On the ultradifferentiable normalization

Hao Wu¹ · Xingdong Xu¹ · Dongfeng Zhang¹

Received: 8 October 2020 / Accepted: 31 January 2021 / Published online: 26 February 2021
© The Author(s) 2021

Abstract

We show the theory of the formal ultradifferentiable normalization. The tools utilized here are KAM methods and Contraction Mapping Principle in the Banach space fixed with weighted norms.

Keywords Normal form · Ultradifferentiable normalization · Small divisor condition

Mathematics Subject Classification 34C20 · 37G05 · 26E10

1 Introduction

The normal form theory founded for dynamical systems by Poincaré is widely used to analyze local dynamical properties, whose main idea is to simplify systems to proper forms. In the modern language, we seek the simple representation of the equivalent class $(\mathcal{X}, \|\cdot\|)/\sim$, where $\mathcal{X}$ is the set of the (local) one-parameter(two-parameters) transformation groups maybe fulfilling some special structures, $\|\cdot\|$ is the norm or topology and $\sim$ refers to the admissible transformations preserving such structures. Many researchers including Dulac, Sternberg, Chen, Birkhoff, Il'yashenko devote their great efforts to the development of the theory. Nowadays, it plays an important role in the study of bifurcations and stabilities and is also widely applied to celestial mechanics, biomathematics, control theory, and so on.

First, let us recall some classical results in the normal form theory. Due to celebrated Poincaré-Dulac reductions, many nonlinear terms can be eliminated by formal coordinates substitutions. Beyond this, in the Poincaré domain systems are always analytically conjugated to their normal forms. However, in the Siegel domain more complicated phenomena arise. Without any small divisor, the analytical normalization was shown in [1] for systems preserving special structures. Then under Siegel's and Bruno's small divisor conditions the corresponding results of the analytical linearization were confirmed. Moreover, by the hyper-

✉ Hao Wu
Haow@seu.edu.cn
Xingdong Xu
xindong.xu@seu.edu.cn
Dongfeng Zhang
zhdf@seu.edu.cn

¹ Department of Mathematics, Southeast University, Nanjing 210096, Jiangsu, People's Republic of China

bolic condition results about C^1 and C^∞ linearization were got by Hartman and Sternberg, together with the C^k and C^∞ normalization by Sternberg and Chen, respectively. See [3] for more details.

Recently, to bridge the gap between analytical and C^∞ conjugations, the ultradifferentiable topology, which is the refinement of C^∞ topology and was once introduced in [10] by Rudin, emerges before our eyes. In the Gevrey smooth case, it was proved in [11] that the Gevrey- α smooth vector fields can be changed into their normal forms by the Gevrey- $(\alpha + \mu + 1)$ smooth coordinates substitutions at the origin for the hyperbolic linear part and the Siegel type small divisor condition with the index μ , which was improved into an accurate one in [13] for the diagonal linear part. Meanwhile, more degenerated formal Gevrey- α vector fields were studied in [6] and see also [2] for the Gevrey linearization. Furthermore, the Siegel-Sternberg linearization theorem for ultradifferentiable systems was given by [7]. So, the task of the work is to explore the theorems about the ultradifferentiable normalization.

In general, the ultradifferentiable functions belong to the C^∞ functional class but their derivatives have certain norm controls with respect to the order. More precisely, for the given positive sequence $\{m(t)\}_{t \in \mathbb{N}}$, the function f is *ultradifferentiable* on the set U in terms of the weights $m(t)$, provided that for any compact set $K \subseteq U$, there exist positive constants M and C such that

$$\sup_{x \in K} |\partial^k f(x)| \leq MC^{|k|} m(|k|)(|k|!), \quad \forall k \in \mathbb{Z}_+^d,$$

where U is an open set in $\mathbb{R}^d$. They are also called to be local Denjoy–Carleman classes of Roumieu type. As shown in [7, 11], the whole proofs of convergence of the conjugations shall contain two steps. The one inherited from the C^∞ topology is to utilize the path method to confirm the Chen type theorem that the formal conjugations lead to the real ones by the hyperbolic non-degenerated condition. The other as the extension of analytic topology is to get the loss of smoothness sharper in the formal norms. Here we restrict our attention on the part of formal conjugations.

Denote $C^d[[x]]$ by the formal Taylor series with complex vector-valued coefficients in $\mathbb{C}^d$. Then the function $f \in C^d[[x]]$ is formal ultradifferentiable with the weight function $E(t)$, provided that there exist positive constant r and

$$\|f\|_{E,r} = \sum_{|k| \geq 1, j} |f_{k,j}| E(|k|) r^{|k|} < \infty, \quad f = \sum_{|k| \geq 1, j} f_{k,j} x^k e_j,$$

where $E(t) = e^{\omega(t)}$ and

(H1) its logarithm $\omega(t): [1, \infty) \rightarrow \mathbb{R}_-$ is the C^1 strictly decreasing function fulfilling that $\omega(1) = 0$, $\omega'(t)$ is decreasing and

$$\lim_{t \rightarrow +\infty} \omega(t) = -\infty.$$

Here e_j refers to the unit vector with the j -th component 1. The above means that the logs of the weights are convex with respect to t . Particularly, the Gevrey- α smooth function f satisfying $f(0) = 0$ is of formal Gevrey- $\alpha + 1$ with weights $E(t) = e^{-(\alpha+1)t \ln t}$. We now generalize results in [7, 11] to deal with formal ultradifferentiable normalization by all kinds of small divisor conditions.

Consider the system

$$\frac{dx}{dt} = Ax + f(x), \quad (1.1)$$

where $f = O(\|x\|^2)$ as $x \rightarrow 0$. Set

$$\Lambda_{nr} = \left\{ (k, j) \mid k \cdot \lambda \neq \lambda_j, |k| = \sum_{i=1}^d |k_i| \geq 2, j = 1, \dots, d \right\},$$

then the small divisor condition is given by

$$|k \cdot \lambda - \lambda_j| \geq \frac{c}{\Omega(|k|)}, \quad (k, j) \in \Lambda_{nr}, \quad (1.2)$$

where $k \cdot \lambda = \sum_{i=1}^d k_i \lambda_i$ and $\Omega(t) : [2, \infty) \rightarrow \mathbb{R}_+$ is the increasing continuous function. By properly choosing the constant c and extending its definition domain, technically we make that Ω is continuous and increasing in $[1, \infty)$ and $\Omega(1) = 1$. Without concerning the case that λ is in the Poincaré domain, $\Omega(t)$ is assumed to be either

$$\Omega(t) = 1, \quad t \geq 1, \quad (1.3)$$

or

$$\lim_{t \rightarrow \infty} \Omega(t) = +\infty. \quad (1.4)$$

When A is in the Jordan normal form, the formal normal form of (1.1) is

$$\frac{dx}{dt} = Ax + g(x)$$

by the Poincaré-Dulac formal normal form reductions. Let

$$\text{Ord}(g) = \min \left\{ |k| \mid j : g_{k,j} \neq 0, g = \sum_{|k| \geq 2, j} g_{k,j} x^k e_j \right\}.$$

The following shows that all small divisors can be formally excluded by the proper loss of the ultradifferentiable smoothness.

Theorem 1.1 Assume that system (1.1) is formally ultradifferentiable with the weight function $E(t) = e^{\omega(t)}$ satisfying (H1), $A = \text{diag}(\lambda_1, \dots, \lambda_d)$ is in the diagonal form and $q = \text{Ord}(g) \geq 2$. Under the small divisor condition (1.2) given by (1.4) there exists a formal coordinates substitution with weight function $\tilde{E}(t) = e^{\omega(t) + \hat{\omega}(t)}$, which turns system (1.1) into its formal normal form, where

$$\hat{\omega}(t) = - \int_1^t \ln \Omega(u + 2) du - \frac{t \ln t}{q - 1}. \quad (1.5)$$

Additionally, if system (1.1) can be linearized, the above result is also valid as $q \rightarrow +\infty$.

The next is generalized from [13] to confirm the stop of the loss of smoothness for the ‘larger’ small divisor. Consider the following small divisor condition (1.2) fixed with

$$\Omega_\mu(t) = e^{\mu t^\tau} \quad (1.6)$$

for $\tau \in (0, \infty)$ and $\mu > 0$. It is of the Liuvillian type fulfilling

$$\limsup_{t \geq 2} \frac{\ln \ln \Omega(t)}{\ln t} < \infty.$$

Let $E_s(t) = e^{\omega_s(t)}$ be the weight function, where the logarithm $\omega_s(t)$ with $s = (s_1, s_2)$ satisfies

- (C1) $\omega_s(t) = -s_1(t^{\tau+1} - 1) - s_2 t \ln t$, $s_1 \geq \frac{\mu}{(\tau+1)(q-1)}$ and $s_2 \geq \frac{1}{q-1}$ for $q = \text{Ord}(g) \geq 2$ and $\tau \in (0, 1)$; or
- (C2) $\omega_s(t) = -s_1(t^{\tau+1} - 1) - s_2(t^\eta - 1)$, $s_1 \geq \frac{\mu}{(\tau+1)(q-1)}$ and $s_2 \geq s_{2,0} > 0$ for $q = \text{Ord}(g) \geq 2$, $\tau \geq 1$, $\eta \in (\tau, \tau + 1]$ and the preassigned positive constant $s_{2,0}$.

Theorem 1.2 Assume that $A = \text{diag}(\lambda_1, \dots, \lambda_d)$ is in the diagonal form and the small divisor condition (1.2) given by (1.6) is satisfied in system (1.1). The following statements hold.

- (i) If system (1.1) is formally ultradifferentiable with the wight function $E_s(t) = e^{\omega_s(t)}$ fulfilling condition (C1) for $\tau \in (0, 1)$, there exists a formally ultradifferentiable coordinates substitution of the same class turning system (1.1) into its normal form.
- (ii) If system (1.1) is formally ultradifferentiable with the wight function $E_s(t) = e^{\omega_s(t)}$ fulfilling condition (C2) for $\tau \geq 1$, there exists a formally ultradifferentiable coordinates substitution with the similar wight function $E_{\hat{s}}(t) = e^{\omega_{\hat{s}}(t)}$ turning system (1.1) into its normal form, where $\hat{s} = (s_1, \hat{s}_2)$ with $\hat{s}_2 > s_2$ and $s = (s_1, s_2)$ is from the log weights $\omega_s(t)$.

Additionally, if system (1.1) can be linearized, above results are also valid as $q \rightarrow +\infty$.

At last, if there is no small divisor, we show that the Gevrey smoothness also does some good to make the clearer classification of nonlinear terms. When A is in the diagonal form, by (1.3) of condition (1.2) the formal normal form of system (1.1) has such further decomposition

$$\frac{dx}{dt} = Ax(1 + \gamma(x)) + \hat{g}(x), \quad (1.7)$$

where $\gamma = O(\|x\|)$ is a scalar-valued function as $x \rightarrow 0$, the other resonant terms are in $\hat{g}$ and $\hat{g} = O(\|x\|^2)$ as $x \rightarrow 0$. Denote the lowest degree of resonant monomials in $\hat{g}$ by $\text{Ord}(\hat{g}) = q$ for $\hat{g}(x) = \sum_{|k| \geq p, j} \hat{g}_{k,j} x^k e_j$. Here e_j refers to the unit vector with the j -th component 1 as above. By results of the classical textbook [1] system (1.1) can be analytic normalization, provided the original system is analytic and $\hat{g}$ vanishes. Particularly, it was proved in [16] that completely integrable systems, which have sufficiently many independent first integrals and are the partial case of [18], shall obey this convergent criterion. Note again that the formal Gevrey- s function f is just formally ultradifferentiable with weight function $E_s(t) = e^{-st \ln t}$. Thus, reviewing it in the ultradifferentiable category, we get the following.

Theorem 1.3 Assume that system (1.1) is formal Gevrey- s , A is in the diagonal form and $\text{Ord}(\hat{g}) = q \geq 2$ in system (1.7). Under (1.3) of condition (1.2) there exists a formal Gevrey- $\hat{s}$ coordinates substitution turning system (1.1) into its normal form, where $\hat{s} = \max\{s, \frac{1}{q-1}\}$. Additionally, if $\hat{g} = 0$, the above result is also valid as $q \rightarrow +\infty$.

Here we do several remarks. First, by proper non-degenerated conditions, which were mentioned by Theorem 8.1 (pp. 24) in [7] for our Theorems 1.1 and 1.2 and Theorem 3.2 (pp. 254) in [11] for our Theorem 1.3, the above formal conjugations can be modified into really ultradifferentiable ones. Next, if the matrix A in system (1.1) has the nilpotent part, similar results as Theorems 1.1 and 1.2 under condition (C2) are valid by using $\|ad_A^{-1}\|_o$ in (6.1) instead. At last, the case admitting Theorem 1.3 appears at the planar degenerated Hopf and homoclinic bifurcations. Since there occurs no small divisor, we think that Gevrey smoothness is enough to catch the order of focus and saddle points quantitatively.

Due to Taguchi's comments in [12], the purpose of Ramis to use formal power series is to overcome some difficulties, when the usual technique of functional analysis breaks down. In the view of our series results, out of the Poincaré domain, it is enough to apply formal Gevrey

conjugacy to classify different resonant terms without any small divisors by Theorem 1.3. When Siegel type small divisors appear, the formal loss of smoothness for the normalization stops in the same formal Gevrey class but with different Gevrey indices by [13]. Now turning to (i) of Theorem 1.2, the formal loss of smoothness for the ultradifferentiable normalization stops in the class with different weight functions, while the slight loss must be allowed in (ii) of Theorem 1.2. At last, the large loss occurs to overcome all kinds of small divisors by Theorem 1.1. Therefore, we think that the ultradifferentiable smoothness is useful to detect fine structures between the analytic and C^∞ normalization.

The rest parts are organized as follows. In Sect. 2, notations, basic definitions, and key lemmas are provided. Then Theorems 1.1 and 1.2 are proved in Sect. 3 by Contraction Mapping Principle, while Sect. 4 contains the proof of Theorem 1.2 via KAM steps. At last, we show the connections between Gevrey conjugations and bifurcations as the application in Sect. 5.

2 Preliminary

In this part, we provide notations, basic definitions, and key lemmas. All notations using frequently in this part are listed as follows.

Denote $C^d[[x]]$ by the formal Taylor series with complex vector-valued coefficients in $\mathbb{C}^d$. Then we restrict our focus on the set of formal series

$$\mathcal{X} = \left\{ f \mid f = \sum_{|k| \geq 1, j} f_{k,j} x^k e_j \in C^d[[x]] \right\}$$

fixed with the ultradifferentiable norm

$$\|f\|_{\omega,r} = \sum_{k,j} |f_{k,j}| e^{\omega(|k|)r^{|k|}}, \quad f \in \mathcal{X}. \quad (2.1)$$

Here the logarithm weight function $\omega(t): [1, \infty) \rightarrow \mathbb{R}_-$ is a C^1 strictly decreasing function such that $\omega(1) = 0$, $\lim_{t \rightarrow +\infty} \omega(t) = -\infty$, $\omega'(t)$ is non-positive and decreasing. Moreover, by setting $E(t) = e^{\omega(t)}$, sometimes we use the notation

$$\|f\|_{E,r} = \sum_{k,j} |f_{k,j}| E(|k|)r^{|k|}, \quad f \in \mathcal{X}. \quad (2.2)$$

Additionally, denote two index sets by

$$\Lambda_r = \{(k, j) \mid \lambda \cdot k = \lambda_j, |k| \geq 2, j = 1, \dots, d\}$$

and

$$\Lambda_{nr} = \{(k, j) \mid \lambda \cdot k \neq \lambda_j, |k| \geq 2, j = 1, \dots, d\}.$$

Then setting $\langle f \rangle_r = \sum_{(k,j) \in \Lambda_r} f_{k,j} x^k e_j$ and $\langle f \rangle_{nr} = f - \langle f \rangle_r$, we make

$$\begin{aligned} \mathcal{X}_q &= \left\{ f \in \mathcal{X} \mid f = \sum_{k,j} f_{k,j} x^k e_j, |k| \geq q \right\}, \\ \mathcal{X}_{nr} &= \left\{ f \in \mathcal{X} \mid f = \sum_{(k,j) \in \Lambda_{nr}} f_{k,j} x^k e_j \right\}, \end{aligned}$$

$\mathcal{X}_r = \mathcal{X} \setminus \mathcal{X}_{nr}$ and $\mathcal{X}_{nr,q} = \mathcal{X}_q \cap \mathcal{X}_{nr}$.

First of all, we study general properties of the formal space $(\mathcal{X}, \|f\|_{\omega,r})$, which is the subspace of formal series from $C^d[[x]]$ having the finite $\|\cdot\|_{\omega,r}$ norm.

Lemma 2.1 *The set $(\mathcal{X}, \|f\|_{\omega,r})$ is a complete Banach space.*

Proof For any $f \in \mathcal{X}$ we build

$$\hat{f} = \sum_{k,j} \hat{f}_{k,j} x^k e_j, \quad \hat{f}_{k,j} = f_{k,j} e^{\omega(|k|)} r^{|k|},$$

which yields a complete Banach space l^1 , when it is fixed with the norm $\|\hat{f}\| = \sum_{k,j} |\hat{f}_{k,j}|$. So it is equivalent to say that the space $(\mathcal{X}, \|f\|_{\omega,r})$ is just weighted l^1 , which confirms the result. $\square$

Especially, comparing with the classical formal Gevrey norm

$$\natural f \natural_{s,r} = \sum_{k,j} |f_{k,j}| \frac{r^{|k|}}{(|k|!)^s}, \quad f \in \mathcal{X},$$

we would like to use the following one instead

$$\|f\|_{s,r} = \sum_{k,j} |f_{k,j}| e^{-s|k| \ln |k|} r^{|k|}, \quad f \in \mathcal{X}.$$

For the completeness of the work, we write details down to confirm the equivalence of these norms.

Lemma 2.2 *There exists a positive constant*

$$c = e^{-s} \min_{t \in \mathbb{N}} \left\{ \frac{2^t}{(2\pi t)^{\frac{s}{2}}} \right\}$$

such that

$$c \|f\|_{s,e^s(r/2)} \leq \natural f \natural_{s,r} \leq \|f\|_{s,e^s r}.$$

Proof From [8] we can take a precise control

$$\sqrt{2\pi} t^{t+\frac{1}{2}} e^{-t} e^{\frac{1}{12t+1}} \leq t! \leq \sqrt{2\pi} t^{t+\frac{1}{2}} e^{-t} e^{\frac{1}{12t}}$$

for all $t \in \mathbb{N}$. Therefore, from the above, it leads to

$$\begin{aligned} \frac{r^{|k|}}{(|k|!)^s} &\leq \frac{r^{|k|}}{\left(\sqrt{2\pi} |k|^{k+\frac{1}{2}} e^{-|k|} e^{\frac{1}{12|k|+1}} \right)^s} \\ &= \frac{1}{\left(\sqrt{2\pi} |k| e^{\frac{1}{12|k|+1}} \right)^s} e^{-s|k| \ln |k|} (e^s r)^{|k|} \leq e^{-s|k| \ln |k|} (e^s r)^{|k|}. \end{aligned}$$

Then from the below, we obtain that

$$\frac{r^{|k|}}{(|k|!)^s} \geq \frac{r^{|k|}}{\left(\sqrt{2\pi} |k|^{k+\frac{1}{2}} e^{-|k|} e^{\frac{1}{12|k|}} \right)^s}$$

$$\begin{aligned}
&= \frac{1}{\left(\sqrt{2\pi}|k|e^{\frac{1}{12|k|}}\right)^s} e^{-s|k|\ln|k|} (e^s r)^{|k|} \\
&\geq \frac{2^{|k|}}{\left(\sqrt{2\pi}|k|\right)^s} e^{-s} e^{-s|k|\ln|k|} \left(e^s \left(\frac{r}{2}\right)\right)^{|k|} \geq c e^{-s|k|\ln|k|} \left(e^s \left(\frac{r}{2}\right)\right)^{|k|}.
\end{aligned}$$

This completes the proof. $\square$

By those arguments we take $(\mathcal{X}, \|\cdot\|_{\omega,r})$ or $(\mathcal{X}, \|\cdot\|_{E,r})$ fixed with the norm given by (2.1) or (2.2) for formal ultradifferentiable functions, respectively. Additionally, we use $\|\cdot\|_{s,r}$ instead for the formal Gevrey- s functional class. Then comes the study of the weight function.

Lemma 2.3 *Let the C^1 function $\omega(t): [1, \infty) \rightarrow \mathbb{R}_-$ be strictly decreasing satisfying $\omega(1) = 0$ and $\lim_{t \rightarrow +\infty} \omega(t) = -\infty$, $\omega'(t)$ is non-positive and decreasing. Set $E(t) = e^{\omega(t)}$. Then the following statements hold.*

(a) *It admits*

$$E\left(\sum_{i=1}^m t_i\right) \leq E(m) \prod_{i=1}^m E(t_i), \quad t_i \geq 1, \quad (2.3)$$

which implies

$$E(t_1 + t_2) \leq E(t_1)E(t_2). \quad (2.4)$$

(b) *When $u \geq \beta \geq 1$, $v \geq \beta \geq 1$, and $0 \leq \gamma < \beta$, we have that*

$$E(u + v - \gamma) \leq c_1 \kappa_1(u + v - \beta)E(u)E(v), \quad (2.5)$$

where $c_1 = e^{-\omega(\beta)}$ and $\kappa_1(u) = e^{\omega'(u)(\beta-\gamma)}$. Additionally, assume that $\omega \in C^l$ for $l \geq 2$ and there exists a constant $M > 0$ such that $|\omega^{(l)}(t)| \leq M$ for all $t \geq 1$, we have that

$$E(u + v - \gamma) \leq c_2 \kappa_2(u + v - \gamma)E(u)E(v), \quad (2.6)$$

where $\kappa_2(u) = e^{P(u)}$,

$$P(u) = \omega'(u)(\beta - \gamma) + \cdots + (-1)^l \frac{\omega^{(l-1)}(u)}{(l-1)!} (\beta - \gamma)^{l-1}$$

and

$$c_2 = \exp(-\omega(\beta) + M(\beta - \gamma)^l/l!).$$

Proof Notice that $\omega(t) = \int_1^t \omega'(u)du$ for $t \geq 1$. Set $s_m = \sum_{i=1}^m t_i$. Since ω' is decreasing, it yields that

$$\omega(t_i) = \int_1^{t_i} \omega'(u)du = \int_v^{t_i+v-1} \omega'(u-v+1)du \geq \int_v^{t_i+v-1} \omega'(u)du$$

for any $v \geq 1$, which implies

$$\begin{aligned}
\omega(t_1) + \omega(t_2) &= \int_1^{t_1} \omega'(u)du + \int_1^{t_2} \omega'(u)du \\
&\geq \int_1^{t_1} \omega'(u)du + \int_{t_1}^{t_1+t_2-1} \omega'(u)du = \omega(t_1 + t_2 - 1)
\end{aligned}$$

and

$$\sum_{i=1}^m \omega(t_i) + \omega(m) \geq \omega(s_m).$$

Therefore, (2.3) of result (a) is verified, which leads to (2.4) by the fact $\omega(m) \leq 0$ and taking $m = 2$.

Now consider the function

$$F_w(t) = \omega(t) + \omega(w - t),$$

where w is a parameter, $\beta \geq 1$, $t \geq \beta$, and $w - t \geq \beta$. Taking derivatives with respect to the variable t , we obtain that $F'_w(t) = \omega'(t) - \omega'(w - t)$, which is non-negative for $t \leq w/2$ and non-positive for $t \geq w/2$. Thus we obtain that

$$F_w(t) \geq \omega(\beta) + \omega(w - \beta). \quad (2.7)$$

Using the Taylor expansion with the Lagrange type remainder, we have that

$$\begin{aligned} \omega(w - \beta) &= \omega(w - \gamma) - \omega'(w - \xi)(\beta - \gamma) \\ &\geq \omega(w - \gamma) - \omega'(w - \beta)(\beta - \gamma), \end{aligned} \quad (2.8)$$

for $0 \leq \gamma < \beta$ and $\xi \in (\gamma, \beta)$. In another form, the above is

$$\omega(w - \gamma) \leq -\omega(\beta) + \omega'(w - \beta)(\beta - \gamma) + F_w(t),$$

which confirms (2.5) of result (b). By additional assumptions we use

$$\begin{aligned} \omega(w - \beta) &= \omega(w - \gamma) - P(w - \gamma) + \frac{(-1)^l \omega^{(l)}(\eta)}{l!} (\beta - \gamma)^l \\ &\geq \omega(w - \gamma) - P(w - \gamma) - \frac{M}{l!} (\beta - \gamma)^l \end{aligned}$$

instead of (2.8), where $\eta \in (w - \beta, w - \gamma)$ and

$$P(u) = \omega'(u)(\beta - \gamma) + \cdots + (-1)^l \frac{\omega^{(l-1)}(u)}{(l-1)!} (\beta - \gamma)^{l-1}.$$

Similarly, we get the control

$$\omega(w - \gamma) \leq -\omega(\beta) + \frac{M}{l!} (\beta - \gamma)^l + P(w - \gamma) + F_w(t),$$

which proves (2.6). $\square$

Next, we deal with norms of multiplicities and compositions of formal functions. As usual, denote by the majorant operator

$$\mathcal{M}_E(f) = \sum_{k,j} |f_{k,j}| E(|k|) x^k e_j.$$

Naturally, for the general $f \in \mathcal{X}$, we have that

$$\|f\|_{E,r} = \|\hat{f}\|_{1,r} = \mathcal{M}_1(\hat{f})(r, \dots, r),$$

where $\hat{f}_{k,j} = |f_{k,j}| E(|k|)$ and

$$\hat{f} = \sum_{k,j} \hat{f}_{k,j} x^k e_j, \quad f \in \mathcal{X}. \quad (2.9)$$

In the book [15, Lemma 5.10, pp. 51], the following lemma was mentioned.

Lemma 2.4 *Let $E \equiv 1$ and $f, g \in \mathcal{X}$. Then the following statements hold.*

- (a) $\mathcal{M}_1(f \cdot g) \leq \mathcal{M}_1(f)\mathcal{M}_1(g)$,
- (b) $\mathcal{M}_1(f \circ g) \leq \mathcal{M}_1(f) \circ \mathcal{M}_1(g)$.

Here the notation $f \leq g$ denotes $|f_{k,j}| \leq |g_{k,j}|$ for all possible k and j .

Lemma 2.5 *Assume that ω and E are the same as the ones in Lemma 2.3. Then the following statements hold.*

- (a) *We have that $\|f \cdot g\|_{E,r} \leq \|f\|_{E,r} \|g\|_{E,r}$, provided that $\|f\|_{E,r} < \infty$ and $\|g\|_{E,r} < \infty$.*
- (b) *We have that $\|f \circ g\|_{E,r} \leq \|f\|_{E,\tau}$, provided that $\|g\|_{E,r} \leq \tau < \infty$ and $\|f\|_{E,\tau} < \infty$.*

Proof From (2.4) of Lemma 2.3(a) we have that

$$|f_{k,j} g_{\hat{k},j}| E(|k + \hat{k}|) \leq |f_{k,j}| E(|k|) |g_{\hat{k},j}| E(|\hat{k}|),$$

which implies $\mathcal{M}_1(\widehat{f \cdot g}) \leq \mathcal{M}_1(\hat{f} \cdot \hat{g})$ and

$$\begin{aligned} \|f \cdot g\|_{E,r} &= \mathcal{M}_1(\widehat{f \cdot g})(r, \cdot, r) \leq \mathcal{M}_1(\hat{f} \cdot \hat{g}) \\ &\leq \mathcal{M}_1(\hat{f}) \mathcal{M}_1(\hat{g})(r, \dots, r) = \|f\|_{E,r} \|g\|_{E,r} \end{aligned}$$

by Lemma 2.4(a). Here $\hat{f}$ is given by formula (2.9).

Then we compare the form of typical terms of $\widehat{f \circ g}$ with ones of $\hat{f} \circ \hat{g}$. Arbitrary choosing k and rewriting g^k in a precise form we obtain that

$$g^k = \prod_t g_t^{k_t} = \sum_{\hat{k}} \left(\prod_t \prod_{u=1}^{k_t} g_{\hat{k}^{(t,u)},t} \right) x^{\hat{k}}$$

for $\hat{k} = \sum_t \sum_{u=1}^{k_t} \hat{k}^{(t,u)}$. From (2.3) of Lemma 2.3(a) we can similarly get that

$$E(|k|) \left(\prod_t \prod_{u=1}^{k_t} |g_{\hat{k}^{(t,u)},t}| E(|\hat{k}^{(t,u)}|) \right) \geq E(|\hat{k}|) \left(\prod_t \prod_{u=1}^{k_t} |g_{\hat{k}^{(t,u)},t}| \right),$$

which implies $\widehat{f \circ g} \leq \hat{f} \circ \hat{g}$ and

$$\begin{aligned} \|f \circ g\|_{E,r} &= \mathcal{M}_1(\widehat{f \circ g})(r, \dots, r) \leq \mathcal{M}_1(\hat{f} \circ \hat{g})(r, \dots, r) \\ &\leq \mathcal{M}_1(\hat{f}) \circ \mathcal{M}_1(\hat{g})(r, \dots, r) \leq \|f\|_{E,\tau} \end{aligned}$$

by Lemma 2.4(b). This completes the proof. $\square$

As usual, $\nabla_x f$ is the gradient function of f with respect to the variable x . The following two are key to this part.

Lemma 2.6 *Assume that ω and E are the same as the ones in Lemma 2.3. Then we have that*

$$\|\nabla_x f \cdot g\|_{E,r} \leq e^{-1} \delta^{-1} r^{-1} \|f\|_{E, r e^\delta} \|g\|_{E,r}$$

for $\delta > 0$.

Proof Using the fact that $\sup_{u \geq 0} \{ue^{-\delta u}\} = (e\delta)^{-1}$ and by (2.5) of Lemma 2.3(b), directly we have that

$$\begin{aligned} \|\nabla_x f \cdot g\|_{E,r} &\leq \sum_{j,t} \sum_{t+1=|\hat{k}|+|k|} |k_j f_{k,j} g_{\hat{k},j}| E(t) r^t, \\ &\leq r^{-1} \sum_{j,t} \sum_{t+1=|\hat{k}|+|k|} |k_j f_{k,j} g_{\hat{k},j}| E(|\hat{k}|) E(|k|) r^{t+1} \\ &\leq r^{-1} \sum_{j,t} \sum_{t+1=|\hat{k}|+|k|} |k_j f_{k,j}| E(|k|) r^{|k|} |g_{\hat{k},j}| E(|\hat{k}|) r^{|\hat{k}|} \\ &\leq r^{-1} \sum_{j,t} \sum_{t+1=|\hat{k}|+|k|} |k_j| e^{-\delta|k|} |f_{k,j}| E(|k|) (r e^\delta)^{|k|} |g_{\hat{k},j}| E(|\hat{k}|) r^{|\hat{k}|} \\ &\leq (re\delta)^{-1} \|f\|_{E, re^\delta} \|g\|_{E,r}. \end{aligned}$$

This completes the proof. $\square$

Denote the operator $\mathcal{P}_\Omega$ by

$$\mathcal{P}_\Omega(f) = \sum_{k,j} \Omega(|k|) |f_{k,j}| x^k e_j, \quad f \in \mathcal{X}.$$

Lemma 2.7 Assume that $\Omega(t) = \Omega_\mu(t)$ is given by (1.6), $E_s(t) = e^{\omega_s(t)}$, $\eta \in (\tau, \tau + 1]$, and $\eta \geq 1$. Then the following statements hold.

(a) Set $\omega_s(t) = -s(t^\eta - 1)$. We have that

$$\|\mathcal{P}_{\Omega_\mu}(f)\|_{E_{s+\delta},r} \leq c_3 e^{\delta^{-\frac{\tau}{\eta-\tau}}} \|f\|_{E_{s,r}}$$

for $0 < \delta \leq 1$ and $c_3 = \exp(1 + (\eta - \tau)(\eta^{-1}\mu)^{\eta/(\eta-\tau)} \tau^{\tau/(\eta-\tau)})$.

(b) In the case that $\eta = 1$, set $\omega_s(t) = -s_1(t^{\tau+1} - 1) - s_2 t \ln t$ for $s = (s_1, s_2)$. When $q = \text{Ord}(f) = \text{Ord}(g) \geq 2$, $s_1 \geq \frac{\mu}{(\tau+1)(q-1)}$, and $s_2 \geq \frac{1}{q-1}$, we have that

$$\|\mathcal{P}_{\Omega_\mu}(\nabla_x f \cdot g)\|_{E_{s,r}} \leq c_4 r^{-1} \|f\|_{E_{s,r}} \|g\|_{E_{s,r}}, \quad (2.10)$$

where $c_4 = c_2(q, 1)$. In the case that $\eta > 1$, set $\omega_s(t) = -s_1(t^{\tau+1} - 1) - s_2(t^\eta - 1)$ for $s = (s_1, s_2)$. When $q = \text{Ord}(f) = \text{Ord}(g) \geq 2$, $s_1 \geq \frac{\mu}{(\tau+1)(q-1)}$, and $s_2 \geq s_{2,0} > 0$ for the preassigned positive constant $s_{2,0}$, then inequality (2.10) is still valid for $c_4 = c_2(q, 1)e^{M_1}$, where

$$M_1 = \max_{u \geq 1} \{(s_1 - s_{2,0})\eta(q-1)u^{\eta-1} + \ln u + P(u) + s_1(\tau+1)(q-1)u^\tau\}.$$

Here $c_2(\beta, \gamma)$ and $P(u)$ are the same as the ones in Lemma 2.3(b).

Proof Taking the derivative to the function $f(x) = \mu x^{\frac{\tau}{\eta}} - \delta x$ for $x \geq 0$, we get the unique extreme point x_* satisfying $f'(x_*) = \frac{\mu\tau}{\eta} x_*^{\frac{\tau}{\eta}-1} - \delta = 0$, which implies that

$$\max_{x \geq 0} \{\mu x^\tau - \delta x^\eta\} = \mu \left(1 - \frac{\tau}{\eta}\right) \left(\frac{\eta\delta}{\mu\tau}\right)^{\frac{\tau}{\tau-\eta}}.$$

So it admits

$$\|\mathcal{P}_{\Omega_\mu}(f)\|_{E_{s+\delta},r} = \sum_{|k| \geq 1, j} e^{\mu|k|^\tau} e^{-(s+\delta)(|k|^\eta-1)} |f_{k,j}| r^{|k|}$$

$$\begin{aligned}
&\leq c_3 e^{\delta^{-\frac{\tau}{\eta-\tau}}} \sum_{|k| \geq 1, j} e^{-s(|k|^\eta - 1)} |f_{k,j}| r^{|k|} \\
&= c_3 e^{\delta^{-\frac{\tau}{\eta-\tau}}} \|f\|_{E_s, r},
\end{aligned}$$

which confirms (a).

When $\eta = 1$, we have that $\tau \in (0, 1)$ and $\tau + 1 \in (1, 2)$. Set $\beta = q$ and $\gamma = 1$ in inequality (2.6) and notice that

$$P(u) = -(s_1(\tau + 1)u^\tau + s_2 \ln u + s_2)(q - 1)$$

of that κ_2 . By careful computations, the estimation in (2.6) of Lemma 2.3(b) leads to

$$\begin{aligned}
\|\mathcal{P}_{\Omega_\mu}(\nabla_x f \cdot g)\|_{E_s, r} &= \|\mathcal{P}_{\Omega_\mu} \left(\sum_j \partial_{x_j} f_j \cdot g_j \right)\|_{E_s, r} \\
&\leq \sum_{j, t} \sum_{|k| + |\hat{k}| = t+1} e^{\mu t^\tau} |k_j| |f_{k,j} g_{\hat{k},j}| E(t) r^t \\
&\leq \sum_{j, t} \sum_{|k| + |\hat{k}| = t+1} e^{\mu t^\tau + \ln t} |f_{k,j} g_{\hat{k},j}| c_4 e^{-(s_1(\tau+1)t^\tau + s_2 \ln t)(q-1)} E(|k|) E(|\hat{k}|) r^t \\
&\leq c_4 r^{-1} \|f\|_{E_s, r} \|g\|_{E_s, r},
\end{aligned}$$

for $|k|, |\hat{k}| \geq q$, $s_1 \geq \frac{\mu}{(\tau+1)(q-1)}$, and $s_2 \geq \frac{1}{q-1}$.

When $\eta > 1$, we have that

$$P(u) = -s_1(\tau + 1)(q - 1)u^\tau - s_2\eta(q - 1)u^{\eta-1} + R(u)$$

and

$$\begin{aligned}
R(u) &= \frac{s_1(q-1)^2}{2!} (\tau + 1) \tau u^{\tau-1} + \dots \\
&\quad + \frac{(-1)^{[\tau]+1} s_1(q-1)^{[\tau]+1}}{([\tau] + 1)!} (\tau + 1) \tau \cdots (\tau - [\tau] + 1) u^{\tau-[\tau]} \\
&\quad + \frac{s_2(q-1)^2}{2!} \eta(\eta-1) u^{\eta-2} + \dots + \frac{(-1)^{[\eta]} s_2(q-1)^{[\eta]}}{[\eta]!} \eta \cdots (\eta - [\eta] + 1) u^{\eta-[\eta]}
\end{aligned}$$

by setting $\beta = q$ and $\gamma = 1$ in inequality (2.6) again. Since $\eta > 1$ and $\eta > \tau$, we have that

$$\lim_{t \rightarrow \infty} -s_{2,0}\eta(q-1)t^{\eta-1} + \ln t + R(t) = -\infty.$$

So we make

$$M_1 = \max_{t \geq 1} \{-s_{2,0}\eta(q-1)t^{\eta-1} + \ln t + R(t)\},$$

which similarly implies

$$\begin{aligned}
\|\mathcal{P}_{\Omega_\mu}(\nabla_x f \cdot g)\|_{E_s, r} &\leq \sum_j \sum_{|k| + |\hat{k}| \geq t+1} e^{\mu t^\tau} |k_j| |f_{k,j} g_{\hat{k},j}| E(t) r^t \\
&\leq \sum_j \sum_{|k| + |\hat{k}| \geq t+1} e^{\mu t^\tau + \ln t} |f_{k,j} g_{\hat{k},j}| c_4 e^{-s_1(\tau+1)(q-1)t^\tau - \ln t} E(|k|) E(|\hat{k}|) r^t \\
&\leq c_4 r^{-1} \|f\|_{E_s, r} \|g\|_{E_s, r},
\end{aligned}$$

where $c_4 = c_2(q, 1)e^{M_1}$. This completes the proof. $\square$

Especially, when we restrict our focus on the normalization with small divisors of the Diophantine type, the classical Gevrey smoothness is proper. In our category, the key lemma in [13] can be represented as follows.

Lemma 2.8 Assume that $\Omega(t) = \Omega_\mu(t)$ is given by $\Omega_\mu(t) = e^{\mu \ln t}$, $E_s(t) = e^{\omega_s(t)}$, and $\omega_s(t) = st \ln t$. When $q = \text{Ord}(f) = \text{Ord}(g) \geq 2$ and $s \geq \frac{\mu+1}{q-1}$, inequality (2.10) is still valid for $c_4 = c_2(q, 1)$.

Moreover, in the general case, we have the following lemma.

Lemma 2.9 Assume that $q = \text{Ord}(f) = \text{Ord}(g) \geq 2$, $\hat{\omega}(t)$ is given by (1.5), E and $\tilde{E}$ are the same as the ones in Theorem 1.1. Then we have that

$$\|\mathcal{P}_\Omega(\nabla_x f \cdot g)\|_{\tilde{E},r} \leq c_4 r^{-1} \|f\|_{E,r} \|g\|_{\tilde{E},r},$$

where $c_4 = c_1(2, 1)c_2(2, 1)e^{-1}$. Here constants c_1 and c_2 are the same as the ones in Lemma 2.3(b).

Proof Let $\omega_1 = -\int_1^t \ln \Omega(u+2)du$, $\omega_2 = -t \ln t / (q-1)$, and $E_i(t) = e^{\omega_i(t)}$ for $i = 1$ and 2 . So $\tilde{E}(t) = E(t)E_1(t)E_2(t)$. Assume that $|k|, |\hat{k}| \geq q = 2$, and $|k| + |\hat{k}| = t + 1$. Then by setting respectively $(\beta, \gamma) = (2, 1)$ and $(q, 1)$ in inequality (2.5) and (2.6), we have that

$$E_1(t) \leq c_{4,1} \kappa_1(t-2) E_1(|k|) E_1(|\hat{k}|)$$

for $c_{4,1} = c_1(2, 1)$ and $\kappa_1(u) = e^{\omega'_1(u)} = e^{-\ln \Omega(u+2)}$ and

$$E_2(t) \leq c_{4,2} \kappa_2(t) E_2(|k|) E_2(|\hat{k}|)$$

for $c_{4,2} = c_2(q, 1)$ and $\kappa_2(u) = e^{-\ln u - 1}$, respectively. Together with the trivial one

$$E(t) = E(|k| + |\hat{k}| - 1) E(1) \leq E(|k|) E(|\hat{k}|)$$

from (2.7) for $\beta = 1$, we can obtain that

$$\tilde{E}(t) \leq c_{4,1} c_{4,2} e^{-1} e^{-\ln \Omega(t) - \ln t} \tilde{E}(|k|) \tilde{E}(|\hat{k}|),$$

which yields that

$$\begin{aligned} \|\mathcal{P}_\Omega(\nabla_x f \cdot g)\|_{\tilde{E},r} &= \|\mathcal{P}_\Omega\left(\sum_j \partial_{x_j} f_j \cdot g_j\right)\|_{\tilde{E},r} \\ &\leq \sum_{j,t} \sum_{|k|+|\hat{k}|=t+1} \Omega(t) |k_j| |f_{k,j} g_{\hat{k},j}| \tilde{E}(t) r^t \\ &\leq \sum_{j,t} \sum_{|k|+|\hat{k}|=t+1} e^{\ln \Omega(t) + \ln t} |f_{k,j} g_{\hat{k},j}| \tilde{E}(t) r^t \\ &\leq c_4 r^{-1} \|f\|_{\tilde{E},r} \|g\|_{\tilde{E},r} \leq c_4 r^{-1} \|f\|_{E,r} \|g\|_{\tilde{E},r} \end{aligned}$$

for $c_4 = c_{4,1} c_{4,2} e^{-1}$. This completes the proof. $\square$

3 Contraction mapping principle in the ultradifferentiable normalization

In this part, we provide the proofs of Theorems 1.1 and 1.3 via the Contraction Mapping Principle.

First, we introduce a criterion to apply the Contraction Mapping Principle. Let F be a function or map on $(\mathcal{X}, \|\cdot\|_{E,r})$. It was mentioned in [15] that the function or map F is *strongly contracting*, provided that

- (i) $\|F(h)\|_{E,r} = O(r^2)$,
- (ii) $\|F(h) - F(h')\|_{E,r} = O(r)\|h - h'\|_{E,r}$ as $r \rightarrow 0$,

for $\|h\|_{E,r} \leq r$ and $\|h'\|_{E,r} \leq r$. As usual, $O(1)$ refers to the bounded quantity as $r \rightarrow 0$. Set $\mathcal{L}(\cdot, \cdot)$ to be a real bilinear form on $(\mathcal{X}, \|\cdot\|_{E,r})$.

Lemma 3.1 *Assume that the bilinear form $\mathcal{L}(\cdot, \cdot)$ satisfies*

$$\|\mathcal{L}(h, h')\|_{E,r} \leq cr^\alpha \|h\|_{E,r} \|h'\|_{E,r}$$

and the map F_i admits

$$\|F_i(0)\|_{E,r} = O(r^{\beta_i+1}), \quad \|F_i(h) - F_i(h')\|_{E,r} \leq O(r^{\beta_i})\|h - h'\|_{E,r} \quad (3.1)$$

with $\|h\|_{E,r} \leq r$ and $\|h'\|_{E,r} \leq r$ for $\alpha \leq 0$, $\beta_i \geq 0$, and $i = 1, 2$ as $r \rightarrow 0$. If we have $\alpha + \beta_1 + \beta_2 \geq 0$, then $\mathcal{L}(F_1(\cdot), F_2(\cdot))$ is strongly contracting.

Proof Making $h' = 0$, we note that

$$\begin{aligned} \|F_i(h)\|_{E,r} &\leq \|F_i(0)\|_{E,r} + \|F_i(h) - F_i(0)\|_{E,r} \\ &= O(r^{\beta_i+1}) + O(r^{\beta_i})\|h\|_{E,r} = O(r^{\beta_i+1}) \end{aligned}$$

as $r \rightarrow 0$. By careful computations, we have that

$$\|\mathcal{L}(F_1(0), F_2(0))\|_{E,r} \leq cr^\alpha \|F_1(0)\|_{E,r} \|F_2(0)\|_{E,r} = O(r^{\alpha+\beta_1+\beta_2+2})$$

as $r \rightarrow 0$. Moreover, by the following estimation

$$\begin{aligned} \text{I} &= \|\mathcal{L}(F_1(h), F_2(h)) - \mathcal{L}(F_1(h), F_2(h'))\|_{E,r} \\ &= \|\mathcal{L}(F_1(h), F_2(h) - F_2(h'))\|_{E,r} \\ &\leq cr^\alpha \|F_1(h)\|_{E,r} \|F_2(h) - F_2(h')\|_{E,r} = O(r^{\alpha+\beta_1+\beta_2+1})\|h - h'\|_{E,r} \end{aligned}$$

and the similar one

$$\begin{aligned} \text{II} &= \|\mathcal{L}(F_1(h), F_2(h')) - \mathcal{L}(F_1(h'), F_2(h'))\|_{E,r} \\ &= O(r^{\alpha+\beta_1+\beta_2+1})\|h - h'\|_{E,r}, \end{aligned}$$

these lead to

$$\begin{aligned} \|\mathcal{L}(F_1(h), F_2(h)) - \mathcal{L}(F_1(h'), F_2(h'))\|_{E,r} &\leq \text{I} \\ &+ \text{II} = O(r^{\alpha+\beta_1+\beta_2+1})\|h - h'\|_{E,r} \end{aligned}$$

for $\|h\|_{E,r} \leq r$ and $\|h'\|_{E,r} \leq r$ as $r \rightarrow 0$. So making $\alpha + \beta_1 + \beta_2 \geq 0$, we complete the proof. $\square$

Now consider system

$$\frac{dx}{dt} = Ax + f(x).$$

Then doing the coordinates substitution $x = y + h(y)$ to the above, it yields that

$$\frac{dy}{dt} = Ay + g(y),$$

where $\langle g \rangle_r = g$. By simple computations, we obtain that

$$\partial h Ay - Ah = f(y + h) - g - \partial h g,$$

which is equivalent to

$$\partial h Ay - Ah = \langle f(y + h) \rangle_{nr} - \partial h \langle f(y + h) \rangle_r \quad (3.2)$$

and

$$g(y) = \langle f(y + h) \rangle_r.$$

3.1 Proof of Theorem 1.1

Denote by the linear operator

$$Ad_A(h) = \partial h Ay - Ah, \quad (3.3)$$

whose inverse is

$$Ad_A^{-1}g = \sum_{(k,j) \in \Lambda_{nr}} \frac{g_{k,j}}{k \cdot \lambda - \lambda_j} x^k e_j, \quad g = \langle g \rangle_{nr}.$$

Then Eq. (3.2) can be represented as

$$h = Ad_A^{-1}(\langle S_f(h) \rangle_{nr} - \mathcal{L}_1(h, \langle S_f(h) \rangle_r)) \quad (3.4)$$

in the space $(\mathcal{X}_{nr}, \|\cdot\|_{\tilde{E},r})$ for any $f \in (\mathcal{X}, \|\cdot\|_{E,r})$, where the f -shifted map S_f is given by

$$S_f(h) = f(\text{id} + h) \quad (3.5)$$

and $\mathcal{L}_1(\cdot, \cdot)$ is the bilinear one as follows

$$\mathcal{L}_1(f, g) = \partial f \cdot g. \quad (3.6)$$

The complete subset utilized here is

$$\hat{\mathcal{X}}_{nr,q,r} = \{h \in \mathcal{X}_{nr,q} \mid \|h\|_{E,r} \leq r\}.$$

Especially, we remark $\hat{\mathcal{X}}_{nr,r}$ instead of $\hat{\mathcal{X}}_{nr,2,r}$.

Lemma 3.2 Assume that $q = \text{Ord}(f) = \text{Ord}(g) \geq 2$ and there exists $r_0 > 0$ such that $\|f\|_{E,r_0} < \infty$ for $E(t) = e^{\omega(t)}$, then the following statements hold.

- (i) The map $S_f(\cdot)$ is strongly contracting in $(\hat{\mathcal{X}}_{nr,r}, \|\cdot\|_{E,r})$.
- (ii) The maps $Ad_A^{-1}\langle S_f(\cdot) \rangle_{nr}$ and $Ad_A^{-1}\mathcal{L}_1(\cdot, \langle S_f(\cdot) \rangle_r)$ are strongly contracting in $(\hat{\mathcal{X}}_{nr,q,r}, \|\cdot\|_{\tilde{E},r})$ for $\tilde{E}(t) = e^{\omega(t) + \hat{\omega}(t)}$, where $\hat{\omega}$ and $\mathcal{L}_1(\cdot, \cdot)$ are given by (1.5) and (3.6), respectively.

Proof First, we show that

$$\begin{aligned}\|Ad_A^{-1}\langle S_f(0)\rangle_{nr}\|_{\tilde{E},r} &= c^{-1} \sum_{(k,j) \in \Lambda_{nr}} \Omega(|k|) |f_{k,j}| e^{\omega(|k|) + \hat{\omega}(|k|)} r^{|k|} \\ &\leq c^{-1} \sum_{(k,j) \in \Lambda_{nr}} |f_{k,j}| e^{\omega(|k|)} r^{|k|} \\ &\leq c^{-1} r^2 r_0^{-2} \|f\|_{E,r_0}\end{aligned}$$

for $r < r_0$ by the fact that

$$\begin{aligned}-\int_1^{|k|} \ln \Omega(u+2) du &\leq -\int_{|k|-1}^{|k|} \ln \Omega(u+2) du \\ &= -\ln \Omega(\xi+2) \leq -\ln \Omega(|k|)\end{aligned}$$

for $\xi \in [|k|-1, |k|]$. Then notice that

$$S_f(h) - S_f(h') = \int_0^1 \nabla_x f(x + uh + (1-u)h')(h - h') du,$$

which by Lemma 2.5(b) and 2.9 implies

$$\begin{aligned}\|Ad_A^{-1}\langle S_f(h) - S_f(h')\rangle_{nr}\|_{\tilde{E},r} &\leq c^{-1} \|\mathcal{P}_\Omega(\langle S_f(h) - S_f(h')\rangle_{nr})\|_{\tilde{E},r} \\ &\leq c^{-1} c_4 r^{-1} \|f\|_{\tilde{E},(d+2)r} \|h - h'\|_{\tilde{E},r} \\ &\leq c^{-1} c_4 r^{-1} \|h - h'\|_{\tilde{E},r} \sum_{|k| \geq 2, j} |f_{k,j}| e^{\omega(|k|) + \hat{\omega}(|k|)} (d+2)^{|k|} r^{|k|} \\ &\leq c^{-1} c_4 (d+2)^2 r_0^{-2} r \|f\|_{E,r_0} \|h - h'\|_{\tilde{E},r}\end{aligned}$$

for $(d+2)r \leq r_0$. So the map $Ad_A^{-1}\langle S_f(\cdot)\rangle_{nr}$ is strongly contracting. These also imply the strong contraction of S_f as Ω takes (1.3), which confirms (i).

At last, since the operator $F_1(h) = h$ is a linear bounded one, the map $F_2(h) = \langle S_f(h)\rangle_r$ is strongly contracting, and the bilinear form $\mathcal{L}(\cdot, \cdot) = Ad_A^{-1}\mathcal{L}_1(\cdot, \cdot)$ satisfies Lemma 2.9, the map $Ad_A^{-1}\mathcal{L}_1(\cdot, \langle S_f(\cdot)\rangle_r)$ is strongly contracting by Lemma 3.1 as taking $\alpha = -1$, $\beta_1 = 0$ and $\beta_2 = 1$. This proves (ii) and completes the proof. $\square$

Now we can prove Theorem 1.1.

Proof of Theorem 1.1 Since $q = \text{Ord}(g) \geq 2$, by Poincaré-Dulac formal normal form reductions we can set $\text{Ord}(f) = \text{Ord}(g)$, provided that $q < \infty$. To change system (1.1) into its normal forms, it is equivalent to solve $h \in (\hat{\mathcal{X}}_{nr,q,r}, \|\cdot\|_{\tilde{E},r})$ from functional equation (3.2), whose precise form is (3.4) for any $f \in (\mathcal{X}, \|\cdot\|_{E,r_0})$. It admits a unique solution for the sufficiently small $r > 0$ by Lemma 3.2(ii). As $q \rightarrow \infty$, the part $\mathcal{L}_1(\cdot, \cdot)$ vanishes. So it completes the proof by Lemma 3.2(i). $\square$

3.2 Proof of Theorem 1.3

In this part, the norm to character the formal Gevrey- s functional class is written as $\|\cdot\|_{s,r}$ for $E(t) = e^{-st \ln t}$. Here we have that either $g = \gamma(x)Ax + \hat{g}$ for $\text{Ord}(\hat{g}) = q < \infty$ or $g = \gamma(x)Ax$ and $\hat{g} = 0$ as $q \rightarrow \infty$.

When $q < \infty$, without loss of generality, we can assume that γ satisfies $\gamma(0) = 0$, which is a polynomial of degree $q-1$ at most. Denote by $\langle f \rangle_{q+} = \sum_{|k| \geq q, j} f_{k,j} x^k e_j$ for

$f = \sum_{k,j} f_{k,j} x^k e_j$. Then we define that $\langle f \rangle_{r,q+} = \langle \langle f \rangle_r \rangle_{q+}$ and $\langle f \rangle_{r,-} = \langle f \rangle_r - \langle f \rangle_{r,q+}$. Submitting $g = \gamma(x)Ax + \hat{g}$ into Eq. (3.2) and doing projections, it yields

$$(1 + \gamma)(\partial Ay - Ah) = f(y + h) - \gamma Ay - \hat{g} - \partial h \hat{g} - \gamma Ah$$

or the equivalent form

$$\partial h Ay - Ah = \frac{1}{1 + \gamma} (\langle f(y + h) \rangle_{nr} - \gamma Ah - \partial h \hat{g}), \quad (3.7)$$

$$\begin{aligned} \langle f(y + h) \rangle_{r,-} &= \gamma Ay, \\ \langle f(y + h) \rangle_{r,q+} &= \hat{g}. \end{aligned} \quad (3.8)$$

Regard the classical scalar multiplication operator as a bilinear one, which is denoted by

$$\mathcal{L}_2(g, f) = gf \quad (3.9)$$

for the function g and the map f on a Banach space. As before, the linear operator Ad_A and the map S_f are the same as the ones given by (3.3) and (3.5), respectively. Note again that the complete subset utilized here is

$$\hat{\mathcal{X}}_{nr,q,r} = \{h \in \mathcal{X}_{nr,q} \mid \|h\|_{s,r} \leq r\}.$$

Notice that $\langle f(y + h) \rangle_{r,-} = f$ for any $h \in \mathcal{X}_{nr,q}$. So system (3.7) can be rewritten as

$$Ad_A h = \mathcal{T}_1^1(h) - \mathcal{T}_2^1(h) - \mathcal{T}_3^1(h) \quad (3.10)$$

for $h \in (\mathcal{X}_{nr,q}, \|\cdot\|_{s,r})$, where

$$\mathcal{T}_1^1(h) = \rho(y) \langle S_f(h) \rangle_{nr}, \quad (3.11)$$

$$\mathcal{T}_2^1(h) = \hat{\rho}(y) Ah, \quad (3.12)$$

$$\mathcal{T}_3^1(h) = \rho(y) \mathcal{L}_1(h, \langle S_f(h) \rangle_{r,q+}). \quad (3.13)$$

Here $\rho = 1/(1 + \gamma)$, $\hat{\rho} = \gamma\rho$, and γ is given by

$$\langle f \rangle_{r,-} = \gamma(y) Ay \quad (3.14)$$

from (3.8).

When $q \rightarrow \infty$, i.e. $\hat{g} = 0$, under condition (1.3) system (3.7) has the following form

$$Ad_A h = \mathcal{T}_1^2(h) - \mathcal{T}_2^2(h) \quad (3.15)$$

for $h \in (\mathcal{X}_{nr}, \|\cdot\|_{s,r})$, where

$$\mathcal{T}_1^2(h) = \mathcal{L}_2(\mathbf{g}(h), \langle S_f(h) \rangle_{nr}), \quad (3.16)$$

$$\mathcal{T}_2^2(h) = \mathcal{L}_2(\hat{\mathbf{g}}(h), Ah), \quad (3.17)$$

the function $\mathbf{g} = 1/(1 + \gamma(h))$, $\hat{\mathbf{g}} = \gamma(h)/(1 + \gamma(h))$, and $\gamma = \gamma(h)$ and $\mathcal{L}_2(\cdot, \cdot)$ are given by (3.8) and (3.9), respectively. Here the subset is $\hat{\mathcal{X}}_{nr,r} := \hat{\mathcal{X}}_{nr,2,r}$.

Then comes the key lemma.

Lemma 3.3 *When $q \rightarrow \infty$, i.e. $\hat{g} = 0$, consider the function $\gamma(h)$ given by (3.8) on $(\hat{\mathcal{X}}_{nr,r}, \|\cdot\|_{s,r})$. If there exists $r_0 > 0$ such that $\|f\|_{s,r_0} < \infty$, then γ and $\hat{\mathbf{g}}$ satisfy (3.1) with $\beta = 0$ and the same s .*

Proof Without loss of generality, we can assume that $\lambda_1 \neq 0$. Let $\mathcal{P}$ be the projection such that $\mathcal{P}x = x_1$ for the vector $x = (x_1, \dots, x_d)$. So Eq. (3.8) can be written as

$$\mathcal{P} \circ \langle S_f(h) \rangle_r = \lambda_1 y_1 \gamma(h).$$

By Lemma 3.2 the map $S_f(h)$ is strongly contracting and so are $\langle S_f(h) \rangle_r$ and $\mathcal{P}\langle S_f(h) \rangle_r$. So taking $h = 0$ and comparing the norm of both sides, we obtain that

$$\|\mathcal{P} \circ \langle f \rangle_r\|_{s,r} = |\lambda_1| r \|\gamma(0)\|_{s,r},$$

which implies $\|\gamma(0)\|_{s,r} = O(r)$ as $r \rightarrow 0$. Similarly, we also have that

$$\|\gamma(h) - \gamma(h')\|_{s,r} \leq |\lambda_1| r^{-1} \|\mathcal{P}\langle S_f(h) - S_f(h') \rangle_r\|_{s,r} = O(1) \|h - h'\|_{s,r}$$

as $r \rightarrow 0$. That is to say, γ regarded as the function of h satisfies (3.1) with $\beta = 0$.

Similarly as shown in the proof of Lemma 3.1, from the above we have that

$$\|\gamma(h)\|_{s,r} = \|\gamma(h) - \gamma(0)\|_{s,r} + \|\gamma(0)\|_{s,r} \leq O(1) \|h\|_{s,r} + O(r) = O(r)$$

for $\|h\|_{s,r} \leq r$ as $r \rightarrow 0$. So there exists $r_1 > 0$ such that $\|\gamma(h)\|_{s,r} \leq 1/2$ for $\|h\|_{s,r} \leq r$ and $r \leq r_1$. By simple computations, we obtain that

$$\|\hat{g}(0)\|_{s,r} \leq \frac{\|\gamma(0)\|_{s,r}}{1 - \|\gamma(0)\|_{s,r}} \leq 2\|\gamma(0)\|_{s,r} = O(r)$$

and

$$\begin{aligned} \|\hat{g}(h) - \hat{g}(h')\|_{s,r} &\leq \frac{\|\gamma(h) - \gamma(h')\|_{s,r}}{(1 - \|\gamma(h)\|_{s,r})(1 - \|\gamma(h')\|_{s,r})} \\ &\leq 4\|\gamma(h) - \gamma(h')\|_{s,r} = O(1) \|h - h'\|_{s,r} \end{aligned}$$

for $\|h\|_{s,r}$ and $\|h'\|_{s,r} \leq r$ as $r \rightarrow 0$. This completes the proof. $\square$

Therefore, we rewrite Theorem 1.3 into a precise form and prove it.

Theorem 3.1 Assume that in (1.1) there exists $r_0 > 0$ such that $\|f\|_{s,r_0} < \infty$, A is in the diagonal form, $\text{Ord}(\hat{g}) = q \geq 2$ in (1.7), and (1.3) of condition (1.2) is satisfied, where $\|\cdot\|_{s,r}$ is the formal Gevrey- s norm as mentioned above. Then following statements hold.

- (i) When $2 \leq q < \infty$, there exists a formal Gevrey- $\hat{s}$ coordinates substitution turning system (1.1) into its normal forms, where $\hat{s} = \max\{s, \frac{1}{q-1}\}$,
- (ii) When $q \rightarrow \infty$, i.e. $\hat{g} = 0$, the above result is valid for $\hat{s} = s$.

Proof Since the linear operator Ad_A has a bounded inverse Ad_A^{-1} by condition (1.3), the key is to verify the strongly contracting of the right sides of Eqs. (3.10) and (3.15).

When $2 \leq q < \infty$, without loss of generality, we can assume that the $(q-1)$ -th order truncated system of (1.1) is in its normal forms by Poincaré-Dulac formal reductions. So we choose the complete subset $(\hat{\mathcal{X}}_{nr,q,r}, \|\cdot\|_{\hat{s},r})$. At first, we control the norms of ρ and $\hat{\rho}$. Similarly as shown in the proof of Lemma 3.3, we can assume that $\lambda_1 \neq 0$ and $\mathcal{P}$ is the projection. Thus from (3.14), it yields that

$$\mathcal{P}(\langle f \rangle_r, -) = \lambda_1 y_1 \gamma(y),$$

which implies

$$\|\gamma\|_{s,r} = \frac{1}{|\lambda_1| r} \|\mathcal{P}(\langle f \rangle_r, -)\|_{s,r} = \frac{1}{|\lambda_1| r} \|\langle f \rangle_r, -\|_{s,r} \leq \frac{1}{|\lambda_1| r} \|f\|_{s,r}$$

$$\begin{aligned}
&= \frac{1}{|\lambda_1|} \sum_{|k| \geq 2, j} |f_{k,j}| e^{-s|k| \ln |k|} r^{|k|-1} \\
&\leq \frac{r}{|\lambda_1| r_0^2} \sum_{|k| \geq 2, j} |f_{k,j}| e^{-s|k| \ln |k|} r_0^{|k|} = c_5 r
\end{aligned}$$

for $r < r_0$ and any possible s , where $c_5 = \|f\|_{s,r_0}/(|\lambda_1| r_0^2)$. Then taking $r \leq 1/(2c_5)$, the above leads to

$$\|\mathcal{V}\|_{s,r} \leq \frac{1}{2}, \quad \|\rho\|_{s,r} \leq \frac{1}{1 - \|\mathcal{V}\|_{s,r}} \leq 2, \quad \|\hat{\rho}\|_{s,r} \leq 2\|\mathcal{V}\|_{s,r} \leq 2c_5 r.$$

At last, by Lemma 3.2 the map $S_f(h)$ given by (3.5) is strongly contracting, and so is $\langle S_f(h) \rangle_{nr}$. Thus $\mathcal{T}_1^1$ given by (3.11) is strongly contracting because that ρ is bounded for $r \leq 1/(2c_5)$. Similarly, $\mathcal{T}_3^1$ given by (3.13) is also strongly contracting by Lemma 3.2(ii) for $s \geq 1/(q-1)$. Moreover, notice that $\mathcal{T}_2^1$ by (3.12) is the linear bounded form satisfying

$$\|\mathcal{T}_2^1(h) - \mathcal{T}_2^1(h')\|_{s,r} \leq \|\hat{\rho}\|_{s,r} \|A(h - h')\|_{s,r} \leq 2c_5 c_6 r \|h - h'\|_{s,r},$$

where $c_6 = \sum |\lambda_i|$. So it is also a strong contracting one, i.e. Eq. (3.10) has the unique solution for the sufficiently small r via Contraction Mapping Principle, which confirms (i).

When $q \rightarrow \infty$, i.e. $\hat{g} = 0$, the subset is $(\hat{\mathcal{X}}_{nr,r}, \|\cdot\|_{s,r})$ and the bilinear form $\mathcal{L}_2(\cdot, \cdot)$ satisfies (3.9), which yields $\alpha = 0$ in Lemma 3.1. Now we do calculations of $\mathcal{T}_1^2$ and $\mathcal{T}_2^2$ one by one. Note that we have known that $\langle S_f(\cdot) \rangle_{nr}$ is strongly contracting. On the one hand, from another form of (3.16)

$$\mathcal{T}_1^2(h) = \langle S_f(h) \rangle_{nr} - \mathcal{L}_2(\hat{g}(h), \langle S_f(h) \rangle_{nr})$$

and Lemmas 3.3 and 3.1, the map $\mathcal{L}_2(\hat{g}(\cdot), \langle S_f(\cdot) \rangle_{nr})$ is strongly contracting and so is $\mathcal{T}_1^2$. On the other hand, $F_1 = \hat{g}$ and $F_2 = Ah$ satisfy condition (3.1) with $\beta = 0$. So the map $\mathcal{T}_2^2$ by (3.17) is also strongly contracting by Lemma 3.3 again. Similarly, Eq. (3.15) has a unique solution, which completes the proof of (ii). $\square$

4 KAM steps in the ultradifferentiable normalization

In this part, the proof of Theorem 1.2 will be shown. More precisely, we begin with solving the homological equation, which yields the control of the one-step transformation and the convergence of the final one by KAM methods. Here we follow the scheme shown in [4, pp. 70–72] for our case.

Now we rewrite system (1.1) in this form

$$\frac{dx}{dt} = Ax + f_r(x) + f_{nr}(x), \quad (4.1)$$

where $f_{nr} = \langle f \rangle_{nr}$ and $f_r = f - f_{nr}$. Doing the coordinates substitution $x = y + h(y)$, which satisfies $\langle h \rangle_{nr} = h$, to the above, it yields that

$$\frac{dy}{dt} = Ay + f_r(y) + [Ay + f_r(y), h] + f_{nr} + R_1 + R_2 + R_3 + R_4, \quad (4.2)$$

where

$$\begin{aligned}
R_1 &= f_r(y + h(y)) - f_r(y) - \partial f_r(y)h(y), \\
R_2 &= f_{nr}(y + h(y)) - f_{nr}(y),
\end{aligned}$$

and

$$\begin{aligned} R_3 &= -\partial h f_{nr} - \partial h \partial f_r h - \partial h A h - \partial h (R_1 + R_2), \\ R_4 &= ((I + \partial h)^{-1} - (I - \partial h))(A y + f_r(y) + A h + f_{nr} + \partial f_r h + R_1 + R_2). \end{aligned}$$

As usual, $[\cdot, \cdot]$ is the classical Lie bracket.

Define $ad_F h := [F, h]$ and simply mark $ad_{A x \cdot}$ as $ad_{A \cdot}$. Under the following conditions of $\omega_s(t)$ with $s = (s_1, s_2)$

$$(C1) \quad \omega_s(t) = -s_1(t^{\tau+1} - 1) - s_2 t \ln t, \quad s_1 \geq \frac{\mu}{(\tau+1)(q-1)}, \text{ and } s_2 \geq \frac{1}{q-1} \text{ for } q = \text{Ord}(f) = \text{Ord}(g) \geq 2,$$

$$(C2) \quad \omega_s(t) = -s_1(t^{\tau+1} - 1) - s_2(t^\eta - 1), \quad s_1 \geq \frac{\mu}{(\tau+1)(q-1)}, \text{ and } s_2 \geq s_{2,0} > 0 \text{ for } q = \text{Ord}(f) = \text{Ord}(g) \geq 2 \text{ and the preassigned positive constant } s_{2,0},$$

KAM steps can be applied.

Lemma 4.1 *The following statements hold.*

(i) *Assume that $\eta = 1$, $\delta \leq 1$, and $4ec^{-1}c_4r^{-1}\|F\|_{E_s,r} \leq 1$. By condition (C1) we have that*

$$\|ad_F^{-1}G\|_{E_s, re^{-\delta}} \leq c_5 e^{\delta^{-\frac{\tau}{\eta-\tau}}} \|G\|_{E_s,r},$$

(ii) *Assume that $\eta > 1$ and $4ec^{-1}c_4r^{-1}\|F\|_{E_s,r} \leq 1$. By condition (C2) we have that*

$$\|ad_F^{-1}G\|_{E_{(s_1, s_2+\delta)}, r} \leq c_5 e^{\delta^{-\frac{\tau}{\eta-\tau}}} \|G\|_{E_s,r},$$

where $c_5 = 2c^{-1}c_3$.

Proof Set $F^{(u)} = \sum_{|k|=u, j} F_{k,j} x^k e_j$. Then we have that

$$F = Ax + \sum_{t \geq q} F^{(t)}, \quad G = \sum_{t \geq q} G^{(t)}, \quad H = \sum_{t \geq q} H^{(t)},$$

which yields

$$G = ad_F H = ad_A H + \sum_{u \geq 2q-1} \sum_{s+t=u+1} ad_{F^{(s)}} H^{(t)}$$

for $s \geq q$ and $t \geq q$. So, we can obtain that

$$H^{(u)} = ad_A^{-1} \left(G^{(u)} - \sum_{s+t=u+1} ad_{F^{(s)}} H^{(t)} \right),$$

which implies the estimate

$$\begin{aligned} \|H^{(u)}\|_{E_s, re^{-\delta}} &\leq \|ad_A^{-1}G^{(u)}\|_{E_s, re^{-\delta}} + \sum_{s+t=u+1} \|ad_A^{-1}(ad_{F^{(s)}}H^{(t)})\|_{E_s, re^{-\delta}} \\ &\leq \|ad_A^{-1}G^{(u)}\|_{E_s, re^{-\delta}} + \sum_{s+t=u+1} 2c^{-1}c_4r^{-1}e^\delta \|F^{(s)}\|_{E_s, re^{-\delta}} \|H^{(t)}\|_{E_s, re^{-\delta}} \\ &\leq c^{-1}c_3e^{\delta^{-\frac{\tau}{1-\tau}}} \|G\|_{E_s,r} + 2ec^{-1}c_4r^{-1}\|F^{(s)}\|_{E_s,r} \|H^{(t)}\|_{E_s, re^{-\delta}} \end{aligned}$$

for $\eta = 1$ and $\delta \leq 1$, and

$$\|H^{(u)}\|_{E_{(s_1, s_2+\delta)}, r} \leq \|ad_A^{-1}G^{(u)}\|_{E_{(s_1, s_2+\delta)}, r} + \sum_{s+t=u+1} \|ad_A^{-1}(ad_{F^{(s)}}H^{(t)})\|_{E_{(s_1, s_2+\delta)}, r}$$

$$\begin{aligned}
&\leq \|ad_A^{-1}G^{(u)}\|_{E_s, re^{-\delta}} + \sum_{s+t=u+1} 2c^{-1}c_4r^{-1}\|F^{(s)}\|_{E_{(s_1, s_2+\delta)}, r}\|H^{(t)}\|_{E_{(s_1, s_2+\delta)}, r} \\
&\leq c^{-1}c_3e^{\delta-\frac{\tau}{\eta-\tau}}\|G\|_{E_s, r} + 2c^{-1}c_4r^{-1}\sum_{s+t=u+1}\|F^{(s)}\|_{E_s, r}\|H^{(t)}\|_{E_{(s_1, s_2+\delta)}, r}
\end{aligned}$$

for $\eta > 1$ by Lemma 2.7.

Denote by $\|\cdot\|_* = \|\cdot\|_{E_s, re^{-\delta}}$ and $\|\cdot\|_* = \|\cdot\|_{E_{(s_1, s_2+\delta)}, r}$, respectively. Choosing a large N and from the above, we have that

$$\begin{aligned}
\sum_{u=q}^N \|H^{(u)}\|_* &\leq c^{-1}c_3e^{\delta-\frac{\tau}{\eta-\tau}}\|G\|_{E_s, r} + 2ec^{-1}c_4r^{-1}\sum_{s=q}^N\|F^{(s)}\|_{E_s, r}\sum_{t=q}^N\|H^{(t)}\|_* \\
&\leq c^{-1}c_3e^{\delta-\frac{\tau}{\eta-\tau}}\|G\|_{E_s, r} + 2ec^{-1}c_4r^{-1}\|F\|_{E_s, r}\sum_{t=q}^N\|H^{(t)}\|_*,
\end{aligned}$$

which implies

$$\sum_{u=q}^N \|H^{(u)}\|_* \leq 2c^{-1}c_3e^{\delta-\frac{\tau}{\eta-\tau}}\|G\|_{E_s, r}.$$

Making $N \rightarrow \infty$ and setting $c_5 = 2c^{-1}c_3$, we complete the proof. $\square$

Denote by $\|\cdot\|_* = \|\cdot\|_{E_s, re^{-\delta}}$ and $\|\cdot\|_* = \|\cdot\|_{E_{(s_1, s_2+\delta)}, r}$, $\|\cdot\|_{**} = \|\cdot\|_{E_s, re^{-2\delta}}$ and $\|\cdot\|_{**} = \|\cdot\|_{E_{(s_1, s_2+2\delta)}, r}$ for (C1) and (C2), respectively. The following is the one-step cancellation of the KAM scheme, which is the key of this part. As mentioned before, doing $x = y + h(y)$ to (4.1), it turns to (4.2), where h is the solution of the following homological equation

$$[Ay + f_r(y), h] = -f_{nr}. \quad (4.3)$$

So we can rewrite system (4.1) into

$$\frac{dx}{dt} = Ax + f_r + f_r^+ + f_{nr}^+, \quad (4.4)$$

where $f^+ = \sum_{i=1}^4 R_i$, $f_r^+ = \langle f^+ \rangle_r$, and $f_{nr}^+ = f^+ - f_r^+$.

Lemma 4.2 Assume that $r \leq 1$, $\delta \leq \min\{\frac{r}{4}, \frac{1}{2ec_4}\}$,

$$\|f_r\|_{E_s, r} \leq \frac{cr}{4ec_4}, \quad (4.5)$$

and

$$\|f_{nr}\|_{E_s, r} \leq \frac{r\delta}{c_5e^{\delta-\frac{\tau}{\eta-\tau}}}. \quad (4.6)$$

Then there exist positive constants K , A , and B such that

$$\|f^+\|_{**} \leq Kr^{-1}\delta^{-2}e^{A\delta^{-B}}\|f_{nr}\|_{E_s, r}^2$$

under (C1) and (C2) in (4.4), respectively. Here $A = 2$ and $B = \tau/(\eta - \tau)$.

Proof First, we give formulas to control norms of R_1 and R_2 by Lemma 2.7. Since $q = \text{Ord}(f) = \text{Ord}(h) \geq 2$, it yields that

$$\|R_2\|_{E_{s,r}} = \left\| \int_0^1 \partial f_{nr}(y + th(y))h(y)dt \right\|_{E_{s,r}} \leq c_4 r^{-1} \|f_{nr}\|_{E_{s,\rho}} \|h\|_{E_{s,r}} \quad (4.7)$$

and

$$\begin{aligned} \|R_1\|_{E_{s,r}} &= \left\| \int_0^1 \int_0^t h(y) \cdot \partial^2 f_r(y + uh(y))h(y)dudt \right\|_{E_{s,r}} \\ &\leq c_4^2 r^{-2} \|f_r\|_{E_{s,\rho}} \|h\|_{E_{s,r}}^2, \end{aligned} \quad (4.8)$$

where $\rho = r + \|h\|_{E_{s,r}}$. Here $\partial \cdot$ and $\partial^2 \cdot$ denote the Jacobian and Hessian matrix, respectively. Notice that by Lemma 4.1 and condition (4.5), we have

$$\|h\|_* \leq c_5 e^{\delta - \frac{\tau}{\eta - \tau}} \|f_{nr}\|_{E_{s,r}} \quad (4.9)$$

and

$$\|h\|_* \leq r\delta \quad (4.10)$$

by condition (4.6), where $\|\cdot\|_* = \|\cdot\|_{E_{s, re^{-\delta}}}$ and $\|\cdot\|_* = \|\cdot\|_{E_{(s_1, s_2 + \delta), r}}$ for (C1) and (C2), respectively. Moreover, regarding ∂h as an operator on $\mathcal{X}$, it yields that

$$\|\partial h\|_{**} \leq c_4 r^{-1} e^{2\delta} \|h\|_{**} \leq c_4 r^{-1} e^{2\delta} \|h\|_* \leq c_4 e\delta \leq \frac{1}{2} \quad (4.11)$$

for $\text{Ord}(g) \geq q$ by making $\mu = 0$ in (2.10) of Lemma 2.7(b) and

$$\|\partial h\|_\diamond \leq e^{-1} \delta^{-1} r^{-1} e^\delta \|h\|_* \leq e^{-\frac{3}{4}} < 1 \quad (4.12)$$

for any g from Lemma 2.6. Here $\|\partial h\|_\diamond = \|\partial h\|_{E_{s, re^{-2\delta}}}$ and $\|\partial h\|_\diamond = \|\partial h\|_{E_{(s_1, s_2 + \delta), re^{-\delta}}}$ for (C1) and (C2), respectively.

Then we can control norms of R_i for $i = 1, \dots, 4$ one by one. For R_1 and R_2 , from inequality (4.9), we notice that

$$\rho = r + \|h\|_{**} \leq r + \|h\|_* \leq r(1 + \delta) \leq re^\delta.$$

This shall lead to

$$\begin{aligned} \|R_1\|_{**} &\leq \|R_1\|_\diamond \leq c_4 r^{-2} e^{4\delta} \|f_r\|_{E_{s,r}} \|h\|_{**}^2 \\ &\leq c_4 r^{-2} e^{4\delta} \|f_r\|_{E_{s,r}} \|h\|_*^2 \leq c c_5^2 r^{-1} e^{2\delta - \frac{\tau}{\eta - \tau}} \|f_{nr}\|_{E_{s,r}}^2 \end{aligned}$$

by inequality (4.8), (4.5), and (4.9) and

$$\begin{aligned} \|R_1\|_{**} &\leq \|R_1\|_\diamond \leq c_4 r^{-1} \|f_{nr}\|_{E_{s,r}} \|h\|_{**} \\ &\leq c_4 r^{-1} \|f_{nr}\|_{E_{s,r}} \|h\|_* \leq c_4 c_5 r^{-1} e^{\delta - \frac{\tau}{\eta - \tau}} \|f_{nr}\|_{E_{s,r}}^2 \end{aligned}$$

by inequality (4.7) and (4.9).

Next comes R_3 . On the one hand, from (4.11), it yields that

$$\|\partial h(R_1 + R_2)\|_{**} \leq \|\partial h\|_{**} (\|R_1\|_{**} + \|R_2\|_{**}) \leq \frac{1}{2} (\|R_1\|_{**} + \|R_2\|_{**}).$$

On the other hand, we have that

$$\begin{aligned} & \|\partial h f_{nr}\|_* + \|\partial h \partial f_r h\|_* + \|\partial h A h\|_* \\ & \leq c_4(r^{-1}e^\delta \|h\|_* \|f_{nr}\|_{E_{s,r}} + r^{-2}e^{2\delta} \|f_r\|_{E_{s,r}} \|h\|_*^2 + M r^{-1}e^\delta \|h\|_*^2) \\ & \leq c_5(c_4e + cc_5 + Mec_4c_5)r^{-1}e^{2\delta - \frac{\tau}{\eta-\tau}} \|f_{nr}\|_{E_{s,r}}^2, \end{aligned}$$

by (4.5) and (4.9), where $M = \sum_i |\lambda_i|$. Naturally, we obtain that

$$\|R_3\|_{**} \leq c_5(2cc_5 + 2ec_4 + Mec_4c_5)r^{-1}e^{2\delta - \frac{\tau}{\eta-\tau}} \|f_{nr}\|_{E_{s,r}}^2.$$

At last, we deal with R_4 . Note that by (4.11) and (4.12) the Newman series can be applied to get

$$(I + \partial h)^{-1} - (I - \partial h) = \sum_{t \geq 2} (-1)^t (\partial h)^i$$

for both $\|\cdot\|_{**}$ and $\|\cdot\|_\diamond$. The first part can be controlled as follows

$$\begin{aligned} & \|((I + \partial h)^{-1} - (I - \partial h))(R_1 + R_2)\|_{**} \leq \left(\sum_{t \geq 2} \|\partial h\|_{**}^i \right) \|R_1 + R_2\|_{**} \\ & \leq \frac{\frac{1}{4}}{1 - \frac{1}{2}} (\|R_1\|_{**} + \|R_2\|_{**}) = \frac{1}{2} (\|R_1\|_{**} + \|R_2\|_{**}) \end{aligned}$$

by (4.11). The second is to use

$$\begin{aligned} & \|((I + \partial h)^{-1} - (I - \partial h))(Ah + f_{nr} + \partial f_r h)\|_{**} \\ & \leq \left(\sum_{t \geq 2} \|\partial h\|_{**}^{i-1} \right) \|\partial h(Ah + f_{nr} + \partial f_r h)\|_{**} \\ & \leq \|\partial h(Ah + f_{nr} + \partial f_r h)\|_* \\ & \leq c_5(c_4e + cc_5 + Mec_4c_5)r^{-1}e^{2\delta - \frac{\tau}{\eta-\tau}} \|f_{nr}\|_{E_{s,r}}^2, \end{aligned}$$

similarly as in the norm estimation of the other hand part of R_3 . The third is

$$\begin{aligned} & \|((I + \partial h)^{-1} - (I - \partial h))f_r(y)\|_{**} \leq \left(\sum_{t \geq 2} \|\partial h\|_{**}^{i-2} \right) \|(\partial h)^2 f_r\|_{**} \\ & \leq 2\|(\partial h)^2 f_r\|_* \leq 2c_4^2 r^{-2} e^{2\delta} \|f_r\|_{E_{s,r}} \|h\|_*^2 \\ & \leq 2cc_5^2 r^{-1} e^{2\delta - \frac{\tau}{\eta-\tau}} \|f_{nr}\|_{E_{s,r}}^2. \end{aligned}$$

While the last one is different, we shall apply (4.12) to get

$$\begin{aligned} & \|((I + \partial h)^{-1} - (I - \partial h))Ay\|_{**} \leq \|((I + \partial h)^{-1} - (I - \partial h))Ay\|_\diamond \\ & \leq \left(\sum_{t \geq 2} \|\partial h\|_\diamond^{i-2} \right) \|(\partial h)^2 Ay\|_\diamond \\ & \leq \frac{1}{1 - e^{-\frac{3}{4}}} e^{-2\delta} \delta^{-2} r^{-2} e^{2\delta} \|h\|_*^2 M r e^\delta \end{aligned}$$

$$\leq c_5^2 \hat{M} r^{-1} \delta^{-2} e^{2\delta^{-\frac{\tau}{\eta-\tau}}} \|f_{nr}\|_{E_{s,r}}^2,$$

where $\hat{M} = M e^{-1}/(1 - e^{-\frac{3}{4}})$. So it leads to

$$\|R_4\|_{**} \leq c_5(4cc_5 + \hat{M}c_5 + 2c_4 + Mec_4)r^{-1}\delta^{-2}e^{2\delta^{-\frac{\tau}{\eta-\tau}}} \|f_{nr}\|_{E_{s,r}}^2.$$

Therefore, summarizing all arguments together, we obtain that

$$\|f^+\|_{**} \leq K r^{-1} \delta^{-2} e^{2\delta^{-\frac{\tau}{\eta-\tau}}} \|f_{nr}\|_{E_{s,r}}^2$$

for the positive constant K . This completes the proof. $\square$

Now we can prove Theorem 1.2.

Proof of Theorem 1.2 Notice that we can set that $\text{Ord}(f_{nr}) \geq q$ by doing at most $q - 1$ times classical normal form reductions degree by degree. Set $\|f\|_{E_{s,\rho}} < \infty$, it yields that

$$\|f\|_{E_{s,r}} = \sum_{|k| \geq q, j} |f_{k,j}| E(|k|) r^{|k|} \leq r^q \rho^{-q} \|f\|_{E_{s,r}}$$

for any $r < \rho$. So we make $\|f\|_{E_{s,r}} = \epsilon_0 r$ with ϵ_0 sufficiently small and $r \leq 1$. Now choose $\delta_n = \delta_0 n^{-2}$. Here δ_0 is a small positive parameter determined below for different cases. Take $r_n = r_{n-1} e^{-\delta_{n-1}}$ with $r_0 = r$ under condition (C1). By inductions, we can assume that $f^{(0)} = f$. Then in the n -th step, it begins with $f^{(n-1)}(x)$ instead of $f(x)$ in system (4.1), goes on with solving homological equation (4.3) by $f_r = f_r^{(n-1)}$, $h = \hat{h}_n$, and $f_{nr} = f_{nr}^{(n-1)}$ in the norm $\|\cdot\|_{E_{s,r_n} e^{-\delta_n}}$ under condition (C1) and $\|\cdot\|_{E_{(s_1, s_2 + \delta_n), r}}$ under condition (C2), respectively, and ends in system (4.4) for $f^+ = f^{(n)}$.

Under condition (C1) we have that $r_{n+1} = r_n e^{-\delta_n} = r e^{-2\delta}$ as in Lemma 4.2. Set $\hat{K} = 4K$ and $\hat{A} = A 2^B$. Take $r < 1$ and fix

$$\delta_0 < \min \left\{ \frac{3r}{2\pi^2}, \frac{3}{ec4\pi^2} \right\}. \quad (4.13)$$

Notice that

$$\delta_n^{-2} \delta_{n-1}^{-2} \cdots \delta_1^{-2n} = (\delta_0^{-2})^{1+2+\cdots+2^{n-1}} n^{2^2} (n-1)^{2^3} \cdots 1^{2^{n+1}} \leq (\delta_0^{-2})^{2^n} e^{2^n \alpha}$$

and

$$e^{\hat{A}(\delta_n^{-B} + 2\delta_{n-1}^{-B} + \cdots + 2^{n-1}\delta_1^{-B})} = e^{\hat{A}\delta_0^{-2B}(n^{2B} + 2(n-1)^{2B} + \cdots + 2^{n-1}1^{2B})} \leq e^{\hat{A}\delta_0^{-2B} 2^n \beta},$$

where $\alpha = \sum_{t=1}^{\infty} 2^{2-t} \ln t$ and $\beta = \sum_{t=1}^{\infty} 2^{-t} t^{2B}$. It yields that

$$\begin{aligned} \|f^{(n)}\|_{E_{s,r_{n+1}}} &= \|f^{(n)}\|_{E_{s,r_n} e^{-\delta_n}} \leq \hat{K} r^{-1} \delta_n^{-2} e^{\hat{A}\delta_n^{-B}} \|f_{nr}^{(n-1)}\|_{E_{s,r_n}}^2 \\ &= \hat{K} r^{-1} \delta_n^{-2} e^{\hat{A}\delta_n^{-B}} \|f_{nr}^{(n-1)}\|_{E_{s,r_{n-1}} e^{-\delta_{n-1}}}^2 \\ &\leq (\hat{K} r^{-1})^{1+2} \delta_n^{-2} \delta_{n-1}^{-4} e^{\hat{A}(\delta_n^{-B} + 2\delta_{n-1}^{-B})} \|f_{nr}^{(n-2)}\|_{E_{s,r_{n-1}}}^4 \\ &\leq \cdots \\ &\leq (\hat{K} r^{-1})^{\sum_{t=0}^{n-1} 2^t} \left(\prod_{t=1}^n \delta_t^{-2^{n+1-t}} \right) e^{\hat{A}(\sum_{t=1}^n 2^{n-t} \delta_t^{-B})} \|f^{(0)}\|_{E_{s,r_1}}^{2^n} \\ &\leq (\hat{K} \delta_0^{-2} e^{\alpha + \hat{A}\delta_0^{-2B} \beta \epsilon_0})^{2^n} r. \end{aligned}$$

Thus, inequality (4.6) and (4.5) lead to

$$(\hat{K}\delta_0^{-2}e^{\alpha+\hat{A}\delta_0^{-2B}\beta}\epsilon_0)^{2^t} \leq \frac{\delta_0(t+1)^{-2}}{c_5e^{\delta_0^{-B}(t+1)^{2B}}}, \quad \forall t$$

and

$$\epsilon_0 + \sum_{t \geq 1} (\hat{K}\delta_0^{-2}e^{\alpha+\hat{A}\delta_0^{-2B}\beta}\epsilon_0)^{2^t} \leq \frac{c}{4ec_4}.$$

Making $Q = \hat{K}\delta_0^{-2}e^{\alpha+\hat{A}\delta_0^{-2B}\beta}$ and taking

$$Z_0 = \inf_{t \geq 0} \left(\frac{\delta_0(t+1)^{-2}}{c_5e^{\delta_0^{-B}(t+1)^{2B}}} \right)^{\frac{1}{2^t}} > 0,$$

we know that

$$\epsilon_0 \leq \min \left\{ \frac{1}{2Q}, \frac{Z_0}{Q}, \frac{c}{(2Q+1)4ec_4} \right\}$$

is enough.

Finally, notice that the constant $\gamma_* = e^{-\delta_0\pi^2/6} < 1$. Set $h_n = \text{id} + \hat{h}_n$ with $h_1 = \text{id}$ and we have that $h^{(n)} = h_n \circ h_{n-1} \circ \cdots \circ h_1$, which implies $h^{(n)} - h^{(n-1)} = \hat{h}_n \circ h_{n-1} \circ \cdots \circ h_1$. We shall choose proper $\hat{r}$ such that $\|h^{(n)}\|_{E_s, \hat{r}} \leq r\gamma_*$ for any n . Making $\hat{r} = r\gamma_*/(2d)$, we have that

$$\|h^{(1)}\|_{E_s, \hat{r}} = \|h_1\|_{E_s, \hat{r}} = \|\text{id}\|_{E_s, \hat{r}} = d\hat{r} \leq \frac{r\gamma_*}{2}.$$

Since $\hat{r} \leq r\gamma_* \leq r_n$ for any n , we have that

$$\|\hat{h}_n\|_{E_s, \hat{r}} \leq \|\hat{h}_n\|_{E_s, r\gamma_*} \leq \|\hat{h}_n\|_{E_s, r_{n-1}}e^{-\gamma_*\delta_{n-1}} \leq r_{n-1}\gamma_*\delta_{n-1} \leq \frac{r\gamma_*\delta_0}{(n-1)^2} \leq \frac{3r\gamma_*}{2\pi^2(n-1)^2}$$

by (4.10) and (4.13) for $\hat{r} \leq r_n = r_{n-1}e^{-\delta_{n-1}} \leq r_{n-1}e^{-\gamma_*\delta_{n-1}} \leq r_{n-1}$. Now assume that $\|h^{(t)}\|_{E_s, \hat{r}} \leq r\gamma_*$ for all $t < n$. Then it yields

$$\begin{aligned} \|h^{(n)}\|_{E_s, \hat{r}} &= \|\hat{h}_n \circ h^{(n-1)} + h^{(n-1)}\|_{E_s, \hat{r}} \leq \|\hat{h}_n \circ h^{(n-1)}\|_{E_s, \hat{r}} + \|h^{(n-1)}\|_{E_s, \hat{r}} \\ &\leq \|\hat{h}_n\|_{E_s, r\gamma_*} + \|h^{(n-1)}\|_{E_s, \hat{r}} \leq \frac{3r\gamma_*}{2\pi^2(n-1)^2} + \|h^{(n-1)}\|_{E_s, \hat{r}} \\ &\leq \cdots \leq \sum_{t=2}^{n-1} \frac{3r\gamma_*}{2\pi^2(t-1)^2} + \frac{r\gamma_*}{2} \leq \frac{3r\gamma_*}{4} \leq r\gamma_* \end{aligned}$$

for any n , which of course implies the inequality

$$\|h^{(n)} - h^{(n-1)}\|_{E_s, \hat{r}} \leq \frac{3r\gamma_*}{2\pi^2(n-1)^2}$$

and the convergence of the sequence $\{h^{(n)}\}$ in $\|\cdot\|_{E_s, \hat{r}}$.

Under condition (C2) the arguments are similar except for several points. First, arbitrarily choose $s_{2,0} = s_2/2 > 0$, take $r < 1$, and fix

$$\delta_0 < \min \left\{ \frac{3r}{2\pi^2}, \frac{3}{ec_4\pi^2}, \frac{3s_2}{\pi^2} \right\}$$

instead of (4.13). Then the norm $\|\cdot\|_{E_s, r_{n+1}}$ changes into $\|\cdot\|_{E_{(s_1, s_2, 0) + \delta_n}, r}$ by the same δ_n . At last, the estimation of ϵ_0 is the same, while $\gamma_* = 1$ and $\hat{r} = r/(2d)$. This completes the proof. $\square$

5 Applications

As shown in [7, 11] and ours, the theory of ultradifferentiable normal forms can make the small divisor ‘visible’ in the smooth category. Moreover, we go further to show that it provides quantitative descriptions of nonlinear terms via conjugations, which bridge the gap between ‘real’ normal forms, the geometry of general leaves, and bifurcations.

We shall indicate that Theorem 1.3 can be applied to characterize the order of planar saddles and foci quantitatively. Consider the real C^∞ vector field

$$\frac{dX}{dz} = V(X), \quad (5.1)$$

where $X = (x, y) \in \mathbb{R}^2$ and $z \in \mathbb{R}$. At the origin O there is a resonant saddle of the following in proper local coordinates

$$\frac{dx}{dz} = px + f_1(x, y), \quad \frac{dy}{dz} = -qy + f_2(x, y), \quad (5.2)$$

where p and $q \in \mathbb{N}$ are co-prime, f_1 and $f_2 = O(|x|^2 + |y|^2)$ as $(x, y) \rightarrow 0$. We call that it is formally non-integrable, associated with its formal normal form

$$\frac{dx}{dz} = px + \sum_l a_l x(x^q y^p)^l, \quad \frac{dy}{dz} = -qy + \sum_l b_l y(x^q y^p)^l,$$

provided that there exists $l \in \mathbb{N}$ such that $b_l p + a_l q \neq 0$. And its order is given by $l_0 = \min\{l \mid b_l p + a_l q \neq 0\}$. Or the origin O is a fine focus in proper local coordinates as follows

$$\frac{dx}{dz} = -y + f_1(x, y), \quad \frac{dy}{dz} = x + f_2(x, y), \quad (5.3)$$

whose formal normal form is

$$\begin{aligned} \frac{dx}{dz} &= -y - y \sum b_m (x^2 + y^2)^m + x \sum a_l (x^2 + y^2)^l, \\ \frac{dy}{dz} &= x + x \sum b_m (x^2 + y^2)^m + y \sum a_l (x^2 + y^2)^l, \end{aligned}$$

where $z \in \mathbb{R}$. Then its order is given by $l_0 = \min\{l \mid a_l \neq 0\}$, provided that the system is formally non-integrable.

Corollary 1 Assume that the origin O is a formally non-integrable resonant saddle of form (5.2) or fine focus of form (5.3) with the order l_0 in system (5.1). Then the following statements hold.

- (i) When the origin O is a resonant saddle, if system (5.1) is Gevrey- s smooth, there exists a Gevrey- $\hat{s}$ smooth coordinates substitution, which turns system (5.1) into its normal form, where $\hat{s} = \max\{s, \frac{l_0(p+q)+1}{l_0(p+q)}\}$.
- (ii) When the origin O is a fine focus, if system (5.1) is of formal Gevrey- s , there exists a formal Gevrey- $\hat{s}$ coordinates substitution, which turns system (5.1) into its normal form, where $\hat{s} = \max\{s, \frac{1}{2l_0}\}$.

Proof System (5.1) is Gevrey- s smooth, which is formal Gevrey- $s - 1$ and admits a normal form transformation of formal Gevrey- s' by Theorem 1.3. Here $s' = \max\{s - 1, 1/l_0(p + q)\}$. Then the corresponding coordinates substitution is Gevrey- $\hat{s}$ smooth for $\hat{s} = s' + 1$ by [11], which confirms (i).

When O is fine focus of (5.3), whose complex form is

$$\frac{du_1}{dz} = \sqrt{-1}u_1 + \hat{f}_1(u_1, u_2), \quad \frac{du_2}{dz} = -\sqrt{-1}u_2 + \hat{f}_2(u_1, u_2), \quad (5.4)$$

for $u_1 = x + \sqrt{-1}y$ and $u_2 = x - \sqrt{-1}y$, the space with the admissible structure is

$$\widetilde{\mathcal{X}}_{nr,q,r} = \{h \in \mathcal{X}_{nr,q} \mid \|h\|_{s,r} \leq r, \quad \overline{h_{(k_1,k_2),1}} = h_{(k_2,k_1),2}\},$$

where

$$\begin{aligned} \hat{f}_1 &= f_1\left(\frac{u_1 + u_2}{2}, \frac{u_1 - u_2}{2\sqrt{-1}}\right) + \sqrt{-1}f_2\left(\frac{u_1 + u_2}{2}, \frac{u_1 - u_2}{2\sqrt{-1}}\right), \\ \hat{f}_2 &= f_1\left(\frac{u_1 + u_2}{2}, \frac{u_1 - u_2}{2\sqrt{-1}}\right) - \sqrt{-1}f_2\left(\frac{u_1 + u_2}{2}, \frac{u_1 - u_2}{2\sqrt{-1}}\right). \end{aligned}$$

and $q = 2l_0 + 1$. By Theorem 1.3 again, we prove statement (i). $\square$

It is known that the order of the resonant saddle and the fine focus decides not only the cyclicity of the bifurcations but also the geometry of the general foliations associated with the non-integrability. Beyond the formal normal forms, now we can detect it quantitatively via conjugations of the ultradifferentiable normalization. Summarizing other results together, it yields the following.

Theorem 5.1 *Assume that the origin O is a formally non-integrable resonant saddle and a fine focus of the order l_0 for system (5.1). Then*

(1) *in the saddle case, the following statements hold.*

- (i) (Homoclinic bifurcation) *Additionally if there is a homoclinic connection at the point O of system (5.1), by the small C^∞ perturbations there are l_0 limit cycles at most bifurcated from this homoclinic orbit.*
- (ii) (Ultradifferentiable normal form) *Additionally if system (5.2) is Gevrey- s ($s \geq 1$) smooth, then there exists a Gevrey- $\hat{s}$ smooth coordinates substitution, which turns system (5.2) into its normal form, where $\hat{s} = \max\{s, \frac{l_0(p+q)+1}{l_0(p+q)}\}$.*
- (iii) (Geometry of the general leaf) *Consider any complex holomorphic approximate system*

$$\frac{d\hat{X}}{dz} = \hat{V}^t(\hat{X}) \quad (5.5)$$

of system (5.2) satisfying $Jet_O^t V = Jet_O^t \hat{V}^t$, where $\hat{X} = (x, y) \in \mathbb{C}^2$, $z \in \mathbb{C}$ and $t \geq l_0(p + q) + 1$. Then there exists $\delta_2 > 0$ such that each complex integral curve $\mathcal{L}$ of (5.5) in $U = \{(x, y) \in \mathbb{C}^2 \mid \delta_1 < \max\{|x|, |y|\} < \delta_2\}$ is immersed into $\mathbb{C}^2$ for any $\delta_1 > 0$ except stable and unstable manifolds, whose box dimension is given by $\dim_B(\mathcal{L} \cap U) = \max\{3 - \frac{1}{l_0q+1}, 3 - \frac{1}{l_0p+1}\}$;

(2) *in the fine focus case, the following statements hold.*

- (i) (Degenerated Hopf bifurcation) *By the small C^∞ perturbations there are $2l_0 + 1$ limit cycles at most bifurcated from the point O of system (5.2).*

- (ii) (*Ultradifferentiable normal form*) Additionally if system (5.1) is of formal Gevrey- s , there exists a formal Gevrey- $\hat{s}$ coordinates substitution, which turns system (5.3) into its normal form, where $\hat{s} = \max\{s, \frac{1}{2\ell_0}\}$.
- (iii) (*Geometry of the leaf*) Then there exists $\delta > 0$ such that each integral curve $\mathcal{L}$ of (5.3) in $U = \{(x, y) \in \mathbb{R}^2 \mid 0 < \sqrt{|x|^2 + |y|^2} < \delta\}$ is immersed into $\mathbb{R}^2$, whose box dimension is given by $\dim_B(\mathcal{L} \cap U) = 2 - \frac{2}{2\ell_0 + 1}$.

Proof 1(i) and 2(i) are from [9] and [5], respectively. Then (ii) of 1 and 2 both are from Corollary 1 straightforwardly, while 3(iii) and 2(iii) are in [14] and [17], respectively. This completes the proof. $\square$

Acknowledgements Hao Wu is partially supported by NSF of China (No. 11871041) and NSF of Jiangsu, China (No. BK20201262). Xingdong Xu is supported by NSF of China (No. 11771077). Dongfeng Zhang is supported by NSF of Jiangsu, China (No. BK20201262).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

Consider the linear operator $ad_A: \partial_x h Ax - Ah$ for $A = D + N$, where $D = \text{diag}(\lambda_1, \dots, \lambda_d)$, N is the nilpotent part, and the maximum block of A is of the size $s \times s$. Denote the space of homogeneous d -dimensional polynomials of degree t in $\mathbb{C}^d$ by

$$H_t^d(\mathbb{C}^d) = \text{span}\{x^k e_j \mid |k| = t, j = 1, \dots, d\}.$$

Then we have that

$$H_{t,nr}^d(\mathbb{C}^d) = \text{span}\{x^k e_j \mid k \cdot \lambda \neq \lambda_j, |k| = t, j = 1, \dots, d\}.$$

By applying l^1 norm in $H_{t,nr}^d$, in this part we explore the connection between the small divisor condition (1.2) and the upper bound of the inverse of the operator ad_A . Here we note again that ad_A is invertible and diagonal in $H_t^d(\mathbb{C}^d) \setminus H_{t,nr}^d(\mathbb{C}^d)$ by the simple computation $ad_A(x^k e_j) = (k \cdot \lambda - \lambda_j)x^k e_j$ and ad_N is given by

$$ad_N(x^k e_j) = \epsilon \sum k_i x^{k-e_i+e_{i+1}} e_{j+1}.$$

Lemma 5.1 *For the given small divisor condition (1.2), we have that*

$$\|ad_A^{-1}\|_o \leq 2c^{-1}\Omega^{t^d+1}(t)t^{t^d} \quad (5.6)$$

for $4\epsilon c^{-1} \leq 1$, where $\|\cdot\|_o$ is the classical operator norm in $(H_{t,nr}^d(\mathbb{C}^d), l^1)$.

Proof Here we identity the operator ad and its matrix representation for the simplicity of notations. First we measure the norm of the ad_N . For any matrix $U = \{u_{i,j}\}$ we have that

$$\|U\|_o = \sup_{\|x\|=1} \|Ux\| = \max_i \sum_j |u_{i,j}|,$$

where $\|x\| = \sum_i |x_i|$ is the classical l^1 norm. So, we obtain that

$$\|ad_N\|_o \leq \epsilon(|k| + 1).$$

Then from $ad_A = ad_D + ad_N$, it yields that

$$ad_A^{-1} = (ad_D + ad_N)^{-1} = ad_D^{-1} - ad_D^{-2}ad_N + \cdots + (-1)^m ad_D^{-(m+1)}ad_N^m$$

for $m \leq \dim H_{t, nr}^d(\mathbb{C}^d) \leq t^d$, because ad_N is also the nilpotent part satisfying

$$ad_N ad_A(x^k e_j) = ad_A ad_N(x^k e_j) = (\lambda \cdot k - \lambda_j) ad_N(x^k e_j).$$

Note that $\|ad_D^{-1}\|_o \leq c^{-1}\Omega(t)$. This implies

$$\begin{aligned} \|ad_A^{-1}\|_o &\leq c^{-1}\Omega(t) + \epsilon c^{-2}\Omega^2(t)(t+1) + \cdots + \epsilon^{t^d} c^{-(t^d+1)}\Omega^{t^d+1}(t)(t+1)^{t^d} \\ &\leq c^{-1}\Omega^{t^d+1}(t)t^{t^d}(1 + \epsilon c^{-1}2 + \cdots (\epsilon c^{-1}2)^{t^d}) \\ &\leq 2c^{-1}\Omega^{t^d+1}(t)t^{t^d} \end{aligned}$$

for $4\epsilon c^{-1} \leq 1$. This completes the proof. $\square$

References

1. Bibikov, Yu.N.: Local theory of nonlinear analytic ordinary differential equations. Lecture Notes in Mathematics, vol. 702. Springer, Berlin (1979)
2. Carletti, T., Marmi, S.: Linearization of analytic and non-analytic germs of diffeomorphisms of $(\mathbb{C}, 0)$. Bull. Soc. Math. Fr. **128**, 59–85 (2000)
3. Chow, S.-N., Li, C., Wang, D.: Normal Forms and Bifurcation of Planar Vector Fields. Cambridge University Press, New York (1994)
4. De La Llave, R.: A tutorial on KAM theory. University Lecture Series, vol. 32. American Mathematical Society, Providence (2008)
5. Farr, W.W., Li, C., Labouriau, I.S., Langford, W.F.: Degenerate Hopf bifurcation formulas and Hilbert's 16th problem. SIAM J. Math. Anal. **20**, 13–30 (1989)
6. Lombardi, E., Stolovitch, L.: Normal forms of analytic perturbations of quasihomogeneous vector fields: rigidity, invariant analytic sets and exponentially small approximation. Ann. Sci. Éc. Norm. Supér. **43**, 659–718 (2010)
7. Pöschel, J.: On the Siegel–Sternberg linearization theorem (2017). [arXiv:1702.03691](https://arxiv.org/abs/1702.03691)
8. Robbins, H.: A remark on Stirling's formula. Am. Math. Mon. **62**(1), 26–29 (1955)
9. Roussarie, R.: On the number of limit cycles which appear by perturbation of separatrix loop of planar vector fields. Bol. Soc. Bras. Math. **17**, 67–101 (1986)
10. Rudin, W.: Division algebras of infinitely differentiable functions. J. Math. Mech. **11**, 797–809 (1962)
11. Stolovitch, L.: Smooth Gevrey normal forms of vector fields near a fixed point. Ann. Inst. Fourier **63**, 241–267 (2013)
12. Taguchi, Y.: Fourier coefficients of periodic functions of Gevrey classes and ultradistributions. Yokohama Math. J. **35**, 51–60 (1987)
13. Wu, H.: Gevrey smooth topology is proper to detect normalization under Siegel type small divisor conditions. Math. Z. **284**, 1223–1243 (2016)
14. Wu, H., Xu, S.: The geometrical demonstration of the order of resonant saddle points in $\mathbb{C}^2$. Qual. Theory Dyn. Syst. **19**, Article number 27 (2020)
15. Yakovenko, S., Ilyashenko, Y.S.: Lectures on analytic differential equation, Graduate Studies in Mathematics, vol. 86. American Mathematical Society, Providence (2008)
16. Zhang, X.: Analytic normalization of analytic integrable systems and the embedding flows. J. Differ. Equ. **244**, 1080–1092 (2008)
17. Žubrinić, D., Žubrinić, V.: Fractal analysis of spiral trajectories of some planar vector fields. Bull. Sci. Math. **129**, 457–485 (2005)

18. Zung, N.T.: Convergence versus integrability in Poincaré–Dulac normal form. *Math. Res. Lett.* **9**, 217–228 (2002)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.