Ordinary differential equations and their exponentials

Anders Kock[1]*, Gonzalo E. Reyes[2]†

[1] *Department of Mathematical Sciences,*
University of Aarhus,
DK 8000 Aarhus C, Denmark

[2] *Department of Mathematics,*
Université de Montréal,
H3C 3J7 Montreal, Quebec, Canada

Abstract: In the context of Synthetic Differential Geometry, we discuss vector fields/ordinary differential equations as actions; in particular, we exploit function space formation (exponential spaces) in the category of actions.

Vector fields or, equivalently, (autonomous, first order) ordinary differential equations, have long been considered, heuristically, to be the same as "infinitesimal (pointed) actions" or "infinitesimal flows", but it is only with the development of Synthetic Differential Geometry (SDG) that we have the tools to formulate these notions and prove their equivalence in a rigourous mathematical way. We exploit this fact to define the exponential of two ordinary differential equations as the exponential of the corresponding infinitesimal actions. The resulting action is seen to be the same as a partial differential equation whose solutions may be obtained by conjugation from the solutions of the differential equations that make up the exponential. Furthermore, we show that this method of conjugation, under some conditions, is an application of the method of change of variables, widely used to solve differential equations.

Our paper has three parts. In the first, we give a brief introduction to the method of

* E-mail: kock@imf.au.dk
† E-mail: reyes@dms.umontreal.ca

Synthetic Differential Geometry (SDG). In the second, we study generalities on actions, and in the third, we describe the exponential of two such actions to obtain the above mentioned result. Some examples illustrate the general method.

1   A synthetic method in differential geometry

What today is called Synthetic Differential Geometry is a method of reasoning which has as one of its origins 20th Century French Algebraic Geometry (Grothendieck et al.), the other origin being certain aspects of modern category/topos theory.

The school of algebraic geometers alluded to insisted on the consideration of *nilpotent* elements in function rings of the geometric objects (schemes).

Recall that an element d of a commutative ring R is called nilpotent (or infinitesimal) of order k if $d^{k+1} = 0$.

For such infinitesimals as increments, Taylor *series* and kth degree Taylor *polynomials* become the same thing, since we do have (to the extent it makes sense)

$$f(x + d) = f(x) + d \cdot f'(x) + \frac{d^2}{2!}f''(x) + \ldots + \frac{d^k}{k!}f^{(k)}(x) \tag{1}$$

exactly, since the remaining terms in the series are zero.

A viewpoint related to the consideration of nilpotent elements is the consideration of the "kth neigbourhood of the diagonal" of a scheme, or manifold, M,

$$M_{(k)} \subseteq M \times M,$$

also a classical concept in 20th Century algebraic, and even differential, geometry, cf. [6, 9]. For instance, if $M = $ affine line, $(x, y) \in M_{(k)}$ precisely when $y - x$ is a kth order infinitesimal, in the above sense.

The potential which this viewpoint evidently has for a new foundation for differential calculus and differential geometry could not be fully realized, until certain category theoretic notions, and the categorical logic inherent in these, became more crystallized - notably through the notion of topos.

Through the work of Lawvere and his collaborators in the 1960's and 1970's, it became gradually clear that any topos behaves so much like the familiar category of sets that most set theoretic reasoning and construction immediately can be carried out in any topos (and toposes have been constructed which contain the category of smooth manifolds, say, in a suitable way). The exception to set theoretic reasoning in general toposes is the *law of excluded middle*, as well as the axiom of choice. These exceptions certainly are to be expected, since they are responsible for constructions of non-smooth entities out of smooth ones; e.g.

$$f(x) = \begin{cases} 1 & \text{if } x \leq 0; \\ 0 & \text{if not.} \end{cases}$$

It has been amply documented (by construction of models in suitable toposes) that at the cost of not employing these two "non-smooth" logical principles, a stronger axiomatic

theory of differential calculus becomes consistent, namely one where the affine "number" line has a sufficient supply of nilpotent elements d such that (1) not only holds for all kth order infinitesimals d, but that this equation *characterizes* $f'(x)$, $f''(x)\ldots,\ldots, f^{(k)}(x)$. A fair amount of differential calculus can be carried out on such axiomatic basis, meaning in particular that it can be carried out without utilizing (or even knowing) any topos theory. Such endeavour is sometimes called "naive" synthetic differential geometry, and the present note is entirely conceived in this naive style.

The only trace of category theoretic notions that are essential for the considerations which we present here, is that of *cartesian closed category*. Any topos, in contrast to the category of schemes or smooth manifolds, is cartesian closed, which means that *function spaces* or *exponential objects* are immediately available: for any pair of "spaces" (objects) A and B, there is a "space" A^B (to be thought of as the "space of maps from B to A") with the property that, for any X, there are natural bijective correspondences between the set of maps $X \to A^B$, the set of maps $X \times B \to A$, and the set of maps $B \to A^X$. (These correspondences are often called λ-*conversions*, a term borrowed from logic; in category theory, they are called "exponential adjointness".)

Lawveres's essential contribution from 1967 (see [8]) was to realize that if D is the set of 1st order infinitesimals on the affine line R (i.e. $D = \{x \in R \mid x^2 = 0\}$), then A^D is the tangent bundle $T(A)$ of the "space" A. Thus he observed that a vector field on A comes by exponential adjointness in three equivalent disguises:

$$\frac{\begin{array}{c} A \qquad \to A^D \\ \hline A \times D \to \ A \end{array}}{D \qquad \to A^A}$$

(in each of the three disguises, there is an equational condition involving $0 \in D$). This is the viewpoint we exploit in the following two sections. Hopefully, the text there will illustrate how the synthetic/axiomatic theory works. (Further illustrations may be found in any of the three standard treatises on the subject, [3, 7, 10].)

To make the character of the axiomatics explicit, we are dealing with an (unspecified) cartesian closed category (whose objects we just call "spaces" or "sets"), and a given particular object R in it, which carries the structure of a commutative ring, to be thought of as the affine number line. This ring object R is assumed to be an algebra over the rationals, i.e. we assume that $1+1$, $1+1+1$, etc. are multiplicatively invertible in R. The main thing is that the Taylor formula holds and characterizes derivatives ("Kock-Lawvere axiom").

The text should be reasonably self-contained, except that we need to consider "spaces" M which have the further technical property of being *microlinear*, cf. e.g. [7] for a lucid exposition of this notion. It essentially means that maps from infinitesimal objects D' into M can be constructed by patching, – as illustrated in the text.

2　Generalities on actions

Recall that an *action* of a set (object) D on a set (object) M is a map $X : D \times M \to M$, and a *homomorphism* of actions $(M, X) \to (N, Y)$ is a map $f : M \to N$ with $f(X(d, m)) = Y(d, f(m))$ for all $m \in M$ and $d \in D$.

The category of actions by a set D form a cartesian closed category (a topos, in fact). We shall be interested in particular in the exponential formation in this category (cf. section 3).

We place ourselves in the context of Synthetic Differential Geometry, as described briefly in Section 1. As there, we take D to be the usual set of square zero elements in the number line R. It is a *pointed* object, pointed by $0 \in D$, and the actions $X : D \times M \to M$ we consider, are *pointed* actions in the sense that $X(0, m) = m$ for all $m \in M$. A pointed action, in this situation, is the same thing as a vector field on M, cf. [8]; namely for $m \in M$, $X(m)$ is the tangent vector at m given by $d \mapsto X(d, m)$.

The pointed actions likewise form a cartesian closed category (again, a topos) and the exponential to be described in section 3 is the same as the exponential in the category of actions (cf. [4]).

For the case of vector fields seen as actions by D, we want to describe the "streamlines" generated by a vector field in abstract action-theoretic terms; this is going to involve a certain "universal" action $(\tilde{R}, \partial/\partial t)$: $\tilde{R}$ is an "infinitesimally open subset" of R, i.e., whenever $t \in \tilde{R}$ then $d + t \in \tilde{R}$ for every $d \in D$; we also assume $0 \in \tilde{R}$. The simplest examples of such subsets are R itself, the non-negative numbers $R_{\geq 0}$, open intervals around 0, and the set D_∞ of all nilpotent elements of the number line. (Another important example is the object Δ described and utilized in [1].) The universal action is the vector field $\partial/\partial t : D \times \tilde{R} \to \tilde{R}$ given by $(d, t) \mapsto d + t$.

If (M, X) is a set with an action, a homomorphism of actions $f : (\tilde{R}, \partial/\partial t) \to (M, X)$ is to be thought of as a particular solution of the differential equation given by X, with initial value $f(0)$, or as a "streamline" for the vector field X, starting at $f(0)$. One wants, however, also to include dependence on initial value into the notion of solution, and so one is led to consider maps

$$F : \tilde{R} \times M \to M,$$

satisfying at least $F(d, m) = X(d, m)$ for all $d \in D$ and all $m \in M$.

We shall for any $t \in \tilde{R}$, $d \in D$ and $m \in M$ consider the following three points in M:

$$F(d + t, m), \quad X(d, F(t, m)) \quad F(t, X(d, m)).$$

We shall consider and compare the three conditions one gets by pairwise equating these (universally quantified over all t, d, m); the first (2) is the fundamental one, expressing that each $F(-, m)$ is a particular solution (streamline) of the ODE given by the vector field X.

$$F(d + t, m)) = X(d, F(t, m)), \tag{2}$$

$$F(d + t, m)) = F(t, X(d, m)), \tag{3}$$

$$F(t, X(d, m)) = X(d, F(t, m)). \tag{4}$$

Writing X_d for the map $X(d, -) : M \to M$, and similarly F_t for $F(t, -) : M \to M$, these three conditions may be rewritten as

$$F_{d+t} = X_d \circ F_t, \tag{5}$$

$$F_{d+t} = F_t \circ X_d, \tag{6}$$

$$F_t \circ X_d = X_d \circ F_t, \tag{7}$$

(these three equations universally quantified over all t, d). The three equations can be reformulated in "classical" terms, i.e. without reference to $d \in D$, using the notion of differential $d(g)$ of a map $g : M \to N$; we return to the exact meaning of the terms occurring here later, but include the equations now for systematic reasons

$$\frac{\partial F}{\partial t}(t, m) = X(F(t, m)) \tag{8}$$

$$\frac{\partial F}{\partial t}(t, m) = d(F_t)(X(m)) \tag{9}$$

$$d(F_t)(X(m)) = X(F_t(m)) \tag{10}$$

(these three equations are universally quantified over all t, m). If M is a suitable vector space (R-module), or a suitable subset hereof, these three last equations have a simpler appearance, via the notion of the *principal part* ξ of the vector field X, see (13), (14), (15) below.

Finally, one may consider the following equation, under the condition that $\tilde{R}$ is a submonoid of $(R, +)$,

$$F(t + s, m) = F(t, F(s, m)). \tag{11}$$

This is the usual condition for an action af a monoid on a set M. Clearly, it implies (2), (3) and (4). But note that we do not in general assume that $\tilde{R}$ is a submonoid of $(R, +)$.

Let X be a vector field on M, thought of as a first-order differential equation, and let $\tilde{R}$ be an infinitesimally open subset of R, containing 0. We say that a map $f : \tilde{R} \to M$ is a *particular solution of* $(M, X, \tilde{R})$ if f satisfies $f(t + d) = X(d, f(t))$ (or equivalently, if f is a homomorphism of actions).

We say that a map $F : \tilde{R} \times M \to M$ is a *complete solution of* $(M, X, \tilde{R})$ if $F_d = X_d$ and F satisfies (2).

From now on, we shall usually omit reference to M and $\tilde{R}$, which will be presupposed, and talk of solutions 'of a vector field X'.

A complete solution does not automatically satisfy the other conditions (3)-(11), but it does, provided that X satisfies a certain axiom (reflecting, synthetically, validity of the uniqueness assertion for solutions of differential equations on M).

The axiom in question is the following:

Uniqueness property for particular solutions of X:

Let X be a pointed D-action on M. If $f, g : \tilde{R} \to M$ are homomorphisms of actions, with $f(0) = g(0)$, then $f = g$.

Note that the validity of the axiom, for a given X, depends on M and the choice of $\tilde{R}, \partial/\partial t$. For instance, we shall prove below that it holds for any microlinear M if $\tilde{R}$ is taken to be D_∞ (and $\partial/\partial t$ given by $(d, t) \mapsto d + t$).

This axiom has the following simple consequence:

Uniqueness property for complete solutions of X:

Proposition 2.1. *Assume that X has the uniqueness property for particular solutions. Then there is at most one complete solution $F : \tilde{R} \times M \to M$*

The converse does not seem to be true, but a weaker result is true, see Proposition 2.8 below.

Proposition 2.2. *Let X be a vector field on M and assume that X satisfies the uniqueness property for particular solutions. Then if $F : \tilde{R} \times M \to M$ is a complete solution of the differential equation X, it satisfies properties (3) and (4). Furthermore, if $\tilde{R}$ is a monoid (under $+$), then F also satisfies (11).*

Proof. Since the proofs are quite similar, we shall do only (3). Fix $m \in M$ and $d_0 \in D$, and define the couple of functions $f, \ g : \tilde{R} \to M$ by the formulas

$$
\begin{cases}
f(t) = F_{d_0+t}(m) \\
g(t) = F_t(X_{d_0}(m))
\end{cases}
$$

Clearly, f and g have the same initial value $f(0) = F(d_0, m) = X(d_0, m) = g(0)$. We have to check that f and g are homomorphisms of D-actions, i.e., they satisfy (2). For g, this is clear. For f,

$$f(d + t) = F_{d_0+(d+t)}(m)$$

$$= F_{d+(d_0+t)}(m)$$

$$= X_d(F_{d_0+t}(m))$$

$$= X_d(f(t)).$$

Thus, the equality of the two expressions follows from the uniqueness property for particular solutions assumed for X. $\qquad\qquad\square$

We recall the notion of the *differential* of a map. Recall that the set M^D is the tangent bundle TM of M; it comes with a map $TM \to M$, (base point map), namely the one

which takes $\tau : D \to M$ to $\tau(0)$. The fibre of TM over $x \in M$ is denoted $T_x M$. If $f : M \to N$ is any map, it induces for each $x \in M$ a map $T_x M \to T_{f(x)} N$, called the *differential* $d_x(f)$ of f at x; it is given by $\tau \mapsto f \circ \tau$. If M and N are microlinear, $d_x(f)$ will be a linear map. (The differentials of f jointly define a map $M^D \to N^D$, which is nothing but the functor $(-)^D$ applied to f.)

We shall also recall some notions that apply to any "Euclidean R-module" $M = V$, or to an "infinitesimally open" subset $M = U \subseteq V$ hereof. These notions are standard in SDG, but let us briefly review them: For any R-module V, we have the map $V \times V \to V^D$ given by $(a, b) \mapsto [d \mapsto a + d \cdot b]$. To say that V is *Euclidean* is to say that this map is a bijection; in other words, every tangent vector $\tau : D \to V$ is uniquely of the form $d \mapsto a + d \cdot b$. The element ("vector") $b \in V$ is called the *principal part* of the tangent vector τ (and a is of course the *base point* of τ). To say that $U \subseteq V$ is *infinitesimally open* is to say that if a tangent vector τ to V, as above, has its base point $\tau(0)$ in U, then $\tau(d)$ is in U for all $d \in D$. Thus $U \times V \cong U^D \ (= T(U))$. For such U and $x \in U$, we may, via the notion of principal part, identify $T_x(U)$ with V.

Recall also that if $\beta : M \to V$ is any map into a Euclidean R-module, and X is a vector field on M, then the *directional derivative* $D_X(\beta)$ of β along X is the composite

$$M \xrightarrow{X} M^D \xrightarrow{\beta^D} V^D \to V,$$

where the last map is the principal part formation. Thus, $D_X(\beta)$ is characterized by validity of the equation

$$\beta(X(d, m)) = \beta(m) + d \cdot D_X(\beta)(m),$$

for all $d \in D$, $m \in M$. Equivalently, $D_X(\beta)(m)$ is the principal part of $d_m(\beta)(X(m))$. (Recall that $d_m \beta$ is the differential of β at m.)

Proposition 2.3. *Assume that X_1, X_2 are vector fields on M_1, M_2, respectively, and that $H : M_1 \to M_2$ is a homomorphism (i.e., it preserves the D-action). Let V be a Euclidean R-module. Then for any $u : M_2 \to V$,*

$$D_{X_1}(u \circ H) = D_{X_2}(u) \circ H.$$

Proof. This is a straightforward computation:

$$u(X_2(d, H(m))) = u(H(m)) + d \cdot D_{X_2}(H(m));$$

on the other hand

$$u(X_2(d, H(m))) = u(H(X_1(d, m))) = u(H(m)) + d \cdot D_{X_1}(u \circ H)(m).$$

By comparing these two expressions we obtain the conclusion of the Proposition. $\square$

If X is a vector field on such Euclidean V, we get, by principal part formation a map $\xi : V \to V$ (i.e. the principal part of the field vector $X(v)$ for $v \in V$ is $\xi(v)$). This map is often called the *principal part* of the vector field. Similarly, a vector field X on an infinitesimally open $U \subseteq V$ may be identified with a map $\xi : U \to V$. In this case, one often writes $D_\xi(\beta)$ instead of $D_X(\beta)$.

The notion of differential discussed above has a variant in case of a map $f : U \to V$ between Euclidean R-modules, due to the identification (via principal part formation) of tangent vectors to V with vectors in V. For x and $u \in U$, write $df(x; u) \in V$ for the principal part of $d_x(f)(\overline{u})$, where $\overline{u}$ is the tangent vector at x whose principal part is u. It depends in a linear way on $u \in U$. (It makes sense also when U is an infinitesimally open subset of a Euclidean R-module.) In this case $D_\xi(\beta)(x)$ can also be described as $d\beta(x; \xi(x))$. In the 1-dimensional case where $U \subseteq R = V$,

$$D_\xi(\beta)(x) = d\beta(x; \xi(x)) = \beta'(x) \cdot \xi(x). \tag{12}$$

If $g : \tilde{R} \times M \to N$ is a map, $\partial g/\partial t(t, x)$ is by definition the tangent vector at $g(t, x) \in N$ given by $d \mapsto g(d + t, x)$, (– this is the usage in the equations (8) and (9)), or its principal part when this makes sense. It may also be written $\dot{g}(t, x)$.

With these notations, (8), (9) and (10) may be reformulated

$$\frac{\partial F}{\partial t}(t, m) = \xi(F(t, m)) \tag{13}$$

$$\frac{\partial F}{\partial t}(t, m) = d(F_t)(m; \xi(m)) \tag{14}$$

$$d(F_t)(m; \xi(m)) = \xi(F_t(m)) \tag{15}$$

Although this latter equation does not tell us how the solution of the vector field varies with the initial value m, (varying m in arbitrary directions), it does tell us the variation of the solution when m varies in the direction prescribed by the vector field. In the one-dimensional case, there is only one direction anyway; so here we get (using (12) with $\beta = F_t$) the following version of (15) for the complete solution $F(t, x)$ of $\dot{x} = \xi(x)$ (i.e. of the vector field with principal part ξ):

$$\frac{\partial F}{\partial x}(t, x) \cdot \xi(x) = \xi(F(t, x)), \tag{16}$$

or equivalently

$$\frac{\partial F}{\partial x}(t, x) \cdot \xi(x) = \frac{\partial F}{\partial t}(t, x). \tag{17}$$

(An elementary proof, for the case of nonvanishing ξ, goes as follows: Let $\ln_\xi$ denote a primitive of $1/\xi$. Then we have

$$\ln_\xi(F(t, x)) = t + \ln_\xi(x);$$

for, both sides yield $\ln_\xi(x)$ for $t = 0$ and have the same t-derivative, namely 1 (using that $F(-, x)$ is a solution of the differential equation). Differentiating this equation after x, we get $(\xi(F(t, x))^{-1}\partial F/\partial x(t, x) = 1/\xi(x)$, which is a rewriting of (16).

Recall that a vector field X on M is called *integrable* if there exists a complete solution $F : \tilde{R} \times M \to M$. If we assume the uniqueness property for particular solutions, the equation (11) holds; if furthermore the commutative monoid structure $+$ on $\tilde{R}$ actually is a group structure, then (11) implies that the action is invertible, with X_{-d} as X_d^{-1} (in fact $F_{-t} = F_t^{-1}$). Of course, both the uniqueness property and the question whether or not the vector field X is integrable, depend on which $\tilde{R}$ is considered. In particular, we shall say that X is *formally integrable* or *has a formal solution* if X is integrable for $\tilde{R} = D_\infty$ (which is a group under addition). For the case of $M = R^n$, this amounts to integration by formal power series, whence the terminology.

Theorem 2.4. *The uniqueness property for particular solutions holds for any vector field on a microlinear object, (for $\tilde{R} = D_\infty$). Furthermore, every vector field on a microlinear object is formally integrable. Thus, every vector field on a microlinear object has a unique complete (formal) solution.*

Proof. This theorem was stated in [1] and a sketch of the proof by induction was indicated. We give here a proof in detail which does not use induction.

We need to recall some infinitesimal objects from the literature on SDG, cf. e.g. [7]. Besides $D \subseteq R$, consisting of $d \in R$ with $d^2 = 0$, we have $D^n \subseteq R^n$, the n-fold product of D with itself. It has the subobject $D(n) \subseteq D^n$ consisting of those n-tuples $(d_1, \ldots, d_n)$ where $d_i \cdot d_j = 0$ for all i, j. There is also the object $D_n \subseteq R$ consisting of $\delta \in R$ with $\delta^{n+1} = 0$; D_∞ is the union of all the D_n's. If $(d_1, \ldots, d_n) \in D^n$, then $d_1 + \ldots + d_n \in D_n$.

Now, let M be a microlinear object, and X a vector field on it. We first recall that if $d_1, d_2 \in D$ have the property that $d_1 + d_2 \in D$, then $X_{d_1} \circ X_{d_2} = X_{d_1+d_2}$. (For microlinear objects perceive $D(2)$ to be a pushout over $\{0\}$ of the two inclusions $D \to D(2)$, and clearly both expressions given agree if either $d_1 = 0$ or $d_2 = 0$.) In particular, X_{d_1} and X_{d_2} commute. But more generally, we have the following lemma.

Lemma 2.5. *If X is a vector field on a microlinear object and $d_1, d_2 \in D$, the maps X_{d_1} and X_{d_2} commute.*

Proof. This is a consequence of the theory of Lie brackets, cf. e.g. [7] 3.2.2, namely $[X, X] = 0$. □

Similarly,

Lemma 2.6. *If X is a vector field on a microlinear object and $d_1, \ldots, d_n \in D$ are such that $d_1 + \cdots + d_n = 0$, then*

$$X_{d_1} \circ \cdots \circ X_{d_n} = 1_M$$

(= the identity map on M). In particular, $(X_d)^{-1} = X_{-d}$.

Proof. We first prove that R, and hence any microlinear object, perceives D_n to be the

orbit space of D^n under the action of the symmetric group $\mathbf{S}_n$ in n letters: Assume that $p : D^n \to R$ coequalizes the action, i.e. is symmetric in the n arguments. By the basic axiom of SDG, p may be written in the form

$$p(d_1, \dots, d_n) = \sum_{Q \subseteq \{1,\dots,n\}} a_Q d^Q$$

for unique a_Q's in R (where d^Q denotes $\prod_{i \in Q} d_i$). We claim that $a_Q = a_{\pi(Q)}$ for every $\pi \in \mathbf{S}_n$. Indeed,

$$\sum_Q a_Q d^Q = (\sum_Q a_Q d^Q) \circ \pi$$

since p is symmetric. But

$$(\sum_Q a_Q d^Q) \circ \pi = \sum a_Q d^{\pi(Q)} = \sum_Q a_{\pi^{-1}(Q)} d^Q.$$

By comparing coefficients and using uniqueness of coefficients, we conclude $a_Q = a_{\pi(Q)}$, and this shows that p is (the restriction to D^n of) a symmetric polynomial $R^n \to R$. By Newton's theorem (which holds internally), p is a polynomial in the elementary symmetric polynomials σ_i. Recall that $\sigma_1(d_1, \dots, d_n) = d_1 + \cdots + d_n$: and each σ_i, when restricted to D^n, is a function of σ_1, since $d_1^2 = 0$; e.g.

$$\sigma_2(d_1, \dots, d_2) = \sum d_i d_j = \frac{1}{2}(d_1 + \cdots + d_n)^2 = \frac{1}{2}(\sigma_1(d_1, \dots, d_n))^2.$$

Now consider, for fixed $m \in M$, the map $p : D^n \to M$ given by $(d_1, \dots, d_n) \mapsto X_{d_1} \circ \cdots \circ X_{d_n}(m)$. By Lemma 2.5, this map is invariant under the symmetric group $\mathbf{S}_n$ (recall that this group is generated by transpositions), so there is a unique $\phi : D_n \to M$ such that

$$\phi(d_1 + \cdots + d_n) = X_{d_1} \circ \cdots \circ X_{d_n}(m).$$

So if $d_1 + \cdots + d_n = 0$, $X_{d_1} \circ \cdots \circ X_{d_n}(m) = \phi(0) = \phi(0 + \cdots + 0) = X_0 \circ \cdots \circ X_0(m) = m$. This proves the Lemma. $\square$

We can now prove the Theorem. We need to define $F_t : M \to M$ when $t \in D_\infty$. Assume for instance that $t \in D_n$. By microlinearity of M, M perceives D_n to be the orbit space of D^n under the action of $\mathbf{S}_n$ (see the proof of Lemma 2.6), via the map $(d_1, \dots, d_n) \mapsto d_1 + \cdots + d_n$, so we are forced to define $F_t = X_{d_1} \circ \dots X_{d_n}$ if F is to extend X and to satisfy (11). The fact that this is well defined independently of the choice of n and the choice of $d_1, \dots, d_n$ that add up to t follows from Lemma 2.6. $\square$

As a Corollary of the proof, we may note the following general "analytic induction principle"

Proposition 2.7. *Let f and g be maps $D_\infty \to M$, where M is a microlinear object. If $f(0) = g(0)$, and if for all $t \in D_\infty$ and $d \in D$, $f(t) = g(t)$ implies $f(d + t) = g(d + t)$, then $f = g$.*

Proof. Since D_∞ is the union of the D_n's, it suffices to prove that f and g agree on any D_n. But, as in the proof above, microlinear objects perceive the addition map $D^n \to D_n$ to be epic, and clearly the assumptions on f and g imply that $f(d_1 + \ldots + d_n) = g(d_1 + \ldots + d_n)$ for any $(d_1, \ldots, d_n) \in D^n$. $\square$

Proposition 2.8. *Assume that there is a complete solution $F : D_\infty \times M \to M$ for a vector field X on M (where M is microlinear). Then X has the uniqueness property for particular solutions.*

Proof. Let $f : D_\infty \to M$ be a particular solution of X. We shall prove that

$$f(t) = F(t, f(0))$$

from which the uniqueness of such f clearly follows. The proof of this equation proceeds 'by analytic induction', i.e. using Proposition 2.7: the equation is obviously true for $t = 0$. Assume that it is true for t. We prove that it is true for $d + t$. In fact,

$$f(d + t) = X(d, f(t)) = X(d, F(t, f(0))) = F(d + t, f(0))$$

The first and last equality hold because f and F are solutions, whereas the middle one holds by the induction assumption. – Notice the following consequence of this proposition: If there is a complete solution for X, then it is unique. $\square$

As a particular case of special importance, we consider a *linear* vector field on a microlinear and Euclidean R-module V. To say that the vector field is linear is to say that its principal-part formation $V \to V$ is a linear map, Δ, say. We have then the following version of a classical result:

Proposition 2.9. *Let a linear vector field on a microlinear Euclidean R-module V be given by the linear map $\Delta : V \to V$. Then the unique formal solution of the corresponding differential equation, i.e., the equation $\dot{F}(t) = \Delta(F(t))$ with initial position v, is the map $D_\infty \times V \to V$ given by*

$$(t, v) \mapsto e^{t \cdot \Delta}(v), \tag{18}$$

where the right hand side here means the sum of the following "series" (which has only finitely many non-vanishing terms, since t is assumed nilpotent):

$$v + t\Delta(v) + \frac{t^2}{2!}\Delta^2(v) + \frac{t^3}{3!}\Delta^3(v) + \ldots$$

Here of course $\Delta^2(v)$ means $\Delta(\Delta(v))$, etc.

Proof. We have to prove that $\dot{F}(t) = \Delta(F(t))$. We calculate the left hand side by differentiating the series term by term (there are only finitely many non-zero terms):

$$\Delta(v) + \frac{2t}{2!} \cdot \Delta^2(v) + \frac{3t^2}{3!}\Delta^3(v) + ... = \Delta(v + t \cdot \Delta(v) + \frac{t^2}{2!} \cdot \Delta^2(v) + ...)$$

using linearity of Δ. But this is just Δ applied to $F(t)$.   □

There is an analogous result (which we utilized in [5]) for second order differential equations of the form $\ddot{F}(t) = \Delta(F(t))$ (with Δ linear); the proof is similar and we omit it:

Proposition 2.10. *The formal solution of this second order differential equation $\ddot{F} = \Delta F$, with initial position v and initial speed w, is given by*

$$F(t) = v + t \cdot w + \frac{t^2}{2!}\Delta(v) + \frac{t^3}{3!}\Delta(w) + \frac{t^4}{4!}\Delta^2(v) + \frac{t^5}{5!}\Delta^2(w) + ...$$

3   Exponential of vector fields

We shall describe the exponential of two D-actions. We do this when the action in the exponent is *invertible*. An action $X : D \times M \to M$ is called *invertible*, if for each $d \in D$, $X(d, -) : M \to M$ is invertible. In this case, the exponential $(N, Y)^{(M,X)}$ may be described as N^M equipped with the following action by D: an element $d \in D$ acts on $\beta : M \to N$ by "conjugation":

$$\beta \mapsto Y_d \circ \beta \circ (X_d)^{-1},$$

where Y_d denotes $Y(d, -) : N \to N$, and similarly for X_d.

It is easy to check that if both actions are pointed, so is the above exponential. This means that exponentials in the category of pointed objects are formed by taking exponentials in the topos of actions.

In this section, we show that solutions of an exponential vector field may be obtained by conjugating solutions of the vector fields that make up the exponential. Furthermore, this method of conjugation is equivalent (under some conditions) to the method of change of variables, widely used to solve differential equations.

Theorem 3.1. *Assume that (M, X) and (N, Y) are vector fields having (complete) solutions $F : \tilde{R} \times M \to M$ and $G : \tilde{R} \times N \to N$, respectively, and assume that all F_t are invertible. Then a (complete) solution $H : \tilde{R} \times M \to M$ of the exponential $(N, Y)^{(M,X)}$ is obtained as the map*

$$H : \tilde{R} \times N^M \to N^M$$

given by conjugation: $H_t(\beta) = G_t \circ \beta \circ F_t^{-1}$.

Proof. This is purely formal. For $\beta \in N^M$, we have

$$(Y^X)_d(H_t(\beta)) = Y_d \circ H_t(\beta) \circ X_d^{-1}$$
$$= Y_d \circ G_t \circ \beta \circ F_t^{-1} \circ X_d^{-1}$$
$$= G_{d+t} \circ \beta \circ F_{d+t}^{-1}$$
$$= H_{d+t}(\beta),$$

where in the third step we used (5) for F as well as for G, together with invertibility of F_s for all s and invertibility of X_d.

A similar argument gives that if each of (3)-(11) holds for both F and G, then the corresponding property holds for H.

In the applications we have considered, the invertibility of the F_t will be secured by subtraction on $\tilde{R}$, with $F_t^{-1} = F_{-t}$. $\square$

Using directional derivatives, we can give a more familiar expression to the vector field (1ODE) Y^X considered above on the object N^M, when the base N is a microlinear Euclidean R-module V (hence also V^M is Euclidean), and the exponent M is microlinear. In fact, letting η be the principal part of the vector field Y on $N = V$, we have, for $u \in V^M$, $m \in M$, $d \in D$ (recall that $(X_d)^{-1} = X_{-d}$)

$$(Y^X)_d(u)(m) = Y_d \circ u \circ X_{-d}(m)$$
$$= u((X_{-d}(m)) + d \cdot \eta(u(X_{-d}(m)))$$
$$= u(m) - d \cdot D_X(u)(m) + d \cdot \eta(u(m))$$
$$= u(m) + d \cdot [-D_X(u)(m) + \eta(u(m))]$$

(at the third equality sign, a cancellation of $d \cdot d$ took place in the last term).

In other words, the principal part of Y^X is $\theta : V^M \to V^M$ given by

$$\theta(u)(m) = \eta(u(m)) - D_X(u)(m). \tag{19}$$

In these terms, the differential equation given by the vector field Y^X may be rewritten (leaving out the m, and modulo some obvious abuse of notation) as

$$\dot{u} = \eta(u) - D_X(u),$$

or

$$\frac{\partial u}{\partial t} + D_X(u) = \eta(u). \tag{20}$$

This is a PDE of first order. Thus, the exponential of two 1ODE's is a 1PDE.

Our Theorem then translates into the following result, formulated entirely in standard terms:

Theorem 3.2. *The complete solution of the PDE (20) is given by*

$$u(t, v)(m) = G(t, v(F(-t, m))),$$

where F and G are complete solutions of $\dot{x} = X(x)$ and $\dot{y} = \eta(y)$, respectively, where v is an any function (initial value) $M \to V$.

In the particular case that $M = \tilde{R}$ and $N = R$, we can give this PDE a more familiar presentation, using (12); the equation (20) becomes the following PDE for a function $u(t, x)$

$$\frac{\partial u}{\partial t} + \xi(x)\frac{\partial u}{\partial x} = \eta(u).$$

The Theorem provides the following solution $u(t, x)$ of it, (with initial condition $u(0, -)$ an arbitrary initial value function $v = v(x)$):

$$u(t, x) = G(t, v(F(-t, x)),$$

where F and G are solutions of $\dot{x} = \xi(x)$, $\dot{y} = \eta(y)$, i.e. satisfy $\frac{\partial}{\partial t}F(t, x) = \xi(F(t, x))$ and $\frac{\partial}{\partial t}G(t, y) = \eta(G(t, y))$.

This can also be verified by plain calculus, using (17).

We shall finish by giving a reformulation (and alternative proof) of the main Theorem; it may be considered as a "change of variable" method, changing variable in the exponent space M. We need some preliminaries.

For any object N, let us consider its "zero vector field" Z , i.e., Z_d is the identity map on N, for all d. For a vector field X on an object M, we then also have the "vertical" vector field $Z \times X$ on $N \times M$, given by $(Z \times X)(d, (n, m)) = (n, X(d, m))$.

If we have a complete solution $F : \tilde{R} \times M \to M$ of a vector field X on M, we may consider the map $\overline{F} : \tilde{R} \times M \to \tilde{R} \times M$ given by $(t, m) \mapsto (t, F(t, m))$

Proposition 3.3. *The map $\overline{F}$ thus described is an automorphism of the vector field $Z \times X$ on $\tilde{R} \times M$.*

Proof. By a straightforward diagram chase, one sees that this is a restatement of condition (4). □

The following is a form of the chain rule. We consider a vector field X on M, with solution $F : \tilde{R} \times M \to M$. Let $U : \tilde{R} \times M \to V$ be any function with values in a Euclidean R-module.

Proposition 3.4. *Under these circumstances, we have*

$$\frac{\partial}{\partial t} U(t, F_t(m)) = \frac{\partial U}{\partial t}(t, F_t(m)) + (D_{Z \times X} U)(t, F_t(m))$$

for all $t \in \tilde{R}$, $m \in M$.

Proof. Since F is a solution of X, $F_{d+t} = X_d \circ F_t$, and so for any $t, t' \in \tilde{R}$, $(Z \times X)_d(t', F_t(m)) = (t', F_{d+t}(m))$. Therefore, by definition of directional derivative,

$$U(t', F_{d+t}(m)) = U(t', F_t(m)) + d \cdot (D_{Z \times X} U)(t', F_t(m)).$$

Putting $t' = d + t$, we thus have

$$U(d + t, F_{d+t}(m)) = U(d + t, F_t(m)) + d \cdot (D_{Z \times X} U)(d + t, F_t(m))$$

$$= U(d + t, F_t(m)) + d \cdot (D_{Z \times X} U)(t, F_t(m))$$

by a standard cancellation of two d's, after Taylor expansion. Expanding the first term, we may continue:

$$= U(t, F_t(m)) + d \cdot \frac{\partial U}{\partial t}(t, F_t(m)) + d \cdot (D_{Z \times X} U)(t, F_t(m)).$$

On the other hand,

$$U(d + t, F_{d+t}(m)) = U(t, F_t(m)) + d \cdot \frac{\partial}{\partial t}(U(t, F_t(m)));$$

comparing these two expressions gives the result. $\qquad \square$

The method of change of variables has been used extensively to solve differential equations. We shall prove that our method for solving the exponential differential equation Y^X, where X is an integrable vector field on M, Y an integrable vector field on a Euclidean R-module, and where $\tilde{R}$ is symmetric with respect to the origin (if $t \in \tilde{R}$, then $-t \in \tilde{R}$), may be seen as an application of the method of change of variables. We let $\eta : V \to V$ denote the principal part of Y, as before. Let $F : \tilde{R} \times M \to M$ be the assumed solution of X. From $F_{-t} = F_t^{-1}$ follows that the map $\overline{F}$ considered in Proposition 3.3 is invertible, with $\overline{F}^{-1}(t, m) = (t, F(-t, m))$. The map $\overline{F}^{-1}(t, m)$ represents the change of variables $\tau = t$, $\mu = F(-t, m)$.

Theorem 3.5. *("Change of variables"). If $u : \tilde{R} \times M \to V$ is a particular solution of Y^X, or, equivalently by (20), of*

$$\frac{\partial u}{\partial t} + D_X(u) = \eta(u), \qquad (21)$$

then the unique map $U : \tilde{R} \times M \to V$ given as the composite

$$\tilde{R} \times M \xrightarrow{\overline{F}} \tilde{R} \times M \xrightarrow{u} V$$

is a particular solution of Y^Z, or, equivalently, of

$$\frac{\partial U}{\partial t} = \eta(U),\qquad(22)$$

and vice versa.

Proof. Since $u(t, m) = U(t, F_{-t}(m))$, we have

$$\frac{\partial u}{\partial t}(t, m) = \frac{\partial}{\partial t}U(t, F_{-t}(m)) = \frac{\partial U}{\partial t}(t, F_{-t}(m)) - (D_{Z \times X}U)(t, F_{-t}(m)),$$

by the chain rule, Proposition 3.4. Writing μ for $F(-t, m)$ (so $\overline{F}(t, \mu) = (t, m)$), we thus have

$$\frac{\partial u}{\partial t}(t, m) = \frac{\partial U}{\partial t}(t, \mu) - (D_{Z \times X}U)(t, \mu).$$

Now $\overline{F}$ is an automorphism of the vector field $Z \times X$, by Proposition 3.3. Also $U = u \circ \overline{F}$. From Proposition 2.3, we therefore get that the second term on the right hand side equals $-(D_{Z \times X}u)(t, m)$, so

$$\frac{\partial u}{\partial t}(t, m) = \frac{\partial U}{\partial t}(t, \mu) - (D_{Z \times X}u)(t, m).$$

On the other hand u is assumed to be a solution of Y^X, which means that

$$\frac{\partial u}{\partial t}(t, m) = \eta(u(t, m)) - (D_{Z \times X}u)(t, m).$$

Comparing these two equations gives us the first equality sign in

$$\frac{\partial U}{\partial t}(t, \mu) = \eta(u(t, m)) = \eta(U(t, \mu)),$$

and since this holds for all μ, $\frac{\partial U}{\partial t} = \eta(U)$, as claimed. (The vice versa part follows because $\overline{F}$ is invertible.) $\square$

Note that when the exponent vector field is the zero field Z, particular solutions $h(t)$ of Y^Z have the property that for each fixed $m \in M$, $h(t)(m)$ is a particular solution of Y; therefore, if Y has the uniqueness property for particular solutions, then so does Y^Z.

Corollary 3.6 (Uniqueness of solutions of the exponential). *Let X be a vector field on M and let Y be a vector field on an Euclidean module V. Assume that $F : \tilde{R} \times M \to M$ is a complete solution of X. If Y has the uniqueness property for particular solutions, then so does Y^X.*

Proof. We let η be the principal part of Y. Let $u, w : \tilde{R} \times M \to V$ be (exponential adjoints of) particular solutions of Y^X such that $u(0, m) = v(0, m)$. For each $m \in M$, define $U(t) = u(t, F(t, m))$ and $W(t) = w(t, F(t, m))$. By the change of variables theorem, both U and W satisfy $\dot{y} = \eta(y)$, i.e. they are particular solutions of Y with the same initial value. By the uniqueness of particular solutions of Y, $U = W$. But m is arbitrary,

i.e., $u(t, F(t, m)) = w(t, F(t, m))$ for every t and every m. Since $F(t, -)$ is bijective, this shows that $u = v$.   □

Corollary 3.7 (Uniqueness of the solution by conjugation). *Assume the hypothesis of the previous corollary. If Y has a complete solution G, then the solution of Y^X obtained by conjugation from F and G is the only complete solution of this exponential vector field.*

Some examples.   The first two are immediate applications of the formula derived in Theorem 3.2. The third is concerned with the tangent bundle $T(M)$ seen as an exponential object M^D.

Example 3.8 ('Simple transport equation').

$$\partial u/\partial t + \partial u/\partial x = 0$$

Here, $\xi(x) = 1$ and $\eta(y) = 0$. The complete solution of $\dot{x} = 1$ is clearly $F(x, t) = x + t$, which is globally defined. Clearly, the complete solution of $\dot{y} = 0$ is $G(t, x) = x$ (also globally defined). Hence

$$H(t, x) = G(t, v(F(-t, x))) = v(x - t)$$

is the only (globally defined) complete solution of the PDE (with initial value the function v).

Example 3.9.
$$\partial u/\partial t + x\partial u/\partial x = u$$

In this case, $\xi(x) = x$ and $\eta(y) = y$ and their complete solutions are the same, namely $F(t, x) = G(t, x) = xe^t$ (globally defined). Therefore,

$$H(t, x) = G(t, v(F(-t, x))) = v(xe^{-t})e^t$$

is the (globally defined) complete solution of the PDE, (with initial value the function v).

Example 3.10. Let D be the set of elements of square zero in R, as usual. It carries a vector field, namely the map $e : D \times D \to D$ given by $(d, \delta) \mapsto (1 + d) \cdot \delta$. It is easy to see that this vector field is integrable, with complete solution $E : R \times D \to D$ given by $(t, \delta) \mapsto e^t \cdot \delta$. Now consider the tangent vector bundle M^D on M. The zero vector field Z on M is certainly integrable, and so we have by the theorem a complete integral for the vector field Z^e on the tangent bundle. We describe the integral explicitly (this then also describes the vector field, by restriction): it is the map $R \times M^D \to M^D$ given by $(t, \beta) \mapsto [d \mapsto \beta(e^{-t} \cdot d)]$.— The vector field on M^D obtained this way is, except for the sign, the *Liouville vector field*, cf. [2], IX.2.

References

[1] M. Bunge and E. Dubuc: "Local Concepts in Synthetic Differential Geometry and Germ Representability", In: D.W. Kueker, E.G.K. Lopez-Escobar and C.H. Smith (Eds.): *Mathematical Logic and Theoretical Computer Science*, Marcel Dekker Inc., 1987, pp. 93–159.

[2] C. Godbillon: *Géométrie Différentielle et Mécanique Analytique*, Hermann, Paris, 1969.

[3] A. Kock: *Synthetic Differential Geometry*, Cambridge University Press, 1981.

[4] A. Kock and G.E. Reyes: "Aspects of Fractional Exponents", *Theor. Appl. Categories*, Vol. 5(10), (1999).

[5] A. Kock and G.E. Reyes: "Some calculus with extensive quantities; wave equation", *Theor. Appl. Categories*, Vol. 11(14), (2003).

[6] A. Kumpera and D. Spencer: *Lie equations*, Vol. 1, Ann. of Math Studies, Vol. 73, Princeton University Press, 1972.

[7] R. Lavendhomme: *Basic Concepts Of Synthetic Differential Geometry*, Kluwer Academic Publishers, 1996.

[8] F.W. Lawvere: "Categorical Dynamics", In: A. Kock (Ed.): *Topos Theoretic Methods in Geometry*, Series 30, Aarhus Various Publ., (1979).

[9] B. Malgrange: "Equations de Lie", *I. J. Diff. Geom.*, Vol. 6, (1972), pp. 503–522.

[10] I. Moerdijk and G.E. Reyes: *Models for Smooth Infinitesimal Analysis*, Springer-Verlag, 1991.