

Gevrey normal form for unfoldings of nilpotent contact points of planar slow fast systems

Peter De Maesschalck ^{*} and Doan Thai Son [†]

Abstract

We present normal forms for unfoldings of regular nilpotent contact points of slow-fast systems in the plane. The normal forms are useful in the treatment of non-generic jump points and non-generic turning points. For non-generic jump points, we prove a normal form of Liénard type, while for non-generic turning points, the normal form is of Liénard type up to exponentially small error. The techniques being used are based on Gevrey estimates on formal power series, summation up to exponentially small error and a Gevrey version of the preparation theorem for right equivalences. This extension of earlier results is based on a Gevrey version of Levinson's preparation theorem, which we prove.

1 Introduction

This paper deals with local normal forms of slow-fast systems in the context of a point of loss of normal hyperbolicity. In the framework of geometric singular perturbation theory, normal form theory usually refers to Fenichel's work [4], which is valid when the critical manifold is normally hyperbolic. Geometrically, this means that slow and fast variables can be meaningfully split up and that the fast variables have hyperbolic singularities. Normal hyperbolicity is lost either when the fast variables have singularities with center directions (for example in the case of slow-fast systems with rapid oscillations, see [7, 9]) or when there is a tangency between the fast dynamics and the critical manifold ('the critical manifold has a fold'). Our interest goes to the second case of loss of normal hyperbolicity, i.e. tangency between

^{*}peter.demaesschalck@uhasselt.be, Hasselt University, Agoralaan, gebouw D, B-3590 Diepenbeek, Belgium

[†]dtson@math.ac.vn, Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, 10307 Ha Noi, Viet Nam

the critical manifold and the fast foliation. At such points, there is no separation between slow and fast variables; our interest goes into locally reducing to a most elementary form, see (1) below for an example.

We focus on slow-fast systems in two variables, experiencing loss of normal hyperbolicity at one point due to the presence of a fold in the critical curve. Generic folds have been dealt with in an earlier paper, here we focus on contacts of order n . Geometrically, higher order contact points are much more difficult to deal with, since they may typically unfold in multiple lower order contact points upon variation of perturbation parameters.

Consider now an (analytic) (ε, λ) -family of local vector fields $X_{\varepsilon, \lambda}$ so that for $(\varepsilon, \lambda) = (0, \lambda_0)$, the system has a simple curve of singular points expressed by the zero set of some analytic function $F = 0$. We assume furthermore that

$$X_{0, \lambda_0} = F \cdot Z_0,$$

where Z_0 is a vector field without singular points (it is the reduced vector field). Normal hyperbolicity can be expressed as $\langle Z_0, \nabla F \rangle \neq 0$ along $\{F = 0\}$, i.e. the nontangency of $\{F = 0\}$ with the orbits of Z_0 . Under the condition that Z_0 has nowhere singular points, X_{0, λ_0} has a nilpotent linearization at any point with loss of normal hyperbolicity, and under these circumstances, the family of vector fields can be brought into the following (pre-normal) form ([3]):

$$\begin{cases} \dot{x} &= y - f(x, \varepsilon, \lambda), \\ \dot{y} &= \varepsilon g(x, y, \varepsilon, \lambda), \end{cases} \quad (1)$$

defined for $(x, y, \varepsilon, \lambda)$ in a neighbourhood of $[-x_0, x_0] \times [-y_0, y_0] \times [0, \varepsilon_0] \times \Lambda$, Λ being a compact subset of parameters inside an euclidean space (and $x_0 > 0$, $y_0 > 0$, $\varepsilon_0 > 0$). For $\lambda = \lambda_0$ the following condition expressed that the origin is a *nilpotent contact point of order n* :

$$\begin{cases} f(0, 0, \lambda_0) = \frac{\partial f}{\partial x}(0, 0, \lambda_0) = \dots = \frac{\partial^{n-1} f}{\partial x^{n-1}}(0, 0, \lambda_0) = 0, \\ \frac{\partial^n f}{\partial x^n}(0, 0, \lambda_0) \neq 0. \end{cases} \quad (2)$$

Our main result is the following theorem, which generalizes the results in [2]:

Theorem 1. *Consider the analytic system (1) and suppose that (2) holds. Then, there exists a local change of coordinates and regular change of time that is analytic in (x, y, λ) and smooth in ε , bringing (1) in the form*

$$\begin{cases} \dot{X} &= Y - X^n - \sum_{k=1}^{n-2} p_k(\lambda, \varepsilon) X^k, \\ \dot{Y} &= \varepsilon [G(X, \lambda, \varepsilon) + R(X, Y, \lambda, \varepsilon)], \end{cases} \quad (3)$$

where $|R(X, Y, \lambda, \varepsilon)| \leq Le^{-\frac{C}{\varepsilon}}$ for some $L, C > 0$. Suppose additionally that $g(0, \lambda_0, 0) \neq 0$. Then for any $N \in \mathbb{N}$ there exists a C^N local change of coordinate and regular change of time, bringing (3) C^N -smoothly in the form

$$\begin{cases} \dot{X} &= Y - X^n - \sum_{k=1}^{n-2} p_k(\lambda, \varepsilon) X^k, \\ \dot{Y} &= \varepsilon G(X, \lambda, \varepsilon). \end{cases} \quad (4)$$

This theorem states that it is possible to write the slow-fast system in Liénard form, up to exponentially small terms. Equation (3) corresponds to singular second-order differential equations

$$\varepsilon X'' + F'(X)X' - G(X, \lambda, \varepsilon) = O(e^{-C/\varepsilon}),$$

after rescaling time, where $F(X) = X^n + \sum_{k=1}^{n-2} p_k(\lambda, \varepsilon) X^k$.

It is an open question to what extent the exponentially small remainder can be removed. We only have a partial result in that direction: when the contact point is of jump type, then it is possible to remove the remainder term in a C^N -smooth way. Two questions hence remain: (1) can one improve the smoothness, (2) can one remove the remainder R for general contact points, not necessarily of jump type? These questions remain unanswered here.

From a formal power series point of view, we conjecture that the normal form in (3) is optimal: let $\hat{G}(X, \lambda, \varepsilon) = \sum_{k=0}^{\infty} G_k(X, \lambda) \varepsilon^k$ be the formal power series expansion of g w.r.t. ε , then we claim that any normal form representation in the same form as (3) of the initial slow-fast system (1) necessarily has the same power series expansion on the \dot{Y} -equation. As a consequence, we can see that not only the order of contact n , but also the formal series \hat{g} would be a formal invariant of the slow-fast system. This is in contrast to the normally hyperbolic case: any pair of planar slow-fast systems, defined near normally hyperbolic point of a slow-fast system (outside a singular point in the slow dynamics) are equivalent to one another (if one allows time reversals to make attracting points equivalent to repelling points), and also to the normal form $\{\dot{X} = -X, \dot{Y} = \varepsilon\}$. In other words, about normally hyperbolic points, slow-fast systems have no invariants. The question whether or not \hat{G} is an invariant near contact points is not resolved in this paper.

The restriction to planar systems reduces the number of obstructions from a formal point of view: as soon as there are three dimensions or more, resonance phenomena may appear either in the slow system, the fast system

or both. We nevertheless envisage future work on normal forms around folded singularities (see [10]) of systems with two slow and one fast variable.

To prove Theorem 1, we strongly use the theory of Gevrey functions, following the ideas put forward in [9] in combination with the use of center manifolds (for the second part of the theorem). This paper is a follow-up paper of [2] where normal forms of contact points of order 2 have been established with the same method. The main difference that we deal in this paper is the degeneracy of the contact point and the possibility that the contact point perturbs w.r.t. parameters λ . As a consequence of this degeneracy, we need an extra element in the construction of the normal form: a Gevrey version of a preparation theorem developed by Levinson [6].

The paper is organized as follows: In Section 2, we first recall the preparation theorem for analytic functions in Section 2.1. As a consequence, we obtain a pre-normal form for (1) in Remark 5. In Section 2.2, we establish a Gevrey version of this preparation theorem (followed by a formal version in Section 2.3). The results in Section 2.2 will be used later to prove Theorem 1 in Section 3.

2 Preparation theorem

Should $f(x, \varepsilon, \lambda) = f_0(x)$ for some reason then, still under assumption (2), it is easy to reduce to the case where $f_0(x) = \pm x^n$. Indeed, in that case $f_0(x) = \pm x^n(c + O(x))$ for some $c > 0$. Writing $X = \varphi(x) := x(c + O(x))^{1/n}$ shows that $f_0(\psi(X)) = \pm X^n$ with $\psi = \varphi^{-1}$. For the slow-fast system (1) it would imply that the coordinate change $X = \varphi(x)$ together with a nonlinear scaling of the time t would reduce the system of equations to a form where the critical curve is given by $y = \pm X^n$ (see the proof of Corollary 4 for details in a more general setting). The crucial step is the fact that we can rewrite $f_0(x)$ as $\pm X^n$ by means of a coordinate change. This is in fact the notion of right equivalence, and is rather trivially exposed in this paragraph but becomes more delicate if we allow for parameters; this is the work of Levinson. There is similarly a notion of left equivalence: we can write $f_0(x) = g(x)x^n$ with $g(0) \neq 0$ and from that point of view say that $f_0(x)$ is left-equivalent to x^n . This trivial observation led to the more general Weierstrass preparation theorem for analytic functions, later generalized to smooth functions by Malgrange-Mather. Quasi-analytic left-equivalence has been studied in [8]. In this paper, we will only be using right equivalence, and while smooth versions most certainly appear in the literature, up to our knowledge there is no Gevrey version of the preparation theorem for right

equivalence.

2.1 Levinson's Preparation theorem

Definition 2. Let $x \in U \subset \mathbb{C}$, U containing the origin and let $\mu \in M \subset \mathbb{C}^s$, M containing a neighbourhood of μ_0 . We say that $f(x, \mu)$ is *analytically right-equivalent* to $\tilde{f}(x, \mu)$ about $(x, \mu) = (0, \mu_0)$ if there exists a locally analytic change of coordinates $x = \psi(X, \mu)$, with $\psi(0, \mu_0) = 0$, such that

$$f(\psi(X, \mu), \mu) = \tilde{f}(X, \mu).$$

Theorem 3 (Levinson [6]). *Let $U \subset \mathbb{C}$ be a neighbourhood of 0 and let $M \subset \mathbb{C}^s$ be a neighbourhood of μ_0 . Let $f: U \times M \rightarrow \mathbb{C}$, $(x, \mu) \mapsto f(x, \mu)$ be an analytic function satisfying the condition*

$$f(0, \mu_0) = \frac{\partial f}{\partial x}(0, \mu_0) = \cdots = \frac{\partial^{n-1} f}{\partial x^{n-1}}(0, \mu_0) = 0, \quad \frac{\partial^n f}{\partial x^n}(0, \mu_0) \neq 0.$$

Then there exists an analytic function $\alpha(x, \mu)$ defined on a (possibly smaller) neighbourhood of $(0, \mu_0)$ such that the coordinate change $x = X + X^2\alpha(X, \mu)$ is a right equivalence between f and a polynomial:

$$f(X + X^2\alpha(X, \mu), \mu) = \sum_{k=0}^n p_k(\mu) X^k,$$

where $p_0(\mu), \dots, p_n(\mu)$ are analytic, vanish at $\mu = \mu_0$ for $k < n$, and does not vanish at $\mu = \mu_0$ for $k = n$.

We apply this theorem to the case of slow-fast systems, using $\mu = (\varepsilon, \lambda)$:

Corollary 4. *Let system (1) be analytic, and assume (2). There exists a real analytic coordinate change and a time rescaling by a strictly positive or strictly negative real analytic function bringing (1) in the form*

$$\begin{cases} \dot{x} &= y - x^n - \sum_{k=1}^{n-2} q_k(\lambda)x^k, \\ \dot{y} &= \varepsilon h(x, y, \varepsilon, \lambda), \end{cases}$$

where h is analytic and for $k = 1, \dots, n-2$ the function q_k is analytic and vanishes at $\lambda = \lambda_0$.

Proof. Let $x = \psi(X, \varepsilon, \lambda)$ be the coordinate change suggested by Theorem 3. By possibly changing ψ to $-\psi$, we may assume without loss of generality that $\psi_X > 0$. Clearly, we have

$$\psi_X \dot{X} = y - f(\psi(X, \varepsilon, \lambda), \varepsilon, \lambda).$$

Since $\psi_X > 0$ a time rescaling brings the family of slow-fast vector fields in the form

$$\begin{cases} \dot{X} &= y - \sum_{k=0}^n p_k(\varepsilon, \lambda) X^k, \\ \dot{y} &= \varepsilon \hat{h}(X, y, \varepsilon, \lambda), \end{cases}$$

where $\hat{h}(X, y, \varepsilon, \lambda) := g(\psi(X, \varepsilon, \lambda), y, \varepsilon, \lambda) / \psi_X(X, \varepsilon, \lambda)$. Letting $Y = y - (\sum_{k=0}^n p_k(\varepsilon, \lambda) x^k - \sum_{k=0}^n p_k(0, \lambda) x^k)$ gives rise to a new system

$$\begin{cases} \dot{X} &= Y - \sum_{k=0}^n q_k(\lambda) X^k, \\ \dot{Y} &= \varepsilon h(X, Y, \varepsilon, \lambda), \end{cases} \quad (5)$$

where $q_k(\lambda) := -p_k(0, \lambda)$ and h is an analytic function. Due to assumptions (2) it is clear that $q_n(\lambda_0) \neq 0$, so an additional scaling of X allows to reduce to the case where $q_n(\lambda) \equiv \pm 1$, depending on the sign of $q_n(\lambda_0)$. In the case of a negative sign, we change $(Y, t) \rightarrow (-Y, -t)$ to obtain the form where the coefficient at x^n is exactly one. The difference between the required form and (5) is that we have to remove the terms in the summation corresponding to $k = 0$, and $k = n - 1$. Replacing X by $X - \frac{q_{n-1}(\lambda)}{n-1}$ allows to remove the term $q_{n-1}(\lambda) X^{n-1}$; afterwards we can remove the constant term by making a translation in the y -direction. This finishes the proof. \square

Remark 5. A normal form for an unfolding would mean that we consider the case

$$\begin{cases} \dot{x} &= y - x^n - \sum_{k=1}^{n-2} \mu_k x^k, \\ \dot{y} &= \varepsilon h(x, y, \varepsilon, \lambda), \end{cases}$$

where the parameter space is now $\Lambda \times M$, with $M = \{(\mu_1, \dots, \mu_{n-2})\}$.

2.2 Preparation theorem in the Gevrey class

A Gevrey version of Theorem 3 could be interpreted in two ways: given functions $f(x, \mu)$, we could consider the Gevrey property w.r.t. the x -variable, or w.r.t. the parameter. In view of this paper, we only need Gevrey-results in the parameter direction, and leave other questions for future research.

2.2.1 Definitions and statement of the result

Throughout this section we will consider functions in $(x, \lambda, \varepsilon)$; typically those functions behave analytically w.r.t. (x, λ) and are Gevrey-1 divergent in ε (definition soon follows). The parameter λ is supposed s -dimensional, $s \geq 1$. Since Gevrey functions are defined on sectors, we introduce a notation for a

sector: for $\rho > 0$ and $(\varphi, \theta) \in (0, 2\pi) \times (0, \pi)$, the complex sector $S_{\rho, \varphi, \theta}$ with vertex 0 is an open subset of \mathbb{C} defined by

$$S_{\rho, \varphi, \theta} := \left\{ z \in \mathbb{C} : \arg(z) \in (\varphi - \theta, \varphi + \theta), 0 < |z| < \rho \right\}.$$

Definition 6. Let $S_{\rho, \varphi, \theta}$ denote a complex sector. Let U be an open neighborhood of \mathbb{C}^s and let $a : U \times S_{\rho, \varphi, \theta} \rightarrow \mathbb{C}$, $(\lambda, \varepsilon) \mapsto a(\lambda, \varepsilon)$, be a bounded and analytic function. The function a is called *Gevrey-1* in ε (uniformly in λ) of type T , if there exists a formal power series $\sum_{k=0}^{\infty} a_k(\lambda) \varepsilon^k$ with analytic coefficient functions, if there exists a positive α , and if for every subsector S' there exists a positive constant $C_{S'}$ such that

$$\left| a(\lambda, \varepsilon) - \sum_{k=0}^{n-1} a_k(\lambda) \varepsilon^k \right| \leq C_{S'} T^n \Gamma(\alpha + n) |\varepsilon|^n, \quad \forall (\lambda, \varepsilon) \in U \times S'.$$

Gevrey-1 functions on sectors S can in principle not be evaluated at $\varepsilon = 0$, since $0 \notin S_{\rho, \varphi, \theta}$, but the Gevrey-property shows that such functions have a smooth extension at $s = 0$ and have a Taylor series there. At different places in the statement of the theorem below, we will therefore substitute $\varepsilon = 0$ without mentioning the smooth extension that is being used:

Theorem 7 (Preparation Theorem for Gevrey Functions). *Let U be an open set of \mathbb{C}^{s+1} containing the point $(0, \lambda_0) \in \mathbb{C} \times \mathbb{C}^s$. Let $f : U \times S \rightarrow \mathbb{C}$, $(x, \lambda, \varepsilon) \mapsto f(x, \lambda, \varepsilon)$ be an bounded and analytic function, with $S \subset \mathbb{C}$ is a sector of the origin with open angle smaller than π . Assume that f is Gevrey-1 in ε (uniformly in (x, λ)) and satisfies*

$$f(0, \lambda_0, 0) = \frac{\partial f}{\partial x}(0, \lambda_0, 0) = \dots = \frac{\partial^{n-1} f}{\partial x^{n-1}}(0, \lambda_0, 0) = 0, \quad \frac{\partial^n f}{\partial x^n}(0, \lambda_0, 0) \neq 0.$$

Then, possibly after shrinking the neighbourhood U and the radius of the sector S , there exists an analytic function $\alpha(x, \lambda, \varepsilon)$ defined on $U \times S$ so that the coordinate change $x = X + X^2 \alpha(X, \lambda, \varepsilon)$ is a right equivalence between f and a polynomial:

$$f(X + X^2 \alpha(X, \lambda, \varepsilon), \lambda, \varepsilon) = \sum_{k=0}^n p_k(\lambda, \varepsilon) X^k.$$

Moreover, the function α is analytic in (x, λ) and Gevrey-1 in ε .

Note: By the right equivalence, one automatically has

$$p_0(\lambda_0, 0) = \dots = p_{n-1}(\lambda_0, 0) = 0 \quad \text{and} \quad p_n(\lambda_0, 0) \neq 0.$$

The proof of Theorem 7 is spread over the next two subsections 2.2.2 and 2.2.3.

2.2.2 Banach spaces of functions defined on sectors

The definition of Gevrey function does not easily permit us to define a proper Banach space with proper norm containing Gevrey functions. Instead we will consider below Banach spaces of *bounded* functions defined on sectors, ignoring at that time the asymptotic properties required to have Gevrey-1 functions. There is a convenient way to obtain Gevrey-asymptotic properties of bounded functions defined on complex sectors, this is the theorem of Ramis-Sibuya. So before stating the Banach space in which we will work during the course of proving Theorem 7, let us first recall the relation between Gevrey asymptotics and bounded analytic functions on sectors.

For this purpose, we introduce a notion of sectorial covering: A *good sectorial covering* of the origin in \mathbb{C} is a finite (ordered) number of complex sectors $S_j := S_{\rho, \varphi_j, \theta_j}, j = 1, \dots, m$, so that the following holds:

- (i) $\cup_{j=1}^m S_j = B(0, \rho) \setminus \{0\}$.
- (ii) $2\theta_j < \pi$ for all $j = 1, \dots, m$ (i.e. no sector has an opening angle wider than π).
- (iii) For $j, k = 1, \dots, m$ with $j < k$, the intersection $S_j \cap S_k$ is not empty if and only if $k = j + 1$ or $(j, k) = (1, m)$.

The following proposition makes it possible for a given Gevrey function $f : U \times S \rightarrow \mathbb{C}$ on a sector S to analytically extend the definition on a full neighborhood of the origin, making a finite number of at most exponentially small jumps:

Theorem 8. *Let U be an open neighborhood of \mathbb{C}^s and S be a complex sector of the origin and $f : U \times S \rightarrow \mathbb{C}$ be analytic and Gevrey-1 asymptotic of type T to some formal series \hat{f} . Let $\tilde{T} > T$ be fixed. Then, there exists a good covering $(S_j)_{j=1, \dots, m}$ and a sequence of functions $(f_j)_{j=1, \dots, m}$ with $f_j : S_j \rightarrow \mathbb{C}$ is analytic and bounded and $(f_1, S_1) = (f, S)$, and all f_j have the following properties:*

- (i) *The functions f_j are also Gevrey-1 asymptotic (of type \tilde{T}) to \hat{f} .*
- (ii) *There is a $C > 0$ so that for all two adjacent sectors (S_j, S_k) one has*

$$|f_j(\lambda, \varepsilon) - f_k(\lambda, \varepsilon)| \leq C e^{-\frac{1}{\tilde{T}|\varepsilon|}} \quad \text{for } \lambda \in U, \varepsilon \in S_j \cap S_k.$$

The following theorem is the inverse of the preceding proposition.

Theorem 9 (Ramis-Sibuya). *Let $(S_j)_{j=1,\dots,m}$ be a good sectorial covering and let $f_j : U \times S_j \rightarrow \mathbb{C}$ be analytic and bounded. Suppose that for all two adjacent sectors (S_j, S_k) we have*

$$|f_j(\lambda, \varepsilon) - f_k(\lambda, \varepsilon)| = O(e^{-\frac{1}{T|\varepsilon|}}) \quad \forall \varepsilon \in S_j \cap S_k, \quad \text{as } \varepsilon \rightarrow 0,$$

for some $T > 0$. Then, all functions f_j are Gevrey-1 asymptotic of type T inside S_j to a common formal power series \hat{f} .

(For the proofs of the above two theorems, see for example Proposition 3.20 and Theorem 4.1 of [5].)

We denote by $B(0; r)$, $B(\lambda_0; \hat{r})$ the open ball with the radius r , \hat{r} centered at $0 \in \mathbb{C}$, $\lambda_0 \in \mathbb{C}^s$, respectively. We finally write $\mathcal{X}_{r, \hat{r}; \rho, (\varphi, \theta)}$ for the space of all bounded analytic functions $\alpha : B(0; r) \times B(\lambda_0; \hat{r}) \times S_{\rho, \varphi, \theta} \rightarrow \mathbb{C}$, endowed with the sup-norm, i.e.

$$\|\alpha\|_\infty := \sup_{(x, \lambda, \varepsilon) \in B(0; r) \times B(\lambda_0; \hat{r}) \times S_{\rho, \varphi, \theta}} |\alpha(x, \lambda, \varepsilon)|.$$

It is well-known that $(\mathcal{X}_{r, \hat{r}; \rho, (\varphi, \theta)}, \|\cdot\|_\infty)$ is a Banach space. Let $\mathcal{K}_{r, \hat{r}; \rho, (\varphi, \theta)}$ be the closed subset of the space $\mathcal{X}_{r, \hat{r}; \rho, (\varphi, \theta)}$ of functions α satisfying that

$$\alpha(0, \lambda, \varepsilon) = 0 \quad \forall \lambda, \varepsilon. \quad (6)$$

We finish this subsection by proving a lemma that will be referenced later on. It describes a sort of shift operator on the space just introduced, in the following sense: a function $\alpha(x, \lambda, \varepsilon)$, written down as a power series in powers of x , is shifted to the left, and the leftmost term (the term with x^1) is cancelled:

Lemma 10. *The map \mathcal{S} defined by*

$$\mathcal{S}\alpha(x, \lambda, \varepsilon) := \begin{cases} \frac{1}{x}\alpha(x, \lambda, \varepsilon) - \frac{\partial \alpha}{\partial x}(0, \lambda, \varepsilon), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0, \end{cases}$$

is a continuous linear map $\mathcal{K}_{r, \hat{r}; \rho, (\varphi, \theta)} \rightarrow \mathcal{K}_{r, \hat{r}; \rho, (\varphi, \theta)}$ with operator norm $\|\mathcal{S}\|_\infty \leq \frac{2}{r}$. Furthermore,

$$\mathcal{S}^n \alpha(x, \lambda, \varepsilon) = \begin{cases} \frac{1}{x^n} \left[\alpha(x, \lambda, \varepsilon) - \sum_{k=1}^n \frac{1}{k!} \frac{\partial^k \alpha}{\partial x^k}(0, \lambda, \varepsilon) x^k \right], & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Proof. Obviously, the map \mathcal{S} is well-defined and linear. It remains to estimate the norm $\|\mathcal{S}\|_\infty$. By the maximum principle, we have

$$\|\mathcal{S}\alpha\|_\infty \leq \frac{1}{r}\|\alpha\|_\infty + \sup_{\lambda \in B(\lambda_0; \hat{r}), \varepsilon \in S_{\rho, (\varphi, \theta)}} \left| \frac{\partial \alpha}{\partial x}(0, \lambda, \varepsilon) \right|. \quad (7)$$

Using Cauchy's integral formula, we obtain that

$$\left| \frac{\partial \alpha}{\partial x}(0, \lambda, \varepsilon) \right| = \left| \frac{1}{2\pi i} \oint \frac{\alpha(x, \lambda, \varepsilon)}{x^2} dx \right| \leq \frac{1}{r}\|\alpha\|_\infty,$$

which together with (7) implies that $\|\mathcal{S}\|_\infty \leq \frac{2}{r}$. The statement on \mathcal{S}^n follows directly from the definition of the operator \mathcal{S} . \square

2.2.3 Proof of Theorem 7

Without loss of generality, and merely for the sake of simplicity, we assume that $\frac{\partial^n f}{\partial x^n}(0, \lambda_0, 0) = 1$.

Let $(S_j)_{j=1, \dots, m}$, where $S_j = S_{\rho, \varphi_j, \theta_j}$, be a good covering and $(f_j)_{j=1, \dots, m}$ a sequence of functions with $f_j : B(0; r) \times B(\lambda_0; \hat{r}) \times S_j \rightarrow \mathbb{C}$ is bounded and analytic and $(f_1, S_1) = (f, S)$ and all f_j satisfy properties (i) and (ii) in Theorem 8, i.e.

- (i) The functions f_j are Gevrey-1 asymptotic to $\hat{f}(x, \lambda, \varepsilon) = \sum_{k=0}^{\infty} a_k(x, \lambda) \varepsilon^k$.
- (ii) There is a $L > 0$ so that for all two adjacent sectors (S_j, S_k) one has

$$|f_j(x, \lambda, \varepsilon) - f_k(x, \lambda, \varepsilon)| \leq L e^{-\frac{1}{T|\varepsilon|}} \quad \text{for } \varepsilon \in S_j \cap S_k,$$

where T is a positive constant.

Since all f_j share the same asymptotic expansion w.r.t. ε with the one from f , Assumption (2) generalizes to all f_j :

$$\frac{\partial^k f_j}{\partial x^k}(0, \lambda_0, 0) = 0 \quad \text{for } k = 0, \dots, n-1 \quad \text{and} \quad \frac{\partial^n f_j}{\partial x^n}(0, \lambda_0, 0) = 1. \quad (8)$$

In particular, it means that

$$f_j \in \mathcal{X}_j := \mathcal{X}_{r, \hat{r}; \rho, (\varphi_j, \theta_j)}.$$

In the sequel of the proof, we will most often work in the Banach space

$$\mathcal{K}_j := \mathcal{K}_{r, \hat{r}; \rho, (\varphi_j, \theta_j)}$$

and look in there for a change of variables of the form $X = x + x\alpha(x, \lambda, \varepsilon)$, thereby exchanging the roles of x and X as compared to the statement of the theorem, and assuming that $\alpha \in \mathcal{K}_j$. The α in the statement is later obtained after dividing by x (we will comment on it at the end of the proof). Note that we do not necessarily have $f_j \in \mathcal{K}_j$.

For each fixed j , we will find an $\alpha_j \in \mathcal{K}_j$ so that $f_j(x + x\alpha, \lambda, \varepsilon)$ is a polynomial in x with coefficient functions in (λ, ε) . In a second step, we show that each two choices α_j and α_k are exponentially close to each other in the intersection of adjacent sectors, hence allowing us to use the Theorem of Ramis-Sibuya to conclude.

Step 1. Fixed point $\alpha_j \in \mathcal{K}_j$. We Taylor expand f_j as follows:

$$f_j(x, \lambda, \varepsilon) = \sum_{k=0}^n \frac{\partial^k f_j}{\partial x^k}(0, \lambda, \varepsilon) \frac{x^k}{k!} + g_j(x, \lambda, \varepsilon). \quad (9)$$

Note that $g_j = x^n \cdot \mathcal{S}^n(f_j)$ is part of \mathcal{K}_j , and satisfies

$$\frac{\partial^k g_j}{\partial x^k}(0, \lambda, \varepsilon) = 0 \quad \text{for } k = 0, \dots, n. \quad (10)$$

Since $g_j \in \mathcal{K}_j$ is bounded, it follows that using the Cauchy's integral formula and upon making the radius r a bit smaller if necessary, that we may assume that $\frac{\partial^k g_j}{\partial x^k}$ are also in the space \mathcal{K}_j . This additional property of the function g_j and (10) implies that there exists a positive constant C_1 such that for all $(x, \lambda, \varepsilon) \in B(0; 2r) \times B(\lambda_0; \hat{r}) \times S_{\rho; \varphi_j, \theta_j}$ we have

$$|g_j(x, \lambda, \varepsilon)| \leq C_1 r^{n+1} \quad \text{and} \quad \left| \frac{\partial g_j}{\partial x}(x, \lambda, \varepsilon) \right| \leq C_1 r^n. \quad (11)$$

By (9) for each $\alpha \in \mathcal{K}_j$, we have the following expansion:

$$f_j(x + x\alpha, \lambda, \varepsilon) = \sum_{k=0}^n \frac{1}{k!} \frac{\partial^k f_j}{\partial x^k}(0, \lambda, \varepsilon) (x + x\alpha)^k + g_j(x + x\alpha, \lambda, \varepsilon).$$

which we split up as the polynomial $\sum_{k=0}^n \frac{1}{k!} \frac{\partial^k f_j}{\partial x^k}(0, \lambda, \varepsilon) x^k$ together with a remainder term

$$g_j(x, \lambda, \varepsilon) + \frac{1}{(n-1)!} \frac{\partial^n f_j}{\partial x^n}(0, \lambda, \varepsilon) x^n \alpha(x, \lambda, \varepsilon) + \mathcal{R}_j(\alpha)(x, \lambda, \varepsilon), \quad (12)$$

where the function $\mathcal{R}_j : \mathcal{K}_j \rightarrow \mathcal{K}_j$ is given by

$$\begin{aligned} \mathcal{R}_j(\alpha) &:= g_j(x + x\alpha, \lambda, \varepsilon) - g_j(x, \lambda, \varepsilon) \\ &+ \sum_{k=1}^{n-1} \frac{1}{k!} \frac{\partial^k f_j}{\partial x^k}(0, \lambda, \varepsilon) \sum_{i=0}^{k-1} C_k^i x^k \alpha^{k-i} \\ &+ \frac{1}{n!} \frac{\partial^n f_j}{\partial x^n}(0, \lambda, \varepsilon) \sum_{i=0}^{n-2} C_n^i x^n \alpha^{n-i}. \end{aligned}$$

Note that the map \mathcal{R}_j is a well-defined map $\mathcal{K}_j \rightarrow \mathcal{K}_j$. In order for the change of variables $X = x + x\alpha$ to lead to a polynomial $f(x + x\alpha, \lambda, \varepsilon)$, we need that \mathcal{S}^n applied to (12) is equal to 0, where \mathcal{S} is the shift-operator introduced in Lemma 10. We therefore have

$$\mathcal{S}^n g_j + \frac{1}{(n-1)!} \frac{\partial^n f_j}{\partial x^n}(0, \lambda, \varepsilon) \alpha + \mathcal{S}^n \mathcal{R}_j(\alpha) = 0.$$

We hence find the following fixed-point formulation for α :

$$\alpha = -\frac{(n-1)!}{\frac{\partial^n f_j}{\partial x^n}(0, \lambda, \varepsilon)} \mathcal{S}^n (g_j + \mathcal{R}_j(\alpha)). \quad (13)$$

Now α is a fixed point on the map $\mathcal{T}_j : \mathcal{K}_j \rightarrow \mathcal{K}_j$ defined by

$$\mathcal{T}_j \alpha := -\frac{(n-1)!}{\frac{\partial^n f_j}{\partial x^n}(0, \lambda, \varepsilon)} \mathcal{S}^n (g_j + \mathcal{R}_j(\alpha)). \quad (14)$$

For a $\gamma > 0$, let $\bar{\mathcal{B}}(\gamma) \subset \mathcal{K}_j$ denote the closed ball of elements α of norm bounded by γ . We will now choose γ, r, \hat{r}, ρ such that the map \mathcal{T}_j maps $\bar{\mathcal{B}}(\gamma)$ into itself and is contractive on $\bar{\mathcal{B}}(\gamma)$. Let $\alpha, \hat{\alpha} \in \bar{\mathcal{B}}(\gamma)$ be arbitrary and $\alpha \neq \hat{\alpha}$. To estimate $\|\mathcal{T}_j \alpha - \mathcal{T}_j \hat{\alpha}\|_\infty$, we consider the difference $\|\mathcal{R}_j(\alpha) - \mathcal{R}_j(\hat{\alpha})\|_\infty$. A direct computation yields that for any $k = 1, \dots, n-1$:

$$\frac{\left\| \sum_{i=0}^{k-1} C_k^i x^k \alpha^{k-i} - \sum_{i=0}^{k-1} C_k^i x^k \hat{\alpha}^{k-i} \right\|_\infty}{\|\alpha - \hat{\alpha}\|_\infty} \leq \sum_{i=0}^{k-1} (k-i) C_k^i r^k \gamma^{k-i-1}.$$

Similarly one has

$$\frac{\left\| \sum_{i=0}^{n-2} C_n^i x^n \alpha^{n-i} - \sum_{i=0}^{n-2} C_n^i x^n \hat{\alpha}^{n-i} \right\|_\infty}{\|\alpha - \hat{\alpha}\|_\infty} \leq \sum_{i=0}^{n-2} (n-i) C_n^i r^n \gamma^{n-i-1}.$$

By the Mean Value Theorem, for any $x \in B(0; r)$, $\lambda \in B(\lambda_0; \hat{r})$ and $\varepsilon \in S_{\rho; \varphi_j, \theta_j}$ we have

$$\frac{|g_j(x + x\alpha, \lambda, \varepsilon) - g_j(x + x\hat{\alpha}, \lambda, \varepsilon)|}{\|\alpha - \hat{\alpha}\|_\infty} \leq r \sup_{(*)} \left| \frac{\partial g_j}{\partial x}(x, \lambda, \varepsilon) \right|,$$

where the supremum $(*)$ is taken for $x \in B(0; r + r\gamma)$, $\lambda \in B(\lambda_0; \hat{r})$, $\varepsilon \in S_{\rho; \varphi_j, \theta_j}$. Let $\gamma := r^{\frac{1}{2}}$ with $r < 1$ which is chosen later. Combining the preceding inequalities yields that

$$\begin{aligned} \frac{\|\mathcal{R}_j(\alpha) - \mathcal{R}_j(\hat{\alpha})\|_\infty}{\|\alpha - \hat{\alpha}\|_\infty} &\leq r \sup_{(*)} \left| \frac{\partial g_j}{\partial x}(x, \lambda, \varepsilon) \right| \\ &+ C_2 \sum_{k=1}^{n-1} \sup_{(**)} \left| \frac{\partial^k f_j}{\partial x^k}(0, \lambda, \varepsilon) \right| \\ &+ C_3 \sup_{(**)} \left| \frac{\partial^n f_j}{\partial x^n}(0, \lambda, \varepsilon) \right| r^{n+\frac{1}{2}}, \end{aligned}$$

where the supremum $(**)$ is taken for $\lambda \in B(\lambda_0; \hat{r})$, $\varepsilon \in S_{\rho; \varphi_j, \theta_j}$, and where C_2 and C_3 depend only on n . Note that $\frac{\partial^k f_j}{\partial x^k}(0, \lambda_0, 0) = 0$ for $k = 1, \dots, n-1$. Hence, choosing $\hat{r} = \rho := r^{n+\frac{1}{2}}$ gives that

$$\sum_{k=1}^{n-1} \sup_{(**)} \left| \frac{\partial^k f_j}{\partial x^k}(0, \lambda, \varepsilon) \right| = O(r^{n+\frac{1}{2}}),$$

which together with (11) implies that there exists a constant C such that

$$\|\mathcal{R}_j(\alpha) - \mathcal{R}_j(\hat{\alpha})\|_\infty \leq C r^{n+\frac{1}{2}} \|\alpha - \hat{\alpha}\|_\infty.$$

Finally, from $\frac{\partial^n f_j}{\partial x^n}(0, \lambda_0, 0) = 1$, we can choose r small enough such that

$$\left| \frac{\partial^n f_j}{\partial x^n}(0, \lambda, \varepsilon) \right| \geq \frac{1}{2} \quad \text{for all } \lambda \in B(\lambda_0, \hat{r}), \varepsilon \in S_{\rho; \varphi_j, \theta_j}.$$

Thus, by the definition of \mathcal{T}_j as in (14) and Lemma 10 we have

$$\begin{aligned} \|\mathcal{T}_j(\alpha) - \mathcal{T}_j(\hat{\alpha})\|_\infty &\leq 2(n-1)! \frac{2^n}{r^n} \|\mathcal{R}_j(\alpha) - \mathcal{R}_j(\hat{\alpha})\|_\infty \\ &\leq C 2^{n+1} (n-1)! r^{\frac{1}{2}} \|\alpha - \hat{\alpha}\|_\infty. \end{aligned}$$

So, letting r small enough so that $\|\mathcal{T}_j(\alpha) - \mathcal{T}_j(\hat{\alpha})\|_\infty \leq \frac{1}{2}\|\alpha - \hat{\alpha}\|_\infty$, the map \mathcal{T}_j is Lipschitz with Lipschitz constant less than $L < \frac{1}{2}$. It remains to choose r such that \mathcal{T}_j maps the ball $\overline{B}(\gamma)$ into itself. To obtain this property, we use

$$\|\mathcal{T}_j(\alpha)\|_\infty \leq \|\mathcal{T}_j(\alpha) - \mathcal{T}_j(0)\|_\infty + \|\mathcal{T}_j(0)\|_\infty \leq L\|\alpha - 0\|_\infty + \|\mathcal{T}_j(0)\|_\infty$$

$$\|\mathcal{T}_j(0)\|_\infty = (n-1)! \left\| \frac{\mathcal{S}^n g_j}{\frac{\partial^n f_j}{\partial x^n}(0, \lambda, \varepsilon)} \right\|_\infty \leq \frac{2^{n+1}(n-1)!}{r^n} \|g_j\|_\infty,$$

which together with (11) implies that

$$\|\mathcal{T}_j(0)\|_\infty \leq C_1 \frac{2^{n+1}(n-1)!}{r^n} r^{n+1} = C_1 2^{n+1}(n-1)! r^{\frac{1}{2}} \gamma.$$

Hence, we can choose r small enough such that $\|\mathcal{T}_j(0)\|_\infty \leq \frac{\gamma}{2}$. So, for any $\alpha \in \overline{B}(\gamma)$ we have $\|\mathcal{T}_j(\alpha)\|_\infty \leq \gamma$.

Step 2. Exponentially small differences. Let $\alpha_j \in \mathcal{K}_j$ be the unique fixed point of \mathcal{T}_j , for $j = 1, \dots, m$. We will prove that α_j is Gevrey-1 in ε (uniformly in (x, λ)), using the Theorem of Ramis-Sibuya. Note that $f_j - f_k$ are exponentially small in adjacent sectors, by construction. By application of Cauchy's lemma it is easily seen that we may assume that any finite number of partial derivatives w.r.t. x have the same property concerning exponential closeness to each other. Similarly, $\mathcal{S}^n f_j - \mathcal{S}^n g_k$ is exponentially small.

As a consequence, not only $g_j - g_k$ is exponentially small (remember that $g_j = x^n \mathcal{S}^n f_j$), but also $\mathcal{R}_j(\alpha) - \mathcal{R}_k(\alpha)$ is exponentially small for all $\|\alpha\|_\infty \leq \gamma$, it suffices to take a look at the definition of \mathcal{R}_j to see this.

From those properties, it easily follows that also $\mathcal{T}_j(\alpha) - \mathcal{T}_k(\alpha)$ is exponentially small. Let us now write

$$\begin{aligned} \|\alpha_j - \alpha_k\|_\infty &= \|\mathcal{T}_j(\alpha_j) - \mathcal{T}_k(\alpha_k)\|_\infty \\ &= \|\mathcal{T}_j(\alpha_j) - \mathcal{T}_j(\alpha_k)\|_\infty + \|\mathcal{T}_j(\alpha_k) - \mathcal{T}_k(\alpha_k)\|_\infty \\ &\leq \frac{1}{2} \|\alpha_j - \alpha_k\|_\infty + M e^{-1/T|\varepsilon|}. \end{aligned}$$

The exponential closeness of α_j and α_k follow directly, and the theorem of Ramis-Sibuya applies, showing all α_j are Gevrey-1 asymptotic on S_j , uniformly in (x, λ) .

The transformation announced in the statement of the theorem can be defined as $\alpha := \frac{1}{x}\alpha_1$. The division by x does not harm the Gevrey estimates since $\alpha_1 \in \mathcal{K}_1$, so it is $O(x)$ anyway (a detailed proof of the Gevrey properties of α would use the maximum principle to deal with the division in x). This completes the proof of Theorem 7.

2.3 A formal version of Theorem 7

Here, we establish a formal series version of the Gevrey Preparation Theorem 7. To that end, let us first introduce a definition

Definition 11. A formal series $\hat{a}(\varepsilon) = \sum_{i=0}^{\infty} a_i \varepsilon^i$ is called *Gevrey-1* in ε of type A , if there exist positive constants C, α such that

$$|a_n| \leq CA^{n/\sigma} \Gamma(\alpha + n).$$

The theorem of Borell-Ritt-Gevrey (see for example See [5, Lemma 3.15]) states that for any Gevrey-1 series \hat{f} of type A and for any sector $S = S_{\rho, \varphi, \theta}$ (with “small” opening angle $2\theta \leq \pi$) there exists a bounded analytic function $f : S \rightarrow \mathbb{C}$ so that f is Gevrey-1 asymptotic to \hat{f} of type $T := A/\cos\theta$ (see definition of Gevrey function). This results allows to relate a statement on formal series to a statement on Gevrey functions and is the key to proving the next theorem:

Theorem 12 (Preparation Theorem for Gevrey Formal Series). *Let U be an open set of \mathbb{C}^{s+1} containing the point $(0, \lambda_0) \in \mathbb{C} \times \mathbb{C}^s$. Let $(a_i(x, \lambda))_{i=0}^{\infty}$ be a sequence of analytical functions from U to \mathbb{C} . Suppose that the formal series $\sum_{i=0}^{\infty} a_i(x, \lambda) \varepsilon^i$ is Gevrey-1 in ε (uniformly in (x, λ)). Assuming additionally that*

$$a_0(0, \lambda_0) = \dots = \frac{\partial^{n-1} a_0}{\partial x^{n-1}}(0, \lambda_0) = 0, \quad \frac{\partial^n a_0}{\partial x^{n-1}}(0, \lambda_0) \neq 0. \quad (15)$$

Then, there exists a Gevrey-1 formal series $\sum_{i=0}^{\infty} \alpha_i(x, \lambda) \varepsilon^i$ in ε such that the change of variable $X = x + x^2 \sum_{i=0}^{\infty} \alpha_i(x, \lambda) \varepsilon^i$ yields that

$$\sum_{i=0}^{\infty} a_i(X, \lambda) \varepsilon^i = \sum_{k=0}^n p_k(\lambda, \varepsilon) x^k,$$

where $p_k(\lambda, \varepsilon) = \sum_{i=0}^{\infty} a_i^{(k)}(\lambda) \varepsilon^i$ is a Gevrey-1 formal series in ε .

Proof. According to the theorem of Borel-Ritt-Gevrey, there exists an analytic function $f : U \times S_{\rho, \varphi, \theta} \rightarrow \mathbb{C}$ that is Gevrey-1 asymptotic to the formal power series $\sum_{i=0}^{\infty} a_i(x, \lambda) \varepsilon^i$. Clearly (15) implies the equivalent condition in the statement of Theorem 7. Hence in the light of that theorem, there exists an analytic function $\alpha(x, \lambda, \varepsilon)$ defined on (a possibly smaller) $U \times S$ such that the coordinate change $x = X + X^2 \alpha(X, \lambda, \varepsilon)$ is a right equivalence between f and a polynomial:

$$f(X + X^2 \alpha(X, \lambda, \varepsilon), \lambda, \varepsilon) = \sum_{k=0}^n p_k(\lambda, \varepsilon) X^k.$$

Then the Gevrey-1 asymptotic $\sum_{i=0}^{\infty} \alpha_i(x, \lambda) \varepsilon^i$ of $\alpha(x, \lambda, \varepsilon)$ is the desired formal series. \square

3 Proof of Theorem 1

In this section, we use the Gevrey preparation theorem to prove Theorem 1. We follow the guidelines of [2]. As a first step, we use Corollary 4 and a linear rescaling of the singular parameter $\varepsilon \mapsto \varepsilon \delta$, to rewrite (1) as

$$\begin{cases} \dot{x} &= y - q(x, \lambda), \\ \dot{y} &= \varepsilon \delta h(x, y, \lambda, \varepsilon \delta), \end{cases} \quad (16)$$

where h is analytic and $q(x, \lambda) = x^n + \sum_{i=1}^{n-2} q_i(\lambda) x^i$ with q_i are analytic and q_i vanish at $\lambda = \lambda_0$, and δ is a small but fixed positive constant.

The dilatation trick $\varepsilon \mapsto \delta \varepsilon$ will become handy since the requested normal form transformation will be seen as a fixed point of a map, whose Lipschitz constant is shown to be $O(\delta)$. Hence by choosing δ small enough, the contraction property is guaranteed. Since for any fixed $\delta > 0$, the requested normal form transformation will be valid for arbitrarily small values of ε in a sector, it suffices to rescale back the sector by a factor δ to find the normal form result for the original system without δ . In other words, we will restrict to proving Theorem 1 for (16), for sufficiently small δ .

Instead of directly establishing an equivalence between (16) and (3), we first establish an intermediary equivalence between (16) and

$$\begin{cases} \dot{x} &= Y - \varphi(x, \varepsilon, \lambda), \\ \dot{Y} &= \varepsilon \tilde{G}(x, \lambda, \varepsilon) + \varepsilon \tilde{R}(x, Y, \lambda, \varepsilon), \end{cases} \quad (17)$$

with \tilde{R} exponentially small in ε . The homological equation arises from applying a coordinate change of the form $Y = y + \varepsilon V(x, y, \lambda, \varepsilon)$ to (17) and imposing that the resulting system is parallel to (16):

$$\begin{vmatrix} y - q(x, \lambda) & (y + \varepsilon V - \varphi)(1 + \varepsilon V_y) \\ \delta h(x, y, \lambda, \delta \varepsilon) & \tilde{G} + \tilde{R}(x, y + \varepsilon V, \lambda, \varepsilon) - V_x(y + \varepsilon V - \varphi) \end{vmatrix} = 0. \quad (18)$$

Resolving (18) in terms of the unknowns φ , \tilde{G} , \tilde{R} and V is the topic of Section 3.2. Section 3.1 prepares a Banach setting to do so.

3.1 Banach spaces and appropriate norms

Below we will consider power series in ε with functions in (x, y, λ) as coefficients. The coefficient function with ε^k will be assigned to a k -dependent Banach space. Let r, \hat{r} be sufficient small and consider for a given $0 < \rho < r$ the scalar function

$$d(y) = \begin{cases} r - |y| & \text{if } |y| \geq \rho, \\ r - \rho & \text{if } |y| < \rho. \end{cases}$$

This function is used to define, for each $k \in \mathbb{N}$, a norm that is equivalent to the sup-norm (on a slightly smaller disk):

$$\|a\|_k := \sup_{|x|, |y| < r, |\lambda - \lambda_0| < \hat{r}} |a(x, y, r)| d(y)^k. \quad (19)$$

Such sequence of norms is called a Nagumo norm, following [9]. Note that the Nagumo norm defined above depends on the choice of ρ but we do not indicate this. We also refer the reader to [1, Section 3] for a list of references relating to this notion. Let us now define

$$\mathcal{A}_{r, \hat{r}} := \{a(x, y, \lambda) : a \text{ is analytic for } |x|, |y| < r, |\lambda - \lambda_0| < \hat{r}, \|a\|_k < \infty\}.$$

In the next lemma, the function q and its coefficient functions q_i are taken from the family of vector fields (16). Using the fact that $q_1(\lambda_0) = \dots = q_{n-1}(\lambda_0) = 0$, for each r we can choose \hat{r} such that

$$|q(x, \lambda)| \leq 2r^n \quad \text{for } |x| \leq r, |\lambda - \lambda_0| \leq \hat{r}. \quad (20)$$

The lemma below contains three parts. The first part is Nagumo's lemma, which shows that one can control the size of a derivative of a function based on the size of the function itself (without shrinking the domain). This is

precisely the reason why the norms are chosen in the above particular way. The second and third parts of the lemma can be understood as follows: suppose one writes the function $a(x, y, \lambda)$ as a series

$$\sum_{k=0}^{\infty} a_k(x, \lambda)(y - q(x, \lambda))^k,$$

then part (ii) of the lemma associates to a function a the zero-order term of the series, namely a_0 , and part (iii) of the lemma shifts the series to the left:

Lemma 13. *Suppose that r, ρ satisfies $0 < 2r^n < \rho < r$ and \hat{r} satisfies (20). Then, the following statements hold for all $k \in \mathbb{N}$:*

(i) *For any function $a \in \mathcal{A}_{r, \hat{r}}$, we have $\|\frac{\partial a}{\partial y}\|_{k+1} \leq (k+1)e\|a\|_k$, where $e := \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n$ is the Euler number.*

(ii) *Define the function $\mathcal{Z} : \mathcal{A}_{r, \hat{r}} \rightarrow \mathcal{A}_{r, \hat{r}}$ by $\mathcal{Z}a(x, y, \lambda) := a(x, q(x, \lambda), \lambda)$. Then, $\|\mathcal{Z}a\|_k \leq \|a\|_k$.*

(iii) *Define the shift function $\mathcal{S} : \mathcal{A}_{r, \hat{r}} \rightarrow \mathcal{A}_{r, \hat{r}}$ by*

$$\mathcal{S}a(x, y, \lambda) := \begin{cases} \frac{a(x, y, \lambda) - a(x, q(x, \lambda), \lambda)}{y - q(x, \lambda)}, & \text{if } y \neq q(x, \lambda), \\ \frac{\partial a}{\partial y}(x, q(x, \lambda), \lambda), & \text{if } y = q(x, \lambda). \end{cases}$$

Then,

$$\|\mathcal{S}a\|_k \leq \frac{2}{\rho - 2r^n} \|a\|_k.$$

Proof. Analog to [2, Lemma 1]. □

Space of formal power series

We consider formal power series in ε with coefficient functions in $\mathcal{A}_{r, \hat{r}}$

$$A(x, y, \lambda, \varepsilon) = \sum_{k=0}^{\infty} A_k(x, y, \lambda) \varepsilon^k.$$

Let us now define a suitable norm on the space of formal power series. Choose $\rho < r$ and define

$$\mathcal{E}_{r, \hat{r}, \rho} := \left\{ A(x, y, \lambda, \varepsilon) = \sum_{k=0}^{\infty} A_k(x, y, \lambda) \varepsilon^k : \|A\|_{r, \hat{r}, \rho} := \sum_{k=0}^{\infty} \frac{\|A_k\|_k}{k!} < \infty \right\}. \quad (21)$$

In the following proposition, we state and prove some fundamental properties of the space $\mathcal{E}_{r,\hat{r},\rho}$.

Proposition 14. *The space $\mathcal{E}_{r,\hat{r},\rho}$ is a Banach space when equipped with the norm $\|\cdot\|_{r,\hat{r},\rho}$ defined as in (21). Furthermore, the following statements hold:*

- (i) *The norm $\|\cdot\|_{r,\hat{r},\rho}$ is sub-multiplicative, i.e. for $A, B \in \mathcal{E}_{r,\hat{r},\rho}$, we have $\|AB\|_{r,\hat{r},\rho} \leq \|A\|_{r,\hat{r},\rho}\|B\|_{r,\hat{r},\rho}$.*
- (ii) *Let $A = \sum_{k=0}^{\infty} A_k(x, y, \lambda)\varepsilon^k \in \mathcal{E}_{r,\hat{r},\rho}$. Then, A is a Gevrey formal series of order 1 in ε (uniformly in (x, y, λ)).*

Proof. Obviously, the space $(\mathcal{E}_{r,\hat{r},\rho}, \|\cdot\|_{r,\hat{r},\rho})$ is a normed vector space, see e.g. [1, Page 102]. It remains to show that $(\mathcal{E}_{r,\hat{r},\rho}, \|\cdot\|_{r,\hat{r},\rho})$ is complete. For this purpose, let $(A^n)_{n \in \mathbb{N}}$ be a Cauchy sequence of elements in $\mathcal{E}_{r,\hat{r},\rho}$. Then, for each $k \in \mathbb{N}$ the sequence of functions $(A_k^n(x, y, \lambda))_{n \in \mathbb{N}}$ is also a Cauchy sequence in the space $(\mathcal{A}_{r,\hat{r}}, \|\cdot\|_k)$. Thus, by (19) for each fixed (x, y, λ) with $|x|, |y| < r, |\lambda - \lambda_0| < \hat{r}$, the limit $A_k(x, y, \lambda) := \lim_{n \rightarrow \infty} A_k^n(x, y, \lambda)$ exists. Due to analyticity of $A_k^n(x, y, \lambda)$, the limit function $A_k(x, y, \lambda)$ is also analytic. Let $\varepsilon > 0$ be arbitrary. Then, there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have

$$\sum_{k=0}^{\infty} \frac{\|A_k^n - A_k^m\|_k}{k!} \leq \frac{\varepsilon}{2},$$

letting $m \rightarrow \infty$ yields that

$$\sum_{k=0}^{\infty} \frac{\|A_k^n - A_k\|_k}{k!} \leq \varepsilon,$$

which proves that $\lim_{n \rightarrow \infty} \|A^n - A\|_{r,\hat{r},\rho} = 0$ and thus we have that the space $(\mathcal{E}_{r,\hat{r},\rho}, \|\cdot\|_{r,\hat{r},\rho})$ is a Banach space. To conclude the proof, we are now going to prove the properties (i) and (ii) of this space:

(i) Let $A = \sum_{k=0}^{\infty} A_k(x, y, \lambda, \varepsilon)\varepsilon^k$ and $B = \sum_{k=0}^{\infty} B_k(x, y, \lambda, \varepsilon)\varepsilon^k$ be elements of $\mathcal{E}_{r,\hat{r},\rho}$. Then,

$$\begin{aligned} \frac{\|(AB)_n\|_n}{n!} &= \frac{\|\sum_{k=0}^n A_k(x, y, \lambda)B_{n-k}(x, y, \lambda)\|_n}{n!} \\ &\leq \sum_{k=0}^n \frac{\|A_k(x, y, \lambda)\|_k}{k!} \frac{\|B_{n-k}(x, y, \lambda)\|_{n-k}}{(n-k)!}, \end{aligned}$$

which implies that $\|AB\|_{r,\hat{r},\rho} \leq \|A\|_{r,\hat{r},\rho}\|B\|_{r,\hat{r},\rho}$.

(ii) Let $A = \sum_{k=0}^{\infty} A_k(x, y, \lambda) \varepsilon^k \in \mathcal{E}_{r, \hat{r}, \rho}$. By (21), there exists $C > 0$ such that

$$\|A_n(x, y, \lambda)\|_n \leq Cn! \quad \text{for all } n \in \mathbb{N}.$$

From definition of Nagumo's norm as in (19), we derive that

$$\|A_n(x, y, \lambda)\|_n \geq \sup_{|x|, |y| < \rho, |\lambda - \lambda| < \hat{r}} |A_n(x, y, \lambda)| (r - \rho)^n,$$

which implies that

$$\sup_{|x|, |y| < \rho, |\lambda - \lambda| < \hat{r}} |A_n(x, y, \lambda)| \leq C \left(\frac{1}{r - \rho} \right)^n n!.$$

Thus, A is a Gevrey formal series of order 1 in ε (uniformly in (x, y, λ)). \square

Some useful lemmas

We define the following operators $\frac{\partial}{\partial y}, \mathcal{S}, \mathcal{I} : \mathcal{E}_{r, \hat{r}, \rho} \rightarrow \mathcal{E}_{r, \hat{r}, \rho}$ as follows: For each $A(x, y, \lambda, \varepsilon) = \sum_{k=0}^{\infty} a_k(x, y, \lambda) \varepsilon^k \in \mathcal{E}_{r, \hat{r}, \rho}$, let

$$\begin{aligned} \frac{\partial}{\partial y} A(x, y, \lambda, \varepsilon) &= \sum_{k=0}^{\infty} \frac{\partial a_k}{\partial y}(x, y, \lambda) \varepsilon^k, \\ \mathcal{S}A(x, y, \lambda, \varepsilon) &= \sum_{k=0}^{\infty} \mathcal{S}a_k(x, y, \lambda) \varepsilon^k \\ \mathcal{I}A(x, y, \lambda, \varepsilon) &:= \sum_{k=0}^{\infty} \left(\int_0^x a_k(u, y, \lambda) du \right) \varepsilon^k. \end{aligned}$$

Lemma 15. *The operators $\varepsilon \frac{\partial}{\partial y}, \mathcal{S}$ and \mathcal{I} are linear operators from $\mathcal{E}_{r, \hat{r}, \rho}$ into itself and*

$$\left\| \varepsilon \frac{\partial}{\partial y} \right\|_{r, \hat{r}, \rho} \leq e, \quad \|\mathcal{S}\|_{r, \hat{r}, \rho} \leq \frac{2}{\rho - 2r^n}, \quad \|\mathcal{I}\|_{r, \hat{r}, \rho} \leq r. \quad (22)$$

Proof. See [2, Proposition 2 & 3]. \square

Lemma 16. *For all $A \in \mathcal{E}_{r, \hat{r}, \rho}$ with $\|A\|_{r, \hat{r}, \rho} < \frac{1}{r}$, then $\frac{1}{1 - \varepsilon A} \in \mathcal{E}_{r, \hat{r}, \rho}$ and*

$$\left\| \frac{1}{1 - \varepsilon A} \right\|_{r, \hat{r}, \rho} \leq \frac{1}{1 - r\|A\|_{r, \hat{r}, \rho}}.$$

The map $A \mapsto \frac{1}{1 - \varepsilon A}$ is Lipschitz continuous with Lipschitz constant $4r$ on the closed ball around 0 with radius $\frac{1}{2r}$.

Proof. See [2, Lemma 2]. \square

3.2 Existence of formal normal form transformation

We first eliminate all but one unknown in the homological equation. It will lead to the fixed-point formula (24). The reduction to this form will be kept concise as it can also be found in [2]. We look at (18) from a formal power series point of view, thereby making one of the unknowns (the exponentially small \tilde{R}) vanish:

$$\begin{vmatrix} y - q(x, \lambda) & (y + \varepsilon V - \varphi)(1 + \varepsilon V_y) \\ \delta h(x, y, \lambda, \delta\varepsilon) & \tilde{G} - V_x(y + \varepsilon V - \varphi) \end{vmatrix} = 0. \quad (23)$$

A necessary condition for (23) to be satisfied is that the zero sets of both elements on the top row are identical (they are nullclines of the two vector fields). It means that $\varphi = q + \varepsilon V(x, q(x, \lambda), \lambda, \varepsilon)$. Written in terms of the operators defined in the previous section, we have $\varphi = q + \varepsilon \mathcal{Z}V$, and we hence also have

$$y + \varepsilon V - \varphi = (y - q) \cdot \mathcal{S}(y + \varepsilon V - \varphi) = (y - q)(1 + \varepsilon \mathcal{S}V).$$

This observation allows to eliminate the unknown φ from (23), which then reduces to

$$\tilde{G} - V_x(y - q)(1 + \varepsilon \mathcal{S}V) = \delta h \cdot (1 + \varepsilon \mathcal{S}V)(1 + \varepsilon V_y)$$

There are still two unknowns here, \tilde{G} and V . However \tilde{G} does not depend on y , implying that $\mathcal{S}(\tilde{G}) = 0$, so applying \mathcal{S} to both sides of the equation eliminates \tilde{G} and reduces (after remembering that \mathcal{S} acts by shifting a power series in $y - q$ to the left) to

$$-V_x(1 + \varepsilon \mathcal{S}V) = \delta \mathcal{S} [h \cdot (1 + \varepsilon \mathcal{S}V)(1 + \varepsilon V_y)].$$

We see that if $V \in \mathcal{E}_{r, \hat{r}, \rho}$ is a fixed point of the map $\mathcal{L} : \mathcal{E}_{r, \hat{r}, \rho} \rightarrow \mathcal{E}_{r, \hat{r}, \rho}$ defined by

$$\mathcal{L}V := -\delta \mathcal{I} \left(\frac{\mathcal{S} \left[h \left(1 + \varepsilon \frac{\partial V}{\partial y} \right) (1 + \varepsilon \mathcal{S}V) \right]}{1 + \varepsilon \mathcal{S}V} \right), \quad (24)$$

then equation (16) is *formally* equivalent to (17).

By virtue of Lemma 15 and Lemma 16, we can choose small but fixed δ, γ such that that $\mathcal{L}(\overline{B}_\gamma(0)) \subset \overline{B}_\gamma(0)$, where $B_\gamma(0)$ denotes the ball of radius γ centered at 0 of the Banach space $\mathcal{E}_{r, \hat{r}, \rho}$, and \mathcal{L} is contractive in this closed ball. Therefore, \mathcal{L} has a fixed point in $\overline{B}_\gamma(0)$ denoted by $V(x, y, \lambda, \varepsilon)$.

Note that using Proposition (14)(ii), the series V is analytic in (x, y, λ) and Gevrey-1 in ε .

The step from (17) to (3) is based on the Gevrey preparation Theorem, Theorem 12, to find a change of variable $x = X + X^2 A(X, \lambda, \varepsilon)$, where A is analytic in (X, λ) and Gevrey-1 in ε , such that

$$\varphi(X + X^2 A(X, \lambda, \varepsilon), \lambda, \varepsilon) = \sum_{k=0}^n p_k(\lambda, \varepsilon) X^k, \quad (25)$$

where $p_k(\lambda, \varepsilon) = \sum_{i=0}^{\infty} a_i^{(k)}(\lambda) \varepsilon^i$ is a Gevrey-1 formal series in ε and

$$a_0^{(0)}(\lambda_0) = \dots = a_0^{(n-1)}(\lambda_0) = 0, \quad a_0^{(n)}(\lambda_0) \neq 0.$$

Using this transformation, we find, from (17), the equation

$$\begin{cases} (1 + X^2 A_X + 2XA) \dot{X} &= Y - q(x, \lambda) - \sum_{k=0}^n p_k(\lambda, \varepsilon) X^k, \\ \dot{Y} &= \varepsilon G(X + X^2 A, \lambda, \varepsilon). \end{cases}$$

Using a time scaling it is equivalent to

$$\begin{cases} \dot{X} &= Y - \sum_{k=0}^n p_k(\lambda, \varepsilon) X^k, \\ \dot{Y} &= \varepsilon H(X, \lambda, \varepsilon), \end{cases}$$

where H is again analytic in (x, λ) and Gevrey-1 in ε . Dividing Y by $q_n(\lambda, \varepsilon)$, we are able to scale the highest coefficient of the polynomial $\sum_{k=0}^n p_k(\lambda, \varepsilon) X^k$ to be equal to 1. To conclude the proof, we need to remove the terms in the summation corresponding to $k = 0, k = n - 1$. However, this work can be done similarly as in the proof of Corollary 4.

3.3 Finishing the proof of Theorem 1

So far, we have proved that under *formal* Gevrey transformations, a slow fast system near a nilpotent contact point of order n can be written as in (3). According to the Theorem of Borel-Ritt-Gevrey, we are able to find Gevrey functions realizing these formal Gevrey transformation. Applying transformations induced by these Gevrey functions, (16) is transformed to a system of Liénard form plus an exponentially small remaining term. Note that the equation of fast variable is still a polynomial of order n . For the case of a jump-point, we can apply an analogous technique as in [2] to remove this

exponentially small remaining term. It is done by means of an extra change of coordinates in the x -variable, of the form $\tilde{X} = X + \Delta(X, Y, \lambda, \varepsilon)$ where Δ is exponentially small. The homological equation characterizes $D = \Delta$ as a graph of a center manifold of a singularly perturbed system in (X, Y, D) -variables, to which one can apply the center manifold theorem. For details we refer to [2]. This finishes the proof of Theorem 1.

Acknowledgements

The authors acknowledge support from the bilateral project between FWO Flanders (G0E7514N) and NAFOSTED Vietnam (FWO. 101.2013.02).

References

- [1] M. Canalis-Durand, J.P. Ramis, R. Schäfke, Y. Sibuya. Gevrey solutions of singularly perturbed differential equations. *J. reine angew. Math.* **588** (2000), 95–129.
- [2] P. De Maesschalck. Gevrey normal form for nilpotent contact points of order two. *Discrete and Continuous Dynamical Systems Series A* **34** (2014), no. 2, 677–688.
- [3] P. De Maesschalck, F. Dumortier and R. Roussarie. Cyclicity of common slow-fast cycles. *Indag. Math. (N.S.)* **22** (2011), no. 3-4, 165–206.
- [4] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *J. Differential Equations* **31** (1979), no. 1, 53–98.
- [5] A. Fruchard and R. Schäfke. *Composite Asymptotic Expansions*. Lecture Notes in Mathematics, **2066**. Springer, Heidelberg, 2013.
- [6] N. Levinson. A canonical form for an analytic function of several variables at a critical point. *Bull. Amer. Math. Soc.* **66** (1960), 68–69.
- [7] A.I. Neishtadt. The separation of motions in systems with rapidly rotating phase. *J. Appl. Math. Mech.* **48** (1985), no. 2, 133–139.
- [8] A. Parusinsky and J-P. Rolin. A note on the Weierstrass preparation theorem in quasianalytic local rings. *Canad. Math. Bull.* **57** (2014), no. 3, 614–620.

- [9] R. Schäfke and J.P. Ramis. Gevrey separation of fast and slow variables. *Nonlinearity* **9** (1996), no. 2, 353–384.
- [10] P. Szmolyan and M. Wechselberger. Canards in \mathbf{R}^3 . *J. Differential Equations* **177** (2001, no. 2, 419–453.