Some differential equations in SDG

Anders Kock and Gonzalo E. Reyes

March 25, 2018

Introduction

We intend to comment on some of those aspects of the theory of differential equations which we think are clarified (for us, at least) by means of the synthetic method. By this, we understand that the objects under consideration are seen as objects in one sufficiently rich category (model for SDG), allowing us, for instance, to work with nilpotent numbers, say $d \in R$ with $d^2 = 0$; but the setting should also permit the formation of function spaces, so that some of the methods of functional analysis, become available, in particular, the theory of distributions.

The specific topics we treat are generalities on vector fields and the solutions of corresponding first- and second-order ordinary differential equations; and also some partial differential equations, which can be seen in this light, the wave- and heat-equation on some simple spaces, like the line R. For these equations, distribution theory is not just a tool, but is rather the essence of the matter, since what develops through time, is a distribution (of heat, say), which, as stressed by Lawvere, is an *extensive* quantity, and as such behaves covariantly, unlike density *functions* (which behave contravariantly); and the distributions may have no density function, in particular in the setting of model for SDG where all functions are *smooth*.

When we consider these partial differential equations, we shall follow an old practice and sometimes denote derivative d/dt with respect to "time" by a dot, $\dot{f}$, whereas differential operators with respect to space variables are denoted $\partial f/\partial x$, f', $\Delta(f)$, etc.

We want to thank Henrik Stetkær for useful conversations on the topic of distributions.

1 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 topos; we shall in particular be interested in the exponent formation in this topos, 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 exponent $(N, Y)^{(M,X)}$ 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.

In the applications below, D is the usual set of square zero elements in 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 $m \in M$, or equivalently, $X_0 : M \to M$ is the identity map on M. A pointed action, in this situation, is the same thing as a vector field on M, cf. [12].

In the above situation, if X and Y are pointed actions, then so is the exponent described. The pointed actions likewise form a topos, and the exponent described is then also the exponent in the category of pointed actions; cf. [9].

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 $(\tilde{R}, \Delta)$: $\tilde{R}$ is an "infinitesimally open subset" of R, i.e., whenever $x \in \tilde{R}$ then $x + d \in \tilde{R}$ for every $d \in D$. The main examples of such subsets are R itself, the non-negative numbers $R_{\geq 0}$, open intervals, and the set D_∞ of all nilpotent elements of the number line. The action Δ is the vector field $\partial/\partial x$, meaning the map $D \times \tilde{R} \to \tilde{R}$ given by $(d, t) \mapsto d + t$. (So it is not to be confused with the Laplace operatot Δ, to be considered later.) The main property to be assumed is that the individual Δ_d's are homomorphisms of D-actions (which is a commutativity requirement); the structure of $\tilde{R}$ could probably be derived from this, but we shall be content with *assuming* that $\tilde{R}$ is an additively written monoid, and that $D \subseteq \tilde{R}$ (with the 0 of D also being the zero of the monoid).

First, if (M, X) is a set with an action, a homomorphism $f : (\tilde{R}, \Delta) \to (M, X)$ 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 in $f(0)$. 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 $m \in M$; we shall consider and compare the following further conditions (universally quantified over all $d \in D$, $t, s \in \tilde{R}$, $m \in M$):

$$F(\Delta(d, t), m)) = X(d, F(t, m)); \tag{1}$$

this is the main one, the two following conditions are included for systematic reasons only:

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

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

Finally, one may consider the following equation

$$F(t + s, m) = F(t, F(s, m)). \tag{4}$$

Writing X_d for the map $X(d, -) : M \to M$, and similarly for F, condition (1) may be rewritten as

$$F_{\Delta(d,t)} = X_d \circ F_t$$

The others may be rewritten in a similar way. For instance (4) may be rewritten as

$$F_{t+s} = F_t \circ F_s$$

Equation (1) expresses that, for each fixed $m \in M$, the map $F(-, m) : \tilde{R} \to M$ is a homomorphism (and thus, by virtue of $F(0, m) = m$, a "solution with initial value m"). Writing the action of D in terms of the symbol $\cdot$, we may write it $F(d \cdot t, m) = d \cdot F(t, m)$. Equation (2) expresses a certain bi-homogeneity condition of F, $F(d \cdot t, m) = F(t, d \cdot m)$; (3) says that for fixed $t \in R$, $F(t, -) : M \to M$ is an endomorphism of D-actions, $F(t, d \cdot m) =$

$d \cdot F(t,m)$. Finally (4) is the usual condition for action af a monoid on a set M. Clearly, it implies all the others.

Let X be a vector field on M, thought of as a first-order differential equation. We say that the map $F : \tilde{R} \times M \to M$ is a *complete solution* or simply a *solution* if $F_d = X_d$ and F satisfies (1). A solution in this sense does not satisfy the other conditions (2)-(4), but it does, provided that M 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 M:
If X is a D-action on M, and $f, g : \tilde{R} \to M$ are homomorphisms of actions, with $f(0) = g(0)$, then $f = g$.

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

Proposition 1 *Let X be a vector field on M and assume that M satisfies the uniqueness axiom. Then any solution $F : \tilde{R} \times M \to M$ of the differential equation X satisfies properties (2) and (3). Furthermore, if $\tilde{R}$ is a monoid (under $+$) then F also satisfies (4).*

Proof. Since the proofs are quite similar, we shall do only (4). Fix $m \in M$ and $s \in \tilde{R}$ and define the couple of functions $f,\ g : \tilde{R} \to M$ by the formulas

$$\begin{cases} f(t) = F_{t+s}(m) \\ g(t) = F_t \circ F_s(m) \end{cases}$$

We have to check that f and g are homomorphisms of D-actions, i.e., they satisfy (1). Let us do this for the first

$$\begin{aligned} f(t+d) &= F_{(t+d)+s}(m) \\ &= F_{d+(t+s)}(m) \\ &= F_d \circ F_{t+s}(m) \\ &= X_d \circ f(t). \end{aligned}$$

The proof that g is a homomorphism is similar. Thus, the equality of the two expressions follows from the uniqueness property assumed for M.

Recall that a vector field X on M is called *integrable* if there exists a solution $F : \tilde{R} \times M \to M$. If we assume the uniqueness property, the equation (4) holds; if further the commutative monoid structure $+$ on $\tilde{R}$ actually is a group structure, then (4) implies that the action is invertible, with X_{-d} as X_d^{-1} (in fact $F_{-d} = F_d^{-1}$). Of course, both the uniqueness property and the question whether or not the vector field X is integrable, depends 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 *The uniqueness property holds for any 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 formal solution.*

Proof. We need to recall some infinitesimal objects from the literature on SDG, cf. e.g. [11]. 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,

Lemma 3 *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. [11] 3.2.2, namely $[X, X] = 0$.

Likewise

Lemma 4 *If X is a vector field on a microlinear object and $d_1, \ldots, d_n \in D$ are such that $d_1 + \ldots + d_n = 0$, then*

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

(= 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

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

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,

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

since p is symmetric. But

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

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 $\sigma_1(d_1, \ldots, d_n) = d_1 + \ldots + d_n$: and each σ_i, when restricted to D^n, is a function of σ_1, since $d_1^2 = 0$; e.g.

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

Now consider, for fixed $m \in M$, the map $p : D^n \to M$ given by $(d_1, \ldots, d_n) \mapsto X_{d_1} \circ \ldots \circ X_{d_n}(m)$. By Lemma 3, 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

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

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

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 4), via the map $(d_1, \ldots, d_n) \mapsto d_1 + \ldots + d_n$, so we are forced to define $F_t = X_{d_1} \circ \ldots X_{d_n}$ if F is to extend X and to satisfy (4). The fact that this is well defined independently of the choice of n and the choice of $d_1, \ldots, d_n$ that add up to t follows from Lemma 4.

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 5 *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*

$$(t, v) \mapsto e^{t \cdot \Delta}(v), \tag{5}$$

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

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

2 Exponent vector fields

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

Theorem 7 *Assume that* (M, X) *and* (N, Y) *are vector fields having 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 solution* $H : \tilde{R} \times M \to M$ *of the exponent* $(N, Y)^{(M,X)}$ *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

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

where in the third step we used the equation (1) for G and F, in the form

$$G_{d+t} = Y_d \circ G_t, \;\; \text{respectively } F_{d+t} = X_d \circ F_t,$$

together with invertibility of F_s for all s and invertibility of X_d.

A similar argument gives that if each of (2)-(4) holds for both F and G, then the corresponding property holds for H.

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

Recall that an R-module V is called *Euclidean* if the canonical map $\alpha : V \times V \to V^D$ given by $\alpha(u, v)(d) = u + d \cdot v$ is invertible; the composite of α^{-1} with projection to the second factor, $V^D \to V \times V \to V$ is called *principal part* formation. If $X : V \to V^D$ is a vector field on a Euclidean module V, we may compose it with principal part formation to get a (not necessarily linear) map $\xi : V \to V$, called the principal part of the vector field X; it is thus characterized by the formula

$$X(v)(d) = v + d \cdot \xi(v).$$

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 principal part formation. Using function theoretic notation, $D_X(\beta)$ is characterized by validity of the equation

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

for all $d \in D$, $m \in M$.

When M itself is a Euclidean R module, and X has principal part ξ, we usually write $D_\xi(\beta)$ instead of $D_X(\beta)$.

Proposition 8 *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:

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

on the other hand

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

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

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 $N \times M$.

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 9 *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 (3).

We now consider solutions $F : \tilde{R} \times V \to V$ for such vector fields, so equation (1) holds: $X_d \circ F_t = F_{t+d}$. In terms of principal parts, this equation may be rewritten as

$$\dot{F}_t(v) = \xi(F_t(v)).$$

Similarly, equation (2) may be written as

$$\dot{F}_t = D_\xi(F_t). \tag{6}$$

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, and the exponent M is mocrolinear. 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 $(X_d)^{-1} = X_{-d}$)

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

(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 $\theta : M \to V$ given by

$$\theta(m) = \eta(u(m)) - D_X(u)(m).$$

Recalling that the 1ODE corresponding to a vector field X on a Euclidean R-module V may be written as $\dot{x} = \xi(x)$ where ξ is the principal part of X. In these terms, the above equation may be rewritten (leaving out the m, and modulo some obvious abuse of notation) as

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

or still, recalling that $(\dot{-})$ is "derivative with respect to time",

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

This is a PDE of first order "in time".

The following may be seen as a generalization of (6), and 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 10 *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, $F_{t+d} = X_d \circ F_t$, and so for any $t, t' \in \tilde{R}$ $(Z \times X)_d(t', F_t(m)) = (t', F_{t+d}(m)$. Therefore, by definition of directional derivative,

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

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

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

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,

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

comparing these two expressions gives the result.

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, and let $\overline{F} : \tilde{R} \times M \to \tilde{R} \times M$ be the map

$$\overline{F}(t, m) = (t, F(-t, m))$$

Then $\overline{F}$ (which represents the change of variables $\tau = t$, $\mu = F(-t, m)$) is invertible.

Theorem 11 *("Change of variables"). If $u : \tilde{R} \times M \to V$ is a particular solution of Y^X, or, equivalently, of*

$$\frac{\partial u}{\partial t} + D_X(u) = \eta(u), \tag{7}$$

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

$$\tilde{R} \times M \stackrel{(\overline{F})^{-1}}{\to} \tilde{R} \times M \stackrel{u}{\to} V$$

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

$$\frac{\partial U}{\partial t} = \eta(U), \tag{8}$$

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 10. On the other hand, $\overline{F}$ is an automorphism of the vector field $Z \times X$, by Proposition 9, and so, by construction of $\overline{F}$ and Proposition 8,

$$D_{Z\times X}(u) = D_{Z\times X}(U \circ \overline{F}) = (D_{Z\times X}U) \circ \overline{F}.$$

Therefore,

$$0 = \frac{\partial u}{\partial t} + D_{Z\times X}(u) - g(u)$$

$$= \frac{\partial U}{\partial t}(t,\mu) - D_{Z\times X}(U)(t,\mu) + D_{Z\times X}(U)(t,\mu) - \eta(U(t,\mu)),$$

where $\mu = F_{-t}(m)$, i.e., U is solution of

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

proving the theorem (the vice versa part follows because $\overline{F}$ is invertible).

Example. 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. [5], IX.2.

3 Generalities on distributions

We want to apply parts of the general theory of ordinary differential equations to some of the basic equations of mathematical physics, the wave- and heat-equations. This takes us by necessity to the realm of distributions. Not primarily as a technique, but because of the nature of these equations: they model evolution through time of (say) a *heat distribution*. A heat distribution is an *extensive* quantity, and does not necessarily have a density *function*, which is an *intensive* quantity; the most important of all distributions, the point distributions (or Dirac distributions), for instance, do not. For the case of the heat equation, it is well known that the evolution through time of any distribution "instantaneously" (i.e., after any *positive* lapse of time, $t > 0$) leads to distributions that do have smooth density functions. But in SDG, we are interested also in what happens after a nilpotent lapse of time. In more computational terms, we are interested in the Taylor expansion of the solutions of evolution equations. For this, it is necessary to stay within one vector space, that of distributions.

The vector space of "distributions of compact support" on any object M can be introduced purely synthetically (see [15] p. 393, or [14] p. 94) as the R-linear dual of the vector space R^M (which internally represents the vector space of smooth functions on M). What follows could, to a certain extent (in particular for the wave equation), be treated purely synthetically.

Presently, we shall only be interested in distributions on R, R^2 , and R^3, so for the presentation, we have chosen to assume that we are working in a sufficiently good "well-adapted" model $\mathbf{E}$ of SDG, containing the category of smooth manifolds as a full subcategory. In such models, for any given manifold M, we could define the linear subspace $\mathcal{D}(M)$ of R^M consisting of functions with compact support, (the "test functions"). Then the vector space of distributions on M, $\mathcal{D}'(M)$, is taken to be the R-linear dual of $\mathcal{D}(M)$.

One could take an alternative, slightly more concrete, approach: namely, take a model $\mathbf{E}$ of SDG which contains the category of smooth manifolds as above, but which also contains the category of Convenient Vector Spaces [4] and the smooth maps between them as a full subcategory. The embedding is to preserve the cartesian closed structure. Such models do exist: we provided in [7], [8] such an embedding of Convenient Vector Spaces into the "Cahiers" topos of Dubuc [2]. Note that the usual topological (Fréchet) vector spaces of smooth functions, test functions, distributions, etc. on a smooth manifold M have canonical structure of Convenient Vector Spaces. In such a model, we can construct internal functions, say curves $f : R \to \mathcal{D}'(M)$, by constructing, externally, a function by an "excluded middle" recipe of the form

$$f(t) = .. \text{ if } t \neq 0 \ ; \ f(t) = .. \text{ if } t = 0,$$

and then proving smoothness of f by a usual limit argument.

We have to resort to this kind of "external" constructions only for the heat equation, and there our embedding from [7], [8] is not quite good enough, since it does not take manifolds with boundary into account; for the heat equation, one constructs externally an “evolution” map

$$R_{\geq 0} \to \mathcal{D}'(R)$$

by an excluded middle recipe.

So, for the justification of our treatment of the heat equation, we need an extension (hopefully forthcoming) of our work [7], [8], i.e., we need to construct a Cahiers-like topos that includes also manifolds *with* boundary, and then to construct an embedding of Convenient Vector Spaces into that “extended” Cahiers Topos. (Maybe even the Cahiers Topos itself will be good enough.)

For what follows about wave equation, the Cahiers Topos, and the embedding of Convenient Vector Spaces into it, is sufficient; in fact, for these equations, a purely synthetic treatment alluded to will be sufficient, since the distributions considered there are all of compact support.

As stressed by Lawvere in [13], distributions should *not* be thought of as generalized functions: functions are intensive quantities, and transform *contravariantly*; distributions are extensive quantities and transform *covariantly*. For functions, this is the fact that the “space” of functions on M, R^M is contravariant in M, by elementary cartesian-closed category theory. Similarly, the “space” of distributions of compact support on M is a subspace of R^{R^M} (carved out by the R linearity condition), and so for similar elementary reasons is covariant in M. We shall write $\mathcal{D}'_c(M)$ for this subspace. The space of functions of compact support on M is only functorial with respect to *proper* smooth maps, (counterimages of compact set required to be compact), and so similarly, the space $\mathcal{D}'(M)$ of all distributions on M is covariant functorial only w.r.to proper maps. The formula for covariant functorality looks the same for $\mathcal{D}'$ and $\mathcal{D}'_c$; let us make it explicit for the $\mathcal{D}'$ case. Let $f : M \to N$ be a proper map. The map $\mathcal{D}'(f) : \mathcal{D}'(M) \to \mathcal{D}'(N)$ is described by declaring

$$< \mathcal{D}'(f)(\mu), \phi >=< \mu, \phi \circ f >, \tag{9}$$

where μ is a distribution on M, and ϕ is a test function on N, (so $\phi \circ f$ is a test function on M, by properness of f). The brackets denote evaluation of distributions on test functions.

We shall also write just $f(\mu)$ instead of $\mathcal{D}'(f)(\mu)$.

Recall that a distribution μ on M may be mulitplied by any function $g : M \to R$, by the recipe

$$< g \cdot \mu, \phi >=< \mu, g \cdot \phi >, \tag{10}$$

observing that $g \cdot \phi$ is a test function (has compact support) if ϕ is.

If X is a vector field on M, one defines the directional derivative $D_X(\mu)$ of a distribution μ on M by the formula

$$< D_X(\mu), \phi >= - < \mu, D_X(\phi) > . \tag{11}$$

This in particular applies to the vector field $\partial/\partial x$ on R, and reads here $< \mu', \phi >= - < \mu, \phi' >$ (ϕ' denoting the ordinary derivative of the function ϕ). One has the following Leibniz rule:

$$D_X(f \cdot \mu) = D_X(f) \cdot \mu + f \cdot D_X(f) \tag{12}$$

for any distribution μ and function f on M. This is an elementary consequence of the Leibniz rule for directional derivatives D_X of functions on M.

Remark. The equation (11) becomes a theorem, rather than a definition, if one takes the following line of reasoning: let F be a covariant functor from microlinear spaces (and invertible maps between them) to Euclidean vector spaces. Then one may define the *Lie derivative* along X, $L_X(\alpha)$, as a map $F(M) \to F(M)$. For the functor $F = \mathcal{D}'$, L_X becomes the D_X described. We shall not pursue this line further here.

Applying D_X twice leads to

$$< D_X(D_X(\mu)), \phi >=< \mu, D_X(D_X(\phi)) >$$

In particular, for μ a distribution on R^n

$$< \partial^2/\partial {x_i}^2(\mu), \phi >=< \mu, \partial^2/\partial {x_i}^2(\phi) >$$

and therefore for the Laplace operator $\Delta = \sum \partial^2/\partial^2 x_i =$ div $\circ$ grad, we put

$$< \Delta(\mu), \phi >=< \mu, \Delta(\phi) > . \tag{13}$$

The following Proposition is an application of the covariant functorality of the functor $\mathcal{D}_c$, which will be used in connection with the wave equation in dimension 2. We consider the (orthogonal) projection $p : R^3 \to R^2$ onto the xy-plane. (It is not a proper map, so functorality only works for compactly supported distributions.)

Proposition 12 *For any distribution S (of compact support) on R^3,*

$$p(\Delta(S)) = \Delta(p(S)).$$

(The same result holds for any orthogonal projection p of R^n onto any linear subspace; the proof is virtually the same, if one uses invariance of Δ under orthogonal transformations.)

Proof. Let ψ be any test function on R^2. Then

$$< p(\Delta(S)), \psi >=< \Delta(S), \psi \circ p >=< S, \Delta(\psi \circ p) > \tag{14}$$

But, with $\psi = \psi(x, y)$, $\psi \circ p$ is just ψ, considered as a function of x, y, z which happens not to depend on z; so

$$\Delta(\psi \circ p) = \frac{\partial \psi}{\partial x} + \frac{\partial \psi}{\partial y} + \frac{\partial \psi}{\partial z};$$

the last term vanishes because ψ does not depend on z, so the equation continues

$$= \frac{\partial \psi}{\partial x} + \frac{\partial \psi}{\partial y} = (\Delta(\psi)) \circ p.$$

So the right hand expression in (14) may be rewritten as

$$=< S, \Delta(\psi) \circ p >=< p(S), \Delta(\psi) >=< \Delta(p(S)), \psi >,$$

from which the result follows.

3.1 Spheres and balls as distributions

For $a, b \in R$, we let $[a, b]$ denote the distribution $f \mapsto \int_a^b f(x)\ dx$. Such distributions on the line, we of course call *intervals*; the *length* of an interval $[a, b]$ is defined to be $b - a$. Note that the interval $[a, b]$ as a distribution is not quite the same as the order theoretic interval, i.e., the subset of R consisting of x with $a \leq x \leq b$. For instance, the order theoretic interval from 0 to 0 contains all nilpotent elements, whereas the distribution $[0, 0]$ is the zero distribution. The distribution theoretic interval $[a, b]$ contains more information about a and b than does the order theoretic one. We consider the question to which extent $[a, b]$ determines the endpoints. The answer is contained in

Proposition 13 *Let $[a_1, b_1]$ and $[a_2, b_2]$ be two intervals in the distribution theoretic sense. They are equal as distributions if and only if they have same length, $b_1 - a_1 = b_2 - a_2$ ($= l$, say), and $l \cdot (a_1 - a_2) = 0$ (this then also implies $l \cdot (b_1 - b_2) = 0$).*

Proof. Assume $[a_1, b_1] = [a_2, b_2]$. The statement about length follows immediately by applying each of these two distributions to the function f which is constant 1. Generally, we have for any function f that

$$\int_{a_1}^{b_1} f(x)\, dx = \int_{a_2}^{b_2} f(x)\, dx$$

$$= \int_{a_1}^{b_1} f(t + a_2 - a_1)\, dt,$$

by making the change of variables $t = x + a_1 - a_2$. Subtracting, we get

$$0 = \int_{a_1}^{b_1} (f(x) - f(x + a_2 - a_1))\, dx.$$

Apply this equation to the function $f(x) = x$, we get

$$0 = \int_{a_1}^{b_1} (x - (x + a_2 - a_1)) = (a_1 - a_2) \cdot (b_1 - a_1) = (a_1 - a_2) \cdot l.$$

Conversely, assume $b_1 - a_1 = b_2 - a_2$ ($= l$, say), and $0 = l \cdot (a_1 - a_2)$. For any function f, we calculate the values of the distribution $[a_1, b_1]$ on f. We have

$$[a_1, b_1](f) = (b_1 - a_1) \int_0^1 f(a_1 + t \cdot (b_1 - a_1))\, dt = l \int_0^1 f(a_1 + t \cdot l)\, dt.$$

Similarly

$$[a_2, b_2](f) = l \int_0^1 f(a_2 + t \cdot l)\, dt.$$

The difference is

$$l \int_0^1 (f(a_1 + t \cdot l) - f(a_2 + t \cdot l))\, dt. \tag{15}$$

By Hadamard's Lemma, $f(a_1 + t \cdot l) - f(a_2 + t \cdot l)$ may be written as $(a_1 - a_2) \cdot g(a_1, a_2, t)$ for some function g, and so the integral (15) can be written as

$$= l \cdot (a_1 - a_2) \int_0^1 g(a_1, a_2, t)\, dt,$$

which vanishes if $l \cdot (a_1 - a_2) = 0$.

The assertions about $b_1 - b_2$ is similar.

Note the following Corollaries: First, if the length $b_1 - a_1$ of an interval $[a_1, b_1]$ is invertible (positive, say), then the endpoints a_1, b_1 are uniquely determined by the distribution $[a_1, b_1]$. Secondly, for any t_1, t_2, we have

$$[-t_1, t_1] = [-t_2, t_2] \text{ implies } t_1 = t_2.$$

In fact, by the Proposition, their lengths must be equal, i.e., $2t_1 = 2t_2$. The distribution $[-t, t]$ will appear below under the name B_t, "the ball of radius t in dimension One".

We shall also consider such "balls" in dimension Two and Three, where, however, t cannot in general be recovered from the distribution, unless t is strictly positive.

We fix a positive integer n. We shall consider the sphere S_t of radius t, and the ball B_t of radius t, for any $t \in R$, as distributions on R^n (of compact support, in fact), in the following sense:

$$< S_t, \psi >= \int_{S_t} \psi(x)dx = t^{n-1} \int_{S_1} \psi(t \cdot u)\, du,$$

$$< B_t, \psi >= \int_{B_t} \psi(x)dx = t^{n} \int_{B_1} \psi(t \cdot u)\, du,$$

where du refers to the surface element of the unit sphere S_1 in the first equation and to the volume element of the unit ball B_1 in the second. The expressions involving $\int_{S_t}$ and $\int_{B_t}$ are to be understood symbolically, unless $t > 0$; if $t > 0$, they make sense literally as integrals over sphere and ball, respectively, of radius t, with dx denoting surface-, resp. volume element. But the expression on the right in both equations make sense for any t, and so the distributions S_t and B_t are defined for all t; in particular, for nilpotent ones.

It is natural to consider also the following distributions S^t and B^t on R^n (likewise of compact support):

$$< S^t, \psi >= \int_{S_1} \psi(t \cdot u)\, du,$$

$$< B^t, \psi >= \int_{B_1} \psi(t \cdot u)\, du.$$

For $t > 0$, they may, modulo factors of the type 4π, be considered as "average over S_t" and "average over B_t", respectively, since S^t differs from S_t by a factor t^{n-1}, which is just the surface area of S_t (modulo the factor of type 4π), and similarly for B^t.

Note that $S^1 = S_1$ and $B^1 = B_1$. And also note that the definition of S^t and B^t can be formulated as

$$S^t = H_t(S_1)\ ,\ B^t = H_t(B_1),$$

where $H_t : R^n \to R^n$ is the homothetic transformation $u \mapsto t \cdot u$, and where we are using the covariant functorality of distributions of compact support.

For low dimensions, we shall describe the distributions S_t, B_t, S^t and B^t explicitly:

Dimension 1

$$< S_t, \psi >= \psi(-t) + \psi(t)$$

$$< B_t, \psi >= \int_{-t}^{t} \psi(s)\, ds$$

$$< S^t, \psi >= \psi(-t) + \psi(t)$$

$$< B^t, \psi >= \int_{-1}^{1} \psi(t \cdot s)\, ds$$

Dimension 2

$$< S_t, \psi >= \int_0^{2\pi} \psi(t\cos\theta, t\sin\theta)\ t\ d\theta$$

$$< B_t, \psi >= \int_0^t \int_0^{2\pi} \psi(s\cos\theta, s\sin\theta)\ s\ d\theta\ ds$$

$$< S^t, \psi >= \int_0^{2\pi} \psi(t\cos\theta, t\sin\theta)\ d\theta$$

$$< B^t, \psi >= \int_0^1 \int_0^{2\pi} \psi(ts\cos\theta, t\ s\sin\theta)\ s\ d\theta\ ds$$

Dimension 3

$$< S_t, \psi >= \int_0^{\pi} \int_0^{2\pi} \psi(t\cos\theta\sin\phi, t\sin\theta\sin\phi, t\cos\phi) t^2 \sin\phi\ d\theta\ d\phi$$

$$< B_t, \psi >= \int_0^t \int_0^\pi \int_0^{2\pi} \psi(s\cos\theta\sin\phi, s\sin\theta\sin\phi, s\cos\phi)\; s^2 \sin\phi \; d\theta \; d\phi \; ds$$

$$< S^t, \psi >= \int_0^\pi \int_0^{2\pi} \psi(t\cos\theta\sin\phi, t\sin\theta\sin\phi, t\cos\phi) \sin\phi \; d\theta \; d\phi$$

$$< B^t, \psi >= \int_0^1 \int_0^\pi \int_0^{2\pi} \psi(ts\cos\theta\sin\phi, ts\sin\theta\sin\phi, ts\cos\phi)\; s^2 \sin\phi \; d\theta \; d\phi \; ds.$$

Notice that these formulas make sense for all t (positive, negative, nilpotent, ...), using the standard convention :$\int_a^b = -\int_b^a$), whereas set-theoretically S_t and B_t (as point sets) only make good sense for $t > 0$.

It is clear from the very definition that $S_t = t^{n-1} S^t$ and $B_t = t^n B^t$ (in any dimension n); but since we are interested also in t's that are not invertible, S_t and S^t cannot be defined in terms of each other.

Note also that $S_0 = B_0 = 0$, whereas S^0 and B^0 are constants times the Dirac distribution at the origin 0. The constants are the "area" of the unit sphere, or the "volume" of the unit ball, in the appropriate dimension. Explicitly,

$$S^0 = 2 \cdot \delta(0),\; 2\pi \cdot \delta(0),\; 4\pi \cdot \delta(0),$$

and

$$B^0 = 2 \cdot \delta(0),\; \pi \cdot \delta(0),\; \frac{4\pi}{3} \cdot \delta(0)$$

in dimensions 1,2, and 3, respectively.

We shall also have occasion to consider the distribution (of compact support) $t \cdot S^t$ on R^3 as well as its projection $p(t \cdot S^t)$ on the xy-plane (using functorality of $\mathcal{D}'_c$ with respect to the projection map $p : R^3 \to R^2$). For $t > 0$ (more generally, for t invertible), we can give an explicit integral expression for it, but note that since S^t and $t \cdot S^t$ are defined for all t, then so is $p(t \cdot S^t)$, whether or not we have such an integral expression. The integral expression (for $t > 0$) goes under the name of *Poisson kernel* for the wave equation in dimension 2 and may be obtained as follows: using the above expression for S^t in dimension 3, we have for a test function ψ that only depends on x, y, but not on z that

$$< t \cdot S^t, \psi >= \int_0^\pi \int_0^{2\pi} \psi(t\cos\theta\sin\phi, t\sin\theta\sin\phi) \cdot t \cdot \sin\phi \; d\theta \; d\phi.$$

We then make the change of variables $\rho = t\sin\phi$, $\phi = \arccos\rho/t$, $d\phi = d\rho/\sqrt{t^2 - \rho^2}$, and then the integral becomes

$$2 \int_0^{\frac{\pi}{2}} \int_0^{2\pi} \frac{\psi(\rho\cos\theta, \rho\sin\theta)\; \rho \; d\theta \; d\rho}{\sqrt{t^2 - \rho^2}},$$

using the explicit form of the ball distribution B_t in dimension 2, we may rewrite the right hand side here as

$$< \frac{2}{\sqrt{t^2 - \rho^2}} \cdot B_t, \psi >,$$

so that we have, for $t > 0$ (or even for t invertible),

$$p(t \cdot S^t) = \frac{2}{\sqrt{t^2 - \rho^2}} \cdot B_t. \tag{16}$$

4 Vector Calculus

The Main Theorem of vector calculus is Stokes' Theorem: $\int_{\partial\gamma} \omega = \int_\gamma d\omega$, for ω an $(n-1)$-form, γ a suitable n-dimensional figure (with appropriate measure on it) and $\partial\gamma$ its geometric boundary. In the synthetic context, the theorem holds at least for any singular cubical chain $\gamma : I^n \to M$ (I^n the n-dimensional coordinate cube), because the theorem may then be reduced to the fundamental theorem of calculus, which is the only way integration enters in the elementary synthetic context; measure theory not being available therein. For an account of Stokes' Theorem in this context, see [14] p.139. Below, we shall apply the result not only for singular *cubes*, but also for singular *boxes*, like the usual $\gamma : [0, 2\pi] \times [0, 1] \to R^2$, parametrizing the unit disk by polar coordinates,

$$\gamma(\theta, r) = (r \cos\theta, r \sin\theta). \tag{17}$$

We shall need from vector calculus the Gauss-Ostrogradsky "Divergence Theorem"

$$\text{flux of } \mathbf{F} \text{ over } \partial\gamma = \int_\gamma (\text{divergence of } \mathbf{F}),$$

with $\mathbf{F}$ a vector field, for the geometric "figure" $\gamma =$ the unit ball in R^n. For the case of the unit ball in R^n, the reduction of the Divergence Theorem to Stokes' Theorem is a matter of the *differential* calculus of vector fields, differential forms, inner products etc. (See e.g. [10] p. 204). For the convenience of the reader, we recall the case $n = 2$.

Given a vector field $\mathbf{F}(x, y) = (F(x, y),\ G(x, y))$ in R^2, apply Stokes' Theorem to the differential form

$$\omega := -G(x, y)dx + F(x, y)dy$$

for the singular rectangle γ given by (17) above. Then

$$\begin{cases} \gamma^*(dx) = \cos\theta dr - r\sin\theta d\theta \\ \gamma^*(dy) = \sin\theta dr + r\cos\theta d\theta \\ \gamma^*(dx \wedge dy) = r\ (dr \wedge d\theta) \end{cases}$$

Since $d\omega = (\partial G/\partial y + \partial F/\partial x)\ dx \wedge dy = \text{ div } (\mathbf{F})\ dx \wedge dy$, then

$$\gamma^*(d\omega) = \text{ div } (\mathbf{F})\ r\ (dr \wedge d\theta)$$

On the other hand,

$$\gamma^*\omega = (F\ \sin\theta - G\ \cos\theta)dr + (F\ r\ \cos\theta + G\ r\ \sin\theta)\ d\theta, \tag{18}$$

(all F, G, and $\mathbf{F}$ to be evaluated ar $(r\cos\theta, r\sin\theta)$). Therefore

$$\int_\gamma d\omega = \int_0^{2\pi}\int_0^1 \text{ div } (\mathbf{F})\ r\ dr\ d\theta;$$

this is $\int_{B_1}$ div $(\mathbf{F})$ dA. On the other hand by Stokes' Theorem $\int_\gamma d\omega = \int_{\partial\gamma}\omega$ which is a curve integral of the 1-form (18) around the boundary of the rectangle $[0, 2\pi] \times [0, 1]$. This curve integral is a sum of four terms corresponding to the four sides of the rectangle. Two of these (corresponding to the sides $\theta = 0$ and $\theta = 2\pi$) cancel, and the term corresponding to the side where $r = 0$ vanishes because of the r in $r\ (dr \wedge d\theta)$, so only the side with $r = 1$, $0 \leq \theta \leq 2\pi$ remains, and its contribution is, with the correct orientation,

$$\int_0^{2\pi}(F(\cos\theta, \sin\theta)\cos\theta + G(\cos\theta, \sin\theta)\sin\theta)\ d\theta = \int_{S_1} \mathbf{F}\cdot\mathbf{n}\ ds$$

where $\mathbf{n}$ is the outward unit normal of the unit circle. This expression is the flux of $\mathbf{F}$ over the unit circle, which thus equals the divergence integral calculated above.

We insert for reference two obvious "change of variables" equations. Recall that $H_t : R^n \to R^n$ is the homothetic transformation "multiplying by t". We have, for any vector field $\mathbf{F}$ on R^n (viewed, via principal part, as a map $R^n \to R^n$):

$$\text{div } (\mathbf{F} \circ H_t) = t \cdot (\text{div } \mathbf{F}) \circ H_t, \tag{19}$$

and

$$t^n \int_{B_1} \phi \circ H_t = \int_{B_t} \phi. \tag{20}$$

We now combine vector calculus with the calculus of the basic ball- and sphere-distributions, as introduced in Section 3, to prove

Theorem 14 *In R^n (for any n), we have, for any t,*

$$\frac{d}{dt}S^t = t \cdot \Delta(B^t),$$

($\Delta =$ the Laplace operator).

Proof. We prove first that

$$t^{n-1} \cdot \frac{d}{dt}S^t = t^n \cdot \Delta(B^t).$$

In fact, for any test function ψ,

$$< t^{n-1} \cdot \frac{d}{dt}S^t, \psi >= t^{n-1} \cdot \frac{d}{dt}\int_{S_1} \psi(tu)\ du = t^{n-1}\int_{S_1} (\nabla\psi)(tu)\ \cdot u\ du,$$

(by differentiating under the integral sign and using the chain rule)

$$= t^{n-1} \cdot \text{ flux of } ((\nabla\psi) \circ H_t) \text{ over } S_1,$$

where $H_t : R^n \to R^n$ is the homothetic transformation "multiplying by t". This, by the Divergence Theorem, may be rewritten as

$$t^{n-1}\int_{B_1} \text{ div } ((\nabla\psi) \circ H_t) = t^n \int_{B_1} (\text{ div } (\nabla\psi)) \circ H_t,$$

(using (19))

$$= t^n \cdot \int_{B_1} (\Delta\psi) \circ H_t = \int_{B_t} \Delta\psi$$

(by a standard change of variables, cf. (20)), so

$$=< B_t, \Delta\psi >=< t^n \cdot B^t, \Delta\psi >= t^n \cdot < \Delta B^t, \psi > .$$

From

$$t^{n-1} < \frac{d}{dt}S^t, \psi >= t^n < \Delta B^t, \psi >,$$

we may of course conclude the desired equality, by cancelling t^{n-1} on both sides, if t is invertible; but we want the equation *for all* t. We can get this from "Lavendhomme's principle", which says that if $f : R \to R$ satisfies $t \cdot f(t) = 0$ for all t, then $f(t)$ is constantly 0. This principle was derived from the integration axiom purely synthetically by Lavendhomme in [11] p.25. So the claim of the Theorem is valid for all t.

We collect information about t-derivatives of the four basic distributions S_t , B_t, S^t and B^t in R^n. The results are valid for any n and any t. For invertible t (say positive t), some of the statements may be simplified by multiplying by t^{-1}, but we prefer having formulae which are universally valid.

Theorem 15 *We have in dimension n for all t:*

$$\frac{d}{dt}(B_t) = S_t, \tag{21}$$

$$t \cdot \frac{d}{dt}(S_t) = (n-1)\ S_t + t \cdot \Delta(B_t), \tag{22}$$

$$t \cdot \frac{d}{dt}(B^t) = S^t - n\ B^t, \tag{23}$$

$$\frac{d}{dt}(S^t) = t \cdot \Delta(B^t), \tag{24}$$

In dimension 1*, we also have*

$$\frac{d}{dt}(S_t) = \Delta B_t \tag{25}$$

Proof. Equation (21) is an immediate consequence of the fundamental theorem of calculus; e.g. for $n = 2$, consider the explicit formula for B_t given above in Section 3 ("Spheres and balls as distributions"). With $\int_0^t$ as the outer integral, the d/dt of it is just the inner integral, i.e., exactly the exhibited formula (idem) for S_t.

For (22), we t-differentiate the equation $S_t = t^{n-1} \cdot S^t$ by the Leibniz rule and get $(n-1) \cdot t^{n-2} \cdot S^t + t^{n-1} \cdot d/dt(S^t)$; so by Theorem 14,

$$d/dt(S_t) = (n-1) \cdot t^{n-2} \cdot S^t + t^{n-1} \cdot t \cdot \Delta(B^t).$$

If we multiply this equation by t, we get

$$t \cdot d/dt(S_t) = (n-1) \cdot t^{n-1} \cdot S^t + t^{n+1} \cdot \Delta(B^t);$$

using $S_t = t^{n-1} S^t$ and $B_t = t^n B^t$, the result follows (note that Δ commutes with multiplication by t).

The proof of (23) is similar: t-differentiating $t^n \cdot B^t = B_t$, we get

$$n \cdot t^{n-1} \cdot B^t + t^n \cdot d/dtB^t = d/dtB_t = S_t,$$

(using (21)), so using $S_t = t^{n-1}S^t$, this equation may be rewritten as

$$t^{n-1} \cdot (n \cdot B^t + t \cdot d/dtB^t) = t^{n-1} \cdot S^t$$

The result now follows by cancelletion of the factor t^{n-1} by Lavendhomme's principle, and rearranging.

Next, (24) is identical to Theorem 14, and is included again for completeness' sake.

Finally, (25) follows from (22): the first term vanishes, since $n - 1 = 0$, and in the remaining equation, we may cancel the factor t by Lavendhomme's principle. Alternatively, (25) can be proved directly, by a very simple calculation.

5 Wave equation

Let Δ denote the Laplace operator $\sum \partial^2/\partial x_i^2$ on R^n. We shall consider the wave equation (WE) in R^n, (for $n = 1, 2, 3$),

$$\frac{\partial^2}{\partial t^2} Q = \Delta Q \tag{26}$$

as a second order ordinary differential equation on the Euclidean vector space $\mathcal{D}'_c(R^n)$ of distributions of compact support; in other words, we are looking for functions

$$Q : R \to \mathcal{D}'_c(R^n)$$

so that for all $t \in R$, $\ddot{Q}(t) = \Delta(Q(t))$ (viewing Δ as a map $\mathcal{D}'_c(R^n) \to \mathcal{D}'_c(R^n)$. We shall only be looking for *particular* solutions, in fact, so called *fundamental* solutions: solutions whose initial value and initial speed is either the Dirac distribution at 0, or 0. Given any other initial value and speed — these being both assumed to be distributions of compact support —, the corresponding particular solution may, as is well known, be obtained from the fundamental solution just by *convolution* $*$ with these fundamental solutions. This follows purely formally from the rules for convolution of distributions P and Q, such as $Q * \delta(0) = Q$, $D(P * Q) = D(P) * Q$, where D is any differential operator on R^n with constant coefficients; and from linearity of the convolution, implying that $d/dt(P_t * Q) = (d/dt\ P_t) * Q$; see e.g. [16], Ch. 3.

Dimension 1

Theorem 16 *The function $R \to \mathcal{D}'_c(R)$ given by*

$$t \mapsto 1/2 \cdot S^t$$

is a solution of the WE in dimension 1; its initial value and speed are, respectively $\delta(0)$ and 0.

The function $R \to \mathcal{D}'_c(R)$ given by

$$t \mapsto 1/2 B_t$$

is a solution of the WE; its initial value and speed are, respectively, 0 and $\delta(0)$.

Proof. The statements about the initial values are immediate from the explicit integral formulas for B_t and S^t (putting $t = 0$). The statements about the initial speeds are equally immediate from the following formulas (27) and (29) for the t-derivatives, (putting $t = 0$). We have by (24)

$$\frac{d}{dt}(S^t) = t \cdot \Delta(B^t), \tag{27}$$

and so by further t differentiation

$$\frac{d^2}{dt^2}(S^t) = \Delta(B^t) + t \cdot \frac{d}{dt}(\Delta(B^t));$$

now, d/dt and Δ commute, so we may continue

$$= \Delta(B^t) + \Delta(t \cdot \frac{d}{dt} B^t) = \Delta(B^t) + \Delta(S^t - 1 \cdot B^t),$$

using (23) with $n = 1$. Now by linearity of Δ, the terms involving B^t in the last expression cancel, and we are left with

$$\frac{d^2}{dt^2}(S^t) = \Delta(S^t), \tag{28}$$

which establishes WE for S^t and hence also for $1/2 \cdot S^t$.

Also, by (21), we have that

$$\frac{d}{dt}(B_t) = S_t, \tag{29}$$

and so by further t differentiation

$$\frac{d^2}{dt^2}(B_t) = \frac{d}{dt}(S_t) = \Delta(B_t),$$

using (25), which establishes WE for B_t and hence for $1/2 \cdot B_t$. So the theorem is proved.

Dimension 3

Theorem 17 *The function $R \to \mathcal{D}'_c(R^3)$ given by*

$$t \mapsto \frac{1}{4\pi} \cdot t \cdot S^t$$

is a solution of the WE in dimension 3; its initial value and speed are, respectively, 0 and $\delta(0)$.

The function $R \to \mathcal{D}'_c(R^3)$ given by

$$t \mapsto \frac{1}{4\pi} \cdot (S^t + t^2 \cdot \Delta(B^t))$$

is a solution of the WE; its initial value and speed are, respectively, $\delta(0)$ and 0.

Proof. We calculate first d/dt of $t \cdot S^t$, using (24):

$$\frac{d}{dt}(t \cdot S^t) = S^t + t^2 \cdot \Delta(B^t), \tag{30}$$

and so by Theorem 14 (= (24)),

$$\frac{d^2}{dt^2}(t \cdot S^t) = t \cdot \Delta(B^t) + 2 \cdot t \cdot \Delta(B^t) + t^2 \cdot \Delta(\frac{d}{dt}B^t)$$

$$= 3 \cdot t \cdot \Delta(B^t) + t \cdot \Delta(t \cdot \frac{d}{dt}B^t)$$

$$= 3 \cdot t \cdot \Delta(B^t) + t \cdot \Delta(S^t - 3B^t),$$

using (23), and now by linearity of Δ, the terms involving $\Delta(B^t)$ cancel, so we are left with the equation

$$\frac{d^2}{dt^2}(t \cdot S^t) = \Delta(t \cdot S^t), \tag{31}$$

which establishes WE for $t \cdot S^t$ and hence for $1/4\pi \ \cdot t \cdot S^t$. The statements about initial value and speed are immediate (using (30) for the speed).

Because d^2/dt^2 and Δ commute, it is clear that if $t \mapsto Q(t)$ is a distributional solution of WE, then so is $t \mapsto d/dtQ(t)$. So since $t \cdot S^t$ is a solution, then so is its t-derivative (calculated in (30) above), i.e. $S^t + t^2 \cdot \Delta(B^t)$ is a solution. Its initial value and its initial speed can be found by putting $t = 0$ in (31) (note δ commutes with multiplication by t).

Dimension 2

Recall that we considered the orthogonal projection $p : R^3 \to R^2$. Applying covariant functorality, we get for any distribution Q on R^3 of compact support a distribution $p(Q)$ on R^2, also of compact support.

Theorem 18 *The function $R \to \mathcal{D}'_c(R^2)$ given by*

$$t \mapsto \frac{1}{4\pi} \cdot p(t \cdot S^t)$$

is a solution of the WE in dimension 2; its initial value and speed are, respectively, 0 *and* $\delta(0)$.

The function $R \to \mathcal{D}'_c(R^2)$ given by $t \mapsto 1/4\pi \cdot p(S^t + t^2 \cdot \Delta(B^t))$ is also a solution of the WE in dimension 2; its initial value and speed are, respectively, $\delta(0)$ *and* 0.

Recall that an explicit integral formula for $p(t \cdot S^t)$, for $t > 0$, was given above, in (16) ("Poisson kernel").

Proof. The fact that the two distributions in question are solutions of the WE is immediate from the Proposition 12 ("p commutes with Δ") and from the fact that $\mathcal{D}'_c(p) : \mathcal{D}'_c(R^3) \to \mathcal{D}'_c(R^2)$ is linear, and hence commutes with formation of d/dt; also, $\mathcal{D}'_c(p)$ sends Dirac distribution at $0 \in R^3$ to Dirac distribution at $0 \in R^2$, so the initial values and speeds are as claimed.

The Taylor Series at $t = 0$ for the solutions given can be calculated directly, but they can more easily be obtained from the formal solution given in Proposition 6.

6 Heat equation.

In this section we deal with distributions that do not have compact support and we only consider the one-dimensional case. We are thus considering solutions for the vector field on the Euclidean vector space $\mathcal{D}'(R)$, whose principal part is given by $\Delta : \mathcal{D}'(R) \to \mathcal{D}'(R)$. We consider the particular solution $K : R_{\geq 0} \to \mathcal{D}'(R)$ whose initial value is the distribution $\delta(0)$. Thus, referring to the general treatment of solutions for (differential equations given by) vector fields, we are considering $\tilde{R} = R_{\geq 0}$; for the heat equation, one cannot do better, as is well known. Also, as mentioned above, we rely on external (classical) calculus; namely, we consider the classical "heat kernel" function, i.e., the function $K : R_{\geq 0} \to \mathcal{D}'(R)$ given by

$$K(t) = \begin{cases} \frac{1}{\sqrt{4\pi t}}\, e^{-\frac{x^2}{4t}} & \text{for } t > 0 \\ \\ \delta(0) & \text{otherwise .} \end{cases} \tag{32}$$

Here, for the case $t > 0$, we described a function rather than a distribution, so here we do make the identification of functions $g(x)$ with distributions $\phi \mapsto \int_{-\infty}^{\infty} g(x)\phi(x)\, dx$. Differentiation of distributions reduces to differentiation of the representing functions. For $t > 0$, we thus have $K_t(x) = K(t, x)$, a smooth function in two variables, described by the above expression. It satisfies the heat equation

$$\frac{\partial K}{\partial t} = \frac{\partial^2 K}{\partial x^2},$$

for $t > 0$. Also the following limit expression is classical:

$$\lim{}_{t \to 0^+} \int_{-\infty}^{\infty} K(t, x)\phi(x)\, dx = \phi(0) \tag{33}$$

for any test function ϕ. More generally,

Proposition 19 *For any integer $n \geq 0$, and any test function ϕ*

$$\lim{}_{t \to 0^+} \frac{\partial^n}{\partial t^n} \int_{-\infty}^{\infty} K(t, x)\phi(x)\, dx = \phi^{(2n)}(0). \tag{34}$$

Proof. The case $n = 0$ is just (33); the general case follows by iteration. Let us do the case $n = 1$. Then

$$\begin{aligned}\frac{\partial}{\partial t}\int_{-\infty}^{\infty} K(t,x)\phi(x)\,dx &= \int_{-\infty}^{\infty}\frac{\partial}{\partial t}K(t,x)\phi(x)\,dx\\ &= \int_{-\infty}^{\infty}\frac{\partial^2}{\partial x^2}K(t,x)\phi(x)\,dx\end{aligned}$$

(by the heat equation for K)

$$= \int_{-\infty}^{\infty} K(t,x)\phi^{(2)}(x)\,dx$$

(by integration by parts.)

We then use (33), for the test function $\phi^{(2)}$ to conclude (34) for $n = 1$.

Proposition 20 *The function $K : R_{\geq 0} \to \mathcal{D}'(R)$ is smooth.*

Here, smoothness is taken in the following sense (appropriate for convenient vector spaces): for each test function ϕ, the function $R_{\geq 0} \to R$ given by $t \mapsto < K(t), \phi >$ is smooth.

Proof. It suffices to prove that K is infinitely often differentiable at 0, since smoothness for $t > 0$ is clear. For fixed $t > 0$, we let K_t denote the function in x described in (the first clause in) (32) above. Thus, $< K(t), \phi >$ is given by the integral

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{4\pi t}}\,e^{-\frac{x^2}{4t}}\cdot\phi(x)\,dx. \tag{35}$$

We first notice that, by Hadamard's Lemma, $\phi(x) = \phi(0) + x\psi(x)$. By linearity, $< K_t, \phi >=< K_t, \phi(0) > + < K_t, x\psi(x) >$. But $< K_t, \phi(0) >= \phi(0)$ and this implies that the derivative of $< K_t, \phi >$ at 0 is

$$\lim_{t\to 0^+}(1/t) < K_t, x\psi(x) > \tag{36}$$

To compute this limit, we use the formulas and notations in Lang's book [10], with the exception that we use $\mathcal{F}$ for the Fourier transform. We also use the following well known formulae, where all the functions under considerations belong to the class S of fast decreasing functions and thus $\mathcal{F}$ works with no limitations. First, for any pair of functions α, β in this class, one has the "adjointness" formula

$$\int_{-\infty}^{\infty} \mathcal{F}(\alpha)\beta = \int_{-\infty}^{\infty} \alpha\mathcal{F}(\beta).$$

Furthermore

$$\mathcal{F}((1/\sqrt{4\pi})e^{-t\xi^2})(x) = (1/\sqrt{4\pi t})e^{-x^2/4t} \tag{37}$$

$$\mathcal{F}(x\psi(x))(\xi) = i(\mathcal{F}(\psi))'(\xi) \tag{38}$$

$$\xi\mathcal{F}(\psi)(\xi) = -i\mathcal{F}(\psi')(\xi) \tag{39}$$

To show the existence of the limit, we compute, using (37), adjointness, and (38)

$$\begin{aligned}
< K_t, x\psi(x) > &= \int_{-\infty}^{\infty} (1/\sqrt{4\pi t})e^{-x^2/4t} x\psi(x)dx \\
&= \int_{\infty}^{\infty} \mathcal{F}((1/\sqrt{4\pi})e^{-t\xi^2})(x)x\psi(x)dx \\
&= \int_{\infty}^{\infty} ((1/\sqrt{4\pi})e^{-t\xi^2})\mathcal{F}(x\psi(x))(\xi)d\xi \\
&= \int_{\infty}^{\infty} ((1/\sqrt{4\pi})e^{-t\xi^2})i(\mathcal{F}(\psi))'(\xi)d\xi \\
&= 2it\int_{-\infty}^{\infty} (1/\sqrt{4\pi})e^{-t\xi^2}\xi\mathcal{F}(\psi)(\xi)d\xi,
\end{aligned}$$

The last step uses integration by parts. Using (39), this may be rewritten as

$$= 2t\int_{-\infty}^{\infty} (1/\sqrt{4\pi})e^{-t\xi^2}\mathcal{F}(\psi')(\xi)\, d\xi$$

$$= 2t\int_{-\infty}^{\infty} (1/\sqrt{4\pi t})e^{-x^2/4t}\psi'(x)\, dx$$

using adjointness and (37) in the last step. Now we divide by t, as requested in (36), and let $t \to 0^+$. Using (33), we thus get that the limit in (36) equals

$$\lim_{t\to 0^+} 2\int_{-\infty}^{\infty} (1/\sqrt{4\pi t})e^{-x^2/4t}\psi'(x)\, dx = 2 \cdot \psi'(0).$$

But since $\phi(x) = \phi(0) + x \cdot \psi(x)$, $2 \cdot \psi'(0) = \phi''(0)$, This proves that the limit in (36) exists and equals $\phi''(0)$; we conclude that

$$\lim_{t\to 0^+} (1/t)[< K_t, \phi > - \phi(0)] = \phi''(0). \tag{40}$$

To better understand what has been done and to develop this matter further, let us define for every $t \geq 0$

$$f(t) = < K_t, \phi >$$

We can summarize the results of this section as follows

$$f'(0) = \phi''(0)$$

Recall from Proposition 19 that

$$f^{(n)}(t) = < K_t, \phi^{(2n)} >$$

and thus, by going to the limit when $t \to 0^+$,

$$\lim_{t \to 0^+} f^{(n)}(t) = \phi^{(2n)}(0)$$

These results suffice to summarize the result in the present Section in the following way:

Corollary 21 *The function f is smooth and, furthermore, $f^{(n)}(0) = \phi^{(2n)}(0)$.*

Proof. Let us show, for instance, that $f''(0)$ exists and equals $\phi''''(0)$. Using the previous results, $(1/t)[f'(t) - f'(0)] = (1/t)[< K_t, \phi'' > - \phi''(0)]$ and this implies the corollary, by going to the limit when $t \to 0^+$ and using (40) with the function ϕ'' instead of ϕ. Now, iterate.

The idea to use Fourier transform to prove smoothness was pointed out to us by H. Stetkær and E. Skibsted.

Summarizing: we have a smooth function $K : R_{\geq 0} \to \mathcal{D}'(R)$, satisfying the heat equation $\partial K / \partial t(t) = \Delta(K(t))$ for all $t \geq 0$; for $t = 0$, this follows from Proposition 19. By the assumed fullness of the embedding of smooth manifolds with boundary and convenient vector spaces into the model of SDG, we have the desired solution internally in the model. We may then ask for the values of K for *nilpotent* t. The answer can be deduced from the Taylor Series at 0 for the function K, and the coefficients can be read off from Proposition 19; alternatively, by the uniqueness of formal solutions (Theorem 2), they can be read off from the formal solution we know already from Proposition 5. In any case, we get for nilpotent t

$$K(t) = \delta(0) + t \cdot \Delta(\delta(0)) + \frac{t^2}{2!} \Delta^2(\delta(0)) + \dots \tag{41}$$

the series being a finite sum, since t is nilpotent. In particular, for d with $d^2 = 0$, we have $K(d) = \delta(0) + d \cdot \Delta(\delta(0))$, or since $\Delta = (\ )''$,

$$K(d) = \delta(0) + d \cdot \delta(0)'' \tag{42}$$

In some sense, the motivation for our study of the heat equation in particular was to see how $\delta(0)$ evolves in nilpotent lapse t of time and specially for $t = d$ with $d^2 = 0$; the answer is (42) (or more generally (41)).

Being an extensive quantity, a distribution like (42) should be drawable. In fact, it can be exhibited as a finite linear combination of Dirac distributions $\delta(a)$ (= "evaluate at a"). This hinges on:

Proposition 22 *Let $h^4 = 0$. Then*

$$h^2 \cdot \delta(0)'' = \delta(-h) - 2\delta(0) + \delta(h).$$

Proof. It suffices to prove, for an arbitrary test function ϕ, that $h^2 \cdot \phi''(0) = \phi(-h) - 2\phi(0) + \phi(h)$; now just Taylor expand the two outer terms in the sum on the right; the terms of odd degree cancel, the terms of even degree (0 and two) give the result. (There is a similar result for higher derivatives of $\delta(0)$: for $h^{n+1} = 0$,

$$h^n \cdot \delta(0)^{(n)} = \sum_{i=0}^{n} (-1)^i (n, i) \delta(i \cdot h),$$

where (n, i) denotes the binomial coefficient $n!/i!(n-i)!$. This hinges on some combinatorics with binomial coefficients, cf. [3] p. 63, Problem 16).

To make a "drawing" of $K(d)$ where $d^2 = 0$, we assume that $d = h^3$ for some h with $h^4 = 0$ (we shall not deal here with the question whether this can always be done). Then

$$K(d) = \delta(0) + d \cdot \delta(0)'' = \delta(0) + h \cdot h^2 \cdot \delta(0)'' = \delta(0) + h \cdot ((\delta(-h) - 2\delta(0) + \delta(h))$$

using (42) and (22). The drawing one can make of $\delta(x)$ (as for any discrete distribution), is a column diagram: erect a column of heigth 1 at x. The distribution above then comes about by removing $2h$ units from the unit column at 0, and placing the small columns of heigth h at $-h$ and h. This is the beginning of the *diffusion* of the Dirac distribution. Several other ways of exhibiting $K(d)$ as linear combination of Dirac distributions are also possible.

Since $\mathcal{D}'(R^n)$ is a microlinear and Euclidean R-module, and $\Delta : \mathcal{D}'(R^n) \to \mathcal{D}'(R^n)$ is linear, we may apply the general results of Propositions 5 and 6 to conclude that the formal solution of the heat equation $\dot{F}(t) = \Delta(F(t))$ with initial value (the distribution) μ, is the series

$$\mu + t\Delta(\mu) + t^2/2!\Delta^2(\mu) + t^3/3!\Delta^3(\mu) + \dots .$$

Similarly, the formal solution of the wave equation $\ddot{F}(t) = \Delta(F(t))$ with initial value (the distribution) μ, and initial speed the distribution ν is the series

$$\mu + t\nu + t^2/2!\Delta(\mu) + t^3/3!\Delta(\nu) + t^4/4!\Delta^2(\mu) + t^5/5!\Delta^2(\nu) \dots .$$

Applying (in the one-variable case, say) these formulas to a test function ϕ in the variable x and to the distributions $\delta(0)$ and $\delta'(0)$ we obtain the following Maclaurin series for the heat equation

$$< F(t), \phi >= \phi(0) + t\phi^{('')}(0) + t^2/2!\phi^{('''')}(0) + \dots .$$

Here $\dot{()}$ refers to the time derivative, whereas $()'$ to the space derivative $\partial/\partial x$. The variable x has been left unexpressed. There is a similar series for the wave equation:

$$< F(t), \phi >= \phi(0) + t\dot{\phi}(0) + t^2/2!\phi^{('')}(0) + t^3/3!\dot{\phi}^{('')}(0) + t^4/4!\phi^{''''}(0) + \dots .$$

6.1 Simple Transport

For the sake of completeness, we also consider the function $\delta : R \to \mathbf{D}'_c(R)$ given by $t \mapsto \delta(t)$, the Dirac distribution at $t \in R$. This is the "fundamental solution" for the equation for "simple transport", cf. e.g. [17].

Proposition 23 *The function δ is the solution for the differential equation for "simple transport",*

$$\frac{d}{dt}(\delta) = (\delta)'$$

with initial value $\delta(0)$.

Proof. For any test function ϕ,

$$\frac{d}{dt} < \delta(t), \phi >= \frac{d}{dt}\phi(t) = \phi'(t) =< \delta(t)', \phi > .$$

References

[1] Bitsadze, A.V. *Equations of Mathematical Physics.* MIR Publishers Moscow 1980

[2] Dubuc, E. Sur les modèles de la Géométrie Différentielle Synthétique. *Cahiers de Top. et Géom Diff.*, (20), 231-279, 1979

[3] Feller, W. *An Introduction to Probability Theory and its Applications volume I* John Wiley and Sons 1965 (Second Edition)

[4] Frölicher, A. and A. Kriegl. *Linear Spaces and Differentiation Theory.* John Wiley and Sons 1988

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

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

[7] Kock, A. Convenient vector spaces embed into the Cahiers topos. *Cahiers de Top. et Géom. Diff.*, (27) 1986

[8] Kock, A. and G.E. Reyes. Corrigendum and addenda to the paper "Convenient vector spaces embed...".*Cahiers de Top. et Géom. Diff.* (28-2), 99-110, 1987

[9] Kock, A. and G.E. Reyes. Aspects of Fractional Exponents. *Theory and Applications of Categories* vol. 5 (10) 1999

[10] Lang, S. *Analysis II.* Addison-Wesley 1969

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

[12] Lawvere, F.W. Categorical Dynamics. *Topos Theoretic Methods in Geometry*, ed. A. Kock, Aarhus Various Publ. Series 30 (1979).

[13] Lawvere, F.W. Categories of Space and of Quantity. *The Space of Mathematics: Philosophical, Epistemological and Historical Explorations. (International Symposium on Structures in Mathematical Theories. San Sebastian, Spain 1990).* p. 14-30. DeGruyter, Berlin 1992

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

[15] Quê, Ngo Van and G.E. Reyes. Smooth functors and synthetic calculus. *The L.E.J. Brouwer Centenary Symposium* p. 377-395. North Holland 1982

[16] Schwartz, L. *Méthodes mathématiques pour les sciences physiques.* Hermann 1998

[17] Strauss, W.A. *Partial Differential Equations.* John Wiley and Sons 1992