This is a preprint of: "Polynomial and linearized normal forms for almost periodic differential systems", Weigu Li, Jaume Llibre, Hao Wu, *Discrete Contin. Dyn. Syst.*, vol. 36(1), 345–360, 2016.
DOI: [10.3934/dcds.2016.36.345]

POLYNOMIAL AND LINEARIZED NORMAL FORMS FOR ALMOST PERIODIC DIFFERENTIAL SYSTEMS

Weigu Li

School of Mathematical Sciences, Peking University
100871, Beijing, China

Jaume Llibre

Departament de Matemàtiques, Universitat Autònoma de Barcelona
08193 Bellaterra, Barcelona, Spain

Hao Wu

Department of Mathematics, Southeast University
Nanjing 210096, Jiangsu, China

Abstract. For almost periodic differential systems $\dot{x} = \varepsilon f(x, t, \varepsilon)$ with $x \in \mathbb{C}^n$, $t \in \mathbb{R}$ and $\varepsilon > 0$ small enough, we get a polynomial normal form in a neighborhood of a hyperbolic singular point of the system $\dot{x} = \varepsilon \lim_{T\to\infty} \dfrac{1}{T}\int_0^T f(x, t, 0)\, dt$, if its eigenvalues are in the Poincaré domain. The normal form linearizes if the real part of the eigenvalues are non–resonant.

1. **Introduction and statement of the main result.** Normal form theory has a long history. The basic idea of simplifying ordinary differential equations through changes of variables can be found in the work of Poincaré [12]. Recently this theory has been developed very rapidly since it plays a very important role in the study of bifurcation, stability and so on. Usually, normal form theory is applied to simplifying a nonlinear system in the neighborhood of a reference solution, which is almost exclusively assumed to be a singular point (sometimes a periodic solution). For an outline of normal form theory, we mentioned the work of Dulac [6], Sternberg [15], Chen [4], Takens [16], and many others.

Normal form theory has a long history. The basic idea of simplifying ordinary differential equations through changes of variables can be found in the work of Poincaré [12]. Recently this theory has been developed very rapidly since it plays a very important role in the study of bifurcation, stability and so on. Usually, normal form theory is applied to simplifying a nonlinear system in the neighborhood of a reference solution, which is almost exclusively assumed to be a singular point (sometimes a periodic solution). For an outline of normal form theory, we mentioned the work of Dulac [6], Sternberg [15], Chen [4], Takens [16], and many others.

2010 *Mathematics Subject Classification.* Primary: 34C29, 34C20.
Key words and phrases. Almost periodic differential systems, normal form, averaging method, linearization.

There are several known equivalent definitions of almost periodic functions. Here we choose the one given by Bochner, which is very direct and useful in its applications to differential equations. Let $f(x,t) \in C(D \times \mathbb{R}, \mathbb{C}^n)$, where D is an open set in $\mathbb{C}^n$(more generally, a separable Banach space), and assume that for any sequence $\{h_k\}$ of real numbers, there exists a subsequence $\{h_{k_j}\}$ such that $\{f(x, t+h_{k_j})\}$ converges uniformly on $S \times \mathbb{R}$, where S is any compact set in D. Then we say $f(x,t)$ is *almost periodic* in t uniformly for $x \in D$. Moreover, for the fixed x let

$$\alpha(\gamma, f) = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(x,t) e^{-\gamma t\sqrt{-1}} dt,$$

the set

$$\Gamma(f) = \{\gamma \in \mathbb{R} : \alpha(\gamma, f) \neq 0\}$$

is called the set of *Fourier exponents* of f. The module generated by $\Gamma(f)$ is defined as the *module* $m(f)$ of f. For instant, if f is periodic of period $2\pi/\omega$, then $\{m(f) = n\omega, n = 0, \pm 1, \cdots, \}$.

Suppose that $x \in \mathbb{C}^n$ and that $\varepsilon \geq 0$ is a real parameter, then we say $f \in \mathbb{F}(D, [0,\infty))$, if

(i) the function f: $D \times \mathbb{R} \times [0,\infty) \to \mathbb{C}^n$ is continuous,
(ii) $f(x,t,\varepsilon)$ is almost periodic in t uniformly with respect to x in compact sets of D for each fixed ε,
(iii) $f(x,t,\varepsilon)$ is analytic with respect to $x \in D$ for fixed t and ε,
(iv) $f(x,t,\varepsilon) \to f(x,t,0)$ as $\varepsilon \to 0$ uniformly for t in $(-\infty,\infty)$, x in compact sets.

We associate to the system of differential equations

$$\dot{x} = \varepsilon f(x,t,\varepsilon), \tag{1}$$

the averaged system

$$\dot{x} = \varepsilon f_0(x), \tag{2}$$

where

$$f_0(x) = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(x,t,0)\, dt.$$

By the classical averaging theorem (see [8, 7]), if $x_0 \in D$ is a hyperbolic singular point of system (2), then system (1) has an almost periodic solution $x^*(t,\varepsilon) \to x_0$ uniformly as $\varepsilon \to 0$.

Let $U_\delta = \{x \in \mathbb{C}^n : \|x\| \leq \delta\}$. Set $\lambda(A) = (\lambda_1, \ldots, \lambda_n)$ be the eigenvalues of $A = \partial f_0(x_0)/\partial x$. Denote by re the real part of a complex. Without loss of generality, we can assume re$\lambda_1 \leq \cdots \leq$ reλ_n. We say that the eigenvalues $\lambda(A)$ are in the *Poincaré domain* if re$\lambda_n < 0$ or re$\lambda_1 > 0$. Moreover, the n–tuple number $\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{C}^n$ is said to be *l–th order non–resonant* if for all $k = (k_1, \ldots, k_n) \in \mathbb{Z}_+^n$ with $l = |k| = \sum_{i=1}^n k_i \geq 2$, we have that

$$\sum_{i=1}^n k_i\mu_i - \mu_j \neq 0, \quad 1 \leq j \leq n. \tag{3}$$

And if the inequalities (3) are satisfied for arbitrary $|k| \geq 2$, μ is called to be *non–resonant*. As usual $\mathbb{Z}_+$ denotes the set of non–negative integers.

We do to system (1) the change of variables $x \mapsto y$ given by

$$x = y + x^*(t,\varepsilon), \tag{4}$$

then the system becomes

$$\dot{y} = \varepsilon F(y,t,\varepsilon), \tag{5}$$

where $F(0,t,\varepsilon) \equiv 0$ and $F(y,t,\varepsilon) = f(y+x^*(t,\varepsilon),t,\varepsilon) - f(x^*(t,\varepsilon),t,\varepsilon)$ satisfies the same conditions as the function $f(x,t,\varepsilon)$.

The following are our main results. One is to deal with formal normal forms of system (5) by the fact that the linear part system

$$\dot{y} = \varepsilon \frac{\partial F}{\partial y}(0,t,\varepsilon) y \tag{6}$$

can be turned into the block diagonal form with respect to the different real parts of eigenvalues in the Jordan normal form of A (See Lemma 3.1). The other is to consider its analytic normal forms in the Poincaré domain.

Theorem 1.1. *Let $\lambda(A) = (\lambda_1, \dots, \lambda_n)$ and $k = (k_1, \dots, k_n) \in \mathbb{Z}_+^n$. Assume A is in the Jordan Normal Form and system (6) is in the block diagonal form with respect to the different real parts of eigenvalues of A. Then for any positive integer $N > 0$ there exists $\varepsilon_0 = \varepsilon_0(N) > 0$ such that for $0 < \varepsilon \le \varepsilon_0$ under the change of coordinates $y = z + \varepsilon w_N(z,t,\varepsilon)$, where $\varepsilon w_N \in \mathcal{F}(U_\delta, [0,\varepsilon_0))$ is a polynomial of degree N with respect to z, $\varepsilon w_N(z,t,\varepsilon) \to 0$ as $\varepsilon \to 0$ uniformly for $(z,t) \in U_\delta \times \mathbb{R}$ and $m(w) \subset m(f)$, system (5) becomes*

$$\dot{z} = \varepsilon Q_N(z,t,\varepsilon) + \varepsilon R(z,t,\varepsilon) \tag{7}$$

where Q_N is a polynomial of degree N in z and the coefficient $q_k^j(t,\varepsilon)$ of terms $x^k e_j$ satisfies $q_k^j(t,\varepsilon) \equiv 0$ if $\mathrm{re}(\sum_{i=1}^n k_i\lambda_i - \lambda_j) \neq 0$ for $1 \le |k| = \sum_{i=1}^n k_i \le N$, $R(z,t,\varepsilon) = O(\|z\|^{N+1})$ as $z \to 0$.

Therefore, the N-th iteration system, i.e.,

$$\dot{z} = \varepsilon Q_N(z,t,\varepsilon),$$

is called the *N-th normal form* of system (7). As usual, $Jet^d_{y=0}F$ is denoted by the set of functional coefficients of the Taylor expansion for the function F of order d with respect to the variable y at $y = 0$. Let $[\cdot]$ be the classical greatest integer function.

Theorem 1.2. *If $\lambda(A)$ is in the Poincaré domain, then there exists $\varepsilon_0 > 0$ and $\delta > 0$ independent of ε such that for $0 < \varepsilon \le \varepsilon_0$ under the change of coordinates $y = z + \varepsilon w(z,t,\varepsilon)$, where $\varepsilon w \in \mathcal{F}(U_\delta, [0,\varepsilon_0))$, $\varepsilon w(z,t,\varepsilon) \to 0$ as $\varepsilon \to 0$ uniformly for $(z,t) \in U_\delta \times \mathbb{R}$ and $m(w) \subset m(f)$, system (5) becomes*

$$\dot{z} = \varepsilon P_d(z,t,\varepsilon), \tag{8}$$

where P_d is a polynomial of degree d with respect to z, $Jet^d_{z=0}P_d = Jet^d_{y=0}F$, the integer $d = \max\{[\mathrm{re}\lambda_1/\mathrm{re}\lambda_n], [\mathrm{re}\lambda_n/\mathrm{re}\lambda_1]\}$. In addition, if the real parts of $\lambda(A)$ are non–resonant, then $d = 1$; i.e.

$$\dot{z} = \varepsilon \frac{\partial F}{\partial y}(0,t,\varepsilon) z.$$

On one hand under the view of normal form theory, there are some obvious difference between our theorem and other theorems. Normally, if the linear part of original system is not an autonomous one, dichotomy spectrum will be applied to character non–resonant conditions. However, how to calculate the exact dichotomy spectrum is still an unsolved technical problem. Whereas for our theorem the resonant condition is given out by the eigenvalues of fixed point of averaged system and it can be verified easily. In this sense, our theorem is more straightforward and clearer. Moreover, if averaging systems are degenerated to the systems independent of the time t, i.e., autonomous systems with small parameter ε, our theorem coincides with the classic one. So our paper also can be seen as the extension of the work in [9] to special non–autonomous systems.

On the other hand under the view of averaging methods, our work gives a new aspect to characterize the close relationship between system (1) and the averaged system. Instead of concerning about the approaching of trajectories of original and corresponding averaged systems in finite time interval, we pay more attention to the long time behavior of solutions. In fact, in Fink's book [7] (pp.266), he regards system (1) as a perturbation of the linear system at the singular point of the averaged system under a stronger condition. Here using our methods, by a more reasonable condition we can obtain a better result, i.e., the linearization and polynomialization of the original system.

The paper is organized as follows. In Section 2, a series lemmas and some useful facts are collected. In Section 3 and 4, we present proofs of main theorems, respectively.

2. **Preliminary results.** In this section we introduce some basic definitions and lemmas, which are important for our proof. First, we describe some of the hyperbolic properties of a non–autonomous system. By studying the exponential dichotomy we can characterize the asymptotic speed of the solutions tending to infinity, which implies a close relationship with the unique bounded solution of a non–autonomous system.

Let $\Phi(t)$ be the fundamental matrix such that $\Phi(0) = I$ (as usual I denotes the identity matrix) for the linear differential equation

$$\dot{x} = A(t)x, \tag{9}$$

where the $n \times n$ coefficient matrix $A(t)$ is continuous in $\mathbb{R}$. Equation (9) is said to possess an *exponential dichotomy* if there exists a *projection* P, that is, a matrix P such that $P^2 = P$, and positive constants K and α such that

$$\begin{array}{ll} \|\Phi(t)P\Phi^{-1}(s)\| \le Ke^{-\alpha(t-s)}, & t \ge s, \\ \|\Phi(t)(I-P)\Phi^{-1}(s)\| \le Ke^{-\alpha(s-t)}, & s \ge t. \end{array}$$

Then, the following lemma show the toughness of the exponential dichotomy. A detailed proof can be found in [5].

Lemma 2.1. *Suppose the linear differential system (9) has an exponential dichotomy and the $n \times n$ coefficient matrix $B(t)$ is continuous in $\mathbb{R}$. If*

$$\underline{=}\sup_{t\in\mathbb{R}} \|B(t)\| < \alpha/(4K^2),$$

then the perturbed system

$$\dot{y} = (A(t) + B(t))y \tag{10}$$

also has an exponential dichotomy

$$\begin{array}{ll} \|\Phi(t)Q\Phi^{-1}(s)\| \le (5/2)K^2e^{-(\alpha-2K\beta)(t-s)}, & t \ge s, \\ \|\Phi(t)(I-Q)\Phi^{-1}(s)\| \le (5/2)K^2e^{-(\alpha-2K\beta)(s-t)}, & s \ge t, \end{array}$$

where $\Phi(t)$ is the fundamental matrix of (10) with the initial condition $\Phi(0) = I$ and the projection Q has the same null space as the projection P.

It is well known that the exponential dichotomy implies the existence of a unique bounded solution, which is almost periodic for almost periodic systems. Here, the statement of the next lemma is from [5], see also [8, 7].

Lemma 2.2. *For the non–homogeneous equation*

$$\dot{x} = A(t)x + f(t),$$

where $A(t)$ and $f(t)$ are almost periodic functions, if the corresponding homogeneous equation $\dot{x} = A(t)x$ has an exponential dichotomy on $\mathbb{R}$, then there exits a unique almost periodic solution ψ of that non–homogeneous equation which satisfies $m(\psi) \subset m(A(t)x + f(t))$.

We recall the next lemma due to Bibikov [3], which plays a key role in the proof of Theorem 1.2.

Lemma 2.3. *Denote by $H_{k,n}(\mathbb{C}^n)$ the linear space of the n–dimensional vector valued homogeneous polynomials in n variables of degree k with complex coefficients. Let A be an $n\times n$ matrix with eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_n)$. Define a linear operator L_k^A on $H_{k,n}(\mathbb{C}^n)$ as follows,*

$$L_k^A h = Ah - \frac{\partial h}{\partial x} Ax,$$

for $h(x) \in H_{k,n}(\mathbb{C}^n)$. Here $\partial h/\partial x$ denotes the Jacobian matrix of h with respect to the variable x. Then the set of eigenvalues of L_k^A is

$$\left\{ \lambda_j - \sum_{i=1}^{n} k_i\lambda_i : k_i \in \mathbb{Z}^+, \sum_{i=1}^{n} k_i \geq 2, 1 \leq j \leq n \right\}.$$

The following is a strong version of the Gronwall type integral inequality, which comes from a result of Sardarly (1965), and its proof can be found in [2].

Lemma 2.4. *Let $u(t)$, $a(t)$, $b(t)$ and $q(t)$ be continuous functions in $J = [\alpha, \beta]$, let $c(t, s)$ be a continuous function for $\alpha \leq s \leq t \leq \beta$, let $b(t)$ and $q(t)$ be non–negative in J and suppose that*

$$u(t) \leq a(t) + \int_\alpha^t (q(t)b(s)u(s) + c(t, s))ds, \text{ for all } t \in J.$$

Then for all $t \in J$ we have that

$$u(t) \leq a(t) + \int_\alpha^t c(t, s)ds + q(t) \int_\alpha^t b(s) \left[a(s) + \int_\alpha^s c(s, \tau)d\tau \right] e^{\int_s^t b(\tau)q(\tau)d\tau} ds.$$

3. **Proof of Theorem 1.1.**

The next two lemmas support the proof of Theorem 1.1.

Consider the linear system

$$\dot{x} = \varepsilon(A + \hat{A}(\varepsilon, t))x, \tag{11}$$

where A is in the Jordan normal form, $\hat{A}$ is almost periodic in the variable t uniformly for the variable ε and $\hat{A}(\varepsilon, t) \to 0$ uniformly as $\varepsilon \to 0$. The following one mainly follows Proposition 1 in [5] (pp.42).

Lemma 3.1. *There exists an invertible transformation $y = (I + S(t, \varepsilon))x$, which turns system (11) into the block diagonal form with respect to the different parts of $(\mathrm{re}\lambda_1, \mathrm{re}\lambda_2, \cdots, \mathrm{re}\lambda_n)$. Here $S(t, \varepsilon)$ is almost periodic in t uniformly for ε, $m(S) \subseteq m(\hat{A})$ and $S \to 0$ uniformly for $t \in \mathbb{R}$ as $\varepsilon \to 0$.*

Proof. Note again that without loss of generality, we can assume $\mathrm{re}\lambda_1 \le \cdots \le \mathrm{re}\lambda_n$. First we consider the case that there exists p such that $\mathrm{re}\lambda_p < 0 < \mathrm{re}\lambda_{p+1}$. Denote two matrix operator M_1 and M_2 by

$$M_1 = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & M_{12} \\ M_{21} & 0 \end{pmatrix}$$

for

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix},$$

where M_{11} and M_{22} are of $p \times p$ and $(n-p) \times (n-p)$, respectively. Then let $P = \mathrm{diag}(I_p, 0)$ be the projection with the identity matrix I_p of $\mathbb{C}^p$.

If the change $x = (I + S(t,\varepsilon))y$ turning system (11) into the form

$$\dot{y} = \varepsilon(A + \{\hat{A}(t,\varepsilon)(I + S(t,\varepsilon))\}_1)y,$$

then S admits

$$\frac{dS}{dt} = \varepsilon(AS - SA + \{\hat{A}(t,\varepsilon)(I+S)\}_2 - S\{\hat{A}(\varepsilon,t)(I+S)\}_1). \tag{12}$$

Consider the operator equation

$$TS = S \tag{13}$$

with

$$\begin{aligned} TS(t) = \varepsilon \int_{-\infty}^{t} & e^{\varepsilon At} P e^{-\varepsilon Au} (I - S(u)) \hat{A}(u) (I + S(u)) e^{\varepsilon Au} (I-P) e^{-\varepsilon At} du \\ -\varepsilon \int_{t}^{\infty} & e^{\varepsilon At} (I-P) e^{-\varepsilon Au} (I - S(u)) \hat{A}(u) (I + H(u)) e^{\varepsilon Au} P e^{-\varepsilon At} du \end{aligned}$$

for $t \in \mathbb{R}$. Here for the simplicity of notations, we omit the parameter ε in function S and $\hat{A}$. Set $\mu = \min_i\{|\mathrm{re}\lambda_i|\}$ and use the matrix norm $||M|| = \sup_{t\in\mathbb{R}} ||M(t)||$. When $||S|| < 1/2$ and $||S'|| < 1/2$, we can obtain that

$$||TS|| \le \varepsilon n^2 (\int_{-\infty}^{t} e^{-2\varepsilon\mu(t-u)} ||\hat{A}|| du + \int_{t}^{\infty} e^{2\varepsilon\mu(u-t)} ||\hat{A}|| du) = n^2 \mu^{-1} ||\hat{A}||,$$

and

$$||T(S - S')|| \le 4n^2\mu^{-1} ||S - S'|| ||\hat{A}||$$

by the fact

$$\begin{aligned} & (I-S)\hat{A}(I+S) - (I-S')\hat{A}(I+S') \\ = & (S'-S)\hat{A} - \hat{A}(S'-S) + (S'-S)\hat{A}S' + S\hat{A}(S'-S). \end{aligned}$$

So if we make $8n^2\mu^{-1}||\hat{A}|| < 1$, then T maps all S satisfying $||S|| < 1/2$ into itself and is a strong contraction. Thus we get the unique solution S such that $TS = S$ by Contracting Mapping Principle, which is also the solution of

$$\frac{dS}{dt} = \hat{A}(t)S - S\hat{A}(t) + \{(I-S)\hat{A}(I+S)\}_2.$$

Note that S always satisfies

$$PSP = 0, \quad (I-P)S(I-P) = 0,$$

which implies

$$S\{\hat{A}(\varepsilon,t)(I+S)\}_1 = \{S\hat{A}(\varepsilon,t)(I+S)\}_2.$$

That is to say, that S is also the solution of equation (12).

At last, we verify the almost periodicity of the solution. Arbitrarily choosing $\{h_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}$ there is a subsequence $\{h_{k_l}\}$ such that $\hat{A}(\varepsilon, t+h_{k_l})$ is convergent uniformly for ε and $t \in \mathbb{R}$ as $l \to \infty$. Then using the notation $\hat{A}_l(\cdot) = \hat{A}(\varepsilon, \cdot + h_{k_l})$ and make $8n^2\mu^{-1}||\hat{A}_l|| < 1$ for any l, we can find the unique solution $T_lS_l = S_l$ for the fixed l, where $||S_l|| < 1/2$ for all l, T_l is similar as equation (13) and the l in T_l refers to $\hat{A}_l$ instead of $\hat{A}$. By the uniqueness of the solution, we know that $S_l(\cdot) = S(\varepsilon, \cdot + h_{k_l})$. Moreover, we obtain that

$$\begin{aligned}
||S_l - S_{l+p}|| &= ||T_lS_l - T_{l+p}S_{l+p}|| \\
&\le \frac{9n^2}{4\mu}||\hat{A}_l - \hat{A}_{l+p}|| + \frac{3n^2(||\hat{A}_l|| + ||\hat{A}_{l+p}||)}{2\mu}||S_l - S_{l+p}|| \\
&\le \frac{9n^2}{4\mu}||\hat{A}_l - \hat{A}_{l+p}|| + \frac{3}{8}||S_l - S_{l+p}||.
\end{aligned}$$

So it admits $||S_l - S_{l+p}|| \le C||\hat{A}_l - \hat{A}_{l+p}||$ with the constant C independent of l and p, which implies the result $m(S) \subseteq m(\hat{A})$.

Anyway, if we have the gap between two real parts of eigenvalues, i.e., $\mathrm{re}\lambda_p < \mathrm{re}\lambda_{p+1}$, considering the $\varepsilon\lambda_0$-shift system

$$\dot{x} = \varepsilon(A + \hat{A}(\varepsilon, t) - \lambda_0 I)x,$$

with $\lambda_0 = (\mathrm{re}\lambda_p + \mathrm{re}\lambda_{p+1})/2$ instead of (11) we can turn it into the two block diagonal form by the similar arguments. Then so do system (11) with the same transformation. Now repeating such steps with respect to the different real parts of eigenvalues, we can get the final block diagonal form. □

Lemma 3.2. *Let $\varepsilon_1 > 0$, there exists a function $\varepsilon u(z,t,\varepsilon) \in \mathcal{F}(U_\delta, [0,\varepsilon_1))$, $\varepsilon u \to 0$ as $\varepsilon \to 0$ uniformly on $U_\delta \times \mathbb{R}$, such that the transformation of variables $y = z + \varepsilon u(z,t,\varepsilon)$ is invertible for $0 < \varepsilon \le \varepsilon_1$ and transforms system (5) into*

$$\dot{z} = \varepsilon F_0(z) + \varepsilon \widetilde{F}(z,t,\varepsilon), \tag{14}$$

where

$$F_0(z) = \lim_{T\to\infty} \frac{1}{T}\int_0^T F(z,t,0)\ dt$$

is analytic in $z \in U_\delta$ and $F_0(0) = 0$, $\widetilde{F} \in \mathcal{F}(U_\delta, [0,\varepsilon_1))$, $\widetilde{F}(z,t,\varepsilon) \to 0$ as $\varepsilon \to 0$ uniformly for $(z,t) \in U_\delta \times \mathbb{R}$ and $\widetilde{F}(0,t,\varepsilon) \equiv 0$. Furthermore, $m(\widetilde{F}) \subset m(F)$.

Proof. For fixed $(x,t) \in U_\delta \times \mathbb{R}$, we define

$$u(z,t,\varepsilon) = \int_{-\infty}^t e^{-\varepsilon(t-s)} Z(z,s)\ ds,$$

where $Z(z,t) = F(z,t,0) - F_0(z)$. By Lemmas 2.1 and 2.2, it can be seen as the unique almost periodic solution of the system

$$\dot{x} = -\varepsilon x + Z(z,t)$$

for a fixed $z \in U_\delta$. So $u(z,t,\varepsilon)$ is analytic in $z \in U_\delta$ for fixed $(t,\varepsilon) \in \mathbb{R} \times [0,\varepsilon_1)$ and so also F_0 and $\widetilde{F}$ are analytic. The rest of the proof can be seen in [5, 7]. □

Proof of Theorem 1.1. By Lemma 3.2 and 3.1, we can turn the original system into

$$\dot{z} = \varepsilon\widetilde{A}(t,\varepsilon)z + \varepsilon G(z,t,\varepsilon),$$

where $\widetilde{A}(t,\varepsilon) \to A = \partial f_0(x_0)/\partial x$ uniformly for $t \in \mathbb{R}$ as $\varepsilon \to 0$, A is in the Jordan normal form, $\widetilde{A}$ is in the block diagonal form with respect to the different real parts of eigenvalues of A and $G = O(\|z\|^2)$ as $z \to 0$ uniformly for $t \in \mathbb{R}$ and fixed $\varepsilon \in (0, \varepsilon_0)$.

By induction assumptions, we can assume that

$$G(z,t,\varepsilon) = NF_{j-1}(z,t,\varepsilon) + G_j(z,t,\varepsilon) + O(\|z\|^{j+1}),$$

where NF_{j-1} is a polynomial of degree $j-1$ with respect to z and contains terms $x^\alpha e_\beta$ only for $2 \le |\alpha| \le j-1$ and $\mathrm{re}(\sum_k \alpha_k\lambda_k - \lambda_\beta) = 0$, G_j is a homogeneous one of degree j with respect to the same variable. Now we do the j–th substitution $z = x + h(x,t,\varepsilon)$, where h is the homogeneous polynomial of degree j but with respect to the variable x, then the transformed system becomes

$$\dot{x} = \varepsilon\widetilde{A}(t,\varepsilon)x + \varepsilon\widetilde{A}(t,\varepsilon)h - \varepsilon\frac{\partial h}{\partial x}\widetilde{A}(t,\varepsilon)x + \varepsilon NF_{j-1} + \varepsilon\hat{G}_j - \frac{\partial h}{\partial t} + O(\|x\|^{j+1}),$$

where $\hat{G}_j$ is the homogenous polynomial of degree j of the following expression

$$G_j + NF_{j-1}(x+h,t,\varepsilon) - \frac{\partial h}{\partial x}NF_{j-1}.$$

So by comparing the j-th order terms, we get that

$$\frac{dh(t,\varepsilon)}{dt} = \varepsilon L_j^{\widetilde{A}}(t,\varepsilon) + \varepsilon\hat{G}_j(t,\varepsilon), \tag{15}$$

where the ε–depending linear operator $L_j^{\widetilde{A}}(t,\varepsilon)$ on $H_{j,n}(\mathbb{C}^n)$ is defined as follows

$$L_j^{\widetilde{A}} : h(x) \mapsto \widetilde{A}(t,\varepsilon)h(x) - \frac{\partial h(x)}{\partial x}\widetilde{A}(t,\varepsilon)x.$$

On one hand, we show that the representation of the operator $L_j^{\widetilde{A}}$ has the block diagonal form. Now $\widetilde{A} = (a_{\mu\nu})$ is in the block diagonal form, i.e., $a_{\mu\nu} = 0$ if $\mathrm{re}\lambda_\mu \ne \mathrm{re}\lambda_\nu$. For the term $x^\alpha e_\beta$, by careful computation we obtain that

$$L_j^{\widetilde{A}} x^\alpha e_\beta = \sum_{\mu=1}^n a_{\mu\beta}x^\alpha e_\mu - \left(\sum_{\nu=1}^n\sum_{\mu=1}^n \alpha_\mu a_{\mu\nu}x^{\alpha-e_\mu}x_\nu\right)e_\beta,$$

where $x^\beta = x_1^{\beta_1}x_2^{\beta_2}\cdots x_n^{\beta_n}$ and e_β is the unit vector with the β-th component 1. Then for typical terms $a_{\mu\beta}x^\alpha e_\mu$ and $\alpha_\mu a_{\mu\nu}x^{\alpha-e_\mu}x_\nu e_\beta$, we calculate

$$\sum_{k=1}^n \alpha_k\mathrm{re}\lambda_k - \mathrm{re}\lambda_\mu = \sum_{k=1}^n \alpha_k\mathrm{re}\lambda_k - \mathrm{re}\lambda_\beta$$

and

$$\sum_{k=1}^n \alpha_k\mathrm{re}\lambda_k - \mathrm{re}\lambda_\mu + \mathrm{re}\lambda_\nu - \mathrm{re}\lambda_\beta = \sum_{k=1}^n \alpha_k\mathrm{re}\lambda_k - \mathrm{re}\lambda_\beta,$$

for $a_{\mu\beta} \ne 0$ and $a_{\mu\nu} \ne 0$. That is to say, by the different values of $\sum_{k=1}^n \alpha_k\mathrm{re}\lambda_k - \mathrm{re}\lambda_\beta$ the representation of the operator $L_j^{\widetilde{A}}$ is in the block diagonal form. Specially making the space decomposition $H_{j,n}(\mathbb{C}^n) = V_1 \oplus V_2$ together with

$$V_1 = \left\{x^\alpha e_\beta \mid \mathrm{re}(\sum_{k=1}^n \alpha_k\lambda_k - \lambda_\beta) \ne 0, \quad |\alpha| = j\right\}$$

and

$$V_2 = \left\{ x^\alpha e_\beta \mid \mathrm{re}(\sum_{k=1}^{n} \alpha_k \lambda_k - \lambda_\beta) = 0, \quad |\alpha| = j \right\},$$

system (15) has the precise form

$$\frac{dh^{(i)}(t,\varepsilon)}{dt} = \varepsilon L_i(t,\varepsilon) + \varepsilon \hat{G}_j^{(i)}(t,\varepsilon)$$

for $i = 1$ and 2. Here

$$L_j^{\widetilde{A}}\Big|_{V_i} = L_i(t,\varepsilon), \quad h|_{V_i} = h^{(i)}, \quad \hat{G}|_{V_i} = \hat{G}_j^{(i)}.$$

On the other hand, we have that $L_j^{\widetilde{A}} \to L_j^A$ uniformly for $t \in \mathbb{R}$ and $\varepsilon \in (0, \varepsilon_0)$. Make $L_j^A = \mathrm{diag}(\hat{L}_1, \hat{L}_2)$ for $L_j^A|_{V_i} = \hat{L}_i$. Thus we set $h^{(2)} = 0$, so all terms in V_2 of degree j are left. By Lemma 2.3 and the induction assumptions, we know that the real parts of eigenvalues of the linear operator $\hat{L}_1$ are different from zero, so the system

$$\frac{dh^{(1)}}{dt} = L_1(t,\varepsilon) h^{(1)}$$

admits an exponential dichotomy. Then doing the time scaling $u = \varepsilon t$, we get the new system

$$\frac{d\overline{h}^{(1)}}{du} = L_1(u/\varepsilon, \varepsilon)\overline{h}^{(1)} + \hat{G}_j^{(1)}(u/\varepsilon, \varepsilon).$$

This system has a unique almost periodic solution $\overline{h}^{(1)}(\varepsilon, u)$ and the module $m(\overline{h}^{(1)}(\varepsilon,u)) \subset m\left(L_1(u/\varepsilon,\varepsilon)\overline{h}^{(1)} + \hat{G}_j^{(1)}(u/\varepsilon,\varepsilon)\right)$ for $0 < \varepsilon < \varepsilon_0$. That is, $h^{(1)}(\varepsilon, t)$ $= \overline{h}^{(1)}(\varepsilon, \varepsilon t)$ solves all terms in V_1 of degree j and $m(h(\varepsilon,t)) \subset m\left(L_j^{\widetilde{A}}(t,\varepsilon)h\right.$ $\left.+ G_j(t,\varepsilon)\right)$ for $0 < \varepsilon \leq \varepsilon_0$. □

4. **Proof of Theorem 1.2.** The next five lemmas prepare the proof of Theorem 1.2. We specially note that the block diagonal condition is not necessary here, so proofs of Theorem 1.1 and 1.2 are independent.

Now using Taylor expansion in z system (14) can be written as

$$\dot{z} = \varepsilon \widetilde{A}(t,\varepsilon) z + \varepsilon G(z,t,\varepsilon), \tag{16}$$

where $\widetilde{A}(t,\varepsilon) \to A = \partial f_0(x_0)/\partial x$ uniformly for $t \in \mathbb{R}$ as $\varepsilon \to 0$ and $G = O(\|z\|^2)$ as $z \to 0$ uniformly for $t \in \mathbb{R}$ and fixed $\varepsilon \in (0, \varepsilon_0)$.

Lemma 4.1. *If the real parts of $\lambda(A)$ are non–resonant until the l–th order, then there exists a coordinate substitution*

$$z = x + h(x,t,\varepsilon),$$

where $h(x,t,\varepsilon)$ is a polynomial of degree l with respect to x and the coefficients of x are all almost periodic functions in t, under which system (16) can be changed into

$$\dot{x} = \varepsilon \widetilde{A}(t,\varepsilon) x + \varepsilon R_{l+1}(x,t,\varepsilon),$$

where $R_{l+1} = O(\|x\|^{l+1})$ as $x \to 0$ uniformly for $t \in \mathbb{R}$ and fixed ε.

Proof. By induction assumptions, we have got the normal form of system (16) till the $(j-1)$–th order; i.e., in system (16) we have

$$G(z,t,\varepsilon) = G_j(z,t,\varepsilon) + O(\|z\|^{j+1}),$$

where G_j is a homogeneous polynomial of degree j with respect to z. Now we do the j–th substitution $z = x + h(x,t,\varepsilon)$, where h is the homogeneous polynomial of degree j but with respect to the variable x, then the transformed system becomes

$$\dot{x} = \varepsilon\widetilde{A}(t,\varepsilon)x + \varepsilon\widetilde{A}(t,\varepsilon)h - \varepsilon\frac{\partial h}{\partial x}\widetilde{A}(t,\varepsilon)x + \varepsilon G_j - \frac{\partial h}{\partial t} + O(\|x\|^{j+1}). \tag{17}$$

Note that the coefficients of x^k are all almost periodic functions of t for a fixed ε, where $k \in \mathbb{Z}_+^n$. By comparing the terms of degree j with respect to the variable x in system (17), we can obtain

$$\frac{dh(t,\varepsilon)}{dt} = \varepsilon L_j^{\widetilde{A}}(t,\varepsilon)h(t,\varepsilon) + \varepsilon G_j(t,\varepsilon), \tag{18}$$

where the ε–depending linear operator $L_j^{\widetilde{A}}(t,\varepsilon)$ on $H_{j,n}(\mathbb{C}^n)$ is defined as follows

$$L_j^{\widetilde{A}} : h(x) \mapsto \widetilde{A}(t,\varepsilon)h(x) - \frac{\partial h(x)}{\partial x}\widetilde{A}(t,\varepsilon)x.$$

Let $\varepsilon \to 0$, then we can get another linear operator L_j^A defined on $H_{j,n}(\mathbb{C}^n)$ as

$$L_j^A : h(x) \mapsto Ah(x) - \frac{\partial h(x)}{\partial x}Ax.$$

Since $\widetilde{A}(t,\varepsilon) \to A$ uniformly for $t \in \mathbb{R}$ and a fixed ε and every entry of the matrix representation of the linear operator $L_j^{\widetilde{A}}(\varepsilon,t)$ is the multiplication and addition of the entries of $\widetilde{A}(t,\varepsilon)$, we have the uniform convergence

$$\sup_{\mathbb{R}} \|L_j^{\widetilde{A}}(\varepsilon,t) - L_j^A\| \to 0, \quad \text{as} \quad \varepsilon \to 0.$$

By Lemma 2.3 and the induction assumptions, we know that the real parts of eigenvalues of the linear operator L_k^A are different from zero, so the system

$$\frac{dh}{dt} = L_j^A h$$

admits an exponential dichotomy together with positive constants K and α. Thus there exists $\varepsilon_0 > 0$ such that

$$\sup_{\mathbb{R}} \|L_j^{\widetilde{A}}(\varepsilon,t) - L_j^A\| \le \alpha/(4K^2), \quad 0 \le \varepsilon \le \varepsilon_0.$$

Doing a time scaling $u = \varepsilon t$ in system (18), we get the new system

$$\frac{d\overline{h}}{du} = L_j^{\widetilde{A}}(u/\varepsilon,\varepsilon)\overline{h} + G_j(u/\varepsilon,\varepsilon).$$

This system has a unique almost periodic solution $\overline{h}(\varepsilon,u)$ and the module $m(\overline{h}(\varepsilon,u)) \subset m\left(L_j^{\widetilde{A}}(u/\varepsilon,\varepsilon)\overline{h} + G_j(u/\varepsilon,\varepsilon)\right)$ for $0 < \varepsilon < \varepsilon_0$. That is, $h(\varepsilon,t) = \overline{h}(\varepsilon,\varepsilon t)$ is the solution of system (18) and $m(h(\varepsilon,t)) \subset m\left(L_j^{\overline{h}}(t,\varepsilon)\overline{h} + G_j(t,\varepsilon)\right)$ for $0 < \varepsilon \le \varepsilon_0$. $\square$

Let $v(x, y, \varepsilon)$ and $w(x, y, \varepsilon)$ be two families of ε–depending continuous skew–product vector fields defined on $\widetilde{D} = D \times \mathbb{R} \subset \mathbb{C}^n \times \mathbb{R}$ and analytic in $x \in D$ for fixed y and ε, where $\varepsilon \in (0, \varepsilon_0)$ is the parameter. More precisely, $v = (v_1(x, y, \varepsilon), v_2(y, \varepsilon))$ and $w = (w_1(x, y, \varepsilon), v_2(y, \varepsilon))$, where $v_2 \in C(\mathbb{R} \times (0, \varepsilon_0), \mathbb{R})$, v_1 and $w_1 \in C(\widetilde{D} \times (0, \varepsilon_0), \mathbb{C}^n)$ are analytic in $x \in D$ for fixed y and ε. Moreover, we say v and w are *analytically equivalent* if there exists an ε–depending coordinate substitution $z = u(x, y, \varepsilon)$, which changes one vector field into the other, where $u \in C(\widetilde{D})$, $\varepsilon \in (0, \varepsilon_0)$ is the parameter and u is analytic in $x \in D$ for fixed y and ε. Let

$$R = w - v, \quad v_s = v + sR,$$

then $v_0 = v$, $v_1 = w$. Consider the ε–depending vector field on the $\widetilde{D} \times \Delta$

$$V(x, y, \varepsilon, s) = (v_s(x, y, \varepsilon), 0), \quad s \in \Delta,$$

where $\Delta = \{x \in \mathbb{C} : \|x\| \leq 2\}$.

Lemma 4.2. *Assume there exists an ε–depending vector field*

$$U(x, y, \varepsilon, s) = (h(x, y, \varepsilon, s), 1), \quad (x, y, s) \in \widetilde{D} \times \Delta,$$

satisfying

$$[h, v_s] = R, \tag{19}$$

where $h \in C(\widetilde{D} \times \Delta)$ is analytic on $D \times \Delta$ for fixed y and ε, $\varepsilon \in (0, \varepsilon_0)$ is the parameter and $[\cdot, \cdot]$ is the Lie bracket taken with respect to the variables x and y. Let $\widetilde{D}_0$, $\widetilde{D}_1 \subset \widetilde{D}$ be two domains satisfying

$$g_U^1(\widetilde{D}_0 \times \{0\}) = \widetilde{D}_1 \times \{1\},$$

where g_U^1 is the time–1 map defined by the vector field U, then the two vector field $v|_{\widetilde{D}_0}$ and $w|_{\widetilde{D}_1}$ are analytically equivalent.

Proof. Note that the set $\{s = constant\}$ is invariant under the vector field V. Moreover, the homological equation (19) implies $[U, V] \equiv 0$, where $[\cdot, \cdot]$ is the Lie bracket taken with respect to the variables x, y and s. Together with the condition that g_U^1 maps $\widetilde{D}_0 \times \{0\}$ into $\widetilde{D}_1 \times \{1\}$, it follows

$$g_U^1 \circ g^t_{V|_{s=0}} = g^t_{V|_{s=1}} \circ g_U^1.$$

Thus we complete the proof by the differentiability on the initial values and the fact that $V|_{s=0} = (v, 0)$ and $V|_{s=1} = (w, 0)$. □

Lemma 4.3. *The function*

$$h(x, y, \varepsilon, s) = -\int_0^\infty X^{-1}(t; x, y, \varepsilon, s) \cdot R \circ g^t(x, y, \varepsilon, s)dt$$

is a formal solution of the homological equation (19), where $g^t(x, y, \varepsilon, s)$ is the time–t map defined by the vector field v_s and the matrix solution $X(t; x, y, \varepsilon, s)$ is defined as

$$X(t; x, y, \varepsilon, s) = \frac{\partial g^t(x, y, \varepsilon, s)}{\partial(x, y)}.$$

Proof. For simplicity of notation, we fix ε, s and denote $\overline{x} = (x, y)$, $g^t_{v_s}\overline{x} = g^t(x, y, \varepsilon, s)$ and $X(t; \overline{x}) = X(t; x, y, \varepsilon, s)$. Let $h^\tau := (g^\tau_{v_s})_* h$ which is defined as

$$h^\tau(\overline{x}) = (X(t, \overline{x})h) \circ g^{-\tau}_{v_s}(\overline{x}) = X(\tau; g^{-\tau}_{v_s}\overline{x})h(g^{-\tau}_{v_s}\overline{x}).$$

Since we have

$$\begin{aligned} g_{v_s}^{\tau}\overline{x} &= x + \tau v_s(\overline{x}) + o(\tau), \\ X(\tau;\overline{x}) &= I + \tau\frac{\partial v_s(\overline{x})}{\partial \overline{x}} + o(\tau). \end{aligned}$$

It follows that

$$\begin{aligned} h^{\tau}(\overline{x}) &= h \circ g_{v_s}^{-t}\overline{x} + \tau(\frac{\partial v_s}{\partial \overline{x}}h) \circ g_{v_s}^{-t}\overline{x} + o(\tau) \\ &= h(\overline{x}) - \tau\frac{\partial h(\overline{x})}{\partial \overline{x}}v_s(\overline{x}) + \tau\frac{\partial v_s(\overline{x})}{\partial \overline{x}}h(\overline{x}) + o(\tau), \end{aligned}$$

which means

$$\left.\frac{dh^{\tau}}{d\tau}\right|_{\tau=0} = [v_s, h].$$

Again by definition, we have

$$\begin{aligned} h^{\tau} &= -\int_0^{\infty} X^{-1}(-\tau;\overline{x})X^{-1}(t;g_{v_s}^{-\tau}\overline{x})R(g_{v_s}^{-\tau+t}\overline{x})\ dt \\ &= -\int_0^{\infty} X(t-\tau;\overline{x})R(g_{v_s}^{t-\tau}\overline{x})\ dt \\ &= -\int_{-\tau}^{\infty} X^{-1}(t;\overline{x})R(g_{v_s}^{t}\overline{x})\ dt. \end{aligned}$$

So

$$\left.\frac{dh^{\tau}}{d\tau}\right|_{\tau=0} = -X^{-1}(0,\overline{x})R(g_{v_s}^{0}\overline{x}) = -R(\overline{x}).$$

This completes the proof of this lemma. $\square$

Let $\lambda(A) = (\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$ be the eigenvalues of the matrix $A = \partial f_0(x_0)/\partial x$. Assume $\mathrm{re}\lambda_1 \leq \ldots \leq \mathrm{re}\lambda_n < 0$. Now we consider the following system

$$\dot{x} = \varepsilon\widetilde{A}(t,\varepsilon)x + \varepsilon\widetilde{G}(x,t,\varepsilon) + s\varepsilon r(x,t,\varepsilon), \tag{20}$$

where $\widetilde{A}(t,\varepsilon) \to A$ uniformly for $t \in \mathbb{R}$ as $\varepsilon \to 0$ and it is almost periodic in t for a fixed ε, $\widetilde{G}$ and $r \in \mathcal{F}(U_\delta, [0,\varepsilon_0))$, $\widetilde{G}(z,t,0) = r(z,t,0) = 0$. Moreover, $\widetilde{G}$ is a polynomial of degree $\widetilde{d} = [\mathrm{re}\lambda_1/\mathrm{re}\lambda_n]$ with respect to x, $\widetilde{G}(x,t,\varepsilon) = Jet_{x=0}^{\widetilde{d}}G(x,t,\varepsilon)$, $r(x,t,\varepsilon) = G(x,t,\varepsilon) - \widetilde{G}(x,t,\varepsilon)$ and $s \in \Delta = \{x \in \mathbb{C} : \|x\| \leq 2\}$.

Lemma 4.4. *Let $W(x,t,\varepsilon,s) = \widetilde{G}(x,t,\varepsilon) + sr(x,t,\varepsilon)$. Consider the following system*

$$\frac{dx}{dt} = \varepsilon\widetilde{A}(t+y,\varepsilon)x + \varepsilon W(x,t+y,\varepsilon,s), \tag{21}$$

where y is a real parameter, $\widetilde{A}$, $\widetilde{G}$ and r are just defined before the statement of the lemma. Let $\mathcal{G}^t(x,y,\varepsilon)$ be the solution of system (21) with the initial condition $\mathcal{G}^0(x,y,\varepsilon) = x$. Then there exists $\varepsilon_1 > 0$ and $\delta_1 > 0$ independent of ε such that

$$\mu(\varepsilon_1,\delta_1) = 2K \underline{+}(5/2)K^2\rho$$

with the dichotomy constant K, $\rho = \sup_{U_{\delta_1}\times\mathbb{R}}\|\partial_x W(x,t,\varepsilon,s)\|$ and $\underline{=}\sup_{\mathbb{R}}\|\widetilde{A}(t,\varepsilon) - A\|$ for $0 < \varepsilon < \varepsilon_1$, fulfilling $\mu(\varepsilon_1,\delta_1) \to 0$ as $(\varepsilon_1,\delta_1) \to 0$, and for a fixed $(y,\varepsilon) \in \mathbb{R}\times(0,\varepsilon_1)$ the following statements hold.

(a) $\|r(x,t,\varepsilon)\| \leq C\|x\|^{\widetilde{d}+1}$ *for all* $(x,t) \in U_{\delta_0}\times\mathbb{R}$.
(b) $\|\mathcal{G}(t;x,y,\varepsilon)\| \leq Ce^{\varepsilon(\mathrm{re}\lambda_n+\mu(\varepsilon_1,\delta_1))t}$ *for all* $(t,x) \in [0,\infty)\times U_{\delta_1}$, $s \in \Delta$.
(c) $\|\partial_x^{-1}\mathcal{G}(t;x,y,\varepsilon)\| \leq Ce^{\varepsilon(-\mathrm{re}\lambda_1+\mu(\varepsilon_1,\delta_1))t}$ *for all* $(t,x) \in [0,\infty)\times U_{\delta_1}$, $s \in \Delta$.

Proof. By the Cauchy's integral representation

$$\begin{aligned}\partial_x^k W(x,t,\varepsilon,s): &= \frac{\partial^{|k|}W(x_1,\dots,x_n,t,\varepsilon,s)}{\partial x_1^{k_1}\cdots\partial x_n^{k_n}}\\ &= \frac{k!}{(2\pi\sqrt{-1})^n}\int_\gamma \frac{W(z,t,\varepsilon,s)\;dz}{(z-x)^{k+e}},\end{aligned}$$

where $|k|=\sum_{i=1}^n k_i$, $e=(1,\dots,1)\in\mathbb{Z}_+^n$ and $\gamma=\{z:|z_i|=r-\chi, i=1,\dots,n\}$ for $0<\chi<<1$, $\partial_x^k W(x,t,\varepsilon,s)$ is an almost periodic function in the variable t uniformly for $\|x\|\le\delta_0<\delta$ and $k\in\mathbb{Z}_+^n$. Furthermore, the following norm estimations are valid for $0<\delta_0<\delta/3$

$$\begin{aligned}&\sup_{U_{\delta_0}\times\mathbb{R}}\|\partial_x W(x,t,\varepsilon,s)\|=\rho\le\frac{C_1M}{\delta}\delta_0,\\ &\|r(x,t,\varepsilon)\|\le\frac{C_1M(\widetilde{d}+1)^n(\widetilde{d}+1)!}{\delta^{\widetilde{d}+1}}\|x\|^{\widetilde{d}+1},\quad (x,t)\in U_{\delta_0}\times\mathbb{R},\end{aligned}$$

where C_1 is a constant depending on n, $\sup_{U_\delta\times\mathbb{R}}\|W\|=M<\infty$ and $\partial_x W$ is the Jaccobian matrix of W with respect to the variable x. This completes the proof of statement (a).

Consider the linear part of system (21) $\dot{x}=\varepsilon\widetilde{A}(t+y,\varepsilon)x$. Let $\Phi(t)$ be its fundamental matrix with the initial condition $\Phi(0)=I$ and denote $\Phi(t,s)=\Phi(t)\Phi^{-1}(s)$. Then by Lemma 2.1, we have

$$\|\Phi(t+y,y)\|\le(5/2)K^2e^{\varepsilon(\mathrm{re}\lambda_n+2K\underline{)}t},$$

where $\underline{=}\sup_{\mathbb{R}}\|\widetilde{A}(t,\varepsilon)-A\|<-\mathrm{re}\lambda_n/2K$ for $0<\varepsilon<\varepsilon_2$. Now we can rewrite system (21) as the integral equation

$$\mathcal{G}(t;x,y,\varepsilon)=\Phi(t+y,y)x+\int_0^t\varepsilon\Phi(t+y,v+y)W(\mathcal{G}(v;x,y,\varepsilon),v+y,\varepsilon,s)\;dv,$$

Since $\rho=\sup_{U_{\delta_0}\times\mathbb{R}}\|\partial_x W(x,t,\varepsilon,s)\|$, we have

$$\begin{aligned}&\|\mathcal{G}(t;x,y,\varepsilon)\|\\ &\le(5/2)K^2\left(e^{\varepsilon(\mathrm{re}\lambda_n+2K\underline{)}t}+\varepsilon\rho\int_0^t e^{\varepsilon(\mathrm{re}\lambda_n+2K\underline{)}(t-v)}\|\mathcal{G}(v;x,y,\varepsilon)\|\;dv\right)\\ &\le(5/2)K^2e^{\varepsilon(\mathrm{re}\lambda_n+2K\underline{)}t}+(5/2)K^2\varepsilon\rho\int_0^t\|\mathcal{G}(v;x,y,\varepsilon)\|\;dv.\end{aligned}$$

Then, by Lemma 2.4, the strong type Gronwall inequality, we obtain

$$\|\mathcal{G}(t;x,y,\varepsilon)\|\le(5/2)K^2e^{\varepsilon(\mathrm{re}\lambda_n+2K\pm(5/2)K^2\rho)t},\ \text{for all } t\ge0,\quad s\in\Delta.$$

So this proves statement (b) for $\mu(\varepsilon_1,\delta_1)=2K\pm(5/2)K^2\rho$ and $C_2=(5/2)K^2$.

Now we study the Jaccobian matrix $\partial_x\mathcal{G}(t;x,y,\varepsilon)$. By take derivative with respect to x in system (21), we can get the matrix differential equation

$$\begin{aligned}\frac{d}{dt}\partial_x^{-1}\mathcal{G}(t;x,y,\varepsilon) &= -\varepsilon\partial_x^{-1}\mathcal{G}(t;x,y,\varepsilon)(\widetilde{A}(t+y,\varepsilon)+\\ &\qquad \partial_x W(\mathcal{G}(t;x,y,\varepsilon),t+y,\varepsilon,s)),\end{aligned}$$

which can also be written as the matrix integral equation

$$\begin{aligned}&\partial_x^{-1}\mathcal{G}(t;x,y,\varepsilon)=\Phi(y,t+y)\\ &\quad-\int_0^t\varepsilon\partial_x^{-1}\mathcal{G}(v;x,y,\varepsilon)\partial_x W(\mathcal{G}(v;x,y,\varepsilon,s),v+y,\varepsilon)\Phi(v+y,t+y)\;dv.\end{aligned}$$

Here, $\Phi(s,t)$ can be seen as the fundamental solution of linear system

$$\frac{d}{dt}\Phi(s,t) = -\Phi(s,t)\varepsilon\widetilde{A}(t,\varepsilon).$$

So, again by Lemma 2.1, we have

$$\|\Phi(s+y,t+y)\| \leq (5/2)K^2 e^{\varepsilon(-\mathrm{re}\lambda_1+2K\underline{)}(t-s)}, \quad t \geq s,$$

which means that

$$\begin{aligned}\|\partial_x^{-1}\mathcal{G}(t;x,y,\varepsilon)\| &\leq (5/2)K^2\bigg(e^{\varepsilon(-\mathrm{re}\lambda_1+2K\underline{)}t} \\ &+ \varepsilon\rho\int_0^t e^{\varepsilon(-\mathrm{re}\lambda_1+2K\underline{)}(t-v)}\|\partial_x^{-1}\mathcal{G}(v;x,y,\varepsilon)\|dv\bigg).\end{aligned}$$

Therefore, again using Lemma 2.4, we have

$$\begin{aligned}&\|\partial_x^{-1}\mathcal{G}(t;x,y,\varepsilon)\| \\ &\leq \left((5/2)K^2 + \frac{(5/2)K^2\rho}{-\mathrm{re}\lambda_1 + 2K\underline{+}(5/2)K^2\rho}\right)e^{\varepsilon(-\mathrm{re}\lambda_1+2K\underline{+}(5/2)K^2\rho)t} \\ &= C_3 e^{\varepsilon(-\mathrm{re}\lambda_1+\mu(\varepsilon_1,\delta_1))t}.\end{aligned}$$

Thus taking $C = \max\{C_1 M(\widetilde{d}+1)^n(\widetilde{d}+1)!/\delta^{\widetilde{d}+1}, C_2, C_3\}$, $\varepsilon_1 = \varepsilon_2$ and $\delta_1 = \delta_0$, we get statement (c). $\square$

Lemma 4.5. *Let $f(x,t)$ be a continuous function, which is almost periodic in the variable t uniformly for x in any compact set D and satisfies*

$$\|f(x_1,t) - f(x_2,t)\| \leq L\|x_1 - x_2\|.$$

Consider the non–autonomous system

$$\dot{x} = f(x, t+y),$$

where y is a real parameter. Let $g^t(x,y)$ be the solution with initial condition $g^0(x,y) = x$, then $g^t(x,y)$ is almost periodic in y for a fixed t and $m(g^t) \subset m(f)$.

Proof. The solution g^t satisfies the integral equation

$$g^t(x,y) = x + \int_0^t f(g^s(x,y), s+y)\, ds.$$

Note that

$$\begin{aligned}&A = \textstyle\int_0^t \|f(g^s(x,y+\alpha_n), s+y+\alpha_n) - f(g^s(x,y+\alpha_n), s+y+\alpha_m)\|\, ds \\ &\leq t\,\sup_{(x,t)\in D\times\mathbb{R}}\|f(x,t+\alpha_n) - f(x,t+\alpha_m)\|,\end{aligned}$$

and

$$\begin{aligned}&B = \textstyle\int_0^t \|f(g^s(x,y+\alpha_n), s+y+\alpha_m) - f(g^s(x,y+\alpha_m), s+y+\alpha_m)\|\, ds \\ &\leq L\textstyle\int_0^t \|g^s(x,y+\alpha_n) - g^s(x,y+\alpha_m)\, ds.\end{aligned}$$

Thus we have

$$\|g^t(x,y+\alpha_n) - g^t(x,y+\alpha_m)\| \leq A + B.$$

By Lemma 2.4, we get

$$\|g^t(x,y+\alpha_n) - g^t(x,y+\alpha_m)\| \leq te^{Lt}\sup_{(x,t)\in D\times\mathbb{R}}\|f(x,t+\alpha_n) - f(x,t+\alpha_m)\|,$$

which means $g^t(x,y)$ is almost periodic in the variable y for a fixed t, and by the definition $m(g^t) \subset m(f)$. $\square$

Proof of Theorem 1.2. First we do the change of variables (4), and we get system (5), which can be transformed into system (16) using Lemma 3.2. In order to eliminate the discrepancy of order great than $\widetilde{d}$, we study system (20) instead of system (16). Considering the corresponding autonomous system of system (20) in higher dimension, we get

$$\dot{z} = \varepsilon \widetilde{A}(y,\varepsilon)z + \varepsilon W(z,y,\varepsilon,s), \quad \dot{y} = 1,$$

where W is the function defined in Lemma 4.4. Applying Lemma 4.2 to this system, we get the homological equation (19) for $v_s = (W,0)$, $h = (h_1,h_2)$ and $R = (r,0)$. By Lemma 4.3, it has the formal solution

$$\begin{aligned} h(x,y,\varepsilon,s) &= -\int_0^\infty X^{-1}(t;x,y,\varepsilon,s)\cdot R\circ g^t(x,y,\varepsilon,s)dt \\ &= \begin{pmatrix} -\displaystyle\int_0^\infty \partial_x^{-1}\mathcal{G}(t;x,y,\varepsilon)\cdot r(\mathcal{G}(t;x,y,\varepsilon),t+y,\varepsilon)dt \\ 0 \end{pmatrix}, \end{aligned}$$

where $\mathcal{G}$ is defined in Lemma 4.4. By the norm estimation of Lemma 4.4, we know that the increasing of the norm of the integrand of h is control by a exponential function $Ce^{\varepsilon\eta t}$, $t \geq 0$, where $\eta = (\widetilde{d}+1)\mathrm{re}\lambda_n - \mathrm{re}\lambda_1 + (\widetilde{d}+2)\mu(\varepsilon_1,\delta_1)$. Since $\widetilde{d} = [\mathrm{re}\lambda_n/\mathrm{re}\lambda_1]$, we can choose ε_1 and δ_1 small enough such that $\eta < 0$. So h converges with the maximum norm. Therefore, by the differentiability of the solutions with respect to the initial value and parameter, the time–1 map w of h is analytic in the domain U_{δ_1} with t and ε fixed. Moreover, by the result of Lemma 4.5 and similar arguments, h is an almost periodic function in y uniformly for $x \in U_{\delta_1}$ with fixed ε, and so the same occurs for the time–1 map w.

In addition, when the real parts of $\lambda(A)$ are non–resonant, we can apply Lemma 4.1 to eliminate any finite order of W with respect to the variable x. By the same arguments, we get the final result of the theorem. $\square$

Acknowledgments. The first author is partially supported by NSFC key program of China (no. 11231001). The second author is partially supported by a MINECO/FEDER grant MINECO grant MTM2013-40998-P, an AGAUR grant number 2014SGR-568, the grants FP7-PEOPLE-2012-IRSES 318999 and 316338, and the MINECO/FEDER grant UNAB13-4E-1604. And the third is supported by NSFC for Young Scientists of China (no. 11001047) and NSF of Jiangsu, China (no. BK20131285). Weigu Li and Hao Wu want to thank to CRM and to the Department of Mathematics of the Universitat Autònoma de Barcelona for their support and hospitality during the period in which this paper was written.

REFERENCES

[1] L. Arnold, *Random Dynamical Systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.
[2] D. Bainov and P. Simenov, *Integral Inequalities and Applications*, Kluwer academic publishers, 1992.
[3] Yn. N. Bibikov, *Local Theory of Nonlinear Analytic Ordinary Differential Equations*, Lecture Notes in Math., **702**, Springer–Verlag, 1979.
[4] K. T. Chen, Equivalence and decomposition of vector fields about an elementary critical point, *Am. J. Math.*, **85** (1963), 693–722.

[5] W. A. Coppel, *Dichotomies in Stability Theory*, Lecture Notes in Mathematics, Vol. 629. Springer-Verlag, Berlin-New York, 1978.
[6] H. Dulac, Solution d'un systeme d'equations differentielles dans le voisinage de valeurs singulieres, *Bull. Soc. Math. Fr.,* **40** (1912), 324–383.
[7] A. M. Fink, *Almost Periodic Differential Equations*, Lecture Notes in Math., **377**, Springer–Verlag, 1974.
[8] J. K. Hale, *Ordinary Differential Equations*, Second edition, Robert E. Krieger Publishing Co., Inc., Huntington, N.Y., 1980.
[9] Yu. S. Il'yashenko and S. Yu. Yakovenko, Finitely–smooth normal forms of local families of diffeomorphisms and vector fields, *Russian Math Sur.,* **46** (1991), 1–43.
[10] W. Li and K. Lu, Poincaré theorems for random dynamical systems, *Ergodic Theory Dynam. Systems,* **25** (2005), 1221–1236.
[11] W. Li and K. Lu, Sternberg theorems for random dynamical systems, *Comm. Pure Appl. Math.,* **58** (2005), 941–988.
[12] H. Poincaré, Thesis, 1879, also *Oeuvres I*, *Gauthier Villars,* Paris, (1928), 59–129.
[13] R. J. Sacker and G. R. Sell, A spectral theory for linear differential systems, *J. Diff. Equ.,* **27** (1978), 320–358.
[14] S. Siegmund, Normal forms for nonautonomous difference equations. Advances in difference equations, IV, *Comput. Math. Appl.,* **45** (2003), 1059–1073.
[15] S. Stenberg, Finite Lie groups and the formal aspects of dynamical systems, *J. Math. Mech.,* **10** (1961), 451–474.
[16] F. Takens, Normal forms for certain singularities of vector fields, *An. Inst. Fourier.,* **23** (1973), 163–195.
[17] A. Vanderbauwhede, Center manifolds and their basic properties. An introduction, *Delft Progr. Rep.,* **12** (1988), 57–78.

Received October 2013; revised May 2014.

E-mail address: weigu@math.pku.edu.cn
E-mail address: jllibre@mat.uab.cat
E-mail address: haow@seu.edu.cn