar{B}_R$. By the Stone-Weierstrass theorem, there exist $N$ and a polynomial $P_j^N(a)$ of degree $N$, depending only on $R$ and $P$, such that

$$|P_j(a)-P_j^N(a)|\le\frac{\xi}{4}, \quad \forall a\in\bar{B}_R.$$

We have got a polynomial vector field $P^N$, for which the assertions of Lemma 3.2 already are proved.

Since $\Phi_{-\Lambda t}a \in \bar{B}_R$ for any $t$, then

$$|y_j^t(a;P_j) - y_j^t(a;P_j^N)| \le \frac{\xi}{4}, \quad \forall t,\ \forall a \in \bar{B}_R.$$

So,

$$|\langle\langle P\rangle\rangle_j^T(a) - \langle\langle P^N\rangle\rangle_j^T(a)| \le \frac{\xi}{4}, \quad \forall T \ne 0. \tag{3.10}$$

By Lemma 3.5, there exists $T' = T'_N > 0$ such that

$$|\langle\langle P^N\rangle\rangle_j^{\tilde{T}}(a) - \langle\langle P^N\rangle\rangle_j(a)| \le \frac{\xi}{4}, \quad \forall |\tilde{T}| \ge T'. \tag{3.11}$$

From (3.10) and (3.11),

$$|\langle\langle P\rangle\rangle_j^{T_1}(a) - \langle\langle P^N\rangle\rangle_j(a)| \le \frac{\xi}{2}, \quad \forall |T_1| \ge T'.$$

The same is true for $T_2$. Therefore (3.9) follows, and the convergence (3.1) is established. The inclusion $\langle\langle P\rangle\rangle \in \mathrm{Lip}_\chi$ is a consequence of Lemma 3.1, while the last assertion of the lemma directly follows from the proof. The lemma is proved. □

**Example 3.6.** Let $\lambda_j > 0$ for all $j$ and $P_j$ be an anti-holomorphic polynomial $P_j(v) = \sum_{|\beta|\le N} C_j^\beta \bar{v}^\beta$. Then no pair $(0,\beta)$ is $(\Lambda, j)$-resonant and $\langle\langle P\rangle\rangle = 0$.

### 3.1. Properties of the operator $\langle\langle\cdot\rangle\rangle$.

**Proposition 3.7.** *Let $P$ and $Q$ be locally Lipschitz vector-fields on $\mathbb{C}^n$. Then*

1) *(linearity): $\langle\langle aP + bQ\rangle\rangle = a\langle\langle P\rangle\rangle + b\langle\langle Q\rangle\rangle$ for any $a, b \in \mathbb{R}$.*
2) *If $P = diag(a_1, \dots, a_n)$, $a_j \in \mathbb{C}$, then $\langle\langle P\rangle\rangle = P$.*
3) *The mapping $v \mapsto \langle\langle P\rangle\rangle(v)$ commutes with all operators $\Phi_{\Lambda\theta}$, $\theta \in \mathbb{R}$.*
4) *The mapping $(\mathbb{C}^n \times (\mathbb{R}\setminus\{0\})^n) \ni (v,\Lambda) \mapsto \langle\langle P\rangle\rangle_\Lambda(v)$ is measurable. So for every $v$ the averaging $\langle\langle P\rangle\rangle_\Lambda(v)$ is a measurable function of $\Lambda$.*

*Proof.* Properties 1) and 2) are obvious. Let us prove 3), assuming for definiteness that $T > 0$. We have:

$$\begin{aligned}
\langle\langle P\rangle\rangle(\Phi_{\Lambda\theta}a) &= \lim_{T\to+\infty} \frac{1}{T}\int_0^T \Phi_{\Lambda t} \circ P(\Phi_{-\Lambda t}\Phi_{\Lambda\theta}a)\mathrm{d}t \\
&= \Phi_{\Lambda\theta} \lim_{T\to+\infty} \frac{1}{T}\int_0^T \Phi_{\Lambda(t-\theta)} \circ P(\Phi_{-\Lambda(t-\theta)}a)\mathrm{d}t \\
&= \Phi_{\Lambda\theta} \lim_{T\to+\infty} \frac{1}{T}\int_{-\theta}^{T-\theta} \Phi_{\Lambda t'} \circ P(\Phi_{-\Lambda t'}a)\mathrm{d}t' \\
&= \Phi_{\Lambda\theta} \lim_{T\to+\infty} \frac{1}{T}\left(\int_0^T + \int_{-\theta}^0 - \int_{T-\theta}^T\right) \Phi_{\Lambda t'} \circ P(\Phi_{-\Lambda t'}a)\mathrm{d}t' \\
&= \Phi_{\Lambda\theta}\langle\langle P\rangle\rangle(a).
\end{aligned}$$

To prove 4), we note that for $T > 0$ the mappings $(a,\Lambda) \to \langle\langle P\rangle\rangle_\Lambda^T(a)$ are continuous, so measurable. By Lemma 3.2, the mapping in question is a point-wise limit, as $T \to \infty$, of the measurable mappings above; so it also is measurable (see [10], Theorem 1.14). □

## 4. Averaging for solutions of equation (2.11)

In this section we get our main result, describing the behaviour, as $\epsilon \to 0$, of solutions of equation (2.11) on time-intervals $|\tau| \le$ const, where *const* does not depend on $\epsilon$. In view of (2.11), this also describes the behaviour of the amplitudes $|v_j(\tau)|$ of solutions (2.8), and accordingly, the behaviour of the amplitudes of solutions for (2.5) on long time-intervals $|t| \le$ const $\epsilon^{-1}$.

Let in eq. (2.8) $P \in \mathrm{Lip}_{\mathcal{X}}(\mathbb{C}^n, \mathbb{C}^n)$ for some function $\mathcal{X}$ as in Definition 2.5, and let $v(\tau)$ be its solution such that $v(0) = v_0$. Denote $|v_0| = R$. Then by Lemma 2.7, $v(\tau) \in \bar{B}_{2R}$ for $|\tau| \le \theta = \frac{R}{\mathcal{X}(2R)}$. For $|\tau| \le \theta$ the curve $a^\epsilon(\tau) = \Phi_{\tau\epsilon^{-1}\Lambda} v(\tau)$ satisfies (2.11), (2.12) and $|a^\epsilon_j(\tau)| = |v_j(\tau)|$ for each $j$. So for $|\tau| \le \theta$ we have:

$$\begin{cases} |a^\epsilon(\tau)| \le 2R, \\ |\frac{\partial a^\epsilon}{\partial \tau}| \le |P(\Phi_{-\tau\epsilon^{-1}\Lambda} a(\tau))| \le \mathcal{X}(2R). \end{cases} \tag{4.1}$$

Consider the collection of curves $a^\epsilon$ (the solutions of equation (2.11)),

$$a^\epsilon \in C([-\theta, \theta], \mathbb{C}^n), \quad \epsilon \in (0, 1].$$

By (4.1) and the Arzela-Ascoli theorem, the family $\{a^\epsilon, 0 < \epsilon \le 1\}$ is precompact in $C([-\theta, \theta], \mathbb{C}^n)$. So there exists a sequence $\epsilon_j \to 0$, such that

$$a^{\epsilon_j} \to a^0 \text{ in } C([-\theta, \theta], \mathbb{C}^n), \qquad \text{as } \epsilon_j \to 0, \tag{4.2}$$

for some curve $a^0 \in C([-\theta, \theta], \mathbb{C}^n)$. By (4.1),

$$|a^\epsilon(\tau_1) - a^\epsilon(\tau_2)| \le \mathcal{X}(2R)|\tau_1 - \tau_2|.$$

Passing in this relation to the limit as $\epsilon_j \to 0$, we obtain

$$|a^0(\tau_1) - a^0(\tau_2)| \le \mathcal{X}(2R)|\tau_1 - \tau_2|, \qquad \forall \tau_1, \tau_2 \in [-\theta, \theta]. \tag{4.3}$$

Now we address the following problem: does the limit $a^0$ depend on $\{\epsilon_j\}$? If it does not, then how to describe it?

A solution $a^\epsilon(\tau)$ of (2.11) satisfies the relation

$$a^\epsilon(\tau) = v_0 + \int_0^\tau \Phi_{s\epsilon^{-1}\Lambda} \circ P(\Phi_{-s\epsilon^{-1}\Lambda} a^\epsilon(s)) \mathrm{d}s, \qquad \forall |\tau| \le \theta, \tag{4.4}$$

and the estimates (4.1). From Lemma 3.2,

$$\frac{1}{|T|} \int_0^T \Phi_{t\epsilon^{-1}\Lambda} \circ P(\Phi_{-t\epsilon^{-1}\Lambda} a^\epsilon(t)) \mathrm{d}t = \langle\langle P \rangle\rangle(a) + o(1), \quad \text{as } T \to \pm\infty, \tag{4.5}$$

where $o(1)$ does not depend on $v_0$ if $v_0 \in \bar{B}_R$.

Consider the following effective equation

$$a(\tau) = v_0 + \int_0^\tau \langle\langle P \rangle\rangle(a(s)) \mathrm{d}s, \tag{4.6}$$

that is

$$\frac{\partial}{\partial \tau} a(\tau) = \langle\langle P \rangle\rangle(a(\tau)), \qquad a(0) = v_0.$$

Since $\langle\langle P \rangle\rangle$ is locally Lipschitz, then a solution for (4.6) is unique and exists at least for small $|\tau|$.

**Lemma 4.1.** *The curve $a^0(\tau)$ is a solution of (4.6) for $|\tau| \le \theta$.*

To prove the lemma we first perform some additional constructions.

Assume for definiteness that $\tau \geq 0$, i.e. $0 \leq \tau \leq \theta$, and consider an intermediate scale $L = \sqrt{\epsilon}$; then $\epsilon \ll L \ll 1$. Denote $N = [\frac{\theta}{L}]$. Let $b_j = jL, 0 \leq j \leq N$, $b_{N+1} = \theta$ and $\triangle_j = [b_{j-1}, b_j], 1 \leq j \leq N+1$. Then $|\triangle_1| = \cdots = |\triangle_N| = L$, $0 \leq |\triangle_{N+1}| < L$.

Let the curves $y^t_\epsilon(a), t \in \mathbb{R}$, be defined as in (3.2) with $a = a^\epsilon$.

**Lemma 4.2.** *For any $0 \leq |\tau| \leq \theta$, denote $I(\tau) = \int_0^\tau G(a^\epsilon(s), s\epsilon^{-1})ds$, where $G(a^\epsilon(s), s\epsilon^{-1}) = y^{s\epsilon^{-1}}(a^\epsilon(s)) - \langle\langle P\rangle\rangle(a^\epsilon(s))$. Then uniformly in $\tau \in [0, \theta]$ we have $|I(\tau)| \leq \kappa(\epsilon)$, where $\kappa(\epsilon) \to 0$ as $\epsilon \to 0$, does not depend on $v_0$ if $|v_0| \leq R$.*

*Proof.* Denote

$$I_j(\tau) = \int_{\triangle_j} G(a^\epsilon(s), s\epsilon^{-1})\mathrm{d}s, \quad 1 \leq j \leq N+1.$$

Then $|I| \leq \sum_{j=1}^{N+1} |I_j|$. The term $I_{N+1}$ is trivially small. Now consider $I_j$ with $1 \leq j \leq N$. We have

$$|I_j| \leq \left| \int_{b_{j-1}}^{b_j} (G(a^\epsilon(s), s\epsilon^{-1}) - G(a^\epsilon(b_{j-1}), s\epsilon^{-1}))\mathrm{d}s \right| + \left| \int_{b_{j-1}}^{b_j} G(a^\epsilon(b_{j-1}), s\epsilon^{-1})\mathrm{d}s \right| =: I_j^1 + I_j^2.$$

Consider the term $I_j^1$. Since $|s - b_{j-1}| \leq L$, then by (4.1), we have

$$|a^\epsilon(s) - a^\epsilon(b_{j-1})| \leq L\mathcal{X}(2R).$$

As for any $t$, $y^t$ and $\langle\langle P\rangle\rangle$ both belong to $\mathrm{Lip}_\chi$, then

$$I_j^1 \leq 2L\mathcal{X}(2R) \cdot L\mathcal{X}(2R) = 2L^2(\mathcal{X}(2R))^2.$$

Now consider the term $I_j^2$. We have

$$\begin{aligned}
I_j^2 &= \left| \int_{b_{j-1}}^{b_j} y^{s\epsilon^{-1}}(a^\epsilon(b_{j-1}))\mathrm{d}s - \int_{b_{j-1}}^{b_j} \langle\langle P\rangle\rangle(a^\epsilon(b_{j-1}))\mathrm{d}s \right| \\
&= \left| \int_{b_{j-1}}^{b_{j-1}+L} \Phi_{\Lambda\epsilon^{-1}\tau} \circ P(\Phi_{-\Lambda\epsilon^{-1}\tau} a^\epsilon(b_{j-1}))\mathrm{d}\tau - L\langle\langle P\rangle\rangle(a^\epsilon(b_{j-1})) \right| \\
&= \left| \int_0^L \Phi_{\Lambda\epsilon^{-1}b_{j-1}} \Phi_{\Lambda\epsilon^{-1}\tilde{\tau}} \circ P(\Phi_{-\Lambda\epsilon^{-1}\tilde{\tau}} \Phi_{-\Lambda\epsilon^{-1}b_{j-1}} a^\epsilon(b_{j-1}))\mathrm{d}\tilde{\tau} - L\langle\langle P\rangle\rangle(a^\epsilon(b_{j-1})) \right| \\
&= \left| \Phi_{\Lambda\epsilon^{-1}b_{j-1}} \int_0^L \Phi_{\Lambda\epsilon^{-1}\tilde{\tau}} \circ P(\Phi_{-\Lambda\epsilon^{-1}\tilde{\tau}} z)\mathrm{d}\tilde{\tau} - L\langle\langle P\rangle\rangle(a^\epsilon(b_{j-1})) \right|,
\end{aligned}$$

where $z := \Phi_{-\Lambda\epsilon^{-1}b_{j-1}} a^\epsilon(b_{j-1}) \in \bar{B}_{2R}$. Making in the last inequality the substitution $\epsilon^{-1}\tilde{\tau} = t$ and noting that $d\tilde{\tau} = \epsilon dt = L^2\, dt$, we obtain:

$$\begin{aligned}
I_j^2 &= \left| \Phi_{\Lambda\epsilon^{-1}b_{j-1}} \frac{L}{L^{-1}} \int_0^{L^{-1}} \Phi_{\Lambda t} \circ P(\Phi_{-\Lambda t} z)\mathrm{d}t - L\langle\langle P\rangle\rangle(a^\epsilon(b_{j-1})) \right| \\
&= \left| L\Phi_{\Lambda\epsilon^{-1}b_{j-1}} \langle\langle P\rangle\rangle^{L^{-1}}(z) - L\langle\langle P\rangle\rangle(a^\epsilon(b_{j-1})) \right|.
\end{aligned}$$

From (4.5), $\langle\langle P\rangle\rangle^{L^{-1}}(z) = \langle\langle P\rangle\rangle(z) + o(1)$ as $\epsilon \to 0$. Since by item 3) of Proposition 3.7

$$\langle\langle P\rangle\rangle(z) = \langle\langle P\rangle\rangle(\Phi_{-\Lambda\epsilon^{-1}b_{j-1}} a^\epsilon(b_{j-1})) = \Phi_{-\Lambda\epsilon^{-1}b_{j-1}} \langle\langle P\rangle\rangle(a^\epsilon(b_{j-1})),$$

and as by Lemma 3.2 the $o(1)$ above does not depend on $z \in \bar{B}_{2R}$, then

$$I_j^2 = |L\langle\langle P\rangle\rangle(a^\epsilon(b_{j-1})) + L \cdot o(1) - L\langle\langle P\rangle\rangle(a^\epsilon(b_{j-1}))| = L \cdot o(1).$$

So

$$|I_j| \le L \cdot o(1) + 2L^2(\mathcal{X}(2R))^2 = L \cdot o(1),$$

and therefore,

$$|I| \le \sum_{j=1}^{N} |I_j| \le NL \cdot o(1) \le \theta \cdot o(1) =: \kappa(\epsilon).$$

The lemma is proved. □

*Proof of Lemma 4.1.* Consider

$$A(\tau) := a^0(\tau) - v_0 - \int_0^\tau \langle\langle P \rangle\rangle(a^0(s))\mathrm{d}s$$

$$= a^0(\tau) - a^{\epsilon_j}(\tau) \tag{4.7}$$

$$+ a^{\epsilon_j}(\tau) - v_0 - \int_0^\tau y(a^{\epsilon_j}(\tau), s\epsilon_j^{-1})\mathrm{d}s \tag{4.8}$$

$$+ \int_0^\tau y(a^{\epsilon_j}(\tau), s\epsilon_j^{-1})\mathrm{d}s - \int_0^\tau \langle\langle P \rangle\rangle(a^{\epsilon_j}(s))\mathrm{d}s \tag{4.9}$$

$$+ \int_0^\tau \langle\langle P \rangle\rangle(a^{\epsilon_j}(s))\mathrm{d}s - \int_0^\tau \langle\langle P \rangle\rangle(a^0(s))\mathrm{d}s. \tag{4.10}$$

The term (4.7) $\to 0$ as $\epsilon_j \to 0$ in view of (4.2), the term (4.8)=0 by (4.4), the term (4.9)$\le \kappa(\epsilon)$ by Lemma 4.2 and (4.10)$\le \tau\mathcal{X}(2R)|a^{\epsilon_j} - a^0| \to 0$ as $\epsilon_j \to 0$ by (4.2). Passing to the limit as $\epsilon_j \to 0$, we see that $A(\tau) \equiv 0$. Therefore, $a^0(\tau)$ is a solution of (4.6) for $0 \le |\tau| \le \theta$. Lemma 4.1 is proved. □

Since a solution of eq. (4.4) is unique, then the convergence (4.2) holds as $\epsilon \to 0$, not only as $\epsilon_j \to 0$. Thus we have proved

**Theorem 4.3.** *Let $a^\epsilon(\tau), |\tau| \le \theta$, be a solution of (4.4). Then*

$$a^\epsilon(\tau) \to a^0(\tau), \qquad \textit{uniformly for } |\tau| \le \theta,$$

*where $a^0(\tau)$ is a solution of (4.6).*

Theorem 4.3 proves the second assertion of Theorem 1.1, whose first assertion was already established in Lemma 3.2.

Since $|v_j^\epsilon(\tau)| \equiv |a_j^\epsilon(\tau)|$, then Theorem 4.3 implies:

**Corollary 4.4.** *The solution $v^\epsilon(t)$ of (2.5) satisfies*

$$\sup_{|t| \le \epsilon^{-1}\theta} \big||v_j^\epsilon(t)| - |a_j^0(\epsilon t)|\big| \to 0 \quad \textit{as } \epsilon \to 0, \ \forall j.$$

**Example 4.5.** By Examples 2.4 and 3.4, the effective equation for the system (2.6), written in the complex variable $v$ and the slow time $\tau$, is

$$\frac{\partial}{\partial \tau} v = v^2 \bar{v}.$$

In the real variables and the fast time the equation reads

$$\dot{x} = \epsilon x(x^2 + y^2), \quad \dot{y} = \epsilon y(x^2 + y^2).$$

For $x_0 = 1, y_0 = 0$ its solution is $x(t) = (1 - 2\epsilon t)^{-1/2}$, $y(t) = 0$. So for $|t| \le \frac{1}{4}\epsilon^{-1}$ the solution of (2.6) with initial data $(1, 0)$ satisfies

$$\left|x^2(t) + y^2(t)\right|^{1/2} = (1 - 2\epsilon t)^{-1/2} + o(1) \qquad \text{as } \epsilon \to 0.$$

## 5. Hamiltonian equations

Let us provide the space $\mathbb{R}^{2n} \sim \mathbb{C}^n$ with the usual symplectic structure, given by the form $\omega_2 = \sum \mathrm{d}x_j \wedge \mathrm{d}y_j$. Then a real-valued Hamiltonian

$$H = h_2(z,\bar z) + \epsilon h(z,\bar z), \quad h_2 = -\frac{1}{2}\sum \lambda_j |z_j|^2, \quad h \in C^1(\mathbb{C}^n),$$

gives rise to the Hamiltonian system

$$\dot z_j = -\mathbf{i}\lambda_j z_j + 2\mathbf{i}\epsilon \frac{\partial h}{\partial \bar z_j}, \qquad 1 \le j \le n,$$

which we rewrite as

$$\frac{\partial z}{\partial \tau} + \mathbf{i}\epsilon^{-1}\mathrm{diag}(\lambda_j) z = 2\mathbf{i}\frac{\partial h}{\partial \bar z} := P(z). \tag{5.1}$$

Assume that $P$ is locally Lipschitz. It means that

$$h \in C^{1,1}_{\mathrm{loc}} = \{u(z) \in C^1(\mathbb{C}^n) : u_z, u_{\bar z} \text{ are locally Lipschitz functions}\}.$$

Then (5.1) is a special case of equation (2.11).

**Lemma 5.1.** *Let $y^t(a)$ be defined as in (3.2). Then*

$$y^t_j(a) = 2\mathbf{i}\frac{\partial}{\partial \bar a_j} h(\Phi_{-\Lambda t} a).$$

*Proof.* Denote $\Phi_{-\Lambda t} a = v$, i.e., $v_j = e^{-\mathbf{i}\lambda_j t} a_j$. Then

$$\begin{aligned} 2\mathbf{i}\frac{\partial h(v(a))}{\partial \bar a_j} &= 2\mathbf{i}\frac{\partial h(v(a))}{\partial \bar v_j}\cdot\frac{\partial \bar v_j}{\partial \bar a_j} = 2\mathbf{i}\frac{\partial h(v(a))}{\partial \bar v_j} e^{\mathbf{i}\lambda_j t} \\ &= 2\mathbf{i} e^{\mathbf{i}\lambda_j t}\frac{\partial h(\Phi_{-\Lambda t} a)}{\partial \bar v_j} = e^{\mathbf{i}\lambda_j t} P_j(\Phi_{-\Lambda t} a) = y^t_j(a), \end{aligned}$$

since $2\mathbf{i}\frac{\partial h}{\partial \bar v_j} = P_j(v)$. □

From Lemma 5.1,

$$\langle\langle P\rangle\rangle^T_j(a) = \frac{2\mathbf{i}}{|T|}\int_0^T \frac{\partial}{\partial \bar a_j} h(\Phi_{-\Lambda t} a)\mathrm{d}t = 2\mathbf{i}\frac{\partial}{\partial \bar a_j}\langle h\rangle^T(a), \tag{5.2}$$

where we denoted $\langle h\rangle^T(a) = \frac{1}{|T|}\int_0^T h(\Phi_{-\Lambda t} a)\mathrm{d}t$.

For the same reason as in Section 3 (see there Lemma 3.2), since $h$ is a locally Lipschitz function, then the limit

$$\lim_{T\to\pm\infty}\langle h\rangle^T(a) =: \langle h\rangle(a) \tag{5.3}$$

exists and is locally Lipschitz (this limit is the averaging of the function $h$ in the direction $\Lambda$). Repeating the proof of Proposition 3.7.3) we get that $\langle h\rangle(a)$ is invariant with respect to the rotations $\Phi_{\Lambda\theta}$:

$$\langle h\rangle(\Phi_{\Lambda\theta} a) = \langle h\rangle(a) \qquad \forall\theta,\ \forall a. \tag{5.4}$$

We claim that $\langle h\rangle$ is a $C^1$-function and

$$2i\frac{\partial}{\partial \bar a_j}\langle h\rangle(a) = \langle\langle P\rangle\rangle_j(a), \tag{5.5}$$

for every $j$. Indeed, by the second assertion of Lemma 2.3 and (5.2)

$$\langle\langle \bar P\rangle\rangle_j^T(a) = -2\mathbf{i}\frac{\partial}{\partial a_j}\langle h\rangle^T(a).$$

From here and (5.2) we conclude that $\langle h\rangle^T$ is a $C^1$ function, and for each $j$ the derivative $(\partial/\partial x_j)\langle h\rangle^T$ is a complex linear combination of $\langle\langle P\rangle\rangle_j^T$ and $\langle\langle \bar P\rangle\rangle_j^T$. So by Lemma 3.1, for any bounded domain $B$ in $\mathbb{C}^n$ functions $(\partial/\partial x_j)\langle h\rangle^T$ are uniformly in $T>0$ bounded and Lipschitz on $B$. Same is true for $(\partial/\partial y_j)\langle h\rangle^T$, for all $j$. Now the uniform in $B$ convergence of $\langle\langle P\rangle\rangle_j^T$ to $\langle\langle P\rangle\rangle_j$ implies the uniform convergence of $(\partial/\partial x_j)\langle h\rangle^T$ and $(\partial/\partial y_j)\langle h\rangle^T$ to limits, for all $j$. In view of the uniform in $B$ convergence of $\langle h^T\rangle$ to $\langle h\rangle$ we get that $\langle h\rangle$ is $C^1$-smooth and satisfies (5.5).

We have proved

**Theorem 5.2.** *If equation* (2.5) *has the hamiltonian form* (5.1)*, where the real Hamiltonian $h$ belongs to the space $C^{1,1}_{loc}$, then the effective equation* (4.6) *also is hamiltonian: the vector field $\langle\langle P\rangle\rangle$ has the hamiltonian form* (5.5)*, where the Hamiltonian $\langle h\rangle \in C^{1,1}_{loc}$ is defined in* (5.3)*.*

**Example 5.3.** If $h$ is a real polynomial,

$$h(a) = \sum_{\alpha,\beta\in\mathbb{Z}^n_+} m_{\alpha\beta}a^\alpha\bar a^\beta, \quad m_{\alpha\beta} = \overline{m_{\alpha\beta}}, \tag{5.6}$$

then $\langle h\rangle(a) = \sum_{\Lambda\cdot\alpha=\Lambda\cdot\beta} m_{\alpha\beta}a^\alpha\bar a^\beta$.

Let us take any $h\in C^{1,1}_{\rm loc}$ and assume that the vector $\Lambda$ is non-resonant, that is $\Lambda\cdot s=0$ for some $s\in\mathbb{Z}^n$ implies that $s=0$. Let us introduce in $\mathbb{C}^n$ the action-angle coordinates $(I,\varphi)$ with $I\in\mathbb{R}^n_+$, $\varphi\in\mathbb{T}^n$, where for a vector $z=(z_1,\dots,z_n)$ with $z_j = r_je^{\mathbf{i}\varphi_j}$ we have $I=(I_1,\dots,I_n), I_j=\frac12 r_j^2$, and $\varphi=(\varphi_1,\dots,\varphi_n)$. Then $\omega_2 = {\rm d}I\wedge{\rm d}\varphi$, and

$$\Phi_{\Lambda t}(I,\varphi) = (I,\varphi+\Lambda t).$$

Since now the curve $t\mapsto\varphi+\Lambda t$ is dense in $\mathbb{T}^n$ for each $\varphi$, then (5.4) implies that $\langle h\rangle$ does not depend on the angles $\varphi$.[2] That is, the Hamiltonian $\langle h\rangle$ is integrable, and the effective equation reads

$$\dot I = 0, \quad \dot\varphi = \nabla_I\langle h\rangle(I).$$

So its solutions $a(\tau)$ are such that $|a_j(\tau)|^2 = {\rm const}$, and Theorem 4.3 implies

**Corollary 5.4.** *If the frequency vector $\Lambda$ is non-resonant and $h\in C^{1,1}_{loc}$, then a solution $z(t)$ of equation* (5.1) *satisfies*

$$\sup_{|t|\le\epsilon^{-1}\theta}\big||z_j(t)|^2-|z_{j0}|^2\big| = o(1) \quad \text{as } \epsilon\to 0,$$

*for a suitable $\theta$ which depends on $|z_0|$.*

[2]For example, if $h$ is the polynomial (5.6), then $\langle h\rangle(a) = \sum m_{\alpha\alpha}a^\alpha\bar a^\alpha$.

If $h \in C^{1,1}_{\rm loc}$, then $2\mathbf{i}\partial_{\bar z} h \in \mathrm{Lip}_{\mathcal{X}}(\mathbb{C}^n, \mathbb{C}^n)$ for a function $\mathcal{X}$ as in Definition 2.5. For $m \in \mathbb{N}$ we write

$$2\mathbf{i}\partial_{\bar z} h(z) = O(z^m),$$

if $\epsilon^{-m} 2\mathbf{i}\partial_{\bar z} h(\epsilon z) \in \mathrm{Lip}_{\mathcal{X}}(\mathbb{C}^n, \mathbb{C}^n)$ for all $0 < \epsilon \le 1$, for a suitable function $\mathcal{X}$, independent from $\epsilon$. For example, if $h$ is a polynomial (5.6) such that the coefficients $m_{\alpha\beta}$ are nonzero only for $|\alpha| + |\beta| \ge M$, then $2\mathbf{i}\partial_{\bar z} h = O(z^{M-1})$. Consider the equation

$$\dot z + \mathbf{i}\,\mathrm{diag}(\lambda_j) z = 2\mathbf{i}\partial_{\bar z} h, \tag{5.7}$$

where $2\mathbf{i}\partial_{\bar z} h \in O(z^m)$. To study its small solutions we substitute $z = \epsilon w$ and get for $w(t)$ equation (5.1) with $\epsilon := \epsilon^{m-1}$. We see that if the frequency vector $\Lambda$ is non-resonant and $z(t)$ is a solution of (5.7) with small initial data $z(0) = \epsilon w_0$, then $\big||z_j(t)|^2 - \epsilon^2 |w_{0j}|^2\big| = o(\epsilon^2)$ for $|t| \le \epsilon^{1-m}\theta$, where $\theta = \theta(|w_0|)$. In [6] a more delicate argument is used to show that if the frequency vector $\Lambda$ satisfies certain diophantine condition and the Hamiltonian $h$ is analytic, then the stability interval is much bigger – it is exponentially long in terms of $\epsilon^{-a}$ for some positive $a$. Our result is significantly weaker, but it only requires that the vector $\Lambda$ is non-resonant and the Hamiltonian vector-field $2\mathbf{i}\partial_{\bar z} h$ is Lipschitz–continuous. We note that the result of [6] generalises to PDEs, e.g. see [3], as well as Theorem 1.1 (and Corollary 5.4), see [7, 8].

## References


[1] V. I. Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations*. Springer-Verlag, New York, second edition, 1988.

[2] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. *Mathematical Aspects of Classical and Celestial Mechanics*. Springer-Verlag, New York, third edition, 2006.

[3] D. Bambusi. An averaging theorem for quasilinear Hamiltonian PDEs. *Ann. Henri Poincaré*, 4(4):685–712, 2003.

[4] N. N. Bogoliubov and Y. A. Mitropolsky. *Asymptotic Methods in the Theory of Non-linear Oscillations*. Gordon and Breach, New York, 1961.

[5] I. M. Gelfand. *Lectures on Linear Algebra*. Dover Publications, Inc., New York, 1989.

[6] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani, and C. Simó. Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three-body problem. *J. Differential Equations*, 77:167–198, 1989.

[7] G. Huang. An averaging theorem for nonlinear Schrödinger equations with small nonlinearities. *Discrete Contin. Dyn. Syst. A*, 34:3555–3574, 2014.

[8] G. Huang, S. Kuksin, and A. Maiocchi. Time-averaging for weakly nonlinear CGL equations with arbitrary potentials. In *Hamiltonian partial differential equations and applications*, volume 75 of *Fields Inst. Commun.*, pages 323–349. Fields Inst. Res. Math. Sci., Toronto, ON, 2015.

[9] R. Khasminski. On the avaraging principle for Ito stochastic differential equations. *Kybernetika*, 4:260–279, 1968. (in Russian).

[10] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, 1987.



School of Mathematical Sciences, Fudan University, Shanghai 200433, P. R. China
*E-mail address*: wwjian16@fudan.edu.cn

Université Paris-Diderot (Paris 7), UFR de Mathématiques - Batiment Sophie Germain, 5 rue Thomas Mann, 75205 Paris, France & School of Mathematics, Shandong University, Jinan, PRC & Saint Petersburg State University, Universitetskaya nab., St. Petersburg, Russia
*E-mail address*: Sergei.Kuksin@imj-prg.fr

School of Mathematical Sciences, Fudan University, Shanghai 200433, P. R. China
*E-mail address*: 14110840003@fudan.edu.cn