Theory and Applications of Categories, Vol. 26, No. 4, 2012, pp. 97–131.

COMMUTATIVE MONADS AS A THEORY OF DISTRIBUTIONS

ANDERS KOCK

ABSTRACT. It is shown how the theory of commutative monads provides an axiomatic framework for several aspects of distribution theory in a broad sense, including probability distributions, physical extensive quantities, and Schwartz distributions of compact support. Among the particular aspects considered here are the notions of convolution, density, expectation, and conditional probability.

Introduction

The word "distribution" has a common, non technical, meaning, roughly synonymous with "dispersion": a given quantity, say a quantity of tomato, may be distributed or dispersed over a given space, say a pizza. A more mathematical use of the word was made precise in functional analysis by L. Schwartz (and predecessors): it is a (continuous linear) functional on a space of functions. Schwartz argues that "the mathematical distributions constitute a correct mathematical definition of the distributions one meets in physics", [19] p. 84.

Our aim here is to present an alternative mathematical theory of distributions (of compact support), applicable also for the "tomato" example, but which does not depend on the "double dualization" construction of functional analysis; and also, to provide a canonical comparison to the distributions of functional analysis.

Distributions of compact support form an important example of an *extensive quantity*, in the sense made precise by Lawvere, cf. e.g. [14], and whose salient feature is the covariant functorial dependence on the space over which it is distributed. Thus, there is a covariant functor T, such that extensive quantities of a given type on a given space X form a new "linear" space $T(X)$. In the theory we present here, T is an endofunctor on a cartesian closed category $\mathcal{E}$, in fact a monad; a linear structure grows, for monads with a certain property, out of the monad structure.

An example of such T is, for suitable $\mathcal{E}$, the double dualization construction S that gives the space of Schwartz distributions of compact support, – except that this S is not commutative. The unit η of the monad S associates to an $x \in X$ the Dirac distribution δ_x at x.

What makes our theory simple is the universal property which the unit map $\eta : X \to T(X)$ is known to have by general monad theory, – in conjunction with the assumed

Received by the editors 2011-09-29 and, in revised form, 2012-01-24.
Transmitted by F. W. Lawvere. Published on 2012-02-24.
2000 Mathematics Subject Classification: 18C15, 18D25, 46T30.
Key words and phrases: monads, distributions, extensive quantities.

commutativity of T (in the sense of [6]). This is what makes for instance the notions of linearity/bilinearity work well.

Generalities: We mainly compose maps from right to left (this is the default, and it is denoted $g \circ f$); but occasionally, in particular in connection with displayed diagrams, we compose from left to right (denoted $f.g$); we remind the reader when this is the case.

In Sections 1-11, the notions and theory (except the remarks on "totals") work for $\mathcal{E}$ a general symmetric monoidal closed category, not necessarily cartesian closed. Here, the reasoning is mainly diagrammatic. In the remainder, we take the liberty to reason with "elements". In particular, the "elements" of $T(X)$ are simply called "distributions on X"

The present text subsumes and simplifies the preliminary arXiv texts, [12], [13], and has been presented in part in Krakow at the conference "Differential Geometry and Mathematical Physics, June-July 2011, in honour of Wlodzimierz Tulczyjew", and at the Nancy Symposium "Sets within Geometry" July 2011. I want to thank the organizers of these meetings for inviting me. I also want to thank Bill Lawvere for fruitful discussions in Nancy, and e-mail correspondence in the spring of 2011.

1. Monads, algebras, and linearity

Recall the notion of monad $T = (T, \eta, \mu)$ on a category $\mathcal{E}$, cf. e.g. [16] 6.1 and 6.2. Recall also the notion of T-algebra $A = (A, \alpha)$ for such monad; here, $\alpha : T(A) \to A$ is the *structure* map for the given algebra. There is a notion of *morphism* of algebras $(A, \alpha) \to (B, \beta)$, cf. loc.cit., so that we have the category $\mathcal{E}^T$ of algebras for the monad T. For each $X \in \mathcal{E}$, we have the T-algebra $T(X)$ with structure map $\mu_X : T^2(X) \to T(X)$. The map $\eta_X : X \to T(X)$ has a universal property, making $(T(X), \mu_X)$ into a *free* algebra on X: to every T-algebra (A, α) and every map $f : X \to A$ in $\mathcal{E}$, there is a unique morphism of algebras $\overline{f} : T(X) \to A$ with $\overline{f} \circ \eta_X = f$ (namely $\overline{f} = \alpha \circ T(f)$).

Recall that any algebra structure α on an object A is by itself an algebra morphism $T(A) \to A$, and in particular μ_X is an algebra morphism. Recall also that any morphism $T(f) : T(X) \to T(Y)$ (for $f : X \to Y$ an arbitrary morphism in $\mathcal{E}$) is an algebra morphism.

For the monads that we are to consider in the present article, namely *commutative* monads (cf. Section 9 below), there are good reasons for an alternative terminology: namely, T-algebras deserve the name *T-linear spaces*, and homomorphisms deserve the name *T-linear maps* (and, if T is understood from the context, the 'T' may even be omitted). This allows us to talk about *partial T-linear maps*, as well as *T-bilinear maps*, as we shall explain.

An example of a commutative monad, with $\mathcal{E}$ the category of sets, is the functor T which to a set X associates (the underlying set of) the free real vector space on X. In this case, the algebras for T are the vector spaces over $\mathbb{R}$, and the morphisms are the $\mathbb{R}$-linear maps.

For a general monad, the "linear" terminology is usually not justified, but we shall nevertheless use this terminology right from the beginning, in so far as maps are concerned. Thus, the map $\overline{f} : T(X) \to A$ considered above will be called *the T-linear extension of* $f : X \to A$ (and $\eta_X : X \to T(X)$ is then to be understood from the context).

2. Enrichment and strength

We henceforth consider the case where $\mathcal{E}$ is a cartesian closed[1] category (cf. e.g. [16] 4.6); so, for any pair of objects X, Y in $\mathcal{E}$, we have the exponential object Y^X, and an "evaluation" map $ev : Y^X \times X \to Y$ (or more pedantically: $ev_{X,Y}$). Since we shall iterate the "exponent formation", it becomes expedient to have an on-line notation. Several such are in use in the literature, like $[X, Y]$, $X \multimap Y$, or $X \pitchfork Y$; we shall use the latter (tex code for $\pitchfork$ is `\pitchfork`; read "$X \pitchfork Y$" as "X hom Y"), so that the evaluation map is a map

$$ev : (X \pitchfork Y) \times X \to Y.$$

Since $\mathcal{E}$ is a monoidal category (with cartesian product as monoidal structure), one has the notion of when a category is *enriched* in $\mathcal{E}$, cf. e.g. [1] II.6.2. And since this monoidal category $\mathcal{E}$ is in fact monoidal closed (with $\pitchfork$ as the closed structure), $\mathcal{E}$ is enriched in itself. For $\mathcal{E}$-enriched categories, one has the notion of enriched functor, as well as enriched natural transformation between such. (cf. e.g. loc.cit. Def. 6.2.3 and 6.2.4). So in particular, it makes sense to ask for an $\mathcal{E}$-enrichment of the functor $T : \mathcal{E} \to \mathcal{E}$ (and also to ask for $\mathcal{E}$-enrichment of η and μ, which we shall, however, not consider until in the next Section).

Specializing the definition of enrichment to the case of an endofunctor $T : \mathcal{E} \to \mathcal{E}$ gives that such enrichment consists in maps in $\mathcal{E}$

$$X \pitchfork Y \xrightarrow{st_{X,Y}} T(X) \pitchfork T(Y),$$

(for any pair of objects X, Y in $\mathcal{E}$) satisfying a composition axiom and a unit axiom. The word "enriched" is sometimes replaced by the word "strong", and "enrichment" by "strength", whence the notation "*st*". We shall, however, need to consider two other equivalent manifestations of such "strength" structure on T, introduced in [6] (and proved equivalent to strength in [10], Theorem 1.3), and in [8], called *tensorial* and *cotensorial* strength, respectively. To give a tensorial strength to the endofunctor T is to give, for any pair of objects X, Y in $\mathcal{E}$ a map

$$X \times T(Y) \xrightarrow{t''_{X,Y}} T(X \times Y),$$

satisfying a couple of equations (cf. [10] (1.7) and (1.8)). The tensorial strength t'' has a "twin sister"

$$T(X) \times Y \xrightarrow{t'_{X,Y}} T(X \times Y),$$

[1] For most of Section 1-11, symmetric monoidal closed will do.

essentially obtained from t'' by conjugation using the twist map $X \times Y \to Y \times X$. Similarly, to give a cotensorial strength to the endofunctor T is to give, for any pair of objects X, Y in $\mathcal{E}$ a map

$$T(X \pitchfork Y) \xrightarrow{\lambda_{X,Y}} X \pitchfork T(Y),$$

(cf. [8]) satisfying a couple of equations. The three manifestations of strength, and also the equations they satisfy, are deduced from one another by simple "exponential adjointness"-transpositions, cf. loc.cit. As an example, let us derive the tensorial strength t'' from the classical "enrichment" strength st: We have the unit u of the adjunction $(- \times Y) \dashv (Y \pitchfork -)$; it is the first map in the following composite; the second is the enrichment strength

$$X \xrightarrow{u} Y \pitchfork (X \times Y) \xrightarrow{st_{Y,X\times Y}} T(Y) \pitchfork T(X \times Y).$$

Now apply exponential adjointness to transform this composite map to the desired $t''_{X,Y} : X \times T(Y) \to T(X \times Y)$.

The strength in its tensorial form t'' (and its twin sister t') will be the main manifestation used in the present paper. The tensorial form has the advantage not of mentioning $\pitchfork$ at all, so makes sense even for endofunctors on monoidal categories that are not closed; this has been exploited e.g. in [17] and [3].

3. Strong monads; partial linearity

The composite of two strong endofunctors on the cartesian closed $\mathcal{E}$ carries a strength derived from the two given strengths. We give it explicitly for the composite $T \circ T$ (with strength of T given in tensorial form t''), namely it is the "composite" strength in the top line in the diagram (1) below.

There is also an evident notion of strong natural transformation between two strong functors; in terms of tensorial strength, and for the special case of the natural transformation $\mu : T \circ T \Rightarrow T$, this means that (not only the naturality squares but also) the top square in (1) commutes, for all X, Y.

The identity functor $I : \mathcal{E} \to \mathcal{E}$ has identity maps id_X as its tensorial strength. The bottom square in (1) expresses that the natural transformation $\eta : I \Rightarrow T$ is strongly natural.

The notion of composite strength, and of strong natural transformation, are equivalent to the classical notion of strength/enrichment (in terms of st), in so far as endofunctors on $\mathcal{E}$ goes, see [6] (and more generally, for functors between categories enriched over $\mathcal{E}$, provided they are tensored over $\mathcal{E}$, cf. [10]).

Here is the diagram which expresses the strength of $\mu : T^2 \Rightarrow T$ (upper part), and of $\eta : I \Rightarrow T$ (lower part):

$$\begin{tikzcd}
X \times T^2Y \arrow[r, "t''_{X,TY}"] \arrow[d, "X \times \mu_Y"'] & T(X \times TY) \arrow[r, "T(t''_{X,Y})"] & T^2(X \times Y) \arrow[d, "\mu_{X \times Y}"] \\
X \times TY \arrow[rr, "t''_{X,Y}"'] & & T(X \times Y) \\
X \times Y \arrow[u, "X \times \eta_Y"] \arrow[rr, "id"'] & & X \times Y \arrow[u, "\eta_{X \times Y}"']
\end{tikzcd} \tag{1}$$

One can equivalently formulate strength of composite functors, and strength of natural transformations, in terms of cotensorial strength, cf. [8]; this will not explicitly play a role here. But one consequence of the cotensorial strength will be important: if (C, γ) is an algebra for the monad T, and X is any object in $\mathcal{E}$, then $X \pitchfork C$ carries a canonical structure of algebra for T, with structure map the composite

$$T(X \pitchfork C) \xrightarrow{\lambda_{X,C}} X \pitchfork T(C) \xrightarrow{X \pitchfork \gamma} X \pitchfork C.$$

The T-algebra $X \pitchfork C$ thus constructed actually witnesses that the category $\mathcal{E}^T$ of T-algebras is *cotensored* over $\mathcal{E}$, cf. e.g. [1] II.6.5.

The tensorial strength makes possible a description (cf. [9]) of the crucial notion of *partial T-linearity* (recall the use of the phrase *T-linear* as synonymous with *T-algebra morphism*). Let (B, β) and (C, γ) be T-algebras, and let $X \in \mathcal{E}$ be an arbitrary object. Then a map $f : X \times B \to C$ is called *T-linear in the second variable* (or *2-linear*, if T can be understood from the context) if the following pentagon commutes:

$$\begin{tikzcd}
X \times T(B) \arrow[r, "t''_{X,B}"] \arrow[d, "X \times \beta"'] & T(X \times B) \arrow[r, "T(f)"] & T(C) \arrow[d, "\gamma"] \\
X \times B \arrow[rr, "f"'] & & C.
\end{tikzcd} \tag{2}$$

In completely analogous way, one describes the notion of 1-linearity of a map $A \times Y \to C$, where A and C are equipped with algebra structures: just apply $t'_{A,Y}$. Finally, a map $A \times B \to C$ is called *bilinear* if it is both 1-linear and 2-linear.

The notion of 2-linearity of $f : X \times B \to C$ can equivalently, and perhaps more intuitively, be described: the exponential transpose of f, i.e. the map $\hat{f} : B \to X \pitchfork C$, is T-linear (= a morphism of T-algebras), with T-structure on $X \pitchfork C$ as given above in terms of the cotensorial strength λ.

We shall prove

3.1. Proposition. *The map $t''_{X,Y} : X \times T(Y) \to T(X \times Y)$ is 2-linear; and it is in fact initial among 2-linear maps from $X \times T(Y)$ into T-algebras. (Similarly, $t'_{X,Y}$ is 1-linear, and is initial among 1-linear maps from $T(X) \times Y$ into T-algebras.)*

Proof. We first argue that $t''_{X,Y}$ is 2-linear. This means by definition that a certain pentagon commutes; for the case of $t''_{X,Y}$, this is precisely the (upper part of) the diagram (1) expressing that μ is strongly natural.

To prove that $t''_{X,Y}$ is initial among 2-linear maps: Given a 2-linear $f : X \times T(Y) \to B$, where $B = (B, \beta)$ is a T-algebra. We must prove unique existence of a T-linear $\overline{f}$ making the right hand triangle in the following diagram commutative:

$$\begin{tikzcd}
X \times Y \arrow[r, "\eta_{X\times Y}"] \arrow[d, "X \times \eta_Y"'] & T(X \times Y) \arrow[d, "\overline{f}"] \\
X \times TY \arrow[ur, "t''_{X,Y}"'] \arrow[r, "f"'] & B.
\end{tikzcd}$$

Since the upper triangle commutes (strong naturality of η, cf. the lower part of (1)), we see that $\overline{f}$ necessarily is the (unique) linear extension over $\eta_{X\times Y}$ of

$$X \times Y \xrightarrow{X \times \eta_Y} X \times TY \xrightarrow{f} B,$$

thus $\overline{f}$ is, by standard monad theory, the composite of the three last arrows in

$$X \times TY \xrightarrow{t''_{X,Y}} T(X \times Y) \xrightarrow{T(X \times \eta_Y)} T(X \times TY) \xrightarrow{Tf} TB \xrightarrow{\beta} B. \tag{3}$$

To prove that this $\overline{f}$ does indeed solve the posed problem, we have to prove that the total composite in (3) equals f. Composing forwards, this follows from the equations

$$\begin{array}{rcll}
t''_{X,Y}.T(X \times \eta_Y).Tf.\beta & = & X \times T(\eta_Y).t''_{X,TY}.Tf.\beta & \text{by naturality of } t'' \text{ w.r.to } \eta_Y \\
& = & X \times T(\eta_Y).X \times \mu_Y.f & \text{by the assumed 2-linearity of } f;
\end{array}$$

but this gives f, by a monad law.

It is a direct consequence of the definitions that postcomposing a 2-linear map $f : X \times B \to C$ with a linear map again yields a 2-linear map. Likewise, precomposing f with a map of the form $X \times g$, with g linear, again yields a 2-linear map. Similarly for 1-linearity and bilinearity.

4. Partially linear extensions

The universal property of η_X (making $T(X)$ the "free T-algebra on X") was mentioned in Section 1. There are similar universal properties of $X \times \eta_Y$ and of $\eta_X \times Y$:

4.1. Proposition. *Let (B, β) be a T-algebra. To any $f : X \times Y \to B$, there exists a unique 2-linear $\overline{f} : X \times T(Y) \to B$ making the triangle*

$$\begin{tikzcd} X \times TY \arrow[r, "\overline{f}"] & B \\ X \times Y \arrow[u, "X \times \eta_Y"] \arrow[ur, "f"'] & \end{tikzcd}$$

commute. (There is a similar universal property of $\eta_X \times Y : X \times Y \to T(X) \times Y$ for 1-linear maps.)

Proof. In essence, this follows by passing to the exponential transpose $\hat{f} : Y \to X \pitchfork B$ of f, and then using the universal property of η_Y. More explicitly, there are natural bijective correspondences

$$\hom(X \times Y, B) \cong \hom(X, Y \pitchfork B) \cong \hom_T(T(X), Y \pitchfork B)$$

(where the second occurrence of $Y \pitchfork B$ is the cotensor $Y \pitchfork B$ in $\mathcal{E}^T$, and where the second bijection is induced by precomposition by η_X); finally, the set $\hom_T(T(X), Y \pitchfork B)$ is in bijective correspondence with the set of 1-linear maps $T(X) \times Y \to B$, by [9] Proposition 1.3 (i).

A direct description of $\overline{f}$ is

$$X \times TY \xrightarrow{t''_{X,Y}} T(X \times Y) \xrightarrow{T(f)} TB \xrightarrow{\beta} B. \tag{4}$$

For, the displayed map is 2-linear, being a composite of the 2-linear $t''_{X,Y}$ and the linear maps $T(f)$ and β; and its precomposition with $X \times \eta_Y$ is easily seen to be f, using that η is a strong natural transformation.

Similarly, an explicit formula for the 1-linear extension of $f : X \times Y \to B$ is

$$T(X) \times Y \xrightarrow{t'_{X,Y}} T(X \times Y) \xrightarrow{T(f)} TB \xrightarrow{\beta} B. \tag{5}$$

A consequence of Proposition 4.1 is that $t''_{X,Y}$ may be characterized "a posteriori" as the unique 2-linear extension over $X \times \eta_Y$ of $\eta_{X \times Y}$. But note that t'' by itself is described independently of η and μ. Similarly for t'.

Also, $\lambda_{X,Y} : T(X \pitchfork Y) \to X \pitchfork T(Y)$ may be characterized as the unique linear extension over $\eta_{X \pitchfork Y}$ of $X \pitchfork \eta_Y$.

5. The two "Fubini" maps

The map $t'_{X,Y} : T(X)\times Y \to T(X\times Y)$ extends by Proposition 4.1 uniquely over $T(X)\times\eta_Y$ to a 2-linear map $T(X) \times T(Y) \to T(X \times Y)$, which we shall denote $\otimes$ or $\otimes_{X,Y}$. Thus, $\otimes$ is characterized as the unique 2-linear map making the following triangle commute:

$$\begin{tikzcd} TX \times TY \arrow[r, "\otimes"] & T(X \times Y) \\ TX \times Y. \arrow[u, "TX \times \eta_Y"] \arrow[ur, "t'_{X,Y}"'] & \end{tikzcd} \tag{6}$$

Now $t'_{X,Y}$ is itself 1-linear; but this 1-linearity may not be inherited by its 2-linear extension $\otimes$; it will be so if the monad is commutative, in the sense explained in Section 9.

Similarly, $t''_{X,Y} : X \times T(Y) \to T(X \times Y)$ extends uniquely over $\eta_X \times TY$ to a 1-linear $\tilde{\otimes} : T(X) \times T(Y) \to T(X \times Y)$, but the 2-linearity of $t''_{X,Y}$ may not be inherited by its 1-linear extension.

An alternative description of the 2-linear $\otimes$ follows by specializing (4); it is

$$TX \times TY \xrightarrow{t''_{TX,Y}} T(TX \times Y) \xrightarrow{T(t'_{X,Y})} T^2(X \times Y) \xrightarrow{\mu_{X\times Y}} T(X \times Y).$$

This is the description in [6], where $\otimes$ is called $\tilde{\psi}$ and $\tilde{\otimes}$ is called ψ. The notion of commutativity of a strong monad was introduced in loc.cit. by the condition $\psi = \tilde{\psi}$. Similarly, the 1-linear $\tilde{\otimes}$ is the composite

$$TX \times TY \xrightarrow{t'_{X,TY}} T(X \times TY) \xrightarrow{T(t''_{X,Y})} T^2(X \times Y) \xrightarrow{\mu_{X\times Y}} T(X \times Y).$$

Both $\otimes$ and $\tilde{\otimes}$ make the endofunctor T into a monoidal functor. The reason why we call these two maps the "Fubini" maps is that when we below interpret "elements" in $T(X)$ (resp. in $T(Y)$) as (Schwartz) distributions, or as (Radon) measures, then the equality of them is a form of Fubini's Theorem.

6. The "integration" pairing

The following structure is the basis for the interpretation of $T(X)$ as a space of distributions or measures, in the "double dualization" sense of functional analysis.

Formally, the pairing consists in maps $T(X)\times(X \pitchfork B) \to B$, available for any $X \in \mathcal{E}$, and any T-algebra $B = (B, \beta)$. We denote it by a "pairing" bracket $\langle -, - \rangle$. Recall that we have the evaluation map $ev : X \times (X \pitchfork B) \to B$, counit of the adjunction $(X \times -) \dashv (X \pitchfork -)$. Then the pairing is the 1-linear extension of it over $\eta_X \times B$, thus

$\langle -, - \rangle$ is 1-linear and makes the following triangle commute:

$$\begin{tikzcd}
T(X) \times (X \pitchfork B) \arrow[r, "\langle -, - \rangle"] & B \\
X \times (X \pitchfork B). \arrow[u, "\eta_X \times B"] \arrow[ur, "ev"']
\end{tikzcd}$$

Thus, the bracket is the composite (composing from left to right) $t'_{X, X \pitchfork B}.T(ev).\beta$.

If $\phi : X \to B$ is an actual map in $\mathcal{E}$, it may be identified with a "global" element $\mathbf{1} \to X \pitchfork B$, and we may consider the map $T(X) \to B$ given by $\langle -, \phi \rangle$. This map is then just the T-linear extension over η_X of $\phi : X \to B$.

The bracket ought to be decorated with symbols X and (B, β), and it will quite evidently be natural in $B \in \mathcal{E}^T$ and in $X \in \mathcal{E}$, the latter naturality in the "extranatural" sense, [16] 9.4, which we shall recall. If we use elements to express equations (a well known and rigorous technique, cf. e.g. [11] II.1, even though objects in a category often are said to have no elements), the extraordinary naturality in X is expressed: for any $f : Y \to X$, $\phi \in X \pitchfork B$ and $P \in T(Y)$,

$$\langle T(f)(P), \phi \rangle = \langle P, \phi \circ f \rangle,$$

or more compactly, using f_* to denote $T(f) : T(Y) \to T(X)$ and f^* to denote $f \pitchfork B$, i.e. precomposition with f,

$$\langle f_*(P), \phi \rangle = \langle P, f^*(\phi) \rangle. \tag{7}$$

To prove this equation, we observe that both sides are 1-linear in P, and therefore it suffices (Proposition 4.1) to prove that their precomposition with $\eta_Y \times B$ are equal; this in turn follows from naturality of η w.r.to f, and the extraordinary naturality of ev.

An alternative notation for the pairing:

$$\langle P, \phi \rangle = \int_X \phi \, dP = \int_X \phi(x) \, dP(x),$$

is motivated by the interpretation of $P \in T(X)$ as a measure (or as a Schwartz distribution) on X, and of $\langle P, \phi \rangle$ as the integral of the function $\phi : X \to B$ w.r.to the measure P, or, as the value of the distribution on the "test function" ϕ. In this context, the main case is where B is the "space of scalars", but any other "vector space" (= T-algebra) B may be the value space for test functions.

7. The "semantics" map

By this, we understand the exponential transpose of the pairing map $T(X) \times (X \pitchfork B) \to B$; it is thus a map

$$T(X) \xrightarrow{\tau} (X \pitchfork B) \pitchfork B.$$

In this context, the T-algebra B may be called "the dualizing object". We could (should) decorate τ with symbols X and/or B, if needed. It is natural in $X \in \mathcal{E}$ (by extra-naturality of the pairing in the variable X), and it is natural in the dualizing object $B \in \mathcal{E}^T$, by the naturality of the pairing in $B \in \mathcal{E}^T$. Why does it deserve the name "semantics"?

Let us for a moment speak about $\mathcal{E}$ as if it were the category of sets. Thus, an element $P \in T(X)$ gets by τ interpreted as an X-ary operation that makes sense for any T-algebra B, and is preserved by morphisms of T-algebras $B \to B'$ in $\mathcal{E}^T$. Here, we are talking about elements in $X \pitchfork B$ as "X-tuples of elements in B", and functions $X \pitchfork B \to B$ are thus "X-ary operations" that, given an X-tuple in B, returns a single element in B. Thus, $\tau_B(P)$ is the interpretation of the "syntactic" (formal) X-ary operation P as an actual X-ary operation on the underlying "set" of B; or, $\tau_B(P)$ is the *semantics* of P in B, and it is preserved by any morphism of T-algebras. (This connection between monads on the category of sets, and universal algebra, - or equivalently, the idea of viewing monads on the category of sets as infinitary algebraic theories (in the sense of Lawvere) - goes back to the early days of monad theory in the mid 60's (Linton, Manes, Beck, Lawvere, ...). Our theory demonstrates that this connection makes sense for strong monads on cartesian closed categories.

Part of the aim of the present study has been to demonstrate that the notion of an extensive quantity P *distributed* over a given space X can be encoded in terms of a strong monad T, but that it need not be encoded in the "double dualization" paradigm of functional analysis, like the double dual $(X \pitchfork B) \pitchfork B$. In slogan form, "*liberate distributions from the yoke of double dualization!*". As a slogan, we shall partially ourselves contradict it, by arguing that "there are enough test functions, provided we let B vary" (cf. Section 11 below). On the other hand, one may push the slogan too far, if one reformulates it into "*liberate syntax from the yoke of semantics!*", which is only occasionally a fruitful endeavour. For, we hold that many distributions, say a distribution of a mass over a space (e.g. the distribution of tomato on a pizza) is not a syntactic entity, and does not in itself depend on the concept of double dualization. Mass distributions will be considered in Section 13 below.

Remark. The τ considered here (with B fixed) is actually a morphism of monads: even without an algebra structure on $B \in \mathcal{E}$, the functor $X \mapsto (X \pitchfork B) \pitchfork B$ is a strong monad on $\mathcal{E}$; if T is another strong monad on T, we have the result that: to give B a T-algebra structure is equivalent to give a strong monad morphism $\tau : T \Rightarrow (- \pitchfork B) \pitchfork B$, cf. [7]. The unit of the double-dualization monad is the exponential transpose of the evaluation map $X \times (X \pitchfork B) \to B$. It is the "Dirac delta map" $\delta_X : X \to (X \pitchfork B) \pitchfork B$. In the category of sets: for $x \in X$, $\delta_X(x)$ is the map $(X \pitchfork B) \to B$ which consists in evaluating at x.

Since τ is the exponential transpose of the pairing, and the pairing was defined as 1-linear extension of evaluation, it follows that τ is the linear extension of the map δ, in formula

$$\tau_X \circ \eta_X = \delta_X, \tag{8}$$

or briefly, $\tau \circ \eta = \delta$, (B being a fixed T-algebra here). In [13], we even denoted $\tau_X(x)$ by δ_x.

8. The object of T-linear maps

The following construction goes back to [2]. Let (B, β) and (C, γ) be two T-algebras, with T a strong monad on $\mathcal{E}$. We assume that $\mathcal{E}$ has equalizers. Then we can out of the object $B \pitchfork C$ carve a subobject $B \pitchfork_T C$ "consisting of" the T-linear maps; precisely, it is the equalizer

$$B \pitchfork_T C \longrightarrow B \pitchfork C \underset{st.(TB \pitchfork \gamma)}{\overset{\beta \pitchfork C}{\rightrightarrows}} T(B) \pitchfork C$$

in the lower map on the right, we compose forwards, so it is, in full

$$B \pitchfork C \xrightarrow{st_{B,C}} T(B) \pitchfork T(C) \xrightarrow{T(B) \pitchfork \gamma} T(B) \pitchfork C,$$

expressing internally that a map $f : B \to C$ is a morphism of algebras if two particular maps $T(B) \to C$ are equal.

Now there is an evident notion of when a subobject $A' \subseteq A$ of (the underlying object of) an algebra (A, α) is a subalgebra. The object $B \pitchfork C$ inherits an algebra structure from that of C, using the cotensorial strength λ, so one may ask, is $B \pitchfork_T C$ a subalgebra ? This is not in general the case. If B and C are groups, the set of group homomorphisms $B \to C$ is not a subgroup of the group of all maps $B \to C$; it is so, however, for *commutative* groups. This leads to the topic of the next Section, commutativity of a monad.

9. Commutative monads

The notion of *commutative* monad (a strong monad with a certain *property*) was introduced in [6]: it is a strong monad for which the two "Fubini maps" $\otimes$ and $\tilde{\otimes}$ agree. There are several equivalent conditions describing commutativity, and we summarize these in the Theorem at the end of the section.

9.1. PROPOSITION. *Let T be commutative. Let $B = (B, \beta)$ and $C = (C, \gamma)$ be T-algebras, and assume that $f : X \times B \to C$ is 2-linear. Then its 1-linear extension $\overline{f} : T(X) \times B \to C$ over $\eta_X \times B$ is 2-linear (hence bilinear). (Similarly for 2-linear extensions of 1-linear maps.)*

Proof. To prove 2-linearity of $\overline{f}$ means to prove commutativity of the following diagram (where the bottom line is $\overline{f}$, according to (5))

$$\begin{tikzcd}
TX \times TB \arrow[r, "t''"] \arrow[d, "TX \times \beta"'] & T(TX \times B) \arrow[r, "T(t')"] & T^2(X \times B) \arrow[r, "T^2 f"] & T^2 C \arrow[r, "T\gamma"] & TC \arrow[d, "\gamma"] \\
TX \times B \arrow[rr, "t'"'] & & T(X \times B) \arrow[r, "Tf"'] & TC \arrow[r, "\gamma"'] & C.
\end{tikzcd}$$

The assumed 2-linearity of f is expressed by the equation (composing from left to right)

$$t''.Tf.\gamma = (X \times \beta).f \tag{9}$$

and the assumed commutativity of the monad is expressed by the equation

$$t''.Tt'.\mu_{X \times B} = t'.Tt''.\mu_{X \times B}, \tag{10}$$

using the explicit formulae for $\otimes$ and $\tilde{\otimes}$ at the end of Section 5. We now calculate, beginning with the clockwise composite in the diagram:

$$\begin{array}{rcll}
t''.Tt'.T^2 f.T\gamma.\gamma & = & t''.Tt'.T^2 f.\mu_C.\gamma & \text{by an algebra law} \\
& = & t''.Tt'.\mu_{X \times B}.Tf.\gamma & \text{by naturality of } \mu \\
& = & t'.Tt''.\mu_{X \times B}.Tf.\gamma & \text{by (10), commutativity of the monad} \\
& = & t'.Tt''.T^2 f.\mu_C.\gamma & \text{by naturality of } \mu \\
& = & t'.Tt''.T^2 f.T\gamma.\gamma & \text{by an algebra law} \\
& = & t'.T(X \times \beta).Tf.\gamma & \text{by (9), 2-linearity of } f \\
& = & (TX \times \beta).t'.Tf.\gamma & \text{by naturality of } t',
\end{array}$$

which is the counterclockwise composite of the diagram. This proves the Proposition.

From Proposition 9.1 now immediately follows that when the monad is commutative, then $\otimes : T(X) \times T(Y) \to T(X \times Y)$ is bilinear; for, it was constructed as the 2-linear extension of the 1-linear $t' : T(X) \times Y \to T(X \times Y)$.

It follows rather easily that, conversely, bilinearity of $\otimes$ implies commutativity of the monad, cf. [9] Proposition 1.5. Thus, also, the extension properties in Proposition 9.1 imply commutativity.

From Proposition 9.1 also follows that when the monad is commutative, the pairing map $T(X) \times (X \pitchfork B) \to B$ is bilinear; for, it was constructed as the 1-linear extension of the evaluation map $X \times (X \pitchfork B) \to B$, which is 2-linear. By passing to the exponential transpose, this 2-linearity implies that its transpose, i.e. the semantics map $\tau : T(X) \to (X \pitchfork B) \pitchfork B$, factors through the subobject $(X \pitchfork B) \pitchfork_T B \subseteq (X \pitchfork B) \pitchfork B$.

For commutative T, and for B and C arbitrary T-algebras, the subobject $B \pitchfork_T C \subseteq B \pitchfork C$ is actually a subalgebra (and in fact provides $\mathcal{E}^T$ with structure of a closed category), cf. [8].

The two last implications argued for here go the other way as well: bilinearity of the pairing implies commutativity; $B \pitchfork_T C$ a subalgebra of $B \pitchfork C$ (for all algebras B, C) implies commutativity. The proofs are not difficult, and will be omitted.

Finally, from [9], Proposition 1.5 (v), we have the equivalence of commutativity with "μ is monoidal"; also, commutativity implies that the monoidal structure $\otimes$ on the endofunctor T is symmetric monoidal, cf. [6] Theorem 3.2). We summarize:

9.2. THEOREM. *Let $T = (T, \eta, \mu)$ be a strong monad on $\mathcal{E}$. Then t.f.a.e., and define the notion of commutativity of T:*

1) Fubini's Theorem holds, i.e. the maps $\otimes$ and $\tilde{\otimes} : T(X) \times T(Y) \to T(X \times Y)$ agree, for all X and Y in $\mathcal{E}$.

2) The map $\otimes : T(X) \times T(Y) \to T(X \times Y)$ is bilinear for all X and Y.

3) For (B, β) and (C, γ) algebras for T, and X arbitrary, the 1-linear extension $T(X) \times B \to C$ of a 2-linear map $X \times B \to C$ is 2-linear (hence bilinear). (Similarly for 2-linear extensions of 1-linear maps.)

4) For (B, β) and (C, γ) algebras for T, the subspace $B \pitchfork_T C \subseteq B \pitchfork C$ is a sub-algebra.

5) For (B, β) an algebra for T, and X arbitrary, the "semantics" map $\tau : T(X) \to (X \pitchfork B) \pitchfork B$ factors through the subspace $(X \pitchfork B) \pitchfork_T B$.

6) For (B, β) an algebra for T, and X arbitrary, the pairing map $T(X) \times (X \pitchfork B) \to B$ is bilinear.

7) The monoidal structure $\otimes$ on the endofunctor T is symmetric monoidal, and with respect to this monoidal structure $\otimes$, the natural transformation $\mu : T \circ T \Rightarrow T$ is monoidal (hence the monad is a symmetric monoidal monad).

Recall the universal property of $X \times \eta_Y : X \times Y \to X \times T(Y)$ and of $\eta_X \times Y : X \times Y \to T(X) \times Y$ (Proposition 4.1). We have an analogous property for $\eta_X \times \eta_Y : X \times Y \to T(X) \times T(Y)$:

9.3. PROPOSITION. *Let T be a commutative monad. Let $B = (B, \beta)$ be a T-algebra. Then any $f : X \times Y \to B$ extends uniquely over $\eta_X \times \eta_Y$ to a bilinear $\overline{f} : T(X) \times T(Y) \to B$.*

PROOF. The extension is performed in two stages, along $\eta_X \times Y$, and then along $T(X) \times \eta_Y$. The first extension is unique, as a 1-linear map, the second is unique as a 2-linear map. However, the 1-linearity of the first extension is preserved by the second extension, using clause 3) in the Theorem. The proof of uniquensess is similar.

Recall from Proposition 3.1 that $t'_{X,Y} : T(X) \times Y \to T(X \times Y)$ is an initial 1-linear map into T-algebras, and similarly $t''_{X,Y} : X \times T(Y) \to T(X \times Y)$ is an initial 2-linear map. These properties join hands when T is commutative:

9.4. PROPOSITION. *Let T be a commutative monad. Then $\otimes : T(X) \times T(Y) \to T(X \times Y)$ is initial among bilinear maps from $T(X) \times T(Y)$ into T-algebras.*

(Thus, $T(X \times Y)$ may be denoted $T(X) \otimes T(Y)$. Tensor products $A \otimes B$ for general T-algebras A and B may not exist, this depends on sufficiently many good coequalizers in $\mathcal{E}^T$.)

Proof. Let $B = (B, \beta)$ be a T-algebra, and let $f : T(X) \times T(Y) \to B$ be bilinear. We must prove unique existence of a linear $\overline{f}$ making the right hand triangle in the following diagram commutative:

$$\begin{tikzcd} X \times Y \arrow[r, "\eta_{X\times Y}"] \arrow[d, "\eta_X \times \eta_Y"'] & T(X \times Y) \arrow[d, "\overline{f}"] \\ TX \times TY \arrow[ur, "\otimes"'] \arrow[r, "f"'] & B. \end{tikzcd}$$

Since the upper triangle commutes, we see that $\overline{f}$ necessarily is the (unique) linear extension over $\eta_{X\times Y}$ of $(\eta_X \times \eta_Y).f$ (composing forwards). To prove that this $\overline{f}$ does indeed solve the posed problem: The two maps $\otimes.\overline{f}$ and f are 2-linear $TX \times TY \to B$, so by the universal property of $TX \times \eta_Y$, to see that they agree, it suffices to see that the the agree when precomposed with $TX \times \eta_Y$. Both of these two maps are 1-linear maps $TX \times Y \to B$, so by the universal property of $\eta_X \times Y$, it suffices to see that the two maps $X \times Y \to B$ which one gets by precomposition with $\eta_X \times Y$ are equal. But they are both equal to $(\eta_X \times \eta_Y).f$, by construction of $\overline{f}$.

10. Convolution; the scalars $R := T(\mathbf{1})$

We henceforth assume that the monad T is commutative. If $m : X \times Y \to Z$ is a map in $\mathcal{E}$, we can use $\otimes$ to manufacture a map $T(X) \times T(Y) \to T(Z)$, namely the composite

$$TX \times TY \xrightarrow{\otimes} T(X \times Y) \xrightarrow{T(m)} TZ,$$

called the *convolution* along m. It is bilinear, since $\otimes$ is so, and $T(m)$ is linear. If $m : Y \times Y \to Y$ is associative, it follows from general equations for monoidal functors that the convolution along m is associative. If $e : \mathbf{1} \to Y$ is a two sided unit for the semigroup Y, m, then it gives rise to a unit for the convolution semigroup $T(Y)$, which thus acquires a monoid structure. If m is furthermore commutative, then so is the convolution monoid; this latter fact depends on T being a *symmetric* monoidal functor, which it is in our case, cf. clause 7) in Theorem 9.2.

Similarly, if a monoid (Y, m, e) acts on an object X in an associative and unitary way, then the convolution monoid $T(Y)$ acts in an associative and unitary way on $T(X)$.

Now $\mathbf{1}$ carries a unique (and trivial) monoid structure, and this trivial monoid acts on any object X, by the projection map $pr : \mathbf{1} \times X \to X$. It follows that $T(\mathbf{1})$ carries a monoid structure, and that this monoid acts on any $T(X)$. Let us for the moment call this action of the *convolution action* on $T(X)$. It is thus the composite

$$T\mathbf{1} \times TX \xrightarrow{\otimes} T(\mathbf{1} \times X) \cong TX,$$

it is bilinear, unitary and associative. It is the 2-linear extension over $T\mathbf{1} \times \eta_X$ of $t'_{\mathbf{1},X}$ (followed by the isomorphism $T(\mathbf{1} \times X) \cong T(X)$); this follows from (6). There is a similar description of a right action of $T(\mathbf{1})$ on $T(X)$, but $T(\mathbf{1})$ is commutative, and these two actions agree.

There is also an action by $T(\mathbf{1})$ on any T-algebra B, derived from the pairing, namely using the canonical isomorphism $i : B \to (\mathbf{1} \pitchfork B)$,

$$T\mathbf{1} \times B \xrightarrow{T\mathbf{1} \times i} T\mathbf{1} \times (\mathbf{1} \pitchfork B) \xrightarrow{\langle -, - \rangle} B.$$

Let us call this the *pairing* action.

10.1. Proposition. *For a T-algebra of the form $T(X)$, the convolution action of $T(\mathbf{1})$ agrees with the pairing action.*

Proof. Consider the diagram (where all the vertical maps are isomorphisms)

$$\begin{tikzcd}
T\mathbf{1} \times TX \arrow[r, "t'"] \arrow[d, "T\mathbf{1} \times i"'] & T(\mathbf{1} \times TX) \arrow[r, "T(t'')"] \arrow[d, "T(\mathbf{1} \times i)"'] & T^2(\mathbf{1} \times X) \arrow[r, "\mu"] \arrow[d, "T^2(pr)"] & T(\mathbf{1} \times X) \arrow[d, "T(pr)"] \\
T\mathbf{1} \times (\mathbf{1} \pitchfork TX) \arrow[r, "t'"'] & T(\mathbf{1} \times (\mathbf{1} \pitchfork TX)) \arrow[r, "T(ev)"'] & T^2X \arrow[r, "\mu"'] & TX.
\end{tikzcd}$$

The top line is the explicit form of (the 1-linear version of) $\otimes$, so the clockwise composite is the convolution action. The counterclockwise composite is the pairing action. The left hand and right hand squares commute by naturality. The middle square comes about by applying T to the square

$$\begin{tikzcd}
\mathbf{1} \times TX \arrow[r, "t''"] \arrow[d, "\mathbf{1} \times i"'] & T(\mathbf{1} \times X) \arrow[d, "T(pr)"] \\
\mathbf{1} \times (\mathbf{1} \pitchfork TX) \arrow[r, "ev"'] & T(X).
\end{tikzcd}$$

We claim that this square commutes. In fact, both composites are 2-linear, so it suffices to see that we have commutativity when we precompose by $\mathbf{1} \times \eta_X$. This commutativity follows by naturality of η, together with the fact that i and pr are the exponential transposes of each other.

We conclude that the total diagram commutes, proving the desired equality of the two actions on an algebra $T(X)$ (the convolution pairing has only been defined for such algebras).

(We note that the proof does not depend on commutativity of T, provided we take the 1-linear version of $\otimes$, as we did. Also, it is valid in any symmetric monoidal closed category, replacing $\mathbf{1}$ by the unit object, and *pr* by the primitively given unit isomorphism $I \otimes X \to X$.)

Since pairing is natural in $B \in \mathcal{E}^T$, it follows that the pairing action has the property that any morphism in $\mathcal{E}^T$, i.e. any T-linear map $B \to C$, is equivariant. The Proposition then has the following Corollary:

10.2. PROPOSITION. *Any T-linear map $T(X) \to T(Y)$ is equivariant for the convolution action by $T(\mathbf{1})$.*

The convolution action on the free T-algebra $T(X)$ by the monoid $T(\mathbf{1})$ is an associative unitary action, because $\otimes$ makes T into a monoidal functor, as we remarked; one can probably prove that the pairing action by $T(\mathbf{1})$, defined for all T-algebras, is likewise unitary and associative.

Since the monoid $T(\mathbf{1})$ acts in a unitary, associative, and bilinear way on any $T(X)$, and T-linear maps preserve the action, it is reasonable to think of $T(\mathbf{1})$ as the (multiplicative) monoid of *scalars*, and to give it a special notation. We denote it by R;

$$R := T(\mathbf{1});$$

we denote its multiplication by a dot, and its unit by 1. It is the map $\eta_{\mathbf{1}} : \mathbf{1} \to T(\mathbf{1})$.

(An addition on R will be considered in Section 14 below.)

ON TOTALS For any $X \in \mathcal{E}$, we have a unique map $!_X : X \to \mathbf{1}$ to the terminal object. Composing it with the map $1 : \mathbf{1} \to R = T(\mathbf{1})$ (i.e. by $\eta_{\mathbf{1}}$) gives a map which we denote $1_X : X \to R$, the "map with constant value 1".

Just by the fact that T is a covariant functor, we have a map $\mathrm{tot}_X : T(X) \to R$, namely $T(!_X) : T(X) \to T(\mathbf{1}) = R$, "formation of total". From uniqueness of maps into the terminal, we have for any $f : X \to Y$ that $\mathrm{tot}_X = \mathrm{tot}_Y \circ T(f)$, or in elementwise terms, writing f_* for $T(f)$:

$$\mathrm{tot}(P) = \mathrm{tot}(f_*(P)).$$

We have

$$\mathrm{tot}_X \circ \eta_X = T(!_X) \circ \eta_X = \eta_{\mathbf{1}} \circ !_X \tag{11}$$

by definition of tot and by naturality of η; but the latter map is 1_X. – In particular, we have, in elementwise terms, for $x \in X$,

$$\mathrm{tot}(\eta_X(x)) = 1. \tag{12}$$

10.3. PROPOSITION. *The map $\mathrm{tot}_X : T(X) \to R$ equals the map $\langle -, 1_X \rangle$.*

Expressed in terms of elements, for $P \in T(X)$,

$$\mathrm{tot}_X(P) = \langle P, 1_X \rangle = \int_X 1_X \, dP \in R$$

(recalling the alternative "integral" notation for the pairing).

PROOF. The two maps $T(X) \to T(\mathbf{1})$ to be compared are T-linear, so it suffices to prove that their precomposites with η_X agree. For tot_X, (11) shows that we get 1_X. For $\langle -, 1_X\rangle$, the precomposition again gives 1_X (cf. the description in Section 6, immediately after the construction of the bracket).

A consequence of the naturality of $\otimes$ w.r.to the maps $! : X \to \mathbf{1}$ and $! : Y \to \mathbf{1}$ is that

$$\mathrm{tot}(P \otimes Q) = \mathrm{tot}(P) \cdot \mathrm{tot}(Q), \tag{13}$$

where the dot denotes the product in $R = T(\mathbf{1})$. Hence also, the convolution $P * Q$ of P and Q (along any map) satisfies $\mathrm{tot}(P * Q) = \mathrm{tot}(P) \cdot \mathrm{tot}(Q)$.

11. Schwartz distributions of compact support

There exist cartesian closed categories $\mathcal{E}$ which contain the category of smooth manifolds, and also contain the category of convenient vector spaces (with smooth, not necessarily linear, maps), for instance the category Lip^∞ of Frölicher and Kriegl, [5]. In the latter, the (convenient) vector space of Schwartz distributions of compact support on a manifold X is represented as $(X \pitchfork \mathbb{R}) \pitchfork_T \mathbb{R}$ for a suitable strong monad T (with $T(\mathbf{1}) = \mathbb{R}$), see [5] Theorem 5.1.1; this was one of the motivations for the present development. We shall not use this material from functional analysis, except for terminology and motivation. Thus, with $R = T(\mathbf{1})$, $X \pitchfork R$ will be called the *T-linear space of (non-compact) test functions* on X, and $(X \pitchfork R) \pitchfork_T R$ is then the space of T-linear functionals on this space of test functions; T-linearity in the example unites the two aspects, the algebraic ($\mathbb{R}$-linear), and the topological/bornological. So in the [5] case of Lip^∞, $(X \pitchfork R) \pitchfork_T R$ is the *space of continuous linear functionals on the space of test functions*, which is how Schwartz distributions[2] (of compact support) are defined, for X a smooth manifold.

Recall that for a commutative monad T on $\mathcal{E}$, we have the "semantics" map $\tau_X : T(X) \to (X \pitchfork B) \pitchfork_T B$, for any T-algebra B, in particular, we have such a τ_X for $B = R = T(\mathbf{1})$,

$$T(X) \to (X \pitchfork R) \pitchfork_T R. \tag{14}$$

The functor $X \mapsto (X \pitchfork R) \pitchfork_T R$ is in fact itself a strong monad on $\mathcal{E}$, (not necessarily commutative), and τ is a strong natural transformation, and it is even a morphism of monads. (More generally, these assertions hold even when R is replaced by any other T-algebra B.)

In the case of the $(\mathcal{E}, T)$ of Frölicher and Kriegl [5], the map $\tau : T(X) \to (X \pitchfork R) \pitchfork_T R$ is an isomorphism for many X, in particular for smooth manifolds, one may express this property verbally by saying that $T(X)$ is *reflexive* w.r.to R in the category of T-algebras. For any X in the $\mathcal{E}$ of [5], it is a monic map; this one may express verbally: "there are enough $\mathbb{R}$-valued valued test functions on X to test equality of elements in $T(X)$".

[2]often with $\mathbb{C}$ rather than $\mathbb{R}$ as dualizing object.

There are many examples of $(\mathcal{E}, T)$ where there are not enough $R = T(\mathbf{1})$-valued test functions; for instance, there are interesting monads T with $T(\mathbf{1}) = \mathbf{1}$ ("affine monads", see Section 16 below).

However, we have the option of choosing other T-algebras B as "dualizing object". Then there are enough test functions, in the following sense

11.1. Proposition. *For any $X \in \mathcal{E}$, there exists a T-algebra B so that*

$$\tau_X : T(X) \to (X \pitchfork B) \pitchfork_T B$$

is monic. The algebra B may be chosen to be free, i.e. of the form $T(Y)$, in fact $T(X)$ suffices.

Proof. Take $B := T(X)$. We claim that the following diagram commutes:

$$\begin{tikzcd} T(X) \arrow[r, "\tau"] \arrow[dr, "="'] & (X \pitchfork T(X)) \pitchfork_T T(X) \arrow[d, "ev_\eta"] \\ & T(X) \end{tikzcd}$$

where ev_η denotes the map "evaluate at the global element $\eta_X \in X \pitchfork T(X)$". For, the maps to be compared are linear, so it suffices to see that they agree when precomposed with η_X. Now η_X followed by τ is the Dirac delta map $\delta : X \to (X \pitchfork B) \pitchfork B$ (cf. (8)), and δ followed by "evaluate at $\phi \in X \pitchfork B$" is the map ϕ itself, by general "λ-calculus". So both maps give η_X when precomposed by η_X.

So not only are there enough test functions; for given X, even *one* single test function suffices, namely $\eta_X : X \to T(X)$.

12. Multiplying a distribution by a function

The action on distributions by (scalar valued) functions, to be described here, has as applications the notion of *density* of one distribution w.r.to another, and also the notion of *conditional* probability distribution.

Except for the subsection on "totals", the considerations so far are valid for any symmetric monoidal closed category $\mathcal{E}$, not just for a cartesian closed categories. The following, however, utilizes the universal property of the cartesian product.

Now that we are using that $\times$ is cartesian product, it becomes more expedient to use elementwise notation, since this will better allow us to "repeat variables", like in the following. It is a description of a "pointwise" monoid structure on $X \pitchfork R$ derived from the monoid structure of R. The multiplication is thus a map $(X \pitchfork R) \times (X \pitchfork R) \to X \pitchfork R$ which in elementwise terms may be described, for ϕ and ψ in $X \pitchfork R$, and for $x \in R$,

$$(\phi \cdot \psi)(x) := \phi(x) \cdot \psi(x).$$

The unit is the previously described 1_X, in elementwise terms the map $x \mapsto 1 \in R$. Similarly, the action of R on $T(Y)$ for any Y gives rise to a "pointwise" action of $X \pitchfork R$ on $X \pitchfork T(Y)$.

We shall construct an action of $X \pitchfork R$ on the space $T(X)$. It has as a special case the "multiplication of a (Schwartz-) distribution on X by a scalar valued function $X \to R$" known from classical distribution theory. We let the action be from the right, and denote it by $\vdash$:

$$T(X) \times (X \pitchfork R) \xrightarrow{\vdash} T(X),$$

where $R = T(\mathbf{1})$; the R here cannot be replaced by other T-algebras B.

To construct it, it suffices to construct a map $\rho : X \times (X \pitchfork R) \to T(X)$ and extend it by 1-linearity over $\eta_X \times (X \pitchfork R)$. The map ρ is constructed as the composite (with pr_1 being projection onto the first factor X of the domain)

$$X \times (X \pitchfork T(\mathbf{1})) \xrightarrow{\langle ev, pr_1 \rangle} T(\mathbf{1}) \times X \xrightarrow{t'_{\mathbf{1},X}} T(\mathbf{1} \times X) \cong T(X).$$

(The map ρ is 2-linear, and 2-linearity is preserved by the extension, by Theorem 9.2, so $\vdash$ is bilinear. The 1-linear action $\vdash$ exists also without the commutativity assumption, but then it cannot be asserted to be bilinear.) The first map in this composite depends on the universal property of the cartesian product: a map into a product can be constructed by a pair of maps into the factors.

In elementwise notation, if $x \in X$ and $\phi \in X \pitchfork R$, we therefore have

$$\eta(x) \vdash \phi = \phi(x) \cdot \eta(x) \tag{15}$$

(recalling the description (6) of the action of $T(\mathbf{1})$ on $T(X)$ in terms of t').

In classical distribution theory, the following formula is the *definition* of how to multiply a distribution P by a function ϕ; in our context, it has to be *proved*. (The classical formula is the special case where $Y = \mathbf{1}$, so $T(Y) = R$.)

12.1. Proposition. *Let $P \in T(X)$, $\phi \in X \pitchfork R$ and $\psi \in X \pitchfork T(Y)$. Then*

$$\langle P \vdash \phi, \psi \rangle = \langle P, \phi \cdot \psi \rangle.$$

Proof. Since both sides of the claimed equation depend in a T-linear way on P, it suffices to prove the equation for the case where $P = \eta(x)$ for some $x \in X$. We calculate the left hand side, using (15):

$$\langle \eta(x) \vdash \phi, \psi \rangle = \langle \phi(x) \cdot \eta(x), \psi \rangle = \phi(x) \cdot \langle \eta(x), \psi \rangle = \phi(x) \cdot \psi(x),$$

using for the middle equality that T-linearity implies equivariance w.r.to the action by R (Proposition 10.2). The right hand side of the desired equation similarly calculates

$$\langle \eta(x), \phi \cdot \psi \rangle = (\phi \cdot \psi)(x),$$

which is likewise $\phi(x) \cdot \psi(x)$, because of the pointwise character of the action of $X \pitchfork R$ on $X \pitchfork T(Y)$.

12.2. Corollary. *The pairing $\langle P, \phi\rangle$ (for $P \in T(X)$ and $\phi \in X \pitchfork R$) can be described in terms of $\vdash$ as follows:*

$$\langle P, \phi\rangle = \mathrm{tot}(P \vdash \phi).$$

Proof. Take $Y = \mathbf{1}$ (so $T(Y) = R$), and take $\psi = 1_X$. Then

$$\mathrm{tot}(P \vdash \phi) = \langle P \vdash \phi, 1_X\rangle = \langle P, \phi \cdot 1_X\rangle$$

using Proposition 10.3, and then the Proposition 12.1. But $\phi \cdot 1_X = \phi$.

Combining the "switch" formula in Proposition 12.1 with Proposition 11.1 ("enough test functions"), we can derive properties of the action $\vdash$:

12.3. Proposition. *The action $\vdash$ is associative and unitary.*

Proof. Let ϕ_1 and ϕ_2 be in $X \pitchfork R$ and let $P \in T(X)$. To see that $(P \vdash \phi_1) \vdash \phi_2 = P \vdash (\phi_1 \cdot \phi_2)$, it suffices by Proposition 11.1 to see that for any free T-algebra B and any $\psi \in X \pitchfork B$, we have

$$\langle (P \vdash \phi_1) \vdash \phi_2, \psi\rangle = \langle P \vdash (\phi_1 \cdot \phi_2), \psi\rangle,$$

but this is immediate using Proposition 12.1 three times, and the associative law for the action of the monoid $X \pitchfork R$ on $X \pitchfork B$. The unitary law is proved in a similar way.

To state the following result, we use for simplicity the notation f_* for $T(f)$ (where $f : Y \to X$) and f^* for $f \pitchfork B$, as in (7).

12.4. Proposition. [Frobenius reciprocity] *For $f : Y \to X$, $P \in T(Y)$, and $\phi \in X \pitchfork R$,*

$$f_*(P) \vdash \phi = f_*(P \vdash f^*(\phi)).$$

Proof. Both sides depend in a T-linear way of P, so it suffices to prove it for P of the form $\eta_Y(y)$, for $y \in Y$. We have

$$\begin{aligned}
f_*(\eta_Y(y)) \vdash \phi &= \eta_X(f(y)) \vdash \phi && \text{by naturality of } \eta \\
&= \phi(f(y)) \cdot \eta_X(f(y)) && \text{by (15)} \\
&= \phi(f(y)) \cdot f_*(\eta_Y(y)) && \text{by naturality of } \eta \\
&= f_*\big(\phi(f(y)) \cdot \eta_Y(y)\big) && \text{by equivariance of } f_*\text{, Prop. 10.2} \\
&= f_*\big((\eta_Y(y)) \vdash f^*(\phi)\big) && \text{by } f_* \text{ applied to (15).}
\end{aligned}$$

An alternative proof is by using the "enough test-functions" technique.

Density functions For P and Q in $T(X)$, it may happen that $Q = P \vdash \phi$ for some $\phi \in X \pitchfork R$, in which case one says that Q has a *density* w.r.to P, namely the (scalar-valued) function ϕ. Such ϕ may not exist, and if it does, it may not be unique. In the case of (non-compact) Schwartz distributions, one case is particularly important, namely where X is $\mathbb{R}^n$, and P is Lebesgue measure; then if Q has a density function ϕ w.r.to P, one sometimes identifies the distribution Q with the function ϕ. Such identification, as stressed by Lawvere, leads to loss of the distinction between the covariant character of $T(X)$ and the contravariant character of $X \pitchfork R$, more specifically, between extensive and intensive quantities.

13. Mass distributions, and other extensive quantities

An extensive quantity of a given type m may, according to Lawvere, be modelled mathematically by a covariant functor M from a category $\mathcal{E}$ of spaces to a "linear" or "additive" category $\mathcal{A}$, with suitable structure and properties. In particular, the category $\mathcal{E}$ should be "lextensive" (cartesian closed categories with finite coproducts have this property), and $\mathcal{A}$ should be "linear" (categories of modules over a rig[3] have this property). In such a category, any object is canonically an abelian semigroup. Also, the functor M should take finite coproducts to bi-products, by a certain canonical map.

It is a reasonable mathematical model that the quantity type of *mass* (in the sense of physics) is such a functor, with $\mathcal{A}$ being the category of modules over the rig of non-negative reals. If $M(X)$ denotes the set of possible mass distributions over the space X, then if P_1 and P_2 are masses which are distributed over the space X, one may, almost by physical construction, combine them into one mass, distributed over X; so $M(X)$ acquires an additive structure; similarly, one may re-scale a mass distribution by non-negative reals. The covariant functorality of M has for its germ the idea that for a mass P distributed over a space X, one can form its *total* $\mathrm{tot}(P)$. Such total mass may be something like 100 gram, so is not in itself a scalar $\in \mathbb{R}_+$, but only becomes so after choice of a *unit* mass, like gram, so the scalar is not *canonically* associated to the total mass.

Intensive quantities are often derived from extensive ones as densities, or *ratios* between extensive ones. Thus, "specific weight" is an intensive quantity of type $m \cdot v^{-1}$ (where v is the quantity type of volume (if the space X is a block of rubber, its volume (Lebesgue measure) may vary; so volume is, like mass, a distribution varying over X).

Here, we shall only be concerned with *pure* intensive quantities; they are those of type $m \cdot m^{-1}$ for some extensive quantity type m. For the example of mass, an intensive quantity of type $m \cdot m^{-1}$ over the space X may be identified with a function $X \to \mathbb{R}_+$. Other extensive quantity types may naturally give rise to other spaces of scalars; e.g. the quantity type of electric charge gives, in the simplest mathematical model, rise to the space $\mathbb{R}$ of *all* real numbers (since charge may also be negative), or, in a more sophisticated model, taking into account that charge is considered to be something quantized (so can be counted by integers), the space $\mathbb{Q}$ should in principle be the correct one.

So every quantity type defines in principle its own rig R of scalars. However, it is a remarkable fact that so many give rise to $\mathbb{R}$, as the simplest mathematical model.

In real life, pure quantities come after the extensive physical ones. The theory presented in the preceding Sections follows, however, the mathematical tradition[4], and puts the theory of *pure* quantities as the basis.

So to use Lawvere's concepts, our cartesian closed category $\mathcal{E}$ is the category of spaces; for X a space, $T(X)$ is an assumed notion of "space of *pure* extensive quantities, varying

[3] commutative semiring with unit

[4] There is a, minor, mathematical tradition in the other direction, namely Geometric Algebra, where one *constructs* a ring of scalars *from* geometry.

over X" (the basic primitive notion of the theory presented here); $R = T(\mathbf{1})$ is the space of scalars, and $X \pitchfork R$ is the space of intensive pure quantities varying over X. The functor T may be seen as taking values in the category $\mathcal{A} := \mathcal{E}^T$ of T-algebras; it is not quite a linear category, but at least, all free objects carry an action by the monoid of scalars, and all maps are equivariant with respect to these, cf. Proposition 10.2. (Additive structure in $\mathcal{A}$ will be considered in the next Section.) Pure extensive quantities are to my knowledge not very common in physics; but $T(X)$ = probability distributions on X is an example (except that this $T(X)$ does not carry an additive structure, only a convex-space structure).

The aim of the remainder of the present Section is to demonstrate how e.g. the theory of a non-pure extensive quantity type, like mass, embeds into the theory considered so far; more precisely, to give a framework which vindicates the idea that "as soon as a unit of mass is chosen, there is no difference between the theory of mass distributions, and the theory of distributed *pure* quantities". The framework is the following:

Let T be a commutative monad on $\mathcal{E}$. Consider another strong endofunctor M on $\mathcal{E}$, equipped with an action ν by T,

$$\nu : T(M(X)) \to M(X)$$

strongly natural in X, and with ν satisfying a unitary and associative law. Then every $M(X)$ is a T-linear space by virtue of $\nu_X : T(M(X)) \to M(X)$, and morphisms of the form $M(f)$ are T-linear. Let M and M' be strong endofunctors equipped with such T-actions. There is an evident notion of when a strong natural transformation $\lambda : M \Rightarrow M'$ is compatible with the T-actions, so we have a category of T-actions. The endofunctor T itself is an object in this category, by virtue of μ. We say that M is a *T-torsor* if it is isomorphic[5] to T in the category of T-actions. Note that no particular isomorphism is chosen.

Our contention is that the category of T-torsors is a mathematical model of (not necessarily pure) quantities M of type T (which is the corresponding pure quantity).

The following Proposition expresses that isomorphisms of actions $\lambda : T \cong M$ are determined by $\lambda_{\mathbf{1}} : T(\mathbf{1}) \to M(\mathbf{1})$; in the example, the latter data means: choosing a *unit* of mass.

13.1. Proposition. *If g and $h : T \Rightarrow M$ are isomorphisms of T-actions, and if $g_{\mathbf{1}} = h_{\mathbf{1}} : T(\mathbf{1}) \to M(\mathbf{1})$, then $g = h$.*

Proof. By replacing h by its inverse $M \Rightarrow T$, it is clear that it suffices to prove that if $\rho : T \Rightarrow T$ is an isomorphism of T-actions, and $\rho_{\mathbf{1}} = id_{T(\mathbf{1})}$, then ρ is the identity transformation. As a morphism of T-actions, ρ is in particular a *strong* natural transformation, which implies that the right hand square in the following diagram commutes for

[5] If $\mathcal{E}$ is a topos, the notion og being *locally* isomorphic makes sense, and our considerations work for M *locally* isomorphic to T; this justifies the "torsor"-terminology. It is quite realistic that non-pure extensive quantities are only *locally* isomorphic to T, since in cosmology, there is not really a physical reasonable way to say that two masses, far away from each other in time and space, are of equal size.

any $X \in \mathcal{E}$; the left hand square commutes by assumption on $\rho_{\mathbf{1}}$:

$$\begin{tikzcd}
X \times \mathbf{1} \arrow[r, "X \times \eta_{\mathbf{1}}"] \arrow[d, "="'] & X \times T(\mathbf{1}) \arrow[r, "t''"] \arrow[d, "X \times \rho_{\mathbf{1}}"'] & T(X \times \mathbf{1}) \arrow[d, "\rho_{X \times \mathbf{1}}"] \\
X \times \mathbf{1} \arrow[r, "X \times \eta_{\mathbf{1}}"'] & X \times T(\mathbf{1}) \arrow[r, "t''"'] & T(X \times \mathbf{1})
\end{tikzcd}$$

Now both the horizontal composites are $\eta_{X \times \mathbf{1}}$, by general theory of tensorial strengths. Also $\rho_{X \times \mathbf{1}}$ is T-linear. Then uniqueness of T-linear extensions over $\eta_{X \times \mathbf{1}}$ implies that the right hand vertical map is the identity map. Using the natural identification of $X \times \mathbf{1}$ with X, we then also get that ρ_X is the identity map of $T(X)$.

14. Additive structure

For the applications of distribution theory, one needs not only that distributions on a space can be multiplied by scalars, but also that they can be *added*. In our context, this will follow if the category T-algebras is an *additive* (or *linear*) category, in a sense which we shall recall. Most textbooks, e.g. [16], [1] define the notion of additive category (with biproducts) as a property of categories $\mathcal{A}$ which are already equipped with the *structure* of enrichment in the category of abelian monoids (or even abelian groups). However, we shall need that such enrichment structure derives canonically from a certain *property* of the category. This is old wisdom, e.g. described in Pareigis' 1969 textbook [18], Section 4.1.

We recall briefly the needed properties: Consider a category $\mathcal{A}$ with finite products and finite coproducts. If the unique map from the initial object to the terminal object is an isomorphism, this object is a zero object 0. The zero object allows one to define a canonical map δ from the coproduct of n objects B_i to their product. If this δ is an isomorphism for all n-tuples of objects, this coproduct (and also, the product) is a biproduct $\oplus_i B_i$. Then every object acquires the structure of an abelian monoid object in $\mathcal{A}$, with the codiagonal map $B \oplus B \to B$ as addition, and with the map unique map $0 \to B$ as unit. This in turn implies a canonical enrichment of $\mathcal{A}$ in the category of abelian monoids. We refer to [18], (except that loc.cit. item 4) includes a property which implies subtraction of morphisms, i.e. an enrichment in the category of abelian groups, but so long subtraction can be dispensed with, then so can item 4).)

Consider now a category $\mathcal{E}$ with finite products and finite coproducts, and consider a monad T on $\mathcal{E}$. Let $F : \mathcal{E} \to \mathcal{E}^T$ be the functor $X \mapsto (T(X), \mu_X)$, i.e. the functor associating to X the free T-algebra on X. If $T(\emptyset)$ is the terminal object $\mathbf{1}$ (where $\emptyset$ is the initial object), then it is clear that $F(\emptyset)$ is a zero object in $\mathcal{E}^T$. Even though $\mathcal{E}^T$ may not have all finite coproducts, at least it has finite coproducts of free algebras, and so one may consider the canonical map $\delta : F(X) + F(Y) \to F(X) \times F(Y)$ (whose underlying map in $\mathcal{E}$ is a map $T(X+Y) \to T(X) \times T(Y)$, easy to describe directly, using 0 and the universal property of η_{X+Y}). If it is an isomorphism for all X and Y, the category of free algebras will therefore be additive with biproducts $\oplus$. The addition map for an object $T(X)$ is thus the inverse of δ, followed by $T(\nabla)$, where $\nabla : X + X \to X$ is the codiagonal. Since δ is T-linear, hence so is its inverse. Also $T(\nabla)$ is T-linear, so the addition map $T(X) \times T(X) \to T(X)$ is T-linear. For a more complete description, see [3] or [12]. (In [3], it is proved that in fact that the category of *all* T-algebras, i.e. $\mathcal{E}^T$, is additive with biproducts.)

Thus, in particular, any T-linear map will be additive. It is tempting to conclude "hence any bilinear (i.e. T-bilinear) map will be bi-additive". This conclusion is true, and similarly 1- or 2-linear will be additive in the first (resp. in the second) variable; but in all three cases, an argument is required. Let us sketch the proof of this for the case of 2-linearity. Since $t''_{X,Y}$ is initial among 2-linear maps out of $X \times T(Y)$ (Proposition 3.1), it suffices to prove that t'' is additive in the second variable, which is to say that the following diagram commutes (where $\tilde{\Delta}$ is "diagonalizing X and then taking middle four interchange"):

$$\begin{tikzcd}
X \times TY \times TY \arrow[r, "\tilde{\Delta}"] \arrow[d, "X \times (+)"'] & X \times TY \times X \times TY \arrow[r, "t'' \times t''"] & T(X \times Y) \times T(X \times Y) \arrow[d, "(+)"] \\
X \times TY \arrow[rr, "t''"'] & & T(X \times Y).
\end{tikzcd}$$

To see this commutativity, it suffices to see that the two composites agree when precomposed with

$$X \times TY \xrightarrow{X \times T(in_i)} X \times T(Y+Y) \xrightarrow[\cong]{\delta} X \times TY \times TY, \tag{16}$$

($i = 1, 2$); it is easy to see that we get t'' in both cases (and for $i = 1$ as well as for $i = 2$). (Hint: use that (16) postcomposed with $X \times (+)$ is the identity map on $X \times TY$, cf. the construction [12], equation (32).)

Let us summarize:

- the addition map $T(X) \times T(X) \to T(X)$ is T-linear
- T-linear maps are additive
- T-bilinear maps (resp. T-1-linear, resp. T-2-linear maps) are bi-additive (resp. additive in the first variable, resp. additive in the second variable)

It follows that the monoid of scalars $R = T(\mathbf{1})$ carries a T-linear addition $+ : R \times R \to R$, and that the T-bilinear multiplication $R \times R \to R$ is bi-additive. So R, with its multiplication and addition, is a rig. Similarly, the multiplicative action of R on any

$T(X)$ is bi-additive. Since any T-linear map $T(X) \to T(Y)$ is equivariant for the action, and, as a T-linear map, is additive, it follows that any $T(X)$ is a module over the rig R, and that T-linear maps $T(X) \to T(Y)$ are R-linear, in the sense of module theory.

Since each object in $\mathcal{E}^T$ carries a canonical structure of abelian monoid, one may ask whether these abelian monoids have the *property* of being abelian *groups*. This is then a property of the monad. It is, more compactly, equivalent to the property: $1 \in R$ has an additive inverse -1. If this is so, difference-formation, for any T-algebra A, will be a T-linear map $A \times A \to A$.

We shall need this further property in Section 17: differential calculus depends on having differences.

15. Distributions on the line R of scalars

We now consider distributions $P \in T(X)$, where X is the space R of scalars. Then $X \pitchfork R = R \pitchfork R$ has some particular (global) elements, to which we can apply $\langle P, - \rangle$, namely the monomials x^n, and also, R has an addition $R \times R \to R$, and we have convolution along this, which we shall denote $* : T(R) \times T(R) \to T(R)$. (There are some quite evident generalizations to the space $X = R^n$, but for simplicity, we stick to the 1-dimensional case.) Examples of elements in $T(R)$ are: probability distributions of random variables, cf. Section 16.

The *nth moment* of $P \in T(R)$ is defined as $\langle P, x^n \rangle \in R$, where $x^n : R \to R$ is the map $x \mapsto x^n$. In particular, the 0th moment of P is $\langle P, 1_R \rangle$, which we have considered already under the name $\mathrm{tot}(P)$, cf. Proposition 10.3. The 1st moment is $\langle P, x \rangle$, i.e. $\langle P, id_R \rangle$, or, in the "integral" notation also used for the pairing, $\int_R x \, dP(x)$. When P is a probability distribution, this scalar is usually called the *expectation* of P. We therefore also denote it $E(P)$, so $E(P) := \langle P, x \rangle$, for arbitrary $P \in T(R)$. (We shall not here consider nth moments for $n \geq 2$.)

Recall that $R := T(\mathbf{1})$, so $T(R) = T^2(\mathbf{1})$, and we therefore have the map $\mu_{\mathbf{1}} : T(R) \to R$.

15.1. Proposition. *The "expectation" map $E : T(R) \to R$ equals $\mu_{\mathbf{1}} : T^2(\mathbf{1}) \to T(\mathbf{1})$.*

Proof. To prove $\langle P, id_R \rangle = \mu_{\mathbf{1}}(P)$, we note that both sides of this equation depend in a T-linear way of P, so it suffices to prove it for the case where $P = \eta_R(x)$ for some $x \in R$. But $\langle \eta_R(x), id_R \rangle$ is "evaluation of id_R on x", by construction of the pairing, so gives x back. On the other hand $\mu_{\mathbf{1}} \circ \eta_R$ is the identity map on R, by a monad law, so this composite likewise returns x.

Recall from (11) that $\mathrm{tot}_X \circ \eta_X = 1_X$ as maps from X to R (for any X). From the proof of Proposition 15.1, we have that $E \circ \eta_R = id_R$. In elementwise terms, for $x \in R$,

$$\mathrm{tot}(\eta(x)) = 1 \quad ; \quad E(\eta(x)) = x, \tag{17}$$

(where η denotes η_R). Using these two facts, we can prove

15.2. Proposition. *For P and Q in $T(R)$, we have*

$$E(P * Q) = E(P) \cdot \mathrm{tot}(Q) + \mathrm{tot}(P) \cdot E(Q).$$

Proof. Convolution is T-bilinear, and E and tot are T-linear. Since addition is a T-linear map, it follows that both sides of the desired equation depend in a T-bilinear way on P, Q. Therefore, by Proposition 9.3, it suffices to prove it for the case where $P = \eta(x)$ and $Q = \eta(y)$ (with x and y in R). Then $E(P) = x$, $E(Q) = y$ and $\mathrm{tot}(P) = \mathrm{tot}(Q) = 1$. Also, since $*$ is convolution along $+$, we have $E(P * Q) = E(\eta(x) * \eta(y)) = E(\eta(x+y)) = x+y$. Then the result is immediate.

Let h be a homothety on R, i.e. a map of the form $x \mapsto b \cdot x$, for some $b \in R$. It is a T-linear map, since the multiplication on R is T-bilinear. We claim

$$E(h_*(P)) = h(E(P)) \tag{18}$$

for any $P \in T(R)$. To see this, note that both sides depend in a T-linear way on P, so it suffices to see it for the case of $P = \eta(x)$ ($x \in R$). But $E(h_*(\eta(x))) = E(\eta(h(x))) = h(x) = h(E(\eta(x)))$.

We also consider the effect of a translation map α on R, i.e. a map of the form $x \mapsto x+a$ for some $a \in R$. We claim

$$\alpha_*(P) = P * (\eta(a)) \tag{19}$$

for any $P \in T(R)$. To see this, we note that both sides here depend in a T-linear way on P, so it suffices to see it for $P = \eta(x)$. Then both sides give $\eta(x + a)$.

Applying Proposition 15.2 to $P * \eta(a)$, we therefore have, for $\alpha =$ translation by $a \in R$,

$$E(\alpha_*(P)) = E(P) + \mathrm{tot}(P) \cdot a.$$

In particular, if P has total 1,

$$E(\alpha_*(P)) = E(P) + a = \alpha(E(P)). \tag{20}$$

Recall that an affine map $R \to R$ is a map f which can be written as a homothety followed by a translation. Combining (18) and (20), we therefore get

15.3. Proposition. *Let $f : R \to R$ be an affine map. Then if $P \in T(R)$ has total 1, we have*

$$E(f_*(P)) = f(E(P)).$$

If $P \in T(R)$ has $\mathrm{tot}(P)$ multiplicatively invertible, we define *the center of gravity* $cg(P)$ by

$$cg(P) := E((\mathrm{tot}(P))^{-1} \cdot P).$$

Since E is T-linear, it preserves multiplication by scalars, so $Q := (\mathrm{tot}(P))^{-1} \cdot P$ has total 1, and the previous Proposition applies. So for an affine $f : R \to R$, $E(f_*(Q)) = f(E(Q))$. The right hand side is $f(cg(P))$, by definition. But f_* preserves formation of totals (any map does), and then it is easy to conclude that the left hand side is $cg(f_*(P))$. Thus we have

15.4. PROPOSITION. *Formation of center of gravity is preserved by affine maps f, $cg(f_*P) = f(cg(P))$.*

Thus, our theory is in concordance with the truth that "center of gravity for a mass distribution on a line does not depend on choice of 0 and choice of unit of length".

16. The affine submonad; probability distributions

A strong monad T on a cartesian closed category $\mathcal{E}$ is called *affine* if $T(\mathbf{1}) = \mathbf{1}$. For algebraic theories (monads on the category of sets), this was introduced in [20]. For strong monads, it was proved in [9] that this is equivalent to the assertion that for all X, Y, the map $\psi_{X,Y} : T(X) \times T(Y) \to T(X \times Y)$ is split monic with $\langle T(pr_1), T(pr_2)\rangle : T(X \times Y) \to T(X) \times T(Y)$ as retraction. In [15], it was proved that if $\mathcal{E}$ has finite limits, any commutative monad T has a maximal affine submonad T_0, the "affine part of T". It is likewise a commutative monad. Speaking in elementwise terms, $T_0(X)$ consists of those distributions whose total is $1 \in T(\mathbf{1})$. We consider in the following a commutative monad T and its affine part T_0.

Probability distributions have by definition total $1 \in R$, and take values in the interval from 0 to 1, in the sense that $0 \leq \langle P, \phi\rangle \leq 1$ if ϕ is a multiplicative idempotent in $X \pitchfork R$. We do not in the present article consider any order relation on R, so there is no "interval from 0 to 1"; so we are stretching terminology a bit when we use the word "probability distribution on X" for the elements of $T_0(X)$, but we shall do so.

If $P \in T_0(X)$ and $Q \in T_0(Y)$, then $P \otimes Q \in T_0(X \times Y)$, cf. (13) and the remark following it. (It also can be seen from the fact that the inclusion of strong monads $T_0 \subseteq T$ is compatible with the monoidal structure $\otimes$.) From this in turn follows that e.g. probability distributions are stable under convolution.

Most of the theory developed presently works for the monad T_0 just as it does for T; however, since $T_0(\mathbf{1}) = \mathbf{1}$ is trivial (but, in good cases, $T_0(\mathbf{2}) = T(\mathbf{1})$), the theory of "multiplying a distribution by a function" (Section 12) trivializes. So we shall consider the T_0 in conjunction with T, and R denotes $T(\mathbf{1})$. Clearly, $T_0(X)$ is not stable under multiplication by scalars $\in R$; for, the total gets multiplied by the scalar as well. In particular, we cannot, for $P \in T_0(X)$, expect that $P \vdash \phi$ is in $T_0(X)$. Nevertheless $P \vdash \phi$ has a probabilistic significance, provided $\langle P, \phi\rangle$ is multiplicatively invertible in R. If this is the case, we may form a distribution denoted $P \,|\, \phi$,

$$P \mid \phi := \langle P, \phi\rangle^{-1} \cdot (P \vdash \phi) = \lambda \cdot (P \vdash \phi),$$

(writing $\lambda \in R$ for $\langle P, \phi\rangle^{-1}$) and so (using Proposition 12.1, and bilinearity) we have

$$\mathrm{tot}(P \,|\, \phi) = \langle P \,|\, \phi, 1_X\rangle = \lambda \cdot \langle P \vdash \phi, 1_X\rangle = \lambda \cdot \langle P, \phi \cdot 1_X\rangle = \lambda \cdot \langle P, \phi\rangle = 1,$$

so that $P \,|\, \phi$ again is a probability distribution. More generally $P \,|\, \phi$, satisfies, for any $\psi \in X \pitchfork R$,

$$\langle P \,|\, \phi, \psi\rangle = \lambda\langle P, \phi \cdot \psi\rangle. \tag{21}$$

Now, multiplicatively idempotent elements in $X \pitchfork R$ may reasonably be seen as *events* in the *outcome space* X, and if ϕ and ψ are such events, also $\phi \cdot \psi$ is an event, and it deserves the notation $\phi \cap \psi$ (simultaneous occurrence of the two events). In this notation, and with the defining equation for λ, the equation (21) reads

$$\langle P \,|\, \phi, \psi \rangle = \frac{\langle P, \phi \cap \psi \rangle}{\langle P, \phi \rangle},$$

so that $P \,|\, \phi$ is "the conditional probability that ψ (and ϕ) occur, given that ϕ does".

Random variables and their distributions A random variable X on an outcome space Ω defines a probability distribution $\in T_0(R)$, and is often identified with this distribution. A *pair* X_1, X_2 of random variables on the same Ω, on the other hand, cannot be identified with a pair of distributions on R, but rather with a distribution $P \in T_0(R^2)$, the *joint* probability distribution of the two random variables. This P *gives rise* to a pair of distributions $P_i := (pr_i)_*(P)$ $(i = 1, 2)$ (the *marginal distributions* of P); so we have $(P_1, P_2) \in T(R) \times T(R)$. We may compare P with $P_1 \otimes P_2$. To say that $P = P_1 \otimes P_2$ is expressed by saying that the two random variables X_1 and X_2 are *independent*.

(On the other hand, the marginal distributions of $P \otimes Q$, for P and $Q \in T_0(R)$, agree with P and Q; this is a general property of affine monads, mentioned earlier.)

A pair of random variables X and Y on Ω has a sum, which is a new random variable; it is often denoted $X + Y$, but the probability distribution $\in T(R)$ of this sum is not a sum of two probability distributions (such a sum would anyway have total 2, not total 1). If X and Y are independent, the sum $X + Y$ has as distribution the convolution along $+$ of the two individual distributions. In general, the distribution of $X + Y$ is $+_*(P)$ where $+ : R^2 \to R$ is the addition map and $P \in T(R^2)$ is the joint distribution of the two random variables.

In text books on probability theory (see e.g. [4]), one sometimes sees the formula

$$E(X + Y) = E(X) + E(Y)$$

for the expectation of a sum of two random variables on Ω (not necessarily independent). The meaning of this non-trivial formula (which looks trivial, because of the notational confusion between a random variable and its distribution) is, in the present framework, a property of distributions on R. Namely it is a way to record the commutativity of the diagram

$$\begin{tikzcd}
T(R \times R) \arrow[r, "T(+)"] \arrow[d, "\beta"'] & T(R) \arrow[d, "E"] \\
R \times R \arrow[r, "+"'] & R.
\end{tikzcd}$$

Here, β is the T-algebra structure on $T(R^2)$, i.e. $\mu_{\mathbf{2}} : T^2(\mathbf{2}) \to T(\mathbf{2})$ (recall that $\mathbf{2} = \mathbf{1} + \mathbf{1}$, so by additivity of the monad, $T(\mathbf{2}) = T(\mathbf{1}) \times T(\mathbf{1}) = R \times R$); similarly, E is $\mu_{\mathbf{1}}$, and the diagram commutes by naturality of μ (recalling that $+$ is $T(\mathbf{2} \to \mathbf{1})$).

17. Differential calculus for distributions

We attempt in this Section to show how some differential calculus of distributions may be developed independently of the standard differential calculus of functions.

For simplicity, we only consider functions and distributions on R (one-variable calculus), but some of the considerations readily generalize to distributions and functions on any space X equipped with a vector field.

For this, we assume that the monad T on $\mathcal{E}$ has the properties described at the end of Section 14, so in particular, R is a commutative ring. To have some differential calculus going for such R, one needs some further assumption: either the possibility to form "limits" in R-modules, or else, availability of the method of synthetic differential geometry (SDG). We present the considerations in the latter form, with (some version of) the KL[6] axiomatics as the base, since this is well suited to make sense in any Cartesian closed category $\mathcal{E}$. In the spirit of SDG, we shall talk about the objects of $\mathcal{E}$ as if they were sets.

Consider a commutative ring $R \in \mathcal{E}$. Let $D \subseteq R$ be a subset containing 0, and let V be an R-module. We say that V satisfies the "KL"-property (relative to D)" if for any $X \in \mathcal{E}$ and any $F : X \times D \to V$, there exists a unique $f' : X \to V$ such that

$$F(x, d) = F(x, 0) + d \cdot f'(x) \quad \text{for all } d \in D.$$

If V is a KL module, then so is $X \pitchfork V$.

Example: 1) models R of synthetic differential geometry, (so D is the set of $d \in R$ with $d^2 = 0$); then the "classical" KL axiom says that at least the module R itself satisfies the above condition. If $X = R$ and $F(x, d) = f(x + d)$ for some function $f : R \to V$, f' is the standard derivative of f.

2) Any commutative ring, with $D = \{0, d\}$, for one single invertible $d \in R$. In this case, for given F, the f' asserted by the axiom is the function

$$f'(x) = \frac{1}{d} \cdot (F(x, d) - F(x, 0));$$

if $X = R$ and $F(x, d) = f(x + d)$ for some function $f : R \to V$, f' is the standard difference quotient.

In either case, we may call f' the *derivative* of f.

It is easy to see that *any* commutative ring R is a model, using $\{0, d\}$ as D, as in Example 2) (and then also, any R-module V satisfies then the axiom); this leads to

[6] Kock-Lawvere

some calculus of finite differences. Also, it is true that if $\mathcal{E}$ is the category of abstract sets, there are *no* non-trivial models of the type in Example 1); but, on the other hand, there are other cartesian closed categories $\mathcal{E}$ (e.g. certain toposes containing the category of smooth manifolds, cf. e.g. [11]), where a rather full-fledged differential calculus for functions emerges from the KL-axiom.

We assume that $R = T(1)$ has the KL property (for some fixed $D \subseteq R$), more generally, that any R-module of the form $T(X)$ has it.

17.1. Proposition. [Cancelling universally quantified ds] *If V is an R-module which satisfies KL, and $v \in V$ has the property that $d \cdot v = 0$ for all $d \in D$, then $v = 0$.*

Proof. Consider the function $f : R \to V$ given by $t \mapsto t \cdot v$. Then for all $x \in R$ and $d \in D$

$$f(x + d) = (x + d) \cdot v = x \cdot v + d \cdot v,$$

so that the constant function with value v will serve as f'. On the other hand, $d \cdot v = d \cdot 0$ by assumption, so that the equation may be continued,

$$= x \cdot v + d \cdot 0$$

so that the constant function with value $0 \in V$ will likewise serve as f'. From the uniqueness of f', as requested by the axiom, then follows that $v = 0$.

We are now going to provide a notion of *derivative* P' for any $P \in T(R)$. Unlike differentiation of distributions in the sense of Schwartz, which is defined in terms of differentiation of test functions ϕ, our construction does not mention test functions, and the Schwartz definition[7] $\langle P', \phi\rangle := \langle P, \phi'\rangle$ comes in our treatment out as a *result*, see Proposition 17.4 below.

For $u \in R$, we let α^u be "translation by u", i.e. the map[8] $x \mapsto x + u$.

For $u = 0$, $\alpha^u_*(P) - P = 0 \in T(R)$. Assuming that the R-module $T(R)$ has the KL property, we therefore have for any $P \in T(R)$ that there exists a unique $P' \in T(R)$ such that for all $d \in D$,

$$d \cdot P' = \alpha^d_*(P) - P.$$

Since $d \cdot P'$ has total 0 for all $d \in D$, it follows from Proposition 17.1 that P' has total 0.

(If V is a KL module, differentiation of functions $R \to V$ can be likewise be described "functorially" as a map $R \pitchfork V \to R \pitchfork V$, namely to $\phi \in R \pitchfork V$, ϕ' is the unique element in $R \pitchfork V$ satisfying $d \cdot \phi' = (\alpha^d)^*(\phi) - \phi$.)

Differentiation is translation-invariant: using

$$\alpha^t \circ \alpha^s = \alpha^{t+s} = \alpha^s \circ \alpha^t,$$

[7]to be precise, Schwartz includes a minus sign in the definition, to accommodate the viewpoint that "*distributions are generalized functions*", so that a minus sign in the definition is needed to have differentiation of distributions *extending* that of functions.

[8]Note that this map is invertible. So we are not using the functor property of T, except for invertible maps. This means that some of the following readily makes sense for distribution notions T which are only functorial for a more restricted class of maps, like proper maps.

it is easy to deduce that

$$(\alpha^t_*(P))' = (\alpha^t)_*(P'). \tag{22}$$

17.2. Proposition. *Differentiation of distributions on R is a T-linear process.*

Proof. Let temporarily $\Delta : T(R) \to T(R)$ denote the differentiation process. Consider a fixed $d \in D$. Then for any $P \in T(R)$, $d \cdot \Delta(P)$ is $\alpha^d_* P - P$; as a function of P, it is a difference of two T-linear maps, namely the identity map on $T(R)$ and $\alpha^d_* = T(\alpha^d)$, and as such is T-linear. Thus for each $d \in D$, the map $d \cdot \Delta : T(R) \to T(R)$ is T-linear. Now to prove T-linearity of Δ means, by monad theory, to prove equality of two maps $T^2(R) \to T(R)$; and since $d \cdot \Delta$ is T-linear, as we proved, it follows that the two desired maps $T^2(R) \to T(R)$ become equal when post-composed with the map "multiplication by d": $T(R) \to T(R)$. Since $d \in D$ was arbitrary, it follows from the principle of cancelling universally quantified ds (Proposition 17.1) that the two desired maps are equal, proving T-linearity.

17.3. Proposition. *Let $P \in T(R)$. Then*

$$E(P') = \mathrm{tot}(P).$$

Proof. The Proposition says that two maps $T(R) \to R$ agree, namely $E \circ \Delta$ and tot, where Δ, as above, is the differentiation process $P \mapsto P'$. Both these maps are T-linear, so it suffices to prove that the equation holds for the case $P = \eta_R(x)$, for each fixed $x \in R$. Let us write δ_x for the distribution $\eta_X(x)$, so that we do not confuse it with a function. So we should prove

$$E(\delta'_x) = \mathrm{tot}(\delta_x).$$

By the principle of cancelling universally quantified ds (Proposition 17.1), it suffices to prove that for all $d \in D$ that

$$d \cdot E(\delta'_x) = d \cdot \mathrm{tot}(\delta_x). \tag{23}$$

The left hand side of (23) is

$$\begin{aligned} E(d \cdot \delta'_x) &= E(\alpha^d_* \delta_x - \delta_x) \\ &= E(\delta_{x+d} - \delta_x) \\ &= E(\delta_{x+d}) - E(\delta_x) = (x+d) - x = d, \end{aligned}$$

by the second equation in (17). The right hand side of (23) is d, by the first equation in (17). This proves the Proposition.

The differentiation process for functions, as a map $R \pitchfork V \to R \pitchfork V$, is likewise T-linear, but this important information cannot be used in the same way as we used T-linearity of the differentiation $T(R) \to T(R)$, since, unlike $T(R)$, $R \pitchfork V$ is not known to be freely generated by elementary quantities like the δ_xs (not even when $V = R$).

Here is an important relationship between differentiation of distributions on R, and of functions $\phi : R \to T(X)$; such functions can be differentiated, since $T(X)$ is assumed to be KL as an R-module. In the Schwartz theory, this relationship, with $X = \mathbf{1}$, serves as *definition* of derivative of distributions (except for a sign change, see an earlier footnote).

17.4. Proposition. *For $P \in T(R)$ and $\phi \in R \pitchfork T(X)$, one has*

$$\langle P', \phi \rangle = \langle P, \phi' \rangle.$$

Proof. We are comparing two maps $T(R) \times (R \pitchfork T(X)) \to T(X)$, both of which are T-linear in the first variable. Therefore, it suffices to prove the equality for the case of $P = \delta_t (= \eta_R(t))$; in fact, by R-bilinearity of the pairing and Proposition 17.1, it suffices to prove that for any $t \in R$ and $d \in D$, we have

$$\langle d \cdot (\delta_t)', \phi \rangle = \langle \delta_t, d \cdot \phi' \rangle.$$

The left hand side is $\langle \alpha^d_*(\delta_t) - \delta_t, \phi \rangle$, and using bi-additivity of the pairing and (7), this gives $((\alpha^d)^*)(\phi)(t) - \phi(t) = \phi(t + d) - \phi(t)$, which is $d \cdot \phi'(t)$.

It is easy to see that if $F : V \to W$ is an R-linear map between KL modules, we have

$$F \circ \phi' = (F \circ \phi)' \tag{24}$$

for any $\phi : R \to V$.

17.5. Proposition. *Let $P \in T(R)$ and $Q \in T(R)$. Then*

$$(P * Q)' = P' * Q = P * Q'.$$

Proof. Let us prove that $(P*Q)' = P'*Q$ (then $(P*Q)' = P*Q'$ follows by commutativity of convolution). Both sides depend in a T-bilinear way on P and Q, so it suffices to see the validity for the case where $P = \delta_a$ and $Q = \delta_b$. To prove $(\delta_a * \delta_b)' = \delta'_a * \delta_b$, it suffices to prove that for all $d \in D$,

$$d \cdot (\delta_a * \delta_b)' = d \cdot \delta'_a * \delta_b,$$

and both sides come out as $\delta_{a+b+d} - \delta_{a+b}$, using that $*$ is R-bilinear.

Primitives of distributions on R We noted already in Section 10 that P and $f_*(P)$ have same total, for any $P \in T(X)$ and $f : X \to Y$. In particular, for $P \in T(R)$ and $d \in D$, $d \cdot P' = \alpha^d_*(P) - P$ has total 0, so cancelling the universally quantified d we get that P' has total 0.

A *primitive* of a distribution $Q \in T(R)$ is a $P \in T(R)$ with $P' = Q$. Since any P' has total 0, a necessary condition that a distribution $Q \in T(R)$ has a primitive is that $\mathrm{tot}(Q) = 0$. Recall that primitives, in ordinary 1-variable calculus, are also called "indefinite *integrals*", whence the following use of the word "integration":

Integration Axiom. *Every $Q \in T(R)$ with* $\mathrm{tot}(Q) = 0$ *has a unique primitive.*

(For contrast: for functions $\phi : R \to R$, the standard integration axiom is that primitives *always* exist, but are *not* unique, only up to an additive constant.)

By R-linearity of the differentiation process $T(R) \to T(R)$, the uniqueness assertion in the Axiom is equivalent to the assertion: *if $P' = 0$, then $P = 0$.* (Note that $P' = 0$ implies that P is invariant under translations $\alpha^d_*(P) = P$ for all $d \in D$.) The reasonableness of this latter assertion is a two-stage argument: 1) if $P' = 0$, P is invariant under *arbitary translations*, $\alpha^u_*(P) = P$. 2) if P is invariant under all translations, and has compact support, it must be 0. (Implicitly here is: R itself is not compact.)

In standard distribution theory, the Dirac distribution δ_a (i.e. $\eta_R(a)$) (where $a \in R$) has a primitive, namely the Heaviside "function"; but this "function" has not compact support - its support is a half line $\subseteq R$.

On the other hand, the integration axiom provides a (unique) primitive for a distribution of the form $\delta_b - \delta_a$, with a and b in R. This primitive is denoted $[a, b]$, the "interval" from a to b; thus, the defining equation for this interval is

$$[a, b]' = \delta_b - \delta_a.$$

Note that the phrase "interval from ... to ..." does not imply that we are considering an ordering or preordering $\leq$ on R (although ultimately, one wants to do so).

17.6. PROPOSITION. *The total of* $[a, b]$ *is* $b - a$.

PROOF. We have

$$\mathrm{tot}([a, b]) = E([a, b]') = E(\delta_b - \delta_a)$$

by Proposition 17.3 and the fact that $[a, b]$ is a primitive of $\delta_b - \delta_a$

$$= E(\delta_b) - E(\delta_a) = b - a,$$

by (17).

It is of some interest to study the sequence of distributions

$$[-a, a], \quad [-a, a] * [-a, a], \quad [-a, a] * [-a, a] * [-a, a], \quad \ldots;$$

they have totals $2a, (2a)^2, (2a)^3, \ldots$; in particular, if $2a = 1$, this is a sequence of probability distributions, approaching a Gauss normal distribution (the latter, however, has presently no place in our context, since it does not have compact support).

The following depends on the Leibniz rule for differentiating a product of two functions; so this is *not* valid under he general assumptions of this Section, but needs the further assumption of Example 2, namely that D consists of $d \in R$ with $d^2 = 0$, as in synthetic differential geometry. We shall then use "test function" technique to prove a "Leibniz rule"[9] for the action $\vdash$:

17.7. PROPOSITION. *For any* $P \in T(R)$ *and* $\phi \in R \pitchfork R$,

$$(P \vdash \phi)' = P' \vdash \phi - P \vdash \phi'.$$

[9] with the Schwartz convention of sign, one gets a plus rather than a minus between the terms.

Proof. Since there are enough B-valued test functions (Proposition 11.1), it suffices to prove for arbitrary test function $\eta : R \to B$ (with B a free T-algebra - hence, by assumption, a KL-module) that we have

$$(\langle P \vdash \phi)', \eta\rangle = \langle P' \vdash \phi - P \vdash \phi', \eta\rangle.$$

We calculate (using that the pairing is bi-additive):

$$\begin{aligned}\langle (P \vdash \phi)', \eta\rangle &= \langle P \vdash \phi, \eta'\rangle \quad \text{(by Proposition 17.4)}\\ &= \langle P, \phi \cdot \eta'\rangle \quad \text{(by Proposition 12.1}\\ &= \langle P, (\phi \cdot \eta)' - \phi' \cdot \eta\rangle\end{aligned}$$

using that Leibniz rule applies, for functions on R, to any bilinear pairing, like the multiplication map $\cdot$,

$$\begin{aligned}&= \langle P, (\phi \cdot \eta)'\rangle - \langle P, \phi' \cdot \eta\rangle\\ &= \langle P', \phi \cdot \eta\rangle - \langle P, \phi' \cdot \eta\rangle\end{aligned}$$

using Proposition 17.4 on the first summand

$$= \langle P' \vdash \phi, \eta\rangle - \langle P \vdash \phi', \eta\rangle$$

using Proposition 12.1 on each summand

$$= \langle P' \vdash \phi - P \vdash \phi', \eta\rangle.$$

In other words, the proof looks formally like the one from books on distribution theory, but does not depend on "sufficiently many test functions with values in R".

References

[1] F. Borceux, *Handbook of Categorical Algebra* Vol. 2, Cambridge University Press 1994.

[2] M. Bunge, Relative functor categories and categories of algebras, *J. Algebra* 11 (1969), 64-101.

[3] D. Coumans and B. Jacobs, Scalars, Monads, and Categories, arXiv [math.RA] 1003.0585 (2010).

[4] H. Cramér, *Sannolikhetskalkylen*, Almqvist & Wiksell, Stockholm 1949.

[5] A. Frölicher and A. Kriegl, *Linear Spaces and Differentiation Theory*, Wiley 1988.

[6] A. Kock, Monads on symmetric monoidal closed categories. *Arch. Math. (Basel)*, 21:1–10, 1970.

[7] A. Kock, On double dualization monads. *Math. Scand.*, 27:151–165, 1970.

[8] A. Kock, Closed categories generated by commutative monads. *J. Austral. Math. Soc.*, 12:405–424, 1971.

[9] A. Kock, Bilinearity and Cartesian closed monads. *Math. Scand.*, 29:161–174, 1971.

[10] A. Kock, Strong functors and monoidal monads. *Arch. Math. (Basel)*, 23:113–120, 1972.

[11] A. Kock, *Synthetic Differential Geometry*, London Math. Soc. Lecture Notes 51, Cambridge University Press 1981; Second Edition London Math. Soc. Lecture Notes 333, Cambridge University Press 2006.

[12] A. Kock, Monads and extensive quantities, arXiv [math.CT] 1103.6009 (2011).

[13] A. Kock, Calculus of extensive quantities, arXiv [math.CT] 1105.3405 (2011).

[14] F.W. Lawvere, Categories of space and of quantity, in: J. Echeverria et al. (eds.), *The Space of Mathematics*, de Gruyter, Berlin, New York (1992)

[15] H. Lindner, Affine parts of monads, *Arch. Math. (Basel)*, 33:437-443, 1979.

[16] S. Mac Lane, *Categories for the Working Mathematician*, Springer 1971.

[17] E. Moggi, Notions of computations and monads, *Information and Computation* 93 (1991), 55-92.

[18] B. Pareigis, *Kategorien und Funktoren*, Teubner Stuttgart 1969.

[19] L. Schwartz, *Méthodes mathématiques pour les sciences physiques*, Hermann Paris 1961.

[20] G.C. Wraith, *Algebraic Theories* (revised version), Aarhus Matematisk Institut Lecture Notes Series 22, 1975.

Department of Mathematics, University of Aarhus, DK 8000 Aarhus C, Denmark
Email: `kock@imf.au.dk`

This article may be accessed at `http://www.tac.mta.ca/tac/` or by anonymous ftp at `ftp://ftp.tac.mta.ca/pub/tac/html/volumes/26/4/26-04.{dvi,ps,pdf}`

THEORY AND APPLICATIONS OF CATEGORIES (ISSN 1201-561X) will disseminate articles that significantly advance the study of categorical algebra or methods, or that make significant new contributions to mathematical science using categorical methods. The scope of the journal includes: all areas of pure category theory, including higher dimensional categories; applications of category theory to algebra, geometry and topology and other areas of mathematics; applications of category theory to computer science, physics and other mathematical sciences; contributions to scientific knowledge that make use of categorical methods.

Articles appearing in the journal have been carefully and critically refereed under the responsibility of members of the Editorial Board. Only papers judged to be both significant and excellent are accepted for publication.

Full text of the journal is freely available in .dvi, Postscript and PDF from the journal's server at `http://www.tac.mta.ca/tac/` and by ftp. It is archived electronically and in printed paper format.

SUBSCRIPTION INFORMATION Individual subscribers receive abstracts of articles by e-mail as they are published. To subscribe, send e-mail to `tac@mta.ca` including a full name and postal address. For institutional subscription, send enquiries to the Managing Editor, Robert Rosebrugh, `rrosebrugh@mta.ca`.

INFORMATION FOR AUTHORS The typesetting language of the journal is TEX, and LATEX2e strongly encouraged. Articles should be submitted by e-mail directly to a Transmitting Editor. Please obtain detailed information on submission format and style files at `http://www.tac.mta.ca/tac/`.