

WEIGHTED LINEARIZATION OF VECTOR FIELDS VIA A FORMAL MOSER TRICK

ARTHUR LEI QIU

ABSTRACT. Many well-known theorems establish sufficient criteria for linearizability of a vector field in terms of the eigenvalues of its linear approximation. By attaching weights to coordinates so that some directions are considered “linear”, others “quadratic”, and so on, one can define the notion of a *weighted linear approximation*. It is thus natural to ask when a vector field is “weighted-linearizable”. In this paper, we formulate a weighted version of the non-resonance condition appearing in the Poincaré and Sternberg linearization theorems and show that it implies weighted linearizability.

Our approach first addresses weighted linearization on the level of formal power series. In doing so, we develop a general framework to make sense of a power series version of *Moser’s trick*, a technique used to prove various normal form results in geometry. This formal Moser trick works over any field of characteristic zero and may be of independent interest.

CONTENTS

1. Introduction	1
2. Formal geometry	5
3. Weightings of $\mathbb{K}^n$	20
4. Weighted linearization	24
References	30

1. INTRODUCTION

1.1. Background. Linearization of vector fields is a fundamental technique with applications throughout mathematics, science, and engineering. It features prominently in the study of local normal forms in differential geometry and dynamical systems, where a natural question arises: when does the first-order approximation furnished by linearization accurately reflect a vector field’s geometry and dynamics near a fixed point? This question is classical, dating back to Poincaré, and by now there is a wealth of literature establishing sufficient conditions for linearizability with respect to various notions of “accuracy”.

In the topological category, the Hartman–Grobman theorem [2, §13] states that the flow of a C^1 vector field on $\mathbb{R}^n$ can be conjugated via a homeomorphism to the flow of its linear approximation near a *hyperbolic* fixed point (one where the vector field’s Jacobian matrix has no purely imaginary eigenvalues). On the level of formal power series, Poincaré’s thesis [19, Part II] investigates when a system of n ordinary differential equations

$$\dot{x} = Ax + O(|x|^2), \quad A \in \text{Mat}_n(\mathbb{C})$$

can be brought to the form $\dot{y} = Ay$ by a formal change of coordinates $y = x + O(|x|^2)$. He formulates the following sufficient condition: call the eigenvalues $\lambda_1, \dots, \lambda_n$ of the matrix A *resonant* if there exists $j \in \{1, \dots, n\}$ and non-negative integers $\alpha_1, \dots, \alpha_n$ such that

$$\langle \lambda, \alpha \rangle = \lambda_j \quad \text{and} \quad \sum_{i=1}^n \alpha_i > 1, \quad \text{where} \quad \langle \lambda, \alpha \rangle := \sum_{i=1}^n \lambda_i \alpha_i.$$

2020 *Mathematics Subject Classification.* 37G05, 34C20 (Primary); 70K45, 13J05, 53D99 (Secondary).

Key words and phrases. Weighted resonance, quasi-homogeneity, normal forms, formal geometry, Moser’s trick.

Poincaré’s linearization theorem states that, if the eigenvalues are *not* resonant, then such a formal change of coordinates exists. Should the formal power series $Ax + O(|x|^2)$ define a holomorphic vector field on $\mathbb{C}^n$, results of Poincaré and Siegel [2, §24] describe when a linearizing formal coordinate transformation converges to a biholomorphism, again in terms of eigenvalues of A .

In the smooth (by which we always mean C^∞) category, Sternberg’s linearization theorem [20, 21] states that a smooth vector field on $\mathbb{R}^n$ can be linearized near a fixed point p by a diffeomorphism if the (possibly complex) eigenvalues of its Jacobian matrix at p satisfy the same non-resonance condition as in Poincaré’s linearization theorem. For example, fix an integer $k > 1$ and consider the vector field on $\mathbb{R}^2$ given by

$$X = (x + y^k) \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}.$$

This vector field is *Euler-like*, meaning its linear approximation at $p = 0$ is the *Euler vector field*

$$\mathcal{E} = \tilde{x} \frac{\partial}{\partial \tilde{x}} + \tilde{y} \frac{\partial}{\partial \tilde{y}}.$$

Since $\mathcal{E}$ has non-resonant eigenvalues $(\lambda_1, \lambda_2) = (1, 1)$, Sternberg’s theorem ensures that there exists a smooth local change of coordinates $(x, y) \mapsto (\tilde{x}, \tilde{y})$ transforming X into $\mathcal{E}$ near 0. In fact, one can explicitly check that the (global) change of coordinates $(\tilde{x}, \tilde{y}) = (x - \frac{1}{k-1}y^k, y)$ works. More generally, a vector field on a smooth manifold is called Euler-like for a submanifold P if it is tangent to P and its linear approximation along P equals the Euler vector field of the normal bundle of P ; Bursztyn–Lima–Meinrenken have proved that every Euler-like vector field is smoothly linearizable along the entire submanifold [6, Lemma 2.4].

In contrast, neither Sternberg’s nor Poincaré’s result apply to the *weighted Euler-like* vector field

$$Y = (x + y^k) \frac{\partial}{\partial x} + 2y \frac{\partial}{\partial y}, \tag{1.1}$$

as the eigenvalues $(\lambda_1, \lambda_2) = (1, 2)$ of its linear part satisfy the resonance relation $2\lambda_1 = \lambda_2$. Nevertheless, Y is linearizable: the change of coordinates $(x', y') = (x - \frac{1}{2k-1}y^k, y)$ takes Y to

$$\mathcal{E}_w = x' \frac{\partial}{\partial x'} + 2y' \frac{\partial}{\partial y'}, \tag{1.2}$$

a so-called *weighted Euler vector field*. Actually, this coordinate transformation takes Y to $\mathcal{E}_w$ even when $k = 1$, in which case the eigenvalues remain resonant but the linear term $y \frac{\partial}{\partial x}$ disappears upon changing coordinates. More generally, Meinrenken has proved that every weighted Euler-like vector field on $\mathbb{R}^n$ can be smoothly conjugated to a corresponding weighted Euler vector field [14, Lemma 2.5], hence is smoothly linearizable, even though many such vector fields are resonant.

Thus, while the non-resonance condition is sufficient for formal and smooth linearizability, it is far from necessary. For resonant vector fields, Poincaré’s theorem is extended by the Poincaré–Dulac theorem [2, §23], which produces a formal change of coordinates that might remove some but not all non-linear terms in the vector field. The particular non-linearities that can appear upon changing coordinates depend on certain non-canonical choices made during the iterative construction of the coordinate transformation, and these in turn depend on the particular resonance relations occurring in the vector field, so there is some ambiguity in what the resulting normal form looks like.

This paper offers one systematic way to gain more control: by accepting a generalized notion of linearity based on assigning positive integer weights to the coordinates of $\mathbb{R}^n$ or $\mathbb{C}^n$, so that coordinates of weight 1 are considered “linear”, coordinates of weight 2 are considered “quadratic”, and so on. The weighted Euler vector field (1.2) corresponds to assigning weights $w = (1, 2)$ to the first and second coordinates of $\mathbb{R}^2$, respectively, which explains why the coordinate transformation $(x, y) \mapsto (x', y')$ removes the term $y^k \frac{\partial}{\partial x}$ from the weighted Euler-like vector field (1.1) even when $k = 1$: it is considered “non-linear” with respect to these weights.

Such *weightings* appear throughout mathematics. For example, weighted projective spaces are objects of study in algebraic geometry, and Loizides–Meinrenken [13] have recently developed a differential-geometric theory of weightings in detail. Weighted linear approximations of vector fields have been previously studied by Algaba–García–Reyes [1] under the name of “like-linearizations”. Weighted Euler-like vector fields play an important role in their work, which provides necessary and sufficient conditions for a vector field to be equivalent to its weighted linear approximation in two different senses: the vector fields themselves being conjugate, and their flows being orbitally equivalent (i.e., the same up to time reparametrization).

1.2. Main results. Algaba–García–Reyes’ characterization of weighted linearizability differs from the other linearization theorems discussed in Section 1.1 in that it is not expressed in terms of eigenvalues, which are straightforward to compute in principle. Our goal is to give an eigenvalue-based sufficient condition akin to the non-resonance condition of Poincaré’s and Sternberg’s linearization theorems. This *weighted non-resonance* condition is more flexible than the Poincaré–Sternberg condition and implies weighted linearizability; the compromise is that weighted linear approximations need not be linear in the usual sense. Nevertheless, one has more control over the non-linear terms that can appear compared to the Poincaré–Dulac theorem, as any remaining non-linearities must be “weighted-linear”. Moreover, one can choose the weighting to their advantage, and sometimes by doing so one obtains weighted linear approximations that *are* linear in the usual sense, as is the case for the weighted Euler-like vector field (1.1) with the weighting $w = (1, 2)$.

By combining the main results of this paper, we offer the following weighted generalization of the Poincaré and Sternberg linearization theorems.

THEOREM 1.1. *Let X be either a formal vector field on $\mathbb{K}^n$ where $\mathbb{K}$ is any field of characteristic zero, or a smooth vector field on an open neighbourhood of $0 \in \mathbb{R}^n$. For any weighting w of $\mathbb{K}^n$ (resp. $\mathbb{R}^n$) with respect to which X is “admissible” and “non-resonant” (Definitions 4.1 and 4.7), there exists a formal (resp. germ of a smooth) diffeomorphism that conjugates X with its weighted linear approximation with respect to w .*

Taking the *trivial weighting* $w = (1, \dots, 1)$ recovers Poincaré’s and Sternberg’s theorems, for in this case admissibility simply means that 0 is a fixed point, non-resonance coincides with the usual notion, and weighted linear approximation is just linear approximation. A special case of this theorem (Proposition 4.4) easily detects that the $y^k \frac{\partial}{\partial x}$ term in the vector field (1.1) can be removed when Poincaré’s and Sternberg’s theorems cannot (one can also see this from [1, Theorem 1.2] or [14, Lemma 2.5]).

To explain our approach to proving Theorem 1.1, it is worth stepping back from weighted linearizability to consider an abstracted problem. A fundamental question in differential geometry is: given tensor fields (or differential forms, or multivector fields) A_0 and A_1 on a smooth manifold M , when does there exist a diffeomorphism $\varphi: M \rightarrow M$ such that the pullback φ^*A_1 is equal to A_0 ? If not globally, then one can ask: supposing $x \in M$ is such that $A_0(x) = A_1(x)$, when does there exist a locally defined diffeomorphism φ such that $\varphi(x) = x$ and $\varphi^*A_1 = A_0$ near x ?

Moser’s trick is a powerful technique for answering this question. In the broadest sense, it consists of searching for a one-parameter family of tensor fields (or differential forms, or multivector fields) $\{A_t\}_{t \in [0,1]}$ joining A_0 and A_1 , and a one-parameter family of diffeomorphisms $\{\varphi_t\}_{t \in [0,1]}$ such that φ_0 is the identity map and $\frac{d}{dt}(\varphi_t^*A_t) = 0$. If such families can be found, then $\varphi_t^*A_t = \varphi_0^*A_0 = A_0$ for all t , so $\varphi := \varphi_1$ solves the problem. Typically, instead of directly looking for the family of diffeomorphisms φ_t , one seeks out the time-dependent vector field X_t that generates φ_t as its flow. The rate of change of $\varphi_t^*A_t$ is related to the Lie derivative $\mathcal{L}_{X_t}$ via the identity

$$\frac{d}{dt}(\varphi_t^*A_t) = \varphi_t^* \left(\frac{dA_t}{dt} + \mathcal{L}_{X_t}A_t \right), \quad (1.3)$$

so it suffices to find A_t and X_t solving the differential equation

$$\frac{dA_t}{dt} + \mathcal{L}_{X_t} A_t = 0. \quad (1.4)$$

By judiciously choosing the interpolating family A_t based on the particular structure of a given problem, Equation (1.4) can often be easily solved for X_t using pointwise (multi-)linear algebra. Integrating X_t to its flow then produces the desired φ_t . Originally developed by Moser in the context of volume-preserving geometry [16], Moser’s trick can be used to give very short proofs of the Darboux theorem [22] and Morse lemma [18]. It also facilitates Bursztyn–Lima–Meinrenken’s proof that Euler-like vector fields are linearizable along submanifolds, and Meinrenken’s proof that weighted Euler-like vector fields on $\mathbb{R}^n$ are linearizable. The Moser-type idea of realizing a diffeomorphism intertwining geometric objects as the flow of a time-dependent vector field finds application, for example, in computational anatomy and medical imaging under the name of *large deformation diffeomorphic metric mapping* [4].

Techniques developed by Sternberg were later modified by Chen [7] to prove a “theorem of equivalence”: two smooth vector fields, both with a hyperbolic fixed point at $0 \in \mathbb{R}^n$, are smoothly conjugate near 0 if and only if they are formally conjugate. Since complex roots of polynomials with real coefficients come in conjugate pairs, non-resonance implies hyperbolicity for smooth vector fields. (If a fixed point is *not* hyperbolic, then either 0 is an eigenvalue of the Jacobian matrix, in which case $0 = 2 \cdot 0$ is a resonance, or there is a pair of non-zero purely imaginary eigenvalues $\pm i\theta$, in which case $i\theta = 2(i\theta) + 1(-i\theta)$ is a resonance.) Thus, one can linearize a non-resonant smooth vector field by first Taylor expanding its component functions; Poincaré’s theorem then produces a formal diffeomorphism that kills off any higher finite-order terms; Chen’s theorem upgrades this to an actual diffeomorphism that kills off any remaining infinite-order terms.

In view of Chen’s theorem of equivalence, we sought to develop a Moser-type argument for Theorem 1.1 that passes through the world of formal power series. Working with functions and vector fields on $\mathbb{R}^n$ or $\mathbb{C}^n$ only insofar as their formal Taylor series mattered led us to naïvely wonder if the time parameter t in Moser’s trick could be treated as a formal variable as well. One obvious problem arises: namely, how to “evaluate” an arbitrary power series at $t = 1$ (or any $t \neq 0$ for that matter). Rather than giving up and reverting to treating t as an actual parameter, we obstinately proceeded to develop criteria describing when it *does* make sense to “evaluate at non-zero t ” in a way that does not use any properties about $\mathbb{R}$ or $\mathbb{C}$ besides being fields of characteristic zero. We thus obtain the following “Moser trick for formal power series”:

FORMAL MOSER TRICK. *Let $\mathbb{K}$ be any field of characteristic zero, and let A_0 and A_1 be formal tensor fields on $\mathbb{K}^n$. Suppose there exists an “evaluative” (Definition 2.12) time-dependent formal tensor field A_t interpolating between A_0 and A_1 , and there exists a time-dependent formal vector field X_t satisfying Equation (1.4) in a certain formal sense and whose flow is “evaluative”. Then there exists a formal diffeomorphism φ of $\mathbb{K}^n$ such that $\varphi^* A_1 = A_0$.*

1.3. Paper overview. The rest of this paper is structured as follows. In Sections 2.1 and 2.2, we define algebraic analogues of standard differential-geometric concepts: vector/tensor fields, diffeomorphisms, related constructions, and one-parameter families thereof. In Section 2.3, we state and prove a version of the existence and uniqueness theorem for formal solutions to ordinary differential equations (Theorem 2.6), allowing us to import Equations (1.3) and (1.4) into this setting. In Section 2.4, we define “evaluative” formal objects, state a mild generalization of the formal Moser trick (Corollary 2.17), and give a sufficient condition for flows of formal vector fields to be evaluative.

Section 3.1 focuses on weightings and the associated notions of quasi-homogeneity of functions and vector fields in an algebraic setting; we also note comparisons to quasi-homogeneity in the differentiable world. Section 3.2 generalizes the sufficient condition for “evaluability” of formal flows from Section 2.4 to accommodate such weightings.

In Section 4.1, we define the class of vector fields for which it is sensible to pose the question of weighted linearizability; sketch an informal Moser-type argument yielding a plausibly sufficient condition in the differentiable world; then reverse the argument to prove that this condition indeed suffices for weighted linearizability on the level of formal power series (Theorem 4.5). We reformulate this condition in terms of eigenvalues by introducing the notion of weighted (non-)resonance in Section 4.2, obtaining a corresponding generalization of Poincaré’s theorem (Theorem 4.8). The paper concludes in Section 4.3 by explaining how Chen’s theorem of equivalence can be invoked in the weighted setting to upgrade from formal to smooth weighted linearizability (Corollary 4.11), thus also generalizing Sternberg’s theorem to the weighted setting.

Throughout this paper, $\mathbb{K}$ denotes a fixed choice of any field of characteristic zero. All algebras and algebra homomorphisms are required to be unital; all derivations are required to be $\mathbb{K}$ -linear. We include 0 in the set of natural numbers $\mathbb{N} := \{0, 1, 2, \dots\}$ to avoid overusing sub-/superscripts. A multi-index of length k is a k -tuple $\alpha = (\alpha_1, \dots, \alpha_k) \in \mathbb{N}^k$; its degree is denoted by $|\alpha| := \sum_i \alpha_i$.

1.4. Acknowledgements. I am greatly indebted to Eckhard Meinrenken for his diligent mentorship while supervising my Master’s project, of which the present work is an expansion. I am also deeply grateful to Yael Karshon for insightful comments and patient support throughout the preparation of this paper. Feedback from Nic Fellini and Charlie Wu has been instrumental in the process of writing and editing the manuscript. Posthumous thanks are owed to Will J. Merry for the profound impact of his teachings on the way I think and write about mathematics.

This work was supported by a University of Toronto Excellence Award, an Ontario Graduate Scholarship, and a doctoral NSERC Canada Graduate Research Scholarship.

2. FORMAL GEOMETRY

We begin by laying out the framework in which we shall obtain a formal version of Moser’s trick. Much of the material in this section may be well-known in the language of ∞ -jets, but we include it for the sake of making this paper self-contained and setting the notation used throughout. We have made an effort to highlight connections to various related concepts in geometry and algebra; in particular, to poly-flow vector fields and locally nilpotent derivations (Remarks 2.22 and 2.23).

2.1. Formal functions, diffeomorphisms, and tensor fields. We write $\mathbb{K}[[x]] := \mathbb{K}[[x^1, \dots, x^n]]$ for the $\mathbb{K}$ -algebra of formal power series in n commuting variables $x^1, \dots, x^n$ with coefficients in $\mathbb{K}$. Elements of $\mathbb{K}[[x]]$ are formal infinite sums

$$\sum_{\alpha \in \mathbb{N}^n} c_\alpha x^\alpha, \quad \text{where } c_\alpha \in \mathbb{K} \text{ and } x^\alpha := (x^1)^{\alpha_1} \dots (x^n)^{\alpha_n}.$$

Such expressions are intended to be evocative of the Taylor series of (germs of) smooth functions,¹ so we will often refer to them as **formal functions** on $\mathbb{K}^n$ and denote them by f or $f(x)$ interchangeably. The “constant term” $c_{(0, \dots, 0)}$ is then denoted by $f(0)$; the map $f \mapsto f(0)$ defines a surjective $\mathbb{K}$ -algebra homomorphism $\mathbb{K}[[x]] \rightarrow \mathbb{K}$. We extend this notation to n -tuples of formal power series $\varphi(x) = (\varphi^1(x), \dots, \varphi^n(x))$ component-wise by setting $\varphi(0) := (\varphi^1(0), \dots, \varphi^n(0))$, and similarly for matrices with entries in $\mathbb{K}[[x]]$.

Since $\mathbb{K}$ is a local integral domain, so too is $\mathbb{K}[[x]]$. Its unique maximal ideal $\mathfrak{m}$ consists of those power series f such that $f(0) = 0$; equivalently, $\mathfrak{m}$ is generated by the variables $x^1, \dots, x^n$. We equip $\mathbb{K}[[x]]$ with the **$\mathfrak{m}$ -adic topology**, which is the translation-invariant topology such that the decreasing filtration by ideals

$$\mathbb{K}[[x]] = \mathfrak{m}^0 \supset \mathfrak{m}^1 \supset \mathfrak{m}^2 \supset \dots$$

¹This is true on the nose for $\mathbb{K} = \mathbb{R}$, as the $\mathbb{R}$ -algebra homomorphism $C^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}[[x]]$ sending a smooth function to its formal Taylor series centred at the origin is surjective by Borel’s lemma [17, §1.5].

forms a basis of open neighbourhoods of the zero power series. We adopt the Big Theta notation

$$\Theta(f) := \max\{k \mid f \in \mathfrak{m}^k\} = \min\{|\alpha| \mid c_\alpha \neq 0\}$$

to denote the *order* of a power series $f(x) = \sum_\alpha c_\alpha x^\alpha$, with the convention that the zero power series has order ∞ . We extend this notation to n -tuples of power series $\varphi(x) = (\varphi^1(x), \dots, \varphi^n(x))$ by defining $\Theta(\varphi) := \min_i \Theta(\varphi^i)$.

The $\mathfrak{m}$ -adic topology on $\mathbb{K}[[x]]$ is induced by a complete metric; for example, the one given by $(f, g) \mapsto 2^{-\Theta(f-g)}$. A sequence of power series $f_k(x) = \sum_\alpha c_{k,\alpha} x^\alpha$ is Cauchy with respect to some (equivalently, every) compatible complete metric if and only if, for each α , the sequence of coefficients $c_{k,\alpha}$ is eventually constant; the coefficient of x^α in the limiting power series is then given by $\lim_{k \rightarrow \infty} c_{k,\alpha}$. An infinite sum $\sum_k^\infty f_k$ converges in the $\mathfrak{m}$ -adic topology if and only if $f_k \rightarrow 0$, or equivalently, $\Theta(f_k) \rightarrow \infty$.

The $\mathbb{K}$ -algebra $\mathbb{K}[[x]]$ together with its $\mathfrak{m}$ -adic topology satisfies a certain universal property, of which we state a special case [5, IV§4.3]:

There is a bijective correspondence

$$\left\{ \begin{array}{l} \text{continuous } \mathbb{K}\text{-algebra endomorphisms} \\ \varphi: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]] \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} n\text{-tuples of elements of } \mathfrak{m} \\ (\varphi^1(x), \dots, \varphi^n(x)) \end{array} \right\}, \quad (2.1)$$

one determining the other by defining $\varphi(x^i) = \varphi^i(x)$ for each i .

The power series $\varphi^i(x)$ is called the i^{th} *component* of φ (or the corresponding n -tuple). Continuity with respect to the $\mathfrak{m}$ -adic topology ensures that the image of any power series under φ may be computed “term-by-term”, even with infinitely many non-zero terms. For example, one always has

$$\sum_{\alpha \in \mathbb{N}^n} c_\alpha x^\alpha = \sum_{k=0}^{\infty} \left(\sum_{|\alpha|=k} c_\alpha x^\alpha \right), \quad (2.2)$$

where the left-hand side is an element of $\mathbb{K}[[x]]$ and the right-hand side is a convergent limit of partial sums of a sequence in $\mathbb{K}[[x]]$; continuity allows us to compute

$$\varphi\left(\sum_\alpha c_\alpha x^\alpha\right) = \sum_{k=0}^{\infty} \varphi\left(\sum_{|\alpha|=k} c_\alpha x^\alpha\right) = \sum_\alpha c_\alpha \varphi(x^1)^{\alpha_1} \cdots \varphi(x^n)^{\alpha_n}.$$

In this sense, φ can be thought of as “pullback of formal functions” (but see Remark 2.2).

Remark 2.1. Actually, for any power series algebras $\mathbb{K}[[x]] = \mathbb{K}[[x^1, \dots, x^n]]$ and $\mathbb{K}[[y]] = \mathbb{K}[[y^1, \dots, y^m]]$ over $\mathbb{K}$ (possibly in different sets of variables), every $\mathbb{K}$ -algebra homomorphism $\varphi: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[y]]$ is *automatically* continuous with respect to the $\mathfrak{m}$ -adic topologies, where $\mathfrak{m}$ refers to the maximal ideals in both $\mathbb{K}[[x]]$ and $\mathbb{K}[[y]]$. To see this, recall that a power series with coefficients in any commutative ring R is invertible if and only if its constant term is invertible in R . For each $i \in \{1, \dots, n\}$, let $c^i \in \mathbb{K}$ be the constant term of $\varphi(x^i)$. Since φ is $\mathbb{K}$ -linear and unital, we have

$$\varphi(x^i - c^i) = \varphi(x^i) - c^i \in \mathfrak{m}_y.$$

Thus $\varphi(x^i - c^i)$ is not invertible in $\mathbb{K}[[y]]$; as φ is a unital ring homomorphism, this implies that $x^i - c^i$ is not invertible in $\mathbb{K}[[x]]$. Since $\mathbb{K}$ is a field, we obtain $c^i = 0$. Thus $\varphi(x^i) \in \mathfrak{m}_y$ for each i ; as the x^i ’s generate $\mathfrak{m}_x$, we conclude that $\varphi(\mathfrak{m}_x) \subseteq \mathfrak{m}_y$, so φ is $\mathfrak{m}$ -adically continuous.

This argument relies on $\mathbb{K}$ being a field. Over more general commutative rings, $\mathfrak{m}$ -adically discontinuous homomorphisms of power series algebras can exist and are more complicated to describe; see [8] and the references within. We explicitly mention continuity in the bijective correspondence (2.1) only to emphasize its utility in future computations. $\blacklozenge$

A diffeomorphism from a smooth manifold to itself defines an automorphism of its algebra of smooth functions via pullback. In this spirit, we define a **formal diffeomorphism**² of $\mathbb{K}^n$ to be a $\mathbb{K}$ -algebra automorphism of $\mathbb{K}[[x]]$. Thus, a formal diffeomorphism determines a unique n -tuple of elements of $\mathfrak{m}$ via the bijective correspondence (2.1). Conversely, given any n -tuple $(\varphi^1(x), \dots, \varphi^n(x))$ of elements of $\mathfrak{m}$, there is a simple criterion to determine when the corresponding endomorphism $\varphi: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ is an automorphism. It mimics the inverse function theorem: φ is invertible if and only if the Jacobian matrix of formal partial derivatives

$$D\varphi(x) := \begin{pmatrix} \frac{\partial \varphi^1}{\partial x^1}(x) & \dots & \frac{\partial \varphi^1}{\partial x^n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi^n}{\partial x^1}(x) & \dots & \frac{\partial \varphi^n}{\partial x^n}(x) \end{pmatrix}$$

is invertible in $\text{Mat}_n(\mathbb{K}[[x]])$. As the constant term of $\det D\varphi(x)$ is $\det D\varphi(0)$, this is also equivalent to $D\varphi(0)$ being invertible in $\text{Mat}_n(\mathbb{K})$ [5, IV§4.7].

Next, we develop the formal versions of smooth tensor fields. The goal is to mimic smooth expressions at every step; as such, the reader would not be led too far astray by importing the smooth version of each definition and replacing the algebra of smooth functions with the algebra of formal functions everywhere. Nevertheless, we shall describe the main players to reduce the guesswork needed and mention some key facts.

Smooth vector fields act as derivations of a manifold's algebra of smooth functions. Analogously, we define a **formal vector field** on $\mathbb{K}^n$ to be a derivation of the algebra of formal functions, i.e., a $\mathbb{K}$ -linear map $X: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ satisfying the Leibniz rule:

$$X(fg) = X(f)g + fX(g), \quad \forall f, g \in \mathbb{K}[[x]].$$

We denote the set of formal vector fields by $\mathfrak{X}$; it is both a module over $\mathbb{K}[[x]]$ and a Lie algebra over $\mathbb{K}$ via the commutator Lie bracket $[X, Y] := X \circ Y - Y \circ X$. Every formal vector field is continuous with respect to the $\mathfrak{m}$ -adic topology; a corollary of this fact is that $\mathfrak{X}$ is free of rank n over $\mathbb{K}[[x]]$, with a basis given by the set of formal partial derivatives $\left\{ \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \right\}$ [5, IV§4.6].

We define a **formal covector field** on $\mathbb{K}^n$ to be an element of the dual $\mathbb{K}[[x]]$ -module

$$\Omega^1 := \text{Hom}_{\mathbb{K}[[x]]}(\mathfrak{X}, \mathbb{K}[[x]]).$$

Every formal function $f \in \mathbb{K}[[x]]$ has a **formal differential** $df \in \Omega^1$, defined by $df(X) = X(f)$ for $X \in \mathfrak{X}$. The set $\{dx^1, \dots, dx^n\}$ is the dual basis to $\left\{ \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \right\}$. Thus, every $X \in \mathfrak{X}$ can be uniquely written in the form $X = \sum_i X^i \frac{\partial}{\partial x^i}$ where $X^i = X(x^i) = dx^i(X)$, and every $\omega \in \Omega^1$ can be uniquely written in the form $\omega = \sum_i \omega_i dx^i$ where $\omega_i = \omega\left(\frac{\partial}{\partial x^i}\right)$. In particular, $df = \sum_i \frac{\partial f}{\partial x^i} dx^i$. (See Remark 2.3 for a comparison of formal covector fields and Kähler differentials.)

Having defined formal vector and covector fields, we define a **formal tensor field of type** (r, s) on $\mathbb{K}^n$ to be an element of the tensor product of $\mathbb{K}[[x]]$ -modules

$$\mathcal{T}^{(r,s)} := \underbrace{\mathfrak{X} \otimes \dots \otimes \mathfrak{X}}_{r \text{ times}} \otimes \underbrace{\Omega^1 \otimes \dots \otimes \Omega^1}_{s \text{ times}},$$

with the convention that $\mathcal{T}^{(0,0)} := \mathbb{K}[[x]]$. Since $\mathfrak{X}$ and Ω^1 are free of finite rank, formal tensor fields $A \in \mathcal{T}^{(r,s)}$ are canonically identifiable with $\mathbb{K}[[x]]$ -multilinear maps

$$A: \underbrace{\Omega^1 \times \dots \times \Omega^1}_{r \text{ times}} \times \underbrace{\mathfrak{X} \times \dots \times \mathfrak{X}}_{s \text{ times}} \rightarrow \mathbb{K}[[x]].$$

²Strictly speaking, what we call a “formal diffeomorphism” more precisely reflects the Taylor series of a smooth diffeomorphism at a *fixed point*.

The tensor product operation $\otimes$ endows

$$\mathcal{T} := \bigoplus_{r,s \geq 0} \mathcal{T}^{(r,s)}$$

with the structure of a graded $\mathbb{K}[[x]]$ -algebra. A **formal tensor derivation** is a $\mathbb{K}$ -linear map $\mathcal{D}: \mathcal{T} \rightarrow \mathcal{T}$ that preserves type, commutes with all contraction homomorphisms $\mathcal{T}^{(r,s)} \rightarrow \mathcal{T}^{(r-1,s-1)}$ (defined by feeding one of the $\mathfrak{X}$ factors to one of the Ω^1 factors), and satisfies

$$\mathcal{D}(A \otimes B) = \mathcal{D}(A) \otimes B + A \otimes \mathcal{D}(B), \quad \forall A, B \in \mathcal{T}.$$

Two formal tensor derivations are identical if they agree on formal functions and formal vector fields. Conversely, every pair of $\mathbb{K}$ -linear maps $\mathcal{D}^{(0,0)}: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ and $\mathcal{D}^{(1,0)}: \mathfrak{X} \rightarrow \mathfrak{X}$ satisfying

$$\mathcal{D}^{(0,0)}(fg) = \mathcal{D}^{(0,0)}(f)g + f\mathcal{D}^{(0,0)}(g), \quad \mathcal{D}^{(1,0)}(fX) = \mathcal{D}^{(0,0)}(f)X + f\mathcal{D}^{(1,0)}(X)$$

for all $f, g \in \mathbb{K}[[x]]$ and $X \in \mathfrak{X}$ extends uniquely to a formal tensor derivation. In particular, given $X \in \mathfrak{X}$, the **formal Lie derivative** $\mathcal{L}_X: \mathcal{T} \rightarrow \mathcal{T}$ is defined to be the unique formal tensor derivation extending $\mathcal{L}_X^{(0,0)}f := X(f)$ and $\mathcal{L}_X^{(1,0)}Y := [X, Y]$. The map $X \mapsto \mathcal{L}_X$ is $\mathbb{K}$ -linear.

Just as diffeomorphisms pull back smooth tensor fields, a formal diffeomorphism $\varphi: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ extends to a graded $\mathbb{K}$ -algebra automorphism $\varphi^*: \mathcal{T} \rightarrow \mathcal{T}$, called the **pullback**, by requiring φ^* to distribute across $\otimes$ and commute with the formal differential $d: \mathbb{K}[[x]] \rightarrow \Omega^1$ and all contractions. That is, for formal functions $f \in \mathcal{T}^{(0,0)}$, define $\varphi^*f = \varphi(f)$. For formal vector fields $X \in \mathcal{T}^{(1,0)}$, define $\varphi^*X \in \mathcal{T}^{(1,0)}$ to be the composition

$$\mathbb{K}[[x]] \xrightarrow{\varphi^{-1}} \mathbb{K}[[x]] \xrightarrow{X} \mathbb{K}[[x]] \xrightarrow{\varphi} \mathbb{K}[[x]],$$

so that $(\varphi^*X)(\varphi^*f) = \varphi^*(X(f))$ for all $f \in \mathbb{K}[[x]]$. For formal covector fields $\omega \in \mathcal{T}^{(0,1)}$, define $\varphi^*\omega \in \mathcal{T}^{(0,1)}$ to be the composition

$$\mathfrak{X} \xrightarrow{(\varphi^*)^{-1}} \mathfrak{X} \xrightarrow{\omega} \mathbb{K}[[x]] \xrightarrow{\varphi} \mathbb{K}[[x]],$$

so that $(\varphi^*\omega)(\varphi^*X) = \varphi^*(\omega(X))$ for all $X \in \mathfrak{X}$. For formal tensor fields of arbitrary type $A \in \mathcal{T}^{(r,s)}$, define $\varphi^*A \in \mathcal{T}^{(r,s)}$ as a $\mathbb{K}[[x]]$ -multilinear map by requiring that

$$(\varphi^*A)(\varphi^*\omega_1, \dots, \varphi^*\omega_r, \varphi^*X_1, \dots, \varphi^*X_s) = \varphi^*(A(\omega_1, \dots, \omega_r, X_1, \dots, X_s))$$

for all $\omega_1, \dots, \omega_r \in \Omega^1$ and $X_1, \dots, X_s \in \mathfrak{X}$.

A few comparisons to the smooth setting assure us that these formal analogues are sensibly defined. Firstly, φ^* restricts to a Lie algebra automorphism $\mathfrak{X} \rightarrow \mathfrak{X}$. Secondly, pullbacks are compatible with Lie derivatives: for any $X \in \mathfrak{X}$, the composition $\varphi^* \circ \mathcal{L}_X \circ (\varphi^*)^{-1}$ is a formal tensor derivation that agrees with $\mathcal{L}_{\varphi^*X}$ on formal functions and formal vector fields, so they must agree on all formal tensor fields; that is,

$$\varphi^*(\mathcal{L}_X A) = \mathcal{L}_{\varphi^*X}(\varphi^*A), \quad \forall A \in \mathcal{T}. \quad (2.3)$$

Lastly, pullbacks look as expected “in coordinates”, which we demonstrate with formal vector fields $X = \sum_i X^i \frac{\partial}{\partial x^i}$ and formal covector fields $\omega = \sum_i \omega_i dx^i$. If we denote by $(\varphi^*X)(x) \in \mathbb{K}[[x]]^n$ the n -tuple of components of φ^*X with respect to the basis $\{\frac{\partial}{\partial x^i}\}$ and we denote by $X(\varphi(x))$ the n -tuple $(\varphi(X^1), \dots, \varphi(X^n)) \in \mathbb{K}[[x]]^n$, then since $\varphi(X^i) = (\varphi^*X)(\varphi^*x^i) = \sum_j (\varphi^*X)(x^j) \frac{\partial \varphi^i}{\partial x^j}$, one finds that

$$X(\varphi(x)) = D\varphi(x)[(\varphi^*X)(x)].$$

As $D\varphi(x)$ is invertible in $\text{Mat}_n(\mathbb{K}[[x]])$, this can be rewritten as

$$(\varphi^*X)(x) = D\varphi(x)^{-1}[X(\varphi(x))],$$

mimicking the usual pointwise formula for pullbacks of smooth vector fields along diffeomorphisms. Similarly, if we denote by $(\varphi^*\omega)(x) \in \mathbb{K}[[x]]^n$ the n -tuple of components of $\varphi^*\omega$ with respect to the

basis $\{dx^i\}$ and we denote by $\omega(\varphi(x))$ the n -tuple $(\varphi(\omega_1), \dots, \varphi(\omega_n)) \in \mathbb{K}[[x]]^n$, then by computing $(\varphi^*\omega)\left(\frac{\partial}{\partial x^i}\right)$, one finds that $\varphi^*\omega$ is given by transposition:

$$(\varphi^*\omega)(x) = D\varphi(x)^t[\omega(\varphi(x))].$$

Remark 2.2 (terminology and notation). The “coordinate” calculations above justify why the map on formal tensor fields $\varphi^*: \mathcal{T} \rightarrow \mathcal{T}$ is called “pullback” and denoted with an upper star. But if ψ is another formal diffeomorphism, then so is $\psi \circ \varphi$, and the pullback maps satisfy $(\psi \circ \varphi)^* = \psi^* \circ \varphi^*$ —in apparent conflict with the usual convention that upper stars denote contravariant actions. We will have little need to consider compositions of formal diffeomorphisms, so this should not pose a serious issue. Nevertheless, this can be reconciled with the usual convention by recognizing that formal diffeomorphisms are defined by how they act on formal functions, not on “formal points”. That is, our notation for pullback is contravariant *with respect to maps of the spectrum of $\mathbb{K}[[x]]$* . ♦

Remark 2.3 (formal covector fields versus Kähler differentials). The definition of Ω^1 and the formal differential $d: \mathbb{K}[[x]] \rightarrow \Omega^1$ differs from the $\mathbb{K}[[x]]$ -module of Kähler differentials $\Omega_{\mathbb{K}[[x]]/\mathbb{K}}$ and the universal derivation $\delta: \mathbb{K}[[x]] \rightarrow \Omega_{\mathbb{K}[[x]]/\mathbb{K}}$ in a few ways. Firstly, $\Omega_{\mathbb{K}[[x]]/\mathbb{K}}$ is neither projective nor finitely generated as a $\mathbb{K}[[x]]$ -module [12, Example 5.5a] whereas Ω^1 is free of finite rank. Secondly, d satisfies $df = \sum_i \frac{\partial f}{\partial x^i} dx^i$ for all $f \in \mathbb{K}[[x]]$ whereas—intuitively speaking— δ is only finitely additive, so one should not expect that $\delta f = \sum_i \frac{\partial f}{\partial x^i} \delta x^i$ when f has infinitely many non-zero coefficients. Thirdly, the pairs (Ω^1, d) and $(\Omega_{\mathbb{K}[[x]]/\mathbb{K}}, \delta)$ satisfy different universal properties [12, Example 12.7].

To explain the relationship, recall the universal property of $(\Omega_{\mathbb{K}[[x]]/\mathbb{K}}, \delta)$: for every $\mathbb{K}[[x]]$ -module M , pre-composition with δ defines an isomorphism of $\mathbb{K}[[x]]$ -modules

$$\delta^\vee: \text{Hom}_{\mathbb{K}[[x]]}(\Omega_{\mathbb{K}[[x]]/\mathbb{K}}, M) \rightarrow \text{Der}_{\mathbb{K}}(\mathbb{K}[[x]], M), \quad T \mapsto T \circ \delta.$$

Taking $M = \mathbb{K}[[x]]$, we see that $\delta^\vee$ identifies the dual $\mathbb{K}[[x]]$ -module $(\Omega_{\mathbb{K}[[x]]/\mathbb{K}})^\vee$ with $\mathfrak{X}$; dualizing again, we see that $\delta^{\vee\vee}$ identifies Ω^1 with the double dual $(\Omega_{\mathbb{K}[[x]]/\mathbb{K}})^{\vee\vee}$. As $d \in \text{Der}_{\mathbb{K}}(\mathbb{K}[[x]], \Omega^1)$, there is a unique $T \in \text{Hom}_{\mathbb{K}[[x]]}(\Omega_{\mathbb{K}[[x]]/\mathbb{K}}, \Omega^1)$ such that $d = T \circ \delta$. If $\varepsilon: \Omega_{\mathbb{K}[[x]]/\mathbb{K}} \rightarrow (\Omega_{\mathbb{K}[[x]]/\mathbb{K}})^{\vee\vee}$ denotes the canonical homomorphism given by $\varepsilon(\omega)(\chi) = \chi(\omega)$ for $\omega \in \Omega_{\mathbb{K}[[x]]/\mathbb{K}}$ and $\chi \in (\Omega_{\mathbb{K}[[x]]/\mathbb{K}})^\vee$, then by unravelling definitions one obtains the following commutative diagram:

$$\begin{array}{ccc} \mathbb{K}[[x]] & \xrightarrow{d} & \Omega^1 \\ \delta \downarrow & \nearrow T & \downarrow \delta^{\vee\vee} \\ \Omega_{\mathbb{K}[[x]]/\mathbb{K}} & \xrightarrow{\varepsilon} & (\Omega_{\mathbb{K}[[x]]/\mathbb{K}})^{\vee\vee} \end{array}$$

Thus, T is just the canonical homomorphism ε up to the identification provided by $\delta^{\vee\vee}$. ♦

2.2. Formal time dependence. We now generalize the formal objects defined so far by allowing them to depend on another formal variable t in addition to the x^i ’s; the development will closely mirror Section 2.1. To begin, we introduce the notation $\mathbb{K}[[t, x]] := \mathbb{K}[[t, x^1, \dots, x^n]]$ for the $\mathbb{K}$ -algebra of “time-dependent” formal functions, from which all other “time-dependent” formal objects shall be built. Analogously to $\mathbb{K}[[x]]$, it is equipped with the $\mathfrak{m}_{t,x}$ -adic topology where $\mathfrak{m}_{t,x}$ is the (unique maximal) ideal generated by t and the x^i ’s. Thus, any time-dependent formal function

$$f(t, x) = \sum_{k, \alpha} c_{k, \alpha} t^k x^\alpha$$

is in particular equal to the convergent limit of partial sums $\sum_k t^k f_k(x)$ where

$$f_k(x) = \sum_{\alpha} c_{k, \alpha} x^\alpha \in \mathbb{K}[[x]] \text{ for each } k \in \mathbb{N},$$

and we can interchange such infinite sums with continuous operations on $\mathbb{K}[[t, x]]$.

We distinguish t from the x^i 's notationally and psychologically because we will eventually try to “evaluate time-dependent formal objects at a fixed time” by replacing t with an element of $\mathbb{K}$ and applying the field operations of $\mathbb{K}$ to make sense of the resulting expressions. The simplest example of this is the (continuous by Remark 2.1) $\mathbb{K}[[x]]$ -algebra homomorphism

$$\text{eval}_0: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[x]], \quad f(t, x) \mapsto f(0, x) = f_0(x).$$

Of course, substituting non-zero values for t is rarely well-defined, even if $\mathbb{K}$ is topologized to help make sense of the resulting infinite sums—for example, one cannot hope to define an “evaluation at time $t = 1$ ” map on all of $\mathbb{K}[[t, x]]$, as demonstrated by $\sum_k t^k$. However, we must initially allow arbitrary power series in t to obtain an existence and uniqueness theorem for flows of formal vector fields (Theorem 2.6); thus we develop the following story with all concepts based on $\mathbb{K}[[t, x]]$.

We define a **formal isotopy** of $\mathbb{K}^n$ to be a $\mathbb{K}$ -algebra automorphism $\varphi: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$ such that $\varphi(t) = t$. To see that this is a good formal analogue of a one-parameter family of diffeomorphisms, let us explicitly characterize formal isotopies similarly to how we characterized formal diffeomorphisms. The universal property of $\mathbb{K}[[t, x]]$ yields a bijective correspondence

$$\left\{ \begin{array}{l} \text{(automatically continuous) } \mathbb{K}\text{-algebra} \\ \text{endomorphisms } \varphi: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]] \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} (n+1)\text{-tuples of elements of } \mathfrak{m}_{t,x} \\ (\varphi^0(t, x), \varphi^1(t, x), \dots, \varphi^n(t, x)) \end{array} \right\},$$

one determining the other via $\varphi(t) = \varphi^0$ and $\varphi(x^i) = \varphi^i$ for $i \in \{1, \dots, n\}$. This restricts to a bijective correspondence between $\mathbb{K}$ -algebra automorphisms of $\mathbb{K}[[t, x]]$ and such $(n+1)$ -tuples satisfying the additional condition that the $(n+1) \times (n+1)$ Jacobian matrix of partial derivatives of $(\varphi^0, \varphi^1, \dots, \varphi^n)$ with respect to t and the x^i 's is invertible in $\text{Mat}_{n+1}(\mathbb{K}[[t, x]])$. If $\varphi^0(t, x) = t$ (i.e., the corresponding endomorphism fixes t), then the top row of this matrix is $(1 \ 0 \ \dots \ 0)$, so this matrix is invertible if and only if the bottom-right $n \times n$ submatrix

$$D_x \varphi(t, x) := \begin{pmatrix} \frac{\partial \varphi^1}{\partial x^1}(t, x) & \dots & \frac{\partial \varphi^1}{\partial x^n}(t, x) \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi^n}{\partial x^1}(t, x) & \dots & \frac{\partial \varphi^n}{\partial x^n}(t, x) \end{pmatrix}$$

is invertible in $\text{Mat}_n(\mathbb{K}[[t, x]])$, or equivalently, $D_x \varphi(0, 0)$ is invertible in $\text{Mat}_n(\mathbb{K})$. Knowing that $\varphi^0(t, x) = t$, we might as well omit it from such $(n+1)$ -tuples; this yields a bijective correspondence

$$\left\{ \begin{array}{l} \text{formal isotopies} \\ \varphi: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]] \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} n\text{-tuples } \varphi(t, x) \in \mathbb{K}[[t, x]]^n \text{ such that} \\ \varphi(0, 0) = 0 \text{ and } D_x \varphi(0, 0) \text{ is invertible} \end{array} \right\}. \quad (2.4)$$

Compare with how formal diffeomorphisms correspond to certain n -tuples in $\mathbb{K}[[x]]^n$: here, the n component power series are allowed to depend on t but the relevant Jacobian matrix still involves partial derivatives only with respect to the x^i 's. We will typically decorate $\mathbb{K}$ -algebra endomorphisms of $\mathbb{K}[[t, x]]$ with a subscript t (φ_t, ψ_t , etc.) to emphasize the allowed dependence on t .

Remark 2.4. Any n -tuple $\varphi(t, x) \in \mathbb{K}[[t, x]]^n$ can be written in the form $\varphi(t, x) = \sum_{k=0}^{\infty} t^k \varphi_k(x)$ where $\varphi_k(x) \in \mathbb{K}[[x]]^n$; written this way, $\varphi(t, x)$ corresponds to a formal isotopy if and only if $\Theta(\varphi_0) \geq 1$ and $D\varphi_0(0)$ is invertible in $\text{Mat}_n(\mathbb{K})$. $\blacklozenge$

Every $\mathbb{K}$ -algebra endomorphism of $\mathbb{K}[[x]]$ (resp. formal diffeomorphism) extends to a “time-independent” $\mathbb{K}$ -algebra endomorphism of $\mathbb{K}[[t, x]]$ (resp. formal isotopy) by declaring it to fix t . On the level of n -tuples, this is simply the inclusion $\mathbb{K}[[x]]^n \subset \mathbb{K}[[t, x]]^n$. Conversely, we define the $t = 0$ *time slice* of any $\mathbb{K}$ -algebra endomorphism φ_t of $\mathbb{K}[[t, x]]$ that fixes t by the composition

$$\varphi_0: \mathbb{K}[[x]] \hookrightarrow \mathbb{K}[[t, x]] \xrightarrow{\varphi_t} \mathbb{K}[[t, x]] \xrightarrow{\text{eval}_0} \mathbb{K}[[x]].$$

The subscript 0 notation is consistent with identifying φ_t with an n -tuple written as in Remark 2.4, as φ_0 corresponds to the n -tuple $\varphi_0(x) \in \mathbb{K}[[x]]^n$. Time zero slices are compatible with composition:

LEMMA 2.5. For all $\mathbb{K}$ -algebra endomorphisms φ_t, ψ_t of $\mathbb{K}[[t, x]]$ fixing t , we have $(\psi_t \circ \varphi_t)_0 = \psi_0 \circ \varphi_0$. In particular, if φ_t is a formal isotopy, then φ_0 is a formal diffeomorphism and $(\varphi_0)^{-1} = (\varphi_t^{-1})_0$.

Proof. Denote the inclusion by $\iota: \mathbb{K}[[x]] \hookrightarrow \mathbb{K}[[t, x]]$. As $\text{eval}_0 \circ \iota$ is the identity map on $\mathbb{K}[[x]]$, we have

$$\text{eval}_0 \circ \varphi_t \circ \iota = \varphi_0 = \text{eval}_0 \circ \iota \circ \varphi_0.$$

Thus, the image of $\varphi_t \circ \iota - \iota \circ \varphi_0$ is contained in the kernel of eval_0 . Since ψ_t is continuous and fixes t , it maps $\ker(\text{eval}_0)$ to itself, so for all $f \in \mathbb{K}[[x]]$,

$$(\psi_t \circ \varphi_t)_0(f) - (\psi_0 \circ \varphi_0)(f) = \text{eval}_0 \left[\psi_t \left((\varphi_t \circ \iota)(f) - (\iota \circ \varphi_0)(f) \right) \right] = 0.$$

Thus $(\psi_t \circ \varphi_t)_0 = \psi_0 \circ \varphi_0$. ■

Before defining time-dependent formal tensor fields on $\mathbb{K}^n$, we note that by treating t as “just another formal variable” no different from the x^i ’s, one obtains formal tensor fields on $\mathbb{K}^{n+1}$: the $\mathbb{K}[[t, x]]$ -module of derivations of $\mathbb{K}[[t, x]]$, its dual, tensor products of these, and the associated definitions of Lie derivatives and pullbacks (along formal diffeomorphisms of $\mathbb{K}^{n+1}$). Morally speaking, time-dependent formal tensor fields on $\mathbb{K}^n$ are simply formal tensor fields on $\mathbb{K}^{n+1}$, except that when written in terms of tensor products of $\frac{\partial}{\partial x^i}$ ’s and dx^i ’s, their component power series are also allowed to depend on t . Equivalently, they are formal tensor fields on $\mathbb{K}^{n+1}$ that do not involve $\frac{\partial}{\partial t}$ or dt .

More precisely, we define a **time-dependent formal tensor field of type (r, s)** on $\mathbb{K}^n$ to be an element of the $\mathbb{K}[[t, x]]$ -module obtained from the $\mathbb{K}[[x]]$ -module $\mathcal{T}^{(r, s)}$ via extension of scalars along the inclusion $\mathbb{K}[[x]] \hookrightarrow \mathbb{K}[[t, x]]$:

$$\mathcal{T}_t^{(r, s)} := \mathbb{K}[[t, x]] \otimes_{\mathbb{K}[[x]]} \mathcal{T}^{(r, s)}.$$

We denote by $\mathfrak{X}_t := \mathcal{T}_t^{(1, 0)}$ and $\Omega_t^1 := \mathcal{T}_t^{(0, 1)}$ the time-dependent formal vector and covector fields, respectively, which are free $\mathbb{K}[[t, x]]$ -modules with bases $\left\{ \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \right\}$ and $\{dx^1, \dots, dx^n\}$ by construction. In particular, $\mathfrak{X}_t$ is identifiable with the set of derivations of $\mathbb{K}[[t, x]]$ that vanish on t , which is a $\mathbb{K}[[t, x]]$ -submodule and Lie subalgebra of the set of all derivations of $\mathbb{K}[[t, x]]$. Since $\mathcal{T}^{(r, s)}$ is free of finite rank over $\mathbb{K}[[x]]$, the definition of $\mathcal{T}_t^{(r, s)}$ is compatible with thinking of it as $(\mathfrak{X}_t)^{\otimes r} \otimes (\Omega_t^1)^{\otimes s}$ (the tensor products being taken over $\mathbb{K}[[t, x]]$), or as $\mathbb{K}[[t, x]]$ -multilinear maps

$$\underbrace{\Omega_t^1 \times \dots \times \Omega_t^1}_{r \text{ times}} \times \underbrace{\mathfrak{X}_t \times \dots \times \mathfrak{X}_t}_{s \text{ times}} \rightarrow \mathbb{K}[[t, x]].$$

As with formal isotopies, time-dependent formal tensor fields may be decorated with a subscript t (X_t, A_t , etc.) to serve as a reminder that their components are power series in both t and the x^i ’s.

Tensoring the inclusion $\iota: \mathbb{K}[[x]] \hookrightarrow \mathbb{K}[[t, x]]$ with the identity maps $\text{id}: \mathcal{T}^{(r, s)} \rightarrow \mathcal{T}^{(r, s)}$ yields inclusions $\iota \otimes \text{id}: \mathcal{T}^{(r, s)} \hookrightarrow \mathcal{T}_t^{(r, s)}$ of “time-independent” formal tensor fields into time-dependent formal tensor fields. For $(r, s) = (1, 0)$, this simply means that every derivation $\mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ can be extended to a derivation $\mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$ by declaring it to vanish on t . For $(r, s) = (0, 1)$, this means that every $\mathbb{K}[[x]]$ -linear map $\mathfrak{X} \rightarrow \mathbb{K}[[x]]$ can be post-composed with ι to obtain an element of $\text{Hom}_{\mathbb{K}[[x]]}(\mathfrak{X}, \mathbb{K}[[t, x]]) \cong \text{Hom}_{\mathbb{K}[[t, x]]}(\mathfrak{X}_t, \mathbb{K}[[t, x]]) \cong \Omega_t^1$ via the universal property of extension of scalars. In a basis, this inclusion means: write out any $A \in \mathcal{T}^{(r, s)}$ in terms of tensor products of $\frac{\partial}{\partial x^i}$ ’s and dx^i ’s, and simply regard each component as an element of $\mathbb{K}[[t, x]]$ instead of $\mathbb{K}[[x]]$.

Conversely, the image of a time-dependent formal tensor field $A_t \in \mathcal{T}_t^{(r, s)}$ under the projection $\text{eval}_0 \otimes \text{id}: \mathcal{T}_t^{(r, s)} \rightarrow \mathcal{T}^{(r, s)}$ is called its $t = 0$ *time slice* and denoted by A_0 . Regarding A_t as a $\mathbb{K}[[t, x]]$ -multilinear map

$$A_t: \underbrace{\Omega_t^1 \times \dots \times \Omega_t^1}_{r \text{ times}} \times \underbrace{\mathfrak{X}_t \times \dots \times \mathfrak{X}_t}_{s \text{ times}} \rightarrow \mathbb{K}[[t, x]],$$

this corresponds to pre-composing each factor in the domain with one of the inclusions $\mathfrak{X} \hookrightarrow \mathfrak{X}_t$ or $\Omega^1 \hookrightarrow \Omega_t^1$ just described, then post-composing with eval_0 to obtain the $\mathbb{K}[[x]]$ -multilinear map

$$A_0: \underbrace{\Omega^1 \times \cdots \times \Omega^1}_{r \text{ times}} \times \underbrace{\mathfrak{X} \times \cdots \times \mathfrak{X}}_{s \text{ times}} \rightarrow \mathbb{K}[[x]].$$

Equivalently, write out A_t in the $\mathbb{K}[[t, x]]$ -basis of tensor products of $\frac{\partial}{\partial x^i}$'s and dx^i 's, then evaluate each of its component power series at $t = 0$; this is $A_0 \in \mathcal{T}^{(r,s)}$.

The identification of $\mathfrak{X}_t$ with a Lie subalgebra of the set of all derivations of $\mathbb{K}[[t, x]]$ implies that, for all $X_t \in \mathfrak{X}_t$, the formal Lie derivative $\mathcal{L}_{X_t}$ (regarded as a formal tensor derivation in one dimension higher) preserves the set of time-dependent formal tensor fields on $\mathbb{K}^n$. In fact, more is true: the Lie bracket $\frac{dX_t}{dt} := [\frac{\partial}{\partial t}, X_t]$ is also in $\mathfrak{X}_t$ (even though $\frac{\partial}{\partial t} \notin \mathfrak{X}_t$), so the *time derivative* operator $\frac{d}{dt} := \mathcal{L}_{\frac{\partial}{\partial t}}$ also preserves $\mathcal{T}_t^{(r,s)}$. We obtain a split short exact sequence of $\mathbb{K}[[x]]$ -modules

$$0 \rightarrow \mathcal{T}^{(r,s)} \xrightarrow{\iota \otimes \text{id}} \mathcal{T}_t^{(r,s)} \xrightarrow{\frac{d}{dt}} \mathcal{T}_t^{(r,s)} \rightarrow 0,$$

which is a fancy way of saying that $A_t \in \mathcal{T}_t^{(r,s)}$ is time-independent if and only if $\frac{dA_t}{dt} = 0$, if and only if the component power series of A_t are all time-independent.

Regarding a formal isotopy $\varphi_t: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$ of $\mathbb{K}^n$ as a formal diffeomorphism of $\mathbb{K}^{n+1}$, the pullback φ_t^* of formal tensor fields on $\mathbb{K}^{n+1}$ preserves the set of time-dependent formal tensor fields on $\mathbb{K}^n$. (For example, regarding time-dependent formal vector fields $X_t \in \mathfrak{X}_t$ as derivations of $\mathbb{K}[[t, x]]$ that vanish on t , we have $(\varphi_t^* X_t)(t) = \varphi_t(X_t(t)) = 0$, so $\varphi_t^* X_t \in \mathfrak{X}_t$.) Given $A_t \in \mathcal{T}_t^{(r,s)}$, let us temporarily denote the $t = 0$ time slice of the pullback $\varphi_t^* A_t$ by $(\varphi_t^* A_t)|_{t=0}$. As one might expect, this coincides with the pullback of the time-independent formal tensor field A_0 along the formal diffeomorphism $\varphi_0: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ of $\mathbb{K}^n$ (Lemma 2.5):

$$(\varphi_t^* A_t)|_{t=0} = \varphi_0^* A_0.$$

Section 2.4 will be dedicated to formulating definitions to make sense of this equality for non-zero values of t , provided that φ_t and A_t are “evaluative” (Proposition 2.15).

2.3. Formal flows. Just as smooth vector fields generate flows, so too do their formal versions. Of course, existence and uniqueness of formal solutions to ordinary differential equations is neither new to state nor difficult to prove: one can simply regard a formal vector field X as a tuple of power series and manually find the power series solution to $\dot{x}(t) = X(x(t))$ with initial value prescribed at $t = 0$ by comparing coefficients on both sides. However, we could not locate a formulation of this theorem in the literature that both regards formal flows as acting on formal functions and also considers the resulting action on formal tensor fields. Moreover, formal flows of time-dependent vector fields seem less commonly discussed, but will be essential for our version of Moser’s trick. As such, we supply a statement and a more conceptual proof.

THEOREM 2.6. *For every $X_t \in \mathfrak{X}_t$, there exists a unique formal isotopy $\varphi_t: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$, called the **flow** of X_t , satisfying the following two properties:*

(IC) *Initial condition: its $t = 0$ time slice $\varphi_0: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ is the identity map.*

(FP) *Flow property: for all $f \in \mathbb{K}[[x]]$, we have*

$$\frac{\partial}{\partial t}(\varphi_t(f)) = \varphi_t(X_t(f)).$$

Remark 2.7. Conversely, every formal isotopy φ_t such that $\varphi_0 = \text{id}_{\mathbb{K}[[x]]}$ is generated by a unique time-dependent formal vector field $X_t \in \mathfrak{X}_t$, but this is easy to see: the formula

$$X_t: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]], \quad X_t(f) := \varphi_t^{-1} \left(\frac{\partial}{\partial t}(\varphi_t(f)) \right) - \frac{\partial f}{\partial t}$$

defines a derivation of $\mathbb{K}[[t, x]]$ that vanishes on t , and the flow property holds for all $f \in \mathbb{K}[[x]]$ by construction. (The significance of the extra $-\frac{\partial f}{\partial t}$ term will be explained by Proposition 2.8.) $\blacklozenge$

Proof. Step 1: uniqueness. Given $m \in \mathbb{N}$ and formal isotopies φ_t, ψ_t satisfying (IC) and (FP), observe that

$$S_m := \{f \in \mathbb{K}[[t, x]] \mid t^m \text{ divides } \varphi_t(f) - \psi_t(f)\}$$

is a $\mathbb{K}$ -subalgebra of $\mathbb{K}[[t, x]]$ by linearity and the identity $\varphi_t(fg) - \psi_t(fg) = \varphi_t(f)[\varphi_t(g) - \psi_t(g)] + \psi_t(g)[\varphi_t(f) - \psi_t(f)]$. Moreover, $\bigcap_m S_m$ is exactly the set on which φ_t and ψ_t agree. We will prove by induction that $S_m = \mathbb{K}[[t, x]]$ for each m , hence $\varphi_t = \psi_t$.

The base case $m = 0$ is vacuously true. For the inductive step, assume $m \in \mathbb{N}$ is such that $S_m = \mathbb{K}[[t, x]]$, so that in particular $X_t(x^i) \in S_m$ for each i . We claim that S_{m+1} contains t and each x^i ; thus S_{m+1} contains the polynomial subalgebra $\mathbb{K}[t, x^1, \dots, x^n]$; linearity and continuity then implies that $S_{m+1} = \mathbb{K}[[t, x]]$. Since formal isotopies fix t , we have $t \in S_{m+1}$. By (IC), each of the power series $\Delta^i(t, x) := \varphi_t(x^i) - \psi_t(x^i)$ satisfy $\Delta^i(0, x) = 0$. By (FP),

$$\frac{\partial \Delta^i}{\partial t} = \varphi_t(X_t(x^i)) - \psi_t(X_t(x^i)),$$

so t^m divides $\frac{\partial \Delta^i}{\partial t}$. Combining these two facts, we see that t^{m+1} divides Δ^i , hence $x^i \in S_{m+1}$.

Step 2: existence for time-independent vector fields X . We claim that the flow of X is given by the formal exponential

$$e^{tX} : \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]], \quad e^{tX}(f) := \sum_{k=0}^{\infty} \frac{t^k}{k!} X^{\circ k}(f),$$

where $X^{\circ k} := X \circ \dots \circ X$ denotes the k -fold iteration of X (the identity for $k = 0$). Each term in the infinite sum is well-defined ($k!$ is invertible as $\mathbb{K}$ is a field of characteristic zero) and the sum converges in the $\mathfrak{m}_{t,x}$ -adic topology as $t^k X^{\circ k}(f) \in \mathfrak{m}_{t,x}^k$ for each k . It is straightforward to verify that e^{tX} is a $\mathbb{K}$ -algebra automorphism with inverse $e^{t(-X)}$; we omit the computations, but note that one uses the higher-order Leibniz rule $X^{\circ k}(fg) = \sum_{i=0}^k \binom{k}{i} X^{\circ i}(f) X^{\circ(k-i)}(g)$ to show that $e^{tX}(fg) = e^{tX}(f)e^{tX}(g)$, and one uses the vanishing of alternating sums of binomial coefficients $\sum_{i=0}^k \frac{(-1)^i}{i!(k-i)!} = 0$ for $k > 0$ to show that $e^{t(-X)} \circ e^{tX} = \text{id}$. Since $X^{\circ k}(t) = 0$ for all $k > 0$, we have $e^{tX}(t) = t$, so e^{tX} is a formal isotopy of $\mathbb{K}^n$. Setting $t = 0$ in the exponential formula immediately shows that e^{tX} satisfies (IC). For (FP), fix $f \in \mathbb{K}[[x]]$. Continuity of $\frac{\partial}{\partial t}$ yields

$$\frac{\partial}{\partial t}(e^{tX}(f)) = \sum_{k=0}^{\infty} \frac{\partial}{\partial t} \left(\frac{t^k}{k!} X^{\circ k}(f) \right) = \sum_{k=0}^{\infty} \frac{k}{k!} t^{k-1} X^{\circ k}(f) + \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{\partial}{\partial t} (X^{\circ k}(f)).$$

The second sum on the right-hand side vanishes since X and f are time-independent: $[\frac{\partial}{\partial t}, X] = 0$, hence $\frac{\partial}{\partial t}(X^{\circ k}(f)) = X^{\circ k}(\frac{\partial f}{\partial t}) = 0$. Re-indexing the first sum then yields

$$\frac{\partial}{\partial t}(e^{tX}(f)) = \sum_{k=0}^{\infty} \frac{t^k}{k!} X^{\circ(k+1)}(f) = e^{tX}(X(f)).$$

Step 3: existence for time-dependent vector fields X_t . We mimic the standard procedure for converting a non-autonomous system of n ordinary differential equations into an autonomous system of $n + 1$ equations. That is, temporarily introduce an additional formal variable s and let X_s denote the same vector field as X_t , but with every instance of t replaced with s . Then $\tilde{X} := \frac{\partial}{\partial s} + X_s$ is a derivation of $\mathbb{K}[[s, x]]$ —a time-independent formal vector field on $\mathbb{K}^{n+1}$ —which

has a unique flow $\Phi_t = e^{t\tilde{X}}: \mathbb{K}[[t, s, x]] \rightarrow \mathbb{K}[[t, s, x]]$ by Steps 1 and 2. We claim that the desired flow of X_t is given by the composition

$$\varphi_t: \mathbb{K}[[t, x]] \hookrightarrow \mathbb{K}[[t, s, x]] \xrightarrow{\Phi_t} \mathbb{K}[[t, s, x]] \xrightarrow{\text{set } s=0} \mathbb{K}[[t, x]].$$

Each map in this composition is a $\mathbb{K}$ -algebra homomorphism that fixes t , so the same is true of φ_t . Since the i^{th} component of φ_t is

$$\Phi_t(x^i)|_{s=0} = e^{t\tilde{X}}(x^i)|_{s=0} = x^i + t\tilde{X}(x^i)|_{s=0} + \frac{t^2}{2!}\tilde{X}^{\circ 2}(x^i)|_{s=0} + \cdots,$$

it follows that $D_x\varphi(0, 0)$ is the identity matrix; thus φ_t is a formal isotopy by (2.4).

To see that φ_t satisfies the initial condition (IC), observe that φ_0 is given by following the top row of arrows left-to-right in the following diagram; it is equal to the identity on $\mathbb{K}[[x]]$ because the diagram commutes (the bottom trapezoid commutes because Φ_t satisfies (IC) on $\mathbb{K}[[s, x]]$):

$$\begin{array}{ccccc} & \mathbb{K}[[t, x]] & \xrightarrow{\varphi_t} & \mathbb{K}[[t, x]] & \\ & \swarrow & & \searrow & \\ \mathbb{K}[[x]] & & \mathbb{K}[[t, s, x]] \xrightarrow{\Phi_t} \mathbb{K}[[t, s, x]] & & \mathbb{K}[[x]] \\ & \searrow & \swarrow & \searrow & \swarrow \\ & \mathbb{K}[[s, x]] & \xrightarrow{\text{id}} & \mathbb{K}[[s, x]] & \end{array}$$

$\begin{array}{ccc} & s=0 & \\ & \swarrow & \searrow \\ & \mathbb{K}[[x]] & \end{array} \quad \begin{array}{ccc} & t=0 & \\ & \swarrow & \searrow \\ & \mathbb{K}[[s, x]] & \end{array}$

To see that φ_t satisfies the flow property (FP), observe that

$$S := \{f \in \mathbb{K}[[x]] \mid s - t \text{ divides } \tilde{X}(f) - X_t(f)\}$$

is a $\mathbb{K}$ -subalgebra of $\mathbb{K}[[x]]$ by linearity and the Leibniz rule. Moreover, $x^i \in S$ for each i as the terms of $\tilde{X}(x^i) - X_t(x^i)$ can be grouped into pairs that all have a factor of $s^k - t^k = (s - t)(s^{k-1} + s^{k-2}t + \cdots + st^{k-2} + t^{k-1})$. Thus S contains the polynomial algebra $\mathbb{K}[x^1, \dots, x^n]$; linearity and continuity then implies that $S = \mathbb{K}[[x]]$. Given $f \in \mathbb{K}[[x]]$, writing $\tilde{X}(f) - X_t(f) = (s - t)g$ for some $g \in \mathbb{K}[[t, s, x]]$ yields

$$\Phi_t(\tilde{X}(f) - X_t(f)) = \Phi_t(s - t)\Phi_t(g) = s\Phi_t(g),$$

the last step using that $\Phi_t(s) = e^{t\tilde{X}}(s) = s + t$. Since $\frac{\partial}{\partial t}$ commutes with setting $s = 0$ and Φ_t satisfies (FP) on $\mathbb{K}[[s, x]]$ (but with the vector field $\tilde{X}$), it follows that

$$\begin{aligned} \frac{\partial}{\partial t}(\varphi_t(f)) &= \frac{\partial}{\partial t}(\Phi_t(f)|_{s=0}) \\ &= \left[\frac{\partial}{\partial t}(\Phi_t(f)) \right] \Big|_{s=0} \\ &= \Phi_t(\tilde{X}(f))|_{s=0} \\ &= [\Phi_t(X_t(f)) + s\Phi_t(g)]|_{s=0} \\ &= \varphi_t(X_t(f)). \end{aligned}$$

■

The flow property for time-independent formal functions generalizes to a flow property for all formal tensor fields (time-independent or dependent), yielding a formal version of Equation (1.3):

PROPOSITION 2.8. *If φ_t is any formal isotopy satisfying the flow property (FP) for X_t , then*

$$\frac{d}{dt}(\varphi_t^* A_t) = \varphi_t^* \left(\frac{dA_t}{dt} + \mathcal{L}_{X_t} A_t \right), \quad \forall A_t \in \mathcal{T}_t^{(r,s)}.$$

Proof. As pullbacks and Lie derivatives are compatible (Equation (2.3)), we have

$$\varphi_t^* \left(\frac{dA_t}{dt} + \mathcal{L}_{X_t} A_t \right) = \varphi_t^* \left(\mathcal{L}_{\frac{\partial}{\partial t} + X_t} A_t \right) = \mathcal{L}_{\varphi_t^* \left(\frac{\partial}{\partial t} + X_t \right)} (\varphi_t^* A_t).$$

Thus it suffices to show that $\frac{\partial}{\partial t} = \varphi_t^* \left(\frac{\partial}{\partial t} + X_t \right)$, or equivalently, that $\frac{\partial}{\partial t} \circ \varphi_t = \varphi_t \circ \left(\frac{\partial}{\partial t} + X_t \right)$. Given $f \in \mathbb{K}[[t, x]]$, write $f(t, x) = \sum_{k=0}^{\infty} t^k f_k(x)$ where each $f_k(x) \in \mathbb{K}[[x]]$ does not depend on t . Since $\frac{\partial}{\partial t}$ and φ_t are continuous and φ_t fixes t , we have

$$\frac{\partial}{\partial t}(\varphi_t(f)) = \sum_k \frac{\partial}{\partial t}(t^k \varphi_t(f_k)) = \sum_k k t^{k-1} \varphi_t(f_k) + \sum_k t^k \frac{\partial}{\partial t}(\varphi_t(f_k)).$$

The first sum is equal to $\varphi_t \left(\frac{\partial f}{\partial t} \right)$. For the second sum, apply (FP) to each f_k :

$$\frac{\partial}{\partial t}(\varphi_t(f)) = \varphi_t \left(\frac{\partial f}{\partial t} \right) + \sum_k t^k \varphi_t(X_t(f_k)) = \varphi_t \left(\frac{\partial f}{\partial t} + \sum_k t^k X_t(f_k) \right) = \varphi_t \left(\frac{\partial f}{\partial t} + X_t(f) \right),$$

the last step using the Leibniz rule, the fact that $X_t(t^k) = 0$ for all $k \in \mathbb{N}$, and continuity of X_t . ■

COROLLARY 2.9. *If φ_t is the flow of X_t , then φ_t^{-1} is the flow of $-\varphi_t^* X_t$.*

Proof. By Lemma 2.5 and Theorem 2.6(IC), we have $(\varphi_t^{-1})_0 = (\varphi_0)^{-1} = \text{id}_{\mathbb{K}[[x]]}$. By Remark 2.7, φ_t^{-1} is the flow of the time-dependent formal vector field

$$Y_t := \varphi_t \circ \frac{\partial}{\partial t} \circ \varphi_t^{-1} - \frac{\partial}{\partial t} = \varphi_t^* \left(\frac{\partial}{\partial t} \right) - \frac{\partial}{\partial t}.$$

The proof of Proposition 2.8 showed that $\frac{\partial}{\partial t} = \varphi_t^* \left(\frac{\partial}{\partial t} + X_t \right) = \varphi_t^* \left(\frac{\partial}{\partial t} \right) + \varphi_t^* X_t$, so $Y_t = -\varphi_t^* X_t$. ■

2.4. Fixed-time evaluation. We now revisit the fundamental question posed in Section 1.2—or rather, its formal analogue:

Given formal tensor fields $A_0, A_1 \in \mathcal{T}^{(r,s)}$, when does there exist a formal diffeomorphism $\varphi: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ such that $\varphi^ A_1 = A_0$?*

To emulate Moser's trick, one would like to view A_0 and A_1 as time slices at $t = 0$ and $t = 1$ of a one-parameter family of formal tensor fields $A_t: \mathbb{K} \rightarrow \mathcal{T}^{(r,s)}$; ditto for φ . However, Moser-type arguments involve time derivatives such as $\frac{dA_t}{dt}$ and $\frac{d}{dt}(\varphi_t^* A_t)$, and it is unclear how to make sense of these for arbitrary functions $\mathbb{K} \rightarrow \mathcal{T}^{(r,s)}$ over an arbitrary field of characteristic zero. On the other hand, the work of Section 2.2 gives meaning to these expressions if $A_t \in \mathcal{T}_t^{(r,s)}$ is a time-dependent formal tensor field and $\varphi_t: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$ is a formal isotopy, and the work of Section 2.3 allows us to say that the equation $\frac{d}{dt}(\varphi_t^* A_t) = 0$ is equivalent to the equation $\frac{dA_t}{dt} + \mathcal{L}_{X_t} A_t = 0$ where $X_t \in \mathfrak{X}_t$ generates φ_t (assuming that $\varphi_0 = \text{id}_{\mathbb{K}[[x]]}$). But then comes another issue, which we have already alluded to: while time slices of A_t , φ_t , and X_t at $t = 0$ are perfectly well-defined, replacing the formal variable t with any non-zero $\tau \in \mathbb{K}$ immediately leads to significant complications. We illustrate these further with some examples in dimension $n = 1$.

EXAMPLE 2.10. Even if substituting $t = \tau$ into a formal isotopy yields well-defined power series in x , there are two distinct ways in which the result can fail to define a formal diffeomorphism. One way is illustrated by the flow of the formal vector field $\frac{\partial}{\partial x}$, which is given by $\varphi_t(x) = x + t$: replacing t with any $\tau \neq 0$ yields a power series with non-zero constant term, which cannot correspond to a $\mathbb{K}$ -algebra endomorphism of $\mathbb{K}[[x]]$ via (2.1) at all. Another way is illustrated by the formal isotopy given by $\psi_t(x) = x(1 - t)$: substituting $t = 1$ yields the zero power series, which defines a $\mathbb{K}$ -algebra endomorphism of $\mathbb{K}[[x]]$ that is not invertible. Relatedly, since $\psi_t^{-1}(x) = x(1 - t)^{-1} = x \sum_{k=0}^{\infty} t^k$, the pullback of $\frac{\partial}{\partial x}$ along ψ_t is $\sum_{k=0}^{\infty} t^k \frac{\partial}{\partial x}$, which is ill-defined at $t = 1$. ♦

EXAMPLE 2.11. Regarding the Euler vector field $x \frac{\partial}{\partial x}$ as a formal vector field, its flow is given by

$$\varphi_t(x) = \sum_{k=0}^{\infty} \frac{t^k}{k!} x. \quad (2.5)$$

One might try to make sense of “substituting $t = 1$ into φ_t ” by declaring φ_1 to correspond via (2.1) to the reasonable-looking expression $\sum_{k=0}^{\infty} \frac{1}{k!} x$. However, *this infinite sum does not converge in the $\mathfrak{m}$ -adic topology on $\mathbb{K}[[x]]$* , as $\Theta(\frac{1}{k!}x) = 1$ for all k . If the base field comes with a topology such that $\sum_k \frac{1}{k!}$ converges to a unique $e \in \mathbb{K}$ (e.g., $\mathbb{K}$ contains the *real number* $e = 2.718\dots$), then one could declare φ_1 to be the formal diffeomorphism given by the power series $\varphi_1(x) := ex \in \mathbb{K}[[x]]$.

We will not do this for both practical and philosophical reasons. Practically speaking, $\mathbb{K}$ need not be topologized; even if it is, there may be power series that diverge for *all* non-zero $\tau \in \mathbb{K}$.³ Philosophically speaking, using a topology on $\mathbb{K}$ is incongruous with our approach: determining whether an infinite sum converges is generally a difficult task, and the entire *raison d’être* for working with formal power series is to blithely ignore such concerns. Indeed, under the identification of sets $\mathbb{K}[[x]] \cong \mathbb{K}^{\mathbb{N}}$ associating $\sum_{\alpha} c_{\alpha} x^{\alpha} \in \mathbb{K}[[x]]$ to the function $\mathbb{N}^n \rightarrow \mathbb{K}$, $\alpha \mapsto c_{\alpha}$, the $\mathfrak{m}$ -adic topology on $\mathbb{K}[[x]]$ corresponds to the product topology on $\mathbb{K}^{\mathbb{N}^n}$ when $\mathbb{K}$ is equipped with the *discrete* topology. Thus $\mathbb{K}[[x]]$ does not see any interesting topology on $\mathbb{K}$ anyways. ♦

While we cannot interpret arbitrary formal power series $f \in \mathbb{K}[[t, x]]$ as “functions” $\mathbb{K} \rightarrow \mathbb{K}[[x]]$, $t \mapsto f(t, x)$ and “evaluate” them at non-zero $\tau \in \mathbb{K}$, a simple way to sidestep this issue is by requiring these “functions” to be polynomial in t , i.e., come from the subalgebra $\mathbb{K}[[x]][t] \subset \mathbb{K}[[t, x]]$. Such polynomials can be regarded as honest-to-goodness functions $\mathbb{K} \rightarrow \mathbb{K}[[x]]$, but this turns out to be too restrictive—for one, non-zero formal vector fields *never* generate $\mathbb{K}$ -algebra automorphisms of $\mathbb{K}[[x]][t]$ (see Remark 2.23). An intermediate approach is to restrict attention to those formal power series such that, *for each fixed* $\alpha \in \mathbb{N}^n$, the coefficient of the monomial x^{α} depends polynomially on t (but the degrees of the polynomials may be unbounded as α varies). This is essentially what we shall do, with some additional conditions for formal isotopies:

DEFINITION 2.12. We define **evaluative** formal tensor fields and formal isotopies as follows. A time-dependent formal function $f(t, x) \in \mathbb{K}[[t, x]]$ is called *evaluative* if $f(t, x) \in \mathbb{K}[t][[x]]$; that is,

$$f(t, x) = \sum_{\alpha \in \mathbb{N}^n} p_{\alpha}(t) x^{\alpha}, \quad \text{where each } p_{\alpha}(t) \text{ is a polynomial (as opposed to a power series).}$$

A time-dependent formal tensor field $A_t \in \mathcal{T}_t^{(r,s)}$ is called *evaluative* if all of its component power series are evaluative. A formal isotopy $\varphi_t: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$, identified with an n -tuple $\varphi(t, x) = (\varphi^1(t, x), \dots, \varphi^n(t, x)) \in \mathbb{K}[[t, x]]^n$ via (2.4), is called *evaluative* if the following two conditions hold:

- (i) All n component power series $\varphi^i(t, x)$ are evaluative and satisfy $\varphi^i(t, 0) = 0$.
- (ii) The matrix $D_x \varphi(t, x)$ is invertible in $\text{Mat}_n(\mathbb{K}[t][[x]])$. Equivalently, $D_x \varphi(t, 0)$ is invertible in $\text{Mat}_n(\mathbb{K}[t])$, or $\det D_x \varphi(t, 0)$ is a non-zero constant polynomial.

EXAMPLE 2.13. For any polynomial $p(t) \in \mathbb{K}[t]$, the time-dependent formal vector field in dimension $n = 1$ given by $X_t = p(t)x^2 \frac{\partial}{\partial x}$ is evaluative. The flow of X_t corresponds via (2.4) to the evaluative power series

$$\varphi(t, x) := x(1 - P(t)x)^{-1} = \sum_{k=0}^{\infty} P(t)^k x^{k+1}, \quad \text{where } P(t) := \int_0^t p(s) ds \in \mathbb{K}[t].$$

Since $\varphi(t, 0) = 0$ and $D_x \varphi(t, 0) = 1$, the flow of X_t is an evaluative formal isotopy. Note that $\varphi(t, x) \notin \mathbb{K}[[x]][t]$ unless $p = 0$. ♦

³For example, Euler noted how the series $\sum_{k=1}^{\infty} (k-1)!x^k$, which diverges for all non-zero $x \in \mathbb{R}$, appears when studying the system of differential equations $\dot{x} = x^2$ and $\dot{y} = y - x$ [2, §24].

For any $\tau \in \mathbb{K}$, the map $\mathbb{K}[t] \rightarrow \mathbb{K}$ given by evaluation of polynomials at $t = \tau$ extends coefficient-wise to a $\mathbb{K}[[x]]$ -algebra homomorphism

$$\text{eval}_\tau: \mathbb{K}[t][[x]] \rightarrow \mathbb{K}[[x]], \quad f(t, x) \mapsto f(\tau, x).$$

Given an evaluative formal tensor field $A_t \in \mathcal{T}_t^{(r,s)}$, we denote the formal tensor field obtained by evaluating each of its components at $t = \tau$ by $A_\tau \in \mathcal{T}^{(r,s)}$ (or $A_t|_{t=\tau}$, if confusion is possible). Given an evaluative formal isotopy φ_t with corresponding n -tuple $\varphi(t, x) = (\varphi^1(t, x), \dots, \varphi^n(t, x)) \in \mathbb{K}[t, x]^n$, condition (i) in Definition 2.12 implies that $\varphi(\tau, x) := (\varphi^1(\tau, x), \dots, \varphi^n(\tau, x)) \in \mathbb{K}[[x]]^n$ is well-defined and corresponds to a unique $\mathbb{K}$ -algebra endomorphism of $\mathbb{K}[[x]]$ via (2.1) that we denote by φ_τ or $\varphi_t|_{t=\tau}$. Condition (ii) guarantees that φ_τ is a formal diffeomorphism. In particular, $A_t|_{t=0}$ and $\varphi_t|_{t=0}$ respectively agree with the time zero slices A_0 and φ_0 as defined in Section 2.2. In general, we call A_τ and φ_τ a *(fixed) time slice* of A_t and φ_t , respectively.

Remark 2.14. One might consider relaxing condition (ii) by only requiring $\det D_x \varphi(\tau, 0)$ to be a non-zero element of $\mathbb{K}$ for each fixed $\tau \in \mathbb{K}$, as this still implies that φ_τ is a formal diffeomorphism when combined with condition (i). However, even if the polynomial function $\tau \mapsto \det D_x \varphi(\tau, 0)$ is nowhere-vanishing on $\mathbb{K}$, it can have zeroes in a field extension of $\mathbb{K}$, and this can lead to undesirable phenomena. For example, the formal isotopy in dimension $n = 1$ given by $\varphi_t(x) = x(t^2 + 1)$ satisfies condition (i) and this weaker version of condition (ii) over $\mathbb{R}$, but not over $\mathbb{C}$; its inverse does not satisfy condition (i) as $\varphi_t^{-1}(x) = x(t^2 + 1)^{-1} = x \sum_{k=0}^{\infty} (-1)^k t^{2k}$ is not evaluative.

For completeness, we also note that there is no benefit to relaxing condition (i) by only requiring that $\varphi^i(\tau, 0) = 0$ for each fixed $\tau \in \mathbb{K}$: since $\mathbb{K}$ is an infinite field, this is equivalent to $\varphi^i(t, 0)$ being the zero polynomial. $\blacklozenge$

A formal isotopy is evaluative if and only if it is the extension of a $\mathbb{K}$ -algebra automorphism of $\mathbb{K}[t][[x]]$ that fixes t , i.e., a $\mathbb{K}[t]$ -algebra automorphism of $\mathbb{K}[t][[x]]$. A time-dependent formal vector field is evaluative if and only if it is the extension of a derivation of $\mathbb{K}[t][[x]]$ that vanishes on t , i.e., a $\mathbb{K}[t]$ -linear derivation of $\mathbb{K}[t][[x]]$. (See Remark 2.16 for more details.) In theory, we could repeat the developments of Sections 2.1 and 2.2 with $\mathbb{K}[t][[x]]$ in mind, *mutatis mutandis*. We list, without proof, some results about evaluative objects obtained by doing so:

PROPOSITION 2.15. Fix $A_t \in \mathcal{T}_t^{(r,s)}$, a formal isotopy $\varphi_t: \mathbb{K}[t, x] \rightarrow \mathbb{K}[t, x]$, and $\tau \in \mathbb{K}$.

- (i) If φ_t is evaluative, then so is φ_t^{-1} . If ψ_t is also an evaluative formal isotopy, then so is $\psi_t \circ \varphi_t$. Evaluation at $t = \tau$ commutes with inversion and composition:

$$(\varphi_t^{-1})|_{t=\tau} = (\varphi_\tau)^{-1} \quad \text{and} \quad (\psi_t \circ \varphi_t)|_{t=\tau} = \psi_\tau \circ \varphi_\tau.$$

- (ii) If φ_t and A_t are both evaluative, then so is $\varphi_t^* A_t$. Evaluation commutes with pullback:

$$(\varphi_t^* A_t)|_{t=\tau} = (\varphi_\tau)^*(A_\tau).$$

- (iii) Tensor products, contractions, Lie derivatives along evaluative formal vector fields, and time derivatives of evaluative formal tensor fields are evaluative. Evaluation commutes with the first three operations (but not with $\frac{d}{dt}$).

Remark 2.16. Let I be the ideal of $\mathbb{K}[t][[x]]$ generated by the x^i 's, and let J be the ideal of $\mathbb{K}[t][[x]]$ generated by t and the x^i 's.⁴ The I -adic topology on $\mathbb{K}[t][[x]]$ is the typical choice for topologizing a power series algebra, while the J -adic topology coincides with the subspace topology inherited from the $\mathfrak{m}_{t,x}$ -adic topology on $\mathbb{K}[[t, x]]$. The I -adic topology is strictly finer than the J -adic topology. For example, the sequence t^k does not converge in the I -adic topology as $k \rightarrow \infty$, but it converges to 0 in the J -adic topology. As a consequence, $\text{eval}_\tau: \mathbb{K}[t][[x]] \rightarrow \mathbb{K}[[x]]$ is continuous when $\mathbb{K}[t][[x]]$

⁴Maximal ideals of $\mathbb{K}[t][[x]]$ are precisely those generated by an irreducible polynomial $p(t) \in \mathbb{K}[t]$ and the x^i 's, so J is maximal while I is not.

is equipped with the I -adic topology, but is discontinuous for $\tau \neq 0$ when $\mathbb{K}[t][x]$ is equipped with the J -adic topology because the sequence $\text{eval}_\tau(t^k) = \tau^k$ does not converge in the $\mathfrak{m}$ -adic topology of $\mathbb{K}[[x]]$ as $k \rightarrow \infty$.

For evaluative objects, the difference turns out to not matter much. For example, every $\mathbb{K}[t]$ -linear derivation X of $\mathbb{K}[t][x]$ is automatically continuous with respect to the I -adic topology, from which it follows that such derivations form a free $\mathbb{K}[t][x]$ -module with basis $\left\{\frac{\partial}{\partial x^i}\right\}$ —the proof is exactly the same as for derivations of $\mathbb{K}[[x]]$ [5, IV§4.6]. Expanding X in this basis extends X to a derivation of $\mathbb{K}[[t, x]]$. This extension is continuous in the $\mathfrak{m}_{t,x}$ -adic topology, so X is also continuous with respect to the subspace (i.e., J -adic) topology.

Similarly, every $\mathbb{K}[t]$ -algebra endomorphism φ of $\mathbb{K}[t][x]$ is automatically continuous with respect to both the I -adic and J -adic topologies, but we cannot justify this using Remark 2.1 as $\mathbb{K}[t]$ is not a field. We use a degree argument instead: for each i , $1 - tx^i$ is invertible in $\mathbb{K}[t][x]$ as its constant term (with respect to the x 's) is invertible in $\mathbb{K}[t]$. Thus $\varphi(1 - tx^i) = 1 - t\varphi(x^i)$ is also invertible in $\mathbb{K}[[t, x]]$. If $p^i(t) \in \mathbb{K}[t]$ denotes the constant term of $\varphi(x^i)$, then the constant term of $1 - t\varphi(x^i)$ is $1 - tp^i(t)$, which is therefore invertible in $\mathbb{K}[t]$. By degree considerations, $p^i(t) = 0$, hence $\varphi(x^i) \in I$, hence $\varphi(I) \subseteq I$. As $\varphi(t) = t$, we also have $\varphi(J) \subseteq J$, so φ is both I -adically and J -adically continuous. See [11] for what can happen if one relaxes $\mathbb{K}[t]$ -linearity to $\mathbb{K}$ -linearity. ♦

By identifying a time-independent (hence evaluative) formal tensor field with all of its fixed time slices, we arrive at a formal version of Moser's trick:

COROLLARY 2.17 (FORMAL MOSER TRICK). *Let $A_t \in \mathcal{T}_t^{(r,s)}$ be an evaluative formal tensor field. If there exists $X_t \in \mathfrak{X}_t$ such that the flow $\varphi_t: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$ of X_t is evaluative and*

$$\frac{dA_t}{dt} + \mathcal{L}_{X_t}A_t = 0,$$

*then for all $\tau \in \mathbb{K}$, the formal tensor fields $A_0, A_\tau \in \mathcal{T}^{(r,s)}$ and the formal diffeomorphism $\varphi_\tau: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ are related by $\varphi_\tau^*A_\tau = A_0$.*

Proof. We have $\frac{d}{dt}(\varphi_t^*A_t) = 0$ by Proposition 2.8, so $\varphi_t^*A_t$ is time-independent. By Proposition 2.15(ii) and Theorem 2.6(IC), $\varphi_\tau^*A_\tau = (\varphi_t^*A_t)|_{t=\tau} = (\varphi_t^*A_t)|_{t=0} = A_0$. ■

Remark 2.18. Differential forms and multivector fields can be imported into the formal setting by considering exterior powers of the $\mathbb{K}[[x]]$ -modules Ω^1 and $\mathfrak{X}$. Naturally, one can then proceed to develop the full apparatus of a formal Cartan calculus (exterior differentials, interior products, Lie derivatives, Schouten brackets, etc.), as well as time-dependent and evaluative versions, thus obtaining a formal Moser trick for these objects as well. ♦

In general, if the flow of $X_t \in \mathfrak{X}_t$ is evaluative, then X_t is evaluative by Remark 2.7 and Proposition 2.15. However, an evaluative (or even time-independent) formal vector field need not have an evaluative flow, as demonstrated by Examples 2.10 and 2.11. To employ the formal Moser trick, we thus seek conditions guaranteeing the flow of X_t to be evaluative. Combining the next two results provides a partial answer.

PROPOSITION 2.19. *Let φ_t be a formal isotopy of $\mathbb{K}^n$ with corresponding n -tuple written in the form $\varphi(t, x) = \sum_{k=0}^\infty t^k \varphi_k(x)$ where $\varphi_k(x) \in \mathbb{K}[[x]]^n$ (c.f. Remark 2.4). If $\Theta(\varphi_k) \geq 2$ for all $k \geq 1$ and $\lim_{k \rightarrow \infty} \Theta(\varphi_k) = \infty$, then φ_t is evaluative.*

Remark 2.20 (a partial converse). If an n -tuple $\varphi(t, x) = \sum_k t^k \varphi_k(x)$ corresponds to an evaluative formal isotopy, then $\Theta(\varphi_k) \geq 1$ for all $k \geq 1$ and $\lim_{k \rightarrow \infty} \Theta(\varphi_k) = \infty$. It is not necessarily true that $\Theta(\varphi_k) \geq 2$ for $k \geq 1$, as demonstrated by the evaluative formal isotopy of $\mathbb{K}^2$ corresponding to the 2-tuple $(x + ty, y)$ (the flow of $y \frac{\partial}{\partial x}$). ♦

PROPOSITION 2.21. Suppose $X_t \in \mathfrak{X}_t$ is written in the form $X_t = \sum_{k=0}^{\infty} t^k X_k$ where $X_k \in \mathfrak{X}$. Identify the flow φ_t of X_t with an n -tuple $\varphi(t, x) = \sum_{k=0}^{\infty} t^k \varphi_k(x)$ where $\varphi_k(x) \in \mathbb{K}[[x]]^n$. If $\Theta(X_k(x^i)) \geq k+2$ for all $k \geq 0$ and $i \in \{1, \dots, n\}$, then $\Theta(\varphi_k) \geq k+1$ for all $k \geq 0$. In particular, φ_t is an evaluative formal isotopy by Proposition 2.19.

Notice that the Euler vector field $x \frac{\partial}{\partial x}$ (Example 2.11) just barely fails the hypothesis of Proposition 2.21 for $k = 0$, but the conclusion drastically fails: the flow (2.5) has $\Theta(\varphi_k) = 1$ for all k . Propositions 2.19 and 2.21 will be generalized in Section 3.2, so we omit their proofs for now.

We conclude Section 2 by drawing some comparisons between vector fields whose flows are evaluative and two well-known types of derivations that generate automorphisms of algebras similar to $\mathbb{K}[[t]][x]$: one obtained by swapping which variables are allowed to appear with unbounded degrees ($\mathbb{K}[[t]][x]$), and one obtained by swapping the order in which the variables are adjoined ($\mathbb{K}[[x]][t]$).

Remark 2.22. A time-independent C^1 vector field X on $\mathbb{R}^n$ is called *poly-flow* if its time t flow φ_t is a polynomial mapping (i.e., the n components of φ_t are all polynomials in x) wherever defined. Bass–Meisters [3, Theorem 4.1] proved that this condition implies the following: X is complete and its component functions are polynomials; the degrees of the components of φ_t are bounded above as $t \in \mathbb{R}$ varies; the coefficients of these polynomials are analytic functions of t . Thus, the flow of a poly-flow vector field corresponds to an automorphism of $\mathbb{R}[[t]][x]$. Bass–Meisters also show that $\det D_x \varphi(t, x)$ depends only on t , not x , hence is equal to $\det D_x \varphi(t, 0)$. This is somewhat orthogonal to our definition of an evaluative formal isotopy, in which we allow the Jacobian determinant to depend on x , but not on t when $x = 0$.

If $\mathbb{K}$ is any field of characteristic zero and $f: \mathbb{K}^n \rightarrow \mathbb{K}^n$ is a bijective function such that f and f^{-1} are both polynomial mappings, then the chain rule implies that $\det Df(x)$ is an invertible element of $\mathbb{K}[x]$, hence a non-zero constant. The famous *Jacobian conjecture* asks: if f is a polynomial mapping and $\det Df(x)$ is a non-zero constant, must f be bijective with its inverse also a polynomial mapping? For $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$, Meisters–Olech [15] have shown that the Jacobian conjecture holding for a given f is equivalent to the assertion that, for each $a \in \mathbb{K}^n$, the flow $\varphi_t(x; a)$ of the time-independent vector field $X_a := Df(x)^{-1}a$ is polynomial in both x and t (where t is a complex time parameter if $\mathbb{K} = \mathbb{C}$). They construct $f^{-1}(a)$ for any given $a = (a^1, \dots, a^n)$ in a fashion similar to evaluation: denoting the standard basis vectors of $\mathbb{K}^n$ by e_i , start at the origin $\alpha_0 = (0, \dots, 0)$ and successively flow along each X_{e_i} for time a^i to obtain the sequence of points $\alpha_i := \varphi_{a^i}(\alpha_{i-1}; e_i)$. Then $f^{-1}(a) = \alpha_n$, which is polynomial in a if each flow is polynomial in both space and time.

We speculate that it might be possible to perform their construction over more general fields of characteristic zero via a formal Moser-type argument, with a caveat: the formal flow of X_a being polynomial in x and t does not imply that it is evaluative. For example, the polynomial mapping $f(x, y) = (x + y^2, y)$ has polynomial inverse $f^{-1}(u, v) = (u - v^2, v)$; for this f , the flow of $X_{(a,b)}$ corresponds to the 2-tuple $(x + (a - 2by)t - b^2t^2, y + bt)$, which satisfies condition (ii) of Definition 2.12 but not condition (i) (unless $(a, b) = (0, 0)$). ♦

Remark 2.23. A derivation X of an algebra A is called *locally nilpotent* [9] if, for each $a \in A$, there exists $k \in \mathbb{N}$ such that $X^{ok}(a) = 0$. Equivalently, extend X to a derivation of $A[[t]]$ by defining $X(\sum_k a_k t^k) = \sum_k X(a_k) t^k$; then X is locally nilpotent if and only if the algebra automorphism $e^{tX}: A[[t]] \rightarrow A[[t]]$ defined exactly as in Theorem 2.6 preserves $A[t]$. However, *the only locally nilpotent derivation of $A = \mathbb{K}[[x]]$ is the zero derivation*.

Indeed, suppose φ_t is any $\mathbb{K}$ -algebra endomorphism of $\mathbb{K}[[t, x]]$ that fixes t and satisfies (IC) (e.g., the flow of any formal vector field). If φ_t preserves $\mathbb{K}[[x]][t]$, then $\varphi_t(1 - x^i) = 1 - \varphi_t(x^i)$ is invertible in $\mathbb{K}[[x]][t]$ as $1 - x^i$ is invertible in $\mathbb{K}[[x]]$. Thus $\varphi_t(x^i)$ is a constant polynomial with respect to t ; in particular, it equals its $t = 0$ time slice, which is x^i by (IC). Since φ_t fixes t and each x^i , it restricts to the identity map on the polynomial algebra $\mathbb{K}[t, x^1, \dots, x^n]$, which is a dense subset of $\mathbb{K}[[t, x]]$ in the (Hausdorff) $\mathfrak{m}_{t,x}$ -adic topology. By continuity, $\varphi_t = \text{id}_{\mathbb{K}[[t, x]]}$; by Remark 2.7, φ_t is the flow of the zero vector field. ♦

3. WEIGHTINGS OF $\mathbb{K}^n$

3.1. The algebra of weights. We continue to use the shorthand notation $\mathbb{K}[[x]] := \mathbb{K}[[x^1, \dots, x^n]]$ for the algebra of formal power series and $\mathbb{K}[x] := \mathbb{K}[x^1, \dots, x^n]$ for the subalgebra of polynomials. A *polynomial vector field* on $\mathbb{K}^n$ is a derivation of $\mathbb{K}[x]$; the poly-flow vector fields discussed in Remark 2.22 provide examples of such. The set $\mathfrak{X}_{\text{pol}}$ of polynomial vector fields is both a module over $\mathbb{K}[x]$ with a basis given by $\left\{ \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \right\}$ and a Lie algebra over $\mathbb{K}$. We may naturally regard $\mathfrak{X}_{\text{pol}} \subset \mathfrak{X}$ as a Lie subalgebra of the formal vector fields, and when $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, also as a Lie subalgebra of the smooth vector fields on $\mathbb{K}^n$.

DEFINITION 3.1. A **weighting** of $\mathbb{K}^n$ is an n -tuple of integers $w = (w_1, w_2, \dots, w_n)$ such that $1 \leq w_1 \leq w_2 \leq \dots \leq w_n$.

We associate the i^{th} weight w_i with both the formal variable $x^i \in \mathbb{K}[x] \subset \mathbb{K}[[x]]$ and the coordinate function $x^i: \mathbb{K}^n \rightarrow \mathbb{K}$. By extension, the *weighted degree* of a monomial $x^\alpha = (x^1)^{\alpha_1} \dots (x^n)^{\alpha_n}$ with respect to w is $\langle w, \alpha \rangle := \sum_i w_i \alpha_i$. The ideal of $\mathbb{K}[[x]]$ generated by monomials x^α such that $\langle w, \alpha \rangle \geq k$ is denoted by $\mathcal{F}^k$. Note that $\mathcal{F}^k = \mathfrak{m}$ for $1 \leq k \leq w_1$, as $\mathcal{F}^1 \subseteq \mathfrak{m} \subseteq \mathcal{F}^{w_1} \subseteq \mathcal{F}^1$.

Requiring weights to be listed from smallest to largest is done out of convenience to simplify proofs, e.g., those of Proposition 3.4 and Lemma 4.6. One can of course start with weights in a different order, permute the coordinates of $\mathbb{K}^n$ to sort them, then modify the upcoming results appropriately. However, the requirement that weights be strictly positive is not arbitrary—for one, it guarantees that the translation-invariant topology on $\mathbb{K}[[x]]$ determined by the decreasing filtration by ideals

$$\mathbb{K}[[x]] = \mathcal{F}^0 \supseteq \mathcal{F}^1 \supseteq \mathcal{F}^2 \supseteq \dots$$

coincides with the $\mathfrak{m}$ -adic topology. From this filtration, we obtain the associated grading

$$\mathbb{K}[x] = \bigoplus_{k=0}^{\infty} \mathcal{P}^k,$$

where the $\mathbb{K}$ -vector space $\mathcal{P}^k$ of *quasi-homogeneous polynomials* of weighted degree k has basis given by $\{x^\alpha \mid \langle w, \alpha \rangle = k\}$. The adjective “unweighted” shall always refer to the *trivial weighting* $w = (1, \dots, 1)$, from which we recover the $\mathfrak{m}$ -adic filtration on $\mathbb{K}[[x]]$ ($\mathcal{F}^k = \mathfrak{m}^k$) and the usual grading on $\mathbb{K}[x]$ by polynomial homogeneity.

Every formal function $f(x) = \sum_{\alpha} c_{\alpha} x^{\alpha} \in \mathbb{K}[[x]]$ can be written as a convergent infinite sum (2.2) of homogeneous polynomials using the $\mathfrak{m}$ -adic topology. As the filtration topology on $\mathbb{K}[[x]]$ induced by w coincides with the $\mathfrak{m}$ -adic topology, we can also write

$$f(x) = \sum_{k=0}^{\infty} f_k(x), \quad \text{where } f_k(x) = \sum_{\langle w, \alpha \rangle = k} c_{\alpha} x^{\alpha} \in \mathcal{P}^k \text{ for each } k. \quad (3.1)$$

We define $\Theta_w(f) := \max\{k \mid f \in \mathcal{F}^k\} = \min\{k \mid f_k \neq 0\}$, with the convention that $\Theta_w(0) := \infty$. If $\Theta_w(f) = k$ and $\Theta_w(g) = \ell$ are both finite, then $fg \in \mathcal{F}^{k+\ell}$, hence $\Theta_w(fg) \geq \Theta_w(f) + \Theta_w(g)$. On the other hand, we have $f_k g_{\ell} \neq 0$ as $\mathbb{K}[x]$ is an integral domain, hence $\Theta_w(fg) \leq \Theta_w(f) + \Theta_w(g)$. For non-trivial weightings, we do *not* extend this “weighted Big Theta” notation to n -tuples of power series, as different components may have different weights.

EXAMPLE 3.2. On $\mathbb{K}^2$ with coordinates (x, y) , we may consider x “quadratic” and y “cubic” by assigning weights $(w_1, w_2) = (2, 3)$. Then the filtration on $\mathbb{K}[[x, y]]$ starts with $\mathcal{F}^1 = \mathcal{F}^2 = \mathfrak{m}$, and

$$\mathcal{F}^3 = \text{span}_{\mathbb{K}[[x, y]]}\{x^2, y\}, \quad \mathcal{F}^4 = \text{span}_{\mathbb{K}[[x, y]]}\{x^2, xy, y^2\}, \quad \mathcal{F}^5 = \text{span}_{\mathbb{K}[[x, y]]}\{x^3, xy, y^2\}.$$

For this weighting, there are no non-zero quasi-homogeneous polynomials of weighted degree 1. ♦

So far, there has been a lot of algebra and not much calculus, so let us compare with analogous definitions in the differentiable world. Recall that a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is called positively homogeneous of degree $s \in \mathbb{R}$ if $f(tx) = t^s f(x)$ for all $x \in \mathbb{R}^n$ and $t > 0$. For a function f that is smooth at $0 \in \mathbb{R}^n$ and positively homogeneous of degree s , one has the following dichotomy:

- If $s < 0$, then f is the zero function.
- If $s \geq 0$, then $s \in \mathbb{N}$ and f is a homogeneous polynomial function of degree s .

Thus, smooth positively homogeneous functions can be studied purely algebraically. Introducing a weighting $w = (w_1, \dots, w_n)$, let us call a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ *positively quasi-homogeneous* of weighted degree $s \in \mathbb{R}$ if $\kappa_t^* f = t^s f$ for all $t > 0$, where $\kappa_t: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the linear map given by

$$\kappa_t(x^1, \dots, x^n) = (t^{w_1} x^1, \dots, t^{w_n} x^n). \quad (3.2)$$

Grabowski–Rotkiewicz [10, Lemma 2.1] have shown that the dichotomy above continues to hold when one replaces “(positively) homogeneous” with “(positively) quasi-homogeneous” and “degree” with “weighted degree”. Thus, smooth positively quasi-homogeneous functions can also be studied through an algebraic lens.

The diffeomorphisms κ_t pull back the coordinate vector fields $\frac{\partial}{\partial x^i}$ on $\mathbb{R}^n$ to $\kappa_t^* \frac{\partial}{\partial x^i} = t^{-w_i} \frac{\partial}{\partial x^i}$. Mnemonically, one thinks of $\frac{\partial}{\partial x^i}$ as having weighted degree $-w_i$ “since x^i is in the denominator”; formally, the filtration on $\mathbb{K}[[x]]$ induces a filtration on $\mathfrak{X}$ that starts in negative degrees and is compatible with the Lie algebra structure. We denote the filtration index with double brackets:

$$\mathfrak{X} = \mathfrak{X}_{[-w_n]} \supseteq \mathfrak{X}_{[-w_n+1]} \supseteq \mathfrak{X}_{[-w_n+2]} \supseteq \dots,$$

where $\mathfrak{X}_{[k]}$ is the $\mathbb{K}[[x]]$ -submodule of formal vector fields that act on $\mathbb{K}[[x]]$ by raising filtration degree by k . Equivalently, $\mathfrak{X}_{[k]}$ is generated by monomial vector fields $x^\alpha \frac{\partial}{\partial x^i}$ such that $\langle w, \alpha \rangle - w_i \geq k$. Similarly, the associated grading on $\mathbb{K}[x]$ makes $\mathfrak{X}_{\text{pol}}$ into a graded Lie algebra

$$\mathfrak{X}_{\text{pol}} = \bigoplus_{k=-w_n}^{\infty} \mathfrak{X}_{[k]}, \quad (3.3)$$

where the $\mathbb{K}$ -linear subspace $\mathfrak{X}_{[k]}$ of *quasi-homogeneous vector fields* of weighted degree k acts on $\mathbb{K}[x]$ by raising grading degree by k . Equivalently, $\mathfrak{X}_{[k]}$ has a basis given by monomial vector fields $x^\alpha \frac{\partial}{\partial x^i}$ such that $\langle w, \alpha \rangle - w_i = k$; thus

$$\mathfrak{X}_{[k]} \cong \bigoplus_{i=1}^n \mathcal{P}^{w_i+k} \text{ as } \mathbb{K}\text{-vector spaces.} \quad (3.4)$$

Remark 3.3. By interpreting the right-hand side of (3.2) as an n -tuple of formal power series, one obtains a $\mathbb{K}$ -algebra endomorphism $\kappa_t: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$ such that $\kappa_t(t) = t$ and $\kappa_t(x^i) = t^{w_i} x^i$. This is *not* a formal isotopy, as its inverse would have to map $x^i \mapsto t^{-w_i} x^i$. But if one extends κ_t to act on Laurent series in t and suitably defines pullbacks in this setting, then $\mathfrak{X}_{[k]}$ is the set of polynomial vector fields X satisfying $\kappa_t^* X = t^k X$. (When $\mathbb{K} = \mathbb{R}$, this is equivalent to requiring $\kappa_t^* X = t^k X$ in the smooth sense for all $t > 0$ by Grabowski–Rotkiewicz.) $\blacklozenge$

3.2. Weights and evaluation. We now state and prove weighted generalizations of Propositions 2.19 and 2.21.

PROPOSITION 3.4. *Let φ_t be a formal isotopy of $\mathbb{K}^n$ with corresponding n -tuple written in the form $\varphi(t, x) = \sum_{k=0}^{\infty} t^k \varphi_k(x)$ where $\varphi_k(x) = (\varphi_k^1(x), \dots, \varphi_k^n(x)) \in \mathbb{K}[[x]]^n$. If there exists a weighting $w = (w_1, \dots, w_n)$ of $\mathbb{K}^n$ such that, for all $i \in \{1, \dots, n\}$, we have*

$$\Theta_w(\varphi_0^i) \geq w_i, \quad \Theta_w(\varphi_k^i) \geq w_i + 1 \text{ for all } k \geq 1, \quad \lim_{k \rightarrow \infty} \Theta(\varphi_k) = \infty,$$

then φ_t is evaluative.

Taking the trivial weighting $w = (1, \dots, 1)$ recovers Proposition 2.19 as the condition on φ_0 reduces to $\Theta(\varphi_0) \geq 1$, which formal isotopies always satisfy (Remark 2.4). Note that the limit condition is stated with *unweighted* Big Theta notation for *tuples*, but it is equivalent to requiring that $\lim_{k \rightarrow \infty} \Theta_w(\varphi_k^i) = \infty$ for each i since the filtration topology defined by w coincides with the $\mathfrak{m}$ -adic topology; both limits are equivalent to saying that $\lim_{k \rightarrow \infty} \varphi_k^i = 0$ for each i in this topology.

Proof. By definition, $\Theta(\varphi_k)$ is the smallest $\ell \in \mathbb{N}$ for which there exists $i \in \{1, \dots, n\}$ and $\alpha \in \mathbb{N}^n$ such that $|\alpha| = \ell$ and the coefficient of x^α in $\varphi_k^i(x)$ is non-zero (unless $\varphi_k(x) = (0, \dots, 0)$). For any particular α , we have $\Theta(\varphi_k) > |\alpha|$ for sufficiently large k since $\Theta(\varphi_k) \rightarrow \infty$. Thus the coefficient of x^α in each $\varphi_k^i(x)$ is non-zero for finitely many k 's, so the coefficient of x^α in $\varphi^i(t, x) = \sum_k t^k \varphi_k^i(x)$ is actually a polynomial in t . As α was arbitrary, each φ^i is an evaluative power series. Since $\Theta_w(\varphi_k^i) \geq w_i \geq 1$ for all $k \geq 0$, we have $\varphi_k^i(x) \in \mathcal{F}^{w_i} \subseteq \mathcal{F}^1 = \mathfrak{m}$, hence $\varphi^i(t, 0) = \sum_k t^k \varphi_k^i(0) = 0$. Thus condition (i) of Definition 2.12 is satisfied, and we are left with checking condition (ii).

We claim that $D\varphi_k(0)$ is a block upper triangular matrix for all $k \geq 0$, and *strictly* block upper triangular for all $k \geq 1$; that is,

$$D\varphi_k(0) = \begin{pmatrix} A_{k,1} & * & \cdots & * \\ O & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ O & \cdots & O & A_{k,m} \end{pmatrix}, \quad \text{where } A_{k,\ell} = O \text{ for } k \geq 1. \quad (3.5)$$

Here, m denotes the number of *distinct* weights and each $A_{k,\ell}$ is a $d_\ell \times d_\ell$ matrix, where d_ℓ is the number of coordinates assigned to the ℓ^{th} smallest weight value.⁵

To prove the claim, recall that $\frac{\partial}{\partial x^j}$ decreases filtration degree on $\mathbb{K}[[x]]$ by w_j , so

$$\Theta_w \left(\frac{\partial \varphi_k^i}{\partial x^j} \right) \geq \Theta_w(\varphi_k^i) - w_j.$$

Since we require weights to be listed from smallest to largest in Definition 3.1 and the partitioning of $D\varphi_k(0)$ illustrated in (3.5) is along blocks corresponding to distinct weights, it follows that entries below the main block diagonal are those indexed by (i, j) where $w_i > w_j$. Since $\Theta_w(\varphi_k^i) \geq w_i$ for all $k \geq 0$, we thus have $\Theta_w \left(\frac{\partial \varphi_k^i}{\partial x^j} \right) \geq w_i - w_j > 0$, hence $\frac{\partial \varphi_k^i}{\partial x^j}(0) = 0$ below the main block diagonal. Entries within the main block diagonal are indexed by (i, j) with $w_i = w_j$, so if $k \geq 1$, then $\Theta_w \left(\frac{\partial \varphi_k^i}{\partial x^j} \right) \geq w_i - w_j + 1 = 1$, hence also $\frac{\partial \varphi_k^i}{\partial x^j}(0) = 0$. This proves the claim.

Now observe that $D_x \varphi(t, x) = \sum_{k=0}^{\infty} t^k D\varphi_k(x)$, so upon setting $x^1 = \dots = x^n = 0$, we obtain

$$D_x \varphi(t, 0) = \sum_{k=0}^{\infty} t^k D\varphi_k(0) = D\varphi_0(0) + \sum_{k=1}^{\infty} t^k \cdot (\text{strictly block upper triangular matrices}).$$

Consequently $\det D_x \varphi(t, 0) = \det D\varphi_0(0)$, which is a non-zero element of $\mathbb{K}$ since φ_t is a formal isotopy (Remark 2.4). Thus condition (ii) of Definition 2.12 is satisfied. $\blacksquare$

PROPOSITION 3.5. *Suppose $X_t \in \mathfrak{X}_t$ is written in the form $X_t = \sum_{k=0}^{\infty} t^k X_k$ where $X_k \in \mathfrak{X}$. Identify the flow φ_t of X_t with an n -tuple $\varphi(t, x) = \sum_{k=0}^{\infty} t^k \varphi_k(x)$ where $\varphi_k(x) \in \mathbb{K}[[x]]^n$. If there exists a weighting $w = (w_1, \dots, w_n)$ of $\mathbb{K}^n$ such that $X_k \in \mathfrak{X}_{[k+1]}$ for all $k \geq 0$, then with respect to this weighting, one has $\Theta_w(\varphi_k^i) \geq w_i + k$ for all $i \in \{1, \dots, n\}$ and $k \geq 0$. In particular, φ_t is an evaluative formal isotopy by Proposition 3.4.*

Proof. We proceed by (strong) induction on k . By Theorem 2.6(IC) we have $\varphi_0^i(x) = x^i$, hence $\Theta_w(\varphi_0^i) = w_i$ for each $i \in \{1, \dots, n\}$, establishing the base case. For the inductive step, assume

⁵For example, if $w = (1, 4, 5, 5)$, then there are $m = 3$ distinct weights and $(d_1, d_2, d_3) = (1, 1, 2)$.

that $\Theta_w(\varphi_k^i) \geq w_i + k$ for all $i \in \{1, \dots, n\}$ and $k \leq m$. The coefficient of t^m on the left-hand side of Theorem 2.6(FP) applied to each x^i is $(m+1)\varphi_{m+1}^i(x)$; we want to find Θ_w of the corresponding coefficient of t^m on the right-hand side. Warning: an unglamorous calculation follows. The reader may wish to skip to the final expression (†) for this coefficient.

As $\Theta_w(x^i) = w_i$ and X_k raises filtration degree by $k+1$, we have $\Theta_w(X_k(x^i)) \geq w_i + k + 1$. Thus

$$X_k(x^i) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ \langle w, \alpha \rangle \geq w_i + k + 1}} c_{k, \alpha}^i x^\alpha, \quad \text{for some coefficients } c_{k, \alpha}^i \in \mathbb{K}.$$

As φ_t is a $\mathbb{K}$ -algebra automorphism fixing t , we thus have

$$\varphi_t(X_t(x^i)) = \sum_{k=0}^{\infty} t^k \varphi_t(X_k(x^i)) = \sum_{k=0}^{\infty} \sum_{\substack{\alpha \in \mathbb{N}^n \\ \langle w, \alpha \rangle \geq w_i + k + 1}} c_{k, \alpha}^i t^k \prod_{j=1}^n \varphi_t(x^j)^{\alpha_j}.$$

For each $j \in \{1, \dots, n\}$, we can express $\varphi_t(x^j)^{\alpha_j}$ as a sum over multi-indices $\beta^j = (\beta_1^j, \dots, \beta_{\alpha_j}^j)$ of length α_j :

$$\varphi_t(x^j)^{\alpha_j} = \left(\sum_{\ell=0}^{\infty} t^\ell \varphi_\ell^j(x) \right)^{\alpha_j} = \sum_{\beta^j \in \mathbb{N}^{\alpha_j}} t^{|\beta^j|} \prod_{\ell=1}^{\alpha_j} \varphi_{\beta_\ell^j}^j(x).$$

The n multi-indices $\beta^j \in \mathbb{N}^{\alpha_j}$ can be concatenated into a single multi-index $\beta := (\beta^1, \dots, \beta^n)$ of total length $|\alpha|$. Conversely, any $\beta \in \mathbb{N}^{|\alpha|} \cong \mathbb{N}^{\alpha_1} \times \dots \times \mathbb{N}^{\alpha_n}$ splits into n multi-indices $\beta^j \in \mathbb{N}^{\alpha_j}$, where β^j is the part of β indexed by $L_\alpha^j := \{\ell \in \mathbb{N} \mid \alpha_1 + \dots + \alpha_{j-1} < \ell \leq \alpha_1 + \dots + \alpha_j\}$. Thus

$$\begin{aligned} \prod_{j=1}^n \varphi_t(x^j)^{\alpha_j} &= \sum_{\beta^1 \in \mathbb{N}^{\alpha_1}} \dots \sum_{\beta^n \in \mathbb{N}^{\alpha_n}} t^{|\beta^1| + \dots + |\beta^n|} \prod_{j=1}^n \prod_{\ell=1}^{\alpha_j} \varphi_{\beta_\ell^j}^j(x) \\ &= \sum_{\beta \in \mathbb{N}^{|\alpha|}} t^{|\beta|} \prod_{j=1}^n \prod_{\ell \in L_\alpha^j} \varphi_{\beta_\ell}^j(x). \end{aligned}$$

It follows that the coefficient of t^m on the right-hand side of Theorem 2.6(FP) applied to each x^i is

$$\sum_{k=0}^{\infty} \sum_{\substack{\alpha \in \mathbb{N}^n \\ \langle w, \alpha \rangle \geq w_i + k + 1}} \sum_{\substack{\beta \in \mathbb{N}^{|\alpha|} \\ |\beta| + k = m}} c_{k, \alpha}^i \prod_{j=1}^n \prod_{\ell \in L_\alpha^j} \varphi_{\beta_\ell}^j(x). \quad (\dagger)$$

The salient point is that for each fixed k , α , and β appearing in (†), we have $\beta_\ell \leq |\beta| \leq m$ for each $\ell \in \{1, \dots, |\alpha|\}$, hence $\Theta_w(\varphi_{\beta_\ell}^j) \geq w_j + \beta_\ell$ by the induction hypothesis. Thus

$$\Theta_w \left(\prod_{j=1}^n \prod_{\ell \in L_\alpha^j} \varphi_{\beta_\ell}^j \right) = \sum_{j=1}^n \sum_{\ell \in L_\alpha^j} \Theta_w(\varphi_{\beta_\ell}^j) \geq \sum_{j=1}^n \sum_{\ell \in L_\alpha^j} (w_j + \beta_\ell).$$

Observing that

$$\sum_{j=1}^n \sum_{\ell \in L_\alpha^j} w_j = \sum_{j=1}^n w_j \alpha_j = \langle w, \alpha \rangle \geq w_i + k + 1, \quad \sum_{j=1}^n \sum_{\ell \in L_\alpha^j} \beta_\ell = |\beta|,$$

we see that $\sum_{j=1}^n \sum_{\ell \in L_\alpha^j} (w_j + \beta_\ell) \geq w_i + k + 1 + |\beta| = w_i + m + 1$. Thus $\Theta_w(\varphi_{m+1}^i) \geq w_i + m + 1$, completing the induction. $\blacksquare$

EXAMPLE 3.6. As noted in Remark 2.20, the flow of $y \frac{\partial}{\partial x}$ is the formal isotopy of $\mathbb{K}^2$ corresponding to the 2-tuple $(x + ty, y)$. We could not conclude that this is evaluative from either Propositions 2.19 or 2.21, but we can from Proposition 3.5 using any weighting (w_1, w_2) such that $w_1 < w_2$. (Note

that with respect to such a weighting, $y \frac{\partial}{\partial x}$ is exactly the “weighted non-linear perturbation” in the vector field (1.1) when $k = 1$.) The author does not know the answer to the following question: if a formal isotopy of $\mathbb{K}^n$ is evaluative, does there exist *some* weighting of $\mathbb{K}^n$ with respect to which the hypotheses of Proposition 3.4 hold, perhaps after permuting coordinates? ♦

Remark 3.7. Based on Example 2.13, we conjecture that replacing the hypothesis of Proposition 3.5 that $X_k \in \mathfrak{X}_{[k+1]}$ for all $k \geq 0$ with the weaker pair of assumptions that $X_k \in \mathfrak{X}_{[1]}$ and $\lim_{k \rightarrow \infty} \Theta(X_k) = \infty$ (the Big Theta notation understood by identifying X_k with its n -tuple of components) is enough to conclude that the flow of X_t is evaluative via Proposition 3.4. Indeed, the only modification to (†) with this weaker hypothesis is that the inequality $\langle w, \alpha \rangle \geq w_i + k + 1$ in the sum over $\alpha \in \mathbb{N}^n$ is replaced with $\langle w, \alpha \rangle \geq w_i + 1$; this suffices to prove by induction that $\Theta_w(\varphi_k^i) \geq w_i + 1$ for all $k \geq 1$. However, proving that $\lim_{k \rightarrow \infty} \Theta(\varphi_k) = \infty$ seems more complicated. It turns out that the vector fields we will flow along in Section 4 satisfy the hypothesis of Proposition 3.5 as written, so we content ourselves with the current statement. ♦

4. WEIGHTED LINEARIZATION

4.1. Weighted linear approximations. Given a weighting $w = (w_1, \dots, w_n)$, every $f \in \mathbb{K}[[x]]$ can be written as an infinite sum of quasi-homogeneous polynomials (3.1). Similarly, every $X \in \mathfrak{X}$ can be written as an infinite sum of quasi-homogeneous vector fields:

$$X = \sum_{k=-w_n}^{\infty} X_{[k]}, \quad \text{where } X_{[k]} \in \mathfrak{X}_{[k]}.$$

DEFINITION 4.1. We call $X_{[0]}$ the **weighted linear approximation** of X with respect to w . A formal vector field $X \in \mathfrak{X}$ is called **admissible** with respect to w (or w -admissible) if $X \in \mathfrak{X}_{[0]}$, or equivalently, if $X_{[k]} = 0$ for all $k < 0$. We use the same terminology for a smooth vector field on an open neighbourhood of $0 \in \mathbb{R}^n$, referring to its formal Taylor series at 0.

Remark 4.2. Extending the $\mathbb{K}$ -algebra endomorphism $\kappa_t: \mathbb{K}[[t, x]] \rightarrow \mathbb{K}[[t, x]]$ from Remark 3.3 to act on Laurent series in t , the pullback of any $X \in \mathfrak{X}$ along κ_t is given by

$$\kappa_t^* X = \sum_{k=-w_n}^{\infty} t^k X_{[k]}.$$

This is generally not a time-dependent formal vector field in the sense of Section 2.2 as the right-hand side can include negative powers of t , but w -admissible formal vector fields are precisely those for which $\kappa_t^* X$ *does* define a time-dependent formal vector field—in fact, an *evaluative* one, with $(\kappa_t^* X)|_{t=0} = X_{[0]}$ and $(\kappa_t^* X)|_{t=1} = X$. For a smooth vector field X on $\mathbb{R}^n$, w -admissibility means that $\lim_{t \rightarrow 0^+} (\kappa_t^* X)$ exists and equals $X_{[0]}$, where $\kappa_t: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the diffeomorphism (3.2). ♦

Remark 4.3. If $X = \sum_i X^i \frac{\partial}{\partial x^i}$ (formal or smooth) is admissible with respect to *any* weighting, then its components X^i satisfy $X^i(0) = 0$ for all i —a consequence of requiring weights to be strictly positive in Definition 3.1. (Conversely, if $X^i(0) = 0$ for all i , then X is admissible with respect to the trivial weighting.) In Loizides–Meinrenken’s differential-geometric framework of weightings [13], one can declare some coordinates to have weight 0; these correspond to directions along a specified submanifold (for us, a single point) of a smooth manifold. One obtains quasi-homogeneous approximations of not only functions and vector fields at a point, but also tensor fields, differential forms, and multivector fields along submanifolds. We envisage generalizations of the present work to quasi-homogeneous normal forms for such objects along submanifolds. ♦

We now turn to the question of formal weighted linearizability for vector fields:

Given a w -admissible vector field $X \in \mathfrak{X}_{[0]}$ with weighted linear approximation $X_{[0]} \in \mathfrak{X}_{[0]}$, when does there exist a formal diffeomorphism $\varphi: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ such that $\varphi^ X = X_{[0]}$?*

Before giving a general answer, we revisit a particularly important special case from Section 1.1: recall that the *weighted Euler vector field* associated to $w = (w_1, \dots, w_n)$ is given by

$$\mathcal{E}_w := \sum_{i=1}^n w_i x^i \frac{\partial}{\partial x^i} \in \mathfrak{X}_{[0]}. \quad (4.1)$$

The weighted Euler vector field captures the notions of quasi-homogeneity associated to w , in that a polynomial $f \in \mathbb{K}[x]$ lies in $\mathcal{P}^k$ if and only if $\mathcal{E}_w(f) = kf$, and a polynomial vector field $X \in \mathfrak{X}_{\text{pol}}$ lies in $\mathfrak{X}_{[k]}$ if and only if $[\mathcal{E}_w, X] = kX$. A (formal or smooth) vector field is called *weighted Euler-like* if it is w -admissible and its weighted linear approximation is equal to $\mathcal{E}_w$. We obtain a formal version of [14, Lemma 2.5] over any field of characteristic zero:

PROPOSITION 4.4. *If $X \in \mathfrak{X}$ is weighted Euler-like with respect to a weighting w of $\mathbb{K}^n$, then there exists a formal diffeomorphism φ of $\mathbb{K}^n$ such that $\varphi^*X = \mathcal{E}_w$.*

Proof. By assumption, we can write $X = \sum_{k=0}^{\infty} X_{[k]}$ where $X_{[k]} \in \mathfrak{X}_{[k]}$ and $X_{[0]} = \mathcal{E}_w$. The time-dependent formal vector fields $X_t := \sum_{k=0}^{\infty} t^k X_{[k]}$ and $U_t := \sum_{k=0}^{\infty} t^k X_{[k+1]}$ thus satisfy

$$\begin{aligned} [\mathcal{E}_w, X_t] &= \sum_{k=0}^{\infty} t^k [\mathcal{E}_w, X_{[k]}] = \sum_{k=0}^{\infty} k t^k X_{[k]} = t \frac{dX_t}{dt}, \\ X_t - \mathcal{E}_w &= \left(\mathcal{E}_w + \sum_{k=1}^{\infty} t^k X_{[k]} \right) - \mathcal{E}_w = t U_t. \end{aligned}$$

It follows that $t \left(\frac{dX_t}{dt} + [U_t, X_t] \right) = 0$, hence $\frac{dX_t}{dt} + [U_t, X_t] = 0$ as $\mathfrak{X}_t$ is torsion-free. As observed in Remark 4.2, X_t is evaluative, with $X_t|_{t=0} = \mathcal{E}_w$ and $X_t|_{t=1} = X$; on the other hand, the flow φ_t of U_t is evaluative by Proposition 3.5. Thus the formal diffeomorphism $\varphi := \varphi_1$ satisfies $\varphi^*X = \mathcal{E}_w$ by the $(r, s) = (1, 0)$ case of the formal Moser trick (Corollary 2.17). ■

To motivate the general answer, suppose that X is a w -admissible smooth vector field on $\mathbb{R}^n$ with weighted linear approximation $X_{[0]}$ and φ is a diffeomorphism of $\mathbb{R}^n$ satisfying $\varphi^*X = X_{[0]}$ as smooth vector fields. Seeking a Moser-type construction for φ , we want to think of $X_{[0]}$ and X as time slices of a one-parameter family of vector fields X_t at $t = 0$ and $t = 1$, respectively, and φ as the $t = 1$ time slice of a one-parameter family of diffeomorphisms φ_t . Remark 4.2 suggests the candidate $X_t := \kappa_t^*X$ where $\kappa_t: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the diffeomorphism (3.2): as $\kappa_1 = \text{id}_{\mathbb{R}^n}$, we have $X_t|_{t=1} = X$, and w -admissibility means that X_t extends smoothly to $t = 0$ by $X_t|_{t=0} = X_{[0]}$. If we define $\varphi_t := \kappa_t^{-1} \circ \varphi \circ \kappa_t$ for $t > 0$, then $\varphi_1 = \varphi$ and⁶

$$\varphi_t^*X_t = (\kappa_t \circ \varphi_t)^*X = (\varphi \circ \kappa_t)^*X = \kappa_t^*(\varphi^*X) = \kappa_t^*X_{[0]} = X_{[0]}, \quad \forall t > 0,$$

the last equality using Remark 3.3 and the fact that $X_{[0]} \in \mathfrak{X}_{[0]}$. Thus $\varphi_t^*X_t$ is constant with respect to t . Differentiating with respect to t yields

$$\frac{dX_t}{dt} + [U_t, X_t] = 0, \quad (4.2)$$

where U_t is the time-dependent vector field that generates φ_t as its flow.

Now observe that $\kappa_{st} = \kappa_s \circ \kappa_t$ for all $s, t > 0$; consequently, the one-parameter family of diffeomorphisms $\{\varphi_t\}_{t>0}$ enjoys the homogeneity property $\varphi_{st} = \kappa_t^{-1} \circ \varphi_s \circ \kappa_t$ for all $s, t > 0$. For fixed t , differentiating both sides of this homogeneity property with respect to s yields

$$tU_{st} = \kappa_t^*U_s.$$

⁶Upper stars reverse order here because we work with actual diffeomorphisms in this discussion, c.f. Remark 2.2.

Assuming U_s is analytic with respect to the time parameter, so that $U_s = \sum_{k=0}^{\infty} s^k U_{[k+1]}$ for some time-independent vector fields $U_{[k+1]}$ (the shifted index to be explained imminently), we obtain

$$\sum_{k=0}^{\infty} s^k t^{k+1} U_{[k+1]} = \sum_{k=0}^{\infty} s^k \kappa_t^* U_{[k+1]}.$$

Comparing the coefficient of s^k on both sides yields $\kappa_t^* U_{[k+1]} = t^{k+1} U_{[k+1]}$; that is, $U_{[k+1]} \in \mathfrak{X}_{[k+1]}$. If we also assume that $X_t = \sum_{k=0}^{\infty} t^k X_{[k]}$ is analytic, then Equation (4.2) becomes

$$\sum_{k=0}^{\infty} t^k \left((k+1) X_{[k+1]} + \sum_{\substack{i,j \geq 0 \\ i+j=k}} [U_{[i+1]}, X_{[j]}] \right) = 0.$$

Thus each coefficient of t^k vanishes, and we obtain a weighted version of the *homological equation*

$$[X_{[0]}, U_{[k+1]}] = (k+1) X_{[k+1]} - \sum_{i=0}^{k-1} [X_{[k-i]}, U_{[i+1]}]. \quad (4.3)$$

Both sides are quasi-homogeneous vector fields of weighted degree $k+1$. In particular, if the adjoint operator $\text{ad}_{X_{[0]}} := [X_{[0]}, \cdot]: \mathfrak{X}_{[k+1]} \rightarrow \mathfrak{X}_{[k+1]}$ is invertible, then Equation (4.3) provides a recursive formula that uniquely determines $U_{[k+1]}$ in terms of $U_{[1]}, \dots, U_{[k]}$ and $X_{[1]}, \dots, X_{[k+1]}$.

These heuristics assumed analyticity with respect to time, but the formal world does not concern itself with such matters of convergence. By turning the above argument on its head, we thus obtain:

THEOREM 4.5. *Suppose $X \in \mathfrak{X}_{[0]}$ is w -admissible, with weighted linear approximation $X_{[0]} \in \mathfrak{X}_{[0]}$. If $\text{ad}_{X_{[0]}}: \mathfrak{X}_{[k]} \rightarrow \mathfrak{X}_{[k]}$ is invertible for all $k \geq 1$, then there exists a formal diffeomorphism φ of $\mathbb{K}^n$ such that $\varphi^* X = X_{[0]}$.*

Proof. As X is w -admissible, we have $X = \sum_{k=0}^{\infty} X_{[k]}$ where $X_{[k]} \in \mathfrak{X}_{[k]}$. Thus $X_t := \sum_{k=0}^{\infty} t^k X_{[k]}$ is an evaluative formal vector field satisfying $X_t|_{t=0} = X_{[0]}$ and $X_t|_{t=1} = X$. Define a time-dependent formal vector field $U_t := \sum_{k=0}^{\infty} t^k U_{[k+1]}$ by declaring $U_{[1]} \in \mathfrak{X}_{[1]}$ to be the unique solution to $[X_{[0]}, U_{[1]}] = X_{[1]}$, then recursively defining $U_{[k+1]} \in \mathfrak{X}_{[k+1]}$ by inverting Equation (4.3).

By construction, X_t and U_t satisfy Equation (4.2) in the formal sense. By Proposition 3.5, the flow φ_t of U_t is evaluative. The $(r, s) = (1, 0)$ case of the formal Moser trick (Corollary 2.17) yields the desired formal diffeomorphism $\varphi := \varphi_1$. $\blacksquare$

As a special case, we note that if X is weighted Euler-like, then $\text{ad}_{X_{[0]}} = \text{ad}_{\mathcal{E}_w}$ acts on $\mathfrak{X}_{[k]}$ via multiplication by k . The unique solution to the weighted homological equation (4.3) is then given by $U_{[k+1]} = X_{[k+1]}$, recovering the proof of Proposition 4.4. In general, Theorem 4.5 is advantageous as each $\mathfrak{X}_{[k]}$ is a finite-dimensional $\mathbb{K}$ -vector space by (3.4), so we can determine invertibility of the adjoint operator on $\mathfrak{X}_{[k]}$ by examining its eigenvalues; such is the topic of the next section.

4.2. Formal weighted linearization. For this section, we fix a weighting $w = (w_1, \dots, w_n)$ of $\mathbb{K}^n$ once and for all. Keeping with the notation used in the proof of Proposition 3.4, we denote the number of *distinct* weights in w by m and the number of coordinates assigned to the ℓ^{th} smallest weight value by d_ℓ . We define $\mu: \{1, \dots, n\} \rightarrow \{1, \dots, m\}$ by declaring that $\mu(i) = \ell$ where w_i is the ℓ^{th} smallest weight value, so that d_ℓ is the cardinality of $\mu^{-1}(\ell)$. We identify the linear polynomials in $\mathbb{K}[x]$ with the dual space $(\mathbb{K}^n)^\vee$ and define a grading

$$(\mathbb{K}^n)^\vee = \bigoplus_{\ell=1}^m V^\ell, \quad (4.4)$$

where $V^\ell = \text{span}_{\mathbb{K}}\{x^i \mid \mu(i) = \ell\}$. Thus, $V^{\mu(i)} = \mathcal{P}^{w_i} \cap (\mathbb{K}^n)^\vee$.

Given a (formal or smooth) vector field X on $\mathbb{K}^n$, we denote by $X_{\text{lin}} \in \mathfrak{X}_{\text{pol}}$ the usual linear approximation of X , i.e., its weighted linear approximation with respect to the trivial weighting. As a derivation of $\mathbb{K}[x]$, X_{lin} preserves (unweighted) polynomial degree, hence leaves $(\mathbb{K}^n)^\vee$ invariant. As explained in Section 1.1, conditions guaranteeing X to be linearizable in a given category are often phrased in terms of the eigenvalues of the Jacobian matrix $DX(0)$ (assuming 0 is a fixed point), where X is identified with its n -tuple of components. Equivalently, these are the eigenvalues of $X_{\text{lin}}: (\mathbb{K}^n)^\vee \rightarrow (\mathbb{K}^n)^\vee$, as the transpose of $DX(0)$ represents X_{lin} in the ordered basis $\{x^1, \dots, x^n\}$.

Assuming X is w -admissible with weighted linear approximation $X_{[0]} \in \mathfrak{X}_{[0]}$, we would like to phrase analogous conditions for weighted linearizability in terms of “the eigenvalues of $X_{[0]}$ on $(\mathbb{K}^n)^\vee$ ”. A priori, it is not immediately clear how to interpret this, as $X_{[0]}$ generally does not preserve $(\mathbb{K}^n)^\vee$: $X_{[0]}$ preserves *weighted* polynomial degree, so the image of each x^i under $X_{[0]}$ may include products of *unweighted* linear polynomials with total weighted degree equal to w_i . One way to return to $(\mathbb{K}^n)^\vee$ is to simply drop the (unweighted) non-linear terms—in other words, to act on $(\mathbb{K}^n)^\vee$ via $(X_{[0]})_{\text{lin}}$. Alternatively, one could first act on $(\mathbb{K}^n)^\vee$ via the usual linear approximation X_{lin} , then drop any terms that are not weighted-linear—in other words, act on $(\mathbb{K}^n)^\vee$ via $(X_{\text{lin}})_{[0]}$. It turns out that the order does not matter:

LEMMA 4.6. *Suppose $X \in \mathfrak{X}$ has weighted linear approximation $X_{[0]} \in \mathfrak{X}_{[0]}$ and unweighted linear approximation $X_{\text{lin}} \in \mathfrak{X}_{\text{pol}}$.*

(i) *Weighted linear approximation commutes with unweighted linear approximation:*

$$(X_{\text{lin}})_{[0]} = (X_{[0]})_{\text{lin}}.$$

(ii) *If X is w -admissible, then $(X_{[0]})_{\text{lin}}$ and X_{lin} share the same characteristic polynomial as operators on $(\mathbb{K}^n)^\vee$, and $(X_{[0]})_{\text{lin}}$ preserves the grading (4.4).*

Proof. (i) By definition, X_{lin} is obtained from X by keeping only the monomial vector fields $x^\alpha \frac{\partial}{\partial x^i}$ with $|\alpha| = 1$, whereas $X_{[0]}$ is obtained from X by keeping only the monomial vector fields with $\langle w, \alpha \rangle = w_i$. Thus $(X_{\text{lin}})_{[0]}$ and $(X_{[0]})_{\text{lin}}$ are both obtained from X by keeping only the terms satisfying both equalities.

(ii) Assuming X is w -admissible, if an (unweighted) linear term $x^j \frac{\partial}{\partial x^i}$ appears in X , then it must be the case that $w_j \geq w_i$. Since we require weights to be listed from smallest to largest in Definition 3.1, it follows that the Jacobian matrix $DX(0)$ is block upper triangular:

$$DX(0) = \begin{pmatrix} A_1 & * & \cdots & * \\ O & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ O & \cdots & O & A_m \end{pmatrix},$$

where each A_ℓ is a $d_\ell \times d_\ell$ matrix. In the weighted linear approximation $X_{[0]}$, one deletes the terms $x^j \frac{\partial}{\partial x^i}$ in X such that $w_j > w_i$, so the Jacobian matrix $DX_{[0]}(0)$ is obtained from $DX(0)$ by replacing the blocks above the diagonal with zeroes. Thus $DX(0)$ and $DX_{[0]}(0)$ have the same characteristic polynomial. The first claim follows as the transposes of these matrices respectively represent X_{lin} and $(X_{[0]})_{\text{lin}}$ in the ordered basis $\{x^1, \dots, x^n\}$ of $(\mathbb{K}^n)^\vee$; the second claim follows as we have just shown that $DX_{[0]}(0)$ is block diagonal, and the transposes of these diagonal blocks correspond precisely to how $(X_{[0]})_{\text{lin}}$ acts on the subspaces V^ℓ . ■

For simplicity, we henceforth assume that this characteristic polynomial splits in $\mathbb{K}$. If not, then pass to its splitting field; the modifications needed are straightforward. As $(X_{[0]})_{\text{lin}}$ respects the grading (4.4), we also require its eigenvalues to be listed in a way that respects this grading for our weighted non-resonance condition:

DEFINITION 4.7. Let $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ be a vector. A **resonance of weighted degree** $k \geq 1$ is a pair (i, α) where $i \in \{1, \dots, n\}$ and $\alpha \in \mathbb{N}^n$ are such that

$$\langle \lambda, \alpha \rangle = \lambda_i, \quad \langle w, \alpha \rangle - w_i = k.$$

We call λ **non-resonant of weighted degree** k if no such pair exists, and we call λ **non-resonant with respect to** w if this holds for all $k \geq 1$.

Let $X \in \mathfrak{X}_{[0]}$ be a w -admissible vector field with weighted linear approximation $X_{[0]} \in \mathfrak{X}_{[0]}$. An ordering of the eigenvalues $(\lambda_1, \dots, \lambda_n)$ of $(X_{[0]})_{\text{lin}}: (\mathbb{K}^n)^\vee \rightarrow (\mathbb{K}^n)^\vee$ is called **w -compatible** if each λ_i is an eigenvalue of the restriction $(X_{[0]})_{\text{lin}}: V^{\mu(i)} \rightarrow V^{\mu(i)}$. We call X **non-resonant of weighted degree** k if some (equivalently, every) w -compatible ordering of the eigenvalues of $(X_{[0]})_{\text{lin}}$ is non-resonant of weighted degree k in the previous sense. We call X **non-resonant with respect to** w if this holds for all $k \geq 1$.

We emphasize that the eigenvalues of $(X_{[0]})_{\text{lin}}$ are the same as those of X_{lin} by Lemma 4.6(ii). Taking the trivial weighting recovers the Poincaré–Sternberg definition of (non-)resonance: in this case the grading (4.4) is trivial, so every ordering of eigenvalues is compatible. But it is possible for X to be simultaneously resonant with respect to the trivial weighting and non-resonant with respect to a non-trivial weighting—for example, a vector field that is weighted Euler-like with respect to a weighting where some weight is a proper divisor of another, or the upcoming Example 4.9.

We now state the main result of this section: the formal half of Theorem 1.1.

THEOREM 4.8. Suppose $X \in \mathfrak{X}_{[0]}$ is w -admissible, with weighted linear approximation $X_{[0]} \in \mathfrak{X}_{[0]}$. If $k \geq 1$ is such that X is non-resonant of weighted degree k , then $\text{ad}_{X_{[0]}}: \mathfrak{X}_{[k]} \rightarrow \mathfrak{X}_{[k]}$ is invertible. In particular, if X is non-resonant with respect to w , then there exists a formal diffeomorphism φ of $\mathbb{K}^n$ such that $\varphi^*X = X_{[0]}$ by Theorem 4.5.

Proof. By Lemma 4.6(ii), each V^ℓ is invariant under $(X_{[0]})_{\text{lin}}$. Concatenating Jordan bases for its restrictions to each V^ℓ produces an ordered Jordan basis $\{p^1, \dots, p^n\}$ of $(\mathbb{K}^n)^\vee$ for $(X_{[0]})_{\text{lin}}$ such that $p^i \in V^{\mu(i)} = \mathcal{P}^{w_i} \cap (\mathbb{K}^n)^\vee$ for each i . Consequently, the linear change of basis $x^i \mapsto p^i$ extends to a formal diffeomorphism $\psi: \mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$ that leaves $\mathbb{K}[x]$ invariant and preserves both weighted and unweighted polynomial degree.⁷ The same facts are true of $\varphi := \psi^{-1}$, so $\varphi^*: \mathfrak{X} \rightarrow \mathfrak{X}$ leaves $\mathfrak{X}_{\text{pol}}$ invariant and preserves both the grading (3.3) and the usual unweighted grading on $\mathfrak{X}_{\text{pol}}$. Thus $\varphi^*X_{[0]} \in \mathfrak{X}_{[0]}$, and it suffices to prove that $\text{ad}_{\varphi^*X_{[0]}} = \varphi^* \circ \text{ad}_{X_{[0]}} \circ (\varphi^*)^{-1}$ is invertible on $\mathfrak{X}_{[k]}$.

All monomial vector fields $x^\beta \frac{\partial}{\partial x^j}$ appearing in $X_{[0]}$ satisfy $\langle w, \beta \rangle = w_j$, and writing $X_{[0]} = (X_{[0]})_{\text{lin}} + [X_{[0]} - (X_{[0]})_{\text{lin}}]$ splits $X_{[0]}$ into those with $|\beta| = 1$ and those with $|\beta| > 1$. As $\{p^1, \dots, p^n\}$ puts $(X_{[0]})_{\text{lin}}$ into Jordan canonical form, we have

$$\varphi^*(X_{[0]})_{\text{lin}} = \lambda_1 x^1 \frac{\partial}{\partial x^1} + \sum_{j=2}^n (\lambda_j x^j + \varepsilon_j x^{j-1}) \frac{\partial}{\partial x^j},$$

where $(\lambda_1, \dots, \lambda_n)$ is a w -compatible ordering of the eigenvalues of $(X_{[0]})_{\text{lin}}$ and $\varepsilon_j \in \{0, 1\}$ corresponds to whether p^j is an eigenvector or a generalized eigenvector of higher rank. Thus we can write $\varphi^*X_{[0]} = \Lambda + J + T$, where

$$\Lambda := \sum_{j=1}^n \lambda_j x^j \frac{\partial}{\partial x^j}, \quad J := \sum_{j=2}^n \varepsilon_j x^{j-1} \frac{\partial}{\partial x^j}, \quad T := \varphi^* (X_{[0]} - (X_{[0]})_{\text{lin}}).$$

A straightforward calculation shows that monomial vector fields $x^\alpha \frac{\partial}{\partial x^i}$ are eigenvectors of ad_Λ with eigenvalues $\langle \lambda, \alpha \rangle - \lambda_i$. We claim that the matrix representations of ad_J and ad_T on $\mathfrak{X}_{[k]}$ are both *strictly* upper triangular when the basis $\left\{ x^\alpha \frac{\partial}{\partial x^i} \mid \langle w, \alpha \rangle - w_i = k \right\}$ of $\mathfrak{X}_{[k]}$ is ordered as follows:

⁷Grabowski–Rotkiewicz call a map preserving weighted polynomial degree a *morphism of graded spaces*.

First, sort the monomial vector fields $x^\alpha \frac{\partial}{\partial x^i}$ in decreasing order of $|\alpha|$. Among those $x^\alpha \frac{\partial}{\partial x^i}$ with a fixed value of $|\alpha|$, sort the ones with $\frac{\partial}{\partial x^i}$ before the ones with $\frac{\partial}{\partial x^j}$ if $i > j$. Among those $x^\alpha \frac{\partial}{\partial x^i}$ with fixed values of $|\alpha|$ and i , sort the x^α 's in decreasing lexicographical order of the multi-indices α .

From this claim, it follows that $\text{ad}_{\varphi^* X_{[0]}} = \text{ad}_\Lambda + \text{ad}_J + \text{ad}_T$ is represented by an upper triangular matrix whose diagonal entries are of the form $\langle \lambda, \alpha \rangle - \lambda_i$ where $\langle w, \alpha \rangle - w_i = k$. As X is non-resonant of weighted degree k , these diagonal entries are non-zero, so $\text{ad}_{\varphi^* X_{[0]}}$ is invertible on $\mathfrak{X}_{[k]}$.

For ad_J , let $j \in \{2, \dots, n\}$ be such that $\varepsilon_j = 1$. (This is only possible if $w_j = w_{j-1}$; otherwise J would have a term $x^{j-1} \frac{\partial}{\partial x^j}$ with negative weighted degree.) We compute

$$\left[x^{j-1} \frac{\partial}{\partial x^j}, x^\alpha \frac{\partial}{\partial x^i} \right] = x^{j-1} \frac{\partial x^\alpha}{\partial x^j} \frac{\partial}{\partial x^i} - x^\alpha \frac{\partial x^{j-1}}{\partial x^i} \frac{\partial}{\partial x^j}.$$

If the first term is non-zero, then $\frac{\partial x^\alpha}{\partial x^j} \neq 0$; writing $\alpha = (\alpha_1, \dots, \alpha_n)$, this implies that $\alpha_j > 0$. Then

$$x^{j-1} \frac{\partial x^\alpha}{\partial x^j} = \alpha_j x^{\alpha'}, \quad \text{where } \alpha' = (\alpha_1, \dots, \alpha_{j-1} + 1, \alpha_j - 1, \dots, \alpha_n).$$

As α' is lexicographically greater than α , the monomial vector field $x^{\alpha'} \frac{\partial}{\partial x^i}$ comes before $x^\alpha \frac{\partial}{\partial x^i}$ in the ordered basis. If the second term is non-zero, then $i = j - 1$, in which case $x^\alpha \frac{\partial x^{j-1}}{\partial x^i} \frac{\partial}{\partial x^j} = x^\alpha \frac{\partial}{\partial x^{i+1}}$, which also comes before $x^\alpha \frac{\partial}{\partial x^i}$ in the prescribed order.

For ad_T , recall that $X_{[0]} - (X_{[0]})_{\text{lin}}$ only involves monomial vector fields $x^\beta \frac{\partial}{\partial x^j}$ such that $|\beta| > 1$. As φ^* preserves the unweighted grading on $\mathfrak{X}_{\text{pol}}$, $T = \varphi^*(X_{[0]} - (X_{[0]})_{\text{lin}})$ also only involves such monomial vector fields. Thus $[T, x^\alpha \frac{\partial}{\partial x^i}]$ is a linear combination of monomial vector fields $x^\gamma \frac{\partial}{\partial x^k}$ such that $|\gamma| > |\alpha|$, all of which come before $x^\alpha \frac{\partial}{\partial x^i}$ in the order described above. ■

EXAMPLE 4.9. Assign the weighting $w = (1, 2, 2, 3, 3)$ to the coordinates (x, y, z, u, v) of $\mathbb{K}^5$. For arbitrary coefficients $a, b, c, d, e, f, g, h \in \mathbb{K}$, the vector field

$$\begin{aligned} X_{[0]} = & x \frac{\partial}{\partial x} + (ax^2 + 4y + z) \frac{\partial}{\partial y} + (bx^2 - 4y) \frac{\partial}{\partial z} \\ & + (cx^3 + dxy + exz + 4u - v) \frac{\partial}{\partial u} + (fx^3 + gxy + hzx + u + 2v) \frac{\partial}{\partial v} \end{aligned}$$

is quasi-homogeneous of weighted degree 0, and its unweighted linear approximation is

$$(X_{[0]})_{\text{lin}} = x \frac{\partial}{\partial x} + (4y + z) \frac{\partial}{\partial y} - 4y \frac{\partial}{\partial z} + (4u - v) \frac{\partial}{\partial u} + (u + 2v) \frac{\partial}{\partial v}.$$

A w -compatible ordering of the eigenvalues of $(X_{[0]})_{\text{lin}}$ is given by $\lambda = (1, 2, 2, 3, 3)$, and a corresponding ordered Jordan basis of $(\mathbb{K}^5)^\vee$ is given by $\{x, 2y + z, y, u - v, u\}$. Thus the formal diffeomorphism ψ from the proof of Theorem 4.8 and its inverse φ act on $\mathbb{K}[[x, y, z, u, v]]$ by

$$\begin{aligned} \psi(x) &= x, & \psi(y) &= 2y + z, & \psi(z) &= y, & \psi(u) &= u - v, & \psi(v) &= u, \\ \varphi(x) &= x, & \varphi(y) &= z, & \varphi(z) &= y - 2z, & \varphi(u) &= v, & \varphi(v) &= -u + v, \end{aligned}$$

and it is straightforward to verify that φ^* puts $(X_{[0]})_{\text{lin}}$ into Jordan canonical form:

$$\varphi^* \left((X_{[0]})_{\text{lin}} \right) = x \frac{\partial}{\partial x} + 2y \frac{\partial}{\partial y} + (y + 2z) \frac{\partial}{\partial z} + 3u \frac{\partial}{\partial u} + (u + 3v) \frac{\partial}{\partial v}.$$

It is a happy coincidence that $\lambda = w$, from which it follows that λ is non-resonant with respect to w (but resonant in the unweighted sense). By Theorem 4.8, any w -admissible formal vector field with weighted linear approximation $X_{[0]}$ can be conjugated to $X_{[0]}$ via a formal diffeomorphism. ♦

4.3. Smooth weighted linearization. We conclude by returning to our original motivating question for developing the formal machinery for Moser’s trick. Let X be a smooth vector field on an open neighbourhood of $0 \in \mathbb{R}^n$. If X is admissible, non-resonant, and has weighted linear approximation $X_{[0]} \in \mathfrak{X}_{[0]}$ with respect to a weighting w of $\mathbb{R}^n$, then there exists a formal diffeomorphism φ of $\mathbb{R}^n$ such that $\varphi^*X = X_{[0]}$ by Theorem 4.8. To upgrade φ to an actual diffeomorphism using Chen’s theorem of equivalence [7], one would like to know if weighted non-resonance implies hyperbolicity. This is indeed the case and the proof is almost the same as the unweighted setting, but it is not immediate as Definition 4.7 involves w -compatible orderings:

LEMMA 4.10. *Let X be a smooth vector field on an open neighbourhood of $0 \in \mathbb{R}^n$. If there exists a weighting of $\mathbb{R}^n$ with respect to which X is admissible and non-resonant, then 0 is a hyperbolic fixed point of X .*

Proof. Suppose that 0 is *not* a hyperbolic fixed point, w is any weighting with respect to which X is admissible, and $\lambda = (\lambda_1, \dots, \lambda_n)$ is any w -compatible ordering of the (possibly complex) eigenvalues of $(X_{[0]})_{\text{lin}}$. If some eigenvalue λ_j vanishes, then $2\lambda_j = \lambda_j$ is a resonance with respect to w as $2w_j > w_j$. If not, then there is a complex conjugate pair of non-zero purely imaginary eigenvalues $\pm i\theta$. By Lemma 4.6(ii), $(X_{[0]})_{\text{lin}}$ preserves the grading $(\mathbb{R}^n)^\vee = \bigoplus_{\ell=1}^m V^\ell$, so this pair must come from the restriction of $(X_{[0]})_{\text{lin}}$ to some V^ℓ : if any such pair came from two *different* V^ℓ ’s, then the characteristic polynomial of some restriction of $(X_{[0]})_{\text{lin}}$ to one of the V^ℓ ’s would have an odd number of non-real roots when counted with multiplicity, which is impossible for a polynomial with real coefficients. Thus $i\theta = 2(i\theta) + 1(-i\theta)$ is a resonance with respect to w . ■

We obtain the smooth half of Theorem 1.1:

COROLLARY 4.11. *Let X be a smooth vector field on an open neighbourhood of $0 \in \mathbb{R}^n$. For any weighting of $\mathbb{R}^n$ with respect to which X is admissible, non-resonant, and has weighted linear approximation $X_{[0]} \in \mathfrak{X}_{[0]}$, there exists a germ of a diffeomorphism φ of $\mathbb{R}^n$ fixing 0 such that $\varphi^*X = X_{[0]}$.*

Proof. Apply Theorem 4.8, Lemma 4.10, and Chen’s theorem of equivalence. ■

In particular, the weighted Euler vector field (4.1) associated to any weighting w of $\mathbb{R}^n$ is admissible and non-resonant with respect to w . We thus recover Algaba–García–Reyes’ and Meinrenken’s result that every smooth weighted Euler-like vector field on $\mathbb{R}^n$ is smoothly linearizable.

REFERENCES

- [1] Antonio Algaba, Cristóbal García, and Manuel Reyes. Like-linearizations of vector fields. *Bulletin des Sciences Mathématiques*, 133(8):806–816, 2009. doi:10.1016/j.bulsci.2009.09.006.
- [2] V. I. Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations*, volume 250 of *Grundlehren der mathematischen Wissenschaften*. Springer, New York, NY, second edition, 1988. doi:10.1007/978-1-4612-1037-5.
- [3] Hyman Bass and Gary Meisters. Polynomial flows in the plane. *Advances in Mathematics*, 55(2):173–208, 1985. doi:10.1016/0001-8708(85)90020-9.
- [4] M. Faisal Beg, Michael I. Miller, Alain Trounev, and Laurent Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005. doi:10.1023/B:VISI.0000043755.93987.aa.
- [5] Nicolas Bourbaki. *Algebra II*. Springer, Berlin, Heidelberg, 2003. doi:10.1007/978-3-642-61698-3.
- [6] Henrique Bursztyn, Hudson Lima, and Eckhard Meinrenken. Splitting theorems for Poisson and related structures. *Journal für die reine und angewandte Mathematik*, 2019(754):281–312, 2019. doi:10.1515/crelle-2017-0014.
- [7] Kuo-Tsai Chen. Equivalence and decomposition of vector fields about an elementary critical point. *American Journal of Mathematics*, 85(4):693–722, 1963. doi:10.2307/2373115.

-
- [8] Paul Eakin and Avinash Sathaye. R -endomorphisms of $R[[X]]$ are essentially continuous. *Pacific Journal of Mathematics*, 66(1):83–87, 1976. doi:[10.2140/pjm.1976.66.83](https://doi.org/10.2140/pjm.1976.66.83).
- [9] Gene Freudenburg. *Algebraic Theory of Locally Nilpotent Derivations*, volume 136 of *Encyclopaedia of Mathematical Sciences*. Springer, Berlin, Heidelberg, second edition, 2017. doi:[10.1007/978-3-662-55350-3](https://doi.org/10.1007/978-3-662-55350-3).
- [10] Janusz Grabowski and Mikołaj Rotkiewicz. Graded bundles and homogeneity structures. *Journal of Geometry and Physics*, 62(1):21–36, 2012. doi:[10.1016/j.geomphys.2011.09.004](https://doi.org/10.1016/j.geomphys.2011.09.004).
- [11] Joong Ho Kim. R -automorphisms of $R[t][[X_1, \dots, X_n]]$. *Journal of Algebra*, 70(1):37–47, 1981. doi:[10.1016/0021-8693\(81\)90242-8](https://doi.org/10.1016/0021-8693(81)90242-8).
- [12] Ernst Kunz. *Kähler Differentials*. Advanced Lectures in Mathematics. Springer Vieweg+Teubner Verlag, Wiesbaden, 1986. doi:[10.1007/978-3-663-14074-0](https://doi.org/10.1007/978-3-663-14074-0).
- [13] Yiannis Loizides and Eckhard Meinrenken. Differential geometry of weightings. *Advances in Mathematics*, 424:109072, 2023. doi:[10.1016/j.aim.2023.109072](https://doi.org/10.1016/j.aim.2023.109072).
- [14] Eckhard Meinrenken. Euler-like vector fields, normal forms, and isotropic embeddings. *Indagationes Mathematicae*, 32(1):224–245, 2021. doi:[10.1016/j.indag.2020.08.006](https://doi.org/10.1016/j.indag.2020.08.006).
- [15] Gary Meisters and Czesław Olech. A poly-flow formulation of the Jacobian conjecture. *Bulletin of the Polish Academy of Sciences*, 35(11-12):725–731, 1987.
- [16] Jürgen Moser. On the volume elements on a manifold. *Transactions of the American Mathematical Society*, 120(2):286–286, 1965. doi:[10.1090/S0002-9947-1965-0182927-5](https://doi.org/10.1090/S0002-9947-1965-0182927-5).
- [17] Raghavan Narasimhan. *Analysis on Real and Complex Manifolds*, volume 35 of *North-Holland Mathematical Library*. Elsevier, Amsterdam, North Holland, 1985.
- [18] Richard S. Palais. The Morse lemma for Banach spaces. *Bulletin of the American Mathematical Society*, 75(5):968–971, 1969. doi:[10.1090/S0002-9904-1969-12318-9](https://doi.org/10.1090/S0002-9904-1969-12318-9).
- [19] Henri Poincaré. *Sur les propriétés des fonctions définies par les équations aux différences partielles*. PhD thesis, University of Paris, 1879.
- [20] Shlomo Sternberg. Local contractions and a theorem of Poincaré. *American Journal of Mathematics*, 79(4):809, 1957. doi:[10.2307/2372437](https://doi.org/10.2307/2372437).
- [21] Shlomo Sternberg. On the structure of local homeomorphisms of Euclidean n -space, II. *American Journal of Mathematics*, 80(3):623, 1958. doi:[10.2307/2372774](https://doi.org/10.2307/2372774).
- [22] Alan Weinstein. Symplectic structures on Banach manifolds. *Bulletin of the American Mathematical Society*, 75(5):1040–1041, 1969. doi:[10.1090/S0002-9904-1969-12353-0](https://doi.org/10.1090/S0002-9904-1969-12353-0).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, TORONTO, CANADA
 Email address: arthur.qiu@mail.utoronto.ca