ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
J. Differential Equations 245 (2008) 2958–2978
Journal of Differential Equations
www.elsevier.com/locate/jde

Poincaré type theorems for non-autonomous systems

Hao Wu [a,*], Weigu Li [b,1]

[a] *Department of Mathematics, Southeast University, Nanjing 210096, Jiangsu, China*
[b] *School of Mathematical Sciences, Peking University, Beijing 100871, China*

Received 22 October 2007; revised 8 April 2008

Available online 16 June 2008

Abstract

In this paper we establish analytic equivalence theorems of Poincaré and Poincaré–Dulac type for analytic non-autonomous differential systems based on the dichotomy spectrum of their linear part. As applications of the theorem, normal forms linearize for two illustrative examples.

Keywords: Normal form; Non-autonomous differential system; Poincaré theorem

1. Introduction

In 1892 Poincaré initiated a technique for simplifying a nonlinear system in the neighborhood of a reference solution by a smooth change of coordinates, today called theory of normal form. Naturally, the reference solution is almost exclusively assumed to be a fixed point (sometimes a periodic solution) and fruitful results for autonomous systems have been obtained such as theorem of Poincaré, Seigel, Chen, Takens, etc. [1]. As we all know that the importance of normal forms goes without saying, it also lays foundations for further study of integrality, stability, bifurcation and so on. The main purpose of our work is to extend formal and analytic normal forms from autonomous systems to non-autonomous ones.

* Corresponding author.
E-mail addresses: haow@seu.edu.cn (H. Wu), weigu@math.pku.edu.cn (W. Li).
1 Partially supported by NSF grant No. 10531010 and NNSF of China (No. 10525104).

0022-0396/$ – see front matter
doi:10.1016/j.jde.2008.05.005

Here we state two well-known theorems in detail. First of all let us recall the classical Poincaré type normal form theorems. Consider the autonomous system

$$\dot{x} = Ax + f(x), \tag{1}$$

where A is a constant matrix and $f(x) = O(\|x\|^2)$ as $x \to 0$. Set $\lambda(A) = \{\lambda_i\}_{i=1}^n$ be the set of eigenvalues of the matrix A. Expanding $f(x)$ in the formal power series, i.e. $f(x) = \sum_{i=2}^{\infty} f_i x^i$, then we can obtain Poincaré's formal normal forms.

Assume the matrix A is in the complex Jordan form, then by a formal coordinate transformation $x = y + \sum_{i=2}^{\infty} h_i y^i$, system (1) *can be changed into system*

$$\dot{y} = Ay + \sum_{i=1}^{\infty} g_i y^i,$$

where $g_i = (g_i^1, \dots, g_i^n)$ and $g_i^j = 0$ if $\sum_{i=1}^n p_i \lambda_i - \lambda_j \neq 0$ for $j = 1, \dots, n$. Here, $p = (p_1, \dots, p_n) \in \mathbb{Z}_+^n$ and $\sum_{i=1}^n p_i \geqslant 2$.

In addition, if $f(x)$ is analytic in the origin instead, then we have Poincaré–Dulac's analytic normal forms.

Assume real parts of eigenvalues of $A \in \mathbb{R}^{n\times n}$ are all positive or negative, then system (1) *can be changed into a polynomial system*

$$\dot{y} = Ay + P(y),$$

by an analytic coordinate transformation $x = y + h(y)$ in some neighborhood of the origin.

The proofs of above theorems can be found in many textbooks such as [1,2]. Furthermore, inspirited by Floquet's beautiful lemma for linear periodic systems, mathematicians have made great efforts to construct theoretic frames of normal forms for more general non-autonomous systems, cf. in [3–5]. In this paper, we mainly deal with normal forms for the following general type non-autonomous differential systems

$$\dot{x} = F(x,t) = A(t)x + f(x,t), \quad x \in \mathbb{R}^n, \tag{2}$$

where F is analytic in $O_\rho \times \mathbb{R}$, $A : \mathbb{R} \to \mathbb{R}^{n\times n}$ is continuous and uniformly bounded on $\mathbb{R}$ and $f = O(\|x\|^2)$ as $\|x\| \to 0$. As usual O_ρ denotes the closed ball centering the origin with the radius ρ. Our goal is to generalize above two theorems from (1) to the non-autonomous systems of the form (2) by studying normal forms in the neighborhood of the zero solution.

As the first step to study normal forms, following the traditional way we shall seek the convenient invariants for the linear part

$$\dot{x} = A(t)x, \quad x \in \mathbb{R}^n. \tag{3}$$

In general it turns out not to be the right thing to study the eigenvalues of $A(t)$ as they have little or nothing to do with the asymptotic properties of solutions. Thus we need to find a dynamical formulation to describe spectral objects in terms of the long-term behavior of solutions instead.

Here we choose the dichotomy spectrum (sometimes called continuous spectrum, or S.S. spectrum [6]) which can be defined as follows. We say that system (3) has an *exponential dichotomy* (for short, E.D.) on $\mathbb{R}$, if there exists a projection $P:\mathbb{R}^n \to \mathbb{R}^n$ and positive constants K, α and β such that

$$\left\|\Phi(t)P\Phi^{-1}(s)\right\| \leqslant Ke^{-\alpha(t-s)}, \quad t \geqslant s,$$
$$\left\|\Phi(t)(I-P)\Phi^{-1}(s)\right\| \leqslant Ke^{-\beta(s-t)}, \quad s \geqslant t,$$

where $\Phi(t)$ is the fundamental matrix of system (3) with $\Phi(0) = I$. To see what an E.D. means in the general case it is convenient to consider two basic aspects. One is that it means a subspace of solutions tending to zero uniformly and exponentially as $t \to \infty$ and a complement subspace of solutions tending to infinity uniformly and exponentially as $t \to \infty$; the other is that the 'angle' between these two subspaces remains bounded away from zero, which nearly leads to the block diagonalization of the fundamental matrix.

In brief, the *dichotomy spectrum* of system (3) is the set

$$\Sigma_A = \left\{\gamma \in \mathbb{R}\colon\ \dot{x} = \left(A(t) - \gamma I\right)x \text{ admits no E.D.}\right\}$$

and the resolvent set $\rho_A = \mathbb{R} \setminus \Sigma_A$ is its complement. Furthermore, by works of [5–7], namely Theorems 2 and 3 in next section, the dichotomy spectrum can be represented by the disjoint union of closed intervals $\bigcup_{i=1}^{p}[a_i, b_i]$ and the corresponding block diagonal theorem is also valid especially when A is continuous and bounded on $\mathbb{R}$. Then by extending L. Arnold's work in [3] of formal normal forms for random dynamic systems, we manage to solve each linear non-homogeneous differential equations generated by Poincaré–Dulac-schemes. Thus following the nature way to eliminate some nonlinear terms degree by degree, we obtain the formal normal forms of system (2). Finally, by applying homotopy method we get the analytic normal forms in the generalized Poincaré domain. Here we specially mention that in fact our partial work of formal normal forms covers Siegmund's results in [5] in a new way.

In detail, a linear time-varying change of variables $x = P(t)y$ is said to be a *Lyapunov–Perron transformation* (for short, *LP transformation*) if $P(t)$ is nonsingular for all $t \in \mathbb{R}$ and P, P^{-1} and $\dot{P} = dP/dt$ are uniform bounded in $t \in \mathbb{R}$. Then the non-autonomous system $\dot{x} = F(x,t)$ is called to be *locally analytically equivalent* to the equation $\dot{y} = G(y,t)$ if there exists a coordinate substitution $x = H(y,t) = P(t)y + h(y,t)$ which transforms one equation to the other, where F, G and H is analytic in $O_\rho \times \mathbb{R}$ for some $\rho > 0$, $F(0,t) = G(0,t) = H(0,t) = 0$, $P(t)$ is an LP transformation and $h(\cdot,t) = O(\|\cdot\|^2)$ as $\|\cdot\| \to 0$. At last, let $A(t) = \operatorname{diag}\{A_1(t), \ldots, A_p(t)\}$ be in the block diagonal form and each block $A_i:\mathbb{R} \to \mathbb{R}^{n_i \times n_i}$ corresponds to a spectral interval $[a_i, b_i]$ for $i = 1, \ldots, p$. Then set $s = (s_0, s_1, \ldots, s_p)$, where $s_0 = 1$, $s_p = n$ and $s_j = \sum_{i=1}^{j} n_i$ for $j = 1, \ldots, p$. Thus a multi-valued index map Θ can be defined as follows

$$\Theta:\mathbb{N}_k^n \to \mathbb{N}_k^p,$$
$$\nu \mapsto \tau,$$

where $\mathbb{N}_k^p = \{\tau = (\tau_1, \ldots, \tau_p) \in \mathbb{Z}_+^p\colon |\tau| = k\}$, $|\tau| = |\nu| = k$ and $\tau_i = \sum_{j=s_{i-1}}^{s_i} \nu_j$ for $i = 1, \ldots, p$. Set e_j be the unit vector with the jth component 1. The vector-valued monomial $x^\nu e_j$ is called *resonant monomial* if τ and l satisfy the condition $0 \in [a_l - \sum_{i=1}^p \tau_i b_i, b_l - \sum_{i=1}^p \tau_i a_i]$, where $e_l = \Theta(e_j)$, $\tau = \Theta(\nu)$. This definition accords with Siegmund's for non-autonomous systems in [5] and will be well illustrated in next section.

Our main results may be summarized as follows.

Theorem 1. *Assume the dichotomy spectrum of system* (3) *to be* $\Sigma_A = [a_1, b_1] \cup \cdots \cup [a_p, b_p]$, *where* $a_1 \leqslant b_1 < \cdots < a_p \leqslant b_p$.

(i) *Equivalence in the jet class. If system* (3) *is in the Poincaré domain, i.e.* $a_1 b_p > 0$, *then any non-autonomous system* $\dot{x} = \tilde{F}(x, t)$ *is locally analytically equivalent to system* (2), *where* $\tilde{F}$ *is analytic in* $O_\rho \times \mathbb{R}$ *and* $\tilde{F} - F = O(\|x\|^d)$ *as* $x \to 0$ *with* $d > \max\{a_1/b_p, b_p/a_1\}$.
(ii) *Poincaré type. If the dichotomy spectrum satisfies* $a_1 b_p > 0$ *and is non-resonant, i.e.*

$$0 \notin \left[\sum_{i=1}^p a_i m_i - a_j, \sum_{i=1}^p b_i m_i - b_j\right], \qquad |m| = \sum_{i=1}^p m_i \geqslant 2,$$

then system (2) *is locally analytically equivalent to its linear part equation with respect to* (*for short, w.r.t.*) x.
(iii) *Poincaré–Dulac type. If* $a_1 b_p > 0$ *and* $A(t)$ *is block diagonal w.r.t. the spectral interval* $[a_i, b_i]$, *then system* (2) *is locally analytically equivalent to system* $\dot{x} = A(t)x + g(x, t)$, *where* g *is a polynomial w.r.t.* x *with the degree* d *not greater than* $\max\{a_1/b_p, b_p/a_1\}$, *which consists resonant monomials only.*
(iv) *Equivalence in the almost periodic case. In addition if* F *is almost periodic in the variable* t *in system* (2), *then so are the transformation and the changed system in* (i), (ii) *and* (iii).

Remark. If system (2) is only continuous but analytic in the variable x for the fixed $t \in \mathbb{R}$, then so are the corresponding equivalence transformation and the changed system in Theorem 1.

In Section 2, we introduce some basic notations and definitions on normal forms. Useful theorems, propositions and lemmas on the spectrum theory for non-autonomous systems, special operators in the tensor space and the Gronwall type inequality are also illustrated. In Sections 3 and 4 we study finite order and analytic normal forms, respectively. In Section 5 the main theorem is proved and in addition we provide two well studied examples as applications.

2. Preliminaries

All useful notations, definitions and technique lemmas are listed in this part. In detail, Theorems 2 and 3 from [5–7] are the foundations of the whole paper. They thoroughly describe properties of dichotomy spectrum for linear non-autonomous systems, by which the basic assumption in Theorem 1 on the spectrum of system (3) is valid and then it is possible to construct normal forms for system (2). Propositions 4, 5 and 6 are preparations for Proposition 9 in next section, which study the special operators generated by Poincaré–Dulac-schemes. Lemma 7 is a Gronwall type inequality and Lemma 8 states the connection between an E.D. and the unique bounded solution.

A nonempty set $\mathcal{M} \subset \mathbb{R}^n \times \mathbb{R}$ is said to be an *integral manifold* of the non-autonomous system (2), if for any point $(x, y) \in \mathcal{M}$, the solution $\varphi(t)$ of the system through (x, y) (namely $\varphi(y) = x$), is such that $(\varphi(t; x, y), t)$ is in $\mathcal{M}$ for all t in the domain of the definition of the solution $\varphi(t; x, y)$. In addition, if for every $y \in \mathbb{R}$, the fiber $\mathcal{M}(y) = \{x \in \mathbb{R}^n,\ (x, y) \in \mathcal{M}\}$ is a linear subspace of $\mathbb{R}^n$, the integral manifold $\mathcal{M}$ is called *linear integral manifold*. Naturally, if $\mathcal{M}_1$ and $\mathcal{M}_2$ are linear integral manifolds, then so are the intersection and the sum

$$\mathcal{M}_1 \cap \mathcal{M}_2 = \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}\colon\ x \in \mathcal{M}_1(y) \cap \mathcal{M}_2(y)\big\},$$
$$\mathcal{M}_1 + \mathcal{M}_2 = \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}\colon\ x \in \mathcal{M}_1(y) + \mathcal{M}_2(y)\big\}.$$

Denote two linear integral manifolds of system (3) by

$$\mathcal{J}_\gamma = \Big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}\colon\ \sup_{t \geqslant 0} e^{-\gamma t} \big\|\varphi(t; x, y)\big\| < \infty\Big\},$$
$$\mathcal{U}_\gamma = \Big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}\colon\ \sup_{t \leqslant 0} e^{-\gamma t} \big\|\varphi(t; x, y)\big\| < \infty\Big\},$$

then the dichotomy spectrum can be characterized as follows. The proof is in [5,6].

Theorem 2. *Assume $A(t)$ is continuous and bounded on $\mathbb{R}$, then the dichotomy spectrum Σ_A of system* (3) *is the disjoint union of p closed intervals* (*called spectrum intervals*) *where $0 < p \leqslant n$, i.e.*

$$\Sigma_A = [a_1, b_1] \cup [a_2, b_2] \cup \cdots \cup [a_p, b_p],$$

where $a_1 \leqslant b_1 < a_2 \leqslant b_2 < \cdots < a_p \leqslant b_p$. Let $b_0 = -\infty$ and $a_{p+1} = \infty$, choose $\gamma_i \in \rho(A)$ with $b_i < \gamma_i < a_{i+1}$ for $i = 0, 1, \ldots, p$. Then for every $i = 1, \ldots, p$ the intersection $\mathcal{M}_i = \mathcal{U}_{\gamma_{i-1}} \cap \mathcal{J}_{\gamma_i}$ is a linear integral manifold with $\dim \mathcal{M}_i > 1$. The linear integral manifolds $\mathcal{M}_i$, $i = 1, \ldots, p$, are called spectral manifolds and they are independent of the choice of γ_i. Moreover,

$$\mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_p = \mathbb{R}^n \times \mathbb{R} \quad (\textit{Whitney sum}),$$

i.e., $\mathcal{M}_i \cap \mathcal{M}_j = \{0\} \times \mathbb{R}$ for $i \neq j$ and $\mathcal{M}_1 + \cdots + \mathcal{M}_p = \mathbb{R}^n \times \mathbb{R}$.

Similar to the block diagonalization of a constant matrix, system (3) can also be changed into the block diagonal form w.r.t. the spectral intervals by an LP transformation. The following theorem is stated and proved in [5,7].

Theorem 3. *Assume $A(t)$ is continuous and bounded on $\mathbb{R}$, then there exists an LP transformation $x = S(t)y$ which turns system* (3) *into*

$$\dot{y} = B(t)y,$$

where $B\colon \mathbb{R} \to \mathbb{R}^{n \times n}$ is in block diagonal form

$$B(t) = \operatorname{diag}\big\{B_1(t), \ldots, B_p(t)\big\} \tag{4}$$

and each block $B_i\colon \mathbb{R} \to \mathbb{R}^{n_i \times n_i}$ corresponds to a spectral interval $[a_i, b_i]$ for $i = 1, \ldots, p$.

Indeed, spectral intervals lead to the block diagonalization of the corresponding linear homogeneous equation, but the inverse is not true because the dichotomy spectrum of each block may contain more than one spectral interval and the intervals of blocks may overlap each other. So we say that p intervals $\bigcup_{i=1}^{p}[a_i, b_i]$, which may overlap, are the *spectrum estimation* of system (3) if there exists an LP transformation which turns it into a block diagonal system and $[a_i, b_i]$ just contains the dichotomy spectrum of the ith block. Here, the word 'just' means that if J is the spectrum dichotomy of the ith block then $J \subset [a_i, b_i]$, $J \not\subset [a_i + \varepsilon, b_i]$ and $J \not\subset [a_i, b_i - \varepsilon]$ for $0 < \varepsilon \ll 1$.

Denote by $V_1, \dots, V_k$ the finite-dimensional real vector spaces of dimensions $\dim V_i = n_i$ for $i = 1, \dots, k$. Then let $V = V_1 \otimes \cdots \otimes V_k$ be their *tensor product*, a vector space of dimension $n = n_1 n_2 \dots n_k$, which is defined to be the vector space $L(V_1^* \times \cdots \times V_k^*, \mathbb{R})$ of k-linear forms on $V_1^* \times \cdots \times V_k^*$. When we restrict our attention to the case $k = 2$, the matrix version of $\Phi_1(t) \otimes \Phi_2(t)$ in $\mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2} \cong \mathbb{R}^{n_1 n_2}$ is called Kronecker product. However, since the Kronecker product cannot maintain block diagonal structures, then we need to find another way to obtain the spectrum estimation of the operator $\Phi_1(t) \otimes \Phi_2(t)$ and the following proposition is an extension of Theorem 5.4.2 in [3].

Proposition 4. *Assume $A_i(t) : \mathbb{R} \to \mathbb{R}^{n_i \times n_i}$ is continuous and bounded on $\mathbb{R}$. Let $\Phi_i(t)$ be the fundamental matrix of system $\dot{x} = A_i(t)x$ with the corresponding spectrum estimation $\Sigma_{A_i} = \bigcup_{j=1}^{p_i}[a_j^{(i)}, b_j^{(i)}]$ for $i = 1, 2$. Then*

(i) *the fundamental matrix $\Phi_1(t) \otimes \Phi_2(t)$ is generated by the non-autonomous linear system*

$$\dot{x} = \big(A_1(t) \otimes I_2 + I_1 \otimes A_2(t)\big)x, \quad x \in \mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2}, \tag{5}$$

(ii) *the spectrum estimation of system* (5) *is*

$$\Sigma = \bigcup_{i,j}\big[a_i^{(1)} + a_j^{(2)}, b_i^{(1)} + b_j^{(2)}\big].$$

Proof. (i) Since $\dot{\Phi}_i(t) = A_i(t)\Phi_i(t)$, then

$$\begin{aligned}\frac{d}{dt}\big(\Phi_1(t) \otimes \Phi_2(t)\big) &= A_1(t)\Phi_1(t) \otimes \Phi_2(t) + \Phi_1(t) \otimes A_2(t)\Phi_2(t) \\ &= \big(A_1(t) \otimes I_2 + I_1 \otimes A_2(t)\big)\big(\Phi_1(t) \otimes \Phi_2(t)\big).\end{aligned}$$

(ii) Without loss of generality, we can assume $\Phi_i(t) = \operatorname{diag}\{\Phi_{i,1}, \dots, \Phi_{i,p_i}\}$ is in the block diagonal form, because LP transformations cannot affect spectrum estimations. Consequently, it naturally implies the $\Phi_i(t)$ invariant splitting of space $\mathbb{R}^{n_i}$, i.e.

$$\mathbb{R}^{n_i} = V_1^i \otimes \cdots \otimes V_{p_i}^i.$$

Let P_j^i be the projection from $\mathbb{R}^{n_i}$ to V_j^i for $j = 1, \dots, p_i$ and $i = 1, 2$, then we have that

$$P_j^i \Phi_i(t) P_k^i = \begin{cases} 0, & j \neq k, \\ \Phi_{i,j}, & j = k. \end{cases}$$

Therefore, it is trivial that

$$\begin{aligned}
&\big(P^1_{j_1} \otimes P^2_{j_2}\big)\big(\Phi_1(t) \otimes \Phi_2(t)\big)\big(P^1_{k_1} \otimes P^2_{k_2}\big) \\
&\quad = \big(P^1_{j_1}\Phi_1(t)P^1_{k_1}\big) \otimes \big(P^2_{j_2}\Phi_2(t)P^2_{k_2}\big) \\
&\quad = \begin{cases} \Phi_{1,j_1} \otimes \Phi_{2,j_2}, & j_1 = k_1,\ j_2 = k_2, \\ 0, & \text{other cases.} \end{cases}
\end{aligned}$$

That is to say, the corresponding matrix representation of operator $\Phi_1(t) \otimes \Phi_2(t)$ is in block diagonal form.

Next we give the spectrum estimation of the operator $\Phi_1 \otimes \Phi_2$ on each invariant linear subspace $V^1_j \otimes V^2_l$. When $t \geqslant s$, for $j = 1, \dots, p_1$, $l = 1, \dots, p_2$ we have

$$\begin{aligned}
&\big\|\big(\Phi_1(t) \otimes \Phi_2(t)\big)\big(\Phi_1(s) \otimes \Phi_2(s)\big)^{-1}(\mu \otimes \nu)\big\| \\
&\quad = \big\|\Phi_1(t)\Phi_1^{-1}(s)\mu\big\|_1 \cdot \big\|\Phi_2(t)\Phi_2^{-1}(s)\nu\big\|_2 \\
&\quad \leqslant K_1 e^{\alpha^{(1)}_j (t-s)}\|\mu\|_1 \cdot K_2 e^{\alpha^{(2)}_l (t-s)}\|\nu\|_2 \\
&\quad = K_1 K_2 e^{(\alpha^{(1)}_j + \alpha^{(2)}_l)(t-s)}\|\mu \otimes \nu\|,
\end{aligned}$$

where $\mu \in V^1_j$, $\nu \in V^2_l$ and $a^{(i)}_j < \alpha^{(i)}_j < a^{(i)}_j + \varepsilon$ for $i = 1, 2$. Similar inequality can be got when $t < s$. Let $\varepsilon \to 0$, that completes the proof. $\square$

Let $x^\tau = x_1^{\tau_1} x_2^{\tau_2} \dots x_n^{\tau_n}$ be the scalar monomial in n variables of degree $|\tau| = k$, then $H_{k,n}(\mathbb{R}^n) = \{f = \sum_{|\tau|=k} x^\tau f_\tau\colon\ f_\tau \in \mathbb{R}^n\}$ is the vector space of homogeneous polynomials of degree k in n variables with values in $\mathbb{R}^n$. Clearly, we have that

$$d = \dim H_{k,n}\big(\mathbb{R}^1\big) = \frac{(k+n-1)!}{k!(n-1)!}, \qquad D = \dim H_{k,n}\big(\mathbb{R}^n\big) = nd.$$

A basis $F = \{u_1, \dots, u_n\}$ in $\mathbb{R}^n$ and the basis $\{x^\tau\}_{|\tau|=k}$ of $H_{k,n}(\mathbb{R}^1) \cong \mathbb{R}^d$ give a basis $x^\tau F = \{x^\tau u_i\}^{i=1,\dots,n}_{|\tau|=k}$ in $H_{k,n}(\mathbb{R}^n)$, while

$$H_{k,n}\big(\mathbb{R}^n\big) \ni f = \sum_{|\tau|=k}\sum_{i=1}^{n} f_{i,\tau} x^\tau u_i \cong \mathcal{K}_F(f) = \mathcal{K}_F(f_\tau)_{|\tau|=k} = f_{i,\tau} \in \mathbb{R}^D$$

(column vectors, ordered lexicographically) identifies $H_{k,n}(\mathbb{R}^n)$ with $\mathbb{R}^D$, where $\mathcal{K}_F$ is the mapping which assigns F coordinates to an element of $\mathbb{R}^D$. Thus, we can define an inner product in $H_{k,n}(\mathbb{R}^n)$ which reduces $H_{k,n}(\mathbb{R}^n) = H_{k,n}(\mathbb{R}^1) \otimes \mathbb{R}^n \cong \mathbb{R}^d \otimes \mathbb{R}^n$. Define a $d \times d$ matrix $N(A)_k$ by

$$N(A)_k = \big(N^{(k)}_{\tau\varsigma}(A)\big), \qquad (Ax)^\tau = \sum_{\sigma_\varsigma = k} N^{(k)}_{\varsigma\tau}(A)x^\varsigma.$$

In particular, the blocks of $N(A)_k$ depend nonlinearly on the entries of blocks in A. Here Proposition 5, whose result (ii) can provide the sharp norm estimation, is a strong version of Lemma 8.1.2 in [3].

Proposition 5. *Let $A(t):\mathbb{R}\to\mathbb{R}^{n\times n}$ and $k\geqslant 2$, then the following statements hold*:

(i) $N(I_2)_k=I_1$, $N(AB)_k=N(B)_kN(A)_k$, *hence* $N(A^{-1})_k=N(A)_k^{-1}$.
(ii) *If $A(t)$ is bounded on $\mathbb{R}$ and in a block diagonal form with blocks $A_i:\mathbb{R}\to\mathbb{R}^{n_i\times n_i}$ for $i=1,\dots,p$, then there exists a permutation matrix P independent of t in $\mathbb{R}^{d\times d}$ which makes $N(A)_k:\mathbb{R}\to\mathbb{R}^{d\times d}$ similar to a matrix in the block diagonal form with blocks $\Lambda_\tau:\mathbb{R}\to\mathbb{R}^{q_\tau\times q_\tau}$, where $\tau=(\tau_1,\dots,\tau_p)$ runs though the set*

$$\mathbb{N}_k^p=\left\{\tau=(\tau_1,\dots,\tau_p)\in\mathbb{Z}_+^p\colon\ |\tau|=k\right\}$$

and $q_\tau=\prod_{i=1}^p\frac{(\tau_i+n_i-1)!}{\tau_i!(n_i-1)!}$. Furthermore, we have

$$\left\|\Lambda_\tau(t)\right\|\leqslant C\prod_{i=1}^p\|A_i\|^{\tau_i},\qquad\tau\in\mathbb{N}_k^p.$$

Here $\|\cdot\|$ denotes the corresponding matrix norm reduced by the vector norm and the constant C only depends on n, k and $\|\cdot\|$.

Proof. (i) By the definition of $N(AB)_k$,

$$\left((AB)x\right)^\tau=\sum_{|\varsigma|=k}N_{\varsigma\tau}(AB)x^\varsigma=\sum_{|\rho|=k}\sum_{|\mu|=k}N_{\rho\tau}(A)N_{\mu\rho}(B)x^\mu.$$

Equating coefficients gives

$$N_{\tau\varsigma}(AB)=\sum_{|\rho|=k}N_{\tau\rho}(B)N_{\rho\varsigma}(A).$$

(ii) The whole proof consists of three parts.

First of all, we define a partial order of $\mathbb{N}_k^n$, which corresponds to the diagonal form of $A(t)$.

Notice that $(n_1,\dots,n_p)\in\mathbb{Z}_+^p$, then set $s=(s_0,s_1,\dots,s_p)$, where $s_0=1$, $s_p=n$ and $s_j=\sum_{i=1}^j n_i$ for $j=1,\dots,p$. Thus a multi-valued index map Θ can be defined as follows

$$\begin{aligned}\Theta\colon\ \mathbb{N}_k^n&\to\mathbb{N}_k^p,\\ \nu&\mapsto\tau,\end{aligned}\tag{6}$$

where $|\tau|=|\nu|=k$ and $\tau_i=\sum_{j=s_{i-1}}^{s_i}\nu_j$ for $i=1,\dots,p$. Consequently, the map Θ naturally reduces a total order of $\mathbb{N}_k^p$ to the partial order of $\mathbb{N}_k^n$ such that $\nu\prec\tilde{\nu}$ if $\tau=\Theta(\nu)\prec\tilde{\tau}=\Theta(\tilde{\nu})$. And we fix the set $\mathbb{N}_k^p$ a total order (lexicographical order), defined as if $\tau_l>\tilde{\tau}_l$ then $\tau\succ\tilde{\tau}$, where $l=\min\{i\colon\ |\tau_i-\tilde{\tau}_i|\neq 0\}$.

Second, we prove that $N(A)_k$ is of the block diagonal form under the above partial order of $\{x^\nu\}_{|\nu|=k}$.

Regard $N(A)_k$ as an operator depending on t on the space $H_{k,n}(\mathbb{R}^1)$. For a fixed $\tau = (\tau_1, \ldots, \tau_p) \in \mathbb{N}_k^p$, define the index set $\Omega(\tau) = \{\nu \in \mathbb{N}_k^n\colon\ \Theta(\nu) = \tau\}$, then $\natural\Omega(\tau) = q_\tau$ and $E_\tau = \operatorname{span}\{x^\nu\}_{\nu\in\Omega(\tau)}$ is a linear subspace of $H_{k,n}(\mathbb{R}^1)$. Moreover, for any fixed $\nu \in \Omega(\tau)$ we have that $(Ax)^\nu = \prod_{i=1}^p (A_i\hat{x}_i)^{\hat{\nu}_i}$, where $\hat{x}_i = (x_{s_{i-1}}, \ldots, x_{s_i-1})$, $\hat{\nu}_i = (\nu_{s_{i-1}}, \ldots, \nu_{s_i-1}) \in \mathbb{Z}_+^{n_i}$ for $i = 1, \ldots, p$, and $\tau = (|\hat{\nu}_1|, \ldots, |\hat{\nu}_p|)$. Note that $(A_i\hat{x}_i)^{\hat{\nu}_i}$ is a homogeneous polynomial of the degree $\tau_i = |\hat{\nu}_i|$ in variables $\hat{x}_i$. So $(Ax)^\nu$ can be represented by vectors in E_τ. Namely, E_τ is $N(A)_k$ invariant, i.e. $N(A)_k$ is in the block diagonal form.

At last, we estimate the norm of each block of $N(A)_k$.

Since any matrix norms deduced by the vector norm are equivalent, it is convenient to choose the matrix norm as $\|A\| = \max|a_{i,j}|$. For any ν and $\mu \in E_\tau$, next we estimate the coefficient $q_{\mu,\nu}$ of the monomial x^μ of the polynomial $(Ax)^\nu = \prod_{i=1}^p (A_i\hat{x}_i)^{\hat{\nu}_i}$. Obviously, the factor $\hat{x}_i^{\hat{\mu}_i}$ is only generated by the polynomial $(A_i\hat{x}_i)^{\hat{\nu}_i}$ such that its coefficient $|q^i_{\hat{\mu}_i,\hat{\nu}_i}| \leqslant (n_i)^{\tau_i}\|A_i\|^{\tau_i}$. Thus, it follows

$$|q_{\mu,\nu}| = \prod_{i=1}^p \left|q^i_{\hat{\mu}_i,\hat{\nu}_i}\right| \leqslant n^k \prod_{i=1}^p \|A_i\|^{\tau_i},$$

which means

$$\left\|\Lambda_\tau(t)\right\| = \max_{\Theta(\nu)=\Theta(\mu)=\tau} |q_{\mu,\nu}| \leqslant n^k \prod_{i=1}^p \|A_i\|^{\tau_i}.$$

This completes the proof. $\square$

Remark. Set e_j be the unit vector with the jth component 1. If $A(t)$ is in the block diagonal form w.r.t. the spectral intervals in system (2), then the vector-valued monomial $x^\nu e_j$ is called *resonant monomial* if τ and l satisfy the condition $0 \in [a_l - \sum_{i=1}^p \tau_i b_i, b_l - \sum_{i=1}^p \tau_i a_i]$, where $e_l = \Theta(e_j)$, $\tau = \Theta(\nu)$ and the definition of Θ is given by (6). This definition accords with Siegmund's for non-autonomous systems in [5].

Let $T(A)_k$ be the matrix describing the linear mapping on $H_{k,n}(\mathbb{R}^1)$ given by

$$h = \sum_{|\tau|=k} h_\tau x^\tau \mapsto \sum_{|\tau|=k}\sum_{j,l=1}^n h_\tau \frac{\partial(x^\tau)}{\partial x_j} a_{jl} x_l =: T(A)_k(h),$$

and denote by $\Phi_{-T(A)_k}$ the principle matrix solution of linear equation $\dot{x} = -T(A)_k(t)x$. In particular, there is a close connection between operators $N(\cdot)_k$ and $T(\cdot)_k$. In Proposition 8.3.4 of [3], Arnold discovered it for random dynamic systems. Here we modify that proposition for the case of ordinary differential equations.

Proposition 6. *Let $A(t)\colon \mathbb{R} \to \mathbb{R}^{n\times n}$ and $k \geqslant 2$, then we have that*

$$N\big(\Phi_A(t)\big)_k^{-1} = \Phi_{-T(A)_k}(t).$$

Proof. We have that

$$\begin{aligned}\frac{d}{dt}h\big(\Phi_A(t)x\big) &= Dh\big(\Phi_A(t)x\big)\frac{d}{dt}\big(\Phi_A(t)x\big)\\ &= Dh\big(\Phi_A(t)x\big)\big(A(t)\Phi_A(t)x\big)\\ &= Dh(y)A(t)y|_{y=\Phi_A(t)x}.\end{aligned}$$

The matrix version of this states that

$$\frac{d}{dt}N\big(\Phi_A(t)\big)_k = N\big(\Phi_A(t)\big)_k T(A)_k(t),$$

which is equivalent to

$$\frac{d}{dt}\big(N\big(\Phi_A(t)\big)_k^{-1}\big) = -T(A)_k(t)\big(N\big(\Phi_A(t)\big)_k^{-1}\big),$$

i.e., $N(\Phi_A(t))_k^{-1} = \Phi_{-T(A)_k}(t)$. $\square$

The following is a strong version Gronwall type integral inequality, which is from a result of Sardarly (1965) and the proof can be found in [8].

Lemma 7. *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 $a \leqslant s \leqslant t \leqslant \beta$, let $b(t)$ and $q(t)$ be nonnegative in J and suppose*

$$u(t) \leqslant a(t) + \int_\alpha^t \big(q(t)b(s)u(s) + c(t,s)\big)\,ds, \quad t \in J.$$

Then for $t \in J$

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

Let $f(x,t) \in C(D\times\mathbb{R},\mathbb{R}^n)$, where D is an open set in $\mathbb{R}^n$ (more generally, a separable Banach space). $f(x,t)$ is said to be *almost periodic* (*a.p.*) in t uniformly for $x \in D$, if for any $\varepsilon > 0$ and any compact set S in D, there exists a positive number $l(\varepsilon,S)$, such that any interval of length $l(\varepsilon,S)$ contains a τ for which

$$\big\|f(x,t+\tau) - f(x,t)\big\| \leqslant \varepsilon$$

for all $t \in \mathbb{R}$ and all $x \in S$. The following lemma illustrates the connection between an E.D. and the unique bounded solution. See [9] for more details.

Lemma 8. *Consider the following inhomogeneous system*

$$\dot{x} = A(t)x + f(t), \tag{7}$$

where $A(t)$ and $f(t)$ are continuous and uniformly bounded on $\mathbb{R}$. Assume that the corresponding homogeneous system $\dot{x} = A(t)x$ has an E.D. on $\mathbb{R}$, then Eq. (7) *has a unique bounded solution ψ. In addition, if $A(t)$ and $f(t)$ are almost periodic on $\mathbb{R}$, so is the unique solution ψ and $m(\psi) \subset m(A, f)$.*

3. Finite order normal forms

In this part we mainly deal with the formal normal forms of system (2) w.r.t. its dichotomy spectrum. The main result is Proposition 9, whose proof greatly relies on Lemma 10.

As usual $\mathrm{Jet}^k_{x=0} F(x,t)$ denotes the Taylor expansion of the function F w.r.t. the variable x at $x = 0$ of order k.

Proposition 9. *Assume the dichotomy spectrum of system* (3) *to be $\Sigma_A = [a_1, b_1] \cup \cdots \cup [a_p, b_p]$, where $a_1 \leqslant b_1 < \cdots < a_p \leqslant b_p$.*

(i) *Non-resonant case. If the dichotomy spectrum is kth non-resonant, i.e.,*

$$0 \notin \left[\sum_{i=1}^{p} a_i m_i - a_j, \sum_{i=1}^{p} b_i m_i - b_j \right], \qquad k = \sum_{i=1}^{p} m_i \geqslant 2,$$

then system (2) *is locally analytically equivalent to the system*

$$\dot{x} = \mathrm{Jet}^{k-1}_{x=0} F(x,t) + O\big(\|x\|^{k+1}\big).$$

(ii) *Resonant case. If $A(t)$ is block diagonal w.r.t. the spectral interval $[a_i, b_i]$ for $i = 1, \ldots, p$, i.e. $A(t) = \mathrm{diag}\{A_1(t), \ldots, A_p(t)\}$, where each block $A_i : \mathbb{R} \to \mathbb{R}^{n_i \times n_i}$ corresponds to a spectral interval $[a_i, b_i]$ for $i = 1, \ldots, p$, then system* (2) *is locally analytically equivalent to the system*

$$\dot{x} = \mathrm{Jet}^{k-1}_{x=0} F(x,t) + g_k(x,t) + O\big(\|x\|^{k+1}\big), \tag{8}$$

where g_k is a polynomial w.r.t. x of the degree k, which consists resonant monomials only.

(iii) *Almost periodic case. In addition if F is almost periodic in the variable t in Eq.* (2)*, then so are the transformation and the changed system in results* (i) *and* (ii)*.*

Lemma 10. *Assume non-autonomous linear system $\dot{x} = A(t)x$ with $A(t)$ continuous and bounded on $\mathbb{R}$ has a dichotomy spectrum $\bigcup_{i=1}^{p} [a_i, b_i]$, where $a_1 \leqslant b_1 < \cdots < a_p \leqslant b_p$, let $H_{k,n}(\mathbb{R}^n)$ be the space of n-dimensional vector-valued homogeneous polynomials of n variables of degree $k \geqslant 2$ as before, and define a t-depending linear operator $L^A_k(t)$ on $H_{k,n}(\mathbb{R}^n)$ as follows*

$$L^A_k : h(x) \mapsto A(t)h(x) - \frac{\partial h(x)}{\partial x} A(t)x, \tag{9}$$

then the non-autonomous linear system $\dot{x} = L_k^A(t)x$ *has the spectrum estimation*

$$\bigcup_{\tau\in\mathbb{N}_k^p}\bigcup_{j=1,\dots,p}\left[a_j-\sum_{i=1}^p\tau_i b_i,\, b_j-\sum_{i=1}^p\tau_i a_i\right].$$

Proof. Together with the space decomposition $H_{k,n}(\mathbb{R}^n) = H_{k,n}(\mathbb{R}^1)\otimes\mathbb{R}^n$, the matrix representation of L_k^A is

$$L_k^A(t) = I_1\otimes A - T(A)_k\otimes I_2,$$

where I_1 and I_2 are identity matrices in $H_{k,n}(\mathbb{R}^1)$ and $\mathbb{R}^n$. See [2,3] for more details. By Propositions 6 and 4 we have that $\Phi_{L_k^A} = \Phi_{-T(A)_k}\otimes\Phi_A = N(\Phi_A)_k^{-1}\otimes\Phi_A$. Consequently, by Theorem 3 we can rewrite $\Phi_{L_k^A}$ in another form

$$\Phi_{L_k^A} = N(\Phi_A)_k^{-1}\otimes\Phi_A = N(S\Psi)_k^{-1}\otimes(S\Psi) = \big(N(S)_k^{-1}\otimes S\big)\cdot\big(N(\Psi)_k^{-1}\otimes\Psi\big),$$

where $x = S(t)y$ is the LP transformation, Ψ is in the block diagonal form with blocks Ψ_i for $i = 1,\dots,p$ and for each block it admits

$$\begin{aligned}\big\|\Psi_i(t)\Psi_i^{-1}(s)\big\| &\leqslant Ke^{\beta_i(t-s)}, \quad t\geqslant s,\\ \big\|\Psi_i(t)\Psi_i^{-1}(s)\big\| &\leqslant Ke^{\alpha_i(t-s)}, \quad t\leqslant s.\end{aligned}$$

Here ε, a positive number, can be chosen arbitrary small and $b_i<\beta_i<b_i+\varepsilon$, $a_i-\varepsilon<\alpha_i<a_i$ for $i=1,\dots,p$. However, notice that the fact

$$\big(N\big(\Psi(t)\big)_k^{-1}\big)\cdot\big(N\big(\Psi(s)\big)_k^{-1}\big)^{-1} = N\big(\Psi(s)\Psi^{-1}(t)\big)_k,$$

and, by Proposition 5, $N(\Psi(s)\Psi^{-1}(t))_k$ is similar to a block diagonal matrix, whose block $\{\Lambda_\tau\}_{\tau\in\mathbb{N}_k^p}$ has a norm control as follows

$$\begin{aligned}\|\Lambda_\tau\| &\leqslant K'e^{-\sum\tau_i\alpha_i(t-s)}, \quad t\geqslant s,\\ \|\Lambda_\tau\| &\leqslant K'e^{-\sum\tau_i\beta_i(t-s)}, \quad t\leqslant s.\end{aligned}$$

Therefore, let $\varepsilon\to 0$ and the spectrum estimation of $N(\Psi)_k^{-1}$ is

$$\bigcup_{\tau\in\mathbb{N}_k^p}\left[-\sum_{i=1}^p\tau_i\beta_i,\, -\sum_{i=1}^p\tau_i\alpha_i\right],$$

which completes the proof because of Proposition 4. □

Proof of Proposition 9. If the system $\dot{x} = A(t)x + f(x,t)$ can be changed into the system $\dot{y} = A(t)y + g(y,t)$, then the coordinate substitution $x = y + h(y,t)$ satisfies equation

$$\frac{\partial h}{\partial t} = A(t)h(y,t) - \frac{\partial h}{\partial y}A(t)y + f(y+h,t) - g(y,t) - \frac{\partial h}{\partial y}g. \tag{10}$$

Expanding h, f and g in the form of Taylor series w.r.t. the variable y, we have that

$$h(y,t) = \sum_{|\nu|\geqslant 2}^{\infty} h_\nu(t)y^\nu, \qquad g(y,t) = \sum_{|\nu|\geqslant 2}^{\infty} g_\nu(t)y^\nu, \qquad f(x,t) = \sum_{|\nu|\geqslant 2}^{\infty} f_\nu(t)x^\nu,$$

where h_ν, g_ν and f_ν are all bounded vector-valued functions from $\mathbb{R}$ to $\mathbb{R}^n$ by the Cauchy estimation for $\rho' < \rho$.

By comparing the monomials of degree k with respect to the variable y in the equality (10), we can obtain

$$\frac{dh_k(t)}{dt} = L_k^A(t)h_k(t) + T_k(t) - g_k(t), \tag{11}$$

where $L_k^A(t)$ is given by (9), h_k and g_k are both vector-valued functions from $\mathbb{R}$ to $H_{k,n}(\mathbb{R}^n)$ and T_k is the vector coefficient of Taylor expansion of the function

$$f\left(y + \sum_{|\nu|=2}^{k-1} h_\nu(t)y^\nu\right)$$

w.r.t. the variable y of degree k, which is known already.

Now we seek bounded solutions h_k of system (11) for the convenient function g_k. If the dichotomy spectrum is kth non-resonant, then by Lemma 10 the linear part system

$$\frac{dh_k(t)}{dt} = L_k^A(t)h_k(t)$$

admits an exponential dichotomy. By Lemma 8 for any bounded function g_k Eq. (11) has a unique bounded solution. It is convenient to choose $g_k = 0$. That completes the proof of (i).

If $A(t)$ is block diagonal w.r.t. the spectral interval $[a_i, b_i]$ for $i = 1, \ldots, p$, then by Propositions 5, 4 and Lemma 10 the vector-valued function $(T_\nu - g_\nu)e_j$ corresponds to the diagonal block of $L_k^A(t)$ with dichotomy spectrum estimation $[a_l - \sum_{i=1}^p \tau_i b_i, b_l - \sum_{i=1}^p \tau_i a_i]$, where $e_l = \Theta(e_j)$, $\tau = \Theta(\nu)$. Moreover, still by Lemma 10 we know that $L_k^A(t)$ is similar to the matrix as follows

$$\begin{pmatrix} M_+(t) & & \\ & M_-(t) & \\ & & M_0(t) \end{pmatrix},$$

where $M_+(t)$ consists of blocks with dichotomy spectrum estimation $a_l - \sum \tau_i b_i > 0$, $M_-(t)$ consists of blocks with spectrum estimation $b_l - \sum \tau_i a_i < 0$ and the others are in $M_0(t)$. Now let $h_k(t) = (h_+, h_-, h_0)$, $T_k(t) = (T_+, T_-, T_0)$, $g_k(t) = (g_+, g_-, g_0)$ and the principle matrix solution $\Phi(t) = (\Phi_+, \Phi_-, \Phi_0)$ corresponding to the decomposition of the operator $L_k^A(t)$. For the first two parts, let $g_+ = g_- = 0$, we have the unique bounded solutions

$$h_+ = -\Phi_+(t)\int_t^{\infty} \Phi_+^{-1}(s)T_+(s)\,ds$$

and

$$h_- = \Phi_-(t)\int_{-\infty}^{t} \Phi_-^{-1}(s)T_-(s)\,ds.$$

And for the third one, we can simply set $h_0 = 0$, which follows $g_0 = T_0$. Therefore the solution $h_k = (h_+, h_-, h_0)$ is bounded on $\mathbb{R}$. That completes the proof of statement (ii).

At last we note that if F is almost periodic in t uniformly w.r.t. the variable x, so are the functions T_k, g_k and by Lemma 8 the unique bounded solution h_k. That completes the proof of statement (iii). $\square$

Remark. Set $\mathcal{F}$ be the set of formal Taylor expansions with bounded functional coefficients, namely,

$$\mathcal{F} = \left\{ f\colon\ f = \sum_{|\nu|=2}^{\infty} f_\nu(t)x^\nu,\ \left|f_\nu(t)\right| \leqslant M_\nu < \infty,\ \nu \in \mathbb{Z}_+^n \right\}.$$

Then the non-autonomous system $\dot{x} = A(t)x + f(x,t)$ is said to be *formally equivalent* to the system $\dot{y} = B(t)y + g(y,t)$ if there exists a coordinate substitution $x = P(t)y + h(x,t)$ which transforms one system to the other, where $P(t)$ is an LP transformation and $f, g, h \in \mathcal{F}$. Therefore, Proposition 9 is also valid provided we use 'formally equivalent' instead of 'locally analytically equivalent,' which covers the work of [5] using a new method. In addition, we briefly introduce the result of Siegmund on normalization of jets in [5] as follows.

Normal Form Theorem *(Siegmund). Assume that $A(t)$ has bounded growth and $F(x,t)$ is a C^k Carathéodory function for some $k \geqslant 2$ in system* (2). *Therefore the dichotomy spectrum consists of p, $1 \leqslant p \leqslant n$, compact intervals $\Sigma_A = [a_1, b_1] \cup \cdots \cup [a_p, b_p]$, where $a_1 \leqslant b_1 < \cdots < a_p \leqslant b_p$. Then system* (2) *is locally C^k equivalent to a differential equation*

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

where $G(x,t)$ is a C^k Carathéodory function and $\tilde{f} = O(\|x\|^2)$ as $x \to 0$ contains resonant monomials only for the order from 2 *to k w.r.t. x.*

4. Analytic normal forms

The main technique in the proof of Theorem 1 is the so-called homotopy method, which reduces the problem of equivalence of families of autonomous vector fields to that of solubility of a system of first-order linear partial differential equations. In general this method is frequently applied to prove the smooth conjugacy of autonomous systems. Now we modify it for our case. One can refer to [10] for more details.

Lemma 11. *Let $\phi_s(x,y) = L(x,y) + sR(x,y)$ be analytic in $O_\rho \times \mathbb{R}$ and $s \in [0,1]$. If there exists a function $r(x,y,s)$ such that*

(i) *$r(x,y,s)$ is analytic in $O_{\rho'} \times \mathbb{R} \times [0,1]$ for some $\rho' < \rho$ and satisfies $\|r(x,y,s)\| = O(\|x\|^2)$ as $x \to 0$,*
(ii) *$r(x,y,s)$ satisfies the following equation*

$$\frac{\partial \phi_s}{\partial x} r - \frac{\partial r}{\partial x}\phi_s - \frac{\partial r}{\partial y} = -R, \tag{12}$$

then system $\frac{dx}{dy} = \phi_1(x,y)$ is locally analytically equivalent to system $\frac{dx}{dy} = \phi_0(x,y)$.

Proof. We introduce two vector fields $V(x,y,s) = (r(x,y,s),0,1)$ and $\psi(x,y,s) = (\phi_s(x,y), 1,0)$, which defined on $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$. Using (12), we obtain that

$$[\psi, V] = 0, \tag{13}$$

where the Lie bracket $[\cdot,\cdot]$ is w.r.t. variables (x,y,s). Denote by $g^t_\cdot$ the local flow generated by the vector fields $\cdot$, then from (13) we have that

$$g^a_\psi \circ g^b_V = g^b_V \circ g^a_\psi, \tag{14}$$

where $0 < a, b \ll 1$. By condition (i) g^1_V is well defined. Write $g^1_V(x,y,0) = (h(x,y),y,1)$ and take the derivative w.r.t. a at $a=0$ in both sides of equality (14), we obtain that

$$\left.\frac{dg^a_\psi \circ g^1_V}{da}\right|_{a=0}(x,y,0) = \psi\big(h(x,y),y,1\big) = \big(\phi_1\big(h(x,y),y\big),1,0\big)^T$$

and

$$\begin{aligned}\left.\frac{g^1_V \circ g^a_\psi}{da}\right|_{a=0}(x,y,0) &= Dg^1_V(x,y,0)\big(\phi_0(x,y),1,0\big)^T \\ &= \big(\partial_x h(x,y)\phi_0(x,y) + \partial_y h(x,y),1,0\big)^T,\end{aligned}$$

which yields

$$\phi_1\big(h(x,y),y\big) = \partial_x h(x,y)\phi_0(x,y) + \partial_y h(x,y).$$

This completes the proof. □

For simplicity of notations, we set $(f_* R)(x) = (DfR) \circ f^{-1}(x)$. Then we have a formal solution of (12).

Lemma 12. *Eq.* (12) *has a formal solution*

$$h(x,y,s)=-\int_0^\infty D_x^{-1}G_s(t;x,y)\cdot R\bigl(G_s(t;x,y),t+y\bigr)\,dt, \tag{15}$$

where $G_s(t;x,y)$ is the solution of the non-autonomous system

$$\frac{dx}{dt}=\phi_s(x,t+y), \tag{16}$$

with the initial condition $G_s(0;x,y)=x$.

Proof. Set $\tilde{v}_s=(\phi_s(x,y),1)$, $\tilde{h}=(r,0)$ and $\tilde{R}=(R,0)$. Together with (12), formula (13) is equivalent to the homological equation

$$[\tilde{h},\tilde{v}_s]=\tilde{R}, \tag{17}$$

where $[\cdot,\cdot]$ is the Lie bracket taken w.r.t. (x,y) and s is the parameter. The trajectories of the field $\tilde{v}_s$ is defined by $g^t(x,y)$ and it has the form $g^t(x,y)=(G_s(t;x,y),t+y)$. Denote by $X_s(t;x,y)$ the matrix solution of the field $\tilde{v}$, i.e.,

$$\begin{aligned}X_s(t;x,y)&=\frac{\partial g^t(x,y)}{\partial(x,y)}\\&=\begin{pmatrix}D_xG_s(t;x,y) & *\\ 0 & 1\end{pmatrix}.\end{aligned}$$

Then we claim that the homological equation (17) has a formal solution

$$\tilde{h}(x,y)=-\int_0^\infty X_s^{-1}(t;x,y)\tilde{R}\bigl(g^t(x,y)\bigr)\,dt.$$

To prove the claim, we set $\tilde{h}^\tau=(g^\tau(x,y))_*\tilde{h}$, which implies

$$\begin{aligned}\tilde{h}^\tau&=X_s\bigl(\tau;g^\tau(x,y)\bigr)\tilde{h}\bigl(g^{-\tau}(x,y)\bigr)\\&=-\int_0^\infty X_s^{-1}(-\tau;x,y)X_s^{-1}\bigl(t;g^{-\tau}(x,y)\bigr)\tilde{R}\bigl(g^{-\tau+t}(x,y)\bigr)\,dt\\&=-\int_0^\infty X_s^{-1}(t-\tau;x,y)\tilde{R}\bigl(g^t(x,y)\bigr)\,dt\\&=-\int_0^\infty X_s^{-1}(t;x,y)\tilde{R}\bigl(g^t(x,y)\bigr)\,dt.\end{aligned}$$

Thus, we obtain that

$$[\tilde{v}_s, \tilde{h}] = \frac{d\tilde{h}^\tau}{d\tau}\bigg|_{\tau=0} = -X_s^{-1}(0; x, y)\tilde{R}\big(g^0(x, y)\big) = -\tilde{R}(x, y).$$

This proves the claim.

Finally it is easy to verify that

$$\begin{aligned}\tilde{h}(x, y, s) &= -\int_0^\infty X_s^{-1}(t; x, y) \cdot \tilde{R}\big(g^t(x, y)\big)\,dt \\ &= \begin{pmatrix} -\int_0^\infty D_x^{-1} G_s(t; x, y) \cdot R(G_s(t; x, y), t + y)\,dt \\ 0 \end{pmatrix}.\end{aligned}$$

By comparing the first component of $\tilde{h}$, this completes the proof. $\square$

The next is the main theorem of this part, which is an extension of Poincaré–Dulac Theorem for autonomous systems.

Theorem 13. *Consider the non-autonomous differential equations*

$$\dot{x} = F(x, t) = A(t)x + f(x, t), \tag{18}$$

where $F(x, t)$ is analytic in $O_\rho \times \mathbb{R}$, $A(t)$ is uniformly bounded on $\mathbb{R}$ and $f = O(\|x\|^2)$ as $x \to 0$. Set the dichotomy spectrum of the corresponding homogeneous equation $\dot{x} = A(t)x$ be $\Sigma_A = [a_1, b_1] \cup \cdots \cup [a_p, b_p]$, where $a_1 \leqslant b_1 < \cdots < a_p \leqslant b_p$. If $a_1 b_p > 0$, then Eq. (18) *is locally analytically equivalent to*

$$\dot{x} = \mathrm{Jet}_{x=0}^{N-1} F(x, t),$$

where N is no greater than $\max\{a_1/b_p, b_p/a_1\}$. *In addition, if F is almost periodic in the variable t, then so are the transformation and the changed equation.*

The key of the proof is to show the formal solution given by (15) is analytic in $O_{\rho'} \times \mathbb{R} \times [0, 1]$ for some $\rho' < \rho$ and almost periodic in y.

Proof of Theorem 13. Without loss of generality, we assume $b_p < 0$. Since $O_\rho \times [-K, K] \subseteq \mathbb{C}^{n+1}$ is compact for any fixed K, we denote the domain

$$U_\delta(K) = \big\{(x, y) \in \mathbb{C}^{n+1} \;\big|\; |x| \leqslant \delta,\ |\mathrm{Re}\, y| \leqslant K + \delta,\ |\mathrm{Im}\, y| \leqslant \delta\big\},$$

which is contained in the domain where $F(x, t)$ is analytic. As usual Re and Im denote the real and imaginary part of a complex, respectively. Set $M_{K,\delta} = \sup_{U_\delta(K)} \|F(x, t)\|$ and $N > a_1/b_p$. Notice that $M_{K,\delta} \leqslant M < \infty$.

By Cauchy's integral representation

$$\partial_x^\tau f(x,t) := \frac{\partial^{|\tau|} f(x_1,\ldots,x_n,t)}{\partial x_1^{\tau_1}\cdots\partial x_n^{\tau_n}} = \frac{\tau!}{(2\pi\sqrt{-1})^n}\int\limits_\gamma \frac{f(z,t)\,dz}{(z-x)^{\tau+e}},$$

where $e=(1,\ldots,1)\in\mathbb{Z}_+^n$ and $\gamma=\{z\colon |z_i|=\delta-\varepsilon,\ i=1,\ldots,n\}$ for $0<\varepsilon\ll 1$, then $\partial_x^\tau f(x,t)$ and $\partial_x^\tau f(0,t)$ are both analytic in $U_{\delta-\varepsilon}(K)$ for all $\tau\in\mathbb{Z}_+^n$, and so are the following functions

$$P(x,t)=\sum_{|\tau|=2}^{N-1}\frac{\partial_x^\tau f(0,t)}{\tau!}x^\tau, \qquad R(x,t)=f(x,t)-P(x,t).$$

Let $\tilde F(x,t,s)=P(x,t)+sR(x,t)$ for $s\in D_2=\{z\in\mathbb{C}\mid |z|\leqslant 2\}$, then we obtain that

$$\sup_{U_\mu(K)\times D_2}\bigl\|D_x\tilde F(x,t,s)\bigr\|=\delta_1\leqslant\frac{CM_{K,\delta}}{\delta}\mu, \quad \mu\leqslant\delta/3,$$

and

$$\bigl\|R(x,t)\bigr\|\leqslant\frac{CM_{K,\delta}N^nN!}{\delta^N}\|x\|^N, \quad (x,t)\in U_\mu(K),$$

where C is a constant depending only on n and $D_xF(x,t)$ is the Jacobi matrix of F w.r.t. the variable x.

Now let $L(x,y)=A(y)x+P(x,y)$. Using Lemmas 11 and 12 next we will show that the formal solution given by (15) is analytic in $U_\mu(K)\times D_2$ for any fixed K.

First, we give the estimation of $G_s(t;x,y)$, which is the solution of Eq. (16). By variation formula we obtain that

$$G_s(t;x,y)=\Phi(t+y,y)x+\int\limits_0^t\Phi(t+y,v+y)F\bigl(G_s(v;x,y),v+y\bigr)\,dv,$$

where $\Phi(t,s)=\Phi(t)\Phi^{-1}(s)$. Notice that $|\mathrm{Re}\,y|\leqslant K+\mu$ and $|\mathrm{Im}\,y|\leqslant\mu$, then for $0\leqslant v\leqslant t$ we have that

$$\begin{aligned}\bigl\|\Phi(t+y,v+y)\bigr\|&=\bigl\|\Phi(t+y)\Phi^{-1}(v+y)\bigr\|\\&=\bigl\|\Phi(t+y,t+\mathrm{Re}\,y)\Phi(t+\mathrm{Re}\,y,v+\mathrm{Re}\,y)\Phi(v+\mathrm{Re}\,y,v+y)\bigr\|\\&\leqslant e^{2\delta M_{K,\delta}}\bigl\|\Phi(t+\mathrm{Re}\,y,v+\mathrm{Re}\,y)\bigr\|\leqslant Ke^{\beta_p(t-v)},\end{aligned}$$

where $b_p<\beta_p<b_p+\varepsilon$. Thus it leads to the inequality

$$\bigl\|G_s(t;x,y)\bigr\|\leqslant Ke^{\beta_pt}+K\delta_1\int\limits_0^t e^{\beta_p(t-v)}\bigl\|G_s(v;x,y)\bigr\|\,dv.$$

Consequently, by Lemma 7, the Gronwall type inequality, we have $\|G_s(t;x,y)\| \leqslant Ke^{(\beta_p+K\delta_1)t}$ for $(x,y) \in U_\delta(K)$ and $s \in D_2$. This inequality also implies that the formal solution h given by (15) is well defined for $t \in [0,\infty]$.

Then, we provide the proof of analyticity of $h(x,y,s)$.

Naturally, for $t \in [0,\infty)$ we have that

$$\big\|R\big(G_s(t;x,y), y+t\big)\big\| \leqslant \frac{CM_{K,\delta}N^n N!}{\delta^N} e^{(\beta_p+K\delta_1)Nt}, \quad \forall (x,y) \in U_\mu(K),\ s \in D_2.$$

Moreover $D_x^{-1}G_s(t;x,y)$ satisfies the matrix differential equation

$$\frac{d}{dt}D_x^{-1}G_s(t;x,y) = -D_x^{-1}G_s(t;x,y)\big(A(t+y) + D_xF\big(G_s(t;x,y), t+y\big)\big),$$

which can also be written as a matrix integral equation

$$\begin{aligned} D_x^{-1}G_s(t;x,y) &= \Phi(y,t+y) \\ &\quad - \int_0^t D_x^{-1}G_s(v;x,y)D_xF\big(G_s(v;x,y), v+y\big)\Phi(v+y,t+y)\,dv. \end{aligned}$$

It follows by Lemma 7 again that $\|D_x^{-1}G_s(t;x,y)\| \leqslant Ke^{(-\alpha_1+K\delta_1)t}$ for $(x,y) \in U_\delta(K)$ and $s \in D_2$, where $a_1 - \varepsilon < \alpha_1 < a_1$. Let $\alpha = \alpha_1 - K\delta_1$ and $\beta = \beta_p + K\delta_1$, by choosing ε and δ small enough, we can make $-\alpha + \beta N < 0$, which means the integral representation of the function $h(x,y,s)$ converges uniformly in the domain $(x,y) \in U_\mu(K)$ and $s \in D_2$ for any fixed K. Since all the above norm estimations are independent of K, therefore $h(x,y,s)$ is analytic in $O_\mu \times \mathbb{R} \times [0,1]$ for $\mu < \rho$.

In addition if $F(x,t)$ is a.p. in t, then so are functions $R(x,y)$ and $G_s(t;x,y)$ in y. Since we have shown that the formal solution h given in (15) converges uniformly in the domain $(x,y) \in U_\mu(K)$ and $s \in D_2$ for any fixed K, then $h(x,y,s)$ is almost periodic in y. This completes the whole proof. □

5. Proof of main theorem and applications

In this part the main theorem is proved. Moreover, as applications we provided two examples, one is a quasi-periodic system, whose spectrum bundles of its linear part are all 1-dimension, and the other is a λ-shift system.

Proof of Theorem 1. The proof of statement (i) is from Theorem 13 straightforwardly. While by Proposition 9(i) and under the condition of statement (ii), system (3) is locally analytically equivalent to

$$\dot{x} = A(t)x + O\big(\|x\|^N\big), \tag{19}$$

where $N > \max\{a_1/b_p, b_p/a_1\}$. Then by Theorem 13 system (19) is locally analytically equivalent to $\dot{x} = A(t)x$. This completes the proof of statement (ii). The arguments for the proof of statement (iii) are similar. At last, Proposition 9(iii) and Theorem 13 imply statement (iv). This completes the proof. □

Applications

In [7] Johnson and Sell study the reducibility of linear quasi-periodic system

$$\frac{dx}{dt} = A(t)x = \tilde{A}(\theta_0 + \omega t)x, \tag{20}$$

where $\tilde{A}(\theta): T^k \to \mathbb{R}^{n\times n}$ is continuous, θ_0 is fixed and $\omega \in \mathbb{R}^k$ is rationally independent. They show that if ω satisfies the non-resonant condition and system (20) has full spectrum and sufficient smoothness, then there exists a quasi-periodic LP transformation changes system (20) into

$$\frac{dy}{dt} = B(t)y, \tag{21}$$

where $B(t) = B$ is a constant. However, in order to consider the linear stability instead we now restrict our attention to the quasi-periodic system

$$\frac{dx}{dt} = F(x,t) = A(t)x + f(x,t), \tag{22}$$

where $F(x,t) = \tilde{F}(x, \theta_0 + \omega t)$ and $f = O(\|x\|^2)$ as $x \to 0$.

Theorem 14. *Assume that the following statements hold*:

- (i) (*Analyticity*) *In system* (22) *the function* $\tilde{F}(x,\theta)$ *is analytic in* $O_\rho \times T^k$.
- (ii) *System* (20) *has the dichotomy spectrum* $[a_1, b_1] \cup \cdots \cup [a_n, b_n]$, *where* $a_1 \leqslant b_1 < \cdots < a_n \leqslant b_n$.
- (iii) (*Non-resonance*) *In addition for the spectrum of system* (20) *we have* $a_1 b_n > 0$ *and* $2a_1 > b_n$ *if* $a_1 > 0$ *or* $2b_n < a_1$ *if* $a_1 < 0$.

Then there exists a quasi-periodic coordinate substitution $x = H(y,t) = P(t)y + h(y,t)$ *that changes system* (22) *to system* (21), *where* $B(t) = \operatorname{diag}(b_1(t), \ldots, b_n(t))$. *Furthermore, the quasi-periodic function* $H(y,t)$ *has the form*

$$H(y,t) = \tilde{H}(y, \overline{\omega}_1 t, \ldots, \overline{\omega}_k t),$$

where $P(t)$ *is an LP transformation,* $\tilde{H}(y,\theta)$ *is analytic in* $\{0\} \times T^k$ *and* $\overline{\omega}_i = \omega_i/2$ *for* $i = 1, \ldots, k$.

Proof. It is proved in [7] that system (20) can be changed into the diagonal form by a quasi-periodic LP transformation $x = P(t)y$ provided condition (ii) is satisfied. However, in addition we have $P(t) = \tilde{P}(\overline{\omega}_1 t, \ldots, \overline{\omega}_k t)$, whose frequencies are half. Finally it is followed by conditions (i) and (iii) that we can directly apply Theorem 1 to get the result. This completes the proof. □

At last we give a one parameter vector fields to illustrate that conditions of Theorem 1 can be naturally archived. Consider the λ-shift system of (2)

$$\frac{dx}{dt} = \big(\lambda I + A(t)\big)x + f(x,t), \tag{23}$$

where λ is a parameter. Obviously if $|\lambda|$ is large enough then conditions of Theorem 1 are automatically fulfilled.

Corollary 15. *For any non-autonomous system of form* (2) *its* λ*-shift system given by* (23) *is locally analytically equivalent to its linear part provided the parameter* $|\lambda|$ *is sufficiently large.*

References

[1] W. Li, Normal Form Theory and Applications, Science Press, Beijing, 2000.
[2] Yu. Ilyashenko, S. Yakovenko, Lectures on Analytic Differential Equations, Amer. Math. Soc., 2007.
[3] L. Arnold, Random Dynamical Systems, Springer-Verlag, Berlin, 1998.
[4] W. Li, J. Llibre, X. Zhang, Extension of Floquet's theory to nonlinear periodic differential systems and embedding diffeomorphisms in differential flows, Amer. J. Math. 124 (2002) 107–127.
[5] S. Siegmund, Normal forms for nonautonomous differential equations, J. Differential Equations 178 (2002) 541–573.
[6] R.J. Sacker, G.R. Sell, A spectral theory for linear systems, J. Differential Equations 27 (1978) 320–358.
[7] R.A. Johnson, G.R. Sell, Smoothness of spectral subbundles and reducibility of quasi-periodic linear differential systems, J. Differential Equations 41 (1981) 262–288.
[8] D. Bainov, P. Simenov, Integral Inequalities and Applications, Kluwer, Dordrecht, 1992.
[9] W.A. Coppel, Dichotomies in Stability Theory, Lecture Notes in Math., vol. 629, Springer-Verlag, Berlin, New York, 1978.
[10] Yu. Il'yashenko, S. Yakovenko, Finitely-smooth normal forms of local families of diffeomorphisms and vector fields, Russian Math. Surveys 46 (1) (1991) 1–43.