Advances in Mathematics 303 (2016) 888–953

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Normal forms and embeddings for power-log transseries ☆

P. Mardešić [a], M. Resman [b], J.-P. Rolin [a,*], V. Županović [b]

[a] *Université de Bourgogne, Départment de Mathématiques, Institut de Mathématiques de Bourgogne, B.P. 47 870, 21078-Dijon Cedex, France*
[b] *University of Zagreb, Department of Applied Mathematics, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia*

ARTICLE INFO

Article history:
Received 10 March 2016
Accepted 24 August 2016
Available online xxxx
Communicated by Vadim Kaloshin

MSC:
34C20
37C10
39B12
46A19
28A75
58K50
26A12

Keywords:
Iteration theory
Dulac map
Normal forms
Embedding in a flow
Transseries

ABSTRACT

Dulac series are asymptotic expansions of first return maps in a neighborhood of a hyperbolic polycycle. In this article, we consider two algebras of power-log transseries (generalized series) which extend the algebra of Dulac series. We give a formal normal form and prove a formal embedding theorem for transseries in these algebras.

☆ This research was supported by: Croatian Science Foundation (HRZZ) project No. 2285, French ANR project STAAVF 11-BS01-009, French–Croatian bilateral Cogito project *Classification de points fixes et de singularités à l'aide d'epsilon-voisinages d'orbites et de courbes*, and University of Zagreb research support for 2015 and 2016.
* Corresponding author.
E-mail address: rolin@u-bourgogne.fr (J.-P. Rolin).

http://dx.doi.org/10.1016/j.aim.2016.08.030
0001-8708/

Contents

1. Introduction and main results

1.1. Description of the problem and motivation

In the study of discrete dynamical systems, two problems are particularly important: the search of a *normal form* and the *embedding* problem. The search of a normal form means a definite process of choosing a representative of the class of conjugacy of the original system under topological, smooth or holomorphic conjugacies. This representative should be simpler than the original one. The embedding problem consists in finding a vector field such that the original system is the time-one map of its flow. These two problems are obviously connected: it should be easier to embed a normal form in a flow than the original system itself.

The study of holomorphic or real analytic systems at the origin of the ambient space naturally leads to the problems of *formal normal form* and *formal embedding*. These questions are discussed in detail, e.g., in [8, Chapter I, Sections 3 and 4]. The case of dimension 1 is well understood. Consider, for example, a *parabolic* formal series

$$f(z) = z + a_1 z^{p+1} + z^{p+1}\varepsilon(z),$$

where $a_1 \in \mathbb{C}^*$, $\varepsilon(z) \in \mathbb{C}[[z]]$ and $\varepsilon(0) = 0$. It is well known that the formal conjugacy class of f has a polynomial representative $f_0(z) = z + z^{p+1} + az^{2p+1}$, where the *residual* *index* $a \in \mathbb{C}$ is a formal invariant (see [9, Prop. 1.3.1], [1, Proposition 3.10] or [11]). The polynomial f_0 is classically called the *formal normal form* of f. Moreover, f_0 embeds formally in the flow of the vector field $X_{p,a} = \left(z^{p+1} + \left(a - \frac{p+1}{2}\right) z^{2p+1}\right) \frac{d}{dz}$, or in the formally equivalent flow of the vector field $X_{p,a} = \frac{z^{p+1}}{1+(\frac{p+1}{2}-a)z^p} \frac{d}{dz}$.

It turns out that there exist interesting maps in one real variable which do not admit asymptotic expansion in integer powers of the variable at the origin. Consider for example a hyperbolic monodromic polycycle Γ of an analytic planar vector field (see [7, Chapter 0] for precise definitions). It has been proved by Dulac in [4] that a semitransversal to Γ can be chosen in such a way that the corresponding *first return map* (or *Poincaré map*) admits at the origin an asymptotic expansion, called a *Dulac series* (see e.g. [7, Chapter 0] or [15, Chapter 3.3]). It is a formal series of the form:

$$D(x) = c_0 x^{\lambda_0} + \sum_{i=1}^{\infty} x^{\lambda_i} P_i(\log x), \quad c_0 > 0, \tag{1.1}$$

where $(\lambda_i)_{i\in\mathbb{N}_0}$ is an increasing sequence of strictly positive real numbers tending to infinity and each P_i, $i \in \mathbb{N}$, is a polynomial with real coefficients. Here, it is understood that $D(x)$ is an asymptotic expansion of $P(x)$ at 0 if, for any $N \in \mathbb{N}$, there exists $k_N \in \mathbb{N}$ such that $P(x) - \sum_{i=1}^{k_N} x^{\lambda_i} P_i(\log x) = o\left(x^N\right)$. The Dulac germs, as well as their asymptotic expansions, form a group for the composition (see [7, Chapter 0]). In particular, notice that the condition $c_0 > 0$ in (1.1) guarantees that the composition of two Dulac series is well defined and is also a Dulac series (no iterated logarithms are generated in the composition). It can be proved that the exponents λ_i, $i \in \mathbb{N}$, in (1.1) for the first return maps of hyperbolic monodromic polycycles belong to a finitely generated additive semigroup of $\mathbb{R}$. We will say, following the terminology of [3, Section 7], that such series is of *finite type*. By [3, Section 7], the *collection of Dulac series of finite type* form a subgroup of all Dulac series for the composition, and is denoted by $\mathfrak{D}$ in the present work.

These examples lead us to consider formal normal form and formal embedding problems for series with real coefficients in the monomials $x^\alpha (\log x)^k$, $\alpha > 0$, $k \in \mathbb{Z}$, considered as germs of functions at 0^+. More precisely, we are looking for classes of formal series which extend the collection $\mathfrak{D}$ and inside which both questions can be solved. It turns out that the set $\mathfrak{D}$ itself does not fit this purpose, mainly because Dulac series contain only *polynomials* in $\log x$, see Example 6.2 in Section 6 for explanation. Hence, we introduce two classes $\mathcal{L}_\mathfrak{D}$ and $\mathcal{L}$ of *generalized series* (or *transseries* if one follows the terminology introduced by Écalle in [6, Chapter 4]), with $\mathfrak{D} \subset \mathcal{L}_\mathfrak{D} \subset \mathcal{L}$, and we prove that both of them have the required properties. Compared to Dulac series (1.1), the elements of $\mathcal{L}_\mathfrak{D}$ and $\mathcal{L}$ are of *transfinite nature*: they involve not only polynomials in $\log x$, but also *infinite series* in $\log x$. We do not address here the question of *summability* of transseries on some domain, nor the meaning of *transseries asymptotic expansions* of germs in general. The question of the existence of an analytic function on some domain with an asymptotic expansion in the form of a given transseries is left for future research (possibly related to Écalle's *accelero-summability of transseries* [5]).

We denote the set of positive real numbers by $(0,\infty)$ or $\mathbb{R}_{>0}$.

The elements of the class $\mathcal{L}_{\mathfrak{D}}$ are the transseries

$$f(x) = \sum_{\alpha \in S} \sum_{k=N_\alpha}^{\infty} a_{\alpha,k} x^\alpha \left(-\frac{1}{\log x}\right)^k, \; a_{\alpha,k} \in \mathbb{R}, \; N_\alpha \in \mathbb{Z}, \tag{1.2}$$

where $S \subset \mathbb{R}_{>0}$, and the pairs (α, k) are contained in a sub-semigroup of the additive semigroup $\mathbb{R}_{>0} \times \mathbb{Z}$ generated by $\{(0,1)\}$ and by finitely many elements of $\mathbb{R}_{>0} \times \mathbb{Z}$. Obviously, $\mathfrak{D} \subseteq \mathcal{L}_{\mathfrak{D}}$. For short, we will say that the *support* of f (namely, the set $\mathcal{S}(f) = \{(\alpha,k) \in \mathbb{R}_{>0} \times \mathbb{Z}\colon a_{\alpha,k} \neq 0\}$) is of *finite type*, or equivalently, that f itself is of finite type. Notice that, while the germ x at the origin 0^+ of $\mathbb{R}_{>0}$ is a positive infinitesimal, the germ $\log x$ at 0^+ is negative and infinitely large. This is why we prefer to work with the germ $-1/\log x$ instead and to introduce the symbol

$$\boldsymbol{\ell} = -\frac{1}{\log x}.$$

We denote the elements of $\mathcal{L}_{\mathfrak{D}}$ as in (1.2) indifferently by f or by $f(x)$. We call a germ $x^\alpha \boldsymbol{\ell}^k$, $\alpha > 0$, $k \in \mathbb{Z}$, a *power-log monomial*. Finally, the series $f(x) = x$ will often be denoted simply by $f = \mathrm{id}$.

In order to define the class $\mathcal{L}$, let us recall that an ordered set X is called *well-ordered* if every non-empty subset of X has a smallest element. This property implies in particular that X is totally ordered. The elements of $\mathcal{L}$ are the formal transseries of the following form:

$$f(x) = \sum_{\alpha \in S} \sum_{k \in \mathbb{Z}} a_{\alpha,k}\, x^\alpha (-\frac{1}{\log x})^k = \sum_{\alpha \in S} \sum_{k \in \mathbb{Z}} a_{\alpha,k} x^\alpha \boldsymbol{\ell}^k, \; a_{\alpha,k} \in \mathbb{R}, \tag{1.3}$$

where $S \subset \mathbb{R}_{>0}$, and the support $\mathcal{S}(f) = \{(\alpha,k) \in \mathbb{R}_{>0} \times \mathbb{Z}\colon a_{\alpha,k} \neq 0\}$ is a well-ordered subset of $\mathbb{R}_{>0} \times \mathbb{Z}$, equipped with the *lexicographic order* $\preceq$. It is equivalent to assume:

- (1) $\mathcal{S}(f)$ is a well-ordered subset of $\mathbb{R}_{>0} \times \mathbb{Z}$;
- (2) S is a well-ordered subset of $\mathbb{R}_{>0}$ and, for every $\alpha \in S$, there exists $N_\alpha \in \mathbb{Z}$, such that a pair (α, k) belongs to $\mathcal{S}(f)$ only if $k \geq N_\alpha$.

In particular, one can easily check that $\mathcal{L}_{\mathfrak{D}} \subset \mathcal{L}$. Recall that a well-ordered subset of $\mathbb{R}$ is countable; hence, the supports of the elements of $\mathcal{L}$ are countable. The lexicographic order $\preceq$ on pairs of exponents corresponds to the usual order $\leq$ on germs of functions at the origin, in the following way: given two pairs (α, k) and (α', k') in $\mathbb{R}_{>0} \times \mathbb{Z}$, we have

$$(\alpha,k) \preceq (\alpha',k') \Longleftrightarrow \lim_{x \to 0^+} \frac{x^{\alpha'} \boldsymbol{\ell}^{k'}}{x^\alpha \boldsymbol{\ell}^k} < \infty \Longleftrightarrow x^{\alpha'} \boldsymbol{\ell}^{k'} \leq x^\alpha \boldsymbol{\ell}^k.$$

Similarly $(\alpha, k) \prec (\alpha', k')$ means that $(\alpha, k) \preceq (\alpha', k')$ and $(\alpha, k) \neq (\alpha', k')$. We call the pair (α, k) the *order* of the monomial $x^\alpha \boldsymbol{\ell}^k$. The *order of a transseries* $f \in \mathcal{L}$, denoted by $\operatorname{ord}(f)$, is the smallest element of $\mathcal{S}(f)$. If (α, k) is the order of f then $x^\alpha \boldsymbol{\ell}^k$ is called the *leading monomial* of f and is denoted by $\operatorname{Lm}(f)$, $a_{\alpha,k}$ is called the *leading coefficient* of f and is denoted by $\operatorname{Lc}(f)$, and $a_{\alpha,k} x^\alpha \boldsymbol{\ell}^k$ is called the *leading term* of f and is denoted by $\operatorname{Lt}(f)$.

Notation. We will sometimes denote by $[f]_{\alpha,k}$ the coefficient of the monomial $x^\alpha \boldsymbol{\ell}^k$ in the series $f \in \mathcal{L}$.

While the questions of formal embeddings and formal normal forms in $\mathcal{L}$ and $\mathcal{L}_{\mathfrak{D}}$ can be considered as natural problems of independent interest, our motivation for this research lies in fractal analysis of orbits of germs. Given an orbit $\mathcal{O}$ of a germ, by its fractal analysis we mean understanding the function $\varepsilon \mapsto A(\varepsilon)$, that assigns to each $\varepsilon > 0$ the Lebesgue measure of the ε-neighborhood of the orbit $\mathcal{O}$. The question that we pose is if we can recognize a germ by fractal properties of its realizations (orbits). Fractal properties of orbits of Poincaré maps around limit periodic sets have been studied in [10] and [19]. In the differentiable cases of elliptic points and limit cycles, it was proven in [19] that fractal analysis of orbits of Poincaré maps gives the multiplicity and the cyclicity. As already mentioned, in the nondifferentiable cases of hyperbolic polycycles, Poincaré maps have an expansion in $\mathcal{L}_{\mathfrak{D}}$. Furthermore, fractal analysis was performed on holomorphic complex germs in [13] and [14]. It was shown in [13] that the function $\varepsilon \mapsto A(\varepsilon)$, $\varepsilon > 0$, characterizes the formal class of a parabolic germ. The analytic class cannot be characterized by $A(\varepsilon)$, since it does not have an asymptotic expansion, as $\varepsilon \to 0$, see [14]. In a subsequent work, we plan to introduce a new definition of the *formal area* $A(\varepsilon)$, based on the formal embedding theorem proven in the present paper. With this new definition, we further hope for a sectorially analytic function which will reveal the analytic class of a germ. This would give a way to *see* the analytic class of a germ by *looking* at its orbits.

1.2. Overview of the results

Our main results (Theorems A and B) hold for a subclass of elements of $\mathcal{L}_{\mathfrak{D}}$ and $\mathcal{L}$. We say that an element f of $\mathcal{L}$ *contains no logarithms in the leading term* $\operatorname{Lt}(f)$ if f is of the form

$$(H) \qquad f(x) = \lambda x^\alpha + \sum_{(\alpha,0) \prec (\beta,k)} a_{\beta,k}\, x^\beta \boldsymbol{\ell}^k, \quad \lambda > 0,\ \alpha_{\beta,k} \in \mathbb{R}.$$

We denote by $\mathcal{L}^H$ the subset of transseries from $\mathcal{L}$ that satisfy (H), and by $\mathcal{L}_{\mathfrak{D}}^H$ the intersection $\mathcal{L}_{\mathfrak{D}} \cap \mathcal{L}^H$. There are two reasons for this additional assumption on the leading monomial. First, we have already mentioned that the Dulac series which are asymptotic expansions of Poincaré maps of hyperbolic polycycles belong to $\mathcal{L}_{\mathfrak{D}}^H$. Second, unlike $\mathcal{L}$, where iterated logarithms may be generated by compositions, the class $\mathcal{L}^H$ is a group for the composition of transseries.

The leading term in the asymptotic expansion at 0 of a germ indicates the rate of convergence of its orbits (or backward orbits) towards 0. According to the standard terminology used for holomorphic diffeomorphisms (see for example [1,11]), we distinguish three cases:

Definition 1.1. Let $f \in \mathcal{L}^H$, $f(x) = \lambda x^\alpha + \cdots$, $\lambda > 0$, $\alpha > 0$. We say that f is

1. *strongly hyperbolic*, if $\alpha \neq 1$,
2. *hyperbolic*, if $\alpha = 1$ and $\lambda \neq 1$,
3. *parabolic*, if $\alpha = 1$ and $\lambda = 1$.

Additionally, we say that a hyperbolic f is a *hyperbolic contraction* if $0 < \lambda < 1$. If $\lambda > 1$, we call f a *hyperbolic expansion*.

Intuitively, strongly hyperbolic cases $\alpha > 1$ correspond to strong contractions in the first term, and cases $\alpha < 1$ to strong expansions. Hyperbolic cases $0 < \lambda < 1$ correspond to exponential contractions, and cases $\lambda > 1$ to exponential expansions.

We denote by $\mathcal{L}^0 \subset \mathcal{L}^H$ the set of *formal changes of variables in* $\mathcal{L}$:

$$\mathcal{L}^0 = \{\varphi \in \mathcal{L}^H : \varphi(x) = ax + \text{h.o.t.},\ a > 0\}.$$

We use here the shortcut “h.o.t.” for “higher order terms”. Similarly, put $\mathcal{L}_{\mathfrak{D}}^0 = \mathcal{L}_{\mathfrak{D}} \cap \mathcal{L}^0$. Unlike $\mathcal{L}$, the classes $\mathcal{L}^H$, $\mathcal{L}_{\mathfrak{D}}^H$, $\mathcal{L}^0$ and $\mathcal{L}_{\mathfrak{D}}^0$ are closed under formal compositions of transseries and they are groups with respect to this operation. We say that $f, g \in \mathcal{L}^H$ (resp. $\mathcal{L}_{\mathfrak{D}}^H$) are *formally equivalent in* $\mathcal{L}^0$ (*resp.* $\mathcal{L}_{\mathfrak{D}}^0$) if there exists a change of variables $\varphi \in \mathcal{L}^0$ (*resp.* $\varphi \in \mathcal{L}_{\mathfrak{D}}^0$) transforming f to g, $g = \varphi^{-1} \circ f \circ \varphi$.

We now recall the definitions needed to state our formal embedding theorem. In the settings of usual power series, similar definitions may be found in, for example, [8] or [9].

Definition 1.2 *(The formal flow of a formal vector field).* Consider a family $(f^t)_{t\in\mathbb{R}}$ of elements of $\mathcal{L}^H$.

1. We say that (f^t) forms a *one-parameter group* (we also say for short: defines a *flow*) if $f^0 = \text{id}$ and $f^s \circ f^t = f^{s+t}$, for all $s, t \in \mathbb{R}$. An element $f \in \mathcal{L}^H$ *embeds* in the flow $(f^t)_{t\in\mathbb{R}}$ if $f = f^1$.
2. The family $(f^t)_{t\in\mathbb{R}}$ is called a $\mathcal{C}^1$-*one-parameter group* or a $\mathcal{C}^1$-*flow* if it defines a flow, and moreover:
   - (i) there exists a well-ordered subset S of $\mathbb{R}_{>0} \times \mathbb{Z}$ such that $\mathcal{S}(f^t) \subseteq S$ for every $t \in \mathbb{R}$, and
   - (ii) for every $(\alpha, m) \in S$, the function $t \mapsto [f^t]_{\alpha,m}$ is $\mathcal{C}^1(\mathbb{R})$.

3. Assume that (f^t) is a $\mathcal{C}^1$-flow and let $\xi := \frac{\mathrm{d}f^t}{\mathrm{d}t}|_{t=0} \in \mathcal{L}$. Then we say that (f^t) is the $\mathcal{C}^1$-*formal flow* of the vector field $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$. In that case, f^t is called the *formal t-map of* X, $t \in \mathbb{R}$.

Remark 1.3. The third point of the former definition means that, if we write

$$f^t(x) = \sum_{\alpha,k} \left[f^t\right]_{\alpha,k} x^\alpha \boldsymbol{\ell}^k, \quad \forall t \in \mathbb{R},$$

then

$$\xi(x) = \sum_{\alpha,k} \frac{\mathrm{d}\left[f^t\right]_{\alpha,m}}{\mathrm{d}t}\bigg|_{t=0} x^\alpha \boldsymbol{\ell}^k.$$

We show in Propositions 5.11, 5.12 and 5.13 in Section 5.2 that a vector field $X = \xi\frac{d}{dx}$, $\xi \in \mathcal{L}$, such that $(1,0) \preceq \operatorname{ord}(\xi)$ admits a *unique* $\mathcal{C}^1$-formal flow $(f^t)_t$ in $\mathcal{L}^H$, which is given by:

$$f^t = \exp(tX)\cdot \mathrm{id} = \mathrm{id} + t\xi + \frac{t^2}{2!}\xi'\xi + \frac{t^3}{3!}(\xi'\xi)'\xi + \cdots, \ t \in \mathbb{R}. \tag{1.4}$$

Remark 1.4. We prove the convergence of formula (1.4) in Propositions 5.11 and 5.12 in Section 5.2. We will actually need two notions of convergence. The first one is relevant of what we call the *formal topology*, see Section 4.2. To describe it roughly, let us say that the formal topology takes into account the orders of monomials, but not the size of coefficients. The series (1.4) converges in this topology when $(1,0) \prec \operatorname{ord}(\xi)$. Nevertheless, it does not converge when $\operatorname{ord}(\xi) = (1,0)$. Hence, we introduce a coarser *weak topology* (later: the product topology with respect to the Euclidean topology), in which the coefficients of the monomials play a role for the convergence of series. In this weak topology, the series (1.4) converges even when $\operatorname{ord}(\xi) = (1,0)$, see Proposition 5.11.

Finally, if $\operatorname{ord}(\xi) \prec (1,0)$, the series (1.4) does not converge in any of these topologies (Proposition 5.28).

We now state the two main theorems of this paper. The precise, but more technical, formulations are given in Sections 4 and 5.

Theorem A. *Let* $f \in \mathcal{L}^H$ (resp. $f \in \mathcal{L}^H_{\mathfrak{D}}$). *Then:*

1. *f is formally equivalent to a normal form $f_0 \in \mathcal{L}^H$* (resp. $f_0 \in \mathcal{L}^H_{\mathfrak{D}}$)*, given as a finite sum of power-log monomials.*
2. *If f is parabolic or hyperbolic, then f is formally equivalent to the formal time-one map $\widehat{f}_0 \in \mathcal{L}^H$* (resp. $\widehat{f}_0 \in \mathcal{L}^H_{\mathfrak{D}}$) *of a (formal) vector field $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$, where $\xi \in \mathcal{L}$ (resp. $\xi \in \mathcal{L}_{\mathfrak{D}}$) is a rational function in power-log monomials.*

The formal normal forms of Theorem A are described by *at most* 4 *scalars*. The actual number of scalars depends on the type (parabolic, hyperbolic or strongly hyperbolic) of the diffeomorphism.

The proof of Theorem A in Section 4 is actually based on a *transfinite algorithm* which transforms any transseries f in $\mathcal{L}^H$ or $\mathcal{L}^H_{\mathfrak{D}}$ into its finite formal normal form f_0.

Theorem B. *Let $f \in \mathcal{L}^H$. Then f embeds in a flow $(f^t)_{t\in\mathbb{R}}$, $f^t \in \mathcal{L}^H$. Moreover, if f is parabolic or hyperbolic, f embeds in the $\mathcal{C}^1$-flow of a unique vector field $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$, $\xi \in \mathcal{L}$ (see Definition 1.2).*

For the detailed statements discussing all cases (parabolic, hyperbolic, strongly hyperbolic) and their proofs, see Sections 4, 5 respectively.

2. Hahn fields and the structures of $\mathcal{L}$, $\mathcal{L}^H$ and $\mathcal{L}^0$ (*resp.* $\mathcal{L}_{\mathfrak{D}}$, $\mathcal{L}^H_{\mathfrak{D}}$ and $\mathcal{L}^0_{\mathfrak{D}}$)

Various descriptions of the notion of *transseries* have been given in several publications. The detailed study of classical operations, such as the operations of fields, as well as derivation, integration or composition, in this setting, is given in detail in [3]. The classes of transseries considered in the present work are proper sub-classes of the general field $\mathbb{R}\left(\left(x^{-1}\right)\right)^{\mathrm{LE}}$ of *logarithmic–exponential series* (or *LE-series*) introduced in [3]. Therefore, the operations we have to deal with are mostly specializations, in our framework, of the similar operations described there. In particular, the proof of the closure of $\mathcal{L}^H$ and $\mathcal{L}^0$ under composition can be checked by a careful, but straightforward, adaptation of the corresponding statement in $\mathbb{R}\left(\left(x^{-1}\right)\right)^{\mathrm{LE}}$.

Hence, we just provide in this section the vary basic notions needed to perform the description of our classes $\mathcal{L}$, $\mathcal{L}^H$ and $\mathcal{L}^0$. We use, as in [3], the language of *Hahn fields*. Recall that given a multiplicative ordered abelian group Γ with unit 1, the *Hahn field* $\mathbb{R}((\Gamma))$ consists of *generalized series* with real coefficients and monomials in Γ. The elements of $\mathbb{R}((\Gamma))$ are the formal sums

$$f = \sum_{\gamma\in\Gamma} f_\gamma \gamma,$$

with coefficients $f_\gamma \in \mathbb{R}$, such that $\mathrm{Supp}(f) = \{\gamma \in \Gamma \colon f_\gamma \neq 0\}$ is a *reverse well-ordered* subset of Γ. If $f \neq 0$ and γ_0 is the biggest element of $\mathrm{Supp}(f)$, then the *leading term* $\mathrm{Lt}(f)$ of f is $f_{\gamma_0}\gamma_0$, its *leading monomial* $\mathrm{Lm}(f)$ is γ_0 and its *leading coefficient* is f_{γ_0}.

One of the most useful tools when dealing with algebraic operations on Hahn fields is a result due to Neumann [12]. Its statement requires the following notations.

Notation 2.1. Consider an ordered (multiplicative) abelian group Γ and two subsets A and B of Γ. We denote:

a) $AB := \{ab \colon a \in A, b \in B\}$,
b) $\langle A \rangle$: the sub-semigroup of Γ generated by A (i.e. the smallest sub-semigroup of G containing all elements of A),
c) $\Gamma_{<\gamma_0} := \{\gamma \in \Gamma : \gamma < \gamma_0\}$, $\Gamma_{\leq\gamma_0} := \{\gamma \in \Gamma : \gamma \leq \gamma_0\}$, $\gamma_0 \in \Gamma$. Note that $\Gamma_{<1}$ denotes the *infinitesimals* of the group Γ.

Lemma 2.2 *(Neumann's Lemma). Consider an ordered (multiplicative) abelian group* Γ *and two reverse well-ordered subsets* A *and* B *of* Γ*. Then:*

1. *The product* AB *is reverse well-ordered.*
2. *For* $g \in AB$*, there are only finitely many pairs* $(a, b) \in A \times B$ *such that* $g = ab$*.*
3. *If* $A \subseteq \Gamma_{<1} = \{g \in \Gamma : g < 1\}$ *is reverse well-ordered, then* $\langle A \rangle$ *is also reverse well-ordered. Moreover, for each* $g \in \langle A \rangle$ *there are only finitely many tuples* $(a_1, \ldots, a_n)$ *with* $n \in \mathbb{N}$, $a_1, \ldots, a_n \in A$*, such that* $g = a_1 \cdots a_n$*.*

These series can be added and multiplied in the following way: if $f = \sum_{\gamma\in\Gamma} f_\gamma \gamma$ and $g = \sum_{\gamma\in\Gamma} g_\gamma \gamma$ belong to $\mathbb{R}((\Gamma))$, then

$$f + g = \sum_{\gamma\in\Gamma} (f_\gamma + g_\gamma)\, \gamma, \quad f \cdot g = \sum_{\gamma\in\Gamma} \left(\sum_{\lambda\mu=\gamma} f_\lambda g_\mu \right) \gamma.$$

Notice that the reverse well-ordering of the supports of f and g guarantees, thanks to Neumann's Lemma (see [12] or [3, p. 64] for example), that the product is well defined. Moreover, it is known that every nonzero element of $\mathbb{R}((\Gamma))$ admits a multiplicative inverse, so that $\mathbb{R}((\Gamma))$ is actually a field. If now Γ' is a sub-semigroup of Γ, then the set

$$\mathbb{R}[[\Gamma']] = \{f \in \mathbb{R}[[\Gamma]] : \operatorname{Supp}(f) \subseteq \Gamma'\}$$

is a subring (actually an $\mathbb{R}$-algebra) of $\mathbb{R}((\Gamma))$.

The LE-series introduced in [3] are generalized series in one variable whose monomials involve the logarithm and the exponential functions. Our classes $\mathcal{L}$ and $\mathcal{L}_{\mathfrak{D}}$ are contained in the field of LE-series (up to the obvious modification which comes from the fact that the variable x is thought as "infinitely big" there, while it is infinitesimal in our work). Let us show how $\mathcal{L}$ and $\mathcal{L}_{\mathfrak{D}}$ can be described by following the above Hahn's construction. Consider the multiplicative group G:

$$G = \{x^\alpha \boldsymbol{\ell}^k : \alpha \in \mathbb{R},\ k \in \mathbb{Z}\},$$

and the multiplicative sub-semigroup $G' = \left\{x^\alpha \boldsymbol{\ell}^k : \alpha \in (0, \infty), k \in \mathbb{Z}\right\}$ of G, equipped with the order $\leq$ introduced in Section 1.1. Then the class $\mathcal{L}$ is equal to the ring $\mathbb{R}[[G']]$. It is a subring of a Hahn field $\mathbb{R}((G))$, which is itself a subfield of the general field of LE-series. Notice that the support $\mathcal{S}(f)$ of a series $f \in \mathcal{L}$ is in Section 1.1 defined as a subset of $\mathbb{R}_{>0} \times \mathbb{Z}$. It differs from the support $\mathrm{Supp}\,(f)$ defined above for elements of general Hahn fields, which would be a set of monomials. The reason is that, in our situation, it is more convenient to work directly with exponents than to work with monomials. For the same reason, we will often refer to the additive version of Neumann's Lemma adapted to sets of exponents rather than to the multiplicative version stated above, which is adapted to sets of monomials.

Finally, as a straightforward consequence of Neumann's Lemma, $\mathcal{L}^H$ is an additive and multiplicative sub-semigroup of $\mathcal{L}$, and $\mathcal{L}^0$ is an additive sub-semigroup of $\mathcal{L}^H$. Furthermore, $\mathcal{L}^H_{\mathfrak{D}}$ is an additive and multiplicative sub-semigroup of $\mathcal{L}_{\mathfrak{D}}$, and $\mathcal{L}^0_{\mathfrak{D}}$ is an additive sub-semigroup of $\mathcal{L}^H_{\mathfrak{D}}$.

We consider now the operation of composition, as an imported operation from the general field of LE-series. As mentioned above, it is proved in [3] that the field of transseries can be equipped with a composition operator, and that each nonzero LE-series admits a compositional inverse. The proof of these facts requires an elaborate construction, which was previously sketched in [6, Chapter 4]. Fortunately, the action of the restriction of these two operators to our classes is much simpler, due to the particular shape of the monomials in G'. To be more precise, the composition of two series is understood by classical *Taylor expansions*. It is mainly based on the following observations, which are used in almost all subsequent computations of this paper. Every series $f \in \mathcal{L}^H$ can be written as:

$$f\,(x) = a x^{\lambda}\,(1 + \varepsilon\,(x)) \text{ with } \varepsilon \in \mathcal{L},\ \varepsilon\,(0) = 0 \text{ and } \lambda > 0.$$

For every real number $\alpha > 0$, the composition defined in [3] leads to:

$$(x^{\alpha} \circ f)\,(x) = (f\,(x))^{\alpha} = a^{\alpha} x^{\lambda\alpha} \sum_{j=0}^{\infty} \binom{\alpha}{j} \varepsilon\,(x)^{j}\,.$$

In the same way, if f is positive (that is, if $a > 0$), we have:

$$\log\,(f\,(x)) = \log a + \lambda \log\,(x) + \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} \varepsilon^{j}\,(x)\,.$$

The analysis made in [3] shows how these formulas extend to composition of series in the following way. If $g\,(x) = \sum_{(\alpha,k)} c_{\alpha,k} x^{\alpha} \boldsymbol{\ell}^{k} \in \mathcal{L}$ and $f \in \mathcal{L}^H$, then

$$(g \circ f)\,(x) = \sum_{(\alpha,k)} c_{\alpha,k} \big(f\,(x)\big)^{\alpha} \Big(\frac{-1}{\log f\,(x)}\Big)^{k}$$

is a well defined element of $\mathcal{L}$. As a consequence of the results proved in [3, Section 7], the similar conclusion holds for $\mathcal{L}_{\mathfrak{D}}$, $\mathcal{L}^H_{\mathfrak{D}}$ and $\mathcal{L}^0_{\mathfrak{D}}$ (the finite type property of the support is preserved). We summarize the former results in the following proposition.

Proposition 2.3 *(Properties of $\mathcal{L}$, $\mathcal{L}^H$ and $\mathcal{L}^0$).*

1. *$\mathcal{L}$ (*resp. *$\mathcal{L}_{\mathfrak{D}}$) is an $\mathbb{R}$-algebra without unity.*
2. *The algebra $\mathcal{L}$ (*resp. *$\mathcal{L}_{\mathfrak{D}}$) is closed under right compositions with elements from $\mathcal{L}^H$ (resp. $\mathcal{L}^H_{\mathfrak{D}}$).*
3. *The sets $\mathcal{L}^H$ and $\mathcal{L}^0$ (*resp. *$\mathcal{L}^H_{\mathfrak{D}}$ and $\mathcal{L}^0_{\mathfrak{D}}$) are groups under composition. In particular, they are closed under compositional inverses.*

The next section is dedicated to adaptations of standard *Lie bracket* techniques to our transseries setting. These techniques are used in the proofs in Sections 4 and 5.

3. Lie brackets in search of normal forms

The method of producing normal forms for analytic or formal diffeomorphisms is an adaptation of the *Lie bracket* technique for normal forms of vector fields, which is described for example in [2] or [17]. As we plan to adapt the method for elements of our algebra $\mathcal{L}$, we first recall briefly how it works in the classical case, more precisely, for formal power series in one variable.

Remark 3.1. In the sequel, the *h.o.t.*, meaning *higher-order terms*, stands for monomials of higher order than the last one written.

3.1. The effect of a change of variables on the elements of $\mathbb{R}[[x]]$

Consider a series $f \in \mathbb{R}[[x]]$ such that $f(0) = 0$. In order to transform f into its normal form, a classical approach consists in describing the effect on f of a change of variables $\varphi \in \mathbb{R}[[x]]$, such that $\varphi(x) = x + \text{h.o.t.}$ The simplest method consists in considering the leading term $\psi = \mathrm{Lt}(f \circ \varphi - \varphi \circ f)$. This leading term ψ is the same as the leading term of the difference $\varphi^{-1} \circ f \circ \varphi - f$. Using Taylor formula, we have:

$$\begin{aligned} f \circ \varphi &= \varphi \circ f + \psi \cdot (1 + \eta), \quad \eta \in x\mathbb{R}[[x]] \\ \varphi^{-1} \circ f \circ \varphi &= \varphi^{-1}(\varphi \circ f + \psi \cdot (1 + \eta)) \\ \varphi^{-1} \circ f \circ \varphi(x) &= f(x) + \left(\varphi^{-1}\right)'(\varphi \circ f(x)) \cdot \psi(x) + \text{h.o.t.} \\ &= f(x) + \psi(x) + \text{h.o.t.}, \end{aligned} \tag{3.1}$$

since $\varphi'(x) = 1 + \text{h.o.t.}$

Recall that the goal of formal normalization is to produce a series in the class of formal conjugacy of f which contains the smallest possible number of terms. So, given a term τ in the expansion of f, the main step consists in removing τ (if possible) via an appropriate change of variables φ. To do this, we choose the change of variables φ

such that the leading term $\mathrm{Lt}(f \circ \varphi - \varphi \circ f)$ is the opposite of τ. This procedure is then repeated term by term.

In particular, if f is *parabolic*, that is if $f(x) = x + \varepsilon(x) = x + ax^p + \text{h.o.t.}$, where $p > 1$, then we look for a change of variables $\varphi(x) = x + \eta(x) = x + cx^m$, $m > 1$, $c \in \mathbb{R}$. We obtain:

$$\begin{aligned}(f \circ \varphi - \varphi \circ f)(x) &= f(x + \eta(x)) - \varphi(x + \varepsilon(x)) \\ &= f(x) + f'(x)\eta(x) - \varphi(x) - \varphi'(x)\varepsilon(x) + \text{h.o.t.} \\ &= x + \varepsilon(x) + (1 + \varepsilon'(x))\eta(x) - x - \eta(x) - (1 + \eta'(x))\varepsilon(x) + \text{h.o.t.} \\ &= \varepsilon'(x)\eta(x) - \varepsilon(x)\eta'(x) + \text{h.o.t.} \qquad (3.2)\end{aligned}$$

The series $\eta\varepsilon' - \eta'\varepsilon$ is called the *Lie bracket (the commutator)* of η and ε and is denoted by $\{\eta, \varepsilon\}$. The leading term ψ of $f \circ \varphi - \varphi \circ f$ is given by the Lie bracket $\{cx^m, ax^p\}$ of the leading terms of η and ε.

3.2. Lie brackets in $\mathbb{R}[[x]]$ *and the homological equation*

The action of the Lie bracket of g is given by the following linear operator on $\mathbb{R}[[x]]$:

$$\mathrm{ad}_g(f) = [f, g], \ f \in \mathbb{R}[[x]]. \qquad (3.3)$$

Denote by H_k the vector space of monomials of degree k, $k \in \mathbb{N}$:

$$H_k = \left\{ax^k : a \in \mathbb{R}\right\}, \ k \geq 1.$$

The *grading* of the space H_k is given by the degree k of its monomials.

Let $f(x) = x + ax^p + \text{h.o.t.}$ be a parabolic element of $\mathbb{R}[[x]]$. It can be reduced to its formal normal form by solving a series of *Lie bracket* (*commutator*) *equations*, considering the action of the Lie bracket of the leading monomial of $f - \mathrm{id}$ to spaces H_l, $l \in \mathbb{N}$. The idea is to work step by step and, in each step, to eliminate the monomial of a given degree, if possible. Here we describe a single step.

Applying the change of variables $\varphi(x) = x + cx^m$, $c \in \mathbb{R}$, $m \in \mathbb{N}$, $m > 1$, we obtain, according to formula (3.2):

$$\begin{aligned}\varphi^{-1} \circ f \circ \varphi &= f + \mathrm{ad}_{ax^m}(cx^m) + \text{h.o.t.} \\ &= f + ac(p - m)x^{p+m-1} + \text{h.o.t.} \qquad (3.4)\end{aligned}$$

Since, $\mathrm{ad}_{x^m}(H_l) \subseteq H_{m+l-1}$, the action of the Lie bracket of any power *preserves the grading*. Moreover, for $m, l \in \mathbb{N}$, we have:

$$H_{m+l-1} = \mathrm{ad}_{x^m}(H_l) \oplus G_{m+l-1}, \quad G_{m+l-1} = \begin{cases}\emptyset, & m \neq l, \\ H_{2m-1}, & m = l.\end{cases}$$

Here, the spaces G_k, $k \in \mathbb{N}$, are subspaces of H_k that are not in the image of ad_{x^m} (consequently, these terms cannot be eliminated by changes of variables).

Consider now a term $\psi = bx^r$ of the expansion of f, with $r > p$. According to (3.4), in order to remove this term, we look for a change of variables $\varphi(x) = x + cx^m$ such that

$$ac(p-m)x^{p+m-1} = -bx^r. \tag{3.5}$$

This equation is classically called the *homological equation*. We find $m = r - p + 1$, so $p - m = 2p - r - 1$. It can be solved if and only if $r \neq 2p - 1$. So the term of degree $r = 2p - 1$ cannot be removed from the expansion of f. In other words, the subspace H_{2p-1} is not in the image of the Lie bracket action operator of the leading monomial of f. In order to remove all possible monomials of f, we proceed with a sequence of changes of variable of the previous type. The normal form appears to be a formal limit of a sequence of elements of $\mathbb{R}[[x]]$, and to have the form:

$$f_0(x) = x + ax^p + \beta x^{2p-1}, \quad \beta \in \mathbb{R}.$$

This procedure is an adaptation of a similar algorithm from [18] for reducing vector fields to their normal forms.

3.3. Lie brackets in $\mathcal{L}$

The general idea of the proof of Theorem A is to mimic the former method for the elements of algebra $\mathcal{L}$. However, because of the presence of logarithms in monomials, several complications are to be expected. Let us first explain the action of Lie brackets in our framework.

By $H_{\gamma,m} \subset \mathcal{L}$ we denote the one-dimensional vector spaces spanned by the monomial $x^\gamma \ell^m$, $\gamma \geq 1$, $m \in \mathbb{Z}$. We introduce the *grading of* $H_{\gamma,m}$ by the order (γ, m) of its monomials. Notice that, for $(\alpha, k) \in (0, \infty) \times \mathbb{Z}$, we have

$$\left(x^\alpha \boldsymbol{\ell}^k\right)' = \alpha x^{\alpha-1}\boldsymbol{\ell}^k + kx^{\alpha-1}\boldsymbol{\ell}^{k+1}.$$

Hence, the action of the Lie bracket, as defined in (3.3), of a monomial $x^\alpha \boldsymbol{\ell}^k$ on a space $H_{\beta,l}$ is given by:

$$\begin{aligned}\mathrm{ad}_{x^\alpha \boldsymbol{\ell}^k}(cx^\beta \boldsymbol{\ell}^l) &= \left[cx^\beta \boldsymbol{\ell}^l, x^\alpha \boldsymbol{\ell}^k\right] \\ &= c(\alpha-\beta)x^{\alpha+\beta-1}\boldsymbol{\ell}^{k+l} + c(l-k)x^{\alpha+\beta-1}\boldsymbol{\ell}^{k+l+1}.\end{aligned} \tag{3.6}$$

We conclude that, on spaces $H_{\gamma,m}$, the action of the Lie bracket of a power-log monomial does not preserve the *grading*, as in power series case. Therefore, we introduce the appropriate quotient spaces. By $K^0_{\gamma,m}$ and $K_{\gamma,m}$, we denote the direct sums:

$$K^0_{\gamma,m} = \bigoplus_{(\gamma,m)\preceq(\gamma',m')} H_{\gamma',m'}, \quad K_{\gamma,m} = \bigoplus_{(\gamma,m)\prec(\gamma',m')} H_{\gamma',m'}.$$

Recall that the order $\preceq$ (*resp.* $\prec$) is the lexicographic order (*resp.* strict lexicographic order) on $\mathbb{R}\times\mathbb{Z}$. We define the quotient spaces:

$$J_{\gamma,m} = \frac{K^0_{\gamma,m}}{K_{\gamma,m}}.$$

Note that the quotient space $J_{\gamma,m}$ can be identified with the vector space $H_{\gamma,m}$ of monomials of order (γ,m). The grading of $J_{\gamma,m}$ is given by the order (γ,m) of any representative. Based on these remarks, we can state the next proposition which claims that the grading is preserved on quotient spaces $J_{\gamma,m}$.

Proposition 3.2 *(Action of the Lie bracket operator on quotient spaces $J_{\gamma,m}$). Let*

$$T_{\alpha,k} \in J_{\alpha,k},\ (\alpha,k)\in\mathbb{R}_{>0}\times\mathbb{Z},\ (1,0)\prec(\alpha,k),$$

be an element of the class $J_{\alpha,k}$ of the monomial $x^\alpha \boldsymbol{\ell}^k$. Let $(\gamma,m)\in\mathbb{R}_{>0}\times\mathbb{Z}$, $(1,0)\prec(\gamma,m)$. The operator $\operatorname{ad}_{T_{\alpha,k}}$ acts on the quotient space $J_{\gamma,m}$ by the following rule:

$$\begin{cases} J_{\alpha+\gamma-1,k+m} = \operatorname{ad}_{T_{\alpha,k}}(J_{\gamma,m}), & \gamma\neq\alpha, \\ J_{2\alpha-1,k+m+1} = \operatorname{ad}_{T_{\alpha,k}}(J_{\alpha,m}) \oplus G_{2\alpha-1,k+l+1}, \end{cases} \tag{3.7}$$

where

$$G_{2\alpha-1,k+m+1} = \begin{cases} \emptyset, & m\neq k, \\ J_{2\alpha-1,2k+1}, & m=k. \end{cases}$$

Obviously, this different behavior of the action of the Lie bracket compared to its behavior in $\mathbb{R}[[x]]$ will induce a different treatment of the homological equation. These aspects are examined in details in the next section, where we give the precise form and the proof of Theorem A.

4. Proof of Theorem A

In this Section we construct changes of variables that transform a transseries from $\mathcal{L}^H$ or $\mathcal{L}^H_{\mathfrak{D}}$ to its formal normal form. These changes of variables will be obtained via *transfinite compositions* of elementary changes of variables. This important difference with the classical case comes from the fact that the supports of the elements of $\mathcal{L}$ are not any more contained in the set of positive integers, but are well-ordered subsets of $(0,\infty)\times\mathbb{Z}$. In order to define properly the notion of a transfinite composition, we recall (for a non-specialized reader) in the next section a few well known facts about well-ordered sets and about transfinite sequences.

4.1. Basic properties of ordinal numbers, well-ordered sets and transfinite sequences

A set is an *ordinal number* (or an *ordinal* for short) if it is transitive and well-ordered by $\in$. Recall that a set X is called *transitive* if every element of X is also a subset of X. Usually, the class of all ordinals is denoted by **On**. It is totally ordered (moreover, well-ordered) by the relation: $\alpha < \beta$ if and only if $\alpha \in \beta$. Recall the *von Neumann* construction of the class **On**. The empty set is the *smallest* ordinal, denoted by 0. Every ordinal α coincides with the set of all ordinals smaller than α, that is $\alpha = \{\beta \in \mathbf{On}\colon \beta < \alpha\}$.

There are two *types* of ordinals:

(1) *The successor ordinal*: The *successor* of an ordinal α, denoted by $\alpha+1$, is the ordinal $\alpha \cup \{\alpha\}$.
(2) *The limit ordinal*: If α is not a successor ordinal, then $\alpha = \sup\{\beta\colon \beta < \alpha\}$. Such α is called a *limit ordinal*.

The smallest limit ordinal is the set of non-negative integers, usually denoted by ω.

The classical principle of induction is generalized by the following principle, called the *principle of transfinite induction*. Consider a class C of ordinals, such that:

1. $0 \in C$;
2. if $\alpha \in C$ then $\alpha + 1 \in C$ (*non-limit case*);
3. if α is a nonzero limit ordinal and $\beta \in C$ for all $\beta < \alpha$, then $\alpha \in C$ (*limit case*).

Then C is the class **On** of *all* ordinals.

Consider now a set A. A *transfinite sequence* (or *θ-sequence*) of elements of A is a function that takes values in A and whose domain is an ordinal $\theta \in \mathbf{On}$. We denote such sequence by $(a_\beta)_{\beta<\theta}$, $a_\beta \in A$. Suppose that A is a topological space. We say that the θ-sequence $\{a_\beta\colon \beta < \theta\}$ of elements of A *converges to $a \in A$* when β goes to θ if, for every neighborhood U of a, there exists an ordinal $\beta_0 < \theta$ such that $a_\beta \in U$ for all β such that $\beta_0 < \beta < \theta$. We put $a := \lim_{\beta\to\theta} a_\beta$ or $a := \lim a_\beta$ for short.

Recall that two totally ordered sets $(P, <)$ and $(Q, <)$ are called *isomorphic* if there exists an order-preserving one-to-one function $f\colon P \to Q$. Finally, the strong connection between well-ordered sets and ordinals is established by the following result: *every well-ordered set is isomorphic to a unique ordinal number.* This ordinal number will be called *the order type* of the well-ordered set. It implies that the elements of a well-ordered set can be enumerated as an increasing θ-sequence (transfinite sequence). The ordinal θ is then its order type. Also conversely: the elements of a well-ordered set can be used as the indices of a transfinite sequence. In particular, given a well-ordered set W and a sequence $(a_w)_{w\in W}$ of elements of a topological space A, we say that (a_w) converges to $a \in A$ if, for every neighborhood U of a, there exists $w_0 \in W$ such that, for every $w \in W$, $w_0 < w$, it holds that $a_w \in U$. We denote this limit by $\lim_{w\in W} a_w$.

Notice that, due to the density of the set of rational numbers in $\mathbb{R}$, every well-ordered subset of $\mathbb{R}$ or of $\mathbb{R} \times \mathbb{Z}$ is countable.

In the sequel, we build transfinite sequences of elements of $\mathcal{L}$ algorithmically, and we study their convergence in $\mathcal{L}$.

4.2. Transfinite sequences of elements of $\mathcal{L}$

In order to study convergence of (transfinite) sequences of elements of $\mathcal{L}$, we endow $\mathcal{L}$ with the following topologies, introduced in order of the decreasing strength.

1. *The formal topology* on $\mathcal{L}$. Consider $f \in \mathcal{L}$ and $(\alpha, k) \in \mathbb{R}_{>0} \times \mathbb{Z}$. Then the (open) *ball* $B(f, (\alpha, k))$ centered at f is the set

$$\{g \in \mathcal{L} \colon \operatorname{ord}(g - f) \succ (\alpha, k)\}.$$

Given two different balls B_1 and B_2 centered at $f \in \mathcal{L}$, either $B_1 \subset B_2$ or $B_2 \subset B_1$. Hence, the collection of balls centered at f form a *fundamental system of neighborhoods*. The family of all balls generates a Hausdorff topology on $\mathcal{L}$.

Consider now an ordinal θ and a transfinite sequence $(f_\mu)_{\mu<\theta}$ of elements of $\mathcal{L}$. Then the sequence (f_μ) *converges* to $f \in \mathcal{L}$ when μ goes to θ in the formal topology if, for every $(\alpha, k) \in \mathbb{R}_{>0} \times \mathbb{Z}$, there exists an ordinal $\mu_0 < \theta$ such that $\operatorname{ord}(f - f_\mu) \succ (\alpha, k)$ for every $\mu_0 < \mu < \theta$.

2. *The product topology on $\mathcal{L}$ with respect to the discrete topology on $\mathbb{R}$.* Let us endow $\mathbb{R}$ with the discrete topology, and the product $\mathbb{R}^{\mathbb{R}_{>0} \times \mathbb{Z}}$ with the product topology. Each transseries $f \in \mathcal{L}$ is understood as a function $f : \mathbb{R}_{>0} \times \mathbb{Z} \to \mathbb{R}$, which assigns to each pair (α, k) the coefficient of the monomial $x^\alpha \boldsymbol{\ell}^k$ in f. We will denote that coefficient by $[f]_{\alpha,k}$. Hence, we can consider $\mathcal{L}$ as a subspace of $\mathbb{R}^{\mathbb{R}_{>0} \times \mathbb{Z}}$, equipped with the induced topology.

Let $(f_\mu)_{\mu<\theta}$ be a transfinite sequence of elements from $\mathcal{L}$. In this product topology, the sequence (f_μ) converges to $f \in \mathcal{L}$ when $\mu \to \theta$ if, for every $(\alpha, k) \in \mathbb{R}_{>0} \times \mathbb{Z}$, there exists an ordinal $\mu_0 < \theta$ such that the coefficient $[f_\mu]_{\alpha,k}$ equals the coefficient $[f]_{\alpha,k}$, for every $\mu_0 < \mu < \theta$.

3. *The weak topology on $\mathcal{L}$ (i.e. the product topology with respect to the Euclidean topology on $\mathbb{R}$).* The topology is similar to the one described in 2. The only difference is that we endow $\mathbb{R}$ with the Euclidean topology instead of the discrete one. The sequence (f_μ) converges to $f \in \mathcal{L}$ when $\mu \to \theta$ in the weak topology if, for every $(\alpha, k) \in \mathbb{R}_{>0} \times \mathbb{Z}$ and $\varepsilon > 0$, there exists an ordinal $\mu_0 < \theta$ such that $\left[f - f_\mu\right]_{\alpha,k} \in (-\varepsilon, \varepsilon)$, for every $\mu_0 < \mu < \theta$.

The three topologies introduced above will be used in this work. We need the *product topology with respect to the discrete topology* in the proof of Theorem A. In the proof of Theorem B, depending on the type of elements of $\mathcal{L}$ considered (parabolic or hyperbolic), we use *formal* or *weak topology*.

Remark 4.1. As has already been mentioned, the above topologies are ordered by their strength. For example, the sequence $(f_n)_{n\in\mathbb{N}} \in \mathcal{L}$,

$$f_n(x) = x^{2-\frac{1}{n}},$$

converges to $f \equiv 0$ in the product topology with respect to the discrete topology, but not in the formal topology.

Likewise, the sequence

$$f_n(x) = \frac{1}{n}x$$

converges to $f \equiv 0$ in the weak topology, but not in the product topology with respect to the discrete topology nor in the formal topology.

In all three cases, we set $f = \lim_{\mu\to\theta} f_\mu$, with an indication of the topology to which we refer. From now on, unless explicitly stated otherwise, we endow $\mathcal{L}$ with *the product topology with respect to the discrete topology*, so every limit or convergence to be mentioned in the sequel is understood with respect to this topology.

Remark 4.2. Given a well-ordered subset W of $\mathbb{R}_{>0} \times \mathbb{Z}$, we can define in the same way, if it exists, the limit $f = \lim_{(\alpha,k)\in W} f_{\alpha,k}$ of a transfinite sequence $(f_{\alpha,k})$ of elements of $\mathcal{L}$. In the rest of this article, we will deal indifferently with sequences indexed by ordinals or by elements of a well-ordered subset of $\mathbb{R}_{>0} \times \mathbb{Z}$.

We define the *elementary changes of variables* in $\mathcal{L}^0$ by:

$$\begin{cases} \varphi_{1,0}(x) = ax, & a \in \mathbb{R},\ a > 0,\ a \neq 1, \\ \varphi_{1,m}(x) = x + cx\boldsymbol{\ell}^m, & m \in \mathbb{N},\ m \neq 0,\ c \in \mathbb{R}, \\ \varphi_{\beta,m}(x) = x + cx^\beta\boldsymbol{\ell}^m, & \beta > 1,\ m \in \mathbb{Z},\ c \in \mathbb{R}. \end{cases} \tag{4.1}$$

Notice that $\mathrm{ord}(\varphi_{\beta,m} - \mathrm{id}) = (\beta, m)$.

We now define the notion of a composition of a transfinite sequence in $\mathcal{L}$. We will apply this notion to transfinite compositions of elementary changes of variables in $\mathcal{L}^0$.

Definition 4.3. Consider an ordinal θ and a transfinite sequence $(\varphi_\mu)_{\mu<\theta}$ of elements from $\mathcal{L}^0$. We say that the *transfinite composition* $\circ_{\mu<\theta}\,\varphi_\mu$ *exists* and is equal to $\varphi \in \mathcal{L}^0$ if the following conditions are satisfied:

1. We can define a sequence $(\psi_\mu)_{\mu<\theta}$ of elements of $\mathcal{L}^0$ (which we call the *partial compositions*) in the following way:
   (a) $\psi_0 := \mathrm{id}$;
   (b) If $\mu = \nu + 1$ is a *successor ordinal*, then $\psi_{\nu+1} := \varphi_\nu \circ \psi_\nu$ (*non-limit case*);

(c) If $\mu < \theta$ is a *limit ordinal*, the sequence $(\psi_\nu)_{\nu<\mu}$ converges to $\psi_\mu \in \mathcal{L}^0$ when ν goes to μ (*limit case*).

2. The sequence $(\psi_\mu)_{\mu<\theta}$ converges to $\varphi \in \mathcal{L}^0$ when μ goes to θ.

We write: $\varphi = \circ_{\mu<\theta}\,\varphi_\mu$.

Proposition 4.4 *(Convergence of* partial normal forms $(f_\mu)_{\mu<\theta} \in \mathcal{L}^H$*). Let $f \in \mathcal{L}^H$. Let $(\varphi_\mu)_{\mu<\theta}$, $\varphi_\mu \in \mathcal{L}^0$, be a transfinite sequence of changes of variables such that the composition $\psi = \circ_{\mu<\theta}\,\varphi_\mu$ exists in $\mathcal{L}^0$ (as the limit of the transfinite sequence $(\psi_\mu)_{\mu<\theta}$ introduced in the former definition). Let $(f_\mu)_{\mu<\theta}$ be a transfinite sequence in $\mathcal{L}^H$, defined by:*

$$f_\mu := \psi_\mu^{-1} \circ f \circ \psi_\mu,\ \mu \le \theta,\ \textit{with } \psi_\theta := \psi.$$

Then $f_\mu \to f_\theta$, as $\mu \to \theta$.

In the proof, we use the following auxiliary lemma. Consider a topology $\mathcal{T}$ on $\mathcal{L}$. We say that an application $F\colon \mathcal{L}^H \to \mathcal{L}^H$ is *transfinitely sequentially continuous with respect to* $\mathcal{T}$ if, for every transfinite sequence $(g_\mu)_{\mu<\theta}$ in $\mathcal{L}^H$ *such that the supports of all the* g_μ *are contained in a well-ordered subset of* $\mathbb{R}_{>0} \times \mathbb{Z}$, and such that $g_\mu \to g$ with respect to $\mathcal{T}$, then $F(g_\mu) \to F(g)$ with respect to $\mathcal{T}$.

Lemma 4.5 *(Transfinite sequential continuity). Assume $\mathcal{L}$ equipped with the product topology (the discrete case).*

1. *Let $h \in \mathcal{L}^H$. The applications defined on $\mathcal{L}^0$ by*

$$(i)\ g \longmapsto g \circ h, g \longmapsto h \circ g, \quad (ii)\ g \longmapsto g^{-1}$$

are transfinitely sequentially continuous.

2. *Consider two transfinite sequences $(h_\mu)_{\mu<\theta}$ in $\mathcal{L}^H$, and $(g_\mu)_{\mu<\theta}$ in $\mathcal{L}^0$, such that $h_\mu \to 0$ as $\mu \to \theta$ and the supports of all the h_μ and g_μ are contained in a common well-ordered subset of $\mathbb{R}_{>0} \times \mathbb{Z}$. Then $h_\mu \circ g_\mu \to 0$ as $\mu \to \theta$.*

Proof. All these statements can be proven by analyzing the supports of the composition and of the inverse as in (4.2) and applying Neumann's Lemma 2.2.3. Concluding similarly as in (4.2), for two transseries $g, h \in \mathcal{L}^H$, such that $\mathrm{ord}(g) = (\alpha_0, 0)$, we obtain:

$$\mathcal{S}(h \circ g) \subset \mathcal{S}(h) \cup H.$$

Here, H is a sub-semigroup of $\mathbb{R}_{\ge 0} \times \mathbb{Z}$ generated by $(\alpha_0\beta, k)$ for $(\beta, k) \in \mathcal{S}(h)$, $(\alpha - \alpha_0, m)$ for $(\alpha, m) \in \mathcal{S}(g)$ and $(0, 1)$. Moreover, every coefficient of the composition is a sum of *finitely many* finite products of coefficients of h and g by Neumann's lemma. Additionally, each product contains exactly one coefficient from h among its factors. (1) (i) and (2) follow.

To prove (1) (ii), due to (1), it suffices to prove the easier statement: if $g_\mu \to \mathrm{id}$ as $\mu \to \theta$, then $g_\mu^{-1} \to \mathrm{id}$, as $\mu \to \theta$. It can be checked that the coefficients of $g_\mu^{-1} - \mathrm{id}$ are sums of finitely many finite products of coefficients of $g_\mu - \mathrm{id}$, which eventually vanish, by Neumann's lemma. Therefore, the coefficients of $g_\mu^{-1} - \mathrm{id}$ eventually vanish. $\square$

Proof of Proposition 4.4. Let $f, (f_\mu)_{\mu<\theta} \in \mathcal{L}^H$ be as defined in the proposition. Knowing that $\psi_\mu \to \psi$, we prove that $\psi_\mu^{-1} \circ f \circ \psi_\mu \to \psi^{-1} \circ f \circ \psi$, as $\mu \to \theta$ (in the product topology, the discrete case).

Since $\psi_\mu \to \psi$, it follows by Lemma 4.5 (1)(i) that $\psi_\mu \circ \psi^{-1} \to \mathrm{id}$, and further by (1)$(ii)$ that $\psi \circ \psi_\mu^{-1} \to \mathrm{id}$. By (1)$(i)$, $f \circ \psi \circ \psi_\mu^{-1} \to f \circ \mathrm{id} = f$ and $\psi \circ \psi_\mu^{-1} \circ f \to \mathrm{id} \circ f = f$. Therefore,

$$\begin{aligned} f \circ \psi \circ \psi_\mu^{-1} - \psi \circ \psi_\mu^{-1} \circ f &\to 0, \\ \overset{(2),\, \psi_\mu \to \psi}{\Longrightarrow} \quad f \circ \psi - \psi \circ \psi_\mu^{-1} \circ f \circ \psi_\mu &\to 0, \\ \psi \circ \psi_\mu^{-1} \circ f \circ \psi_\mu &\to f \circ \psi, \\ \psi_\mu^{-1} \circ f \circ \psi_\mu &\to \psi^{-1} \circ f \circ \psi. \qquad \square \end{aligned}$$

As we did above for a composition of a transfinite sequence of elements of $\mathcal{L}^0$, we can define in particular, if it exists, a composition of a sequence $(\varphi_{\beta,m})$ of elementary changes of variables indexed by elements of a well-ordered subset $W \subset \mathbb{R}_{>0} \times \mathbb{Z}$, via the sequence $(\psi_{\beta,m})$ of partial compositions. Again, we have to consider non-limit cases and limit cases.

The following proposition gives an important characterization of elements of $\mathcal{L}^0$ or $\mathcal{L}^0_{\mathfrak{D}}$ in terms of transfinite compositions of elementary changes of variables.

Proposition 4.6 *(Characterization of changes of variables in $\mathcal{L}^0$ or $\mathcal{L}^0_{\mathfrak{D}}$). Let $\mathcal{L}^0$ be endowed with the product topology with respect to the discrete topology.*

1. *Let $W \subset \mathbb{R}_{>0} \times \mathbb{Z}$ be well-ordered. Let $(\varphi_{\alpha,m})_{(\alpha,m)\in W}$ be a transfinite sequence of elementary changes of variables, such that the sequence of orders* $\mathrm{ord}(\varphi_{\alpha,m} - \mathrm{id}) = (\alpha, m)$ *is strictly increasing. Then the transfinite composition $\varphi = \circ_{(\alpha,m)\in W} \varphi_{\alpha,m}$ is well defined in $\mathcal{L}^0$. Moreover, if $W \subset \mathbb{R}_{>0} \times \mathbb{Z}$ is of finite type, then $\varphi \in \mathcal{L}^0_{\mathfrak{D}}$.*
2. *For every transseries $\varphi \in \mathcal{L}^0$ (*resp. *$\varphi \in \mathcal{L}^0_{\mathfrak{D}}$) there exist a well-ordered subset (*resp. *a subset of finite type) $W \subset \mathbb{R}_{>0} \times \mathbb{Z}$ and a transfinite sequence $(\varphi_{\alpha,m})_{(\alpha,m)\in W}$ of elementary changes of variables such that $\varphi = \circ_{(\alpha,m)\in W} \varphi_{\alpha,m}$.*

Proof. (1) We first give a preliminary computation which describes the change in the support of an element of $\mathcal{L}^0$ after composition with an elementary change of variables, and prove implicitly that the composition remains in $\mathcal{L}^0$. Let $h = \mathrm{id} + \varepsilon \in \mathcal{L}^0$ (i.e.

$\operatorname{ord}(\varepsilon) \succ (1,0)$). Consider an elementary change of variables $\varphi_{\beta,\ell}(x) = x + cx^{\beta}\boldsymbol{\ell}^{\ell}$, $c \in \mathbb{R}$, $(\beta,\ell) \succ (1,0)$. A straightforward computation shows that, for every integer $p \geq 1$, the support of the p-th derivative $\left(x^{\beta}\boldsymbol{\ell}^{\ell}\right)^{(p)}$ is contained in the set $\{(\beta-p,\ell),(\beta-p,\ell+1),\ldots,(\beta-p,\ell+p)\}$. Hence, it follows from Taylor formula that

$$\begin{aligned}(\varphi_{\beta,\ell}\circ h)(x) &= \left(x + cx^{\beta}\boldsymbol{\ell}^{\ell}\right)\circ(x+\varepsilon(x)) \\ &= h(x) + cx^{\beta}\boldsymbol{\ell}^{\ell} + \sum_{p=1}^{\infty}\sum_{j_p=0}^{p} b_{p,j_p}x^{\beta-p}\boldsymbol{\ell}^{\ell+j_p}\varepsilon(x)^p, \quad b_{p,j_p}\in\mathbb{R}. \qquad (4.2)\end{aligned}$$

Notice that, for each $p \geq 1$, every element (γ,r) of the support of $\varepsilon(x)^p$ has the form

$$(\gamma,r) = \left(\alpha_{i_1}+\cdots+\alpha_{i_p}, k_{i_1}+\cdots+k_{i_p}\right),$$

where the exponents (α_{i_s},k_{i_s}), $s=1,\ldots,p$, belong to $\mathcal{S}(\varepsilon)$. Hence, every element of the support of the double sum in the formula (4.2) has the form

$$(\beta,\ell) + \left(\alpha_{i_1}-1+\cdots+\alpha_{i_p}-1, k_{i_1}+\cdots+k_{i_p}\right) + \left(0, j_{i_p}\right), \qquad (4.3)$$

where $p \geq 1$, $(\alpha_{i_s},k_{i_s}) \in \mathcal{S}(h)$ and $j_{i_p} \in \{0,\ldots,p\}$.

Two main facts follow from this computation. Denote by H the (additive) sub-semigroup of $\mathbb{R}_{\geq 0}\times\mathbb{Z}$ generated by (β,ℓ), the elements $(\alpha-1,k)$ for $(\alpha,k)\in\mathcal{S}(h)$ and $(0,1)$. First, by (4.2), the support $\mathcal{S}(\varphi_{\beta,\ell}\circ h)$ of the composition is contained in the union $\mathcal{S}(h)\cup H$. Since $\mathcal{S}(h)$ is well-ordered, and since, by Neumann's Lemma 2.2, H is well-ordered, the support $\mathcal{S}(\varphi_{\beta,\ell}\circ h)$ is also well-ordered. Second, by Neumann's Lemma 2.2, the composition $\varphi_{\beta,\ell}\circ h$ is well-defined, meaning that every monomial in the support of $\varphi_{\beta,\ell}\circ h$ has a well-defined coefficient. This in particular implies that $\varphi_{\beta,\ell}\circ h \in \mathcal{L}^0$. Now, assume additionally that $h\in\mathcal{L}^0_{\mathfrak{D}}$, so $\mathcal{S}(h)$ is contained in a (additive) sub-semigroup of $\mathbb{R}_{>0}\times\mathbb{Z}$ generated by finitely many elements $(\gamma_1,p_1),\ldots,(\gamma_r,p_r)$ of $\mathbb{R}_{>0}\times\mathbb{Z}$. For $n\in\mathbb{N}$, we can write $(n\gamma_i-1,np_i)=(n-1)(\gamma_i,p_i)+(\gamma_i-1,p_i)$. Hence, $\mathcal{S}(\varphi_{\beta,\ell}\circ h)$ is contained in the (additive) sub-semigroup of $\mathbb{R}_{\geq 0}\times\mathbb{Z}$ generated by (β,ℓ), $(0,1)$ and the elements (γ_i,p_i), (γ_i-1,p_i) for $i=1,\ldots,r$. In particular, $\varphi_{\beta,\ell}\circ h\in\mathcal{L}^0_{\mathfrak{D}}$.

Consider now the sequence $(\varphi_{\alpha,m})_{(\alpha,m)\in W}$ given in the statement of the proposition. We prove the existence of the composition $\circ_{(\alpha,m)\in W}\varphi_{\alpha,m}$ by transfinite induction. Let (α_0,m_0) be the smallest element of W. Put $\psi_{\alpha_0,m_0} := \varphi_{(\alpha_0,m_0)}\in\mathcal{L}^0$. Consider the sub-semigroup $\overline{W}\subset\mathbb{R}_{\geq 0}\times\mathbb{Z}$ generated by the elements of W, the elements $(\alpha-1,p)$ for $(\alpha,p)\in W$, and $(0,1)$. We already know that $\overline{W}$ is well-ordered, and of finite type if W is of finite type.

Existence of partial compositions in $\mathcal{L}^0$ in the *non-limit case* follows directly by the above considerations. Moreover, the support of the partial compositions is contained in $\overline{W}$. Consider the *limit case*. Suppose (α_θ,m_θ) is a limit ordinal (or the order type of W), and for every $(\beta,l)\in W$, $(\beta,l)\prec(\alpha_\theta,m_\theta)$, it holds that $\psi_{\beta,l}\in\mathcal{L}^0$

and $\mathcal{S}(\psi_{\beta,l}) \subset \overline{W}$. We prove that $\psi_{\beta,l}$ converge in $\mathcal{L}^0$ in the product topology, as $(\beta,l) \to (\alpha_\theta, m_\theta)$, and that the support of the limit belongs to $\overline{W}$.

By (4.2), $\mathcal{S}(\psi) \subset \mathcal{S}(\varphi_{\beta,l} \circ \psi) \subseteq \overline{W}$ for every partial sum ψ and every change of variables $\varphi_{\beta,l}$, $(\beta,l) \in W$. Thus, if $(\gamma,k) \in \mathcal{S}(\psi_{\beta,l})$, then $(\gamma,k) \in \mathcal{S}(\psi_{\alpha,m})$, for all $(\beta,l) \prec (\alpha,m) \prec (\alpha_\theta,m_\theta)$. To prove the convergence in the product topology, we have to prove that the coefficient of monomial $x^\gamma \ell^k$ eventually stabilizes in the sequence of partial sums $(\psi_{\beta,l})_{(\beta,l)\prec(\alpha_\theta,m_\theta)}$. Since (β,l) is a summand of (4.3) and the sequence $(\beta,l) \in W$ is strictly increasing, it follows from Neumann's Lemma 2.2.3. that each $(\gamma,k) \in \overline{W}$ is realized at most finitely many times as sum of the type (4.3). That is, the coefficient of monomial $x^\gamma \ell^k$ in the support of $(\psi_{\beta,l})$ changes only *finitely many* times in the course of compositions $\circ_{(\beta,l)\prec(\alpha_\theta,k_\theta)}\varphi_{\beta,l}$. This guarantees the convergence in $\mathcal{L}^0$ of partial compositions in the limit case, for the product topology. The limit is the partial composition for the limit ordinal $(\alpha_\theta, m_\theta)$:

$$\psi_{\alpha_\theta,m_\theta} = \circ_{(\beta,l)\prec(\alpha_\theta,m_\theta)}\,\varphi_{\beta,l} := \lim_{(\beta,l)\to(\alpha_\theta,m_\theta)} \psi_{\beta,l},$$

with $\mathcal{S}(\psi_{\alpha_\theta,m_\theta}) \subseteq \overline{W}$ by construction.

(2) Let $\varphi \in \mathcal{L}^0$, $\varphi(x) = ax + \text{h.o.t.}$, $a \in \mathbb{R}$. Obviously, $\varphi(x) = ax \circ \varphi_0(x)$, where $\varphi_0(x) = x + ax^{\alpha_0}\boldsymbol{\ell}^{m_0} + \text{h.o.t.}$ tangent to the identity. By Neumann's Lemma 2.2, the sub-semigroup W of $\mathbb{R}_{>0} \times \mathbb{Z}$ generated by $\mathcal{S}(\varphi - a\cdot\text{id})$ is well-ordered, with (α_0,m_0) its smallest element. We prove that φ_0 can be decomposed in a transfinite composition of elementary changes of variables $(\varphi_{\beta,m})_{(\beta,m)\in V}$, for some $V \subseteq W$. More precisely: we build, by transfinite induction, a sequence $(\varphi_{\beta,m})_{(\beta,m)\in V}$, $V \subseteq W$, of elementary changes of variables, such that, for every $(\alpha,k) \in W$, there exists $(\beta(\alpha,k), m(\alpha,k)) \in V$ and a partial composition

$$\psi_{\beta(\alpha,k),m(\alpha,k)} = \circ_{(\beta,m)\prec(\beta(\alpha,k),m(\alpha,k))}\varphi_{(\beta,m)}$$

with $\operatorname{ord}\left(\varphi - \psi_{\beta(\alpha,k),m(\alpha,k)}\right) \succ (\alpha,k)$. Moreover, the function $(\alpha,m) \in W \mapsto (\beta(\alpha,k), m(\alpha,k)) \in V$ is increasing, but not necessarily strictly. Since W contains arbitrarily big elements with respect to the order topology, it means that the sequence $(\psi_{\beta,m})_{(\beta,m)\in V}$ converges towards φ in the formal topology. In particular, it converges towards φ in the product topology with respect to the discrete topology.

Put $\psi_{\alpha_0,m_0}(x) := \varphi_{\alpha_0,m_0}(x) = x + ax^{\alpha_0}\ell^{m_0}$. Consider $(\alpha,k) \in W$. By the induction hypothesis, for all $(\gamma,r) \in W$, $(\gamma,r) \prec (\alpha,k)$, there exists a transfinite composition $\psi_{\beta(\gamma,r),m(\gamma,r)} \in \mathcal{L}^0$, such that $(\gamma,r) \prec \operatorname{ord}(\varphi - \psi_{\beta(\gamma,r),m(\gamma,r)})$. We prove that then there exists a transfinite composition $\psi_{\beta(\alpha,k),m(\alpha,k)}$, such that $(\alpha,k) \prec \operatorname{ord}(\varphi - \psi_{\beta(\alpha,k),m(\alpha,k)})$.

In the *non-limit case*, consider the predecessor (α',k') of (α,k) in W, and the partial composition $\psi_{\beta(\alpha',k'),m(\alpha',k')} \in \mathcal{L}^0$. Then either $\psi_{\beta(\alpha',k'),m(\alpha',k')} = \psi_{\beta(\alpha,k),m(\alpha,k)}$, or there exists an elementary change of variables $\varphi_{\alpha,k}(x) = x + cx^\alpha\boldsymbol{\ell}^k$, with $c \in \mathbb{R}$ such that the term of order (α,k) from $\varphi - \psi_{\beta(\alpha',k'),m(\alpha',k')}$ is cancelled in $\varphi - \psi_{\beta(\alpha,k),m(\alpha,k)}$:

$$\begin{aligned}\varphi - \psi_{\beta(\alpha,k),m(\alpha,k)} &= \varphi - \varphi_{\alpha,k} \circ \psi_{\beta(\alpha',k'),m(\alpha',k')} \\ &= \left(\varphi - \psi_{\beta(\alpha',k'),m(\alpha',k')}\right) - cx^{\alpha}\boldsymbol{\ell}^{k} + \text{h.o.t.}\end{aligned}$$

Obviously, $\psi_{\beta(\alpha,k),m(\alpha,k)} = \varphi_{\alpha,k} \circ \psi_{\beta(\alpha',k'),m(\alpha',k')} \in \mathcal{L}^0$. Hence, the claim is proved in the non-limit case.

In the *limit case*, when (α, k) is a limit ordinal, we put $\psi_{\beta(\alpha,k),m(\alpha,k)} = \lim_{(\gamma,r)\prec(\alpha,k)} \psi_{\beta(\gamma,r),m(\gamma,r)}$, as in Definition 4.3. By (1), this limit exists and belongs to $\mathcal{L}^0$.

We conclude the proof by noticing that, if $\varphi \in \mathcal{L}^0_{\mathfrak{D}}$, that is, if $\mathcal{S}(\varphi)$ is of finite type, so is the set W. □

Propositions 4.4 and 4.6 will be used in the proof of Theorem A to derive the formal normal forms of $f \in \mathcal{L}^H$ by *transfinite induction*: eliminating terms from f, step by step, by elementary changes of variables, when possible.

By Definition 4.3, the composition of a transfinite sequence $(f_\mu)_{\mu<\theta} \in \mathcal{L}^0$ *exists*, if the transsequence of partial compositions $(\psi_\mu)_{\mu<\nu}$ at any limit ordinal $\nu \le \theta$ converges in the product topology in $\mathcal{L}^0$. That is if, for every $(\beta, l) \in \mathbb{R}_{>0} \times \mathbb{Z}$, there exists an index $\mu_{\beta,l}$ such that, for $\mu_{\beta,l} < \mu < \nu$, the coefficient $[\psi_\mu]_{\beta,l}$ remains constant. In the proof of Proposition 4.6, for transfinite compositions of *elementary changes of variables* we have proved (by Neumann's lemma) even more: for every $(\beta, l) \in \mathbb{R}_{>0} \times \mathbb{Z}$, the coefficient $[\psi_\mu]_{\beta,l}$ changes in the sequence of partial compositions $(\psi_\mu)_{\mu<\nu}$ at most at *finitely many* indices.

4.3. The precise form of Theorem A

We now give the precise statement of Theorem A, which was given with less details on page 894.

Theorem A *(Formal normal forms). Let $f \in \mathcal{L}^H$ (*resp. *$f \in \mathcal{L}^H_{\mathfrak{D}}$).*

1. *f is formally equivalent in $\mathcal{L}^0$ (resp. in $\mathcal{L}^0_{\mathfrak{D}}$) to the finite normal form $f_0 \in \mathcal{L}^H$ (actually in $\mathcal{L}^H_{\mathfrak{D}}$):*
   (a) (Parabolic case)

$$f(x) = x + ax^{\alpha}\boldsymbol{\ell}^{k} + \text{h.o.t.},\ \ \alpha \ge 1,\ \ k \in \mathbb{Z},\ \ (\alpha, k) \succ (1, 0);\ \ a \in \mathbb{R},\ \ a \ne 0;$$
$$f_0(x) = x + ax^{\alpha}\boldsymbol{\ell}^{k} + bx^{2\alpha-1}\boldsymbol{\ell}^{2k+1},\ \ b \in \mathbb{R}.$$

   (b) (Hyperbolic case)

$$f(x) = \lambda x + ax\boldsymbol{\ell} + \text{h.o.t.},\ \ a \in \mathbb{R},\ \ \lambda > 0,\ \ \lambda \ne 1;$$
$$f_0(x) = \lambda x + ax\boldsymbol{\ell}.$$

(c) (Strongly hyperbolic case)

$$f(x)=\lambda x^{\alpha}+\text{h.o.t.},\ \lambda>0,\ \alpha\neq 1;\qquad f_0(x)=x^{\alpha}.$$

2. *Let f be hyperbolic or parabolic. Then f is formally equivalent in $\mathcal{L}^0$ (*resp. $\in\mathcal{L}^0_{\mathfrak{D}}$*) to $\widehat{f}_0\in\mathcal{L}^H$ (*resp. *in* $\mathcal{L}^H_{\mathfrak{D}}$*), given as the formal time-one map of the following vector fields:*
   (a) (Parabolic case)

$$f_0(x)=\exp(X_{\alpha,k,a,b}).\mathrm{id},$$
$$X_{\alpha,k,a,b}=\frac{ax^{\alpha}\boldsymbol{\ell}^k}{1+\frac{a\alpha}{2}x^{\alpha-1}\boldsymbol{\ell}^k-(\frac{ak}{2}+\frac{b}{a})x^{\alpha-1}\boldsymbol{\ell}^{k+1}}\frac{\mathrm{d}}{\mathrm{d}x}.$$

   (b) (Hyperbolic case)

$$\widehat{f}_0(x)=\exp(X_{\lambda,a}).\mathrm{id},\quad X_{\lambda,a}=\frac{\log\lambda\cdot x}{1+\frac{a}{2(\lambda-1)}\boldsymbol{\ell}}\frac{\mathrm{d}}{\mathrm{d}x}.$$

In the parabolic case, the formal normal forms are described by the quadruples:

$$(\alpha,k,a,b);\ \alpha\geq 1,\ k\in\mathbb{Z},\ b\in\mathbb{R},\ a\neq 0.$$

Additionally, if $\alpha>1$, a can be replaced by $\mathrm{sgn}(a)$*, up to a linear change.*

It is worth recalling that in the *hyperbolic case*, the series $\exp(X_{\lambda,a})\cdot\mathrm{id}$ does not converge in $\mathcal{L}$ neither for the formal topology nor for the product topology with respect to the discrete topology. Nevertheless, it converges for the weak topology (the product topology with respect to the Euclidean topology), which takes into account not only the supports, but also the size of their coefficients. For details, see the proof of Proposition 5.11.

Note that the formal normal form of a *strongly hyperbolic* transseries cannot be expressed as the formal time-one map of a vector field in $\mathcal{L}$. The exponential of a parabolic vector field does not converge in $\mathcal{L}$ in any of the three topologies that we mentioned on page 904. The formula (1.4) for the formal-time map of a parabolic field does not make sense in $\mathcal{L}$. The detailed description of this phenomenon is given in Proposition 5.28.

Furthermore, notice that if $f\in\mathbb{R}[[x]]$ is a parabolic formal power series, its formal normal form f_0 in $\mathcal{L}^0$ is different from the standard formal normal form (recalled in Section 3.2). Indeed, due to the fact that we use a wider class of changes of variables, the residual term can also be eliminated. See Example 6.3 for details.

4.4. Proof of the precise form of Theorem A

The proof is divided in three parts. Let $f\in\mathcal{L}^H$ (resp. $f\in\mathcal{L}^H_{\mathfrak{D}}$).

1. *Part* 1 is *the step* of the algorithm. We describe a process which allows, by an appropriate elementary change of variables as defined on page 904, to eliminate the smallest possible monomial of $\mathcal{S}(f)$.
2. *Part* 2 is *the convergence* of the algorithm. We prove that the collection of consecutive changes of variables made in *Part 1* is actually a *transfinite sequence*, which can be indexed by a well-ordered subset of $\mathbb{R}_{>0} \times \mathbb{Z}$. The main difficulty here is the following one: each execution of a local step of the algorithm, while eliminating a single monomial of the support of the transseries to which it is applied, may at the same time add infinitely many new monomials to the support. Hence, we have to prove that, nevertheless, all the monomials which appear during the process (except at most finitely many of them) will be ultimately eliminated by a transfinite sequence of elementary changes of variables.
3. Finally, in *Part 3*, we show how to obtain another normal form, which is the formal time-one map of a vector field in the sense of Definition 1.2.

Part 1 (the step of the algorithm).

Let $f \in \mathcal{L}^H$. We examine three possible situations: f parabolic, hyperbolic, or strongly hyperbolic. In each case, we show how to construct an elementary change of variables which will eliminate the smallest possible monomial of $\mathcal{S}(f)$. We give a detailed description for f parabolic. The other cases follow the similar scheme.

(a) f *parabolic.* Let us write:

$$f(x) = x + ax^{\alpha}\boldsymbol{\ell}^{k} + a_1 x^{\alpha}\boldsymbol{\ell}^{k+1} + \text{h.o.t.},$$
$$a \in \mathbb{R},\ a \neq 0,\ (\alpha, k) \succ (1, 0).$$

In *Part 2* of the proof we want to eliminate the term of smallest possible order in the expansion of f, and proceed by induction. To see which terms can be eliminated, we examine the action of an elementary change of variables $\varphi_{\beta,m}$:

$$\varphi_{\beta,m}(x) = x + cx^{\beta}\boldsymbol{\ell}^{m}, \quad c \in \mathbb{R},\ c \neq 0,\ (\beta, m) \in \mathbb{R}_{>0} \times \mathbb{Z},\ (1, 0) \prec (\beta, l).$$

We apply the method described in Section 3. Recall from (3.1) that if ψ is the leading term of the difference $f \circ \varphi_{\beta,m} - \varphi_{\beta,m} \circ f$, then

$$\varphi_{\beta,m}^{-1} \circ f \circ \varphi_{\beta,m} = f + \psi + \text{h.o.t.} \tag{4.4}$$

We prove that ψ is exactly the leading term of $ad_{ax^{\alpha}\boldsymbol{\ell}^{k}}(cx^{\beta}\boldsymbol{\ell}^{m})$, that is,

$$\psi = \text{Lt}\left(ad_{ax^{\alpha}\boldsymbol{\ell}^{k}}(cx^{\beta}\boldsymbol{\ell}^{m})\right),$$

where Lt() denotes the leading term of expression in brackets. Indeed, write

$$f(x) = x + \varepsilon(x), \quad \varphi_{\beta,m}(x) = x + \eta(x),$$

with $(1,0) \prec \operatorname{ord}(\varepsilon)$ and $(1,0) \prec \operatorname{ord}(\eta)$. We obtain, using Taylor formula:

$$\begin{aligned}
\left(\varphi_{\beta,m}\circ f - f\circ\varphi_{\beta,m}\right)(x) &= \varphi_{\beta,m}(x+\varepsilon(x)) - f(x+\eta(x)) \\
&= \varphi_{\beta,m}(x) + \varphi'_{\beta,m}(x)\,\varepsilon(x) + \frac{1}{2}\varphi''_{\beta,m}(x)\,\varepsilon^2(x) \\
&\quad - f(x) - f'(x)\,\eta(x) - \frac{1}{2}f''(x)\,\eta^2(x) + \text{h.o.t.} \\
&= x + \eta(x) + (1+\eta'(x))\,\varepsilon(x) + \frac{1}{2}\eta''(x)\,\varepsilon^2(x) \\
&\quad - x - \varepsilon(x) - (1+\varepsilon'(x))\,\eta(x) - \frac{1}{2}\varepsilon''(x)\,\eta^2(x) + \text{h.o.t.} \\
&= \eta'(x)\,\varepsilon(x) - \eta(x)\,\varepsilon'(x) \\
&\quad + \frac{1}{2}\left(\eta''(x)\,\varepsilon^2(x) - \varepsilon''(x)\,\eta^2(x)\right) + \text{h.o.t.}
\end{aligned}$$

The expansion of this expression gives

$$\begin{aligned}
&\left(\varphi_{\beta,m}\circ f - f\circ\varphi_{\beta,m}\right)(x) \\
&\quad = ca(\beta-\alpha)\,x^{\alpha+\beta-1}\boldsymbol{\ell}^{m+k} \\
&\qquad + ca_1(\beta-\alpha)\,x^{\alpha+\beta-1}\boldsymbol{\ell}^{m+k+1} + ca(m-k)\,x^{\alpha+\beta-1}\boldsymbol{\ell}^{m+k+1} \\
&\qquad + \frac{1}{2}\left(ca^2\beta(\beta-1)\,x^{\alpha+\beta-1+(\alpha-1)}\boldsymbol{\ell}^{m+2k} - c^2a\alpha(\alpha-1)\,x^{\alpha+\beta-1+(\beta-1)}\boldsymbol{\ell}^{2m+k}\right) \\
&\qquad + \text{h.o.t.}
\end{aligned}$$

Let us examine various possibilities for the leading term, depending on α, β, k, m. If $\beta \neq \alpha$, then the leading term of this expression is $ca(\beta-\alpha)\,x^{\alpha+\beta-1}\boldsymbol{\ell}^{m+k}$. If $\beta = \alpha$, notice that one of the terms of the second line could contribute to the order $(\alpha+\beta-1, m+k+1)$, when $\alpha = 1$, or when $\beta = 1$. But in this case, since $\alpha = \beta = 1$, the coefficients of these terms vanish. So, in any case, the leading term ψ of the former expression is exactly the leading term of $\operatorname{ad}_{ax^\alpha\boldsymbol{\ell}^k}(cx^\beta\boldsymbol{\ell}^m) = ca(\beta-\alpha)\,x^{\alpha+\beta-1}\boldsymbol{\ell}^{m+k} + ca(m-k)\,x^{\alpha+\beta-1}\boldsymbol{\ell}^{m+k+1}$. By (4.4), we now have:

$$\varphi_{\beta,m}^{-1}\circ f\circ\varphi_{\beta,m} = f + \operatorname{Lt}\left(\operatorname{ad}_{ax^\alpha\boldsymbol{\ell}^k}(cx^\beta\boldsymbol{\ell}^m)\right) + \text{h.o.t.}$$

In order to find the change of variables $\varphi_{\beta,m}$ whose action would eliminate a given monomial $dx^\gamma\boldsymbol{\ell}^l$ in the expansion of f, we need to solve the *homological equation*:

$$\operatorname{Lt}\left(\operatorname{ad}_{ax^\alpha\boldsymbol{\ell}^k}(cx^\beta\boldsymbol{\ell}^m)\right) = dx^\gamma\boldsymbol{\ell}^r. \tag{4.5}$$

That is, as in the proof of Proposition 3.2, we want to see which monomials are not in the image of $\operatorname{ad}_{ax^\alpha\boldsymbol{\ell}^k}(J_{\beta,m})$ for any $(\beta,m) \succ (1,0)$, since these cannot be eliminated by elementary changes of variables. The homological equation (4.5) leads to $\alpha+\beta-1=\gamma$. That is, to $\beta = \gamma-\alpha+1$. We have three possibilities:

(i) If $\beta \neq \alpha$, i.e. if $\gamma \neq 2\alpha - 1$, then we put $m = r - k$ and we can solve the homological equation.

(ii) If $\beta = \alpha$, i.e. if $\gamma = 2\alpha - 1$, then the homological equation becomes $ca\,(m-k)\,x^{\gamma}\boldsymbol{\ell}^{m+k+1} = dx^{\gamma}\boldsymbol{\ell}^{r}$. This equation leads to $m + k + 1 = r$, that is, $m = r - k - 1$. If $m \neq k$, i.e. if $r \neq 2k + 1$, the homological equation can be solved.

(iii) If $r = 2k + 1$, then the homological equation cannot be solved, so the term $dx^{\gamma}\boldsymbol{\ell}^{r}$ cannot be eliminated from f. We have

$$dx^{\gamma}\boldsymbol{\ell}^{r} \notin \mathrm{ad}_{\,ax^{\alpha}\boldsymbol{\ell}^{k}}\Big(\bigcup_{(1,0)\prec(\beta,m)} J_{\beta,m}\Big).$$

Note that the assumption $(1,0) \prec (\beta, m)$ on order of elementary changes of variables $\varphi_{\beta,m} - \mathrm{id}$ is necessary so that $\varphi_{\beta,m} \in \mathcal{L}^{0}$.

If $\alpha > 1$, it follows from our computations that all the terms in the expansion of f can be eliminated except for the first term $ax^{\alpha}\boldsymbol{\ell}^{k}$, and the *residual term* $dx^{2\alpha-1}\boldsymbol{\ell}^{2k+1}$.

If $\alpha = 1$, along with these two terms, we observe that the term $a_1 x\boldsymbol{\ell}^{k+1}$ is not in the image of $\mathrm{ad}_{\,ax\boldsymbol{\ell}^{k}}$. Indeed, to solve the homological equation, we need a change of variables $\varphi_{1,0}$, which is impossible by the comment above. Nevertheless, in that case, as initial step we apply the appropriate linear change of variables, $\varphi_{1,0}(x) = cx$, $c \neq 0$. The action of the linear change of variables on the first terms of f is explained in the following computation:

$$\varphi_{1,0}^{-1} \circ f \circ \varphi_{1,0} = x + ac^{\alpha-1}x^{\alpha}\boldsymbol{\ell}^{k} + \\ + c^{\alpha-1}\big(a_1 - ka\log c\big)x^{\alpha}\boldsymbol{\ell}^{k+1} + \text{h.o.t.} \tag{4.6}$$

Hence, if $(1, k+1) \in \mathcal{S}(f)$, we eliminate the term $a_1 x\boldsymbol{\ell}^{k+1}$ with the linear change of variables $\varphi_{1,0}(x) = e^{\frac{a_1}{k\cdot a}}x$.

Notice that, if $\alpha > 1$, we can use the linear change of variables $\varphi_{1,0}(x) = cx$, with $c = |a|^{-\frac{1}{\alpha-1}}$, to *normalize* the coefficient a to $\mathrm{sign}\,(a)$.

(b) *f hyperbolic.* Let

$$f(x) = \lambda x + ax\boldsymbol{\ell} + \text{h.o.t.}$$

Applying the change of variables $\varphi_{\beta,m}(x) = x + cx^{\beta}\boldsymbol{\ell}^{m}$, $c \in \mathbb{R}$, $(1,0) \prec (\beta, m)$, we obtain:

$$\begin{aligned} &\varphi_{\beta,m} \circ f - f \circ \varphi_{\beta,m} = \\ &\quad = c(\lambda^{\beta} - \lambda)x^{\beta}\boldsymbol{\ell}^{m} + \\ &\qquad - \big(ac(1 - \beta\lambda^{\beta-1}) + cm\lambda^{\beta}\log\lambda\big)\,x^{\beta}\boldsymbol{\ell}^{m+1} + \text{h.o.t.} \end{aligned}$$

Then we proceed as in (a) above: every term can be eliminated, except for the terms of order $(1,0)$ and $(1,1)$.

(c) f *strongly hyperbolic.* First, by the linear elementary change of variables $\varphi_{1,0}(x) = \lambda^{-\frac{1}{\alpha-1}}x$, we *normalize* the first term. Then, as in the parabolic case, we want to remove the other monomials by appropriate elementary changes of variables $\varphi_{\beta,m}(x) = x + cx^\beta \boldsymbol{\ell}^m$, $c \in \mathbb{R}$, $(1,0) \prec (\beta, m)$. Let:

$$f(x) = x^\alpha + dx^\gamma \boldsymbol{\ell}^r + \text{h.o.t.}, \quad (\alpha, 0) \prec (\gamma, r),\ d \in \mathbb{R},\ d \neq 0.$$

As above, we consider the difference $(\varphi_{\beta,m} \circ f - f \circ \varphi_{\beta,m})(x) = t_{\varphi_{\beta,m}}(x) + \text{h.o.t.}$ Here, the leading monomial $t_{\varphi_{\beta,m}}$ is given by:

$$t_{\varphi_{\beta,m}}(x) = \begin{cases} c(\alpha^m - \alpha)x^\alpha \boldsymbol{\ell}^m, & \beta = 1, \\ -c\alpha x^{\alpha+\beta-1}\boldsymbol{\ell}^m, & \alpha > 1,\ \beta \neq 1, \\ c\alpha^m x^{\alpha\beta}\boldsymbol{\ell}^m, & \alpha < 1,\ \beta \neq 1. \end{cases} \tag{4.7}$$

By the change of variables $\varphi_{\beta,m}$ which solves the equation $t_{\varphi_{\beta,m}}(x) = dx^\gamma \boldsymbol{\ell}^r$, we eliminate the term $x^\gamma \boldsymbol{\ell}^r$ from f. Notice that, unlike in the former cases, in the strongly hyperbolic case *all the monomials except for the first one can be eliminated.*

Part 2 (the convergence of the algorithm).

Let $f \in \mathcal{L}^H$. We repeatedly apply to f the changes of variables built in local *Part 1* of the proof. This step by step process leads to some *collection* $(\varphi_\mu)_{\mu\in I}$ of elementary changes of variables from $\mathcal{L}_0$, indexed by some initial segment I of the ordinals:

$$f \underset{\varphi_0}{\longrightarrow} f_1 = \varphi_0^{-1} \circ f \circ \varphi_0 \underset{\varphi_1}{\longrightarrow} f_2 = \varphi_1^{-1} \circ f_1 \circ \varphi_1 \to \cdots .$$

For each step μ, the change of variables φ_μ is designed to eliminate the smallest possible monomial of the support $\mathcal{S}(f_\mu)$. We have to prove that the collections (φ_μ) and (f_μ) obtained in the process are *transfinite sequences*. That is, that there exists a bounding ordinal θ such that I is the set of ordinals $\{\mu < \theta\}$. The idea is to analyze the orders of elementary changes of variables used in step by step eliminations of all possible monomials from f. The analysis for f parabolic (other two cases can be done similarly) is given in Section 7, Subsections 7.1 and 7.2. For $f(x) = x + ax^\alpha \boldsymbol{\ell}^k + \text{h.o.t.}$, $a \neq 0$, we prove that the supports of all $(f_\mu - \text{id})$ belong to the set $\mathcal{R} \subset \mathbb{R}_{>0} \times \mathbb{Z}$:

$$\mathcal{R} = \Big\langle \mathcal{S}(f - \text{id}) \setminus \{\alpha, k\} - (\alpha, k+1) \Big\rangle + \mathbb{N}_* (\alpha - 1, k) + \{1\} \times \mathbb{N}_*.$$

The orders of the elementary changes of variables used for normalization thus belong to the set $\mathcal{R}_1 \subset \mathbb{R}_{>0} \times \mathbb{Z}$ explicitly obtained from $\mathcal{R}$:

$$\mathcal{R}_1 = \Big\langle \mathcal{S}(f - \text{id}) \setminus \{\alpha, k\} - (\alpha, k+1) \Big\rangle + \mathbb{N} (\alpha - 1, k) + \{0\} \times \mathbb{N}.$$

Both $\mathcal{R}$ and $\mathcal{R}_1$ are well-ordered by Neumann's Lemma 2.2, since $\mathcal{S}(f - \text{id})$ is well-ordered. The computations in *Part 1* of the proof show that, in each step, not only the monomial of smallest order in f_μ is eliminated, but no other monomial of smaller order is added to the support, so that the orders of $(f_\mu - \mathrm{id})$ and accordingly of $(\varphi_\mu - \mathrm{id})$ strictly increase and at the same time stay inside well-ordered sets $\mathcal{R}$ resp. $\mathcal{R}_1$. The steps of eliminations can be carried through, since we know from Propositions 4.4 and 4.6 that the partial compositions ψ_ν and f_ν at any limit ordinal ν do exist in $\mathcal{L}^0$.

Let us now index the collection (φ_μ) of elementary changes of variables by the orders $\mathrm{ord}(\varphi_\mu - \mathrm{id})$. The orders form a *strictly increasing*, *well-ordered* subset W of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$. Therefore, we obtain a transfinite sequence of elementary changes of variables and we use the notation $(\varphi_{\beta,m})_{(\beta,m)\in W}$. According to Proposition 4.6, the transfinite composition $\varphi = \circ_{(\beta,m)\in W}\varphi_{\beta,m}$ is a well-defined element of $\mathcal{L}^0$. On the other hand, Proposition 4.4 guarantees that our transfinite process of eliminations *ends*, that is, *converges to an element from* $\mathcal{L}^H$. By construction in the algorithm, the limit is the normal form $f_0 \in \mathcal{L}^H$.

If additionally $f \in \mathcal{L}^H_{\mathfrak{D}}$, that is if f is of finite type, we prove that the normalizing change of variables φ is also of finite type. The proof of this fact is quite long and technical. It is described in detail for the parabolic case in Section 7: Appendix. The proof of the hyperbolic and the strongly hyperbolic case follow the same lines and are left to the reader.

Part 3 (the 2nd normal form).

(i) Take the vector field X as in $(2.a)$ or $(2.b)$ of the theorem. Expanding the *coefficient* $\xi(x)$ of the vector field in the geometric series, we see that $\xi \in \mathcal{L}$ and that $(1,0) \preceq \mathrm{ord}\,(\xi)$. By Propositions 5.11 and 5.12 of Sections 5.2 and 5.4, the exponential of X (formula (1.4)) converges in $\mathcal{L}$ and gives a normal form as the formal time-one map $\widehat{f_0} \in \mathcal{L}^H$. It should be mentioned that the appropriate topology for the convergence of the series in (1.4) depends on whether $\mathrm{ord}\,(\xi)$ is equal to or bigger than $(1,0)$. This is why the proof of Proposition 5.11 is split between Subsection 5.2 and 5.4. Finally, we simply observe from this expansion that $\widehat{f_0} = f_0 + \text{h.o.t.}$

(ii) Let $f \in \mathcal{L}^H$, and let f_0 and $\widehat{f_0}$ be as in the statements 1. and 2. of the theorem. We show that there exists a change of variables $\widehat{\varphi} \in \mathcal{L}^0$ that conjugates f to $\widehat{f_0}$. Indeed, by Theorem A(1), there exists $\varphi \in \mathcal{L}^0$ such that

$$f = \varphi \circ f_0 \circ \varphi^{-1}. \tag{4.8}$$

On the other hand, by (i) above, $\widehat{f_0} \in \mathcal{L}^H$ and $\widehat{f_0} = f_0 + \text{h.o.t.}$ Applying the transfinite algorithm described in *Parts 1–2* of the proof to $\widehat{f_0} \in \mathcal{L}^H$, we obtain an element $\psi \in \mathcal{L}^0$ such that:

$$\widehat{f_0} = \psi \circ f_0 \circ \psi^{-1}. \tag{4.9}$$

By (4.8) and (4.9), it follows that f can be transformed into $\widehat{f_0}$ by composition $\widehat{\varphi} = \varphi \circ \psi^{-1}$, $\widehat{\varphi} \in \mathcal{L}^0$. That is, f is conjugated to $\widehat{f_0}$ in $\mathcal{L}^0$. Note that by Proposition 4.6(2) $\widehat{\varphi}$ can be considered as a transfinite step by step process of elementary changes of variables applied to f. $\square$

Remark 4.7. Let $f_0 \in \mathcal{L}^H$ be already in the form as in Theorem A(1), (a), (b) or (c). Let the beginning of $f \in \mathcal{L}^H$ coincide with f_0:

$$f = f_0 + \text{h.o.t.}$$

It is easy to see that the algorithm described in the proof of Theorem A transforms f to f_0 itself. In other words, every transseries from $\mathcal{L}^H$ which begins by one of the normal forms f_0 from Theorem A has f_0 itself as its normal form in $\mathcal{L}$.

5. Proof of Theorem B

In this section we state and prove the precise form of Theorem B. It turns out that, unlike in the proof of Theorem A, the techniques involved depend strongly on the nature (parabolic, hyperbolic, or strongly hyperbolic) of the element $f \in \mathcal{L}^H$. Hence, we divide the statement and the proof in different subsections. In Subsection 5.1, we recall and state some useful facts about *linear* operators on $\mathcal{L}$, more specifically about isomorphisms and derivations on $\mathcal{L}$. An important part is the relationship between elements of $\mathcal{L}^H$ and linear operators acting on $\mathcal{L}$. Subsection 5.2 is dedicated to vector fields, that present a particular class of derivations on $\mathcal{L}$. Finally, Subsections 5.3, 5.4, 5.5 contain the statement and the proof of Theorem B respectively in parabolic, hyperbolic and strongly hyperbolic case.

5.1. Operators acting on $\mathcal{L}$, isomorphisms and derivations

Some notions considered in this section are similar to [8, Chapter I.3] for formal power series.

By an *operator on $\mathcal{L}$ (respectively $\mathcal{L}_{\mathfrak{D}}$)*, we denote a strongly linear map $B\colon \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$) $\to \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$). By *strongly linear*, we mean that

$$B\left(\sum_{\alpha,k} c_{\alpha,k} x^{\alpha} \boldsymbol{\ell}^{k}\right) = \sum_{\alpha,k} c_{\alpha,k} B\left(x^{\alpha}\boldsymbol{\ell}^{k}\right), \quad c_{\alpha,k} \in \mathbb{R},$$

for every transseries $\sum_{\alpha,k} c_{\alpha,k} x^{\alpha} \boldsymbol{\ell}^{k} \in \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$).

We denote by $L(\mathcal{L})$, respectively $L(\mathcal{L}_{\mathfrak{D}})$, the set of all operators $B : \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$) $\to \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$).

For an operator $B \in L(\mathcal{L})$ and an element $f \in \mathcal{L}$, we denote indifferently by $B \cdot f$ or $B(f)$ the image of f under B. The identity operator will be denoted by Id.

Definition 5.1 *(Operators defined as a series of operators).* Let $(B_j)_{j\in\mathbb{N}}$ be a sequence of operators in $L(\mathcal{L})$ (resp. in $L(\mathcal{L}_{\mathfrak{D}})$).

1. We say that the operator $B \in L(\mathcal{L})$ (resp. $B \in L(\mathcal{L}_{\mathfrak{D}})$) is *well-defined by the series* $\sum_{j=0}^{\infty} B_j$ if, for every $f \in \mathcal{L}$ (resp. $f \in \mathcal{L}_{\mathfrak{D}}$), the sequence $\sum_{j=0}^{N} B_j \cdot f$ converges towards $B \cdot f$ in the *formal topology*, as $N \to \infty$.
2. If for every $f \in \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$) the sequence $\sum_{j=0}^{N} B_j \cdot f$ converges towards $B \cdot f$ in the *weak topology*, we say that B is *weakly well-defined by* $\sum_{j=0}^{\infty} B_j$.

In both cases, we write $B := \sum_{j=0}^{\infty} B_j$.

The notion *weakly* used throughout the article indicates relation to the weak topology on $\mathcal{L}$, see also the Definition 5.22 of the *small operator in the weak sense* in Section 5.4.

Note that *well-defined* is a stronger notion than *weakly well-defined*, since it relates to the stronger formal topology. That is, if an operator B is well-defined by a series of operators, then it is also weakly well-defined by the same series. Note also that the operator series defines an operator in $L(\mathcal{L})$ *as soon as the convergence of the series is weak*.

Definition 5.2 *(Formal differential operator in $L(\mathcal{L})$).* We say that an operator $B \in L(\mathcal{L})$ (resp. $B \in L(\mathcal{L}_{\mathfrak{D}})$) is a *formal differential operator* if there exists a sequence $(h_j)_{j\in\mathbb{N}}$ of elements of $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$) such that B is (weakly) well-defined by the series

$$B = \sum_{j=0}^{\infty} h_j \frac{\mathrm{d}^j}{\mathrm{d}x^j}.$$

The following definition of a *small operator* is inspired by [3, Section 1.3].

Definition 5.3. An operator $B : \mathcal{L} \to \mathcal{L}$ is *small* if there exists a well-ordered set $R \subseteq \mathbb{R}_{\geq 0} \times \mathbb{Z}$ of exponents strictly bigger than $(0, 0)$ such that $\mathcal{S}(B.f) \subseteq \mathcal{S}(f) + R$, for every $f \in \mathcal{L}$. An operator $B : \mathcal{L}_{\mathfrak{D}} \to \mathcal{L}_{\mathfrak{D}}$ is small if the set R is in addition of finite type.

Proposition 5.4. *Let B be a small operator on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$) and let $(c_k)_{k\in\mathbb{N}\cup\{0\}}$ be a sequence of real numbers. The sum*

$$S := \sum_{k=0}^{\infty} c_k B^k \tag{5.1}$$

is a well-defined operator on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$). Here, B^k denotes the k-th iterate of B.

Proposition 5.4 is a special case of a more general fact used repeatedly in [3]. Note that the proof is based on the *smallness property* of operator B. It implies indeed that $\mathcal{S}(S.f) \subseteq \mathcal{S}(f) + \langle R\rangle$, where $\langle R\rangle$ denotes the (additive) sub-semigroup of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$ generated by R. Furthermore, for any $f \in \mathcal{L}$, the order $\mathrm{ord}(B^k.f)$ strictly increases as k increases, by at least $\min\{R\}$ in every step. Consequently, the series

$S_n := \sum_{k=0}^{n} c_k B^k.f \in \mathcal{L}$ converges to $S \in \mathcal{L}$ in the *formal topology*. We omit the details of the proof.

Proposition 5.5. *Let B be a small operator on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$). Then*

$$\exp(B) := \sum_{k=0}^{\infty} \frac{B^k}{k!}, \quad \log(\mathrm{Id}+B) := \sum_{k=1}^{\infty} \frac{(-1)^{k+1} B^k}{k},$$

$\exp\big(\log(\mathrm{Id}+B)\big)$ *and* $\log\exp(B)$ *are well-defined operators on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$). Moreover,*

$$\exp\big(\log(\mathrm{Id}+B)\big) = \mathrm{Id}+B \ \textit{and} \ \log\exp(B) = B. \tag{5.2}$$

Proof. Since B is a small operator, by Proposition 5.4, $\log(\mathrm{Id}+B)$ and $\exp B$ are well-defined operators on $\mathcal{L}$ ($\mathcal{L}_{\mathfrak{D}}$). Moreover, by Definition 5.3 of small operators on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$), we obtain inductively:

$$\mathcal{S}(B.f) \subseteq \mathcal{S}(f) + R,\ \mathcal{S}(B^k \cdot f) \subseteq \mathcal{S}(f) + \langle R \rangle,\ k \in \mathbb{N}_0,$$

where R is as in Definition 5.3. Therefore,

$$\mathcal{S}(\log(\mathrm{Id}+B)),\ \mathcal{S}(\exp(B)) \subseteq \mathcal{S}(f) + \langle R \rangle.$$

The operators $\exp B$ and $\log(\mathrm{Id}+B)$ are small in $L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$). It follows from Proposition 5.4 that $\exp\big(\log(\mathrm{Id}+B)\big)$ and $\log\exp(B)$ are also well-defined operators in $L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$). The equality (5.2) now follows by symbolic computation from the standard properties of formal exp–log series, similarly as in the proof of Proposition 5.9. □

Definition 5.6. Let $B\colon \mathcal{L} \to \mathcal{L}$ be an operator on $\mathcal{L}$.

1. We say that B is a *derivation* if it satisfies the usual Leibniz's rule.
2. We say that $B : \mathcal{L} \to \mathcal{L}$ is a *morphism* if it satisfies the *morphism property*: $B(f \cdot g) = B(f) \cdot B(g)$, $f, g \in \mathcal{L}$.
3. We say that $B : \mathcal{L} \to \mathcal{L}$ is an *isomorphism* of $\mathcal{L}$ if B is a bijective morphism.

Remark 5.7 *(Isomorphisms associated with $f \in \mathcal{L}^H$ parabolic or hyperbolic).* Let $f \in \mathcal{L}^H$ be parabolic or hyperbolic. The map $F\colon \mathcal{L} \to \mathcal{L}$ defined by

$$F(g) = g \circ f,\ g \in \mathcal{L}, \tag{5.3}$$

is an isomorphism of $\mathcal{L}$. Moreover, the same conclusion holds in finitely generated case. If $f \in \mathcal{L}_{\mathfrak{D}}^H$ is parabolic or hyperbolic, then F defined by (5.3) is an isomorphism of $\mathcal{L}_{\mathfrak{D}}^H$.

We call such F the *isomorphism associated with* f and denote it by

$$F = \mathrm{iso}(f).$$

The morphism property is easily checked. Moreover, since a parabolic (resp. hyperbolic) element $f \in \mathcal{L}^H$ admits a parabolic (resp. hyperbolic) compositional inverse $f^{-1} \in \mathcal{L}^H$, then F is bijective, with the inverse $F^{-1} : \mathcal{L} \to \mathcal{L}$, $F^{-1}(g) = g \circ f^{-1}$, $g \in \mathcal{L}$.

Lemma 5.8. *Let $f \in \mathcal{L}^H$ (resp. $f \in \mathcal{L}^H_{\mathfrak{D}}$) be parabolic or hyperbolic contraction. Let the operator F be defined as in (5.3). Then the formal operators $\log F \in L(\mathcal{L})$, $\exp\log F \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$) are weakly well-defined. Moreover,*

$$\exp\log F = F.$$

Finally, if f is parabolic, these operators are well-defined.

Proof of Lemma 5.8. We prove here the lemma for f parabolic. For f hyperbolic, the proof is postponed to Section 5.4. We write $f(x) = x + \varepsilon(x)$, with $\mathrm{ord}(\varepsilon) \succ (1, 0)$. By Taylor expansion, for every $g \in \mathcal{L}$, we have:

$$\begin{aligned} g(f(x)) &= g(x + \varepsilon(x)) \\ &= g(x) + \sum_{k=1}^{\infty} \frac{g^{(k)}(x)}{k!} \varepsilon(x)^k . \end{aligned} \tag{5.4}$$

Hence, we can write $F = \mathrm{Id} + P = \mathrm{Id} + \sum_{k=1}^{\infty} \frac{\varepsilon(x)^k}{k!} \frac{\mathrm{d}^k}{\mathrm{d}x^k}$. Obviously, $P \cdot g = g \circ f - g \in \mathcal{L}$, $g \in \mathcal{L}$. We show that P is a small operator. By (5.4), we have:

$$\mathcal{S}(P \cdot g) = \bigcup_{k \in \mathbb{N}} \mathcal{S}\big(g^{(k)} \varepsilon^k\big).$$

The support $\mathcal{S}\big(g^{(k)}\varepsilon^k\big)$ contains pairs of the form:

$$\Big((\beta_1 - 1) + \cdots + (\beta_k - 1) + \alpha,\ l_1 + \cdots + l_k + m + j\Big),$$

where $(\beta_i, l_i) \in \mathcal{S}(\varepsilon)$, $i = 1, \ldots, k$, $(\alpha, m) \in \mathcal{S}(g)$ and $j \in \{0, \ldots, k\}$. Therefore,

$$\mathcal{S}(P \cdot g) \subseteq \mathcal{S}(g) + R, \tag{5.5}$$

where R is a sub-semigroup $R \subseteq \mathbb{R}_{\geq 0} \times \mathbb{Z}$ generated by elements $(\beta - 1, l)$ for $(\beta, l) \in \mathcal{S}(\varepsilon)$, and $(0, 1)$. By Neumann's lemma and since $(1, 0) \prec \mathrm{ord}(\varepsilon)$, R is well ordered and its elements are of order strictly greater than $(0, 0)$. Therefore, the operator P is small. By Proposition 5.5, the operators $\log F$ and $\exp(\log F) : \mathcal{L} \to \mathcal{L}$ are well-defined.

It remains to be proven that

$$\exp(\log F).f = F.f,\ \ f \in \mathcal{L}. \tag{5.6}$$

But once formal convergence is proven, this property follows from the well-known formal identities concerning $\exp - \log$ series.

In the finitely generated case ($f \in \mathcal{L}_{\mathfrak{D}}$), the semigroup R above is in addition of finite type (a subset of a finitely generated sub-semigroup of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$), for details see the "finite part" of the proof of Proposition 4.6. The operator P is small in $L(\mathcal{L}_{\mathfrak{D}})$, and the result follows by Proposition 5.5. □

We suspect that the next statement is already known, but we could not find it in the literature. Therefore, we give a short proof.

Proposition 5.9. *Let $A : \mathcal{L} \to \mathcal{L}$ be a linear morphism. Assume that the operator $\log A : \mathcal{L} \to \mathcal{L}$ is (weakly) well-defined. Then $\log A$ is a derivation.*

Proof of Proposition 5.9. Take any $f, g \in \mathcal{L}$. We prove the Newton–Leibniz rule, that is,

$$\log A(fg) = \log A\,(f)\,g + f\log A\,(g)\,.$$

Put $H = A - Id$. Using the fact that A is a morphism acting on $\mathcal{L}$, we compute:

$$\begin{aligned} H(fg) = A(fg) - fg = A\,(f)\,A\,(g) - fg = (f + H\,(f)) \cdot (g + H\,(g)) - fg = \\ = H\,(f)\,g + fH\,(g) + H\,(f)\,H\,(g)\,. \end{aligned} \tag{5.7}$$

Using the linearity of H and (5.7), we compute $H^2(fg)$:

$$H^2(fg) = H^2\,(f)\,g + 2H^2\,(f)\,H\,(g) + 2H\,(f)\,H\,(g) + 2H\,(f)\,H^2\,(g) + fH^2\,(g)\,.$$

We proceed by symbolic computation. We substitute

$$x^i \text{ for } H^i\,(f)\,,\ \ y^i \text{ for } H^i\,(g)\,,\ \ i \in \mathbb{N}_0.$$

By induction, the symbolic computation allows to substitute

$$(x + xy + y)^k \text{ for } H^k(fg),\ \ k \in \mathbb{N}_0.$$

Hence, we have:

$$\begin{aligned} \log A(fg) = (\text{substitution}) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}(x + xy + y)^i = \\ = \log(1 + x + xy + y) = \log\big((1 + x)(1 + y)\big) = \end{aligned}$$

$$= \log(1+x) + \log(1+y) =$$
$$= \text{(substitution)} = \log A(f)\, g + f \log A(g),$$

which proves that $\log A$ is a derivation. □

5.2. Vector fields and differential operators

We focus in this subsection on a special type of derivations. We denote by $\frac{\mathrm{d}}{\mathrm{d}x}$ the usual derivation on germs of functions. Note that, by strong linearity, the derivation $\frac{\mathrm{d}}{\mathrm{d}x}$ can be extended as an *operator on* $\mathcal{L}$.

Definition 5.10. An operator B on $\mathcal{L}$ is a *vector field* if there exists $\xi \in \mathcal{L}$ such that $B = \xi \frac{\mathrm{d}}{\mathrm{d}x}$.

Notice that there is an important difference here between $\mathcal{L}$ and $\mathbb{R}[[x]]$. A vector field is determined by its value on the element $x \in \mathcal{L}$. But since $\mathcal{L}$ contains infinitely many elements, which are, on $\mathbb{R}$, algebraically independent of x (such as, for example, the powers $x^\alpha, \alpha \in \mathbb{R}_{>0} \setminus \mathbb{Q}$), then all the derivations on $\mathcal{L}$ cannot be vector fields.

Theorem B discusses the possible embedding of an element $f \in \mathcal{L}^H$ (in the three cases) in a formal flow of a vector field from $\mathcal{L}$. Let us recall the definition of a formal flow, adapted from the standard definition in the usual setting of formal power series to our class $\mathcal{L}$. The next discussions follow the lines of similar results for usual power series (see [8, Chapter I.3], for example).

Proposition 5.11 *(The existence of a formal flow of a formal vector field in $\mathcal{L}$, the parabolic case). Let $X = \xi \frac{\mathrm{d}}{\mathrm{d}x}$, $\xi \in \mathcal{L}$, be a vector field in $\mathcal{L}$ such that* $(1,0) \prec \mathrm{ord}(\xi)$. *Then the vector field X admits the $\mathcal{C}^1$-formal flow $\{h^t \in \mathcal{L}^0 : t \in \mathbb{R}\}$ defined by* $h^t := H^t \cdot \mathrm{id}$, *where $\{H^t \in L(\mathcal{L}) : t \in \mathbb{R}\}$ is the one-parameter group of isomorphisms of $\mathcal{L}$ well-defined by:*

$$H^t := \exp(tX) = \sum_{k=0}^{\infty} \frac{t^k}{k!} X^k. \tag{5.8}$$

Moreover, H^t are the isomorphisms associated to h_t, $t \in \mathbb{R}$, in the sense of Remark 5.7.

If, in addition, $\xi \in \mathcal{L}_{\mathfrak{D}}$, then $H^t \in L(\mathcal{L}_{\mathfrak{D}})$, $h^t \in \mathcal{L}^0_{\mathfrak{D}}$, $t \in \mathbb{R}$.

Proposition 5.12 *(The existence of a formal flow of a formal vector field, the hyperbolic case). Let $X = \xi \frac{\mathrm{d}}{\mathrm{d}x}$, $\xi \in \mathcal{L}$, be a vector field in $\mathcal{L}$ such that* $\mathrm{ord}(\xi) = (1,0)$. *Then the statements of Proposition 5.11 hold in this case as well, with the difference that H^t is just* weakly *well-defined by (5.8).*

Note that the time-t map h^t of X in Propositions 5.11 and 5.12 is given by the following formula:

$$h^t = H^t \cdot \mathrm{id} = \mathrm{id} + t\xi + \frac{t^2}{2!}\xi'\xi + \frac{t^3}{3!}(\xi'\xi)'\xi + \cdots . \tag{5.9}$$

Note also that in the case $(1,0) \prec \mathrm{ord}(\xi)$, $h^t \in \mathcal{L}$ are *parabolic*, while in the case $\mathrm{ord}(\xi) = (1,0)$ they are *hyperbolic*. Moreover, the formula (5.9) converges in the formal topology if $(1,0) \prec \mathrm{ord}(\xi)$, and in the weak topology if $\mathrm{ord}(\xi) = (1,0)$.

Proof of Proposition 5.11. The assumption $\mathrm{ord}(\xi)$ guarantees that $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$ is a *small operator* in the sense of Definition 5.3. It is easy to check that

$$\mathcal{S}(X.g) = \mathcal{S}(\xi \cdot g') \subseteq \mathcal{S}(g) + R,\ g \in \mathcal{L}, \tag{5.10}$$

where R is a sub-semigroup of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$ generated by elements $(\beta - 1, l)$, $(\beta, l) \in \mathcal{S}(\xi)$, and $(0,1)$. All elements of R are of order strictly bigger than $(0,0)$. Hence, the sum (5.8) gives by Proposition 5.4 a *well-defined* operator H^t for all $t \in \mathbb{R}$.

The statement in $\mathcal{L}_{\mathfrak{D}}$ follows as in the proof of finite part of Lemma 5.8. By (5.10), we have that $\mathcal{S}(H^t.g) \subseteq \mathcal{S}(g) + R$, $g \in \mathcal{L}$, $t \in \mathbb{R}$.

Finally, the proof of the morphism property of operators H^t, $t \in \mathbb{R}$, and the proof that the family $(h^t)_t$ is a flow of X (see Definition 1.2) are routine, following the lines of similar results for formal power series, see for example [8, Chapter I.3].

To prove that $H^t.f = f \circ h_t$, $f \in \mathcal{L}$, we combine Lemma 5.17 below in this section and Proposition 5.18. □

The proof of Proposition 5.12 (the case $\mathrm{ord}\,(\xi) = (1,0)$) is more involved and is postponed to Subsection 5.4. In fact, in this case X is not a small operator in the sense of Definition 5.3. Therefore, H^t is not well-defined by (5.8). Nevertheless, we prove in Subsection 5.4 that it is *weakly well-defined* by (5.8).

Proposition 5.13 *(Uniqueness of the $\mathcal{C}^1$-formal flow of a vector field). Consider a vector field $X = \xi\frac{d}{dx}$, $\xi \in \mathcal{L}$.*

1. *If $(1,0) \preceq \mathrm{ord}\,(\xi)$, then there exists a* unique *$\mathcal{C}^1$-formal flow $(f^t)_{t\in\mathbb{R}}$ of X, $f^t \in \mathcal{L}^H$, in the sense of Definition 1.2. Moreover:*
   (i) *if $(1,0) \prec \mathrm{ord}(\xi)$, then the $f^t \in \mathcal{L}^H$ are parabolic;*
   (ii) *if $\mathrm{ord}(\xi) = (1,0)$, then the $f^t \in \mathcal{L}^H$ are hyperbolic.*
2. *If $\mathrm{ord}\,(\xi) \prec (1,0)$, then X does not admit any $\mathcal{C}^1$-flow.*

Proof. 1. The existence of a $\mathcal{C}^1$-flow for the vector field X is shown by an explicit construction in Propositions 5.11 and 5.12. Suppose now that X admits two $\mathcal{C}^1$-flows $(f^t)_{t\in\mathbb{R}}$ and $(g^t)_{t\in\mathbb{R}}$ in $\mathcal{L}^H$. Let S be a well-ordered subset of $\mathbb{R}_{>0} \times \mathbb{Z}$ such that $\mathcal{S}(f^t)$ and $\mathcal{S}(g^t)$ are contained in S for all $t \in \mathbb{R}$. Let (α, m) be the smallest element of S (which exists since S is well-ordered) such that the coefficient $h(t)$ of $x^\alpha \boldsymbol{\ell}^m$ in $f^t(x) - g^t(x)$ does not vanish identically. Since (f^t) and (g^t) are both $\mathcal{C}^1$-formal flows of $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$, we have the integral equation in $\mathcal{L}$:

$$f^t(x) - g^t(x) = \int_0^t \big(\xi(f^s)(x) - \xi(g^s)(x)\big)\mathrm{d}s, \quad \forall t \in \mathbb{R}, \tag{5.11}$$

where the integral on (5.11) is applied on each coefficient of the integrand. The coefficient of the monomial $x^\alpha \boldsymbol{\ell}^m$ on the left-hand side of (5.11) is $h(t)$. Hence, in order to estimate the coefficient of the same monomial on the right-hand side of this equation, we write, based on the definition of (α, m):

$$f^s(x) = M(s;x) + h_1(s)\, x^\alpha \boldsymbol{\ell}^m + \text{h.o.t.}$$
$$g^s(x) = M(s;x) + h_2(s)\, x^\alpha \boldsymbol{\ell}^m + \text{h.o.t.}$$

Here, $M(s;x) = b(s)\, x^\beta + \text{h.o.t.}$, $(\beta, 0) \prec (\alpha, m)$, is a transseries with monomials in S and coefficients in $\mathcal{C}^1(\mathbb{R})$ such that b is not identically zero, and h_1, h_2 are $\mathcal{C}^1$-functions. Obviously, $h = h_1 - h_2$.

Let $ax^\gamma \boldsymbol{\ell}^n$ be the leading term of ξ. We see that the leading term of the difference $\xi(f^s) - \xi(g^s)$ is

$$a\left(\frac{1}{\beta}\right)^n \gamma b(s)^{\gamma-1}(h_1(s) - h_2(s))\, x^{\alpha+\beta(\gamma-1)} \boldsymbol{\ell}^{m+n}$$
$$= a\left(\frac{1}{\beta}\right)^n \gamma b(s)^{\gamma-1}\, h(s)\, x^{\alpha+\beta(\gamma-1)} \boldsymbol{\ell}^{m+n}.$$

If $(1,0) \prec (\gamma, n)$, the order of the right-hand side is bigger than (α, m). It would imply $h \equiv 0$, which is a contradiction. On the other hand, if $(\gamma, n) = (1,0)$, by comparing the coefficients of $x^\alpha \boldsymbol{\ell}^m$ on both sides of (5.11), we see that:

$$h(t) = a\int_0^t h(s)\,\mathrm{d}s, \text{ so } |h(t)| \leq |a| \int_0^t |h(s)|\,\mathrm{d}s, \quad a \in \mathbb{R}.$$

It follows from *Gronwall's lemma* applied to $|h|$ that $h \equiv 0$, which is again a contradiction.

The points (i) and (ii) follow by uniqueness on one hand, and by the explicit construction of the flow done in the proof of Propositions 5.11 and 5.12 on the other hand.

2. Assume now that $(\beta, m) := \operatorname{ord}(\xi) \prec (1,0)$ and that X admits a $\mathcal{C}^1$-flow $(f^t)_{t\in\mathbb{R}}$. We show that this assumption leads to a contradiction. As above, let S be a well-ordered subset of $\mathbb{R}_{>0} \times \mathbb{Z}$ such that $\mathcal{S}(f^t) \subseteq S$ for all $t \in \mathbb{R}$. Let (α, m) be the smallest element of S such that the coefficient $h(t)$ of $x^\alpha \boldsymbol{\ell}^m$ in f^t does not vanish identically. Let $t_0 \in \mathbb{R}$ with $h(t_0) \neq 0$. In particular, $\operatorname{ord}(f^{t_0}(x)) = (\alpha, m)$. We have:

$$\left.\frac{\mathrm{d}f^t}{\mathrm{d}t}\right|_{t=t_0}(x) = \xi\left(f^{t_0}(x)\right).$$

The order of the left-hand side of this equation is bigger than or equal to (α, m). But since $\mathrm{ord}\,(\xi) \prec (1,0)$, the order of the right-hand side is strictly smaller than (α, m), and we get a contradiction. □

Corollary 5.14. *Let $X = \xi \frac{d}{dx}$, $\xi \in \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$) and let $(1,0) \preceq \mathrm{ord}(\xi)$. Then its $\mathcal{C}^1$-flow $(f^t)_t$, $f^t \in \mathcal{L}^H$ (resp. $\mathcal{L}^H_{\mathfrak{D}}$), is given uniquely by the formula:*

$$f^t := \exp(tX).\mathrm{id},\ \ t \in \mathbb{R}.$$

Proof. The proof follows by Propositions 5.11 and 5.12 and the uniqueness result in Proposition 5.13. □

Lemma 5.15. *Let $X = \xi \frac{\mathrm{d}}{\mathrm{d}x}$, $\xi \in \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$), be such that $(1,0) \prec \mathrm{ord}(\xi)$ or $\xi(x) = \lambda x + \mathrm{h.o.t.}$ with $\lambda < 0$. The operators $\exp(X)$ and $\log\exp(X)$ are weakly well-defined in $L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$) and*

$$\log\, \exp(X) = X.$$

Moreover, in the case $\mathrm{ord}(\xi) \succ (1,0)$, the operators are well-defined.

Proof. The result in the case $\mathrm{ord}(\xi) \succ (1,0)$ follows directly from Proposition 5.5, since X is a small operator in this case. The case $\mathrm{ord}(\xi) = (1,0)$ is proven in Section 5.4. □

Proposition 5.16 *(The convergence of the Taylor expansion). Let $f \in \mathcal{L}^H$ (resp. $\mathcal{L}^H_{\mathfrak{D}}$) be parabolic or hyperbolic contraction. Let $F = \mathrm{iso}(f) \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$). Put $f = \mathrm{id}+\varepsilon$. Then F is weakly well-defined as the formal differential operator:*

$$F = \mathrm{Id} + \sum_{k=1}^{\infty} \frac{\varepsilon^k}{k!} \frac{\mathrm{d}^k}{\mathrm{d}x^k}.$$

Moreover, if f is parabolic, then F is well-defined by the above series.

Note that Proposition 5.16 claims that in parabolic and hyperbolic cases the *Taylor expansions converge in $\mathcal{L}$* (in the respective topologies). That is, for every $g \in \mathcal{L}$, we can write:

$$F.g\,(x) = g \circ f(x) = g\,(x + \varepsilon\,(x)) = g\,(x) + \sum_{k=1}^{\infty} \frac{\varepsilon\,(x)^k}{k!} g^{(k)}\,(x)\,. \tag{5.12}$$

Proof. (*i*) *f parabolic.* Since $\mathrm{ord}(\varepsilon) \succ (1,0)$, the Taylor expansion (5.12) converges in the formal topology to $g \circ f$. Indeed, the orders $\mathrm{ord}(\varepsilon^k g^{(k)})$ strictly increase by a fixed value $(0,0) \prec \mathrm{ord}(h) - (1,0)$, as k increases.

(ii) f a hyperbolic contraction. We prove that if $\mathrm{ord}(\varepsilon) = (1, 0)$ the Taylor expansion (5.12) converges in the weak topology. Additionally, we prove that *the coefficients of respective monomials converge absolutely.*

Let $f(x) = \lambda x + \text{h.o.t.}$ be hyperbolic, with $0 < \lambda < 1$. Then $\varepsilon(x) = f(x) - x = (\lambda-1)x+\Psi(x)$, $(1,0) \prec \mathrm{ord}(\Psi)$. We prove that the Taylor expansion (5.12) for monomials $g(x) = x^\alpha$, $\alpha > 0$, and $g(x) = \big(\frac{1}{-\log x}\big)^m$, $m \in \mathbb{Z}$, converges in the weak topology. The convergence is then deduced for all elements $g \in \mathcal{L}$, since products of absolutely convergent series converge absolutely.

1. $g(x) = x^\alpha$. By definition of compositions in $\mathcal{L}$, see Section 2, we have:

$$\big(f(x)\big)^\alpha = \big(\lambda x + \Psi(x)\big)^\alpha = \lambda^\alpha x^\alpha \Big(1 + \frac{\Psi(x)}{\lambda x}\Big)^\alpha = \lambda^\alpha x^\alpha \sum_{k=0}^{\infty} \binom{\alpha}{k} \lambda^{-k} \Big(\frac{\Psi(x)}{x}\Big)^k. \tag{5.13}$$

Since $\mathrm{ord}(\frac{\Psi(x)}{x}) \succ (1,0)$, the above series converges in the formal topology.

Consider the series corresponding to the Taylor expansion (5.12):

$$\begin{aligned}
&\sum_{k=0}^{\infty} \frac{\big(x^\alpha\big)^{(k)}}{k!} \big((\lambda-1)x + \Psi(x)\big)^k \\
&= \sum_{k=0}^{\infty} \frac{\big(x^\alpha\big)^{(k)} \cdot x^k}{k!} (\lambda-1)^k \Big(1 + \frac{\Psi(x)}{(\lambda-1)x}\Big)^k = \\
&= \sum_{k=0}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)\cdot x^\alpha}{k!} (\lambda-1)^k \Big(1 + \frac{\Psi(x)}{(\lambda-1)x}\Big)^k = \\
&= \sum_{k=0}^{\infty} \binom{\alpha}{k} x^\alpha (\lambda-1)^k \Big[\sum_{l=0}^{k} \binom{k}{l} \Big(\frac{\Psi(x)}{(\lambda-1)x}\Big)^l\Big].
\end{aligned} \tag{5.14}$$

We show that the series converges in $\mathcal{L}$ in the weak topology to $\big(f(x)\big)^\alpha$ above. It can be easily seen that both the monomials of (5.13) and of (5.14) belong to $S = \bigcup_{k\in\mathbb{N}_0} \mathcal{S}\Big(\big(\frac{\Psi(x)}{x}\big)^k x^\alpha\Big)$. For every $k_0 \in \mathbb{N}_0$, $x^\alpha \Big(\frac{\Psi(x)}{x}\Big)^{k_0}$ is present in infinitely many elements of (5.14), but, due to the fact that $|\lambda - 1| < 1$, its coefficient converges to the coefficient of $x^\alpha \Big(\frac{\Psi(x)}{x}\Big)^{k_0}$ in (5.13):

$$\begin{aligned}
\sum_{k=k_0}^{\infty} \binom{\alpha}{k} (\lambda-1)^{k-k_0} \binom{k}{k_0} &= \sum_{k=k_0}^{\infty} \binom{\alpha}{k_0} \binom{\alpha-k_0}{k-k_0} (\lambda-1)^{k-k_0} \\
&= \binom{\alpha}{k_0} (1+\lambda-1)^{\alpha-k_0} = \binom{\alpha}{k_0} \lambda^{\alpha-k_0}.
\end{aligned}$$

Moreover, the convergence is absolute.

On the other hand, since $(1,0) \prec \mathrm{ord}\big(\frac{\Psi(x)}{x}\big)$, for every monomial $x^\beta \boldsymbol{\ell}^n \in S$ there exists $N \in \mathbb{N}$ such that $x^\beta \boldsymbol{\ell}^n \notin \mathcal{S}\Big(\big(\frac{\Psi(x)}{x}\big)^k x^\alpha\Big)$ for $k > N$. Together with the above analysis, this proves the convergence of coefficients of every monomial in (5.14) to its coefficient in (5.13). That is, the Taylor expansion (5.14) converges in the weak topology.

2. $g(x) = \left(\frac{1}{-\log x}\right)^m$. The proof of convergence of Taylor expansion for $g(x) = \left(\frac{1}{-\log x}\right)^m$, $m \in \mathbb{Z}$, follows the same idea, so we omit it. $\square$

The next Lemma 5.17 is a weaker version of the well-known diffeomorphism–isomorphism correspondence for the algebra $\mathbb{C}[[x]]$ of formal power series, see [8, Section 3A]. It is used, together with Proposition 5.18 below, to finish the proof of Propositions 5.11 and 5.12 concerning the correspondence $h_t \leftrightarrow H^t$. Their Corollary 5.19 is used to prove uniqueness in Theorem B.

Lemma 5.17 *(Formal diffeomorphism–isomorphism correspondence for $\mathcal{L}$). Let $B \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$) be a morphism which is also a (weakly) well-defined formal differential operator (in the sense of Definition 5.2), and such that $h := B \cdot \mathrm{id} \in \mathcal{L}^0$ is parabolic or a hyperbolic contraction. Then $B = \mathrm{iso}(h)$, $h \in \mathcal{L}^0$ (resp. $\mathcal{L}_{\mathfrak{D}}^0$).*

Proof. Since B is a formal differential operator, put

$$B = \mathrm{Id} + h_1 \frac{\mathrm{d}}{\mathrm{d}x} + h_2 \frac{\mathrm{d}^2}{\mathrm{d}x^2} + \cdots, \quad h_i \in \mathcal{L} \text{ (resp. } \mathcal{L}_{\mathfrak{D}}), \ i \in \mathbb{N}. \tag{5.15}$$

Given an integer $p > 1$, we compute $B(x^p)$ in two different ways and compare. First, by (5.15), we have:

$$B(x^p) = x^p + \sum_{n=1}^{p} h_n(x) \frac{\mathrm{d}^n(x^p)}{\mathrm{d}x^n} = x^p + \sum_{n=1}^{p} h_n(x) n! \binom{p}{n} x^{p-n}.$$

Since B is a morphism, we have:

$$B(x^p) = (B \cdot x)^p = (x + h_1(x))^p = x^p + \sum_{n=1}^{p} \binom{p}{n} h_1^n(x) x^{p-n}.$$

Identifying these two expressions for every integer $p > 1$, we see that

$$h_n = \frac{h_1^n}{n!}, \ n \in \mathbb{N} \cup \{0\}. \tag{5.16}$$

Let $h = B \cdot \mathrm{id} = \mathrm{id} + h_1$, $h \in \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$). Let $H = \mathrm{iso}(h)$, as defined in Remark 5.7. It follows from Proposition 5.16 that:

$$H = \mathrm{Id} + \sum_{n=1}^{\infty} \frac{h_1^n}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}.$$

By (5.15) and (5.16), $B = H$. Note that B is additionally well-defined by differential series (5.15) if f is parabolic. $\square$

Proposition 5.18. *Let* $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$, $\xi \in \mathcal{L}$ *(resp.* $\mathcal{L}_{\mathfrak{D}}$*) with* $(1,0) \preceq \mathrm{ord}(\xi)$*. The operators* $H^t = \exp(tX)$ *from Propositions 5.11 and 5.12 are weakly well-defined formal differential operators. If moreover* $\mathrm{ord}(\xi) = (1,0)$*, they are well-defined formal differential operators.*

Proof. Let $f \in \mathcal{L}$. Then by (5.8) we have:

$$\begin{aligned} H^t \cdot f = \exp(tX) \cdot f &= f + t\xi f' + \frac{t^2}{2!}(\xi f')'\xi + \frac{t^3}{3!}\big((\xi f')'\xi\big)'\xi + \cdots = \\ &= f + t\xi f' + \frac{t^2}{2!}(\xi\xi' f' + \xi^2 f'') \\ &\quad + \frac{t^3}{3!}\big(\xi(\xi')^2 f' + \xi^2\xi'' f' + \xi^2\xi' f'' + 2\xi^2\xi' f'' + \xi^3 f'''\big) + \cdots . \end{aligned} \tag{5.17}$$

We prove in both cases ($(1,0) \prec \mathrm{ord}(\xi)$ and $\mathrm{ord}(\xi) = (1,0)$) that we can *change the order of the summation* in the respective topologies so that we group the terms multiplying f, f', f'', etc:

$$\begin{aligned} H^t \cdot f &= f + \big(t\xi + \frac{t^2}{2!}\xi\xi' + \frac{t^3}{3!}\xi(\xi')^2 + \frac{t^3}{3!}\xi^2\xi'' + \cdots\big)f' + \big(\frac{t^2}{2!}\xi^2 + \frac{t^3}{3!}3\xi^2\xi' + \cdots\big)f'' \\ &\quad + \big(\frac{t^3}{3!}\xi^3 + \cdots\big)f''' + \cdots \\ &= f + h_1 f' + h_2 f'' + \text{h.o.t.} \end{aligned} \tag{5.18}$$

Obviously, by (5.17), $h_1 = H^t.\mathrm{id} \in \mathcal{L}$, $h_2 = \frac{1}{2}(H^t.x^2 - x^2) - xh_1 \in \mathcal{L}$, etc. Thus, $h_n \in \mathcal{L}$, $n \in \mathbb{N}$.

(i) $(1,0) \prec \mathrm{ord}(\xi)$. The orders of the summands in (5.18) increase by the fixed value $\mathrm{ord}(\xi) - (1,0) \succ (0,0)$, so (5.18) converges in the formal topology in $\mathcal{L}$ to an element of $\mathcal{L}$. Moreover, it converges to the same limit as (5.17), since the difference of partial sums of (5.17) and (5.18) converges to zero in the formal topology. Indeed, by Proposition 5.11 (5.17) the order of summands increases by the fixed value $\mathrm{ord}(\xi) - (1,0) \succ (0,0)$ also in (5.17).

(ii) $\mathrm{ord}(\xi) = (1,0)$. Let us represent $H^t.f$ by the following grid:

$$\begin{array}{llll} f & & & \\ * f' & & & \\ * f' & * f'' & & \\ * f' & * f'' & * f''' & \\ * f' & * f'' & * f''' & * f^{(4)} \\ \vdots & \dots & & \end{array} \tag{5.19}$$

Here, $*$ denotes the *coefficients* (transseries in ξ) of the respective powers of f in (5.17). The first row represents the first bracket in (5.17), the second row the second bracket in (5.17) etc.

Let us fix a monomial from the support $\mathcal{S}(H^t.f)$. The order of terms remains the same by rows and by columns, in contrast with the parabolic case. Therefore, a fixed monomial may appear in every term of every row and of every column of (5.19). Nevertheless, we have proven in Proposition 5.12 that (5.17) converges in the weak topology, meaning exactly that the coefficients of the given monomial converge when summation is done by rows. We prove that we can change the order of the summation of coefficients of the chosen monomial from summation by rows as in (5.17) to summation by columns as in (5.18). It would give us the convergence of (5.18) in the weak topology in $\mathcal{L}$ (to the same limit as (5.17)).

By the proof of Proposition 5.12, the coefficient of a fixed monomial of the support $\mathcal{S}(H^t.f)$ converges *absolutely* in (5.17). Moreover, we see that each row of (5.19) contains only *finitely many* elements of $\mathcal{L}$. Consequently, a fixed monomial can appear only finitely many times in each row. By the *Moore–Osgood theorem* stated on p. 939 (or see [16, Theorem 8.3]), we are allowed to change the order of the summation in our double sum and to sum coefficients by columns.

The finitely generated case follows easily. □

Corollary 5.19. *Let* $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$*,* $\xi \in \mathcal{L}$ *(resp.* $\mathcal{L}_{\mathfrak{D}}$*), be such that* $(1,0) \prec \mathrm{ord}(\xi)$ *or* $\xi(x) = \lambda x + \text{h.o.t.}$ *with* $\lambda < 0$*. Then, for any* $t \neq 0$*, the following two statements are equivalent:*

1. $\exp(tX) \cdot \mathrm{id} = f$*,*
2. $\exp(tX) \cdot h = h \circ f$*,* $h \in \mathcal{L}$ *(resp.* $\mathcal{L}_{\mathfrak{D}}$*).*

Proof. By Proposition 5.18, the operator $\exp(tX)$ is a (weakly) well-defined formal differential operator. (2) ⇒ (1) is obvious. We prove (1) ⇒ (2). Suppose (1) holds. By Lemma 5.17, $\exp(tX)$ is the isomorphism associated with $\exp(tX) \cdot \mathrm{id} = f$, which proves (2). □

5.3. Theorem B in the parabolic case

This section is dedicated to the precise statement and the proof of Theorem B for parabolic elements of $\mathcal{L}$.

Theorem *(Precise form of Theorem B for parabolic elements). Let* $f \in \mathcal{L}^H$ *(*resp. $f \in \mathcal{L}_{\mathfrak{D}}^H$*) be parabolic. Then there exists a unique formal vector field*

$$X = \xi\frac{\mathrm{d}}{\mathrm{d}x},\ \xi \in \mathcal{L}\ (\textit{resp. } \xi \in \mathcal{L}_{\mathfrak{D}}),$$

such that f *embeds in its* $\mathcal{C}^1$*-flow. Moreover,*

$$f = \exp(X) \cdot \mathrm{id}.$$

Here, $(1,0) \prec \mathrm{ord}(\xi)$*, and* $\exp(X)$ *is well-defined in* $L(\mathcal{L})$ *(resp.* $L(\mathcal{L}_{\mathfrak{D}})$*).*

Let $f \in \mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$) be a parabolic element as in the statement of Theorem B. Let $F = \mathrm{iso}(f)$:

$$F.h = h \circ f,\ h \in \mathcal{L} \text{ (resp. } \mathcal{L}_{\mathfrak{D}}).$$

We prove that the vector field X is given by $X = \log F$. Note that by Lemma 5.8 and Proposition 5.9, the operator $X = \log F$ is a well-defined operator on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$) and a derivation. The proof is now given in three steps:

(1) We prove in Lemma 5.21 that $X = \log F$ is a formal differential operator $\sum_k h_k \frac{\mathrm{d}^k}{\mathrm{d}x^k}$, $h_k \in \mathcal{L}$ (resp. $\mathcal{L}_D$).
(2) Since X is a derivation and at the same time a formal differential operator of the above form, we prove that X is necessarily a vector field. Moreover, we prove that f is the time-one map of X.
(3) We prove the uniqueness of the formal vector field whose time-one map is f.

Lemma 5.20. *Let $f \in \mathcal{L}^0$ (resp. $\mathcal{L}^0_{\mathfrak{D}}$) be parabolic and let $F = \mathrm{iso}(f) \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$). Let $H = F - \mathrm{Id}$. Then all the iterates H^k can be written as well-defined formal differential operators on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$):*

$$H^k = \sum_{\ell=1}^{\infty} h_\ell^k \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell},\ h_\ell^k \in \mathcal{L} \ (\textit{resp. } \mathcal{L}_{\mathfrak{D}}),\ k \in \mathbb{N}. \tag{5.20}$$

Proof. Let $f = \mathrm{id} + h$, $h \in \mathcal{L}$ with $(1,0) \prec \mathrm{ord}(h)$. The lemma is proven by induction. The induction basis ($k = 1$) follows easily by Taylor expansion:

$$H \cdot g = g \circ f - g = \sum_{\ell=1}^{\infty} \frac{h^\ell}{\ell!} \frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell},\ g \in \mathcal{L}.$$

Thus, $H = \sum_{\ell=1}^{\infty} h_\ell^0 \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}$, with the coefficients $h_\ell^0 = \frac{h^\ell}{\ell!} \in \mathcal{L}$, $\ell \in \mathbb{N}$. Assume that the operators H^m, $m \le k$ can be written in the form (5.20), with formal convergence on $\mathcal{L}$. Note that the formal convergence of series $H^m \cdot g$ from (5.20) is equivalent to asking that the orders of summands $\mathrm{ord}(h_\ell^m g^{(\ell)})$ infinitely increase as $\ell \to \infty$. We prove (5.20) for the operator H^{k+1}. By Taylor expansion, we have:

$$\begin{aligned} H^{k+1} \cdot g\,(x) &= H(H^k \cdot g)\,(x) = H^k \cdot g\,(x + h(x)) - H^k \cdot g\,(x) \\ &= \sum_{i=1}^{\infty} \frac{h(x)^i}{i!} \frac{\mathrm{d}^i (H^k.g)}{\mathrm{d}x^i} \\ &= \sum_{i=1}^{\infty} \frac{h^i}{i!} \frac{\mathrm{d}^i}{\mathrm{d}x^i} \Big(\sum_{\ell=1}^{\infty} h_\ell^k \frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell} \Big) = \sum_{i=1}^{\infty} \Big(\sum_{\ell=1}^{\infty} h_{i\ell}^k \frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell} \Big), \end{aligned} \tag{5.21}$$

with $h^k_{i\ell} \in \mathcal{L}$, $i, \ell \in \mathbb{N}$. We represent the double sum by the following grid:

$$H^{k+1} \cdot g \ : \qquad \begin{array}{cccccc} & & \overset{\ell}{\rightarrow} & & & \\ i\downarrow & h_{11}\frac{\mathrm{d}g}{\mathrm{d}x} & h_{12}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2} & h_{13}\frac{\mathrm{d}^3 g}{\mathrm{d}x^3} & \ldots \\ & h_{21}\frac{\mathrm{d}g}{\mathrm{d}x} & h_{22}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2} & h_{23}\frac{\mathrm{d}^3 g}{\mathrm{d}x^3} & \ldots \\ & h_{31}\frac{\mathrm{d}g}{\mathrm{d}x} & h_{32}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2} & h_{33}\frac{\mathrm{d}^3 g}{\mathrm{d}x^3} & \ldots \\ & \vdots & \vdots & \vdots & \end{array} \tag{5.22}$$

The order of the summation in (5.21) is by rows. Since f is parabolic, the Taylor expansion in (5.21) converges in the formal topology. Moreover, we assumed formal convergence of the differential expansion of $H^k.g$. Therefore, the order of the terms increases indefinitely along the rows and along the columns of (5.22). The monomials up to some fixed order exist only in finitely many first rows and columns. Consequently, we are allowed to change the order of the summation from summation by rows to summation by columns, and the following sum converges in $\mathcal{L}$ in the formal topology:

$$\sum_{\ell=1}^{\infty}\Big(\sum_{i=1}^{\infty} h^k_{i\ell}\frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell}\Big) = \sum_{\ell=1}^{\infty} h^{k+1}_\ell \frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell}. \tag{5.23}$$

Here, $h^{k+1}_\ell := \sum_{i=1}^{\infty} h^k_{i\ell}$. By increasing orders, we immediately obtain $h^{k+1}_\ell \in \mathcal{L}$. The difference of partial sums of (5.21) and (5.23) converges to zero in the formal topology, so the limit of both series is the same, that is, $H^{k+1} \cdot g$. Thus we have:

$$H^{k+1} = \sum_{\ell=1}^{\infty} h^{k+1}_\ell \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell},\ h^{k+1}_\ell \in \mathcal{L},$$

with the formal convergence.

Note additionally that from (5.20) we have that $h^k_1 = H^k \cdot \mathrm{id}$, $h^k_2 = \frac{1}{2}H^k \cdot x^2 - xh^k_1$, etc. by induction, $k \in \mathbb{N}$. The finitely generated case follows directly. □

Lemma 5.21. *Let $f \in \mathcal{L}^0$ (resp. $\mathcal{L}^0_D$) be parabolic. Let $F = \mathrm{iso}(f) \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$) and $H = F - \mathrm{Id}$. Let $X = \log F = \log(\mathrm{Id} + H)$. Then X can be written as a well-defined formal differential operator on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$):*

$$X = \sum_{\ell=1}^{\infty} h_\ell \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell},\ h_\ell \in \mathcal{L}\ (\textit{resp. } \mathcal{L}_{\mathfrak{D}}). \tag{5.24}$$

Proof of Step (1). By Lemma 5.20, all operators H^k, $k \in \mathbb{N}$, can be written as differential operators, with convergence in the formal topology in $\mathcal{L}$. By Lemma 5.8, the operator $X \cdot g$ given by the logarithmic series

$$X \cdot g = \log(\mathrm{Id} + H) \cdot g = H \cdot g - \frac{1}{2}H^2 \cdot g + \frac{1}{3}H^3 \cdot g + \cdots \tag{5.25}$$

also converges in the formal topology. We put the convergent expansions (5.20) for $H^k \cdot g$ in (5.25). Due to the formal convergence of all series, proceeding exactly as in Lemma 5.20, we are allowed to change the order of the summation *from rows to columns*, that is, to group together the terms in front of the same derivative of g. The new sum again converges formally to the same limit:

$$\begin{aligned} X \cdot g &= H \cdot g - \frac{1}{2}H^2 \cdot g + \frac{1}{3}H^3 \cdot g + \cdots = \\ &= \left(h_1^1 g' + h_2^1 g'' + h_3^1 g''' + \cdots\right) + \left(h_1^2 g' + h_2^2 g'' + h_3^2 g''' + \cdots\right) \\ &\quad + \left(h_1^3 g' + h_2^3 g'' + h_3^3 g''' + \cdots\right) + \cdots = \\ &= \left(h_1^1 + h_1^2 + h_1^3 + \cdots\right) \cdot g' + \left(h_2^1 + h_2^2 + h_2^3 + \cdots\right) \cdot g'' \\ &\quad + \left(h_3^1 + h_3^2 + h_3^3 + \cdots\right) \cdot g''' + \cdots = \sum_{\ell=1}^{\infty} h_\ell g^{(\ell)}. \end{aligned}$$

Here, $h_\ell := \sum_{k=1}^{\infty} h_\ell^k \in \mathcal{L}$, since the orders of the terms increase indefinitely. □

Proof of Step (2). We now finish the proof of Theorem B. By Lemma 5.21, we have that X is a formal differential operator:

$$X = \sum_{\ell=1}^{\infty} h_\ell \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}, \; h_\ell \in \mathcal{L} \text{ (resp. } \mathcal{L}_{\mathfrak{D}}). \tag{5.26}$$

We now prove that, due to the Leibniz's property of X (X is a derivation by Proposition 5.9), all the h_ℓ except h_1 *vanish*. We apply (5.26) successively to *test monomials* x^n, $n \in \mathbb{N}$, and use the Leibniz's rule. We deduce from (5.26) applied to $g = \mathrm{id}$ that:

$$h_1 = X \cdot \mathrm{id}.$$

We then apply (5.26) to $g(x) = x^2$. It follows from Leibniz's rule that:

$$X \cdot x^2 = 2x h_1(x) = h_1(x) \cdot 2x + 2h_2(x).$$

It follows that $h_2 \equiv 0$, and, by induction, that $h_i \equiv 0$, $i \geq 2$. Putting $\xi := h_1$, (5.26) becomes:

$$X = \xi \frac{\mathrm{d}}{\mathrm{d}x}, \; \xi \in \mathcal{L},$$

which is the desired vector field. By Lemma 5.8, we have:

$$\exp(X) \cdot \mathrm{id} = \exp(\log F) \cdot \mathrm{id} = F \cdot \mathrm{id} = f,$$

so f is the time-one map of X in the sense of Definition 1.2. The finitely generated case follows easily. □

Proof of Step (3). Let $X = \xi \frac{d}{dx}$, $\xi \in \mathcal{L}$, be any vector field in whose $\mathcal{C}^1$-flow f embeds. Since f is parabolic, it follows from Proposition 5.13 that $(1, 0) \prec \mathrm{ord}(\xi)$. By Proposition 5.11, $\exp(tX) \cdot \mathrm{id}$ defines a $\mathcal{C}^1$-flow of X. Since by Proposition 5.13 the $\mathcal{C}^1$-flow of X is unique, it follows that

$$f = \exp(X) \cdot \mathrm{id}.$$

By Corollary 5.19, it now necessarily follows that

$$\exp(X) \cdot h = h \circ f = F \cdot h, \;\; h \in \mathcal{L},$$

so $\exp(X) = F$. By Lemma 5.15, X is uniquely given by $X = \log F$. □

5.4. Theorem B in the hyperbolic case

Theorem *(Precise form of Theorem B for hyperbolic elements). Let $f \in \mathcal{L}^H$ (*resp. *$f \in \mathcal{L}^H_{\mathfrak{D}}$) be hyperbolic. Then there exists a unique formal vector field on $\mathbb{R}$,*

$$X = \xi \frac{\mathrm{d}}{\mathrm{d}x}, \;\; \xi \in \mathcal{L} \;(\textit{resp. } \xi \in \mathcal{L}_{\mathfrak{D}}),$$

such that f embeds in the flow of X as its time-one map in the sense of Definition 1.2. Moreover,

$$f = \exp(X) \cdot \mathrm{id}.$$

Here, $\mathrm{ord}(\xi) = (1, 0)$ *and* $\exp(X)$ *is a weakly well-defined operator in $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$).*

Let $f(x) = \lambda x + \text{h.o.t.} \in \mathcal{L}^H$, $\lambda > 0$, $\lambda \neq 1$. In the proof of the theorem we suppose without loss of generality that f is a formal contraction, that is $0 < \lambda < 1$. If $\lambda > 1$ (a formal expansion), we consider its formal inverse $f^{-1} \in \mathcal{L}^H$, which is a formal contraction. Obviously, f^{-1} embeds in the flow of X (in the sense of Theorem B) if and only if f embeds in the flow of $-X$.

In the previous section, the standard notion of a *small operator* was used to prove the convergence of operator power series in $L(\mathcal{L})$, see Proposition 5.5. The convergence of the series was in *formal topology* on $\mathcal{L}$. In the parabolic case, this notion was sufficient to obtain the embedding result of Theorem B. Here, we introduce the definition of *small operator in the weak sense* with the aim of giving a meaning to operator power series in $L(\mathcal{L})$ under weaker assumptions, needed for case. The convergence of operator power series will be in the *weak topology* on $\mathcal{L}$.

Definition 5.22 *(Small operator* in the weak sense *with respect to a sequence).* An operator $B : \mathcal{L} \to \mathcal{L}$ is *small in the weak sense with respect to the sequence* $(c_k)_{k \in \mathbb{N}_0}$ *of real numbers* if:

1. there exists a well-ordered set $R \subseteq \mathbb{R}_{\geq 0} \times \mathbb{Z}$ of exponents *equal to* or strictly bigger than $(0,0)$ such that $\mathcal{S}(B \cdot f) \subseteq \mathcal{S}(f) + R$, for every $f \in \mathcal{L}$;
2. For every $f \in \mathcal{L}$ and for every $(\alpha, m) \in \mathcal{S}(f) + \langle R \rangle$, there exists a sequence $(C^k_{\alpha,m})_{k \in \mathbb{N}_0}$ of strictly positive numbers such that:

$$\left|[B^k \cdot f]_{\alpha,m}\right| \leq C^k_{\alpha,m}, \tag{5.27}$$

   and such that the series

$$\sum_{k=0}^{\infty} c_k C^k_{\alpha,m} \tag{5.28}$$

   converges absolutely. Here, $[B^k \cdot f]_{\alpha,m}$ denotes the *coefficient* of monomial $x^\alpha \boldsymbol{\ell}^m$ in $B^k \cdot f$ (the notation introduced in Section 4.2).

An operator $B : \mathcal{L}_{\mathcal{D}} \to \mathcal{L}_{\mathcal{D}}$ is *small in the weak sense with respect to the sequence* $(c_k)_k$ if the set R is in addition of finite type.

Notice that from (1) we obtain by induction that $\mathcal{S}(B^k \cdot f) \subseteq \mathcal{S}(f) + \langle R \rangle$, $k \in \mathbb{N}$, so (2) makes sense.

Proposition 5.23 *(A version of Proposition 5.4 in the weak sense). Let $(c_k)_{k \in \mathbb{N}_0}$ be a sequence of real numbers and let $B \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_D)$) be a small operator in the weak sense with respect to the sequence (c_k). Then an operator $B \in L(\mathcal{L})$ $(L(\mathcal{L}_D))$ is* weakly well-defined *by the series*

$$B := \sum_{k=0}^{\infty} c_k B^k. \tag{5.29}$$

Proof. The proof is straightforward by (5.27) and the absolute convergence of the series $\sum_{k=0}^{\infty} c_k C^k_{\alpha,m}$. □

We state and prove in this section the analogous of Lemmas 5.8, 5.15, 5.20 and 5.21. We prove Proposition 5.12. All of them are needed for the proof of Theorem B in the hyperbolic case. Then the proof in the hyperbolic case follows the same steps as the proof in the parabolic case, but using the corresponding weak notions.

Lemma 5.24 *(Lemma 5.8 for hyperbolic elements). Let $f = \lambda x + \text{h.o.t.} \in \mathcal{L}^H$ (resp. $f \in \mathcal{L}^H_{\mathfrak{D}}$) be a hyperbolic contraction $(0 < \lambda < 1)$. Let $F = \text{iso}(f) \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$), as in Remark 5.7. Then $\log F$ is a* weakly well-defined *operator in $L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$).*

Proof. Put $F = \text{Id} + H$, $H : \mathcal{L} \to \mathcal{L}$. Then

$$\log F = \log(\text{Id} + H) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{H^k}{k}. \tag{5.30}$$

We prove that operator $H \in L(\mathcal{L})$ is small in the weak sense (see Definition 5.22) with respect to the sequence $\left(\frac{(-1)^{k+1}}{k}\right)_k$. Applying Proposition 5.23, we conclude that the series $\log F$ is a weakly well-defined operator $\log F : \mathcal{L} \to \mathcal{L}$.

Take $g \in \mathcal{L}$. Then $H \cdot g = g \circ f - g = g\big(\lambda x + \psi(x)\big) - g$. For $x^\alpha \boldsymbol{\ell}^m \in \mathcal{S}(g)$, we compute:

$$\begin{aligned}
&(\lambda x + \psi(x))^\alpha \boldsymbol{\ell}(\lambda x + \psi(x))^m - x^\alpha \ell^m = \\
&\quad = \lambda^\alpha x^\alpha \boldsymbol{\ell}^m \Big(1 + \lambda^{-1} x^{-1} \psi(x)\Big)^\alpha \Big(1 - \log \lambda \boldsymbol{\ell} - \boldsymbol{\ell} \log\big(1 + \lambda^{-1} x^{-1} \psi(x)\big)\Big)^{-m} - x^\alpha \ell^m.
\end{aligned} \tag{5.31}$$

Here, $\psi(x) = f(x) - \lambda x$, with $(1, 0) \prec \operatorname{ord}(\psi)$. We conclude that, for every $g \in \mathcal{L}$,

$$\mathcal{S}(H \cdot g) \subseteq \mathcal{S}(g) + R,$$

where R is a sub-semigroup of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$ generated by $(0, 1)$ and $(\beta - 1, \ell)$ for $(\beta, \ell) \in \mathcal{S}(\psi)$ and containing $(0, 0)$. Obviously, since $(1, 0) \prec \operatorname{ord}(\psi)$, all the elements of R except $(0, 0)$ are of order strictly bigger than $(0, 0)$. By Neumann's lemma, R is well-ordered.

Now take any $(\alpha, m) \in R + \mathcal{S}(g)$. If $(\alpha, m) \notin \mathcal{S}(H^k.g)$, for any $k \in \mathbb{N}$, then (5.27) holds for every $C_{\alpha,m} > 0$. Suppose that there exists some $\ell \in \mathbb{N}$ such that $(\alpha, m) \in \mathcal{S}(H^\ell.g)$. It can be seen from (5.31), since $\lambda \neq 1$, that (α, m) then appears in every $H^k \cdot g$, $k \geq \ell$ (unlike the parabolic case). We prove nevertheless that the coefficient of $x^\alpha \boldsymbol{\ell}^m$ evolves *controllably* with k in the sense of (5.27) and (5.28).

To this end, we analyze in what ways we can obtain the monomial $x^\alpha \boldsymbol{\ell}^m$ in iterates $H^k \cdot g$, $k \in \mathbb{N}$. The monomial $x^\alpha \boldsymbol{\ell}^m$ *evolves* from some *initial monomial* $x^\beta \boldsymbol{\ell}^n \in \mathcal{S}(g)$ through iterates $H \cdot g$, $H^2 \cdot g$, etc. The evolution of coefficients and exponents from one iterate $H^k \cdot g$ to the next one $H^{k+1} \cdot g$ is described by (5.31). Hence, the rule which governs this evolution is the following: *in each step, the corresponding monomial $x^\gamma \boldsymbol{\ell}^n$ either stays the same while its coefficient is multiplied by $\lambda^\gamma - 1$, or it is multiplied by a monomial from $R \setminus \{(0, 0)\}$*. We can thus write a finite *chain of changes* corresponding to this evolution:

$$x^\beta \boldsymbol{\ell}^n \to x^{\beta+\gamma_1} \boldsymbol{\ell}^{n+n_1} \to x^{\beta+\gamma_1+\gamma_2} \boldsymbol{\ell}^{n+n_1+n_2} \to \cdots \to x^{\beta+\gamma_1+\cdots+\gamma_r} \boldsymbol{\ell}^{n+n_1+\cdots+n_r} = x^\alpha \boldsymbol{\ell}^m, \tag{5.32}$$

where $(\beta, n) \in \mathcal{S}(g)$, $(\gamma_i, n_i) \in R \setminus \{(0, 0)\}$, $i = 1, \ldots, r$, $r \in \mathbb{N}_0$. By Neumann's lemma, for every $(\alpha, m) \in R + \mathcal{S}(g)$, there exist only *finitely many* chains describing the evolution of monomial $x^\alpha \ell^m$ from elements of $\mathcal{S} \cdot g$. Given a chain as in (5.32), the integer $r \in \mathbb{N}_0$ as well as the pairs (β, n) and (γ_i, n_i), $i = 1, \ldots, r$, do not depend on k. We fix one such chain (5.32) and prove (5.27) and (5.28) only for this chain (in the end we sum up the coefficients of finitely many chains contributing to $x^\alpha \boldsymbol{\ell}^m$ and conclude for the whole $x^\alpha \boldsymbol{\ell}^m$). For the coefficient of $x^\alpha \boldsymbol{\ell}^m$ obtained by this chain in $H^k \cdot g$, we obtain, using (5.31):

$$\left|[H^k \cdot g]_{\alpha,m}\right| = \binom{k}{r} \cdot a(\lambda^\beta - 1)^{k_1} \cdot C_1 \cdot (\lambda^{\beta+\gamma_1} - 1)^{k_2} \cdot C_2 \cdots (\lambda^{\beta+\gamma_1+\cdots+\gamma_{r-1}} - 1)^{k_r} \cdot C_r \cdot (\lambda^\alpha - 1)^{k_{r+1}}, \quad k \geq k_0.$$

Here, k_0 is the index of the first iterate $H^{k_0} \cdot g$ in which $x^\alpha \boldsymbol{\ell}^m$ obtained by chain (5.32) appears, and $a \in \mathbb{R}$ is the coefficient of the initial monomial, $ax^\beta \boldsymbol{\ell}^n$. We choose r iterates (out of k in total) in which the monomial changes: $H^{k_1} \cdot g$, $H^{k_1+k_2} \cdot g, \ldots, H^{k_1+\cdots+k_r} \cdot g$. Note also that $k_1 + \cdots + k_r + k_{r+1} = k - r$. Note also that the change of the coefficient in r steps in which the monomial changes depends only on the fixed chain and not on k. That is why, in the above formula, we multiply by numbers $C_1, \ldots, C_r$ which depend on f and on the given chain, but not on k. The remaining $(k - r)$ steps are characterized by multiplications by respective $\lambda^\gamma - 1$, $k - r$ times in total. Therefore,

$$\left|[H^k \cdot g]_{\alpha,m}\right| \leq \binom{k}{r} \cdot C_f^1 \cdot |\lambda^\alpha - 1|^{k-r} \leq \binom{k}{r} C_f |\lambda^\alpha - 1|^k, \quad k \geq k_0. \tag{5.33}$$

Here, C_f^1, C_f are *constants* that depend only on initial f, on g and on the fixed chain; they are independent of k.

For the given chain (5.32) contributing to the coefficient of $x^\alpha \boldsymbol{\ell}^m$, the inequality (5.27) in Definition 5.22 is, by (5.33), satisfied with

$$C_{\alpha,m}^k = \binom{k}{r} \cdot C_f \cdot |\lambda^\alpha - 1|^k, \quad k \geq k_0.$$

Moreover, the series

$$\sum_{k=k_0}^{\infty} \frac{(-1)^{k+1}}{k} C_{\alpha,m}^k = \sum_{k=k_0}^{\infty} \frac{(-1)^{k+1}}{k} \binom{k}{r} \cdot C_f \cdot |\lambda^\alpha - 1|^k$$

converges absolutely (which can easily be checked by, for example, the ratio test) since $0 < \lambda < 1$. The operator $H \in L(\mathcal{L})$ is therefore *small in the weak sense* with respect to the sequence $\left(\frac{(-1)^{k+1}}{k}\right)_k$. By Proposition 5.23, the series $(\log F) \cdot g$, $g \in \mathcal{L}$, converges in $\mathcal{L}$ in the weak topology to an element of $\mathcal{L}$. The operator $\log F \in L(\mathcal{L})$ is thus weakly well-defined.

Finally, the claim in finitely generated case $(\mathcal{L}_{\mathfrak{D}})$ follows easily, since R is then finitely generated. □

We prove here the Proposition 5.12 stated in Subsection 5.2. The problem consists in giving a meaning to the formal time-one map of a vector field $X = \xi \frac{\mathrm{d}}{\mathrm{d}x}$ in the case where $\xi = \lambda x + \text{h.o.t.}$ is *hyperbolic*. The question is: does X admit a formal one-parameter flow? Hence we need to study the convergence in $\mathcal{L}$ of the exponential of $X = \xi \frac{\mathrm{d}}{\mathrm{d}x}$. The problem, compared to the case $(1, 0) \prec \operatorname{ord}(\xi)$ proven in Subsection 5.2, is that the operator $X = \xi \frac{\mathrm{d}}{\mathrm{d}x}$ is not *small* any more if $\operatorname{ord}(\xi) = (1, 0)$. Hence, its exponential

$H^t = \exp(tX)$, $t \in \mathbb{R}$, is not a well-defined operator in $L(\mathcal{L})$. Moreover, given $f \in \mathcal{L}$, the formula:

$$H^t \cdot f = \exp(tX) \cdot f = f + tf'\xi + \frac{t^2}{2!}(f'\xi)'\xi + \frac{t^3}{3!}\Big((f'\xi)'\xi\Big)'\xi + \cdots \tag{5.34}$$

does not converge in $\mathcal{L}$ even with respect to the weaker product topology with respect to the discrete topology. Indeed, a fixed monomial is present in infinitely many terms of the sum (5.34). Nevertheless, we show here that operator X is *small in the weak sense with respect to the sequence* $\left(\frac{t^n}{n!}\right)_n$, see Definition 5.22. Applying Proposition 5.23, we conclude that $H^t = \exp(tX)$ is a *weakly well-defined* operator in $L(\mathcal{L})$. In other words, the series (5.34) converges in $\mathcal{L}$ in the weak topology.

Proof of Proposition 5.12. Let $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$, with

$$\xi = \lambda \cdot \mathrm{id} + \psi,\ \ \lambda \neq 0.$$

Here, $\psi \in \mathcal{L}$ such that $(1,0) \prec \mathrm{ord}(\psi)$. Recall that

$$H^t \cdot f = \sum_{n=0}^{\infty} \frac{t^n}{n!} X^n \cdot f.$$

Since $X \cdot f = \xi f'$, $f \in \mathcal{L}$, we conclude that $\mathcal{S}(X \cdot f) \subseteq \mathcal{S}(f) + R$, where R is a sub-semigroup of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$ generated by $(0,1)$, $(\beta-1,k)$ for $(\beta,k) \in \mathcal{S}(\psi)$, and containing $(0,0)$. Note that $(0,0) \prec \mathrm{ord}(\beta-1,k)$, since $(1,0) \prec \mathrm{ord}(\psi)$. Each $X^n \cdot f$ is by (5.34) a sum of 2^n terms of the type

$$\big(((f' \cdot *)' \cdot *)' \ldots\big)' \cdot *, \tag{5.35}$$

with n stars representing either λx or ψ.

Fix any $(\alpha,m) \in \cup_{n\in\mathbb{N}_0}\mathcal{S}(X^n \cdot f) \subseteq \mathcal{S}(f) + R$. Then $(\alpha,m) \in \mathcal{S}(X^{n_0} \cdot f)$, for some $n_0 \in \mathbb{N}$. Since f is hyperbolic, we see by (5.35) that $(\alpha,m) \in \mathcal{S}(X^n \cdot f)$, for infinitely many $n \geq n_0$ in general (the coefficient of $x^\alpha \boldsymbol{\ell}^m$ may sometimes vanish due to cancellations). In order to prove (5.27) and (5.28), we analyze the coefficient of $x^\alpha \boldsymbol{\ell}^m$ in $X^n \cdot f$, $n \in \mathbb{N}$, $n \geq n_0$.

Let $n \geq n_0$. The fixed monomial $x^\alpha \boldsymbol{\ell}^m$ in $X^n \cdot f$ is obtained in the course of iterates $X^\ell \cdot f$, $0 \leq \ell \leq n$, from some *initial monomials* $ax^\beta \boldsymbol{\ell}^p \in \mathcal{S}(f)$, $a \in \mathbb{R}$, which evolve in n steps of iteration to $x^\alpha \boldsymbol{\ell}^m$. By (5.35), we see that in each step ℓ we differentiate the respective monomial and then multiply by either a monomial from ψ or by λx. After say k multiplications by monomials from ψ and the remaining $n-k$ multiplications by λx, each following one differentiation, the initial monomial $x^\beta \boldsymbol{\ell}^p \in \mathcal{S}(f)$ transforms to:

$$x^{\beta+(\alpha_1-1)+\cdots+(\alpha_k-1)}\boldsymbol{\ell}^{p+p_1+\cdots+p_k+r},$$

where $x^{\alpha_i}\boldsymbol{\ell}^{p_i} \in \mathcal{S}(\psi)$, $i = 1, \dots, k$, and $r \in \mathbb{N}_0$, $0 \le r \le n$ (r corresponding to the number of differentiations of the logarithm part). In order to obtain all chains of changes of monomials resulting in $x^\alpha\boldsymbol{\ell}^m \in \mathcal{S}(X^n \cdot f)$, whose coefficients then add up to the coefficient of $x^\alpha\boldsymbol{\ell}^m$ in $X^n \cdot f$, we search for all $(\beta, p) \in \mathcal{S}(f)$, $k, r \in \mathbb{N}_0$ and $(\alpha_i, p_i) \in \mathcal{S}(\psi)$, $i = 1, \dots, k$, such that:

$$x^{\beta+(\alpha_1-1)+\cdots+(\alpha_k-1)}\boldsymbol{\ell}^{p+p_1+\cdots+p_k+r} = x^\alpha\boldsymbol{\ell}^m.$$

By Neumann's lemma, there are only *finitely many* such choices.

For any such choice (of finitely many), put $L := (k+1) \cdot \max_{i=1\dots k}\{|p|, |p_i|\} + r$. Its contribution to the coefficient of $x^\alpha\boldsymbol{\ell}^m$ in $X^n \cdot f$ is absolutely bounded by:

$$\big|[H^n.f]_{\alpha,m}\big| \le a \cdot \alpha^{n-r}L^r\binom{n}{k}\lambda^{n-k} \le C_{\xi,f} \cdot \binom{n}{k}(\alpha\lambda)^n,\ n \ge N.$$

Here, N is the smallest iterate $X^N \cdot f$ containing $x^\alpha\ell^m$ obtained in this chain, and $C_{\xi,f} > 0$ is a coefficient depending on the coefficients of ξ and f and on the chosen chain, but independent of n. The term $\alpha^{n-r}L^r$ comes from differentiating n times the initial monomial $x^\beta\boldsymbol{\ell}^p$. The term $\binom{n}{k}\lambda^{n-k}$ comes from $n-k$ multiplications by λx. For the given chain, we put $C^n_{\alpha,m} = C_{\xi,f} \cdot \binom{n}{k}(\alpha\lambda)^n$, $n \ge N$. The series

$$\sum_{n=N}^{\infty}\frac{t^n}{n!}C^n_{\alpha,m} = \sum_{n=N}^{\infty}\frac{t^n}{n!}C_{\xi,f} \cdot \binom{n}{k}(\alpha\lambda)^n$$

converges absolutely. Summing contributions to the coefficient of $x^\alpha\ell^m$ of all (finitely many) possible chains, we conclude the same for the absolute convergence of the whole coefficient of $x^\alpha\ell^m$. By Proposition 5.23, the operator $\exp(tX) : \mathcal{L} \to \mathcal{L}$ is weakly well-defined.

The final statements are proven in the same way as in the proof of Proposition 5.11 from Section 5.2. Finally, the finitely generated case follows easily, since the semigroup R is then finitely generated. □

Lemma 5.25 *(Lemma 5.20 in the hyperbolic case). Let $f \in \mathcal{L}^0$ (resp. $\mathcal{L}^0_{\mathfrak{D}}$) be a hyperbolic contraction. Let $F = \mathrm{iso}(f) \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$) and let $H = F - \mathrm{Id}$. Then all the iterates H^k can be written as weakly well-defined formal differential operators on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$):*

$$H^k = \sum_{\ell=1}^{\infty} h^k_\ell \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell},\ h^k_\ell \in \mathcal{L}\ (resp.\ \mathcal{L}_D),\ k \in \mathbb{N}; \tag{5.36}$$

Proof. Let $f = \lambda \cdot \mathrm{id} + \psi$, $\psi \in \mathcal{L}$, $\mathrm{ord}(\psi) \succ (1,0)$ and $0 < \lambda < 1$. Let $h = f - \mathrm{id} = (\lambda-1)\cdot\mathrm{id}+\psi$. Note that $\mathrm{ord}(h) = (1,0)$. The proof is by induction. It is a more elaborate version of the proof of Lemma 5.21 in the parabolic case. The induction basis ($k = 1$)

follows easily by Taylor expansion, which we have proven to converge to $H.g \in \mathcal{L}$ in the weak topology:

$$H \cdot g = g(x+h) - g(x) = \sum_{\ell=1}^{\infty} \frac{h^\ell}{\ell!} \frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell}, \ g \in \mathcal{L}. \tag{5.37}$$

Thus, $H = \sum_{\ell=1}^{\infty} h_\ell^0 \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}$, with the coefficients $h_\ell^0 := \frac{h^\ell}{\ell!} \in \mathcal{L}$, $\ell \in \mathbb{N}$. Note that, unlike the parabolic case, all summands of the series are *of the same order* $\mathrm{ord}(g)$. For every monomial of Taylor expansion (5.37), the series of its coefficients converges *absolutely*. Assume now that operators H^m, $m \le k$, can be written in the form (5.36) of a differential operator, where the series converges in the weak topology on $\mathcal{L}$. That is, if a monomial appears in infinitely many summands, the series of its coefficients is convergent. Suppose additionally that the series of coefficients of every monomial in the expansion (5.36) converges *absolutely*.

The induction step: we prove that the operator H^{k+1} can be written in the differential form (5.37). Moreover, we prove the *absolute* convergence of series of coefficients of every monomial of the support of $H^{k+1} \cdot g$ in this formula. By Taylor expansion, we obtain:

$$\begin{aligned} H^{k+1} \cdot g\,(x) &= H(H^k \cdot g)\,(x) = H^k \cdot g\,(x+h(x)) - H^k \cdot g\,(x) \\ &= \sum_{i=1}^{\infty} \frac{h(x)^i}{i!} \frac{\mathrm{d}^i (H^k \cdot g)}{\mathrm{d}x^i} \\ &= \sum_{i=1}^{\infty} \frac{h^i}{i!} \frac{\mathrm{d}^i}{\mathrm{d}x^i} \Big(\sum_{\ell=1}^{\infty} h_\ell^k \frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell} \Big) = \sum_{i=1}^{\infty} \Big(\sum_{\ell=1}^{\infty} h_{i\ell}^k \frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell} \Big), \end{aligned} \tag{5.38}$$

with $h_{i\ell}^k \in \mathcal{L}$, $i, \ell \in \mathbb{N}$. The elements of the double sum (5.21) can be represented by the grid:

$$\begin{array}{r|cccc} H^k \cdot g & h_1^k g' & h_2^k g'' & h_3^k g''' & \dots \\ \hline h \frac{\mathrm{d}}{\mathrm{d}x} (H^k \cdot g) & h \frac{\mathrm{d}}{\mathrm{d}x} (h_1^k g') & h \frac{\mathrm{d}}{\mathrm{d}x} (h_2^k g'') & \dots & \\ \frac{h^2}{2!} \frac{\mathrm{d}^2}{\mathrm{d}x^2} (H^k \cdot g) & \frac{h^2}{2!} \frac{\mathrm{d}^2}{\mathrm{d}x^2} (h_1^k g') & \frac{h^2}{2!} \frac{\mathrm{d}^2}{\mathrm{d}x^2} (h_2^k g'') & \dots & \\ \frac{h^3}{3!} \frac{\mathrm{d}^3}{\mathrm{d}x^2} (H^k \cdot g) & & & & \\ \vdots & \vdots & \vdots & \vdots & \end{array} \tag{5.39}$$

Unlike the parabolic case, the order of the terms *stays the same* along the rows and along the columns, so that one monomial from the support may appear in every term of every row and of every column. The order of the summation in (5.38) is by rows. The double sum (5.38) converges in this order of the summation in the weak topology on $\mathcal{L}$ (by assumption and by convergence of Taylor expansions). In the sequel, we prove that the coefficient of a fixed monomial of the support of $H^{k+1}.g$ converges (to the same limit) if we change the order of summation from the summation by rows to the summation by columns. Since each derivative of g appears only in finitely many first columns, we have thus proven that the following sum converges in $\mathcal{L}$ (in the weak topology), to the same limit $H^{k+1}\cdot g$:

$$H^{k+1}\cdot g=\sum_{\ell=1}^{\infty}\Big(\sum_{i=1}^{\infty}h_{i\ell}^{k}\frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell}\Big)=\sum_{\ell=1}^{\infty}h_{\ell}^{k+1}\frac{\mathrm{d}^\ell g}{\mathrm{d}x^\ell}. \tag{5.40}$$

We also have that $h_\ell^{k+1}:=\sum_{i=1}^{\infty}h_{i\ell}^{k}\in\mathcal{L}$.

We use the following version of the *Moore–Osgood theorem* (see for example [16, Theorem 8.3]): given a real double sequence $a_{m,n}$, if the sum

$$\sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}}|a_{m,n}|$$

converges, then the following iterated sums exist and commute:

$$\sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}}a_{m,n}=\sum_{n\in\mathbb{N}}\sum_{m\in\mathbb{N}}a_{m,n}.$$

Therefore, in order to prove the step of the induction, we need to prove that the double sum of absolute values of coefficients of every monomial of (5.38) converges in this order of the summation.

By the induction assumption, the convergence of series of coefficients of every monomial along the first row in (5.39) is absolute. Fix a monomial $x^\alpha\boldsymbol{\ell}^m$ from the support of (5.38). We prove here that the convergence of its respective coefficient along every other row is absolute, and that these limits converge by columns. By the *Moore–Osgood theorem* stated above, this will prove the step of the induction.

Note that $\mathcal{S}(h_j^k g^{(j)})$, $\mathcal{S}\big(\frac{h^\ell}{\ell!}\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}(h_j^k g^{(j)})\big)\subseteq\mathcal{S}(g)+R$, for every $j,\ell\in\mathbb{N}$. Here, R is a sub-semigroup generated by $(0,1)$, $(\beta-1,p)$ for $(\beta,p)\in\mathcal{S}(\psi)$, and containing $(0,0)$. We denote, for a monomial $x^\alpha\boldsymbol{\ell}^m\in\mathcal{S}(g)+R$,

$$c_j^0(\alpha,m):=[h_j^k g^{(j)}]_{\alpha,m},\quad c_j^\ell(\alpha,m):=\Big[\frac{h^\ell}{\ell!}\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}(h_j^k g^{(j)})\Big]_{\alpha,m},\quad j,\ell\in\mathbb{N}.$$

We prove that

$$\sum_{\ell\in\mathbb{N}}\sum_{j\in\mathbb{N}}|c_j^\ell(\alpha,m)|<\infty,\ (\alpha,m)\in\mathcal{S}(g)+R.$$

In order to bound $|c_j^\ell(\alpha,m)|$, note that the monomial $x^\alpha\boldsymbol{\ell}^m\in\mathcal{S}\big(\frac{h^\ell}{\ell!}\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}(h_j^k g^{(j)})\big)$, $\ell\in\mathbb{N}$, is obtained from some *initial* monomial $b_j^0x^\beta\boldsymbol{\ell}^n\in\mathcal{S}(h_j^k g^{(j)})$, $b_j^0=[h_j^k g^{(j)}]_{\beta,n}$, undergoing ℓ differentiations and the multiplication by

$$\frac{h^\ell}{\ell!}=\frac{1}{\ell!}\big((\lambda-1)x+\psi(x)\big)^\ell=\frac{(\lambda-1)^\ell}{\ell!}x^\ell\Big(1+x^{-1}\frac{\psi(x)}{\lambda-1}\Big)^\ell.$$

We obtain:

$$b_j^0 \cdot \beta(\beta-1)\cdots(\beta-(\ell-r)+1)\cdot n(n+1)\cdots(n+r-1)x^{\beta-\ell}\boldsymbol{\ell}^{n+r}\cdot \\ \cdot \frac{(\lambda-1)^\ell}{\ell!}x^\ell\binom{\ell}{s}(\lambda-1)^{-s}b_1x^{\beta_1-1}\boldsymbol{\ell}^{p_1}\cdots b_sx^{\beta_s-1}\boldsymbol{\ell}^{p_s} =_\star x^\alpha\boldsymbol{\ell}^m. \tag{5.41}$$

Here, $0 \le r \le \ell$ is the number of derivatives applied on the logarithmic components, $s \in \mathbb{N}_0$, $0 \le s \le \ell$, $b_ix^{\beta_i}\boldsymbol{\ell}^{p_i} \in \mathcal{S}(\psi)$, $i = 1\ldots s$. By Neumann's lemma, there are only finitely many choices for $r, s \in \mathbb{N}_0$, $(\beta, n) \in \mathcal{S}(g) + R$, $(\beta_i, p_i) \in \mathcal{S}(\psi)$, $i = 1, \ldots, s$. The choices are independent of ℓ, k, j. We analyze here only one of the combinations (afterwards, we sum up finitely many bounds to obtain a bound on the whole coefficient $c_j^\ell(\alpha, m)$):

$$|c_j^\ell(\alpha,m)| \le C\cdot|c_j^0(\beta,n)|\cdot\left|\binom{\beta}{\ell-r}\right|\cdot\binom{\ell}{s}\cdot\frac{(1-\lambda)^{\ell-s}}{\ell!},$$

where $C \ge 0$ depends only on the given combination (independent of j, ℓ or k). Therefore,

$$\sum_{\ell\ge r,s}\sum_{j\in\mathbb{N}}|c_j^\ell(\alpha,m)| \le C\cdot\sum_{\ell\ge r,s}\left|\binom{\beta}{\ell-r}\right|\binom{\ell}{s}\cdot\frac{(1-\lambda)^{\ell-s}}{\ell!}\sum_{j\in\mathbb{N}}|c_j^0(\beta,n)|.$$

By the induction assumption, $C(\beta,n) := \sum_{j\in\mathbb{N}}|c_j^0(\beta,n)| < \infty$. We have now:

$$\sum_{\ell\ge r,s}\sum_{j\in\mathbb{N}}|c_j^\ell(\alpha,m)| \le C\cdot C(\beta,n)\sum_{\ell\ge r,s}\left|\binom{\beta}{\ell-r}\right|\binom{\ell}{s}\frac{(1-\lambda)^{\ell-s}}{\ell!} \\ \le C\cdot C(\beta,n)\sum_{\ell\ge r,s}\left|\binom{\beta}{\ell-r}\right|(1-\lambda)^{\ell-s} < \infty. \tag{5.42}$$

The last sum converges by the ratio test, since $0 < 1-\lambda < 1$. We have thus proven that the summation in (5.39) may be done by columns instead of by rows, while the sum $H^{k+1}.g$ remains the same. The formula (5.36) for $H^{k+1}.g$, $g \in \mathcal{L}$, thus converges in the weak topology. Moreover, in this formula, the series of absolute values of coefficients of every fixed monomial $x^\alpha\boldsymbol{\ell}^m$ converges to (5.42) (more accurately, to a *finite sum of sums of the type* (5.42), each for every possible combination).

Note additionally that from (5.36) we have that $h_1^k = H^k\cdot\mathrm{id}$, $h_2^k = \frac{1}{2}H^k\cdot x^2 - xh_1^k$, etc., for every $k \in \mathbb{N}$. The finitely generated case follows directly. □

Lemma 5.26 *(Lemma 5.21 in the hyperbolic case). Let $f \in \mathcal{L}^0$ (resp. $\mathcal{L}^0_{\mathfrak{D}}$) be a hyperbolic contraction. Let $F = \mathrm{iso}(f) \in L(\mathcal{L})$ (resp. $L(\mathcal{L}_{\mathfrak{D}})$) and $H = F - \mathrm{Id}$. Let $X = \log F =$*

$\log(\mathrm{Id}+H)$. *Then X can be written as a weakly well-defined formal differential operator on $\mathcal{L}$ (resp. $\mathcal{L}_{\mathfrak{D}}$):*

$$X=\sum_{\ell=1}^{\infty}h_\ell\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell},\ h_\ell\in\mathcal{L}\ (\textit{resp. }\mathcal{L}_{\mathfrak{D}}). \tag{5.43}$$

Proof. By Lemma 5.24, the operator $X.g$ is given by the logarithmic series which converges in the weak topology:

$$\begin{aligned}X\cdot g&=\log(\mathrm{Id}+H)\cdot g=H\cdot g-\frac{1}{2}H^2\cdot g+\frac{1}{3}H^3\cdot g+\cdots\\&=\left(h_1^1g'+h_2^1g''+h_3^1g'''+\cdots\right)-\frac{1}{2}\left(h_1^2g'+h_2^2g''+h_3^2g'''+\cdots\right)\\&\quad+\frac{1}{3}\left(h_1^3g'+h_2^3g''+h_3^3g'''+\cdots\right)+\cdots.\end{aligned} \tag{5.44}$$

By Lemma 5.25, all operators H^k, $k\in\mathbb{N}$, can be written as differential operators (5.36), with convergence in the weak topology in $\mathcal{L}$. Thus the *double sum* (5.44) converges to $X.g$ in this order of the summation, in the same topology.

Let us consider a fixed monomial from the support of (5.44), $x^\alpha\ell^m\in\mathcal{S}(g)+R$, see Lemma 5.25. The monomial may appear in every term of the double sum (5.44). By the proof of Lemma 5.25, in each bracket of (5.44), the series of coefficients of $x^\alpha\boldsymbol{\ell}^m$ converges *absolutely*. We denote the absolute limit in the k-th bracket by $A_k>0$. By the *Moore–Osgood theorem* stated in the proof of Lemma 5.25, to prove that we are allowed to change the order of the summation in (5.44), that is, to group the terms in front of every derivative of g, we need to prove the convergence of the sum:

$$\sum_{k\in\mathbb{N}}\frac{A_k}{k}. \tag{5.45}$$

The following argument is similar as in the proof of Proposition 5.8. In the proof of Lemma 5.25, we have described the iterative step in which the $(k+1)$-st bracket is deduced from the k-th bracket of (5.44) (that is, the differential form for $H^{k+1}.g$ from the differential form for $H^k.g$). All monomials from (5.44) belong to $\mathcal{S}(g)+R$. By Neumann's lemma, there are only finitely many ways in which a fixed monomial $x^\alpha\boldsymbol{\ell}^m$ belonging to some bracket of (5.44) is obtained from previous brackets and, initially, from monomials of $\mathcal{S}(g)$, *independently of the bracket*. We adopt the notion of *chains* to describe the evolution of monomials, similarly as in the proof of Lemma 5.8. We fix one (of finitely many chains): an initial monomial $x^{\beta_0}\boldsymbol{\ell}^{n_0}\in\mathcal{S}(g)$ evolves in k steps to $x^\alpha\boldsymbol{\ell}^m$, through (necessarily distinct) monomials $x^{\beta_i}\boldsymbol{\ell}^{p_i}$, $i=1,\ldots,r$, $r\in\mathbb{N}_0$, as described in (5.41):

$$x^{\beta_0}\boldsymbol{\ell}^{n_0}\to x^{\beta_1}\boldsymbol{\ell}^{p_1}\to x^{\beta_2}\boldsymbol{\ell}^{p_2}\to\cdots\to x^{\beta_r}\boldsymbol{\ell}^{p_r}=x^\alpha\boldsymbol{\ell}^m. \tag{5.46}$$

To estimate A_k, $k \geq r$, for this fixed combination, we use the estimate (5.42) from the proof of Lemma 5.25, where the sum of absolute values of coefficients of a monomial in the $(\ell+1)$-st bracket is estimated by the sum of absolute values of coefficients of its corresponding (for the given chain) monomial in the ℓ-th bracket, $\ell \in \mathbb{N}$. Note that the estimate (5.42) is *independent of* ℓ. For the above combination (5.46), in $k-r$ of total k steps the monomial remains the same, and changes in the remaining r steps. Let $C_\ell(\beta_i, p_i)$ denote the sum of absolute values of coefficients of the monomial $x^{\beta_i}\ell^{p_i}$ in the ℓ-th bracket, $\ell \in \mathbb{N}$. By (5.42), in the steps where the respective monomial stays the same (then, $r = s = 0$, $C = 1$), we have the estimate:

$$C_{\ell+1}(\beta_i, p_i) \leq C_\ell(\beta_i, p_i) \cdot \sum_{j \geq 0} \left| \binom{\beta_i}{j} \right| (1-\lambda)^j, \ \ell \in \mathbb{N}.$$

Note that it is sufficient to prove that the terms of (5.44) starting from some fixed derivative can be regrouped as in the statement of the lemma (the terms with first finitely many derivatives form a finite sum of series, so the order of the summation can be changed trivially). Therefore, in the above sum, without loss of generality we can take $j \geq j_0$ instead of $j \geq 0$, for any $j_0 \in \mathbb{N}$. To each chain, we associate a number $0 < A < 1$ and an integer $j_0 \in \mathbb{N}_0$, such that the sum $\sum_{j \geq j_0} \left| \binom{\beta_i}{j} \right| (1-\lambda)^j$ above is bounded by A, for every (β_i, p_i) of the given chain. This follows from the convergence of the series and the fact that there exist only finitely many β_i-s in the given chain. Notice that the constant $0 < A < 1$ depends only on the chain. Therefore, for the steps in which the monomial remains the same, we have:

$$C_{\ell+1}(\beta_i, p_i) \leq A \cdot C_\ell(\beta_i, p_i), \ \ell \in \mathbb{N}, \ 0 < A < 1.$$

On the other hand, there are only finitely many (r) steps (for the given chain) in which the corresponding monomial $x^{\beta_i}\boldsymbol{\ell}^{p_i}$ changes to $x^{\beta_{i+1}}\boldsymbol{\ell}^{p_{i+1}}$, $i = 0 \dots r-1$. By (5.42), we have a simple estimate:

$$C_{\ell+1}(\beta_{i+1}, p_{i+1}) \leq D C_\ell(\beta_i, p_i), \ i = 0, \dots, r-1, \ \ell \in \mathbb{N},$$

where $D > 0$ depends only on the chain. We obtain the estimate:

$$A_k \leq a \cdot D^r A^{k-r} \leq C A^k.$$

Here, $a \in \mathbb{R}$ is the coefficient of $x^{\beta_0}\boldsymbol{\ell}^{p_0}$ in g, and $C > 0$ and $0 < A < 1$ depend only on the chain. The series (5.45) thus converges. Since there are only finitely many chains contributing to the coefficient of $x^\alpha \boldsymbol{\ell}^m$ in (5.44), the result follows. □

Proof of Theorem B in the hyperbolic case. Let $f(x) = \lambda x + \text{h.o.t.} \in \mathcal{L}$, $0 < \lambda < 1$. Let $F = \text{iso}(f) \in L(\mathcal{L})$, see Remark 5.7. By Lemma 5.24, the operator $\log F \in L(\mathcal{L})$ is weakly well-defined. Using Proposition 5.9 and Lemma 5.26, we prove (as in the parabolic case in Section 5.3) that the operator $X = \log F$ is a vector field: $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$, $\xi \in \mathcal{L}$ with $\mathrm{ord}(\xi) = (1, 0)$. It follows from Lemma 5.8, that:

$$\exp(X) \cdot \mathrm{id} = \exp(\log F) \cdot \mathrm{id} = F \cdot \mathrm{id} = f,$$

so f is the time-one map of X.

We now prove the *uniqueness* of X. Let $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$ be *any* vector field such that f (a hyperbolic contraction) is its time-one map. Since f is hyperbolic, by Proposition 5.13 it follows that $\mathrm{ord}(\xi) = (1, 0)$. By Proposition 5.12, the family $\exp(tX)$ defines a $\mathcal{C}^1$-flow of X. By Proposition 5.13, the $\mathcal{C}^1$-flow of X is unique, so it follows that:

$$f = \exp\left(\xi\frac{\mathrm{d}}{\mathrm{d}x}\right) \cdot \mathrm{id} = \xi + \frac{1}{2!}\xi'\xi + \frac{1}{3!}(\xi'\xi)'\xi + \cdots .$$

Using the above expansion, we additionally conclude that $\xi(x) = \lambda x + \text{h.o.t.}$ if and only if $f(x) = e^{\lambda} x + \text{h.o.t.}$ Since f is a hyperbolic contraction, it follows that $\xi(x) = \lambda x + \text{h.o.t.}$ with $\lambda < 0$. By Corollary 5.19, we have:

$$\exp(X) \cdot h = h \circ f = F \cdot h, \ h \in \mathcal{L}.$$

That is, $\exp X = F$. By Lemma 5.15, $X = \log F$ and uniqueness follows. □

Finally, we prove Lemma 5.8 and Lemma 5.15 from Sections 5.1 and 5.2 for the hyperbolic case.

Proof of Lemma 5.8 in the hyperbolic case. By Lemma 5.24, the operator $\log F \in L(\mathcal{L})$ is weakly well-defined. As in the proof of Theorem B above, using Proposition 5.9 and Lemma 5.26, the operator $\log F$ is a vector field. Thus, $\log F = \xi\frac{\mathrm{d}}{\mathrm{d}x}$, $\xi \in \mathcal{L}$, with $\mathrm{ord}(\xi) = (1, 0)$. By Proposition 5.12, $\exp(\log F) \in L(\mathcal{L})$ is weakly well-defined. Having proven that all operators are weakly well-defined, the equality follows by symbolic computation with formal exp–log series. □

Proof of Lemma 5.15 in the case $\mathbf{ord(\xi) = (1, 0)}$. By Proposition 5.12, $\exp(X) \in L(\mathcal{L})$ is a weakly well-defined operator and an isomorphism associated with $f = \exp(X) \cdot \mathrm{id}$. $f \in \mathcal{L}$ is hyperbolic since $\mathrm{ord}(\xi) = (1, 0)$. More precisely, we compute:

$$\begin{aligned} f(x) = \exp(X) \cdot \mathrm{id} = x + \xi + \frac{1}{2!}\xi'\xi + \frac{1}{3!}(\xi'\xi)'\xi + \cdots \\ &= x + (\lambda x + \text{h.o.t.}) + \frac{1}{2!}(\lambda^2 x + \text{h.o.t.}) + \frac{1}{3!}(\lambda^3 x + \text{h.o.t.}) + \cdots \\ &= e^{\lambda} x + \text{h.o.t.} \end{aligned} \tag{5.47}$$

Since $\lambda < 0$, we have $0 < e^{\lambda} < 1$, so f is a hyperbolic contraction. By Lemma 5.24, the operator $\log F$ is a weakly well-defined operator. The equality follows by symbolic computation with formal exp–log series. □

We illustrate the convergence of coefficients (that is, the convergence in the weak topology in $\mathcal{L}$ of respective series) in the hyperbolic case on the simplest hyperbolic elements of $\mathcal{L}$:

Example 5.27.

1. Let $f(x) = \lambda x$, $0 < \lambda < 1$. By Theorem B (hyperbolic case), the embedding vector field for f is given by $X = \xi \frac{\mathrm{d}}{\mathrm{d}x}$, where

$$\xi(x) = \log(F) \cdot \mathrm{id} = (\lambda - 1)x - \frac{1}{2}(\lambda - 1)^2 x + \frac{1}{3}(\lambda - 1)^3 x + \cdots = \log \lambda \cdot x \in \mathcal{L}.$$

2. Let $X = ax\frac{\mathrm{d}}{\mathrm{d}x}$, $a \in \mathbb{R}$. By Proposition 5.12, the field X admits a flow $\{f_t : t \in \mathbb{R}\} \subset \mathcal{L}$:

$$\begin{aligned} f_t(x) &= \exp(tX) \cdot \mathrm{id} = x + t\xi + \frac{t^2}{2!}\xi'\xi + \cdots = x + tax + \frac{t^2 a^2}{2!}x + \frac{t^3 a^3}{3!}x + \cdots \\ &= e^{ta} \cdot x, \quad t \in \mathbb{R}. \end{aligned}$$

The above series converge in $\mathcal{L}$ in the weak topology. Note that they *do not* converge in $\mathcal{L}$ neither in the formal topology nor in the product topology with respect to the discrete topology, since the monomial x appears in every term.

5.5. Theorem B in the strongly hyperbolic case

We first observe that a strongly hyperbolic element of $\mathcal{L}^H$ does not embed in the $\mathcal{C}^1$-flow of a vector field. Indeed, let $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$. If $(1, 0) \preceq \mathrm{ord}\,(\xi)$, it follows from Propositions 5.11 and 5.12 that all the elements of the $\mathcal{C}^1$-flow of X are either parabolic or hyperbolic. If $\mathrm{ord}\,(\xi) \prec (1, 0)$, it follows from Proposition 5.13 that X does not admit any $\mathcal{C}^1$-flow. We have moreover the following *negative version* of Propositions 5.11 and 5.12:

Proposition 5.28. *Let* $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$*,* $\xi \in \mathcal{L}$*, such that* $\mathrm{ord}\,(\xi) \prec (1, 0)$*. Then the exponential operator* $\exp\,(tX)$*,* $t \in \mathbb{R}$*, is not weakly well-defined.*

Proof. Consider the expansion

$$\exp\,(tX) \cdot \mathrm{id} = \mathrm{id} + t\xi + \frac{t^2}{2!}\xi'\xi + \frac{t^3}{3!}\,(\xi'\xi)'\,\xi + \cdots . \tag{5.48}$$

We observe that the orders of the terms in this expansion are unboundedly increasing instead of decreasing. Hence this exponential series does not converge in $\mathcal{L}$ in any of the topologies considered in this work (see Subsection 4.2). □

However, the results of Section 4 lead to the following embedding statement, which is a *weak* version of Theorem B for strongly hyperbolic elements:

Theorem *(Weaker version of Theorem B, the strongly hyperbolic case). Let $f \in \mathcal{L}^H$ (resp. $f \in \mathcal{L}^H_{\mathfrak{D}}$) be strongly hyperbolic. Then f embeds in a flow $(f^t)_{t\in\mathbb{R}}$ of elements of $\mathcal{L}^H$ (resp. $\mathcal{L}^H_{\mathfrak{D}}$).*

Proof. Write $f(x) = \lambda x^\alpha + \text{h.o.t.}$, $\lambda \neq 0$, $\alpha \neq 1$. According to Theorem A (*c*), there exists a change of variables $\varphi \in \mathcal{L}^0$ such that $f_0(x) = \varphi^{-1} \circ f \circ \varphi(x) = x^\alpha$. Obviously, $f_0(x) = x^\alpha$ embeds in the $f_0^t(x) = x^{(\alpha^t)}$, $t \in \mathbb{R}$. Hence, f embeds in the flow:

$$f^t(x) = \left(\varphi \circ f_0^t \circ \varphi^{-1}\right)(x), \ t \in \mathbb{R}. \tag{5.49}$$

The claim in the finitely generated case follows easily. □

We notice here an important difference between the parabolic or hyperbolic case and the strongly hyperbolic case. If $f \in \mathcal{L}^H$ is parabolic or hyperbolic, there exists a well-ordered subset $S \subseteq \mathbb{R}_{>0} \times \mathbb{Z}$ which contains the supports of all the elements of the $\mathcal{C}^1$-flow in which f embeds. It is not the case anymore if f is strongly hyperbolic. Moreover, in this case, the monomials of the $(f^t)_t$, and not only their coefficients, depend on $t \in \mathbb{R}$. The following example illustrates these facts, as well as other specific features of the strongly hyperbolic situation.

Example 5.29 *(A counterexample to the exponential formula for the flow in the strongly hyperbolic case).* Consider the flow $f_0^t(x) = x^{(\alpha^t)}$, $t \in \mathbb{R}$. The strongly hyperbolic element $f_0^1(t) = x^\alpha$ embeds in this flow, and all the elements f_0^t are strongly hyperbolic. Since $\mathcal{S}(f_0^t) = \{(\alpha^t, 0)\}$, these supports are not contained in a common well-ordered subset of $\mathbb{R}_{>0} \times \mathbb{Z}$. Hence the family (f_0^t), $t \in \mathbb{R}$, is not a $\mathcal{C}^1$-*flow* in the sense of Definition 1.2.

Let us now consider

$$\xi(x) := \frac{\mathrm{d}f_0^t(x)}{\mathrm{d}t}\Big|_{t=0} = -\log\alpha \cdot x\boldsymbol{\ell}^{-1} \in \mathcal{L}.$$

It would seem that (f_0^t) is a flow of the vector field $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$. But we have just noticed that (f_0^t) is not a $\mathcal{C}^1$-flow. Moreover, since $\operatorname{ord}(\xi) = (1,-1) \prec (1,0)$, we have seen in Proposition 5.13 that X does not admit any $\mathcal{C}^1$-flow.

It is nevertheless interesting to observe the result of the exponential formula (5.8) for the field $X = \xi\frac{\mathrm{d}}{\mathrm{d}x}$. We obtain:

$$\begin{aligned}\exp\big(-\log\alpha \cdot x\boldsymbol{\ell}^{-1} \cdot \frac{\mathrm{d}}{\mathrm{d}x}\big) \cdot \mathrm{id} = x - \log\alpha \cdot x\boldsymbol{\ell}^{-1} + &\\ + \frac{\log^2\alpha}{2!}(x\boldsymbol{\ell}^{-2} - x\boldsymbol{\ell}^{-1}) + &\\ + \frac{\log^3\alpha}{3!}(-x\boldsymbol{\ell}^{-3}x + 3x\boldsymbol{\ell}^{-2} - x\boldsymbol{\ell}^{-1}) + &\\ + \frac{\log^4\alpha}{4!}(x\boldsymbol{\ell}^{-4} - 6x\boldsymbol{\ell}^{-3} + 7x\boldsymbol{\ell}^{-2} - x\boldsymbol{\ell}^{-1}) + \cdots .&\end{aligned}$$

The order of terms obviously increases, and the above sum *does not converge in $\mathcal{L}$ in any of the mentioned topologies*. Formally, it is not even an element of $\mathcal{L}$. However, if we regroup and sum the terms along the diagonals going from bottom to top, using the convergence of the exponential series, we obtain:

$$\begin{aligned}
&\exp\big(-\log\alpha\cdot x\boldsymbol{\ell}^{-1}\frac{\mathrm{d}}{\mathrm{d}x}\big)\cdot \mathrm{id} = \\
&= x - x\boldsymbol{\ell}^{-1}\cdot(e^{\log\alpha}-1) + x\boldsymbol{\ell}^{-2}\cdot\frac{1}{2!}\cdot(e^{\log\alpha}-1)^2 - x\boldsymbol{\ell}^{-3}\cdot\frac{1}{3!}\cdot(e^{\log\alpha}-1)^3+\cdots = \\
&= x - x\boldsymbol{\ell}^{-1}\cdot(\alpha-1)+\frac{1}{2!}x\boldsymbol{\ell}^{-2}\cdot(\alpha-1)^2-\frac{1}{3!}x\boldsymbol{\ell}^{-3}\cdot(\alpha-1)^3+\cdots = x\cdot e^{(\alpha-1)\log x} \\
&= x^\alpha\in\mathcal{L}.
\end{aligned}$$

Hence in some sense f_0 embeds in the flow of X, but not as it is defined in the present work. We intend to give a precise meaning to the above computations in a subsequent work.

6. Examples

Example 6.1.

$$f(x) = x + x\boldsymbol{\ell} + \text{h.o.t.}$$

By Theorem A, we obtain the formal normal form f_0 and its embedding vector field X_0:

$$\begin{aligned}
&f_0(x) = x + x\boldsymbol{\ell} + bx\boldsymbol{\ell}^3,\\
&\widehat{f_0} = \exp(X_0).\mathrm{id},\quad X_0 = \frac{x}{\boldsymbol{\ell}^{-1}+1/2+(1/2+b)\boldsymbol{\ell}}\frac{\mathrm{d}}{\mathrm{d}x}.
\end{aligned}$$

Here, $b\in\mathbb{R}$ depends on the terms of f up to $x\boldsymbol{\ell}^3$.

In the next example we explain on a very simple example of a Dulac germ why we need a *transfinite sequence* of power-logarithmic changes of variables to derive the finite formal normal form from Theorem A. That is, we illustrate why a standard sequence of changes of variables is not sufficient for elimination.

Example 6.2 *(Dulac germ).* Take $f(x) = x + x^2\boldsymbol{\ell}^{-1} + x^2$. This germ is of *Dulac type* – it has the expansion $f(x) = x + x^2P_1(-\log x)$, where $P_1(x) = x + 1$. By Theorem A, the finite formal normal form of f in $\mathcal{L}$ is:

$$f_0(x) = x + x^2\boldsymbol{\ell}^{-1} + bx^3\boldsymbol{\ell}^{-1},\ b\in\mathbb{R}.$$

Let us illustrate on this example the process used in the proof of Theorem A. We first eliminate the term x^2 from f. Computing the first finitely many (important) terms of

$f \circ \varphi - \varphi \circ f$, for a change of variables $\varphi(x) = x + cx^\beta \boldsymbol{\ell}^\ell$, $(\beta, \ell) \succ (1, 0)$, $c \in \mathbb{R}$, we obtain:

$$f \circ \varphi - \varphi \circ f = c(\beta - 2)x^{\beta+1}\boldsymbol{\ell}^{\ell-1} + c(\ell + 1)x^{\beta+1}\boldsymbol{\ell}^{\ell} - c^2 x^{\beta+1}\boldsymbol{\ell}^{2\ell-1} + c(2 - \beta)(-1)^m x^{\beta+1}\boldsymbol{\ell}^{\ell-m} + \text{h.o.t.}$$

We conclude: by a change of variables $\varphi(x) = x + cx\boldsymbol{\ell}^{-m+1}$, $m \leq 0$, for an appropriate $c \in \mathbb{R}$, we eliminate the term $x^2\boldsymbol{\ell}^{-m}$, but at the same time we generate the *next one*: $x^2\boldsymbol{\ell}^{-m+1}$. Thus we need a transfinite sequence of changes of variables:

$$f(x) = x + x^2\boldsymbol{\ell}^{-1} + x^2 \overset{\varphi_1(x)=x+c_1x\boldsymbol{\ell}}{\longrightarrow} f_1(x) = x + x^2\boldsymbol{\ell}^{-1} + a_1x^2\boldsymbol{\ell} + \text{h.o.t.} \overset{\varphi_2(x)=x+c_2x\boldsymbol{\ell}^2}{\longrightarrow} f_2(x) = x + x^2\boldsymbol{\ell}^{-1} + a_2x^2\boldsymbol{\ell}^2 + \text{h.o.t.} \longrightarrow \cdots.$$

Example 6.3 *(Formal normal forms in $\mathcal{L}$ of formal power series).* Let $f \in \mathbb{R}[[x]]$ be a parabolic formal diffeomorphism,

$$f(x) = x + x^{k+1} + o(x^{k+1}),\ k \in \mathbb{N}.$$

The standard formal normal form in $\mathbb{R}[[x]]$ is equal to:

$$f_s(x) = x + x^{k+1} + bx^{2k+1},\ b \in \mathbb{R}.$$

On the other hand, Theorem A gives a normal form f_0 of f in the wider class $\mathcal{L}^0$ of changes of variables. Note that $\mathbb{R}[[x]] \subset \mathcal{L}$. We prove here that f_0 is equal to

$$f_0(x) = x + x^{k+1}. \tag{6.1}$$

Note that, allowing wider class of logarithmic changes of variables, we remove also the residual term x^{2k+1} from f_s.

Let us now prove (6.1). By Theorem A, we have:

$$f_0(x) = x + x^{k+1} + bx^{2k+1}\boldsymbol{\ell},\ b \in \mathbb{R}.$$

By the algorithm in the proof of Theorem A applied to parabolic power series f, we show that its residual coefficient b is actually equal to 0. Indeed, in order to eliminate all the terms before the residual one, we use in the algorithm non-logarithmic changes of variables $\varphi_{m,0}(x) = x + cx^m$, $c \in \mathbb{R}$, $m \in \{2, 3, \ldots, k\}$. By these changes of variables, no logarithmic terms are generated in f. Therefore, f is transformed into:

$$f(x) = x + x^{k+1} + bx^{2k+1} + dx^{2k+2} + \text{h.o.t.},\ b, d \in \mathbb{R}. \tag{6.2}$$

In the next step, we remove the residual term x^{2k+1} from f. By (3.7), we use a logarithmic change of variables $\varphi_{k+1,-1}(x) = x + cx^{k+1}\boldsymbol{\ell}^{-1}$, $c \in \mathbb{R}$. We compute:

$$\varphi_{k+1}^{-1} \circ f \circ \varphi_{k+1} = f + cx^{2k+1} + r(x). \tag{6.3}$$

Here, $r \in \mathcal{L}$ may contain logarithmic terms, but its leading monomial is of order at least $(2k+2)$ in x. Choosing $c = -b$, the term x^{2k+1} is eliminated from F. Therefore, by (6.2) and (6.3),

$$\varphi_{k+1}^{-1} \circ f \circ \varphi_{k+1} = x + x^{k+1} + (dx^{2k+2} + \text{h.o.t.}) + r(x).$$

The terms after x^{k+1} are of strictly higher order than the residual order $(2k+1, 1)$, so they are eliminated by changes of variables from $\mathcal{L}^0$.

7. Appendix

In the proof of Theorem A in Subsection 4.4, *Part 2*, in order to prove the existence of a formal normalizing change of variables $\varphi \in \mathcal{L}^0$ for $f \in \mathcal{L}^H$ as the composition of a *transfinite sequence* of elementary changes of variables, we index the set of all elementary changes of variables used in the normalization by their orders. We describe here explicitly the set of orders of all elementary changes needed for reduction of f to a formal normal form f_0. In addition, this description allows us to prove easily that, if $f \in \mathcal{L}^H_{\mathfrak{D}}$ (that is, if f is of finite type), then φ is also of finite type.

In Subsection 7.1, we give the main lemma of the appendix. It explains how the support of a transseries behaves under the action of an elementary change of variables. In Subsection 7.2, we use this lemma to control the orders of the normalizing elementary changes of variables. Finally, we discuss the finite type cases in Subsection 7.3.

We analyze only the case when f is *parabolic*. The analysis for other two cases can be done similarly and we omit it.

7.1. The action of an elementary change of variables on the support

Consider $f \in \mathcal{L}^H$ parabolic, $f(x) = x + ax^{\alpha}\boldsymbol{\ell}^p + \text{h.o.t.}$, $(\alpha, p) \succ (1, 0)$. Let $S = \mathcal{S}(f - \text{id})$, $(\alpha, p) = \min(S)$, and $\overline{S} = S \setminus \{(\alpha, p)\}$. Recall that we denote by $\langle A \rangle$ the additive semigroup generated by a subset A of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$. We introduce the set

$$\mathcal{R} = \left\langle \overline{S} - (\alpha, p+1) \right\rangle + \mathbb{N}_* (\alpha - 1, p) + \{1\} \times \mathbb{N}_*,$$

where $\mathbb{N}_*$ means $\mathbb{N} \setminus \{0\}$. It follows from Neumann's Lemma that $\mathcal{R}$ is well-ordered. Moreover, it is easily seen that all elements of $\overline{S} - (\alpha, p+1)$ are bigger than or equal to $(0, 0)$.

Notice that $S \subseteq \mathcal{R}$ (this remark will allow us to initiate a transfinite induction in the next subsection). Indeed, if $(\alpha_1, p_1) \in S$, we write:

$$\begin{aligned}(\alpha_1, p_1) &= (\alpha_1 - \alpha, p_1 - p - 1) + (\alpha, p+1) \\ &= (\alpha_1 - \alpha, p_1 - p - 1) + (\alpha - 1, p) + (1, 1) \in \mathcal{R}.\end{aligned}$$

We now prove the main lemma of the appendix.

Lemma 7.1. *Consider a parabolic series* $f(x) = x + ax^{\alpha}\boldsymbol{\ell}^{p} + a_1 x^{\gamma_1}\boldsymbol{\ell}^{r_1} + \cdots \in \mathcal{L}^H$ *such that all exponents* (γ_i, r_i) *belong to* $\mathcal{R}$*. Let* $\varphi(x) = x + cx^{\beta}\boldsymbol{\ell}^{m}$*,* $(\beta, m) \succ (1, 0)$*, be such that*

$$(\beta, m) = (\gamma_1 - \alpha + 1, r_1 - p) \ \text{or} \ (\gamma_1 - \alpha + 1, r_1 - p - 1). \tag{7.1}$$

Then $\mathcal{S}\left(\varphi^{-1} \circ f \circ \varphi - \mathrm{id} - ax^{\alpha}\boldsymbol{\ell}^{p}\right)$ *is contained in* $\mathcal{R}$*.*

Remark 7.2. The change of variables $\varphi(x) = x + cx^{\beta}\boldsymbol{\ell}^{m}$ with (β, m) as in (7.1) eliminates the term $a_1 x^{\gamma_1}\boldsymbol{\ell}^{r_1}$ from f. The exponent (β, m) is given by the homological equation (4.5), as the proof of Theorem A. Notice that we denote the elementary change of variables here by φ instead of $\varphi_{\beta,m}$ for easier reading of the forthcoming computations.

Proof of Lemma 7.1. We proceed in several steps: we have to control the supports $\mathcal{S}(\varphi)$, $\mathcal{S}\left(\varphi^{-1}\right)$, then $\mathcal{S}\left(\varphi^{-1} \circ f\right)$, and finally of $\mathcal{S}\left(\varphi^{-1} \circ f \circ \varphi\right)$, for f and φ as in Lemma 7.1.

1. **Control of the support $\mathcal{S}(\varphi)$.** Assume $(\beta, m) = (\gamma_1 - \alpha + 1, r_1 - p)$ or $(\gamma_1 - \alpha + 1, r_1 - p - 1)$. We claim that

$$(\beta, m) \in \mathcal{R}_1 := \left\langle \overline{S} - (\alpha, p+1) \right\rangle + \mathbb{N}_0 (\alpha - 1, p) + \{0\} \times \mathbb{N}_0.$$

Notice that $\mathcal{R}_1$ is well-ordered.

We observe that, if $(\beta, m) = (\gamma_1, r_1) - (\alpha, p+1) + (1, 0)$, then, for all $k \in \mathbb{N}_0$, $(\beta, m) + (-1, k) \in \mathcal{R}_1$. In the same way, if $(\beta, m) = (\gamma_1, r_1) - (\alpha, p) + (1, 0)$, then

$$(\beta, m) + (-1, k) = (\gamma_1, r_1) - (\alpha, p+1) + (0, k+1) \in \mathcal{R}_1.$$

In particular, we have that $(\beta, m) \in \mathcal{R}_1$ and $(\beta - 1, m) \in \mathcal{R}_1$.

From now on, we suppose that $(\beta, m) = (\gamma_1, r_1) - (\alpha, p+1) + (1, 0)$.

2. **Control of the support $\mathcal{S}\left(\varphi^{-1}\right)$.** For this purpose, we consider the isomorphism Φ associated to the change of variables φ. It holds that $\Phi.\mathrm{id} = \varphi$. We analyze φ^{-1} using the inverse operator: $\varphi^{-1} = \Phi^{-1}.\mathrm{id}$. Since $\varphi(x) = x + cx^{\beta}\boldsymbol{\ell}^{m}$, we have:

$$\Phi(h)(x) = h\big(\varphi(x)\big) = h\left(x + cx^{\beta}\boldsymbol{\ell}^{m}\right) = h(x) + \sum_{i=1}^{\infty} c_i h^{(i)}(x)\, x^{i\beta}\boldsymbol{\ell}^{im}, \ h \in \mathcal{L}.$$

Hence, $\Phi = \mathrm{Id} + \sum_{i=1}^{\infty} c_i x^{i\beta}\boldsymbol{\ell}^{im}\frac{\mathrm{d}^i}{\mathrm{d}x^i} = \mathrm{Id} + P$. It is easily seen that P is a small operator, so that Φ^{-1} is well-defined by the convergent series $\Phi^{-1} = \sum_{k=0}^{\infty} (-1)^k P^k$ and $\varphi^{-1}(x) = x + \sum_{k=1}^{\infty} (-1)^k P^k(x)$.

We claim that:

$$\mathcal{S}\left(\varphi^{-1}(x) - x\right) \subseteq \mathbb{N}_* (\beta - 1, m) + \{1\} \times \mathbb{N}.$$

Indeed, $P(x) = c_1 x^{\beta}\boldsymbol{\ell}^{m}$, and we can write $(\beta, m) = (\beta - 1, m) + \{1\} \times \mathbb{N}$. Inductively, a consecutive action of P leads to a series in terms:

$$x^{i\beta}\boldsymbol{\ell}^{im}\frac{\mathrm{d}^i}{\mathrm{d}x^i}\left(x^{k(\beta-1)+1}\boldsymbol{\ell}^{km+j}\right),\quad i\geq 0,\ k\geq 1,\ j\in\mathbb{N}.$$

Hence, the elements of their support can be written

$$\begin{aligned}&(i\beta+k(\beta-1)+1-i, im+km+j+s)\\&\quad=((i+k)(\beta-1)+1,(i+k)m+j+s)\\&\quad=(i+k)(\beta-1,m)+(1,j+s)\in\mathbb{N}_*(\beta-1,m)+\{1\}\times\mathbb{N},\end{aligned}$$

which proves the claim. Since we have proven above that $(\beta-1,m)\in\mathcal{R}_1$, it follows that

$$\mathcal{S}(\varphi^{-1}-\mathrm{id})\subseteq\mathcal{R}_1.$$

3. **Control of the support $\mathcal{S}\left(\varphi^{-1}\circ f\right)$.** Let us write $f(x)=x+\varepsilon(x)$, $\mathrm{ord}(\varepsilon)\succ(1,0)$, and $\varphi^{-1}(x)=x+\sum b_{\mu s}x^{\mu}\boldsymbol{\ell}^s$. Then

$$\begin{aligned}\varphi^{-1}(f(x))&=\varphi^{-1}(x+\varepsilon(x))=\varphi^{-1}(x)+\sum_{k=1}^{\infty}\frac{1}{k!}\left(\varphi^{-1}\right)^{(k)}(x)\,\varepsilon^k\\&=\varphi^{-1}(x)+\varepsilon(x)+\left(\sum b_{\mu s}x^{\mu}\boldsymbol{\ell}^s\right)'\varepsilon(x)+\sum_{k=2}^{\infty}\frac{1}{k!}\left(\sum b_{\mu s}x^{\mu}\boldsymbol{\ell}^s\right)^{(k)}\varepsilon(x)^k.\end{aligned}$$

We already know that $\mathcal{S}(\varepsilon)\subseteq\mathcal{R}$. Let us study the exponents of the series $(x^{\mu}\boldsymbol{\ell}^s)^{(k)}\,\varepsilon(x)^k$. They are of the form:

$$\begin{aligned}&(\mu-k,s+\nu)+\left(\left(\gamma_{i_1},r_{i_1}\right)+\cdots+\left(\gamma_{i_k},r_{i_k}\right)\right)\\&\quad=(\mu,s)+\left(\gamma_{i_1}-1,r_{i_1}\right)+\cdots+\left(\gamma_{i_k}-1,r_{i_k}\right)+(0,\nu)\\&\quad=(\mu,s)+\left(\gamma_{i_1}-\alpha,r_{i_1}-p-1\right)+\cdots+\left(\gamma_{i_k}-\alpha,r_{i_k}-p-1\right)\\&\qquad+k(\alpha-1,p)+(0,\nu+k),\end{aligned}$$

where $\nu\in\mathbb{N}$, $k\geq 1$ and $\left(\gamma_{i_1},r_{i_1}\right),\ldots,\left(\gamma_{i_k},r_{i_k}\right)\in\mathcal{S}(\varepsilon)\subseteq\mathcal{R}$. Each pair $(\gamma_{i_j}-\alpha, r_{i_j}-p-1)$ belongs to $\mathcal{R}_1$. Recall from the previous step that $\mathcal{S}(\varphi^{-1}-\mathrm{id})\subseteq\mathcal{R}_1$. Hence, these exponents can be written as

$$(\bar{\mu},\bar{s})+k(\alpha-1,p)+(1,\nu+k),$$

where $(\bar{\mu},\bar{s})\in\mathcal{R}_1$. So they belong to $\mathcal{R}$. Hence, we can write $\varphi^{-1}(f(x))=x+g(x)$, with $\mathcal{S}(g)\subseteq\mathcal{R}$.

4. **Control of the support $\mathcal{S}\left(\varphi^{-1}\circ f\circ\varphi\right)$.** We have

$$\varphi^{-1}(f(\varphi(x)))=x+g(\varphi(x))=x+g\left(x+cx^{\beta}\boldsymbol{\ell}^m\right),$$

where $\mathcal{S}(g) \subseteq \mathcal{R}$ from the previous step. The elements of the support of $\varphi^{-1} \circ f \circ \varphi$ can be written

$$\begin{aligned}(\tau, l) &= \big(\mu + k(\beta+1), s + km + j\big) \\ &= (\mu, s) + k\,(\beta - 1, m) + (0, j)\,, \quad (\mu, s) \in \mathcal{S}(g) \subseteq \mathcal{R},\ k \geq 0,\ j \in \mathbb{N}.\end{aligned}$$

Since we have shown above that $(\beta - 1, m) \in \mathcal{R}_1$, so is $(\tau, l) \in \mathcal{R}$. □

7.2. The control of the orders of the normalizing elementary changes of variables

Let $(\varphi_{\beta,m})$ be a transfinite sequence of elementary changes of variables used to normalize f. Let $(\psi_{\beta,m})$ denote their partial compositions and φ the limit of these partial compositions (hence φ is the change of variables which normalizes f). We prove first that the supports of $(f_{\beta,m} - \mathrm{id})$, $f_{\beta,m} := \psi_{\beta,m}^{-1} \circ f \circ \psi_{\beta,m}$, are all contained in the set $\mathcal{R}$ of the previous subsection. This is based on a straightforward transfinite induction:

i) We have already noticed that $\mathcal{S}\,(f - \mathrm{id})$ is contained in $\mathcal{R}$.
ii) The non-limit case follows directly by Lemma 7.1: if the support of f is contained in $\mathcal{R}$, so is the support of $\varphi_{\beta,m}^{-1} \circ f \circ \varphi_{\beta,m}$.
iii) The limit case comes from the obvious classical fact in Hahn fields: *Consider a transfinite sequence $(g_\mu)_{\mu<\theta}$ of elements of a Hahn field, which admits a limit g and whose supports are contained in a common well-ordered set W. Then $\mathcal{S}\,(g) \subseteq W$.*

By Lemma 7.1 and Remark 7.2, the supports of all elementary changes, $\mathcal{S}(\varphi_{\beta,m} - \mathrm{id})$, are contained in $\mathcal{R}_1$. By an easy computation in the non-limit case and using the classical result mentioned under (iii) above in the limit case, we conclude that $\mathcal{S}(\psi_{\beta,m} - \mathrm{id})$, $\mathcal{S}(\varphi - \mathrm{id}) \subseteq \mathcal{R}_1$.

7.3. Finite type cases

The goal of this subsection is to show that if a parabolic series f is in addition of finite type, then so is the normalizing change of variables φ built in the proof of Theorem A in Section 4.4. We keep all the notations as above.

Assume that f is of finite type. We prove that in this case the sets $\mathcal{R}$ and $\mathcal{R}_1$, as well as the support of the composition $\mathcal{S}(\varphi)$, are of finite type. These claims follow from the following easy result:

Lemma 7.3. *Consider a subset A of finite type of $\mathbb{R}_{>0} \times \mathbb{Z}$. Let $(\alpha, p) \in \mathbb{R}_{>0} \times \mathbb{Z}$ and $\overline{A} := \{(\beta, m) \in A\colon\ (\beta, m) \succeq (\alpha, p)\}$. Then the set $\overline{A} - (\alpha, p)$ is also of finite type.*

Proof. Suppose A is contained in the sub-semigroup G of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$ generated by the elements $(\alpha_1, p_1)\,, \dots, (\alpha_k, p_k)$ of $\mathbb{R}_{>0} \times \mathbb{Z}$. For each $i = 1, \dots, k$, denote by N_i the smallest positive integer such that $N_i(\alpha_i, p_i) \succeq (\alpha, p)$. Consider an element $(\gamma, r) \in A \subseteq G$, $(\gamma, r) \succeq (\alpha, p)$. Then $(\gamma, r) = \sum_{i=1}^{k} n_i(\alpha_i, p_i)$, $n_i \in \mathbb{N}_0$. There exist only *finitely many* elements $(\gamma, r) \in A$ such that the respective n_i-s satisfy $n_i < N_i$, for all $i = 1, \ldots, k$. On the other hand, if one of the n_i's, say n_1, is greater than or equal to N_1, we write

$$(\gamma, r) - (\alpha, p) = (n_1 - N_1)(\alpha_1, p_1) + n_2(\alpha_2, p_2) + \cdots \\ + n_k(\alpha_k, p_k) + N_1(\alpha_1, p_1) - (\alpha, p).$$

This shows that $\overline{A} - (\alpha, p)$ is contained in the sub-semigroup of $\mathbb{R}_{\geq 0} \times \mathbb{Z}$ finitely generated by the elements $(\alpha_1, p_1), \ldots, (\alpha_k, p_k)$, $N_i(\alpha_i, p_i) - (\alpha, p)$, $i = 1, \ldots, k$ (note that $N_i(\alpha_i, p_i) - (\alpha, p) \succeq (0, 0)$), and the elements $(\gamma, r) - (\alpha, p)$, for the *finitely many* $(\gamma, r) \succeq (\alpha, p)$ for which the respective n_i-s satisfy $n_i < N_i$, for all $i = 1, \ldots, k$. □

If f is of the finite type, we apply Lemma 7.3 to the set $\overline{S} - (\alpha, p + 1)$ from the previous section. It follows that the sets $\mathcal{R}$ and $\mathcal{R}_1$ are also of the finite type. Since, by Sections 7.1 and 7.2, we have that $\mathcal{S}(\varphi - \mathrm{id}) \in \mathcal{R}_1$ and $\mathcal{S}(f_0 - \mathrm{id}) \in \mathcal{R}$, we deduce that φ and f_0 are of the finite type.

References

[1] M. Abate, Discrete holomorphic local dynamical systems, in: Holomorphic Dynamical Systems, in: Lecture Notes in Math., vol. 1998, Springer, Berlin, 2010, pp. 1–55.
[2] V.I. Arnold, Geometrical Methods in the Theory of Ordinary Differential Equations, 2nd edition, Grundlehren Math. Wiss., vol. 250, Springer-Verlag, New York, 1988.
[3] L. van den Dries, A. Macintyre, D. Marker, Logarithmic–exponential series, in: Proceedings of the International Conference 'Analyse & Logique', Mons, 1997, Ann. Pure Appl. Logic 111 (1–2) (2001) 61–113.
[4] H. Dulac, Sur les cycles limites, Bull. Soc. Math. France 51 (1923) 45–188.
[5] J. Écalle, The acceleration operators and their applications to differential equations, quasianalytic functions, and the constructive proof of Dulac's conjecture, in: Proceedings of the International Congress of Mathematicians, vols. I, II, Kyoto, 1990, Math. Soc. Japan, Tokyo, 1991, pp. 1249–1258.
[6] J. Écalle, Introduction aux fonctions analysables et preuve constructive de la conjecture de Dulac, Actual. Math. (Current Mathematical Topics), Hermann, Paris, 1992.
[7] Y. Ilyashenko, Finiteness theorems for limit cycles, Russian Math. Surveys 45 (2) (1990) 143–200; Transl. Amer. Math. Soc. 94 (1991).
[8] Y. Ilyashenko, S. Yakovenko, Lectures on Analytic Differential Equations, Grad. Stud. Math., vol. 86, American Mathematical Society, Providence, RI, 2008, xiv+625 pp.
[9] F. Loray, Pseudo-Groupe d'une Singularité de Feuilletage Holomorphe en Dimension Deux, Prépublication IRMAR, ccsd-00016434, 2005.
[10] P. Mardešić, M. Resman, V. Županović, Multiplicity of fixed points and ε-neighborhoods of orbits, J. Differential Equations 253 (2012) 2493–2514.
[11] J. Milnor, Dynamics in One Complex Variable. Introductory Lectures, 2nd edition, Friedr. Vieweg. & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiesbaden, 2000.
[12] B.H. Neumann, On ordered division rings, Trans. Amer. Math. Soc. 66 (1949) 202–252.
[13] M. Resman, ε-neighborhoods of orbits and formal classification of parabolic diffeomorphisms, Discrete Contin. Dyn. Syst. 33 (8) (2013) 3767–3790.
[14] M. Resman, ε-neighborhoods of orbits of parabolic diffeomorphisms and cohomological equations, Nonlinearity 27 (2014) 3005–3029.
[15] R. Roussarie, Bifurcations of Planar Vector Fields and Hilbert's Sixteenth Problem, Birkhäuser Verlag, Basel, 1998.
[16] W. Rudin, Principles of Mathematical Analysis, 3rd edition, McGraw–Hill, Inc., USA, 1976.

[17] F. Takens, Normal forms for certain singularities of vector fields, Ann. Inst. Fourier 23 (2) (1973) 163–195.
[18] F. Takens, Singularities of vector fields, Publ. Math. Inst. Hautes Études Sci. 43 (1974) 47–100.
[19] D. Žubrinić, V. Županović, Poincaré map in fractal analysis of spiral trajectories of planar vector fields, Bull. Belg. Math. Soc. Simon Stevin 15 (2008) 947–960.