Journal of Applied Analysis and Computation
Volume 14, Number 2, April 2024, 1097–1110
Website:http://www.jaac-online.com
DOI:10.11948/20230355

AVERAGING PRINCIPLE FOR NONLINEAR DIFFERENTIAL SYSTEMS WITH JORDAN BLOCKS

Shuyuan Xiao[1] and Zhicheng Tong[2,†]

Abstract This paper studies a perturbed differential system

$$\frac{\partial v}{\partial t} = Av + \varepsilon H(v),\ v(0) = v_0,\ \varepsilon \in (0, 1],$$

where A is a linear operator having purely imaginary eigenvalues with Jordan blocks, and H is an analytic perturbation satisfying $H(v) = \mathcal{O}(|v|^2)$ as $|v| \to 0$. Such a case cannot be dealt with straightforwardly by the averaging principle due to the difficulties presenting by A. To this end, by employing the Poincaré normal form with nilpotent term for nonlinear quasiperiodic system to simplify the above differential system, we extend the classical Krylov-Bogoliubov averaging method to nonlinear systems admitting Jordan blocks.

Keywords Krylov-Bogoliubov averaging method, Jordan blocks, Poincaré normal form, nilpotent term.

MSC(2010) 34C29.

1. Introduction

Averaging theory is powerful in the study of dynamical systems. The concept of average appeared in the perturbation theory due to Clairaut, Laplace and Lagrange in the 18th century. At the end of the 19th century, it was mainly used in the field of celestial mechanics. Around 1920, van der Pol extended the use of averaging method to equations that appeared in electronic circuit theory. In 1928, Fatou first formalized the averaging theory of smooth differential systems in [1]. Krylov and Bogoliubov proposed a strict averaging method for nonlinear oscillations in [11], known as the Krylov-Bogoliubov averaging method. In 1958, Sanders and Verhulst gave a review of the averaging theorem in [13]. In 1961, Stratonovich [14] proposed the random averaging method based on physical considerations, which was later mathematically proved by Khasminskii [9] in 1964. Since then, there has been widespread research interest in random averaging in the fields of both mathematics and mechanical engineering. Inspired by the classical averaging principle, ones also established some random versions of the Bogoliubov averaging principle, see Vrkoč [15], Kolomiets et al [10], N'Goran and N'zi [12], Xu et al [18], Gao [2, 3] and the references therein for instance. Recently, Gao and Li studied the averaging

†The corresponding author.
[1]School of Mathematics and Statistics, Center for Mathematics and Interdisciplinary Sciences, Northeast Normal University, Changchun 130024, China
[2]School of Mathematics, Jilin University, Changchun 130012, China
Email: xiaosy716@nenu.edu.cn(S. Xiao), tongzc20@mails.jlu.edu.cn(Z. Tong)

principle of the cubic nonlinear Schrödinger equation with fast oscillation potential and fast oscillation force on finite but large time interval and the entire time axis in [4]; Xing, Yang and Li gave the averaging method of higher order perturbed systems in [17]. In [8], Jian et al established the Krylov-Bogoliubov averaging theorem concerning with Lipschitz perturbations of linear systems. However, it is difficult to touch averaging theory when considering the unperturbed part with Jordan blocks due to nilpotent terms, and it seems that there does not exist any result on this aspect. Thereby, this paper is devoted to present a Krylov-Bogoliubov type theorem under this setting.

To be more precise, let us consider a perturbed differential system

$$\frac{\partial v}{\partial t} = Av + \varepsilon H(v),\ v(0) = v_0,\ \varepsilon \in (0, 1], \tag{1.1}$$

on $\mathbb{R}^{2n}$ with $n \in \mathbb{N}^+$, where A is a linear operator admitting purely imaginary eigenvalues with Jordan blocks, and H is analytic with $H(v) = \mathcal{O}(|v|^2)$ as $|v| \to 0$. One can assume that A has $2n$ purely imaginary eigenvalues without loss of generality. Then these eigenvalues are in pairs $\pm \mathbf{i}\lambda_j$, where $0 \neq \lambda_j \in \mathbb{R}$, and λ_i may equal to λ_j for some $i \neq j$. We aim to touch the asymptotic behaviors for solutions of (1.1). In view of the difficulties brought by Jordan blocks, we shall develop a new approach to deal with (1.1) via the averaging method in this paper, namely applying the Poincaré normal form that contains the nilpotent term for the perturbed system.

It should be mentioned that there are few results on the nilpotent type normal forms of differential systems with multiple eigenvalues for linear operators, let along involving purely imaginary eigenvalues. For instance, Zung [19] studied the Birkhoff normal form of Hamiltonian systems admitting nilpotent terms in 2005. Very recently, Xiao and Li investigated the nilpotent Poincaré normal forms for general nonlinear quasiperiodic systems in [16]. In what follows, we shall employ the latter to simplify the original perturbed differential system (1.1).

For the sake of statements, let us denote by $\langle \cdot, \cdot \rangle$ the real scalar product in the complex notation:

$$\langle z, z' \rangle := \mathrm{Re} \sum_j z_j \overline{z_j'} := \mathrm{Re}(z \cdot \overline{z'}) \text{ for } z, z' \in \mathbb{C}^n,$$

and define the set

$$\mathbb{N}_+^{2n} := \{\alpha = (\alpha_1, \ldots, \alpha_{2n}) : \alpha_j \in \mathbb{N}^+,\ j = 1, \ldots, 2n\}$$

as usual, where $|\alpha| := |\alpha_1| + \cdots + |\alpha_{2n}|$ for $\alpha \in \mathbb{N}_+^{2n}$. We also define $u^\beta := \prod_{j=1}^n u_j^{\beta_j}$ and $\bar{u}^\beta := \prod_{j=1}^n \bar{u}_j^{\beta_j}$ for $u \in \mathbb{C}^n$ and $\beta \in \mathbb{N}_+^n$. The previous work [16] illustrates that if

$$|\langle \alpha, \tilde{\Lambda} \rangle \pm \lambda_j| \neq 0,\ \ 1 \leq j \leq n, \tag{1.2}$$

where $\alpha \in \mathbb{N}_+^{2n}$, $\tilde{\Lambda} = (\lambda_1, -\lambda_1, \ldots, \lambda_n, -\lambda_n)$ and $2 \leq |\alpha| \leq l$, then system (1.1) we are aiming at can be changed into

$$\frac{\partial u}{\partial t} = (B + \eta \tilde{B})u + \varepsilon C^{-1} P(Cu) + \varepsilon^2 g(u, \varepsilon),\ u(0) = C^{-1} v_0 = u_0 \tag{1.3}$$

through an analytic transformation, where $0 < \eta = \varepsilon\varpi \leq 1$ is a constant, $1 + (\eta t)^2 + \cdots + \frac{(\eta t)^{2(i_{\max}-1)}}{((i_{\max}-1)!)^2} \leq 2$ (here the range of t will be specified later), $\varepsilon \in (0,1]$, $\varpi > 0$, P and g are analytic, C^{-1} is the inverse of $C, P = (P^1, \ldots, P^{2n}), g = (g^1, \ldots, g^{2n}), P^j = o(|Cu|^l), g = o(|u|^l), 1 \leq j \leq 2n, l \geq 2$, and the operators read

$$B = \begin{pmatrix} B_1 & & & & & \\ & -B_1 & & & & \\ & & \ddots & & & \\ & & & B_s & \\ & & & & -B_s \end{pmatrix}, \ B_c = \begin{pmatrix} \mathbf{i}\lambda_c & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \mathbf{i}\lambda_c \end{pmatrix}_{i_c \times i_c},$$

$$\tilde{B} = \begin{pmatrix} \tilde{B}_1 & & & & \\ & \tilde{B}_1 & & & \\ & & \ddots & & \\ & & & \tilde{B}_s & \\ & & & & \tilde{B}_s \end{pmatrix}, \ \tilde{B}_c = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{i_c \times i_c},$$

$$C = \begin{pmatrix} C_1 & & & & \\ & C_1 & & & \\ & & \ddots & & \\ & & & C_s & \\ & & & & C_s \end{pmatrix}, \ C_c = \begin{pmatrix} 1 & & & \\ & \eta & & \\ & & \ddots & \\ & & & \eta^{i_c-1} \end{pmatrix}_{i_c \times i_c},$$

and

$$\begin{aligned} C^{-1}P(Cu) = (&\eta^{\varsigma\alpha l}P^1 u, \ldots, \eta^{\varsigma\alpha l - i_1 + 1}P^{i_1}u, \eta^{\varsigma\alpha l}P^{i_1+1}u, \ldots, \eta^{\varsigma\alpha l - i_1 + 1}P^{2i_1}u, \\ &\ldots, \eta^{\varsigma\alpha l}P^{2(i_1+\cdots+i_{s-1})+1}u, \ldots, \eta^{\varsigma\alpha l - i_s + 1}P^{2(i_1+\cdots+i_{s-1})+i_s}u, \\ &\eta^{\varsigma\alpha l}P^{2(i_1+\cdots+i_{s-1})+i_s+1}u, \ldots, \eta^{\varsigma\alpha l - i_s + 1}P^{2(i_1+\cdots+i_s)}u)^T \\ := &\tilde{P}(u) + \varepsilon\tilde{\tilde{P}}(u,\varepsilon), \end{aligned}$$

with

$$\begin{aligned} &\varsigma = (0, \ldots, i_1 - 1, 0, \ldots, i_1 - 1, \ldots, 0, \ldots, i_s - 1, 0, \ldots, i_s - 1), \\ &l \geq 2, \quad i_{\max} = \max\{i_1, \ldots, i_s\}, \quad \varsigma\alpha l - i_{\max} + 1 \geq 0. \end{aligned}$$

The degrees of $B_c, \tilde{B}_c$ and $C_c (c = 1, \ldots, s)$ are i_c, and $i_1 + \cdots + i_s = n$. As a consequence, we arrive at the following perturbed differential system in the Poincaré normal form:

$$\frac{\partial u}{\partial t} = (B + \eta\tilde{B})u + \varepsilon\tilde{P}(u) + \varepsilon^2\tilde{g}(u,\varepsilon), \ u(0) = u_0, \tag{1.4}$$

where $\tilde{P}$ and $\tilde{g}$ are analytic, $\tilde{P}=o(|u|^l)$, $\tilde{g}=o(|u|^l)$ and $\alpha\in\mathbb{N}_+^{2n}$. We shall emphasize that the Poincaré normal form of the original system (1.1) is somewhat difficult to deal with, because the operator $\tilde{B}$ in (1.4) is a *nilpotent term*, and therefore the classical averaging theory seems unable to work.

We wish to study the asymptotic behaviors of solutions of (1.4) on a time interval of length ε^{-1} with $0<\varepsilon\ll 1$. Let us define the interaction representation variable $\phi(t)=e^{-t(B+\eta\tilde{B})}u(t)$. Using system (1.4), we obtain for $\phi=(\phi_1,\dots,\phi_{2n})$ the system of equation

$$\frac{\partial\phi}{\partial t}=\varepsilon e^{-t(B+\eta\tilde{B})}\tilde{P}(e^{t(B+\eta\tilde{B})}\phi)+\varepsilon^2 e^{-t(B+\eta\tilde{B})}\tilde{g}(e^{t(B+\eta\tilde{B})}\phi,\varepsilon),\ \phi(0)=u_0. \tag{1.5}$$

We can prove that the limit

$$Q^o(\phi)=\lim_{T\to\pm\infty}\frac{1}{T}\int_0^T e^{-\upsilon(B+\eta\tilde{B})}\tilde{P}(e^{\upsilon(B+\eta\tilde{B})}\phi)d\upsilon \tag{1.6}$$

exists for any $\phi\in\mathbb{R}^{2n}$ and is locally Lipschitz continuous with respect to ϕ, where the vector field $\tilde{P}$ is locally Lipschitz continuous. We will present the detailed proof in Section 3. Our Krylov-Bogolyubov type averaging theorem under the above setting is given below.

Theorem 1.1. *There exists some $b=b(|u_0|,\varpi)>0$ such that for all $|t|\le\varepsilon^{-1}b$, $0<\varepsilon\ll 1$, a solution $\phi^\varepsilon(t)$ of system (1.5) is $o(1)-$close to the solution of the system*

$$\frac{\partial\phi^0}{\partial t}=\varepsilon Q^o(\phi^0),\ \phi^0(0)=u_0.$$

Let us define the variable $\tilde{\phi}(t)=C\phi(t)$ with the operator C given before. Then by system (1.5), we have the following system of equation

$$\begin{aligned}\frac{\partial\tilde{\phi}}{\partial t}=&\varepsilon Ce^{-t(B+\eta\tilde{B})}\tilde{P}(e^{t(B+\eta\tilde{B})}C^{-1}\tilde{\phi})\\&+\varepsilon^2Ce^{-t(B+\eta\tilde{B})}\tilde{g}(e^{t(B+\eta\tilde{B})}C^{-1}\tilde{\phi},\varepsilon),\ \tilde{\phi}(0)=v_0.\end{aligned} \tag{1.7}$$

Theorem 1.2. *There exists some $b=b(|v_0|,\varpi)>0$ such that for all $|t|\le\varepsilon^{-1}b$, $0<\varepsilon\ll 1$, a solution $\tilde{\phi}^\varepsilon(t)$ of system (1.7) is $o(1)-$close to the solution of the system*

$$\frac{\partial\tilde{\phi}^0}{\partial t}=\varepsilon CQ^o(C^{-1}\tilde{\phi}^0),\ \tilde{\phi}^0(0)=v_0.$$

The idea of the normal form theory is to find a transformation which changes the original system into the simplest one. In this paper, by using the Poincaré normal form for nonlinear quasiperiodic system and basing on the works of Huang et al in [6,7] and Jian et al in [8], we establish the Krylov-Bogolyubov averaging theorem of the nonlinear system with Jordan blocks. Significantly, A is the operator with purely imaginary eigenvalues and the nonlinear system satisfies the nonresonant conditions (1.2). In Section 2, we provide the corresponding matrix forms of linear operators under different bases, some basic concepts and complex variable systems. Averaging of vector fields is shown in Section 3. Finally, Section 4 presents the proofs of Theorems 1.1 and 1.2.

2. Preliminaries

In this section, let us recall some notations.

For $\lambda_1,\dots,\lambda_n\in\mathbb{R}$ and real vectors $w=(w_1,\dots,w_n),\tilde{w}=(\tilde{w}_1,\dots,\tilde{w}_n)\in\mathbb{R}^n$, set

$$\begin{aligned}&\mathrm{diag}\{\mathbf{i}\lambda_j\}:\mathbb{C}^n\to\mathbb{C}^n,\\&(z_1,\dots,z_n)\mapsto(\mathbf{i}\lambda_1z_1,\dots,\mathbf{i}\lambda_nz_n),\\&\text{sub-diag}\{\eta\}:\mathbb{C}^n\to\mathbb{C}^n,\\&(z_1,\dots,z_n)\mapsto(\eta z_2,\dots,\eta z_{i_1},0,\dots,\eta z_{i_1+\cdots+i_{s-1}+2},\dots,\eta z_{i_1+\cdots+i_s},0),\end{aligned}$$

and

$$\Phi_w z=\ \mathrm{diag}\{e^{\mathbf{i}w_1},\dots,e^{\mathbf{i}w_n}\}z=(e^{\mathbf{i}w_1}z_1,\dots,e^{\mathbf{i}w_n}z_n)\text{ for }z\in\mathbb{C}^n,$$

$$\Psi_{\tilde{w}}=\begin{pmatrix}\Psi_1&&\\&\ddots&\\&&\Psi_s\end{pmatrix},\quad\Psi_c=\begin{pmatrix}1&\tilde{w}&\cdots&\frac{\tilde{w}^{i_c-1}}{(i_c-1)!}\\&\ddots&\ddots&\vdots\\&&\ddots&\tilde{w}\\&&&1\end{pmatrix}_{i_c\times i_c},\tag{2.1}$$

$$(\Phi\circ\Psi)_{w,\tilde{w}}:=\ \Phi_w\circ\Psi_{\tilde{w}}.$$

Obviously, $\mathrm{diag}\{\mathbf{i}\lambda_j\},\text{sub-diag}\{\eta\},\Phi_w,\Psi_{\tilde{w}}$ and $(\Phi\circ\Psi)_{w,\tilde{w}}$ are all linear operators. According to [5,8], the linear operators B and $\tilde{B}$ satisfy the following conditions:

(i) $\mathrm{Ker}B=\{0\}$;

(ii) in $\mathbb{R}^{2n}$ there is a basis $\{e_1^+,e_1^-,\dots,e_n^+,e_n^-\}$ such that in the corresponding coordinates $\{x_1,y_1,\dots,x_n,y_n\}$, the matrix of the linear operator B has the form

$$\left(\begin{array}{cc|c|cc}0&-\lambda_1&&&\\\lambda_1&0&&&\\\hline&&\ddots&&\\\hline&&&0&-\lambda_n\\&&&\lambda_n&0\end{array}\right),\tag{2.2}$$

and $\tilde{B}$ has the form

$$\begin{pmatrix}\tilde{B}'_1&&&\\&\tilde{B}'_2&&\\&&\ddots&\\&&&\tilde{B}'_s\end{pmatrix},\quad\tilde{B}'_c=\begin{pmatrix}0&0&1&&\\&\ddots&\ddots&\ddots&\\&&\ddots&\ddots&1\\&&&\ddots&0\\&&&&0\end{pmatrix}_{2i_c\times 2i_c},\tag{2.3}$$

where $\lambda_{i_1+\cdots+i_{c-1}+1}=\cdots=\lambda_{i_1+\cdots+i_c}=\lambda_c,1\le c\le s,i_1+\cdots+i_s=n,i_0=0.$

Thus, by the perturbed system (1.4), the unperturbed linear system reads

$$\begin{aligned}
&\dot{x}_1 = -\lambda_1 y_1 + \eta x_2,\\
&\dot{y}_1 = \lambda_1 x_1 + \eta y_2,\\
&\vdots\\
&\dot{x}_{i_1-1} = -\lambda_{i_1-1} y_{i_1-1} + \eta x_{i_1},\\
&\dot{y}_{i_1-1} = \lambda_{i_1-1} x_{i_1-1} + \eta y_{i_1},\\
&\dot{x}_{i_1} = -\lambda_{i_1} y_{i_1},\\
&\dot{y}_{i_1} = \lambda_{i_1} x_{i_1},\\
&\vdots\\
&\dot{x}_{i_1+\cdots+i_{s-1}+1} = -\lambda_{i_1+\cdots+i_{s-1}+1} y_{i_1+\cdots+i_{s-1}+1} + \eta x_{i_1+\cdots+i_{s-1}+2},\\
&\dot{y}_{i_1+\cdots+i_{s-1}+1} = \lambda_{i_1+\cdots+i_{s-1}+1} x_{i_1+\cdots+i_{s-1}+1} + \eta y_{i_1+\cdots+i_{s-1}+2},\\
&\vdots\\
&\dot{x}_{i_1+\cdots+i_s-1} = -\lambda_{i_1+\cdots+i_s-1} y_{i_1+\cdots+i_s-1} + \eta x_{i_1+\cdots+i_s},\\
&\dot{y}_{i_1+\cdots+i_s-1} = \lambda_{i_1+\cdots+i_s-1} x_{i_1+\cdots+i_s-1} + \eta y_{i_1+\cdots+i_s},\\
&\dot{x}_{i_1+\cdots+i_s} = -\lambda_{i_1+\cdots+i_s} y_{i_1+\cdots+i_s},\\
&\dot{y}_{i_1+\cdots+i_s} = \lambda_{i_1+\cdots+i_s} x_{i_1+\cdots+i_s}.
\end{aligned} \tag{2.4}$$

Definition 2.1. Let $\mathfrak{Z} : \mathbb{R}^+ \to \mathbb{R}^+$ be a non-decreasing continuous function and let $f : \mathbb{C}^n \to \mathbb{C}^n$ be a continuous vector field. We say that $f \in \mathrm{Lip}_{\mathfrak{Z}}(\mathbb{C}^n, \mathbb{C}^n)$ if for any $r \geq 0$,

$$|f|_{B_r} \leq \mathfrak{Z}(r) \quad \text{and} \quad \mathrm{Lip}(f|_{B_r}) \leq \mathfrak{Z}(r),$$

where Lip g denotes the Lipschitz constant of the mapping g.

2.1. The complex variable system

Let a vector $(x_1, y_1, \ldots, x_n, y_n) \in \mathbb{R}^{2n}$ be given. We introduce a complex structure in $\mathbb{R}^{2n}$ by the notation

$$z_1 = x_1 + \mathbf{i}y_1, \ \ldots, \ z_n = x_n + \mathbf{i}y_n. \tag{2.5}$$

In the complex coordinates, the operator B with the matrix (2.2) is the operator diag$\{\mathbf{i}\lambda_j\}$ and the operator $\eta\tilde{B}$ with the matrix (2.3) is the operator sub-diag$\{\eta\}$. Therefore, the system of linear equation $(1.4)|_{\varepsilon=0} = (2.4)$ can be reduced to a complex system as

$$\dot{u}_j = \mathbf{i}\lambda_j u_j + \eta_j u_{j+1}, \quad 1 \leq j \leq n,$$

where $\eta_j = \begin{cases} \eta, \, j \neq i_1 + \cdots + i_c, \\ 0, \, j = i_1 + \cdots + i_c, \end{cases}$ with $1 \leq c \leq s$.

In the complex notation, the perturbed system (1.4) turns to

$$\begin{aligned}&\dot{u}_j = \mathbf{i}\lambda_j u_j + \eta_j u_{j+1} + \varepsilon \tilde{P}_j(u) + \varepsilon^2 \tilde{g}_j(u,\varepsilon), \quad u(0) = u_0, \quad |u_0| \le r,\\ &u = (u_1, \dots, u_n) \in \mathbb{C}^n, \quad 1 \le j \le n,\\ &\eta_j = \begin{cases} \eta, \ j \ne i_1 + \cdots + i_c, \\ 0, \ j = i_1 + \cdots + i_c, \end{cases} \quad 1 \le c \le s.\end{aligned} \tag{2.6}$$

Remark 2.1. In the complex notation, the perturbed system (1.1) turns to

$$\dot{v}_j = \mathbf{i}\lambda_j v_j + v_{j+1} + \varepsilon \tilde{H}_j(u), \quad v(0) = v_0, \tag{2.7}$$

where $v = (v_1, \dots, v_n) \in \mathbb{C}^n, 1 \le j \le n$.

Since the vector fields $\tilde{P}$ and $\tilde{g}$ are analytic, they are automatically locally Lipschitz, that is, their restrictions to bounded ball B_{2r} with any $r > 0$, are Lipschitz continuous. Therefore, we can let

$$\tilde{P} \in \mathrm{Lip}_{\mathfrak{X}}(\mathbb{C}^n, \mathbb{C}^n), \quad \tilde{g} \in \mathrm{Lip}_{\mathfrak{Y}}(\mathbb{C}^n, \mathbb{C}^n), \tag{2.8}$$

where $\mathfrak{X} : \mathbb{R}^+ \to \mathbb{R}^+$ and $\mathfrak{Y} : \mathbb{R}^+ \to \mathbb{R}^+$ are non-decreasing continuous functions.

Lemma 2.1. *Let $u_0 \in \bar{B}_r$, and $b = \frac{r}{\mathfrak{X}(2r)+\mathfrak{Y}(2r)+2r\varpi}$, where $\mathfrak{X} : \mathbb{R}^+ \to \mathbb{R}^+$ and $\mathfrak{Y} : \mathbb{R}^+ \to \mathbb{R}^+$ are non-decreasing continuous functions. Then a solution $u(t)$ of system* (2.6) *exists for $|t| \le \varepsilon^{-1} b$ and stays in the ball $\bar{B}_{2r}$.*

Proof. By the Lipschitz properties of $\tilde{P}$ and $\tilde{g}$, a solution $u(t)$ of (2.6) exists up to the blow-up time. Taking the scalar product of system (2.6) with $u(t)$ and comparing the real part, we get

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|u(t)|^2 &= \mathbf{i}\langle \mathrm{diag}\{\lambda_j\} u, u\rangle + \langle \text{sub-diag}\{\eta\} u, u\rangle + \varepsilon \langle \tilde{P}(u), u\rangle + \varepsilon^2 \langle \tilde{g}(u,\varepsilon), u\rangle\\ &= \langle \text{sub-diag}\{\eta\} u, u\rangle + \varepsilon \langle \tilde{P}(u), u\rangle + \varepsilon^2 \langle \tilde{g}(u,\varepsilon), u\rangle.\end{aligned}$$

Let $T = \inf\{t \in [0, \varepsilon^{-1} b] : |u(t)| \ge 2r\}$, where T equals to $\varepsilon^{-1} b$ if the set under the inf-sign is empty. Then for $0 < t \le T$ we have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|u(t)|^2 &\le \eta |u|^2 + \varepsilon |u||\tilde{P}(u)| + \varepsilon^2 |u||\tilde{g}(u,\varepsilon)|\\ &\le \eta |u(t)|^2 + \varepsilon \mathfrak{X}(2r)|u(t)| + \varepsilon^2 \mathfrak{Y}(2r)|u(t)|,\end{aligned}$$

which implies $\frac{d|u(t)|}{dt} \le \eta |u(t)| + \varepsilon \mathfrak{X}(2r) + \varepsilon^2 \mathfrak{Y}(2r)$. Thus, $|u(t)| \le \int_0^t \varepsilon\varpi |u(s)| ds + r + \varepsilon \mathfrak{X}(2r) t + \varepsilon \mathfrak{Y}(2r) t < 2r$ due to $\eta = \varepsilon\varpi, |u(0)| \le r$ and $0 < \varepsilon \le 1$, i.e., $2r\varepsilon\varpi t + \varepsilon \mathfrak{X}(2r) t + \varepsilon \mathfrak{Y}(2r) t < r$ for all $0 < t < \varepsilon^{-1} b$ with b given before. Therefore, $T = \varepsilon^{-1} b$ and the result follows. $\square$

Let $\tau = \varepsilon t$. Then $\frac{\partial u}{\partial t} = \varepsilon \frac{\partial u}{\partial \tau}$, and system (2.6) can be reduced to

$$\frac{\partial u_j}{\partial \tau} = \mathbf{i}\varepsilon^{-1}\lambda_j u_j + \varepsilon^{-1}\eta_j u_{j+1} + \tilde{P}_j + \varepsilon \tilde{g}_j, \quad 1 \le j \le n. \tag{2.9}$$

Let us define

$$\phi_j = e^{-\mathbf{i}\varepsilon^{-1}\lambda_j \tau} \sum_{\substack{k=0\\ i_1+\cdots+i_{c-1} < j \le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j} \frac{(-\varepsilon^{-1}\eta\tau)^k}{k!} u_{j+k}(\tau), \quad 1 \le j \le n,$$

and

$$\begin{aligned}\tilde{\phi}_j &= \eta^{j-(i_1+\cdots+i_{c-1}+1)}\phi_j\\ &= e^{-\mathbf{i}\varepsilon^{-1}\lambda_j\tau}\sum_{\substack{k=0\\ i_1+\cdots+i_{c-1}<j\le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j}\eta^{j-(i_1+\cdots+i_{c-1}+1)}\frac{(-\varepsilon^{-1}\eta\tau)^k}{k!}u_{j+k}(\tau)\\ &= e^{-\mathbf{i}\varepsilon^{-1}\lambda_j\tau}\sum_{\substack{k=0\\ i_1+\cdots+i_{c-1}<j\le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j}\eta^{j-(i_1+\cdots+i_{c-1}+1)}\frac{(-\varepsilon\varpi^2\tau)^k}{k!}u_{j+k}(\tau),\ 1\le j\le n.\end{aligned}$$

Denote $\Lambda=(\lambda_1,\dots,\lambda_n)$. Then using (2.1) and (2.9), we obtain

$$\begin{aligned}\frac{\partial\phi_j}{\partial\tau} = e^{-\mathbf{i}\varepsilon^{-1}\lambda_j\tau}&\sum_{\substack{k=0\\ i_1+\cdots+i_{c-1}<j\le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j}\frac{(-\varepsilon^{-1}\eta\tau)^k}{k!}(\tilde{P}_{j+k}((\Phi\circ\Psi)_{\tau\varepsilon^{-1}\Lambda,\tau\varepsilon^{-1}\eta}\phi(\tau))\\ &+\varepsilon e^{-\mathbf{i}\lambda_j\tau\varepsilon^{-1}}\tilde{g}_{j+k}((\Phi\circ\Psi)_{\tau\varepsilon^{-1}\Lambda,\tau\varepsilon^{-1}\eta}\phi(\tau),\varepsilon))\end{aligned}\tag{2.10}$$

for $\phi(\tau)=(\phi_1(\tau),\dots,\phi_n(\tau))\in\mathbb{C}^n$ in the time scale $|\tau|\le b$ due to Lemma 2.1. Thus, system (2.10) takes the form

$$\begin{aligned}\frac{\partial\phi}{\partial\tau} =&(\Phi\circ\Psi)_{-\tau\varepsilon^{-1}\Lambda,-\tau\varepsilon^{-1}\eta}\circ(\tilde{P}((\Phi\circ\Psi)_{\tau\varepsilon^{-1}\Lambda,\tau\varepsilon^{-1}\eta}\phi(\tau))\\ &+\varepsilon\tilde{g}((\Phi\circ\Psi)_{\tau\varepsilon^{-1}\Lambda,\tau\varepsilon^{-1}\eta}\phi(\tau),\varepsilon))\end{aligned}\tag{2.11}$$

with the initial condition

$$\phi(0)=u_0,\quad |u_0|:=r.\tag{2.12}$$

3. Averaging of vector fields

We recall that a diffeomorphism $G:\mathbb{R}^{2n}\to\mathbb{R}^{2n}$ transforms a vector field W on $\mathbb{R}^{2n}$ into the vector field $(G_*W)(u):=dG(\tilde{u})(W(\tilde{u})),\tilde{u}=G^{-1}(u)$. Accordingly, the linear isomorphism $(\Phi\circ\Psi)_{t\Lambda,t\eta}$ with $t\in\mathbb{R}$ transforms the vector field $\tilde{P}$ into $(((\Phi\circ\Psi)_{-t\Lambda,-t\eta})_*\tilde{P})(u)=(\Phi\circ\Psi)_{-t\Lambda,-t\eta}\circ\tilde{P}((\Phi\circ\Psi)_{t\Lambda,t\eta}u)$. Our goal in this section is to study the averaging of this field with respect to t.

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

$$Q_T^o(\phi)=\frac{1}{T}\int_0^T(\Phi\circ\Psi)_{-t\Lambda,-t\eta}\circ\tilde{P}((\Phi\circ\Psi)_{t\Lambda,t\eta}\phi)dt,\tag{3.1}$$

and if the limit exists as $t\to\pm\infty$(for $T<0$ we understand $\int_0^T\dots dt$ as the integral $\int_T^0\dots dt$). We denote $\psi^t(\phi)=(\Phi\circ\Psi)_{-t\Lambda,-t\eta}\circ\tilde{P}((\Phi\circ\Psi)_{t\Lambda,t\eta}\phi)$. For $T\neq0$, we have $Q_T^o(\phi)=\frac{1}{T}\int_0^T\psi^t(\phi)dt$. We denote

$$Q^o(\phi)=\lim_{T\to\pm\infty}Q_T^o(\phi),$$

and $Q_T^o(\phi)$ is called the local average of Q. The matrices of operator B and operator $\tilde{B}$ take the forms (2.2) and (2.3) in the special basis, respectively. By introducing the complex structure (2.5) in $\mathbb{R}^{2n}$, the matrices become $\mathrm{diag}\{\mathbf{i}\lambda_j\}$ and sub-diag$\{\eta\}$. Since $(\Phi\circ\Psi)_{t\Lambda,t\eta}=e^{\mathrm{diag}\{\mathbf{i}\lambda_j\}t+\text{sub-diag}\{\eta\}t}$, the definition of Q^o agrees with (1.6).

Lemma 3.1. *$Q_T^o \in \mathrm{Lip}_{4\mathfrak{X}}(\mathbb{C}^n, \mathbb{C}^n)$ for any $T \neq 0$.*

Proof. If $\phi \in \bar{B}_r$, then $(\Phi \circ \Psi)_{t\Lambda, t\eta}\phi \in \bar{B}_{2r}$. By (2.8), we have $|\tilde{P}((\Phi \circ \Psi)_{t\Lambda,t\eta}\phi)| \leq \mathfrak{X}(2r)$ for each t, and therefore

$$|\psi^t(\phi)| \leq 2|\tilde{P}((\Phi \circ \Psi)_{t\Lambda,t\eta}\phi)| \leq 2\mathfrak{X}, \quad \forall t \in \mathbb{R}. \tag{3.2}$$

Similarly, for any $\phi^1, \phi^2 \in \bar{B}_r$ we have

$$\begin{aligned}|\psi^t(\phi^1) - \psi^t(\phi^2)| &\leq 2|\tilde{P}((\Phi \circ \Psi)_{t\Lambda,t\eta}\phi^1) - \tilde{P}((\Phi \circ \Psi)_{t\Lambda,t\eta}\phi^2)| \\ &\leq 2\mathfrak{X}|(\Phi \circ \Psi)_{t\Lambda,t\eta}\phi^1 - (\Phi \circ \Psi)_{t\Lambda,t\eta}\phi^2| \\ &\leq 4\mathfrak{X}|\phi^1 - \phi^2|, \quad \forall t \in \mathbb{R}. \end{aligned}\tag{3.3}$$

From (3.2) and (3.3), one arrives at $|Q_T^o(\phi)| \leq \sup\limits_{|t|\leq|T|} |\psi^t(\phi)| \leq 2\mathfrak{X} \leq 4\mathfrak{X}$ and $|Q_T^o(\phi^1) - Q_T^o(\phi^2)| \leq \sup\limits_{|t|\leq|T|} |\psi^t(\phi^1) - \psi^t(\phi^2)| \leq 4\mathfrak{X}|\phi^1 - \phi^2|$. As a consequence, $Q_T^o \in \mathrm{Lip}_{4\mathfrak{X}}(\mathbb{C}^n, \mathbb{C}^n)$ for $T \neq 0$. □

Definition 3.1. If $\lambda_j - \Lambda \cdot \gamma + \Lambda \cdot \beta = 0$, where $\gamma = (\gamma_1, \ldots, \gamma_n)$, $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{Z}_+^n$, $\Lambda = (\lambda_1, \ldots, \lambda_n)$, we say a pair (γ, β) is $(\Lambda, j)-$resonant.

Lemma 3.2. *(Main averaging lemma) For $\Lambda \in (\mathbb{R} \setminus \{0\})^n$, analytic $\tilde{P}$ satisfying $\tilde{P} = o(|u|^l)$ with $l \geq 2$ and $\tilde{P} \in \mathrm{Lip}_{\mathfrak{X}}(\mathbb{C}^n, \mathbb{C}^n)$, the limit (namely Q^o) for* (3.1) *exists for any $\phi \in \mathbb{C}^n$, and $Q^o \in \mathrm{Lip}_{4\mathfrak{X}}(\mathbb{C}^n, \mathbb{C}^n)$, where $\mathfrak{X} : \mathbb{R}^+ \to \mathbb{R}^+$ is a non-decreasing continuous function. If $\phi \in \bar{B}_r$, then the rate of convergence in* (3.1) *only depends on r, Λ and $\tilde{P}$.*

Proof. Recall the analyticity of $\tilde{P}$. Then using the Taylor expansion for $\tilde{P}_j$ (component of $\tilde{P}$), we have

$$\tilde{P}_j(u) = \sum_{|\gamma^1|+|\gamma^2|=1} C_j^{\gamma^1\gamma^2} u^{\gamma^1}\bar{u}^{\gamma^2} + \sum_{|\beta^1|+|\beta^2|>l} C_j^{\beta^1\beta^2} u^{\beta^1}\bar{u}^{\beta^2}, \quad 1 \leq j \leq n \tag{3.4}$$

due to the normal form, where $\gamma^1 = (\gamma_1^1, \ldots, \gamma_n^1)$, $\gamma^2 = (\gamma_1^2, \ldots, \gamma_n^2)$, $\beta^1 = (\beta_1^1, \ldots, \beta_n^1)$ and $\beta^2 = (\beta_1^2, \ldots, \beta_n^2)$. Then

$$\begin{aligned}&\psi_j^t(u) \\ &= \sum_{\substack{k=0 \\ i_1+\cdots+i_{c-1}<j\leq i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j} \frac{(-t\eta)^k}{k!}\Bigg[\sum_{|\gamma^1|+|\gamma^2|=1} C_j^{\gamma^1\gamma^2} e^{\mathbf{i}t(-\lambda_j+\Lambda\cdot\gamma^1-\Lambda\cdot\gamma^2)} f_{\gamma_{j'}^1,\gamma_{j'}^2}(t,u,\bar{u}) \\ &\quad + \sum_{|\beta^1|+|\beta^2|>l} C_j^{\beta^1\beta^2} e^{\mathbf{i}t(-\lambda_j+\Lambda\cdot\beta^1-\Lambda\cdot\beta^2)} f_{\beta_{j'}^1,\beta_{j'}^2}(t,u,\bar{u})\Bigg]\end{aligned}$$

for $1 \leq j \leq n$, provided with

$$\begin{aligned}f_{\gamma_{j'}^1,\gamma_{j'}^2}(t,u,\bar{u}) = \prod_{1\leq j'\leq n}&\Big(\sum_{\substack{k'=0 \\ i_1+\cdots+i_{c-1}<j'\leq i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j'} \frac{(t\eta)^{k'}}{k'!} u_{j'+k'}\Big)^{\gamma_{j'}^1} \\ &\Big(\sum_{\substack{k'=0 \\ i_1+\cdots+i_{c-1}<j'\leq i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j'} \frac{(t\eta)^{k'}}{k'!} \bar{u}_{j'+k'}\Big)^{\gamma_{j'}^2},\end{aligned}$$

$$f_{\beta^1_{j'},\beta^2_{j'}}(t,u,\bar{u}) = \prod_{1\le j'\le n}\Big(\sum_{\substack{k'=0\\ i_1+\cdots+i_{c-1}<j'\le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j'} \frac{(t\eta)^{k'}}{k'!}u_{j'+k'}\Big)^{\beta^1_{j'}}$$
$$\Big(\sum_{\substack{k'=0\\ i_1+\cdots+i_{c-1}<j'\le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j'} \frac{(t\eta)^{k'}}{k'!}\bar{u}_{j'+k'}\Big)^{\beta^2_{j'}}.$$

It follows that

$$\begin{aligned} Q^o_{T,j}(u) = &\sum_{\substack{k=0\\ i_1+\cdots+i_{c-1}<j\le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j} \sum_{|\gamma^1|+|\gamma^2|=1} C_j^{\gamma^1\gamma^2} \\ &\left(\frac{1}{T}\int_0^T \frac{(-t\eta)^k}{k!} e^{\mathbf{i}t(-\lambda_j+\Lambda\cdot\gamma^1-\Lambda\cdot\gamma^2)} f_{\gamma^1_{j'},\gamma^2_{j'}}(t,u,\bar{u})dt\right) \\ &+ \sum_{\substack{k=0\\ i_1+\cdots+i_{c-1}<j\le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j} \sum_{|\beta^1|+|\beta^2|>l} C_j^{\beta^1\beta^2} \\ &\left(\frac{1}{T}\int_0^T \frac{(-t\eta)^k}{k!} e^{\mathbf{i}t(-\lambda_j+\Lambda\cdot\beta^1-\Lambda\cdot\beta^2)} f_{\beta^1_{j'},\beta^2_{j'}}(t,u,\bar{u})dt\right). \end{aligned}$$

We set

$$\begin{aligned} Q^r_j(u) &= \sum_{|\gamma|+|\beta|>l} \sum_{\substack{k=0\\ i_1+\cdots+i_{c-1}<j\le i_1+\cdots+i_c}}^{i_1+\cdots+i_c-j} \sum_{\tilde{k}=0}^{\tilde{c}} C_j^{\gamma\beta}(-1)^{\tilde{k}+k}\frac{(\varpi b)^{\tilde{k}+k}}{(\tilde{k}+k+1)!}\frac{D_t^{\tilde{k}} f_{\gamma_{j'}\beta_{j'}}|_{t=\frac{b\varpi}{\eta}}}{\eta^{\tilde{k}}} \\ &= \sum_{|\tilde{\gamma}|+|\tilde{\beta}|>l} C_j^{\tilde{\gamma}\tilde{\beta}} u^{\tilde{\gamma}}\bar{u}^{\tilde{\beta}} \end{aligned}$$

for $1\le j\le n$, where (γ,β) is a $(\Lambda,j)-$resonant pair, $\tilde{c} = \sum_{c=1}^{s}\sum_{k'=0}^{i_c-1}(i_c-k')(\beta_{k'+1}+\gamma_{k'+1})$, $D_t^m f_{\gamma_{j'}\beta_{j'}} = \frac{\partial^m f_{\gamma_{j'}\beta_{j'}}}{\partial t^m}$ and $\eta = \varepsilon\varpi$. If $1\le|\gamma|+|\beta|\le l$, then $\lambda_j - \Lambda\cdot\gamma + \Lambda\cdot\beta \ne 0$ for $1\le j\le n$ by (1.2) (i.e., are nonresonant).

One can observe that

$$\lim_{T\to\pm\infty} Q^o_{T,j}(u) = \begin{cases} Q^r_j(u), \text{ if } (\gamma,\beta) \text{ is } (\Lambda,j)-\text{resonant}, \\ 0, \text{ otherwise.} \end{cases} \tag{3.5}$$

Thus, we have

$$Q^o_{T,j}(u) \to Q^r_j(u) \text{ as } T\to\pm\infty. \tag{3.6}$$

By Lemma 3.1, we have its Lipschitz continuity. This completes the proof. □

Proposition 3.1. *Let Q^1 and Q^2 be locally Lipschitz vector fields on $\mathbb{C}^n$. Then the mapping $u\mapsto Q^o(u)$ commutes with all the operators $(\Phi\circ\Psi)_{b\Lambda,b\eta}$ for $b\in\mathbb{R}$.*

Proof. This proof is similar to the proof in [8]. □

4. Proof of the averaging theorems

This section is devoted to the proofs of Theorems 1.1 and 1.2.

Recalling (2.8) and (2.9), we have $Q^o \in \mathrm{Lip}_{4\mathfrak{X}}(\mathbb{C}^n, \mathbb{C}^n)$. Let $|u_0| = r$. Then by Lemma 2.1, $u(\tau) \in \bar{B}_{2r}$ for $|\tau| \le b = \frac{r}{\mathfrak{X}(2r)+\mathfrak{Y}(2r)+2r\varpi}$, where $\varpi > 0$. For $|\tau| \le b$ the curve $\phi^\varepsilon(\tau) = (\Phi \circ \Psi)_{-\varepsilon^{-1}\tau\Lambda, -\varepsilon^{-1}\tau\eta} u(\tau)$ satisfies (2.11) and (2.12). Hence, for $|\tau| \le b$, we have

$$|\phi^\varepsilon(\tau)| \le 4r, \ \left|\frac{\partial \phi^\varepsilon}{\partial \tau}\right| \le 2\left|\tilde{P}((\Phi \circ \Psi)_{\varepsilon^{-1}\tau\Lambda, \varepsilon^{-1}\tau\eta}\phi(\tau)) + \tilde{g}((\Phi \circ \Psi)_{\varepsilon^{-1}\tau\Lambda, \varepsilon^{-1}\tau\eta}\phi(\tau))\right| \le 2\mathfrak{X}(4r) + \mathfrak{Y}(4r). \tag{4.1}$$

By (4.1), it holds

$$|\phi^\varepsilon(\tau_1) - \phi^\varepsilon(\tau_2)| \le (2\mathfrak{X}(4r) + \mathfrak{Y}(4r))|\tau_1 - \tau_2|. \tag{4.2}$$

For all $|\tau| \le b$, a solution $\phi^\varepsilon(\tau)$ of (2.11) satisfies the relation

$$\begin{aligned}\phi^\varepsilon(\tau) =& u_0 + \int_0^\tau (\Phi \circ \Psi)_{-\varepsilon^{-1}\upsilon\Lambda, -\varepsilon^{-1}\upsilon\eta} \circ \tilde{P}((\Phi \circ \Psi)_{\varepsilon^{-1}\upsilon\Lambda, \varepsilon^{-1}\upsilon\eta}\phi^\varepsilon(\upsilon))d\upsilon \\ &+ \int_0^\tau \varepsilon(\Phi \circ \Psi)_{-\varepsilon^{-1}\upsilon\Lambda, -\varepsilon^{-1}\upsilon\eta} \circ \tilde{g}((\Phi \circ \Psi)_{\varepsilon^{-1}\upsilon\Lambda, \varepsilon^{-1}\upsilon\eta}\phi^\varepsilon(\upsilon), \varepsilon)d\upsilon\end{aligned} \tag{4.3}$$

and the estimates in (4.1).

Lemma 4.1. *Let $\phi^\varepsilon(\tau)$ be a solution of (4.3), where $\varepsilon \in (0, 1]$, $|\tau| \le b$. Then*

$$\phi^\varepsilon \to \phi^0 \text{ in } C([-b, b], \mathbb{C}^n) \text{ as } \varepsilon \to 0^+. \tag{4.4}$$

Proof. By (4.1) and (4.2), the family $\{\phi^\varepsilon, 0 < \varepsilon \le 1\}$ satisfies the Arzelà-Ascoli theorem. Then $\{\phi^\varepsilon, 0 < \varepsilon \le 1\}$ is precompact in $C([-b, b], \mathbb{C}^n)$. Therefore, there exists a sequence $\varepsilon_j \to 0^+$ such that $\phi^{\varepsilon_j} \to \phi^0$ in $C([-b, b], \mathbb{C}^n)$ as $\varepsilon_j \to 0^+$. Since the solution of (4.3) is unique, the limit ϕ^0 does not depend on the sequence $\varepsilon_j \to 0^+$, then the convergence in (4.4) holds as $\varepsilon \to 0^+$. □

In view of this convergence, we obtain $\left|\phi^0(\tau_1) - \phi^0(\tau_2)\right| \le (2\mathfrak{X}(4r)+\mathfrak{Y}(4r))|\tau_1 - \tau_2|$ for all $\tau_1, \tau_2 \in [-b, b]$, as long as $\varepsilon \to 0^+$.

Lemma 4.2. *For any $|\tau| \le b$, there exists a function $\iota(\varepsilon)$ satisfying $\iota(\varepsilon) \to 0$ as $\varepsilon \to 0^+$, such that*

$$\left|Y(\phi^\varepsilon(\tau), \tau\varepsilon^{-1})\right| \le \iota(\varepsilon), \tag{4.5}$$

where $Y(\phi^\varepsilon(\tau), \tau\varepsilon^{-1}) = \int_0^\tau (\psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(\upsilon)) - Q^o(\phi^\varepsilon(\upsilon)))d\upsilon$.

Proof. It is sufficient to prove the conclusion for $\tau > 0$. We divide the time interval $[0, b]$ into subintervals $[b_{j-1}, b_j]$, $j = 1, \ldots, N'+1$ of length $L = \varepsilon^{\frac{1}{2}}$, $b_{\tilde{r}} = \tilde{r}L$ for $\tilde{r} = 0, \ldots, N'$, $b_{N'} = b$ and $0 \le b_{N'+1} - b_{N'} < L$, where $N' = [\frac{b}{L}]$ ($[b']$ is the largest integer less or equal to b').

In view of (3.2), (3.3) and Lemma 3.2, we get

$$\left|\int_{b_{N'}}^{b_{N'+1}} (\psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(\upsilon)) - Q^o(\phi^\varepsilon(\upsilon)))d\upsilon\right| \le 6L\mathfrak{X}. \tag{4.6}$$

Similarly, if $\tau \in [b_{\tilde{k}}, b_{\tilde{k}+1})$ for some $0 \le \tilde{k} \le N'+1$, then $\int_{b_{\tilde{k}}}^{\tau} (\psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(\upsilon)) - Q^o(\phi^\varepsilon(\upsilon)))d\upsilon$ is also bounded by the R.H.S. of (4.6).

Now we estimate the integral of Y over any segment $[b_j, b_{j+1}]$, where $1 \le j \le N'-1$. Firstly, we have

$$\begin{aligned}
&\left| \int_{b_j}^{b_{j+1}} (\psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(\upsilon)) - Q^o(\phi^\varepsilon(\upsilon)))d\upsilon \right| \\
&\le \left| \int_{b_j}^{b_{j+1}} (\psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(b_j)) - Q^o(\phi^\varepsilon(b_j)))d\upsilon \right| \\
&\quad + \left| \int_{b_j}^{b_{j+1}} (\psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(\upsilon)) - \psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(b_j)))d\upsilon \right| \\
&\quad + \left| \int_{b_j}^{b_{j+1}} (Q^o(\phi^\varepsilon(b_j)) - Q^o(\phi^\varepsilon(\upsilon)))d\upsilon \right|.
\end{aligned}$$

By virtue of Lemma 3.2 and (4.2), the second and third terms in the R.H.S. are bounded by $4L^2\mathfrak{X}(2\mathfrak{X}+\mathfrak{Y})$. Since

$$\begin{aligned}
&\left| \int_{b_j}^{b_{j+1}} (\psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(b_j)) - Q^o(\phi^\varepsilon(b_j)))d\upsilon \right| \\
&= \left| L(\Phi \circ \Psi)_{-\varepsilon^{-1}b_j\Lambda, -\varepsilon^{-1}b_j\eta} Q^o_{L^{-1}}(z) - LQ^o(\phi^\varepsilon(b_j)) \right|,
\end{aligned}$$

where $z = (\Phi \circ \Psi)_{\varepsilon^{-1}b_j\Lambda, \varepsilon^{-1}b_j\eta}\phi^\varepsilon(b_j) \in \bar{B}_{2r}$, using Lemma 3.2 and Proposition 3.1 we see that

$$\begin{aligned}
Q^o_{L^{-1}}(z) &= Q^o(z) + o(1) \\
&= (\Phi \circ \Psi)_{\varepsilon^{-1}b_j\Lambda, \varepsilon^{-1}b_j\eta} Q^o(\phi^\varepsilon(b_j)) + o(1) \text{ as } L^{-1} \to +\infty \ \text{ (i.e., } \varepsilon \to 0^+).
\end{aligned}$$

Now we have arrived at the estimate

$$\left| \int_{b_j}^{b_{j+1}} (\psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(\upsilon)) - Q^o(\phi^\varepsilon(\upsilon)))d\upsilon \right| \le L \cdot o(1) + 8L^2\mathfrak{X}(2\mathfrak{X}+\mathfrak{Y}) = L \cdot o(1). \quad (4.7)$$

By (4.6) and (4.7), the L.H.S. of (4.5) is bounded by $b \cdot o(1) + 6\varepsilon^{\frac{1}{2}}(\mathfrak{X} + \tilde{\mathfrak{X}}) := \iota(\varepsilon)$. This completes the proof of Lemma 4.1. □

Consider the following effective system:

$$\phi^0(\tau) = u_0 + \int_0^\tau Q^o(\phi^0(\upsilon))d\upsilon, \ \phi^0(0) = u_0, \quad (4.8)$$

that is, $\frac{\partial \phi^0(\tau)}{\partial \tau} = Q^o(\phi^0(\tau))$ with $\phi^0(0) = u_0$. By Lemma 3.2, Q^o is locally Lipschitz, a solution of (4.8) exists and is unique, at least for small τ.

Theorem 4.1. *Let $\phi^\varepsilon(\tau)$ with $|\tau| \le b$ be a solution of* (4.3). *Then $\phi^\varepsilon(\tau) \to \phi^0(\tau)$ uniformly for $|\tau| \le b$, where $\phi^0(\tau)$ is the unique solution of* (4.8).

Proof. By Lemma 4.1, we can get $\phi^\varepsilon \to \phi^0$ in $C([-b,b], \mathbb{C}^n)$ as $\varepsilon \to 0^+$. Then in view of (2.8), (4.3), (4.4), Lemma 4.2, and the fact that Q^o is locally Lipschitz, we have

$$\phi^0(\tau) - \phi^\varepsilon(\tau) \to 0,$$

$$
\begin{aligned}
&\phi^\varepsilon(\tau) - u_0 - \int_0^\tau \psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(\upsilon))d\upsilon \\
&- \varepsilon \int_0^\tau (\Phi \circ \Psi)_{-\varepsilon^{-1}\upsilon\Lambda, -\varepsilon^{-1}\upsilon\eta} \circ \tilde{g}((\Phi \circ \Psi)_{\varepsilon^{-1}\upsilon\Lambda, \varepsilon^{-1}\upsilon\eta}\phi^\varepsilon(\upsilon), \varepsilon)d\upsilon \\
\to& 0, \\
&\varepsilon \int_0^\tau (\Phi \circ \Psi)_{-\varepsilon^{-1}\upsilon\Lambda, -\varepsilon^{-1}\upsilon\eta} \circ \tilde{g}((\Phi \circ \Psi)_{\varepsilon^{-1}\upsilon\Lambda, \varepsilon^{-1}\upsilon\eta}\phi^\varepsilon(\upsilon), \varepsilon)d\upsilon \\
&+ \int_0^\tau \psi^{\upsilon\varepsilon^{-1}}(\phi^\varepsilon(\upsilon)) - Q^o(\phi^\varepsilon(\upsilon)))d\upsilon \\
\to& 0, \\
&\int_0^\tau (Q^o(\phi^\varepsilon(\upsilon)) - Q^o(\phi^0(\upsilon)))d\upsilon \to 0,
\end{aligned}
$$

whenever $\varepsilon \to 0^+$. Therefore, combining the above estimates we prove that ϕ^0 is a solution of (4.8) for $|\tau| \leq b$. □

Remark 4.1. Since Q^o is continuous in $\mathbb{C}^n$ and $Q^o \in \mathrm{Lip}_{4\mathfrak{X}}(\mathbb{C}^n, \mathbb{C}^n)$, the existence interval of the solution of (4.8) ϕ^0 is $(-\infty, +\infty)$.

By Theorem 4.1, we have

$$
\|\tilde{\phi}^\varepsilon(\tau) - \tilde{\phi}^0(\tau)\| = \|C\phi^\varepsilon(\tau) - C\phi^0(\tau)\| \leq \|\phi^\varepsilon(\tau) - \phi^0(\tau)\| \to 0 \text{ as } \varepsilon \to 0. \tag{4.9}
$$

Hence, by (4.9) and the proof of Theorem 4.1, we can derive Theorem 1.2 directly.

Corollary 4.1. *The solution $v^\varepsilon(t)$ of equation (2.7) satisfies*

$$
\sup_{|t|\leq\varepsilon^{-1}b} |v_j^\varepsilon(t) - \tilde{\phi}_j^0(t)| \to 0 \textit{ as } \varepsilon \to 0^+, \quad 1 \leq j \leq n.
$$

Acknowledgements

We would like to thank Professor Yong Li (Jilin University) for many useful discussions along this work. We are also very grateful to the editor and referees for their valuable suggestions and comments, which led to a significant improvement of the paper.

References

[1] P. Fatou, *Sur le mouvement d'un systéme soumis à des forces à courte période*, Bull. Soc. Math. Fr., 1928, 56, 98–139.

[2] P. Gao, *Averaging principle for stochastic Kuramoto-Sivashinsky equation with a fast oscillation*, Discrete Contin. Dyn. Syst., 2018, 38(11), 5649–5684. DOI: 10.3934/dcds.2018247.

[3] P. Gao, *Averaging principles for stochastic 2D Navier-Stokes equations*, J. Stat. Phys., 2022, 186(2), 28, 29 pp. DOI: 10.1007/s10955-022-02876-9.

[4] P. Gao and Y. Li, *Averaging principle for the Schrödinger equations*, Discrete Contin. Dyn. Syst. Ser. B, 2017, 22(6), 2147–2168. DOI: 10.3934/dcdsb.2017089.

[5] I. M. Gel'fand, *Lectures on Linear Algebra*, Interscience Tracts in Pure and Applied Mathematics, Vol. 9., Interscience Publishers, New York-London, 1961.

[6] G. Huang, *An averaging theorem for nonlinear Schrödinger equations with small nonlinearities*, Discrete Contin. Dyn. Sys., 2014, 34(9), 3555–3574. DOI: 10.3934/dcds.2014.34.3555.

[7] G. Huang, S. B. Kuksin and A. M. Maiocchi, *Time-averaging for weakly nonlinear CGL equations with arbitrary potentials*, Hamiltonian Partial Differential Equations and Applications, Vol. 75, Springer, New York, 2015, 323–349. DOI: 10.1007/978-1-4939-2950-4_11.

[8] W. W. Jian, S. B. Kuksin and Y. Wu, *Krylov-Bogolyubov averaging*, Uspekhi Mat. Nauk, 2020, 75(3), 37–54. DOI: 10.4213/rm9933.

[9] R. Z. Khas'minskii, *The behavior of a conservative system under the action of slight friction and slight random noise*, Prikl. Mat. Meh., 1964, 28(5), 931–935. DOI: 10.1016/0021-8928(64)90017-6.

[10] V. G. Kolomiets, A. I. Mel'nikov and T. M. Muhitdinov, *Averaging of stochastic systems of integral-differential equations with Poisson noise*, Ukrain. Mat. Zh., 1991, 43(2), 273–277. DOI: 10.1007/BF01060515.

[11] N. M. Krylov and N. N. Bogoliubov, *Introduction to Nonlinear Mechanics*, Princeton University Press, Princeton, 1947.

[12] L. N'Goran and M. N'zi, *Averaging principle for multivalued stochastic differential equations*, Random Oper. Stoch. Equ., 2001, 9(4), 399–407. DOI: 10.1515/rose.2001.9.4.399.

[13] J. A. Sanders and F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems*, Applied Mathematical Sciences, Vol. 59, Springer-Verlag, New York, 1985. DOI: 10.1007/978-1-4757-4575-7.

[14] R. L. Stratnovich, *Izbrannye Voprosy Teorii Fluktuatsii v Radiotekhnike*, 1961.

[15] I. Vrkoč, *Extension of the averaging method to stochastic equations*, Czechoslovak Math. J., 1966, 16(91), 518–544.

[16] S. Y. Xiao and Y. Li, *Normal form for nonlinear quasiperiodic systems*, submitted.

[17] J. M. Xing, X. Yang and Y. Li, *Affine-periodic solutions by averaging methods*, Sci. China Math., 2018, 61(3), 439–452. DOI: 10.1007/s11425-016-0455-1.

[18] Y. Xu, J. Q. Duan and W. Xu, *An averaging principle for stochastic dynamical systems with Lévy noise*, Phys. D, 2011, 240, 1395–1401. DOI: 10.1016/j.physd.2011.06.001.

[19] N. T. Zung, *Convergence versus integrability in Birkhoff normal form*, Ann. Math., 2005, 161(1), 141–156. DOI: 10.4007/annals.2005.161.141.