Mathematical Structures in Computer Science (2026), **35**, e41, pp. 1–49
doi:10.1017/S0960129525100212

CAMBRIDGE UNIVERSITY PRESS

PAPER

From differential linear logic to coherent differentiation

Thomas Ehrhard

Université Paris Cité, CNRS, Inria, IRIF, F-75013, Paris, France
Email: thomas.ehrhard@inria.fr

(Received 26 April 2024; revised 20 July 2025; accepted 25 July 2025)

Abstract

In thise survey, we present in a unified way the categorical and syntactical settings of coherent differentiation introduced recently, which shows that the basic ideas of differential linear logic and of the differential lambda-calculus are compatible with determinism. Indeed, due to the Leibniz rule of the differential calculus, differential linear logic and the differential lambda-calculus feature an operation of addition of proofs or terms operationally interpreted as a strong form of nondeterminism. The main idea of coherent differentiation is that the summations required by differentiation can be controlled and kept deterministic, in the denotational models, as well as in the syntax and operational semantics themselves. In sharp contrast with the differential lambda-calculus, coherent differentiation indeed allows one to design differential extensions of Turing complete functional programming languages, such as PCF, endowed with a deterministic evaluation mechanism, typically implemented in a Krivine abstract machine.

Keywords: Denotational semantics; linear logic; differential lambda-calculus

1. Introduction

During his development of the denotational semantics of System F in the cartesian closed category of qualitative domains, and more specifically of coherence spaces, Girard observed in Girard (1986) that a specific class of morphisms arises naturally among the general morphisms of this model (the stable maps). These particular stable functions are characterized by an additional preservation property: they commute with compatible unions.

Girard not only recognized the relevance of these morphisms to the description of the stable semantics, he also understood that they should play a more fundamental role than the general stable functions themselves. He called them *linear maps* because they collectively behave very much like linear maps in algebra, forming a symmetric monoidal closed category which is even $*$-autonomous (a general categorical notion introduced in Barr (1979) describing categories of linear morphisms where all objects are reflexive, i.e., canonically isomorphic to their bidual).

He understood that this observation, *a priori* relative to the stable semantics, was the denotational shadow of a fundamental and hitherto hidden structure of Intuitionistic Logic itself: Linear Logic (LL), see Girard (1987). This essential discovery had a major impact in Logic and Computer Science, notably on the study and design of programming languages.

One aspect of linearity which is not directly addressed in LL – although it is contemplated in the concluding section of Girard (1987) — is its central role in the differential calculus where differentiation consists in extracting from a morphism its "best linear approximation". The purpose of Differential LL (DiLL for short) is to take this role of linearity into account. This logical system

© The Author(s), 2026. Published by Cambridge University Press.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

was introduced by the author and Laurent Regnier in Ehrhard and Regnier (2003, 2006b) and is summarized in Ehrhard (2018). It extends LL without adding new connectives, but by adding a set of new deduction rules that we can classify as follows.

- There are 3 rules relative to the exponential modality !_, dual to the standard rules of weakening, contraction, and dereliction.
- And there are two rules expressing that any finite family of proofs of the same formula has a sum, which is again a proof of that formula. This includes the case of an empty family, meaning that any formula is provable by a 0 proof.

These latter rules mean that the proofs in DiLL are essentially partial: the 0 proof is very similar to the term Ω in the theory of Böhm trees, which is a completely undefined term. More importantly, the unrestricted ability of adding proofs of the same formula means that DiLL features a fundamental nondeterminism in the sense that it is possible to add the two normal proofs of $1 \oplus 1$, which corresponds to the type of booleans, that is, intuitively, the two booleans **t** and **f**. The most natural operational understanding of this "boolean" $\mathbf{t} + \mathbf{f} : 1 \oplus 1$ consists in considering it as a program, which can nondeterministically reduce to **t** or **f**. We provide in Section 3.2 a more precise description of the λ-calculus account of this extension of LL, which features a typing rule allowing one to add any two terms having the same type.

1.1 Determinism and differentiation

In Ehrhard and Regnier (2008, 2006a), we developed a Taylor expansion of λ-terms, which is based on the differential λ-calculus. This expansion consists in translating a λ-term M into a (generally infinite) sum M^* of "resource terms," that is, of differential λ-terms whose only use of the standard λ-calculus operation of application is application to 0. In other words, all the applications $(P)Q$ are replaced hereditarily by differential terms:

$$((P)Q)^* = \sum_{n=0}^{\infty} \frac{1}{n!} (\mathsf{D}^n P^* \cdot (\overbrace{Q^*, \ldots, Q^*}^{n}))0 \,.$$

In this expression, if $M : A \Rightarrow B$ and $N_1, \ldots, N_n : A$ are differential λ-term, then

$$\mathsf{D}^n M \cdot (N_1, \ldots, N_n) : A \Rightarrow B$$

is the nth derivative of M applied to linear argument $N_1, \ldots, N_n$, see Section 3.2.

Analyzing the execution of standard λ-terms in the Krivine Machine we showed that, if a standard λ-term M is β-equivalent to a variable x then there is exactly one term s in the Taylor expansion M^* of M which does not reduce to 0, and s reduces to $(M^*_s)^{-1}\, x$, if we take multiplicities into account.[1] This resource term s is a trace of the execution of M in the machine, or more precisely, s provides a precise account of the quantitative use of the various subterms of M during the execution.

In other words, in the infinite sum M^*, only one term is nonzero, although for knowing which one, one needs to reduce M to its normal form. This means that this infinite sum is only apparently nondeterministic. But for proving this property, we strongly use the fact that M is a *standard* λ-term, that is, contains no differential construct of shape $\mathsf{D}P \cdot Q$.

So beyond this first encouraging observation, the question remained open of whether differentiation (in the sense of DiLL) can be made compatible with determinism. Toward a positive answer to this question, the second crucial observation was that, in the model of LL based on probabilistic coherence spaces (PCSs), nonlinear morphisms are functions which are analytic in the sense that they are defined by power series and hence should be differentiable, in spite of the fact that, in the corresponding category **Pcoh**, it is not always possible to add two morphisms

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

of the same type. For instance, the sum of two subprobability distributions on the natural numbers is not always a subprobability distribution (the global mass can become > 1). The simplest interesting object of this category is the closed unit interval $[0, 1]$ that we denote as 1, and a morphism $1 \to 1$ is a function $f : [0, 1] \to [0, 1]$ such that there is a (necessarily uniquely defined) sequence $(t_n)_{n\in\mathbb{N}}$ of nonnegative real numbers such that $\forall x \in [0, 1]\ f(x) = \sum_{n=0}^{\infty} t_n x^n \in [0, 1]$, a condition which simply means $\sum_{n=0}^{\infty} t_n \leq 1$. Such a function has a derivative, defined on $[0, 1)$ by $f'(x) = \sum_{n=0}^{\infty} (n + 1)t_{n+1}x^n$ and this derivative cannot always be extended to $[0, 1]$: think of $f(x) = 1 - \sqrt{1 - x}$ whose derivative is $f'(x) = \frac{1}{2\sqrt{1-x}}$ for $x < 1$. At this point, it is useful to remember that this basic notion of derivative that we have learned at school hides an ingredient that is essential when one considers more general situations: a linear argument. Indeed, the derivative of a (sufficiently regular) function $h : E \to F$ where E and F are vector spaces is a function $h' : E \to (E \multimap F)$ where $E \multimap F$ is the vector space of linear functions from E to F. In the case where $E = \mathbb{R}$ it is natural to identify $E \multimap F$ with F and then $h' : \mathbb{R} \to F$ as we are used to. This identification amounts to saying that the derivative $h'(x) \in \mathbb{R} \multimap F$ is always given $1 \in \mathbb{R}$ as linear argument which is perfectly fine when one deals with vector spaces or similar structures. But the interval $[0, 1]$ is not a vector space, and in particular when $x, u \in [0, 1]$, it is not always true that $x + u \in [0, 1]$. If we set $S = \{(x, u) \in [0, 1] \mid x + u \in [0, 1]\}$ and consider the function $\mathsf{D}f : S \to \mathbb{R}_{\geq 0}$ given by $\mathsf{D}f(x, u) = f'(x)u$, then it is easy to check that $\forall (x, u) \in S \mathsf{D}f(x, u) \leq f(x + u) - f(x) \leq 1$ so that $\mathsf{D}f$ is actually an analytic morphism $S \to [0, 1]$. We can even say that the map $\mathsf{T}f : S \to S$ given by $\mathsf{T}f(x, u) = (f(x), f'(x)u)$ is analytic. This simple observation, together with the fact that the correspondence $f \mapsto \mathsf{T}f$ is functorial, is the starting point of a new setting for differentiation in the λ-calculus and in LL called *coherent differentiation*.

1.2 Content

The paper starts with a historical description of some ideas leading to coherent differentiation, that we divide into 3 phases:

- a first unpublished attempt by the author at defining the derivative of a stable function on coherence spaces in the 1980's;
- the introduction of the differential λ-calculus and of DiLL in the 2000's;
- and last the discovery of coherent differentiation in 2021, which gives a clear status to the first attempt and makes it work completely.

From Section 5, we describe the categorical and syntactic setting of coherent differentiation. In Section 5, we introduce the basic structure of summability, which allows one to consider categories where hom-sets are commutative partial monoids and more precisely axiomatizes functorially an operation which maps an object X to the object $\mathsf{S}X$ of pairs (x_0, x_1) such that the sum $x_0 + x_1$ exists. This operation is presented as a functor $\mathsf{S} : \mathcal{L} \to \mathcal{L}$ equipped with three natural transformations where, intuitively, $\mathcal{L}$ is a "linear" category. Elaborating on this infrastructure, we introduce in Section 6 the basic idea of coherent differentiation which is to represent differentiation as a distributive law between the !-comonad of a (weak) structure of model of LL on $\mathcal{L}$ and the functor S that we equip with a monad structure canonically induced by the summability structure. In Section 7, we consider the case where the category $\mathcal{L}$ is symmetric monoidal closed and hence induces a model of the λ-calculus, explaining how the differential structure interacts with the closed structure, and with fixpoint operators when available (Section 8).

In these developments, we consider a particularly important situation, called the *representable* situation[2] where the functor S can be described on objects by $\mathsf{S}X = (1 \,\&\, 1 \multimap X)$ and similarly on morphisms. In that case, the differential structure boils down to a !-coalgebra structure on $1 \,\&\, 1$. It turns out that all the concrete models of coherent differentiation known so far are representable.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Last in Section 9, we outline a syntax incorporating in a functional language the categorical structures developed in the previous section. This functional language is an extension of Scott-Milner-Plotkin's PCF Plotkin (1977). Here are the main features of this extension.

- The only ground type (integers) is equipped with a let construct making it possible to use call-by-value on integers, which is crucial when the language is extended with probabilistic choice (this refinement of PCF was introduced in Ehrhard et al. (2014));
- There is a type constructor corresponding to the functor S mentioned above and associated term constructors corresponding to the main categorical ingredients of the axiomatization of coherent differentiation;
- The operational semantics is described by means of an abstract machine.

We give the main results about this operational semantics, which have been proven in Ehrhard (2023a): soundness and adequacy, and we explain how the denotational semantics shows that this operational semantics is deterministic.

1.3 Comparison with the previous articles on coherent differentiation

The present article is a survey, so it does not introduce essentially new notions or results with respect to the two articles it is based on and where coherent differentiation was introduced: semantically in Ehrhard (2023b) and syntactically in Ehrhard (2023a). More than these articles, this survey tries to articulate coherent differentiation with two closely related topics: tangent categories and differential LL. It summarizes rather faithfully the content of these two papers with more emphasis on Ehrhard (2023b), omitting technical details and proofs which are not essential for understanding the main ideas.

The presentation of summability structures in Section 5.2 is very close to that of Ehrhard (2023b). As in that paper the associated functor S is shown in Section 5.3 to have a monad structure similar to that of the tangent bundle functor in differential geometry. We describe moreover the comonad structure of S whose comultiplication corresponds to the vertical lift of the tangent bundle functor, see for instance Cockett and Cruttwell (2014). The compatibility of summability structures with the monoidal and the cartesian products are presented in Section 5.4 and 5.5, in a way which is very close to Ehrhard (2023b).

Section 6, about differential structures on summability structures, follows Ehrhard (2023b) faithfully. Concerning the representable situation, whose similarity with synthetic differential geometry is discussed in Remark 4.12 and where the functor S is defined on objects by $\mathsf{S}X = (1 \,\&\, 1 \multimap X)$, the first difference of presentation wrt. Ehrhard (2023b) lies in the terminology. This representable situation was called "elementary" in that paper, this change of terminology follows a very sensible suggestion of a reviewer of Ehrhard and Walch 2025). The presentation itself is drastically simplified wrt. Ehrhard (2023b), mainly by the omission of proofs and intermediate lemmas so that, to fully understand the connection between the representable situation and general coherent differential categories, the reader will probably have to read Section 5.4 of Ehrhard (2023b) (which uses implicitly the concept of *mate* in a way which is systematized in Ehrhard and Walch 2025)).

We describe a syntax for coherent differentiation in Section 9, borrowing material from the rather long and technical Ehrhard (2023a). We tried to make the presentation of this language, of its denotational semantics and of its operational semantics self-contained, and, in the same time, to avoid technicalities as much as possible. So, for instance, to fully understand the proof of a statement like Theorem 8.4, the reader would certainly need to spend some time on Ehrhard (2023a). The author hopes that Section 9 may serve as a motivation and as an introduction to such a reading.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

To illustrate coherent differentiation, we present the examples of coherence spaces as in Ehrhard (2023b) and of PCSs as in Ehrhard (2023a), this latter one being particularly meaningful because of its similarity with ordinary differentiation in Analysis. To increase readability, examples are interleaved with more abstract developments.

There is also a way of presenting the ideas of coherent differentiation in the line of *differential categories* (see for instance Blute et al. (2009); Cockett and Cruttwell (2014); Blute et al. (2020)) that we initiated in Ehrhard and Walch (2023a), the main idea being to modify the axiomatization of cartesian differential categories by requiring that the homsets are equipped with a *partial* commutative monoid structure. We do not address this topic in the present survey; let us mention that this approach is also developed in the framework of Taylor expansion in Ehrhard and Walch (2025) in the present volume.

2. Notations and Terminology

A finite multiset of elements of a set A is a function $m : A \to \mathbb{N}$ such that $\mathsf{supp}(m) = \{a \in A \mid m(a) \neq 0\}$ is finite. We use $[\,]$ for the empty multiset such that $\mathsf{supp}([\,]) = \emptyset$ and standard algebraic notations $m_1 + m_2$ and $\sum_{i\in I} m_i$ (for I finite) for the pointwise addition of multisets.

Given $m = [a_1, \ldots, a_k] \in \mathcal{M}_{\mathrm{fin}}(A)$, the multiset defined by $m(a) = \#\{j \mid a_j = a\}$, and given $i \in I$, we define $i \cdot m = [(i, a_1), \ldots, (i, a_k)] \in \mathcal{M}_{\mathrm{fin}}(I \times A)$.

If $M = [m_1, \ldots, m_k] \in \mathcal{M}_{\mathrm{fin}}(\mathcal{M}_{\mathrm{fin}}(A))$, we set $\sum M = \sum_{i=1}^{k} m_k \in \mathcal{M}_{\mathrm{fin}}(A)$.

Let $\mathcal{C}$ be a category. A family of morphisms $(h_i \in \mathcal{C}(X, Y_i))_{i\in I}$ is *jointly monic* if for any $f, f' \in \mathcal{C}(Z, X)$, if $(h_i f = h_i f')_{i\in I}$ then $f = f'$. And $(h_i \in \mathcal{C}(X_i, Y))_{i\in I}$ is *jointly epic* if for any $f, f' \in \mathcal{C}(Y, Z)$, if $(f\, h_i = f'\, h_i)_{i\in I}$ then $f = f'$.

3. Differentiation in LL

A coherence space (CS) is a pair $E = (|E|, \frown_E)$ where $|E|$ is a set, the *web*, and $\frown_E$ is a binary, reflexive, and symmetric relation on $|E|$, the *coherence relation*. A clique of E is a subset x of $|E|$ such that $\forall a, a' \in x\ a \frown_E a'$. Given CSs E and F, one defines a CS $E \multimap F$ by $|E \multimap F|$ and $(a, b) \frown_{E\multimap F} (a', b')$ if $a \frown_E a' \Rightarrow (b \frown_F b'$ and $b = b' \Rightarrow a = a')$, and the category **Coh** has CSs as objects, and $\mathbf{Coh}(E, F) = \mathrm{Cl}(E \multimap F)$, identity morphisms being the diagonal relations and composition being the standard composition of relations.

The category **Coh** is a model of LL and the exponential considered first by Girard $!_{\mathrm{g}}E$ is defined as follows: $\left|!_{\mathrm{g}}E\right| = \{\{a_1, \ldots, a_n\} \mid n \in \mathbb{N} \text{ and } \{a_1, \ldots, a_n\} \in \mathrm{Cl}(E)\}$. The main feature of this exponential is that its Kleisli category $\mathbf{Coh}_{!_{\mathrm{g}}}$ is isomorphic to the category of coherence spaces and stable functions by the following correspondence: with any $t \in \mathbf{Coh}_{!_{\mathrm{g}}}(E, F) = \mathrm{Cl}(!_{\mathrm{g}}E \multimap F)$ one associates the stable function $\hat{t} : \mathrm{Cl}(E) \to \mathrm{Cl}(F)$ defined by $\hat{t}(x) = \{b \in |F| \mid \exists x_0 \subseteq x\ (x_0, b) \in t\}$ and the mapping $t \mapsto \hat{t}$ is a bijection between $\mathbf{Coh}_{!_{\mathrm{g}}}(E, F)$ and the set of stable functions $\mathrm{Cl}(E) \to \mathrm{Cl}(F)$.

Such a $t \in \mathbf{Coh}_{!_{\mathrm{g}}}(E, F)$ is linear if all its elements (x_0, b) are such that x_0 is a singleton, so that these linear stable maps are just the same thing as elements of $\mathbf{Coh}(E, F)$. We use the linear algebraic notation $t \cdot x$ to denote the application of such a linear t to $x \in \mathrm{Cl}(E)$, that is $t \cdot x = \{b \in |F| \mid \exists a \in x\ (a, b) \in t\}$.

3.1 A first attempt: the derivative of a stable function

In light of these definitions, and in view of the role of linearity in analysis and geometry, a natural question appeared to the author: is it possible to turn such a general stable morphism $t \in \mathbf{Coh}_{!_{\mathrm{g}}}(E, F)$ into a linear map by an operation similar to differentiation?

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

This question should be made a bit more precise and is actually twofold: given $x \in \mathrm{Cl}(E)$,

- can we define a coherence space E_x of all possible "extensions" of x, that is, such that for all $u \in \mathrm{Cl}(E_x)$, the set $x \cup u \in \mathrm{Cl}(E)$;
- and is there a linear $t'(x) \in \mathrm{Cl}(E_x \multimap F)$ such that, for all $u \in \mathrm{Cl}(E_x)$, $\hat{t}(x) \cup t'(x) \cdot u$ exists, is a subset of $\hat{t}(x \cup u)$ and is the "best" possible approximation of that set by means of a linear map?

A natural tentative answer to the first question is to take $|E_x| = \{a' \in |E| \mid \forall a \in x \; a \frown_E a'\}$ (with $a_1 \frown_{E_x} a_2$ if $a_1 \frown_E a_2$) and then $t'(x) = \{(a', b) \in |E_x| \times |F| \mid \exists x_0 \subseteq x \; (x_0 \cup \{a'\}, b) \in t\}$.

This works quite well if we take $x = \emptyset$, in that case $E_x = E$ and $t'(\emptyset) \in \mathbf{Coh}(E, F)$ and is characterized by

$$t'(\emptyset) \cdot u = \bigcup_{a \in u} \hat{t}(\{a\})$$

and notice also that we have a morphism $\overline{\mathsf{der}}_E = \{(a, \{a\}) \mid a \in |E|\} \in \mathbf{Coh}(E, !_{\mathrm{g}}E)$ such that $t'(\emptyset) = t\, \overline{\mathsf{der}}_E$.

Imagine now that $x = \{a, a'\}$ for some $a \frown_E a'$, and that moreover $t = \{(x, b)\}$. Notice that $a, a' \in |E_x|$. Then our definition of $t'(x)$ yields $(a, b), (a', b) \in t'(x)$ and hence $t'(x) \notin \mathrm{Cl}(E_x \multimap F)$. We stopped our investigation of this idea at this point in 1986-1987, but we could have tried to push this line of ideas a little bit further as we explain now.

We can interpret this failure as meaning that our definition of E_x is not satisfactory, we can try $|E_x| = \{a' \in |E| \mid \forall a \in x \; a \frown_E a'\}$ (with $a_1 \frown_{E_x} a_2$ if $a_1 \frown_E a_2$), that is, if $u \in \mathrm{Cl}(E_x)$, not only $x \cup u \in \mathrm{Cl}(E)$ but also $x \cap u = \emptyset$.

With this definition of E_x, let $((a_i, b_i) \in t'(x))_{i=1,2}$ and assume that $a_1 \frown_{E_x} a_2$, that is $a_1 \frown_E a_2$. This means that there are $(x_i \subseteq x)_{i=1,2}$ with $((x_i \cup \{a_i\}, b_i) \in t)_{i=1,2}$, which implies $b_1 \frown_F b_2$ because $x_1 \cup x_2 \cup \{a_1, a_2\} \in \mathrm{Cl}(E)$ by definition of E_x. If moreover $b_1 = b_2$, we know that $x_1 \cup \{a_1\} = x_2 \cup \{a_2\}$ which implies $a_1 = a_2$ because $a_1 \notin x_2$ and $a_2 \notin x_1$ (remember that $x_1, x_2 \subseteq x$). So we do have $t'(x) \in \mathbf{Coh}(E_x, F)$. We can be even more precise: let $(a', b') \in t'(x)$ and let $b \in \hat{t}(x)$, so that there is $x_0 \subseteq x$ such that $(x_0, b) \in t$. Let $x'_0 \subseteq x$ be such that $(x'_0 \cup \{a'\}, b') \in t$, then we have $x_0 \frown_{!_{\mathrm{g}}E} x'_0 \cup \{a'\}$ because we know that $a' \notin x_0$. Therefore $b \frown_F b'$ and we have shown that $b' \in \left|F_{\hat{t}(x)}\right|$ from which it follows that $t'(x) \in \mathbf{Coh}(E_x, F_{\hat{t}(x)})$.

Let us adopt the following convention introduced by Girard: given $(x_i \in \mathrm{Cl}(E))_{i \in I}$, we use the notation $\sum_{i \in I} x_i$ to denote $\bigcup_{i \in I} x_i$ and to express at the same time that the x_i's are pairwise disjoint and that $\bigcup_{i \in I} x_i \in \mathrm{Cl}(X)$. Then our definition of $t'(x)$ satisfies

$$t'(x) \cdot u = \sum_{a \in u} (\hat{t}(x + \{a\}) \setminus \hat{t}(x)) \,.$$

Indeed, let first $b \in t'(x) \cdot u$, so let $a \in u$ be such that $(a, b) \in t'(x)$. There is $x_0 \subseteq x$ such that $(x_0 + \{a\}, b) \in t$ from which it follows that $b \in \hat{t}(x + \{a\})$. Next since $x_0 + \{a\}$ is minimal such that $b \in \hat{t}(x_0 + \{a\})$ and $a \notin x$, we cannot have $b \in \hat{t}(x)$ from which the $\subseteq$ inclusion follows. Let now $b \in \sum_{a \in u} (\hat{t}(x + \{a\}) \setminus \hat{t}(x))$, so let $a \in u$ be such that $b \in \hat{t}(x + \{a\}) \setminus \hat{t}(x)$. Let $x'_0 \subseteq x + \{a\}$ be such that $(x'_0, b) \in t$. Since $b \notin \hat{t}(x)$, we cannot have $x'_0 \subseteq x$ and hence we must have $x'_0 = x_0 + \{a\}$ for some $x_0 \subseteq x$. Since $a \in u \in \mathrm{Cl}(E_x)$, we have $(a, b) \in t'(x)$ from which the $\supseteq$ inclusion follows.

This shows in particular that $\hat{t}(x) + t'(x) \cdot u \subseteq \hat{t}(x + u)$. Let now $h \in \mathbf{Coh}(E_x, F_{\hat{t}(x)})$ be such that $\forall u \in \mathrm{Cl}(E_x) \; \hat{t}(x) + h \cdot u \subseteq \hat{t}(x + u)$, that is $h \cdot u \subseteq \hat{t}(x + u) \setminus \hat{t}(x)$. In particular, for $u = \{a\}$ with $a \in |E_x|$ we get $h \cdot \{a\} \subseteq t'(x) \cdot \{a\}$ which means that $h \subseteq t'(x)$. In that precise sense $t'(x)$ is the best linear under-approximation of the map $u \mapsto \hat{t}(x + u) \setminus \hat{t}(x)$ so can reasonably be called the derivative of t at x.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Let $s \in \mathbf{Coh}_{!_g}(E, F)$ and $t \in \mathbf{Coh}_{!_g}(F, G)$ so that $t \circ s \in \mathbf{Coh}_{!_g}(E, G)$ is the composition in the Kleisli category which can be described as follows:

$$t \circ s = \{(x_1 \cup \cdots \cup x_n, c) \in \left|!_g E\right| \times |G| \mid \exists b_1, \ldots, b_n \in |F| \\ ((x_i, b_i) \in s)_{i=1}^n \text{ and } (\{b_1, \ldots, b_n\}, c) \in t\}$$

and is fully characterized by $\widehat{t \circ s}(x) = \hat{t}(\hat{s}(x))$.

Remember that the coherence space 1 & 1 has $\{(1, *), (2, *)\}$ as web, with $(1, *) \frown_{1\&1} (2, *)$. Let $s \in \mathbf{Coh}_{!_g}(1, 1 \,\&\, 1)$ and $t \in \mathbf{Coh}_{!_g}(1 \,\&\, 1, 1)$ be given by

$$s = \{(\{*\}, (1, *)), (\{*\}, (2, *))\}$$
$$t = \{(\{(1, *), (2, *)\}, *)\}\,.$$

Then we have $t \circ s = \{(\{*\}, *)\}$. Therefore, $(t \circ s)'(\emptyset) \cdot \{*\} = \{*\}$ and on the other hand we have

$$\begin{aligned} t'(\hat{s}(\emptyset)) \cdot (s'(\emptyset) \cdot \{*\}) &= t'(\emptyset) \cdot (s'(\emptyset) \cdot \{*\}) \\ &= t'(\emptyset) \cdot \{(1, *), (2, *)\} \\ &= \hat{t}\{(1, *)\}) \cup \hat{t}\{(2, *)\}) \\ &= \emptyset\,. \end{aligned}$$

This means that the chain rule

$$(t \circ s)'(x) = t'(s(x))\, s'(x)$$

does not hold for this differentiation of stable function, we only have a weak version thereof $(t \circ s)'(x) \supseteq t'(s(x))\, s'(x)$.

The reason for the failure of the chain rule is clear: the morphism t is nonlinear (it needs to be fed with $\{(1, *), (2, *)\}$ to output the atomic result $*$, i.e., it uses its parameter at least twice) so in the computation of $\widehat{t \circ s}(\{*\})$, the atomic data $*$ is actually used at least twice, but this nonlinearity does not appear in $t \circ s = \{(\{*\}, *)\}$, which turns out to be a linear morphism in **Coh**: the two used copies of $*$ have been "merged" because set union is an idempotent operation.

Another difficulty in this concrete approach to differentiation of stable maps is that it is not clear how to express the regularity (and hopefully, the stability) of the derivative $t'(x)$ with respect to x (the author was not aware of the – at that time recently introduced – tangent categories of Rosický (1984)), and therefore, it was unclear how to define higher derivatives in this kind of setting.

3.2 Coming back to differentiation in LL

In the early 2000's, motivated by the phase space parameterized LL models of Bucciarelli and Ehrhard (2001), the author explored categorical models of LL where formulas are interpreted as vector spaces and morphisms as linear maps (in the usual sense of Linear Algebra). In such categories, the only known resource modalities yield infinite dimensional vector spaces as soon as they are applied to a non 0 space, and so the vector spaces under consideration must be equipped with a topology compatible with their algebraic structure. Two such models were developed by the author: Köthe sequence spaces Ehrhard (2002) and finiteness spaces Ehrhard (2005).

The objects of these models are topological vector spaces (tvs), which admit a simple description based on the existence of a "web" (in the sense of coherence spaces) which, in this algebraic context, can be understood as a Schauder basis, that is, not exactly a basis in the usual algebraic sense (Hamel basis), but a natural topological adaptation thereof, in which all the elements of the vector space can be written uniquely as possibly infinite linear combinations of base vectors (these infinite sums being defined as limits in the sense of the topology the vector space is endowed with). One important feature of these models is that webs are not used in the definition of morphisms, which are just linear and continuous maps.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

The obtained categories $\mathcal{L}$ are models of LL such that the Kleisli category $\mathcal{L}_!$ can be described as a category of analytic (or more precisely entire) functions. Moreover, the homsets of these categories have a natural tvs structure (because $\mathcal{L}$ is a symmetric monoidal closed category) and in particular have an operation of addition: they are additive categories.

The fact that the morphisms of such CCCs $\mathcal{L}_!$ are analytic and therefore infinitely differentiable led to the idea of extending the typed[3] λ-calculus with differential operations, developed in Ehrhard and Regnier (2003). The basic differential typing rule of this calculus is

$$\frac{\Gamma \vdash M : A \Rightarrow B \quad \Gamma \vdash N : A}{\Gamma \vdash \mathsf{D}M \cdot N : A \Rightarrow B}$$

the intuition being that A and B denote some kind of tvs E (of the symmetric monoidal closed category $\mathcal{L}$) and F, and M an entire function $f : E \to F$, more precisely $f \in \mathcal{L}_!(E, F)$ (possibly depending on additional parameters listed in Γ). Such a function can be differentiated into another entire function $f' \in \mathcal{L}_!(E, E \multimap F)$ such that, for any $x \in E$, the map $u \mapsto f(x) + f'(x) \cdot u$ is the best affine approximation of the map $u \mapsto f(x + u)$ (in the sense of the topology our tvs are endowed with). Then, if N denotes $u \in E$, the term $\mathsf{D}M \cdot N$ denotes the analytic function $E \to F$ which maps x to $f'(x) \cdot u$. This slightly unusual writing of differentials makes it easy to iterate derivatives: given $(\Gamma \vdash N_i : A)_{i=1,2}$, we have $\Gamma \vdash \mathsf{D}(\mathsf{D}M \cdot N_1) \cdot N_2 : A \Rightarrow B$, the second derivative of M, a bilinear morphism applied to its two linear arguments N_1 and N_2.

This differential application induces a new redex, in the case where $M = \lambda x^A\, P$ with $\Gamma, x : A \vdash P : B$, similar to a β-redex. The corresponding reduction is

$$\mathsf{D}(\lambda x^A\, P) \cdot N \to \lambda x^A \left(\frac{\partial P}{\partial x} \cdot N\right)$$

where the term $\frac{\partial P}{\partial x} \cdot N$ is the *differential substitution* of N for x in P, defined by induction on P, which is typed as follows

$$\Gamma, x : A \vdash \frac{\partial P}{\partial x} \cdot N : B$$

which shows that, in the term $\frac{\partial P}{\partial x} \cdot N$, the variable x can still be free. This is due to the fact that in that term only one linear copy is substituted with N. This linear substitution operation performs a nontrivial operation on P, creating linear occurrences of x to be substituted by N on request. The most important case in the definition of $\frac{\partial P}{\partial x} \cdot N$ is when P is an (ordinary) application $P = (Q)R$ with $\Gamma, x : A \vdash Q : C \Rightarrow B$ and $\Gamma, x : A \vdash R : C$. The inductive definition of this linear substitution stipulates that

$$\frac{\partial (Q)R}{\partial x} \cdot N = \left(\frac{\partial Q}{\partial x} \cdot N\right)R + \left(\mathsf{D}Q \cdot \left(\frac{\partial R}{\partial x} \cdot N\right)\right)R$$

and this definition involves the $+$ operation on terms, subject to the following typing rule

$$\frac{\Gamma \vdash M_1 : A \quad \Gamma \vdash M_2 : A}{\Gamma \vdash M_1 + M_2 : A} \tag{1}$$

(for each context Γ and each type A) which has an obvious denotational interpretation in our tvs models and should, operationally, be understood as a nondeterministic superposition. The case of a variable is also interesting, we set

$$\frac{\partial y}{\partial x} \cdot N = \begin{cases} N & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

where we see a 0 which is the neutral element of the $+$ above. The meaning of this 0 is that if $y \neq x$, then "y does not depend on x" and so we are taking the derivative of a constant function; a more operational understanding is that y has no linear occurrence of x and hence the linear

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

substitution fails. Addition is made possible by the typing rule (1) because the variable x may have several potential linear occurrences in P. Consider for instance $P = (x)(x)y$ typed as follows:

$$\Gamma, y : A, x : A \Rightarrow A \vdash (x)(x)y : A$$

and let N be a term such that $\Gamma \vdash \ : NA \Rightarrow A$, we have

$$\frac{\partial (x)(x)y}{\partial x} \cdot N = (N)(x)y + (\mathsf{D}x \cdot (N)y)(x)y$$

where some intuitively clear equations on terms (such as $\mathsf{D}M \cdot 0 = 0$) have been used implicitly. Let us write the term $P = (x_1)(x_2)y$, using x_1, x_2 to distinguish the two occurrences of x in P. Only the occurrence x_1 is linear,[4] the occurrence x_2 is not because the occurrence x_1 might take as value a nonlinear function, using x_2 in a nonlinear way. This is why for substituting N linearly for x_2 we need first to make the function x_1 use its argument linearly (or more precisely extract a linear copy of its argument); this is exactly the purpose of the $\mathsf{D}x \cdot _$ in the second term of the sum.

It is worth observing that these differential reduction rules produce nontrivial sums even if the term we start from does not contain such sums. Consider for instance, in a typed calculus with a base type of booleans, the following term

$$x : \mathbf{Bool} \vdash M = \mathsf{if}\,(x, \mathsf{if}\,(x, \mathbf{f}, \mathbf{t}), \mathsf{if}\,(x, \mathbf{f}, \mathbf{t})) : \mathbf{Bool}$$

then the definition of differential substitution leads to the reduction

$$(\mathsf{D}(\mathsf{D}(\lambda x^{\mathbf{Bool}}\, M) \cdot \mathbf{t}) \cdot \mathbf{f})0 \to \mathbf{t} + \mathbf{f}\,.$$

This results from the fact that, using iterated differential application, we manage to give the variable x two incompatible values $\mathbf{t}$ and $\mathbf{f}$ and the term M is written in such a way that, when x changes value during the computation (an impossible scenario in a deterministic setting), M issues $\mathbf{t}$ or $\mathbf{f}$, depending on the scheduling of this change of value. The order of the differential substitution being essentially irrelevant (this corresponds to the Schwarz rule of Calculus: the second derivative is a *symmetric* bilinear function), the computation is necessarily nondeterministic and leads to this nontrivial sum $\mathbf{t} + \mathbf{f}$.

This kind of example strongly suggested that extending the λ-calculus with differential constructs necessarily leads to essentially nondeterministic systems.

Later on, together with Laurent Regnier, the author developed a differential extension of LL in Ehrhard and Regnier (2004); Ehrhard (2018), fully compatible with these semantic and syntactic ideas. The beauty of this differential LL is that the new differential logical structure does not require new connectives, but introduces new deduction rules relative to the resource modality ! of LL, dual to the standard rules of dereliction, weakening and contraction. We already mentioned codereliction $\overline{\mathsf{der}}_E$ in the setting of coherence spaces in Section 3.1; coweakening and cocontraction are similar (the first also exists in coherence spaces, the second not). The associated new cut elimination rules preserve this new symmetry. The recent paper Kerjean and Lemay (2023) even extends this symmetry to promotion, the fundamentally infinitary rule of LL, which becomes then a bimonad.

4. Coherence and Determinism

At the most fundamental level, this was the author's understanding of differential λ-calculus and LL in May 2021. Of course, many results have been obtained, and many notions have been introduced concerning these systems, their applications, and their semantics since they have been introduced in the early 2000s; it would not be possible to cite all of them here. We can mention in particular many important advances on the categorical semantics of differentiation in cartesian categories, which are not necessarily models of LL (differential categories), on the 2-categorical meaning of differentiation and on applications of the syntactic Taylor expansion associated with

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

differential LL and the differential λ-calculus. Nevertheless, as far as we know, none of these developments questions the assumption that it is always possible to add terms, proofs or morphisms of the same type. And as explained above, there are very good reasons for such an assumption.

However, the first observations summarized in Section 3.1, suggesting the possibility of giving a meaning to derivatives in categorical models of LL where addition is a partial operation on morphisms, were strongly reinforced (more than 30 years later!) by the study of probabilistic coherence spaces (PCSs) developed in Ehrhard (2022), based on the fact that the Kleisli morphisms can clearly be understood as analytic functions in a very standard sense. In that paper, it is shown that the endeavor of Section 3.1 can be carried out successfully in PCSs, in the sense that the chain rule, which failed in **Coh** as we showed, perfectly holds in this setting.

As already observed, beyond the failure of the chain rule, one major puzzling question in Section 3.1 was: how can we express that $s'(x) \in \mathbf{Coh}(E_x, F_{\widehat{s}(x)})$ depends stably on x? This kind of question already arises in differential geometry where the derivative of a map from a real manifold X to another one Y at a point $x \in X$ is a linear map $f'(x) : \mathsf{T}_x X \to \mathsf{T}_{f(x)} Y$ where $\mathsf{T}_x X$ is the tangent space to X at x, which is a vector space. To express that this derivative has, for instance, a derivative at each point x, one introduces a new manifold $\mathsf{T}X$ (the tangent bundle of X) whose elements are the pairs (x, u) such that $u \in \mathsf{T}_x X$ and one turns this operation into a functor, mapping $f : X \to Y$ to the function $\mathsf{T}X \to \mathsf{T}Y$ defined by $\mathsf{T}f(x, u) = (f(x), f'(x) \cdot u)$ and then one can speak of the regularity (for instance, the differentiability) of this compound map $\mathsf{T}f$. This standard construction was categorically axiomatized for the first time in Rosický (1984). Under the name of *tangent category*, it has been generalized in Cockett and Cruttwell (2014). In that paper, differential categories are also shown to provide very natural examples of tangent categories, where $\mathsf{T}X = X \times X$ for any object X. This definition is compatible with the fact that, in differential categories, objects have to be thought of as some kind of "vector spaces" in accordance with the fact that such categories are left-additive. This means that hom-sets have a structure of commutative monoid, and that composition preserves this structure only on the left: we have $(f_1 + f_2) \circ g = (f_1 \circ g) + (f_2 \circ g)$ but $f \circ (g_1 + g_2) = (f \circ g_1) + (f \circ g_2)$ does not always hold because, intuitively, morphisms are smooth but not linear in general. In tangent categories, the functoriality of the operation T expresses exactly the chain rule.

The main idea of coherent differentiation is very close to that of tangent categories, with one major difference which required some time to the author to fully understand. In the tangent bundle construct, the manifold X and the tangent space at $x \in X$ are typically of very different natures: the tangent spaces are usually all isomorphic to $\mathbb{R}^d$ where d is the dimension of the manifold, and so they are trivial geometric objects, whereas the manifold itself can have a rich geometrical structure (defined typically by systems of equations, gluing, quotient, etc). In our setting that arises from the LL analysis of denotational semantics, the manifold is replaced by a "domain" E, a coherence space for instance, and, given $x \in \mathrm{Cl}(E)$, an element of $\mathsf{T}_x E = E_x$ should be an $y \in \mathrm{Cl}(E)$ such that $x + y$ makes sense, that is, $x \cap y = \emptyset$ and $x \cup y \in \mathrm{Cl}(E)$, for the quasi-example developed in Section 3.1. Since the LL analysis of denotational semantics is based on a fundamental analogy between domains and vector spaces, this means that here the "tangent bundle" functor already applies nontrivially to objects of the linear category and implements – functorially as we shall see – a notion of partial summability. In sharp contrast, in the tangent bundle case, when X is a vector space, the associated tangent bundle is simply $\mathsf{T}X = X \times X$ equipped with the first projection.

Remark 3.1. This also means that some room is left for developing a notion of "manifold" for coherent differentiation, or of a notion of coherent tangent categories where the "tangent spaces" would only be partially additive. Such a generalization requires motivations coming from concrete computational situations or from coherent differential situations arising in geometry; as far as we know such situations are still to be discovered.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

The first ingredient of coherent differentiation is therefore an axiomatization of categories where morphisms are only partially summable. It is perfectly meaningful, although not really necessary[5] to assume that such a category $\mathcal{L}$ is a "linear category," that is a symmetric monoidal closed category with possibly additional properties and structures (cartesian products, resource modality etc). The present paper makes this kind of assumption about $\mathcal{L}$.

We could assume that $\mathcal{L}$ is enriched in some kind of "partial commutative monoids," but this would not be really sufficient, because we also need to associate with each object X of $\mathcal{L}$ an object $\mathsf{S}X$ whose elements are, intuitively, the pairs $(x_0, x_1) \in X^2$ such that $x_0 + x_1$ is well defined. Therefore our partial summability structure is axiomatized as a functor $\mathsf{S} : \mathcal{L} \to \mathcal{L}$ equipped with three natural transformations $\pi_0, \pi_1, \sigma : \mathsf{S}X \to X$ which intuitively map such a summable pair (x_0, x_1) to x_0, x_1 and $x_0 + x_1$, respectively. Then saying that two morphisms $f_0, f_1 \in \mathcal{L}(X, Y)$ are summable simply means that there is a morphism $h \in \mathcal{L}(X, \mathsf{S}Y)$ such that $\pi_i \ h = f_i$ for $i = 0, 1$ and since it is important for us that h, the "witness of summability" of f_0 and f_1, be unique, we assume π_0 and π_1 to be jointly monic. Thanks to this uniqueness, we can set $f_0 + f_1 = \sigma \ h$. Suitable axioms on this structure allow one to show that $\mathcal{L}$ is enriched in partial commutative monoids.

5. Summability Structures in a Linear Setting

5.1 Partial monoids

We first describe the kind of partial commutative monoids that our axiomatization of summability induces.

Definition 4.1. *A partial commutative monoid is a triple $(M, 0, +)$ where M is a set, $0 \in M$ and $+ : M^2 \to M$ is a partial function such that*

- *$0 + a$ is defined for all $a \in M$ and $0 + a = a$;*
- *if $a + b$ is defined then $b + a$ is defined and $a + b = b + a$;*
- *if $a + b$ and $(a + b) + c$ are defined then $b + c$ and $a + (b + c)$ are defined, and $(a + b) + c = a + (b + c)$.*

Remark 4.1. This notion of partial commutative monoid is stronger than it might seem at first sight and involves some kind of "positivity". For instance the set $M = \{0, 1\} \subseteq \mathbb{Z}$ with addition $a + b$ defined as in $\mathbb{Z}$ if $a + b \in M$ and undefined otherwise, is a partial commutative monoid. But if we apply a similar definition to $M = \{-1, 0, 1\}$, the obtained structure does not satisfy the associativity condition of partial commutative monoids (take $a = -1$ and $b = c = 1$).

This notion of partial monoid is perfectly adapted to the kind of denotational situations we are abstracting on — which are essentially positive — but more general notions of partial monoid which do not enforce positivity have been introduced and lead to interesting notions of summability, see for instance Hines (2013). See also the discussion in Ehrhard and Walch (2023a), Section 3.1.

Definition 4.2. *Let $(M, 0, +)$ be a partial commutative monoid. Then we define by induction on n what it means for a sequence $\vec{a} \in M^n$ to be summable, and the value $\sum_{i=1}^{n} a_i$ of its sum:*

- *if $n = 0$, the empty sequence is summable and has 0 as sum;*
- *if $n > 0$, a sequence $a_1, \ldots, a_n$ is summable if $a_1, \ldots, a_{n-1}$ is summable and $\sum_{i=1}^{n-1} a_i$ and a_n are summable, and then $\sum_{i=1}^{n} a_i = (\sum_{i=1}^{n-1} a_i) + a_n$.*

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Lemma 4.1. *Let $(M, 0, +)$ be a partial commutative monoid. Let $\vec{a} \in M^n$ and let $f : \{1, \dots, n\} \to \{1, \dots, n\}$ be a bijection. The sequence $\vec{a}$ is summable iff $(a_{f(1)}, \dots, a_{f(n)})$ is summable, and then $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_{f(i)}$.*

Thanks to that lemma, the following definition makes sense.

Definition 4.3. *Let $(M, 0, +)$ be a partial commutative monoid and I be a finite set. One says that $\vec{a} \in M^I$ is summable if there is an enumeration without repetitions $\vec{i} = (i_1, \dots, i_n)$ of the elements of I (so that $n = \#I$) such that $(a_{i_1}, \dots, a_{i_n})$ is summable, and if this is the case the sum $\sum \vec{a} = \sum_{i \in I} a_i$ of $\vec{a}$ is defined as $\sum_{j=1}^{n} a_{i_j}$. Indeed, by* Lemma 4.1*, the summability and sum of $(a_{i_1}, \dots, a_{i_n})$ does not depend on the enumeration $\vec{i}$ of I.*

Theorem 4.1. *Let $(M, 0, +)$ is a partial commutative monoid. Let I be a finite set and $(I_j)_{j \in J}$ be a finite family of pairwise disjoint sets such that $\bigcup_{j \in J} I_j = I$. Let $\vec{a} \in M^I$. The following statements are equivalent*

- *$\vec{a}$ is summable;*
- *for all $j \in J$ the family $(a_i)_{i \in I_j}$ is summable and the family $(\sum_{i \in I_j} a_i)_{j \in J}$ is summable.*

When these two equivalent conditions hold, one has $\sum_{i \in I} a_i = \sum_{j \in J} \sum_{i \in I_j} a_i$.

The proofs of these facts are standard and can also be found in Ehrhard (2023b).

Lemma 4.2. *If $(a_i)_{i \in I}$ is a finite summable family in a partial commutative monoid $(M, 0, +)$ and $I' \subseteq I$, then $(a_i)_{i \in I'}$ is summable.*

Proof. Immediate consequence of Theorem 4.1 (take $J = \{1, 2\}$, $I_1 = I'$ and $I_2 = I \setminus I'$). □

5.2 Summability structures

Let $\mathcal{L}$ be a category with zero-morphisms, that is, $\mathcal{L}$ is enriched over the category of pointed sets. We use $0_{X,Y}$ or simply 0 for the distinguished zero-element of $\mathcal{L}(X, Y)$, so that $f\,0 = 0\,f = 0$. If $\mathcal{L}$ is a symmetric monoidal category (SMC), we also assume that $0 \otimes f = 0$ and $f \otimes 0 = 0$. If $\mathcal{L}$ has a terminal object $\top$, notice that $\mathcal{L}(X, \top) = \{0\}$ for any object X.

The first structure we assume $\mathcal{L}$ to be equipped with is a functor $\mathsf{S} : \mathcal{L} \to \mathcal{L}$ whose intuitive meaning is to map any object X to the object $\mathsf{S}X$ of all pairs (x_0, x_1) of elements of X for which the sum $x_0 + x_1$ exists. In accordance with this intuition, this functor is equipped with three natural transformations $\pi_0, \pi_1, \sigma \in \mathcal{L}(\mathsf{S}X, X)$ intuitively mapping such a pair to x_0, x_1 and $x_0 + x_1$ respectively.

Definition 4.4. *A pre-summability structure on $\mathcal{L}$ is a triple $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ where $\mathsf{S} : \mathcal{L} \to \mathcal{L}$ is a functor and π_0, π_1 and σ are natural transformations with $\pi_{0,X}, \pi_{1,X}, \sigma_X \in \mathcal{L}(\mathsf{S}X, X)$ such that $\pi_{0,X}$ and $\pi_{1,X}$ are jointly monic for all objects X of $\mathcal{L}$.*

Remark 4.2. When dealing with natural transformations such as $\pi_{0,X}, \pi_{1,X}$, and σ_X, we will most often omit the indexing objects X as they can always be recovered from the context.

Remember that the latter condition of Definition 4.4 means that if $f, g \in \mathcal{L}(Y, \mathsf{S}X)$ satisfy $(\pi_i\, f = \pi_i\, g)_{i=0,1}$ then $f = g$; this is the categorical way to say that $\mathsf{S}X$ is an object of pairs. From now on we fix a pre-summability structure on $\mathcal{L}$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Definition 4.5. *We say that two morphisms* $(f_i \in \mathcal{L}(Y, X))_{i=0,1}$ *are summable if there is an* $h \in \mathcal{L}(Y, \mathsf{S}X)$ *such that* $(\pi_i \ h = f_i)_{i=0,1}$. *If such an* h *exists, it is unique by joint monicity of the* π_i*'s and we set* $h = \langle\!\langle f_0, f_1 \rangle\!\rangle$ *and* $f_0 + f_1 = \sigma \ \langle\!\langle f_0, f_1 \rangle\!\rangle \in \mathcal{L}(Y, X)$. *The morphism* $\langle\!\langle f_0, f_1 \rangle\!\rangle$ *is the witness of the summability of* f_0 *and* f_1, *and* $f_0 + f_1$ *is their sum.*

Lemma 4.3. *The morphisms* $\pi_0, \pi_1 \in \mathcal{L}(\mathsf{S}X, X)$ *are summable, with* $\langle\!\langle \pi_0, \pi_1 \rangle\!\rangle = \mathsf{Id}_{\mathsf{S}X}$ *and* $\pi_0 + \pi_1 = \sigma$.

Proof. This is tautological. □

Definition 4.6. *If* $\vec{i} \in \{0, 1\}^n$, *we set* $\pi_{\vec{i}} = \pi_{i_1} \ \cdots \ \pi_{i_n} \in \mathcal{L}(\mathsf{S}^n X, X)$.

Lemma 4.4. *The morphisms* $\pi_{\vec{i}} \in \mathcal{L}(\mathsf{S}^n X, X)$, *for* $\vec{i} \in \{0, 1\}^n$, *are jointly monic.*

Proof. Simple induction on n. □

Lemma 4.5. *If* $f_0, f_1 \in \mathcal{L}(X, Y)$ *are summable and* $g \in \mathcal{L}(U, X)$ *and* $h \in \mathcal{L}(Y, V)$ *then* $h f_0 g$ *and* $h f_1 g$ *are summable with* $\langle\!\langle h f_0 g, h f_1 g \rangle\!\rangle = \mathsf{S}h \ \langle\!\langle f_0, f_1 \rangle\!\rangle \ g$ *and* $h f_0 g + h f_1 g = h (f_0 + f_1) g$.

Proof. By naturality. □

The additional assumptions on pre-summability structures that we will introduce now for defining summability structures will turn each hom-set of $\mathcal{L}$, equipped with 0 and the operation "+" defined above, into a partial commutative monoid in the sense of Section 5.2.

Definition 4.7. *(*S*-com) The pre-summability structure* $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ *is commutative if* π_1 *and* π_0 *are summable and* $\pi_1 + \pi_0 = \sigma$.

Lemma 4.6. *If the pre-summability structure* $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ *is commutative and* $f_0, f_1 \in \mathcal{L}(Y, X)$ *are summable, then* f_1, f_0 *are summable and* $f_1 + f_0 = f_0 + f_1$.

Remark 4.3. The (S-com) axiom corresponds to the commutativity condition in Definition 4.1. It should be noticed that even if $f_0 + f_1 = f_1 + f_0$, it is of course not the case that $\langle\!\langle f_0, f_1 \rangle\!\rangle = \langle\!\langle f_1, f_0 \rangle\!\rangle$ (unless $f_0 = f_1$) and the fact that these witnesses are distinct is an essential aspect of the theory.

Definition 4.8. *(*S*-zero) The pre-summability structure* $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ *has zero if* $0_{X,X}$ *and* Id_X *are summable and* $0 + \mathsf{Id}_X = \mathsf{Id}_X$.

Lemma 4.7. *If the pre-summability structure* $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ *has zero then, for any* $f \in \mathcal{L}(X, Y)$, 0 *and* f *are summable and* $0 + f = f$.

Proof. Easy consequence of Lemma 4.5. □

Definition 4.9. *(*S*-wit) The pre-summability structure* $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ *has witnesses if, for any* $f_0, f_1 \in \mathcal{L}(X, \mathsf{S}Y)$, *if* $\sigma \ f_0$ *and* $\sigma \ f_1$ *are summable, then* f_0 *and* f_1 *are summable.*

Lemma 4.8. *If the pre-summability structure* $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ *satisfies (*S*-wit), then there is a unique natural* $\mathsf{c}_X \in \mathcal{L}(\mathsf{S}^2 X, \mathsf{S}^2 X)$ *such that* $\pi_i \ \pi_j \ \mathsf{c}_X = \pi_j \ \pi_i$ *for all* $i, j \in \{0, 1\}$. *Moreover* $\mathsf{c}_X^2 = \mathsf{Id}_{\mathsf{S}^2 X}$ *and* $\mathsf{c}_X = \langle\!\langle \mathsf{S}\pi_0, \mathsf{S}\pi_1 \rangle\!\rangle$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Proof. Since $\pi_0, \pi_1 \in \mathcal{L}(\mathsf{S}^2X, \mathsf{S}X)$ are summable, we know by Lemma 4.5 that $\pi_i\ \pi_0, \pi_i\ \pi_1 \in \mathcal{L}(\mathsf{S}^2X, X)$ are summable for $i = 0, 1$ which gives us witnesses

$$(f_i = \langle\!\langle \pi_i\ \pi_0, \pi_i\ \pi_1 \rangle\!\rangle = (\mathsf{S}\,\pi_i)\langle\!\langle \pi_0, \pi_1 \rangle\!\rangle = \mathsf{S}\,\pi_i \in \mathcal{L}(\mathsf{S}^2X, \mathsf{S}X))_{i=0,1}\,.$$

We have $(\sigma\ f_i = \pi_i\ \sigma_{\mathsf{S}X})_{i=0,1}$ by naturality of σ so that σf_0 and σf_1 are summable by Lemma 4.5 again. So by 4.9 (S-wit) the morphisms f_0 and f_1 are summable. We set

$$\mathsf{c}_X = \langle\!\langle f_0, f_1 \rangle\!\rangle = \langle\!\langle \langle\!\langle \pi_0\ \pi_0, \pi_0\ \pi_1 \rangle\!\rangle, \langle\!\langle \pi_1\ \pi_0, \pi_1\ \pi_1 \rangle\!\rangle \rangle\!\rangle = \langle\!\langle \mathsf{S}\,\pi_0, \mathsf{S}\,\pi_1 \rangle\!\rangle\,.$$

This definition implies immediately that $\pi_i\ \pi_j\ \mathsf{c}_X = \pi_j\ \pi_i$ for all $i, j \in \{0, 1\}$ and this characterizes c_X by Lemma 4.4. One proves in the same way naturality, as well as the equation $\mathsf{c}_X^2 = \mathsf{Id}_{\mathsf{S}^2X}$. □

Lemma 4.9. *The following diagram commutes*

$$\begin{array}{ccc} \mathsf{S}^2X & \xrightarrow{\mathsf{c}_X} & \mathsf{S}^2X \\ & {\scriptstyle \sigma_{\mathsf{S}X}} \searrow \quad \swarrow {\scriptstyle \mathsf{S}\sigma_X} & \\ & \mathsf{S}X & \end{array}$$

Proof. For $i \in \{0, 1\}$, we have

$$\begin{aligned} \pi_i\ (\mathsf{S}\sigma_X)\ \mathsf{c}_X &= \sigma_X\ \pi_i\ \mathsf{c}_X \quad \text{by naturality of } \pi_i \\ &= \sigma_X\,\mathsf{S}\,\pi_i \quad \text{by definition of } \mathsf{c}_X \\ &= \pi_i\ \sigma_{\mathsf{S}X} \quad \text{by naturality of } \sigma \end{aligned}$$

and the triangle commutes by joint monicity of the π_i's. □

Now we use these properties of c to prove associativity of our partially defined addition on $\mathcal{L}(X, Y)$.

Lemma 4.10. *Let $(f_{ij} \in \mathcal{L}(X, Y))_{(i,j)\in\{0,1\}}$ be such that f_{i0} and f_{i1} are summable for $i = 0, 1$, and the corresponding sums $(f_{i0} + f_{i1})_{i=0,1}$ are summable. Then, f_{0j} and f_{1j} are summable for $j = 0, 1$, the corresponding sums $(f_{0j} + f_{1j})_{j=0,1}$ are summable, and $(f_{00} + f_{01}) + (f_{10} + f_{11}) = (f_{00} + f_{10}) + (f_{01} + f_{11})$.*

Proof. By 4.9 (S-wit), the morphisms $\langle\!\langle f_{00}, f_{01} \rangle\!\rangle, \langle\!\langle f_{10}, f_{11} \rangle\!\rangle \in \mathcal{L}(X, \mathsf{S}Y)$ are summable; this summability has a witness $\langle\!\langle \langle\!\langle f_{00}, f_{01} \rangle\!\rangle, \langle\!\langle f_{10}, f_{11} \rangle\!\rangle \rangle\!\rangle \in \mathcal{L}(X, \mathsf{S}^2Y)$ so that we can define

$$h = \mathsf{c}_Y\ \langle\!\langle \langle\!\langle f_{00}, f_{01} \rangle\!\rangle, \langle\!\langle f_{10}, f_{11} \rangle\!\rangle \rangle\!\rangle \in \mathcal{L}(X, \mathsf{S}^2Y)\,.$$

For $i \in \{0, 1\}$, let $h_i = \pi_i\ h \in \mathcal{L}(X, \mathsf{S}Y)$. We have $\pi_j\ h_i = \pi_i\ \pi_j\ \langle\!\langle \langle\!\langle f_{00}, f_{01} \rangle\!\rangle, \langle\!\langle f_{10}, f_{11} \rangle\!\rangle \rangle\!\rangle = f_{ji}$ and hence f_{0i} and f_{1i} are summable with $\langle\!\langle f_{0i}, f_{1i} \rangle\!\rangle = h_i$ and sum

$$\begin{aligned} f_{0i} + f_{1i} &= \sigma_Y\ h_i \\ &= \sigma_Y\ \pi_i\,\mathsf{c}_Y\ \langle\!\langle \langle\!\langle f_{00}, f_{01} \rangle\!\rangle, \langle\!\langle f_{10}, f_{11} \rangle\!\rangle \rangle\!\rangle \\ &= \sigma_Y\,(\mathsf{S}\,\pi_i)\ \langle\!\langle \langle\!\langle f_{00}, f_{01} \rangle\!\rangle, \langle\!\langle f_{10}, f_{11} \rangle\!\rangle \rangle\!\rangle \quad \text{by definition of } \mathsf{c}_Y \\ &= \pi_i\ \sigma_{\mathsf{S}Y}\ \langle\!\langle \langle\!\langle f_{00}, f_{01} \rangle\!\rangle, \langle\!\langle f_{10}, f_{11} \rangle\!\rangle \rangle\!\rangle \quad \text{by naturality of } \sigma \end{aligned}$$

so that $f_{00} + f_{10}$ and $f_{01} + f_{11}$ are summable with

$$\langle\!\langle f_{00} + f_{10}, f_{01} + f_{11} \rangle\!\rangle = \sigma_{\mathsf{S}Y}\ \langle\!\langle \langle\!\langle f_{00}, f_{01} \rangle\!\rangle, \langle\!\langle f_{10}, f_{11} \rangle\!\rangle \rangle\!\rangle\,.$$

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

We have

$$\begin{aligned}(f_{00}+f_{10})+(f_{01}+f_{11}) &= \sigma_Y\ \sigma_{\mathsf{S}Y}\ \langle\!\langle \langle\!\langle f_{00}, f_{01}\rangle\!\rangle, \langle\!\langle f_{10}, f_{11}\rangle\!\rangle\rangle\!\rangle \\ &= \sigma_Y\ (\mathsf{S}\sigma_Y)\ \langle\!\langle \langle\!\langle f_{00}, f_{01}\rangle\!\rangle, \langle\!\langle f_{10}, f_{11}\rangle\!\rangle\rangle\!\rangle \quad \text{by naturality of } \sigma \\ &= \sigma_Y\ \langle\!\langle \sigma_Y\ \langle\!\langle f_{00}, f_{01}\rangle\!\rangle, \sigma_Y\ \langle\!\langle f_{10}, f_{11}\rangle\!\rangle\rangle\!\rangle \quad \text{by Lemma 4.5} \\ &= \sigma_Y\ \langle\!\langle f_{00}+f_{01}, f_{10}+f_{11}\rangle\!\rangle \\ &= (f_{00}+f_{01})+(f_{10}+f_{11})\,.\end{aligned}$$

□

Lemma 4.11. *If $f_0, f_1 \in \mathcal{L}(X, Y)$ are summable, then so are $\mathsf{S}f_0, \mathsf{S}f_1 \in \mathcal{L}(\mathsf{S}X, \mathsf{S}Y)$ with $\langle\!\langle \mathsf{S}f_0, \mathsf{S}f_1\rangle\!\rangle = \mathsf{c}_Y\ \mathsf{S}\langle\!\langle f_0, f_1\rangle\!\rangle$ and $\mathsf{S}f_0 + \mathsf{S}f_1 = \mathsf{S}(f_0 + f_1)$.*

Proof. We have

$$\begin{aligned}\pi_i\ \pi_j\ \mathsf{c}_Y\ \mathsf{S}\langle\!\langle f_0, f_1\rangle\!\rangle &= \pi_j\ \pi_i\ \mathsf{S}\langle\!\langle f_0, f_1\rangle\!\rangle \\ &= \pi_j\ \langle\!\langle f_0, f_1\rangle\!\rangle\ \pi_i \quad \text{by naturality of } \pi_i \\ &= f_j\ \pi_i \\ &= \pi_i\ \mathsf{S}f_j \\ &= \pi_i\ \pi_j\ \langle\!\langle \mathsf{S}f_0, \mathsf{S}f_1\rangle\!\rangle\end{aligned}$$

which proves the first statement. The second follows easily from Lemma 4.5 using as usual the joint monicity of the π_i's. □

Definition 4.10. *A pre-summability structure $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ which satisfies (S-com), (S-zero) and (S-wit) is called a summability structure and a category $\mathcal{L}$ equipped with a summability structure is called a summable category.*

Theorem 4.2. *Any summable category is enriched in partial commutative monoids by the partial addition operation induced by the summability structure (with neutral elements $0_{X,Y}$).*

Proof. It suffices to prove the associativity condition in Definition 4.1. It results from Lemma 4.10 and 4.8 (S-zero), upon taking $f_{01} = 0$. □

In particular, we can speak of finite summable families of morphisms $\vec{f} = (f_i \in \mathcal{L}(X, Y))_{i\in I}$ and of their sum $\sum \vec{f} = \sum_{i\in I} f_i$ without particular cautions.

Theorem 4.3. *If $(f_i \in \mathcal{L}(X, Y))_{i\in I}$ and $(g_j \in \mathcal{L}(Y, Z))_{j\in J}$ are finite summable families of morphisms, then the family $(g_j f_i \in \mathcal{L}(X, Z))_{(i,j)\in I\times J}$ is summable and*

$$\sum_{(i,j)\in I\times J} g_j f_i = \Big(\sum_{j\in J} g_j\Big)\Big(\sum_{i\in I} f_i\Big)\,.$$

Proof sketch. One takes repetition-free enumerations $(i_l)_{l=1}^m$ and $(j_k)_{k=1}^n$ of I and J and proves the result by induction on $m + n$ coming back to Definition 4.2 and using Lemma 4.5. □

5.3 The monad structure of S

We assume that $\mathcal{L}$ is equipped with a summability structure, and we use the notations introduced above. By 4.8 (S-zero) and 4.7 (S-com), there are natural morphisms $(\iota_i \in \mathcal{L}(X, \mathsf{S}X))_{i=0,1}$ given by $\iota_0 = \langle\!\langle \mathsf{Id}_X, 0\rangle\!\rangle$ and $\iota_1 = \langle\!\langle 0, \mathsf{Id}_X\rangle\!\rangle$, and we have $\sigma\ \iota_i = \mathsf{Id}_X$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Lemma 4.12. *There is a natural morphism* $\tau_X \in \mathcal{L}(\mathsf{S}^2X, \mathsf{S}X)$ *such that* $\pi_0\ \tau = \pi_0\ \pi_0$ *and* $\pi_1\ \tau = \pi_0\ \pi_1 + \pi_1\ \pi_0$. *In particular, we have* $\tau_X\ \mathsf{c}_X = \tau_X$.

Proof. By Theorem 4.3, the family $(\pi_i\ \pi_j)_{(i,j)\in\{0,1\}^2}$ is summable and hence by Lemma 4.2 the morphisms $\pi_0\ \pi_0, \pi_0\ \pi_1, \pi_1\ \pi_0 \in \mathcal{L}(\mathsf{S}^2X, X)$ are summable so that $\pi_0\ \pi_0, \pi_0\ \pi_1 + \pi_1\ \pi_0$ are summable. We take $\tau_X = \langle\!\langle \pi_0\ \pi_0, \pi_0\ \pi_1 + \pi_1\ \pi_0 \rangle\!\rangle$. Let us prove for instance naturality, so let $f \in \mathcal{L}(X, Y)$, we have

$$\begin{aligned}
\tau_Y\ (\mathsf{S}^2 f) &= \langle\!\langle \pi_0\ \pi_0, \pi_0\ \pi_1 + \pi_1\ \pi_0 \rangle\!\rangle\ (\mathsf{S}^2 f) \\
&= \langle\!\langle \pi_0\ \pi_0\ (\mathsf{S}^2 f), (\pi_0\ \pi_1 + \pi_1\ \pi_0)\ (\mathsf{S}^2 f) \rangle\!\rangle \quad \text{by Lemma 4.5} \\
&= \langle\!\langle \pi_0\ \pi_0\ (\mathsf{S}^2 f), \pi_0\ \pi_1\ (\mathsf{S}^2 f) + \pi_1\ \pi_0\ (\mathsf{S}^2 f) \rangle\!\rangle \quad \text{by Lemma 4.5} \\
&= \langle\!\langle f\ \pi_0\ \pi_0, f\ \pi_0\ \pi_1 + f\ \pi_1\ \pi_0 \rangle\!\rangle \quad \text{by naturality of } \pi_0 \text{ and } \pi_1 \\
&= (\mathsf{S}f)\ \langle\!\langle \pi_0\ \pi_0, \pi_0\ \pi_1 + \pi_1\ \pi_0 \rangle\!\rangle \quad \text{by Lemma 4.5} \\
&= (\mathsf{S}f)\ \tau_X\ .
\end{aligned}$$

Last we have $\pi_0\ \tau_X\ \mathsf{c}_X = \pi_0\ \pi_0\ \mathsf{c}_X = \pi_0\ \pi_0 = \pi_0\ \tau_X$ by definition of c and

$$\begin{aligned}
\pi_1\ \tau_X\ \mathsf{c}_X &= (\pi_1\ \pi_0 + \pi_0\ \pi_1)\ \mathsf{c}_X \\
&= \pi_1\ \pi_0\ \mathsf{c}_X + \pi_0\ \pi_1\ \mathsf{c}_X \quad \text{by Lemma 4.5} \\
&= \pi_0\ \pi_1 + \pi_1\ \pi_0 \quad \text{by definition of } \mathsf{c} \\
&= \pi_1\ \tau_X \quad \text{by Theorem 4.2}\,.
\end{aligned}$$

□

Theorem 4.4. *The triple* $(\mathsf{S}, \iota_0, \tau)$ *is a monad.*

Proof. The monad commutations are easily checked, using as usual the joint monicity of the π_i's. For instance, one proves easily that

$$\begin{aligned}
(\mathsf{S}\,\tau_X)\ \tau_X &= \tau_{\mathsf{S}X}\ \tau_X \\
&= \langle\!\langle \pi_0\ \pi_0\ \pi_0, \pi_1\ \pi_0\ \pi_0 + \pi_0\ \pi_1\ \pi_0 + \pi_0\ \pi_0\ \pi_1 \rangle\!\rangle\,.
\end{aligned}$$

□

Another important structure map of S was overlooked in Ehrhard (2022).

Lemma 4.13. *There is a natural lift morphism* $\mathsf{l}_X \in \mathcal{L}(\mathsf{S}X, \mathsf{S}^2X)$ *characterized by* $\pi_i\ \pi_i\ \mathsf{l} = \pi_i$ *and* $\pi_i\ \pi_{1-i}\ \mathsf{l} = 0$.

Proof. We have $\iota_0 = \langle\!\langle \mathsf{Id}_X, 0 \rangle\!\rangle$ and hence $\iota_0\ \pi_0 = \langle\!\langle \pi_0, 0 \rangle\!\rangle \in \mathcal{L}(\mathsf{S}X, \mathsf{S}X)$ and similarly $\iota_1\ \pi_1 = \langle\!\langle 0, \pi_1 \rangle\!\rangle \in \mathcal{L}(\mathsf{S}X, \mathsf{S}X)$. Therefore by 4.8 (S-zero), we have $\sigma\ \iota_i\ \pi_i = \pi_i$ for $i = 0, 1$ and since π_0 and π_1 are summable, 4.9 (S-wit) implies that $\iota_0\ \pi_0$ and $\iota_1\ \pi_1$ are summable and we set

$$\mathsf{l} = \langle\!\langle \iota_0\ \pi_0, \iota_1\ \pi_1 \rangle\!\rangle \in \mathcal{L}(\mathsf{S}X, \mathsf{S}^2X)$$

which satisfies the announced condition. □

Remark 4.4. The lift morphism l_X plays an important role in tangent categories as it makes it possible to express that the differential is linear in a certain sense. It will play the very same role here.

Equipped with σ as counit and l as comultiplication, it is easy to check that S is also a comonad; it is actually a bimonad (with c as distributive law) but contrarily to the monad structure, the differential distributive law that we will introduce soon does not allow one to extend the comonad structure of S to the Kleisli category of $!_$. Such an extension will be possible in the coherent theory of Taylor expansion developed in Ehrhard and Walch (2025), and more precisely in the analytic

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

situation where any morphism of the Kleisli category of !_ is the sum of its Taylor expansion. See that paper for more information about bimonads and about the extension of the bimonad structure of S when it represents a countable (instead of binary, as here) notion of summability.

5.4 Compatibility of the summability structure with the tensor product and the internal hom

We assume that $\mathcal{L}$ has an SMC structure, that is, a distinguished object 1 (tensor unit) and a binary functor $\otimes\colon \mathcal{L}^2 \to \mathcal{L}$ (tensor product) together with natural isomorphisms $\lambda_X \in \mathcal{L}(1 \otimes X, X)$, $\rho_X \in \mathcal{L}(X \otimes 1, X)$, $\alpha_{X_1,X_2,X_3} \in \mathcal{L}((X_1 \otimes X_2) \otimes X_3, X_1 \otimes (X_2 \otimes X_3))$ and $\gamma_{X_1,X_2} \in \mathcal{L}(X_1 \otimes X_2, X_2 \otimes X_1)$ subject to the well-known Mac Lane coherence conditions.

Definition 4.11. *(S-⊗) A pre-summability structure $(\mathcal{L}, \pi_0, \pi_1, \sigma)$ on an SMC $\mathcal{L}$ satisfies (S-⊗) if for any summable morphisms $f_0, f_1 \in \mathcal{L}(X, Y)$ and $g \in \mathcal{L}(U, V)$, the morphisms $f_0 \otimes g, f_1 \otimes g \in \mathcal{L}(X \otimes U, Y \otimes V)$ are summable and satisfy $(f_0 + f_1) \otimes g = f_0 \otimes g + f_1 \otimes g$.*

Lemma 4.14. *Assume that $(\mathcal{L}, \pi_0, \pi_1, \sigma)$ satisfies (S-⊗). Let $(f_i \in \mathcal{L}(X, Y))_{i\in I}$ and $(g_j \in \mathcal{L}(U, V))_{j\in J}$ be finite summable families of morphisms. Then the family $(f_i \otimes g_j \in \mathcal{L}(X \otimes U, Y \otimes V))_{(i,j)\in I\times J}$ is summable, and we have $(\sum_{i\in I} f_i) \otimes (\sum_{j\in J} g_j) = \sum_{(i,j)\in I\times J} f_i \otimes g_j$.*

Proof sketch. Without loss of generality, we can assume that $I = \{1, \dots, n\}$ and $J = \{1, \dots, p\}$ for $n, p \in \mathbb{N}$ and one proves the result by induction on (n, p) ordered lexicographically (for instance). Notice that when I or J is empty, our assumption that the 0 morphisms satisfy $0 \otimes g = 0$ is essential (first paragraph of Section 5.2). □

Definition 4.12. *An SMC $\mathcal{L}$ is a summable SMC if it is equipped with a summability structure which satisfies (S-⊗).*

Lemma 4.15. *Let $\mathcal{L}$ be a summable SMC. There is a natural morphism*

$$\varphi^1_{X_1,X_2} \in \mathcal{L}(\mathsf{S}X_1 \otimes X_2, \mathsf{S}(X_1 \otimes X_2))$$

such that $(\pi_i \ \varphi^1 = \pi_i \otimes X_2)_{i=0,1}$.

Proof. We know that $\pi_0, \pi_1 \in \mathcal{L}(\mathsf{S}X_1, X_1)$ are summable, so by 4.11 (S-⊗) the morphisms $\pi_0 \otimes X_2, \pi_1 \otimes X_2 \in \mathcal{L}(\mathsf{S}X_1 \otimes X_2, X_1 \otimes X_2)$ are summable, we set $\varphi^1 = \langle\!\langle \pi_0 \otimes X_2, \pi_1 \otimes X_2 \rangle\!\rangle$. Let $(f_i \in \mathcal{L}(X_i, Y_i))_{i=1,2}$, we have

$$\begin{aligned}
\varphi^1 \, (\mathsf{S}f_1 \otimes f_2) &= \langle\!\langle (\pi_0 \otimes X_2) \, (\mathsf{S}f_1 \otimes f_2), (\pi_1 \otimes X_2) \, (\mathsf{S}f_1 \otimes f_2) \rangle\!\rangle \quad \text{by Lemma 4.5} \\
&= \langle\!\langle f_1 \ \pi_0 \otimes f_2, f_1 \ \pi_1 \otimes f_2 \rangle\!\rangle \\
&= \mathsf{S}(f_1 \otimes f_2) \, \varphi^1 \quad \text{by Lemma 4.5}
\end{aligned}$$

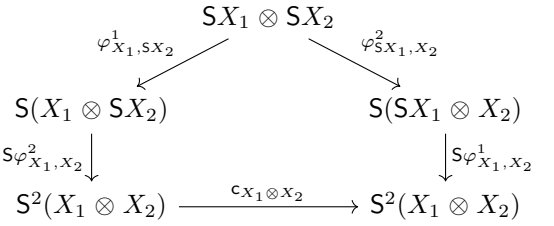
which shows that φ^1 is natural. □

One defines $\varphi^2_{X_1,X_2} \in \mathcal{L}(X_1 \otimes \mathsf{S}X_2, \mathsf{S}(X_1 \otimes X_2))$ by $\varphi^2_{X_1,X_2} = (\mathsf{S}\gamma_{X_2,X_1}) \, \varphi^1_{X_2,X_1} \gamma_{X_1,\mathsf{S}X_2}$.

Theorem 4.5. *The following diagrams commute*

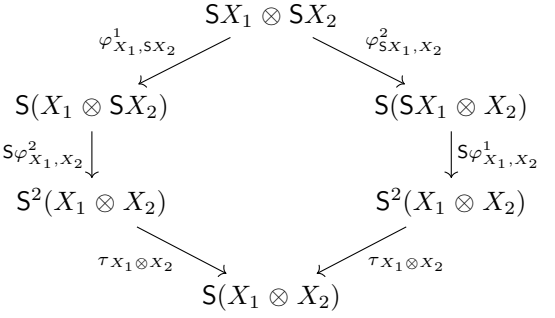
$$\begin{array}{ccc}
X_1 \otimes X_2 & \xrightarrow{\iota_0 \otimes X_2} & \mathsf{S}X_1 \otimes X_2 \\
& \searrow^{\iota_0} & \downarrow{\scriptstyle \varphi^1} \\
& & \mathsf{S}(X_1 \otimes X_2)
\end{array}
\qquad
\begin{array}{ccccc}
\mathsf{S}^2X_1 \otimes X_2 & \xrightarrow{\varphi^1} & \mathsf{S}(\mathsf{S}X_1 \otimes X_2) & \xrightarrow{\mathsf{S}\varphi^1} & \mathsf{S}^2(X_1 \otimes X_2) \\
\downarrow{\scriptstyle \tau \otimes X_2} & & & & \downarrow{\scriptstyle \tau} \\
\mathsf{S}X_1 \otimes X_2 & & \xrightarrow{\qquad\qquad \varphi^1 \qquad\qquad} & & \mathsf{S}(X_1 \otimes X_2)
\end{array}$$

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press



and hence (S, φ^1) *is a commutative strong monad.*

Proof. The diagrams are proven commutative using the joint monicity of the π_i's. Commutativity of the monad means that the following diagram commutes



which results from Lemma 4.12. □

Remark 4.5. It is a standard fact that the common value $\widetilde{\varphi}_{X_1,X_2} = \tau_{X_1\otimes X_2}\ (\mathsf{S}\varphi^1_{X_1,X_2})\,\varphi^2_{\mathsf{S}X_1,X_2} = \tau_{X_1\otimes X_2}\ (\mathsf{S}\varphi^2_{X_1,X_2})\,\varphi^1_{X_1,\mathsf{S}X_2} \in \mathcal{L}(\mathsf{S}X_1 \otimes \mathsf{S}X_2, \mathsf{S}(X_1 \otimes X_2))$ (and associated unit $\iota_0 \in \mathcal{L}(1, \mathsf{S}1)$) turns S into a lax monoidal monad. Notice that this morphism is characterized by

$$\pi_0\ \widetilde{\varphi} = \pi_0 \otimes \pi_0 \quad \text{and} \quad \pi_1\ \widetilde{\varphi} = \pi_1 \otimes \pi_0 + \pi_0 \otimes \pi_1\,.$$

If the SMC $\mathcal{L}$ is closed, with internal hom of X and Y denoted as $(X \multimap Y, \mathsf{ev})$ where $\mathsf{ev} \in \mathcal{L}((X \multimap Y) \otimes X, Y)$ is the evaluation morphism (and, given $f \in \mathcal{L}(Z \otimes X, Y)$, we use $\mathsf{cur}\, f \in \mathcal{L}(Z, X \multimap Y)$ for the currying of f), then we need a further assumption on the summability structure expressing that $\mathsf{S}(X \multimap Y)$ and $X \multimap \mathsf{S}Y$ are isomorphic. More precisely, notice that thanks to 4.11 (S-$\otimes$) we have a morphism $\varphi_{X,Y}$ defined as the following composition of morphisms

$$\mathsf{S}(X \multimap Y) \otimes X \xrightarrow{\varphi^1_{X\multimap Y,X}} \mathsf{S}((X \multimap Y) \otimes X) \xrightarrow{\mathsf{S}\,\mathsf{ev}} \mathsf{S}X$$

so that $\varphi^{\multimap}_{X,Y} = \mathsf{cur}\,\varphi_{X,Y} \in \mathcal{L}(\mathsf{S}(X \multimap Y), X \multimap \mathsf{S}Y)$.

Definition 4.13. *(S-$\multimap$) We say that $\mathcal{L}$ satisfies (S-$\multimap$) if the morphism $\varphi^{\multimap}_{X,Y}$ is an iso. If a summable SMC is closed, we always assume that it satisfies (S-$\multimap$).*

Lemma 4.16. *Let $\mathcal{L}$ be a summable symmetric monoidal closed category (SMCC). If $(f_i \in \mathcal{L}(Z \otimes X, Y))_{i=0,1}$ are summable, then so are $(\mathsf{cur}\, f_i \in \mathcal{L}(Z, X \multimap Y))_{i=0,1}$ and we have $\mathsf{cur}\,(f_0 + f_1) = \mathsf{cur}\, f_0 + \mathsf{cur}\, f_1$.*

Proof. We have $\mathsf{cur}\,\langle\!\langle f_0, f_1\rangle\!\rangle \in \mathcal{L}(Z, X \multimap \mathsf{S}Y)$ and hence $(\varphi^{\multimap})^{-1}\ \mathsf{cur}\,\langle\!\langle f_0, f_1\rangle\!\rangle \in \mathcal{L}(Z, \mathsf{S}(X \multimap Y))$. By naturality of $\varphi^{\multimap}$ we have $\pi_i\ (\varphi^{\multimap})^{-1}\ \mathsf{cur}\,\langle\!\langle f_0, f_1\rangle\!\rangle = (X \multimap \pi_i)\ \mathsf{cur}\,\langle\!\langle f_0, f_1\rangle\!\rangle =$

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

$\mathsf{cur}\,(\,\pi_i\,\langle\!\langle f_0, f_1\rangle\!\rangle) = \mathsf{cur}\, f_i$ for $i = 0, 1$. Hence, $(\,\mathsf{cur}\, f_i)_{i=0,1}$ are summable with $\mathsf{cur}\, f_0 + \mathsf{cur}\, f_1 = \sigma\,(\varphi^{-\circ})^{-1}\ \mathsf{cur}\,\langle\!\langle f_0, f_1\rangle\!\rangle = \mathsf{cur}\,(f_0 + f_1)$ by the same kind of computation. □

5.5 Summability in a cartesian category

We assume now that $\mathcal{L}$ is cartesian, that is, any finite family $(X_i)_{i\in I}$ has a cartesian product $(\,\&_{i\in I}\, X_i, (\mathsf{pr}_i)_{i\in I})$ where the $\mathsf{pr}_j \in \mathcal{L}(\,\&_{i\in I}\, X_i, X_j)$ are the projections; we use $\top$ for the terminal object. When $(f_i \in \mathcal{L}(Y, X_i))_{i\in I}$, we use $\langle f_i\rangle_{i\in I}$ for the unique morphism $Y \to \&_{i\in I}\, X_i$ which, post-composed with pr_j, yields f_j.

Definition 4.14. *(S-&) A pre-summability structure $(\mathcal{L}, \pi_0, \pi_1, \sigma)$ satisfies (S-&) or is cartesian if, for any finite family $\vec{X} = (X_i)_{i\in I}$ of objects, the morphism $\langle \mathsf{Spr}_i\rangle_{i\in I} \in \mathcal{L}(\mathsf{S}(\,\&_{i\in I}\, X_i), \&_{i\in I}\, \mathsf{S}X_i)$ is an iso. We use then $\widetilde{\psi}_{\vec{X}} \in \mathcal{L}(\,\&_{i\in I}\, \mathsf{S}X_i), \mathsf{S}(\,\&_{i\in I}\, X_i))$ for the inverse of $\langle \mathsf{Spr}_i\rangle_{i\in I}$, which is a strong monoidal structure for the functor S wrt. the monoidal structure induced on $\mathcal{L}$ by its cartesian product.*

In the sequel, when dealing with a (pre-)summability structure on a cartesian category, we always assume that 4.14 (S-&) holds.

5.6 Summability in the representable situation

It turns out that most nontrivial summability structures result from very simple properties of the category $\mathcal{L}$ that we describe now: such summability structures will be called *representable*.[6] We assume to be given an SMC category $\mathcal{L}$ with zero-morphisms which is cartesian.[7]

Remark 4.6. The internal hom $(1 \multimap X, \mathsf{ev}\,)$ exists: we can take $(1 \multimap X) = X$ and $\mathsf{ev} = \rho_X \in \mathcal{L}(X \otimes 1, X)$. We will always use this particular version of this internal hom.

We set $\mathbb{D} = 1 \,\&\, 1$, which will play a role similar to that of an object of infinitesimals in synthetic differential geometry Kock (2010); Cockett and Cruttwell (2014), see also Remark 4.12.

We assume that, for all objects X of $\mathcal{L}$, the internal hom $(\mathbb{D} \multimap X, \mathsf{ev}\,)$ exists, where $\mathsf{ev} \in \mathcal{L}((\mathbb{D} \multimap)\, X \otimes \mathbb{D}, X)$ is the evaluation morphisms. In that way we define a functor $\mathsf{S} = (\mathbb{D} \multimap _) : \mathcal{L} \to \mathcal{L}$. Since $\mathcal{L}$ has zero-morphisms, we can define $\overline{\pi}_0^{\&} = \langle \mathsf{Id}_1, 0\rangle$, $\overline{\pi}_1^{\&} = \langle 0, \mathsf{Id}_1\rangle$ and $\Delta^{\&} = \langle \mathsf{Id}_1, \mathsf{Id}_1\rangle$ which all belong to $\mathcal{L}(1, \mathbb{D})$. Notice that if $f \in \mathcal{L}(1, \mathbb{D})$, then $f \multimap X \in \mathcal{L}(\mathsf{S}X, X)$ is a natural transformation.

Definition 4.15. *($\mathbb{D}$-epi) The category $\mathcal{L}$ is representably pre-summable if, for any object X of $\mathcal{L}$, the morphisms $X \otimes \overline{\pi}_0^{\&}$ and $X \otimes \overline{\pi}_1^{\&}$ are jointly epic.*

Remark 4.7. If $\mathcal{L}$ satisfies ($\mathbb{D}$-epi), then $\overline{\pi}_0^{\&}$ and $\overline{\pi}_1^{\&}$ are jointly epic, and if $\mathcal{L}$ is an SMCC, then this latter property implies ($\mathbb{D}$-epi).

Lemma 4.17. *If $\mathcal{L}$ satisfies 4.15 ($\mathbb{D}$-epi), then $\left(\mathsf{S}, \pi_0 = \left(\overline{\pi}_0^{\&} \multimap X\right), \pi_1 = \left(\overline{\pi}_1^{\&} \multimap X\right), \sigma = \left(\Delta^{\&} \multimap X\right)\right)$ is a pre-summability structure on $\mathcal{L}$, that is, π_0 and π_1 are jointly monic. Moreover, the conditions (S-com) and (S-zero) hold.*

In this pre-summability structure, saying that two morphisms $f_0, f_1 \in \mathcal{L}(X, Y)$ are summable means that there is an $f \in \mathcal{L}(X \otimes \mathbb{D}, Y)$ such that $f_i = f\,(\mathbb{D} \otimes \overline{\pi}_i^{\&})\,\rho^{-1}$, and then we have $f_0 + f_1 = f\,(\mathbb{D} \otimes \Delta^{\&})\,\rho^{-1}$. We set $\langle\!\langle f_0, f_1\rangle\!\rangle^{\otimes} = f$, so that $\langle\!\langle f_0, f_1\rangle\!\rangle = \mathsf{cur}\langle\!\langle f_0, f_1\rangle\!\rangle^{\otimes}$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Lemma 4.18. *For any object X and any $n \in \mathbb{N}$, the morphisms $(X \otimes \overline{\pi}^{\&}_{i_1} \otimes \cdots \otimes \overline{\pi}^{\&}_{i_n})_{\vec{i} \in \{0,1\}^n}$ are jointly epic.*

Proof. Simple induction on n. □

Definition 4.16. *We say that $\mathcal{L}$ is representably summable if it is representably pre-summable, and the induced summability structure satisfies* (S-wit).

It is not particularly enlightening to unfold this definition and express directly the condition 4.9 (S-wit) in terms of $\mathbb{D}$, see Ehrhard (2023b), beginning of Section 4 and more precisely Theorem 12.

Remark 4.8. Being representably summable is a property of a category $\mathcal{L}$, and not an additional structure (contrarily to the general notion of summability structure).

► **Example. 4.1.** *Remember that the cartesian product in* **Coh** *of a family $(E_i)_{i\in I}$ is $E = \&_{i\in I} E_i$ defined by $|E| = \bigcup_{i\in I}\{i\} \times |E_i|$ with $(i,a) \frown_E (i',a')$ if $i = i' \Rightarrow a \frown_{E_i} a'$ and projections $\mathsf{pr}_i = \{((i,a),a) \mid i \in I \text{ and } a \in |E_i|\} \in \mathbf{Coh}(E, E_i)$ as easily checked. Given $(t_i \in \mathbf{Coh}(F,E_i))_{i\in I}$, the unique $\langle t_i\rangle_{i\in I} \mathbf{Coh}(F,E)$ such that $(\mathsf{pr}_j\ \langle t_i\rangle_{i\in I} = t_j)_{j\in I}$ is $\langle t_i\rangle_{i\in I} = \{(b,(i,a)) \mid i \in I \text{ and } (b,a) \in t_i\}$.*

The category **Coh** *has zero-morphisms (with $0_{E,F} = \emptyset \in \mathrm{Cl}(E \multimap F) = \mathbf{Coh}(E,F)$). The object $\mathbb{D} = 1 \,\&\, 1$ can be described by $|\mathbb{D}| = \{0,1\}$ with $0 \frown_{\mathbb{D}} 1$ and we have $\overline{\pi}^{\&}_i = \{(*,i)\}$ and $\Delta^{\&} = \{(*,0),(*,1)\}$. It is easy to check that* **Coh** *is representably summable. The induced functor $\mathsf{S} : \mathbf{Coh} \to \mathbf{Coh}$ can be described directly as follows: $|\mathsf{S}E| = \{0,1\} \times |E|$ and $(i,a) \frown_{\mathsf{S}E} (i',a')$ if $a \frown_E a'$ and $i \neq i' \Rightarrow a \neq a'$. Therefore, up to a trivial order isomorphism, $\mathrm{Cl}(\mathsf{S}E) = \{(x_0,x_1) \in \mathrm{Cl}(E)^2 \mid x_0 \cup x_1 \in \mathrm{Cl}(E)$ and $x_0 \cap x_1 = \emptyset\}$ (with the product order). Therefore, two morphisms $t_0, t_1 \in \mathbf{Coh}(E,F)$ are summable iff $t_0 \cup t_1 \in \mathbf{Coh}(E,F)$ (which is not surprising) and $t_0 \cap t_1 = \emptyset$ (which is more), and then their sum is $t_0 \cup t_1$. The corresponding witness is $\langle\!\langle t_0, t_1\rangle\!\rangle = \{(a,(i,b)) \mid i \in \{0,1\}$ and $(a,b) \in t_i\} \in \mathbf{Coh}(E, \mathsf{S}F)$. Notice that the second definition of E_x in Section 3.1 is directly related to this notion of summability: $E_x = \{x' \mid (x,x') \in \mathsf{S}E\}$.*

► **Example. 4.2.** *Another crucial example is that of probabilistic coherence space (PCS) introduced in Girard (2004); Danos and Ehrhard (2011). A PCS is a pair $X = (|X|, \mathsf{P}X)$ where $|X|$ is a set and $\mathsf{P}X \subseteq (\mathbb{R}_{\geq 0})^{|X|}$ that we consider as a poset (whose order relation is the product order) and that we assume to satisfy*

- *$\mathsf{P}X$ is nonempty;*
- *for all $a \in |X|$, the set $\{x_a \mid x \in \mathsf{P}X\} \subseteq \mathbb{R}_{\geq 0}$ is bounded and is not reduced to $\{0\}$;*
- *$\mathsf{P}X$ is down-closed (i.e., if $x \in \mathsf{P}X$ and $y \in (\mathbb{R}_{\geq 0})^{|X|}$ satisfy $y \leq x$ then $y \in \mathsf{P}X$) and closed under the lubs of monotone chains (i.e., if $(x(n) \in \mathsf{P}X)_{n\in\mathbb{N}}$ is monotone for the pointwise order, then $(\sup_{n\in\mathbb{N}} x(n)_a)_{a\in|X|} \in \mathsf{P}X$);*
- *$\mathsf{P}X$ is closed under barycentric combinations, that is, if $x, y \in \mathsf{P}X$ and $p \in [0,1]$, then $(1-p)x + py \in \mathsf{P}X$ (the algebraic operations being defined pointwise).*

Remark 4.9. In the literature, the definition of PCSs is most often based on a duality typical of LL; the present definition is equivalent to the duality-based definition as shown in Girard (2004); Ehrhard (2022), and is perhaps more intuitive. More precisely, given $\mathcal{P} \subseteq (\mathbb{R}_{\geq 0})^I$ for some set I, we can define $\mathcal{P}^\perp = \{x' \in (\mathbb{R}_{\geq 0})^I \mid \forall x \in \mathcal{P}\ \langle x,x'\rangle \leq 1\} \subseteq (\mathbb{R}_{\geq 0})^I$ where $\langle x,x'\rangle = \sum_{i\in I} x_i x'_i$. Then, under the assumption that $\mathsf{P}X$ fulfills the two first conditions above, the two last conditions are equivalent to $\mathsf{P}X = \mathsf{P}X^{\perp\perp}$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

A morphism from X to Y is a matrix $t \in (\mathbb{R}_{\geq 0})^{|X|\times|Y|}$ such that for all $x \in \mathsf{P}X$, one has $t \cdot x \in \mathsf{P}Y$ where $t \cdot x = (\sum_{a\in|X|} t_{a,b}x_a)_{b\in|Y|} \in (\mathbb{R}_{\geq 0})^{|Y|}$. These morphisms are in bijective correspondence with the functions $f : \mathsf{P}X \to \mathsf{P}Y$ which are monotone, commute with lubs of monotone sequences and satisfy $f((1-p)x+py) = (1-p)f(x)+pf(y)$: given such an f we define its matrix $t \in (\mathbb{R}_{\geq 0})^{|X|\times|Y|}$ as follows. Let $(a,b) \in |X| \times |Y|$, then there is an $\varepsilon \in \mathbb{R}_{>0}$ such that $\varepsilon \mathsf{e}_a \in \mathsf{P}X$ (where $\mathsf{e}_a \in (\mathbb{R}_{\geq 0})^{|X|}$ is defined by $(\mathsf{e}_a)_{a'} = \boldsymbol{\delta}_{a,a'}$), and we set $t_{a,b} = \varepsilon^{-1} f(\varepsilon \mathsf{e}_a)_b$, which does not depend on the choice of ε. In that way, we have defined a category **Pcoh** where composition is defined by the usual product of matrices: if $s \in \mathbf{Pcoh}(X,Y)$ and $t \in \mathbf{Pcoh}(Y,Z)$, then $t\,s \in \mathbf{Pcoh}(X,Z)$ is defined by $(t\,s)_{a,c} = \sum_{b\in|Y|} s_{a,b}t_{b,c}$. The identity morphism $\mathsf{Id} \in \mathbf{Pcoh}(X,X)$ is the diagonal matrix, $\mathsf{Id}_{a,a'} = \boldsymbol{\delta}_{a,a'}$.

The category **Pcoh** is an SMC. The tensor unit is $1 = (\{*\}, [0,1])$ (upon identifying $(\mathbb{R}_{\geq 0})^{\{*\}}$ with $\mathbb{R}_{\geq 0}$). Given $(x(i) \in \mathsf{P}X_i)_{i=1,2}$, we set $x(1) \otimes x(2) = (x(1)_{a_1}x(2)_{a_2})_{(a_1,a_2)\in|X_1|\times|X_2|} \in (\mathbb{R}_{\geq 0})^{|X_1|\times|X_2|}$. We can define $X_1 \otimes X_2$ as $(|X_1| \times |X_2|, P)$ where P is the least subset of $(\mathbb{R}_{\geq 0})^{|X_1|\times|X_2|}$ which contains all the $x(1)\otimes x(2)$ for $(x(i) \in \mathsf{P}X_i)_{i=1,2}$ and satisfies the two closure properties in the definition of PCS (the two last conditions). To describe more explicitly this operation, it is convenient to introduce the PCS $X \multimap Y = (|X| \times |Y|, Q)$ where $Q = \mathbf{Pcoh}(X,Y)$. Then $X \multimap 1$ is (trivially) isomorphic to the PCS $X^\perp$ where $|X|^\perp = |X|$ and $x' \in (\mathbb{R}_{\geq 0})^{|X|}$ belongs to $\mathsf{P}X^\perp$ if, for all $x \in \mathsf{P}X$, one has $\langle x, x'\rangle = \sum_{a\in|X|} x_a x'_a \leq 1$. It is then possible to prove that $X^{\perp\perp} = X$ (a kind of "bipolar theorem", see Remark 4.9). Then one has $X_1 \otimes X_2 = \left(X_1 \multimap X_2{}^\perp\right)^\perp$ and, based on this property, when equipped with 1 and $\otimes$, the category **Pcoh** is an SMC. This SMC is closed with $(X \multimap Y, \mathsf{ev})$ as internal hom from X to Y, with $\mathsf{ev}_{((a,b),a'),b'} = \boldsymbol{\delta}_{a,a'}\boldsymbol{\delta}_{b,b'} \in \mathcal{L}((X \multimap Y) \otimes X, Y)$ as evaluation morphism, which satisfies of course $\mathsf{ev} \cdot (t \otimes x) = t \cdot x$ for $t \in \mathbf{Pcoh}(X,Y)$ and $x \in \mathsf{P}X$. This SMCC is even $*$-autonomous with $\perp = 1$ as dualizing object (this essentially boils down to the fact that $X^{\perp\perp} = X$).

The category **Pcoh** has all products: given a family $(X_i)_{i\in I}$ of PCSs, let X be defined by $|X| = \bigcup_{i\in I}\{i\} \times |X_i|$ and $x \in (\mathbb{R}_{\geq 0})^{|X|}$ belongs to $\mathsf{P}X$ if $((x_{i,a})_{a\in|X_i|} \in \mathsf{P}X_i)_{i\in I}$. Equipped with $(\mathsf{pr}_i \in \mathbf{Pcoh}(X, X_i))_{i\in I}$ defined by $(\mathsf{pr}_i)_{(j,a),a'} = \boldsymbol{\delta}_{j,i}\boldsymbol{\delta}_{a,a'}$, X is easily seen to be the cartesian product of the X_i's, and we set $\&_{i\in I}\, X_i = X$. Notice that $\mathsf{P}(\&_{i\in I}\, X_i)$ is isomorphic (for the order and for the barycentric structures) to $\prod_{i\in I} \mathsf{P}X_i$. Given $(t_i \in \mathbf{Pcoh}(Y, X_i))_{i\in I}$, the unique $\langle t_i\rangle_{i\in I} \in \mathbf{Pcoh}(Y, \&_{i\in I}\, X_i)$ such that $(\mathsf{pr}_j\, \langle t_i\rangle_{i\in I} = t_j)_{j\in I}$ is given by $(\langle t_i\rangle_{i\in I})_{b,(j,a)} = (t_j)_{b,a}$ for all $j \in I$, $b \in |Y|$ and $a \in |X_j|$.

The category **Pcoh** has zero morphisms (namely the 0 matrix which obviously belongs to all homsets $\mathbf{Pcoh}(X,Y)$). The object $\mathbb{D} = 1 \,\&\, 1$ is described by $|\mathbb{D}| = \{0,1\}$ and $\mathsf{P}\mathbb{D} = [0,1]^2$ (upon identifying $(\mathbb{R}_{\geq 0})^{|\mathbb{D}|}$ with $(\mathbb{R}_{\geq 0})^2$). Then $(\overline{\pi}_0^{\&}, \overline{\pi}_1^{\&}, \Delta^{\&} \in \mathbf{Pcoh}(1, \mathbb{D}))$ are given by $\overline{\pi}_0^{\&} \cdot u = (u, 0)$, $\overline{\pi}_1^{\&} \cdot u = (0, u)$ and $\Delta^{\&} \cdot u = (u, u)$, for all $u \in [0,1]$. Given a PCS X, the PCS $\mathsf{S}X = (\mathbb{D} \multimap X)$ is given by $|\mathsf{S}X| = \{0,1\} \times |X|$ and an element of $\mathsf{P}(\mathsf{S}X)$ is a $t \in (\mathbb{R}_{\geq 0})^{\{0,1\}\times|X|}$ such that $t \cdot (\mathsf{e}_0 + \mathsf{e}_1) \in \mathsf{P}X$ where $\mathsf{e}_0 + \mathsf{e}_1 \in \mathsf{P}\mathbb{D}$ corresponds to the pair $(1,1) \in [0,1]^2$. In other words,

$$\mathsf{P}(\mathsf{S}X) \simeq \{(x(0), x(1)) \in \mathsf{P}X^2 \mid x(0) + x(1) \in \mathsf{P}X\} \tag{2}$$

for the poset and barycentric structures. With this identification, the morphisms $\pi_0, \pi_1, \sigma \in \mathbf{Pcoh}(\mathsf{S}X, X)$ are characterized by $\pi_i \cdot (x(0), x(1)) = x(i)$ and $\sigma \cdot (x(0), x(1)) = x(0) + x(1)$. It follows easily from these observations that two morphisms $t(0), t(1) \in \mathbf{Pcoh}(Y, X)$ are summable iff $\forall y \in \mathsf{P}Y\ t(0) \cdot y + t(1) \cdot y \in \mathsf{P}X$, and the corresponding witness $\langle\!\langle t(0), t(1)\rangle\!\rangle \in \mathbf{Pcoh}(Y, \mathsf{S}X)$ is given as a matrix by $\langle\!\langle t(0), t(1)\rangle\!\rangle_{b,(i,a)} = t(i)_{b,a}$ and characterized by $\langle\!\langle t(0), t(1)\rangle\!\rangle \cdot y = (t(0) \cdot y, t(1) \cdot y)$ if we consider Equation (2) as an equality. From this characterization of summability and witnesses, it follows easily that **Pcoh** is a representably summable category.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Remark 4.10. If $\mathcal{L}$ is additive, that is, enriched in commutative monoids, then it is well known that finite cartesian products are also coproducts. It follows that, in that case, $\mathbb{D} = 1 \oplus 1$ and hence $\mathsf{S}X = (\mathbb{D} \multimap)\, X$ is canonically isomorphic to $X \,\&\, X$ and the summability structure is trivial: all pairs of morphisms $f_0, f_1 \in \mathcal{L}(X, Y)$ are summable with witness $\langle f_0, f_1\rangle \in \mathcal{L}(X, Y \,\&\, Y)$ and sum $f_0 + f_1$ (the addition provided by the enrichment).

Theorem 4.6. *Any representably summable category satisfies* (S-&) *and* (S-⊗).

Proof sketch. The first property results from the fact that S is the right adjoint to the functor $_ \otimes \mathbb{D}$ and as such preserves all existing limits. Next take $(f_i \in \mathcal{L}(X, Y))_{i\in\{0,1\}}$ be summable so that $\langle\!\langle f_0, f_1\rangle\!\rangle^{\otimes} \in \mathcal{L}(X \otimes \mathbb{D}, Y)$, and let $g \in \mathcal{L}(U, V)$. We have $(\langle\!\langle f_0, f_1\rangle\!\rangle^{\otimes} \otimes g)\,(X \otimes \gamma_{U,\mathbb{D}}) \in \mathcal{L}(X \otimes U \otimes \mathbb{D}, Y \otimes V)$. It is easily checked that $(\langle\!\langle f_0, f_1\rangle\!\rangle^{\otimes} \otimes g)\,(X \otimes \gamma_{U,\mathbb{D}}) = \langle\!\langle f_0 \otimes g, f_1 \otimes g\rangle\!\rangle^{\otimes}$. □

We know by Theorem 4.4 that S has a canonical (bi)monad structure. In the representable situation, we can describe more directly this structure by means of a (bi)monoid structure on $\mathbb{D}$. We describe first the comonoid structure.

Proposition 4.1. *There is a unique* $\overline{\tau} \in \mathcal{L}(\mathbb{D}, \mathbb{D} \otimes \mathbb{D})$ *such that* $\overline{\tau}\,\overline{\pi}_0^{\&} = \overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&}$ *and* $\overline{\tau}\,\overline{\pi}_1^{\&} = \overline{\pi}_1^{\&} \otimes \overline{\pi}_0^{\&} + \overline{\pi}_0^{\&} \otimes \overline{\pi}_1^{\&}$ *(keeping the iso* $\lambda_1 = \rho_1 \in \mathcal{L}(1 \otimes 1, 1)$ *implicit). The triple* $(\mathbb{D}, \mathsf{pr}_0, \overline{\tau})$ *is a commutative comonoid.*

Proof. Clearly $\overline{\pi}_0^{\&}, \overline{\pi}_1^{\&} \in \mathcal{L}(1, \mathbb{D})$ are summable (with $\langle\!\langle \overline{\pi}_0^{\&}, \overline{\pi}_1^{\&}\rangle\!\rangle^{\otimes} = \mathsf{Id}_{\mathbb{D}}$ and $\overline{\pi}_0^{\&} + \overline{\pi}_1^{\&} = \Delta^{\&}$). So by Lemma 4.14 the morphisms $(\overline{\pi}_i^{\&} \otimes \overline{\pi}_j^{\&} \in \mathcal{L}(1 \otimes 1, \mathbb{D} \otimes \mathbb{D}))_{(i,j)\in\{0,1\}^2}$ are summable. By Lemma 4.2, the morphisms $\overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&}$, $\overline{\pi}_1^{\&} \otimes \overline{\pi}_0^{\&}$ and $\overline{\pi}_0^{\&} \otimes \overline{\pi}_1^{\&}$ are summable, and hence the morphisms $\overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&}$ and $\overline{\pi}_1^{\&} \otimes \overline{\pi}_0^{\&} + \overline{\pi}_0^{\&} \otimes \overline{\pi}_1^{\&}$ are summable. This gives us a witness $\langle\!\langle \overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&}, \overline{\pi}_1^{\&} \otimes \overline{\pi}_0^{\&} + \overline{\pi}_0^{\&} \otimes \overline{\pi}_1^{\&}\rangle\!\rangle^{\otimes} \in \mathcal{L}(1 \otimes 1 \otimes \mathbb{D}, \mathbb{D} \otimes \mathbb{D})$ and we set

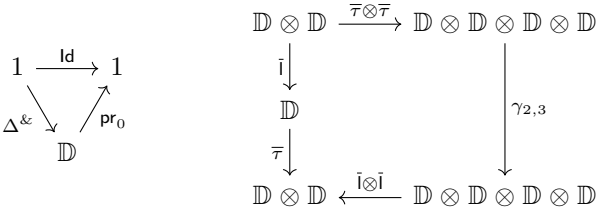
$$\overline{\tau} = \langle\!\langle \overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&}, \overline{\pi}_1^{\&} \otimes \overline{\pi}_0^{\&} + \overline{\pi}_0^{\&} \otimes \overline{\pi}_1^{\&}\rangle\!\rangle^{\otimes}\, \lambda^{-1}\, \lambda^{-1} \in \mathcal{L}(\mathbb{D}, \mathbb{D} \otimes \mathbb{D})$$

which obviously satisfies the announced equations. Uniqueness results from the joint epicity of $\overline{\pi}_0^{\&}$ and $\overline{\pi}_1^{\&}$. The fact that we define in that way a commutative comonoid is easily checked, again by joint epicity of $\overline{\pi}_0^{\&}$ and $\overline{\pi}_1^{\&}$. For instance (keeping the associator α implicit), the morphism $\overline{\tau}^{(3)} = (\overline{\tau} \otimes \mathbb{D})\,\overline{\tau} = (\mathbb{D} \otimes \overline{\tau})\,\overline{\tau} \in \mathcal{L}(\mathbb{D}, \mathbb{D} \otimes \mathbb{D} \otimes \mathbb{D})$ is characterized by $\overline{\tau}^{(3)}\,\overline{\pi}_0^{\&} = \overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&}$ and $\overline{\tau}^{(3)}\,\overline{\pi}_1^{\&} = \overline{\pi}_1^{\&} \otimes \overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&} + \overline{\pi}_0^{\&} \otimes \overline{\pi}_1^{\&} \otimes \overline{\pi}_0^{\&} + \overline{\pi}_0^{\&} \otimes \overline{\pi}_0^{\&} \otimes \overline{\pi}_1^{\&}$. The commutations involving the counit pr_0 result from the fact that $\mathsf{pr}_i\,\overline{\pi}_j^{\&} = \delta_{i,j}\,\mathsf{Id}_1$. □

Theorem 4.7. *There is a unique* $\overline{\mathsf{l}} \in \mathcal{L}(\mathbb{D} \otimes \mathbb{D}, \mathbb{D})$ *such that* $\overline{\mathsf{l}}\,(\overline{\pi}_i^{\&} \otimes \overline{\pi}_j^{\&}) = \delta_{i,j}\overline{\pi}_i^{\&}$ *(keeping the iso* $\lambda_1 = \rho_1 \in \mathcal{L}(1 \otimes 1, 1)$ *implicit). The triple* $(\mathbb{D}, \Delta^{\&}, \overline{\mathsf{l}})$ *is a commutative monoid, and* $(\mathbb{D}, \Delta^{\&}, \overline{\mathsf{l}}, \mathsf{pr}_0, \overline{\tau})$ *is a bicommutative bimonoid.*

Proof. Remember that $\mathbb{D} = 1 \,\&\, 1$, so we can set $\overline{\mathsf{l}} = \langle \lambda_1\,(\mathsf{pr}_0 \otimes \mathsf{pr}_0), \lambda_1\,(\mathsf{pr}_1 \otimes \mathsf{pr}_1)\rangle$ which obviously satisfies the announced property, which characterizes $\overline{\mathsf{l}}$ uniquely by Lemma 4.18. The fact that we define in that way a commutative monoid is easy to check. For instance (keeping implicit the associator α and the λ isos) we have $\overline{\mathsf{l}}^{(3)} = \overline{\mathsf{l}}\,(\overline{\mathsf{l}} \otimes \mathbb{D}) = \overline{\mathsf{l}}\,(\mathbb{D} \otimes \overline{\mathsf{l}}) = \langle \mathsf{pr}_0 \otimes \mathsf{pr}_0 \otimes \mathsf{pr}_0, \mathsf{pr}_1 \otimes \mathsf{pr}_1 \otimes \mathsf{pr}_1\rangle \in \mathcal{L}(\mathbb{D} \otimes \mathbb{D} \otimes \mathbb{D}, \mathbb{D})$. The commutations of the diagram involving the unit $\Delta^{\&}$ result from the fact that $\mathsf{pr}_i\,\Delta^{\&} = \mathsf{Id}_1$ for $i = 0, 1$. The last statement means that the following diagrams commute

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press



where $\gamma_{2,3}$ is the Mac Lane iso which exchanges the 2 central factors of the 4-fold tensor product. The first diagram obviously commutes. For the second we use Lemma 4.18. We have

$$\begin{aligned}\overline{\tau}\,\bar{\mathsf{l}}\,(\overline{\pi}_i^{\&}\otimes\overline{\pi}_j^{\&}) &= \overline{\tau}\,\boldsymbol{\delta}_{i,j}\overline{\pi}_i^{\&}\\ &= \boldsymbol{\delta}_{i,j}\sum_{\substack{(l,r)\in\{0,1\}^2\\ l+r=i}}\overline{\pi}_l^{\&}\otimes\overline{\pi}_r^{\&}\end{aligned}$$

and

$$\begin{aligned}(\bar{\mathsf{l}}\otimes\bar{\mathsf{l}})\,\gamma_{2,3}\,(\overline{\tau}\otimes\overline{\tau})\,(\overline{\pi}_i^{\&}\otimes\overline{\pi}_j^{\&}) &= (\bar{\mathsf{l}}\otimes\bar{\mathsf{l}})\,\gamma_{2,3}\Big(\Big(\sum_{\substack{(l,r)\in\{0,1\}^2\\ l+r=i}}\overline{\pi}_l^{\&}\otimes\overline{\pi}_r^{\&}\Big)\otimes\Big)\Big(\sum_{\substack{(l',r')\in\{0,1\}^2\\ l'+r'=j}}\overline{\pi}_l^{\&}\otimes\overline{\pi}_r^{\&}\Big)\\ &= (\bar{\mathsf{l}}\otimes\bar{\mathsf{l}})\,\gamma_{2,3}\sum_{\substack{(l,r,l',r')\in\{0,1\}^4\\ l+r=i,\ l'+r'=j}}\overline{\pi}_l^{\&}\otimes\overline{\pi}_r^{\&}\otimes\overline{\pi}_{l'}^{\&}\otimes\overline{\pi}_{r'}^{\&}\\ &= (\bar{\mathsf{l}}\otimes\bar{\mathsf{l}})\sum_{\substack{(l,r,l',r')\in\{0,1\}^4\\ l+r=i,\ l'+r'=j}}\overline{\pi}_l^{\&}\otimes\overline{\pi}_{l'}^{\&}\otimes\overline{\pi}_r^{\&}\otimes\overline{\pi}_{r'}^{\&}\\ &= \sum_{\substack{(l,r,l',r')\in\{0,1\}^4\\ l+r=i,\ l'+r'=j}}\boldsymbol{\delta}_{l,l'}\boldsymbol{\delta}_{r,r'}\overline{\pi}_l^{\&}\otimes\overline{\pi}_r^{\&}\\ &= \boldsymbol{\delta}_{i,j}\sum_{\substack{(l,r)\in\{0,1\}^2\\ l+r=i}}\overline{\pi}_l^{\&}\otimes\overline{\pi}_r^{\&}\end{aligned}$$

□

Then it can be checked that the bimonad structure of S is induced by this bimonoid structure of $\mathbb{D}$: the unit of the monad is $\iota_0 = \big(\mathsf{pr}_0 \multimap X\big) \in \mathcal{L}(\mathsf{S}X, X)$ (identifying X and $1 \multimap X$) and its multiplication is $\tau = (\overline{\tau} \multimap X) \in \mathcal{L}(\mathsf{S}X, \mathsf{S}^2X)$ (identifying $\mathbb{D}\otimes\mathbb{D}\multimap X$ and $\mathsf{S}^2X = (\mathbb{D}\multimap(\mathbb{D}\multimap X))$ which are canonically isomorphic). The comonad structure of S can be described similarly using the unit $\Delta^{\&}$ and the product $\bar{\mathsf{l}}$ of the monoid structure of $\mathbb{D}$. The distributive law of the bimonad, which is the flip isomorphism $\mathsf{c}_X \in \mathcal{L}(\mathsf{S}^2X, \mathsf{S}^2X)$, is obtained similarly from the braiding of the SMC structure of $\mathcal{L}$: $\mathsf{c}_X = \big(\gamma_{\mathbb{D},\mathbb{D}} \multimap X\big)$ (leaving again implicit the canonical iso between $\mathbb{D}\otimes\mathbb{D}\multimap X$ and S^2X).

Remark 4.11. The general categorical concept of mate provides a systematic understanding of this correspondence. For instance, the commutative comonoid structure of $\mathbb{D}$ induces very easily a comonad $\mathsf{S}^{\otimes}$ on $\mathcal{L}$ such that $\mathsf{S}^{\otimes}X = X\otimes\mathbb{D}$. Then the monad $(\mathsf{S}, \iota_0, \tau)$ is the mate of the comonad $\mathsf{S}^{\otimes}$ through the adjunction $_\otimes\mathbb{D} \dashv \mathbb{D}\multimap_$. This point of view is developed in Ehrhard and Walch (2023a).

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

► **Example. 4.3.** *In* **Coh**, *the bimonoid structure of* $\mathbb{D}$ *can be described as follows.*

- *Counit* $\mathsf{pr}_0 = \{(0, *)\} \in \mathbf{Coh}(\mathbb{D}, 1)$;
- *Comultiplication* $\overline{\tau} = \{(0, (0, 0)), (1, (1, 0)), (1, (0, 1))\} \in \mathbf{Coh}(\mathbb{D}, \mathbb{D} \otimes \mathbb{D})$;
- *Unit* $\Delta^{\&} = \{(*, 0), (*, 1)\} \in \mathbf{Coh}(1, \mathbb{D})$;
- *Multiplication* $\overline{\mathsf{l}} = \{((0, 0), 0), ((1, 1), 1))\} \in \mathbf{Coh}(\mathbb{D} \otimes \mathbb{D}, \mathbb{D})$.

► **Example. 4.4.** *Let us describe the bimonoid structure of* $\mathbb{D}$ *in* **Pcoh**. *An element* u *of* $\mathsf{P}\mathbb{D}$ *can be written uniquely* $u = u_0\mathsf{e}_0 + u_1\mathsf{e}_1$ *where* $u_0, u_1 \in [0, 1]$, *so that* $\mathsf{P}\mathbb{D} \simeq [0, 1]^2$ *as already mentioned.*

- *The unit* $\mathsf{pr}_0 \in \mathbf{Pcoh}(\mathbb{D}, 1)$ *is characterized by* $\mathsf{pr}_0 \cdot u = u_0$;
- *the comultiplication* $\overline{\tau} \in \mathbf{Pcoh}(\mathbb{D}, \mathbb{D} \otimes \mathbb{D})$ *is characterized by* $\overline{\tau} \cdot u = u_0\mathsf{e}_{(0,0)} + u_1(\mathsf{e}_{(1,0)} + \mathsf{e}_{(0,1)})$;
- *the unit* $\Delta^{\&} \in \mathbf{Pcoh}(1, \mathbb{D})$ *is characterized by* $\Delta^{\&} \cdot p = p(\mathsf{e}_0 + \mathsf{e}_1)$;
- *the multiplication* $\overline{\mathsf{l}} \in \mathbf{Pcoh}(\mathbb{D} \otimes \mathbb{D}, \mathbb{D})$ *is characterized by* $\overline{\mathsf{l}} \cdot (u \otimes v) = u_0v_0\mathsf{e}_0 + u_1v_1\mathsf{e}_1$.

Remark 4.12. There is a strong similarity between our object $\mathbb{D}$ and the infinitesimal object D introduced in synthetic differential geometry (SDG) for describing the tangent bundle functor as a representable functor $M \mapsto M^D$, see Cockett and Cruttwell (2014) Section 5.2. This similarity is not surprising since the purpose of $\mathbb{D}$ is precisely to present the summability functor S as $\mathbb{D} \multimap _$. More precisely, the properties of $\mathbb{D}$ listed in the present section and in Section 6.3 have strong similarities with the properties of D listed in Definition 5.6 of Cockett and Cruttwell (2014). As an example, the comonoid structure of D (Condition [**Infsml.3**]) corresponds to the comonoid structure of $\mathbb{D}$ of Proposition 4.1 and its purpose is similar: it allows one to endow S with a natural transformation which is the multiplication of a monad structure playing the same role as the local addition natural transformation of the tangent bundle functor, called bundle addition in Cockett and Cruttwell (2014).

The main difference between the two settings lies in the ambient category itself (the category $\mathcal{L}$ in our case and the category $\mathbb{X}$ in Cockett and Cruttwell (2014)) which are of different natures and carry different intuitions.

The objects of $\mathbb{X}$ are of a geometric nature and should be thought of as manifolds, the only property we can always reasonably expect to hold in $\mathbb{X}$ is the existence of all finite cartesian products. Therefore $\mathbb{X}$ has typically no "multi-linear morphisms" which would be associated with a specific monoidal structure, distinct from the cartesian product as in a (multi)category of vector spaces and (multi)linear morphisms. Our category $\mathcal{L}$ is typically cartesian and co-cartesian, although products do typically not distribute over coproducts. And $\mathcal{L}$ has also a symmetric monoidal structure for which it is typically closed and its objects are of an algebraic nature, close to vector spaces, see Example 4.2. This difference appears clearly in the comonoid structures mentioned above: for the infinitesimal object D of $\mathbb{X}$, this comonoid is relative to a kind of coproduct of D with itself in $\mathbb{X}$ (see Cockett and Cruttwell (2014), Definition 5.6 [**Infsml.3**]), whereas the comonoid structure of $\mathbb{D}$ in $\mathcal{L}$ is relative to the monoidal product of $\mathcal{L}$.

Let us mention last that an infinitesimal object D is defined as a subobject of a rig of line type in $\mathbb{X}$, see Cockett and Cruttwell (2014) Definition 5.1, and it is explained just after this definition that "a category with a rig of line type is necessarily rather special" and quite difficult to build. In contrast, it seems very easy to find categories $\mathcal{L}$ where the object $\mathbb{D} = 1 \,\&\, 1$ satisfies our requirements.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

6. The Differential Structure

Using the summability structures described above, we introduce now an operation of differentiation. We adopt the viewpoint of differential LL that consists in associating such an operation with the *exponential modality* of LL, which is a unary connective $!A$ interpreted categorically as a comonad. Indeed, in differential LL, differentiation is represented as a morphism of type $A \multimap !A$, which allows one to transform by precomposition a morphism $f : !A \multimap B$ (to be understood as a nonlinear morphism $A \to B$) into a linear morphism $A \multimap B$: the derivative of f at 0.

In the present setting where we insist on making differentiation compatible with determinism as explained in Section 4, it turns out that differentiation can be represented as a distributive law between S and the exponential modality. We first give a minimal set of axioms on a SMC $\mathcal{L}$ to make this axiomatization of coherent differentiation possible.

Definition 5.1. *A resource category is a category $\mathcal{L}$ such that*

- *$\mathcal{L}$ is an SMC with zero-morphisms;*
- *$\mathcal{L}$ is cartesian;*
- *$\mathcal{L}$ is equipped with a resource modality, that is, a tuple $(!, \mathsf{der}, \mathsf{dig}, \mathsf{m}^0, \mathsf{m}^2)$ where $! : \mathcal{L} \to \mathcal{L}$ is a functor, der (dereliction) and dig (digging) are respectively the counit and the comultiplication of a comonad structure on this functor and $(\mathsf{m}^0 \in \mathcal{L}(1, !\top))$, an iso, and $\mathsf{m}^2_{X_1,X_2} \in \mathcal{L}(!X_1 \otimes !X_2)$, a natural iso, turn ! into a symmetric monoidal comonad from the SMC $(\mathcal{L}, \top, \&)$ to the SMC $(\mathcal{L}, 1, \otimes)$. These isos are called the Seely isomorphisms of $\mathcal{L}$.*

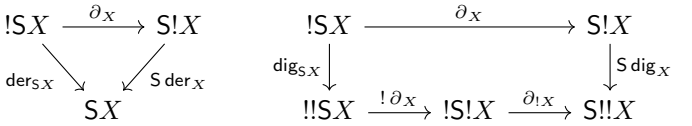
We assume that $\mathcal{L}$ is such a resource category. The resource modality induces a Kleisli category $\mathcal{L}_!$ whose objects are those of $\mathcal{L}$ and where $\mathcal{L}_!(X, Y) = \mathcal{L}(!X, Y)$. In this category, the identity morphisms are $\mathsf{Id}^{(!)}_X = \mathsf{der}_X$, and composition of $f \in \mathcal{L}_!(X, Y)$ and $g \in \mathcal{L}_!(Y, Z)$ is defined by $g \circ f = g\ (!f)\ \mathsf{dig}_X$.

The basic intuition in this situation is that the morphisms of $\mathcal{L}$ are linear whereas $\mathcal{L}_!$ is a category of nonlinear morphisms. Here the word "linear" can be used in its algebraic and its computer science meaning. This intuition is supported by the fact that there is a functor $\mathsf{Der} : \mathcal{L} \to \mathcal{L}_!$ which acts as the identity on objects and maps $f \in \mathcal{L}(X, Y)$ to $\mathsf{Der}\,(f) = f\ \mathsf{der}_X \in \mathcal{L}_!(X, Y)$. This functor is not necessarily faithful (it is, in most known categorical models of LL), but it should nevertheless be considered as a kind of "inclusion" of $\mathcal{L}$ into the larger $\mathcal{L}_!$.

We assume from now on that $\mathcal{L}$ is equipped with a summability structure (remember that this means in particular that 4.11 (S-⊗) and 4.14 (S-&) hold).

The main idea of coherent differentiation is to associate with any nonlinear morphism $f \in \mathcal{L}_!(X, Y)$ a "derivative" $\mathsf{T}f \in \mathcal{L}_!(\mathsf{S}X, \mathsf{S}Y)$, which intuitively maps a summable pair (x_0, x_1) of elements of X to the summable pair $(f(x_0), f'(x_0) \cdot x_1)$, and the chain rule of Calculus tells us that this operation T should be functorial. In other words, T should be an extension of the functor S to $\mathcal{L}_!$ in the sense that if $f \in \mathcal{L}(X, Y)$, one has $\mathsf{T}(\mathsf{Der}\,(f)) = \mathsf{Der}\,(\mathsf{S}f)$. Intuitively this condition means that the derivative of a linear map is the map itself. It is known that such extensions are in one-to-one correspondence with *distributive laws* between the functor S and the comonad !_.

Definition 5.2. *(∂-chain) A predifferential structure on $\mathcal{L}$ is a distributive law between the functor* S *and the comonad* !_*, that is, a natural transformation $\partial_X \in \mathcal{L}(!\mathsf{S}X, \mathsf{S}!X)$ such that the following diagrams commute*



https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Then the extended functor $\mathsf{T} : \mathcal{L}_! \to \mathcal{L}_!$ is defined by $\mathsf{T}X = \mathsf{S}X$ and $\mathsf{T}f = (\mathsf{S}f)\ \partial_X \in \mathcal{L}(!\mathsf{S}X, \mathsf{S}Y)$ for $f \in \mathcal{L}(!X, Y)$.

This simple condition is not sufficient for specifying a differential operation. Here are the additional conditions.

Definition 5.3. *(∂-local)*

$$\begin{array}{ccc} !\mathsf{S}X & \xrightarrow{\partial_X} & \mathsf{S}!X \\ & {\scriptstyle !\pi_0}\searrow \quad \swarrow{\scriptstyle \pi_0} & \\ & !X & \end{array}$$

This condition means that our differentiation operator, when applied to a nonlinear morphism $f : X \to Y$, will yield a nonlinear morphism $\mathsf{T}f : \mathsf{S}X \to \mathsf{S}Y$ whose first component $\mathsf{S}X \to Y$ maps all summable pairs (x_0, x_1) to $f(x_0)$. The tangent bundle functor behaves exactly in the same way, and this condition expresses that the differential is computed locally (in this differential geometric setting, x_0 is an element of the manifold X, and x_1 is a tangent vector at x_0). This justifies our terminology.

Definition 5.4. *(∂-add) The natural transformation ∂ is also a distributive law between the functor* ! *and the monad* $(\mathsf{S}, \iota_0, \tau)$*, that is,*

$$\begin{array}{ccc} & !X & \\ {\scriptstyle !\iota_0}\swarrow & & \searrow{\scriptstyle \iota_0} \\ !\mathsf{S}X & \xrightarrow{\partial_X} & \mathsf{S}!X \end{array} \qquad \begin{array}{ccccc} !\mathsf{S}^2X & \xrightarrow{\partial_{\mathsf{S}X}} & \mathsf{S}!\mathsf{S}X & \xrightarrow{\mathsf{S}\,\partial_X} & \mathsf{S}^2!X \\ {\scriptstyle !\tau_X}\downarrow & & & & \downarrow{\scriptstyle \tau_{!X}} \\ !\mathsf{S}X & & \xrightarrow{\qquad\partial_X\qquad} & & \mathsf{S}!X \end{array}$$

This means that the comonad ! can be extended to the Kleisli category of the monad S. Due to 4.11 (S-⊗) and 4.14 (S-&), this latter Kleisli category is monoidal and cartesian so that, when 5.4 (∂-add) holds, it becomes a resource category which can be understood as a categorical version of Clifford's ring of *dual numbers*.

More concretely, the condition 5.4 (∂-add) means that derivatives are additive morphisms, that is, preserve 0 and (the partially defined) addition of morphisms.

Definition 5.5. *(∂-&)*

$$\begin{array}{ccc} 1 & \xrightarrow{\iota_0} & \mathsf{S}1 \\ {\scriptstyle \mathsf{m}^0}\downarrow & & \downarrow{\scriptstyle \mathsf{S}\mathsf{m}^0} \\ !\top & & \\ {\scriptstyle !0}\downarrow & & \\ !\mathsf{S}\top & \xrightarrow{\partial_\top} & \mathsf{S}!\top \end{array} \qquad \begin{array}{ccccc} !\mathsf{S}X_1 \otimes !\mathsf{S}X_1 & \xrightarrow{\partial_{X_1}\otimes\partial_{X_2}} & \mathsf{S}!X_1 \otimes \mathsf{S}!X_2 & \xrightarrow{\widetilde{\varphi}_{!X_1,!X_2}} & \mathsf{S}(!X_1 \otimes !X_2) \\ {\scriptstyle \mathsf{m}^2_{\mathsf{S}X_1,\mathsf{S}X_2}}\downarrow & & & & \\ !\,(\mathsf{S}X_1 \,\&\, \mathsf{S}X_2) & & & & \downarrow{\scriptstyle \mathsf{S}\mathsf{m}^2_{X_1,X_2}} \\ {\scriptstyle !\widetilde{\psi}_{X_1,X_2}}\downarrow & & & & \\ !\mathsf{S}\,(X_1 \,\&\, X_2) & & \xrightarrow{\qquad\qquad\partial_{X_1\&X_2}\qquad\qquad} & & \mathsf{S}!\,(X_1 \,\&\, X_2) \end{array}$$

This condition means that the differential structure is compatible with the strong monoidal structure of the resource category $\mathcal{L}$. It becomes quite important when the SMC $\mathcal{L}$ is assumed to be closed since, in that situation, this strong monoidal structure turns $\mathcal{L}_!$ into a cartesian closed category.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Definition 5.6. *(∂-Schwarz)*

$$\begin{array}{ccccc} !\mathsf{S}^2X & \xrightarrow{\partial_{\mathsf{S}X}} & \mathsf{S}!\mathsf{S}X & \xrightarrow{\mathsf{S}\,\partial_X} & \mathsf{S}^2!X \\ {\scriptstyle !\,\mathsf{c}_X}\downarrow & & & & \downarrow{\scriptstyle \mathsf{c}_{!X}} \\ !\mathsf{S}^2X & \xrightarrow{\partial_{\mathsf{S}X}} & \mathsf{S}!\mathsf{S}X & \xrightarrow{\mathsf{S}\,\partial_X} & \mathsf{S}^2!X \end{array}$$

This condition means that the second derivative is a *symmetric* bilinear function: $\frac{\partial^2 f(x_1,x_2)}{\partial x_1 \partial x_2} \cdot (u_1, u_2) = \frac{\partial^2 f(x_1,x_2)}{\partial x_2 \partial x_1} \cdot (u_2, u_1)$.

The last condition was overlooked in Ehrhard (2023b), but the corresponding condition was already recognized as important in the theory of tangent categories Rosický (1984); Cockett and Cruttwell (2014).

Definition 5.7. *(∂-lift)*

$$\begin{array}{ccccc} !\mathsf{S}X & & \xrightarrow{\qquad\partial_X\qquad} & & \mathsf{S}!X \\ {\scriptstyle !\,\mathsf{l}_X}\downarrow & & & & \downarrow{\scriptstyle \mathsf{l}_{!X}} \\ !\mathsf{S}^2X & \xrightarrow{\partial_{\mathsf{S}X}} & \mathsf{S}!\mathsf{S}X & \xrightarrow{\mathsf{S}\,\partial_X} & \mathsf{S}^2!X \end{array}$$

Keeping in mind that l is the comultiplication of the bimonad S, one might expect the commutation corresponding to the counit σ of that comonad to commute, that is

$$\begin{array}{ccc} !\mathsf{S}X & \xrightarrow{\partial_X} & \mathsf{S}!X \\ & {\scriptstyle !\sigma_X}\searrow \quad \swarrow{\scriptstyle \sigma_{!X}} & \\ & X & \end{array}$$

but this would be too strong a requirement in the present setting as it would require intuitively that all morphisms $f \in \mathcal{L}_!(X, Y)$ satisfy $f(x_0 + x_1) = f(x_0) + f'(x_0) \cdot x_1$ (for all summable pair (x_0, x_1) of elements of X), that is, are affine. So, in coherent differentiation, this latter commutation is not required. In the infinitary setting of Ehrhard and Walch (2025), it expresses that morphisms are analytic in the sense that they coincide with their Taylor expansion, so this commutation will be an essential ingredient in the definition of a *coherent analytic category*.

6.1 The induced differentiation monad

The axiom 5.2 (∂-chain) exactly means that $\mathsf{S} : \mathcal{L} \to \mathcal{L}$ can be extended to a functor $\mathsf{T} : \mathcal{L}_! \to \mathcal{L}_!$. Remember that one speaks here of extension because we consider $\mathcal{L}$ as a "subcategory" of $\mathcal{L}_!$ (the inclusion being the functor $\mathsf{Der} : \mathcal{L} \to \mathcal{L}_!$). Concretely, T is defined on objects by $\mathsf{T}X = \mathsf{S}X$, and given $f \in \mathcal{L}_!(X, Y) = \mathcal{L}(!X, Y)$, one sets

$$\mathsf{T}f = (\mathsf{S}f)\ \partial_X\ .$$

Proposition 5.1. *The operation* T *is a functor* $\mathcal{L}_! \to \mathcal{L}_!$ *which extends* S *in the sense that for any* $f \in \mathcal{L}(X, Y)$*, one has* $\mathsf{T}\left(\mathsf{Der} f\right) = \mathsf{Der}\left(\mathsf{S}f\right)$.

This is completely standard in the theory of distributive laws.

We define $\zeta_X = \mathsf{Der}\,\iota_0 \in \mathcal{L}_!(X, \mathsf{T}X)$ and $\theta_X = \mathsf{Der}\,\tau_X \in \mathcal{L}_!(\mathsf{T}^2X, \mathsf{T}X)$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Proposition 5.2. *The morphisms $\zeta_X \in \mathcal{L}_!(X, \mathsf{T}X)$ and $\theta_X \in \mathcal{L}_!(\mathsf{T}^2X, \mathsf{T}X)$ are natural in X and turn the functor T into a monad on $\mathcal{L}_!$.*

Proof sketch. The only nontrivial properties are the naturality of ζ and θ. They result from 5.4 (∂-add). □

Remark 5.1. Intuitively,

$$\mathsf{T}f(x_0, x_1) = (f(x_0), f'(x_0) \cdot x_1)$$
$$\zeta_X(x) = (x, 0)$$
$$\theta_X((x_{00}, x_{01}), (x_{10}, x_{11})) = (x_{00}, x_{01} + x_{10})$$

and the naturality of ζ and θ means that $f'(x) \cdot 0 = 0$ and $f'(x_{00}) \cdot (x_{01} + x_{10}) = f'(x_{00}) \cdot x_{01} + f'(x_{00}) \cdot x_{10}$.

6.2 Partial derivatives

Given $f \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, X_n, Y)$, we have seen how to define the global differential $\mathsf{T}f \in \mathcal{L}_!(\mathsf{S}(X_1 \,\&\, \cdots \,\&\, X_n), \mathsf{S}Y)$ of f, that is (up to the iso stipulated by 4.14 (S-&)), $\mathsf{T}f \in \mathcal{L}_!(\mathsf{S}X_1 \,\&\, \cdots \,\&\, \mathsf{S}X_n, \mathsf{S}Y)$, which intuitively maps $(x_1, u_1), \ldots, (x_n, u_n)$ to $(f(x_1, \ldots, x_n), f'(x_1, \ldots, x_n) \cdot (u_1, \ldots, u_n))$. For any $i \in \{1, \ldots, n\}$, we also need to be able to define an ith *partial derivative* $\mathsf{T}_i f \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, \mathsf{S}X_i \,\&\, \cdots \,\&\, X_n, \mathsf{S}Y)$ which intuitively maps $(x_1, \ldots, (x_i, u), \ldots, x_n)$ to $(f(x_1, \ldots, x_n), f_i'(x_1, \ldots, x_n) \cdot u)$. We also expect these partial derivatives to satisfy

$$f'(x_1, \ldots, x_n) \cdot (u_1, \ldots, u_n) = \sum_{i=1}^{n} f_i'(x_1, \ldots, x_n) \cdot u_i = \sum_{i=1}^{n} \frac{\partial f(x_1, \ldots, x_n)}{\partial x_i} \cdot u_i \,.$$

The conditions introduced so far allow us to define such partial derivatives and prove their expected properties without further assumptions as we explain now.

We take $n = 2$ to simplify notations but the general case is not more complicated conceptually. Then it is possible to define $\psi^1_{X_1,X_2} \in \mathcal{L}(\mathsf{S}X_1 \,\&\, X_2, \mathsf{S}(X_1 \,\&\, X_2))$ as the following composition of morphisms

$$\mathsf{S}X_1 \,\&\, X_2 \xrightarrow{\mathsf{S}X_1 \& \iota_0} \mathsf{S}X_1 \,\&\, \mathsf{S}X_2 \xrightarrow{\widetilde{\psi}_{X_1,X_2}} \mathsf{S}(X_1 \,\&\, X_2)$$

and we use $\Psi^1_{X_1,X_2}$ for the associated morphism $\mathsf{Der}\psi^1_{X_1,X_2} \in \mathcal{L}_!(\mathsf{T}X_1 \,\&\, X_2, \mathsf{T}(X_1 \,\&\, X_2))$ in the Kleisli category $\mathcal{L}_!$. We define similarly $\Psi^2_{X_1,X_2} \in \mathcal{L}_!(X_1 \,\&\, \mathsf{T}X_2, \mathsf{T}(X_1 \,\&\, X_2))$. Intuitively $\Psi^1((x_1, u_1), x_2) = ((x_1, x_2), (u_1, 0))$ and $\Psi^2(x_1, (x_2, u_2)) = ((x_1, x_2), (0, u_2))$. It is also easily checked that Ψ^2 can be obtained from Ψ^1 using the symmetry isomorphism associated with &: $\Psi^2 = \mathsf{T}\langle \mathsf{pr}_2, \mathsf{pr}_1 \rangle \circ \Psi^1 \circ \langle \mathsf{pr}_2, \mathsf{pr}_1 \rangle$.

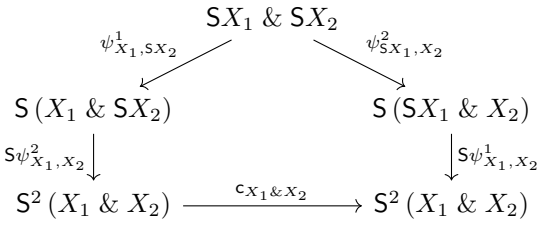
Theorem 5.1. *The morphisms $\Psi^1_{X_1,X_2}$ and $\Psi^2_{X_1,X_2}$ of $\mathcal{L}_!$ are natural in X_1 and X_2 and define a commutative strength on the monad T. More precisely, the following diagram commutes in $\mathcal{L}_!$*

$$\begin{array}{ccc}
 & \mathsf{T}X_1 \,\&\, \mathsf{T}X_2 & \\
{\scriptstyle \Psi^1_{X_1,\mathsf{T}X_2}}\swarrow & & \searrow{\scriptstyle \Psi^2_{\mathsf{T}X_1,X_2}} \\
\mathsf{T}(X_1 \,\&\, \mathsf{T}X_2) & & \mathsf{T}(\mathsf{T}X_1 \,\&\, X_2) \\
{\scriptstyle \mathsf{T}\Psi^2_{X_1,X_2}}\downarrow & & \downarrow{\scriptstyle \mathsf{T}\Psi^1_{X_1,X_2}} \\
\mathsf{T}^2(X_1 \,\&\, X_2) & \xrightarrow{\mathsf{c}_{X_1 \& X_2}} & \mathsf{T}^2(X_1 \,\&\, X_2)
\end{array} \tag{3}$$

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

and the induced monoidality $\widetilde{\Psi}_{X_1,X_2} = \theta_{X_1 \& X_2} \circ \mathsf{T}\Psi^2_{X_1,X_2} \circ \Psi^1_{\mathsf{T}X_1,X_2} = \theta_{X_1 \& X_2} \circ \mathsf{T}\Psi^1_{X_1,X_2} \circ \Psi^2_{X_1,\mathsf{T}X_2}$ *coincides with* $\mathsf{Der}\widetilde{\psi}_{X_1,X_2}$ *which is an iso in* $\mathcal{L}_!$.

Proof sketch. This is essentially trivial. For instance, Equation (3) is the image by Der of the diagram



whose commutation is easily proven using the joint monicity of $(\,\pi_i\ \pi_j\,)_{(i,j)\in\{0,1\}}$. The naturality of Ψ^1 and Ψ^2 boils down to the commutativity in $\mathcal{L}$ of

$$\begin{array}{ccc} !\mathsf{S}(X_1 \,\&\, X_2) & \xrightarrow{\partial_{X_1 \& X_2}} & \mathsf{S}!(X_1 \,\&\, X_2) \\ {\scriptstyle \langle !\mathsf{S}\mathsf{pr}_1, !\mathsf{S}\mathsf{pr}_2\rangle}\downarrow & & \downarrow{\scriptstyle \langle \mathsf{S}!\mathsf{pr}_1, \mathsf{S}!\mathsf{pr}_2\rangle} \\ !\mathsf{S}X_1 \,\&\, !\mathsf{S}X_2 & \xrightarrow{\partial_{X_1} \& \partial_{X_2}} & \mathsf{S}!X_1 \,\&\, \mathsf{S}!X_2 \end{array}$$

which results from the naturality of ∂_X in X and the joint monicity of pr_1 and pr_2. □

Definition 5.8. *The two partial derivatives of* $f \in \mathcal{L}_!(X_1 \,\&\, X_2, Y)$ *are*

$$\mathsf{T}_1 f = \mathsf{T}f \circ \Psi^1_{X_1,X_2} \in \mathcal{L}_!(\mathsf{T}X_1 \,\&\, X_2, \mathsf{T}Y)$$
$$\mathsf{T}_2 f = \mathsf{T}f \circ \Psi^2_{X_1,X_2} \in \mathcal{L}_!(X_1 \,\&\, \mathsf{T}X_2, \mathsf{T}Y)\,.$$

Proposition 5.3. *If* $f \in \mathcal{L}_!(X_1 \,\&\, X_2, Y)$, *we have*

$$\mathsf{T}f = \theta \circ \mathsf{T}_1\mathsf{T}_2 f = \theta \circ \mathsf{T}_2\mathsf{T}_1 f\,.$$

Proof. Apply Theorem 5.1. □

This means intuitively, as expected, that $f'(x_1,x_2)\cdot(u_1,u_2) = f'_1(x_1,x_2)\cdot u_1 + f'_2(x_1,x_2)\cdot u_2$.

Composing these morphisms Ψ^1 and Ψ^2, one can define $\Psi^i \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, \mathsf{T}X_i \,\&\, \cdots \,\&\, X_n, \mathsf{T}\,(X_1 \,\&\, \cdots \,\&\, X_n)\,)$ that one can also define directly as $\Psi^i = \mathsf{Der}\psi^i$ where $\psi^i = \widetilde{\psi}_{X_1,\ldots,X_n}$ $(\iota_0 \,\&\, \cdots \,\&\, \mathsf{S}X_i \,\&\, \cdots \,\&\, \iota_0\,)$ (see Definition 4.14 (S-&)), that is, intuitively,

$$\Psi^i(x_1,\ldots,(x_i,u),\ldots,x_n) = ((x_1,\ldots,x_n),(0,\ldots,u,\ldots,0))\,.$$

Given $f \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, X_n, Y)$, we can define the ith partial derivative of f as

$$\mathsf{T}_i f = \mathsf{T}f \circ \Psi^i \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, \mathsf{T}X_i \,\&\, \cdots \,\&\, X_n, \mathsf{T}Y)\,.$$

and given any repetition-free enumeration $(i_1,\ldots,i_n)$ of $\{1,\ldots,n\}$ we have

$$\mathsf{T}f = \theta^{n-1} \circ \mathsf{T}_{i_1} \cdots \mathsf{T}_{i_n} f\,.$$

6.3 The differential structure, in the representable case

Let $\mathcal{L}$ be a representably summable category (see Definition 4.16) and let $(\mathsf{S}, \pi_0, \pi_1, \sigma)$ be the associated summability structure.

Let $\widetilde{\partial} \in \mathcal{L}(\mathbb{D}, !\mathbb{D})$ be a morphism.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Definition 5.9. *($\widetilde{\partial}$-chain) We say that $\widetilde{\partial}$ satisfies ($\widetilde{\partial}$-chain) if it is a !-coalgebra structure on $\mathbb{D}$, that is, the two following diagrams commute*

$$\begin{array}{ccc} \mathbb{D} & \xrightarrow{\widetilde{\partial}} & !\mathbb{D} \\ & \searrow^{\mathsf{Id}_{\mathbb{D}}} & \downarrow{\scriptstyle \mathsf{der}_{\mathbb{D}}} \\ & & \mathbb{D} \end{array} \qquad \begin{array}{ccc} \mathbb{D} & \xrightarrow{\widetilde{\partial}} & !\mathbb{D} \\ {\scriptstyle \widetilde{\partial}}\downarrow & & \downarrow{\scriptstyle \mathsf{dig}_{\mathbb{D}}} \\ !\mathbb{D} & \xrightarrow{!\widetilde{\partial}} & !!\mathbb{D} \end{array}$$

Remember (from Melliès (2009), for instance) that the functor !_ inherits, from its strict symmetric monoidal structure $(\mathsf{m}^0, \mathsf{m}^2)$ from the SMC $(\mathcal{L}, \&)$ to the SMC $(\mathcal{L}, \otimes)$, a lax symmetric monoidal structure (μ^0, μ^2) from the SMC $(\mathcal{L}, \otimes)$ to itself. This means that $\mu^0 \in \mathcal{L}(1, !1)$ and $\mu^2_{X,Y} \in \mathcal{L}(!X \otimes !Y, !(X \otimes Y))$ satisfy some coherence diagrams that we do not record here. These morphisms are defined as follows:

$$1 \xrightarrow{\mathsf{m}^0} !\top \xrightarrow{\mathsf{dig}_\top} !!\top \xrightarrow{!(\mathsf{m}^0)^{-1}} !1$$

$$!X \otimes !Y \xrightarrow{\mathsf{m}^2_{X,Y}} !(X \,\&\, Y) \xrightarrow{\mathsf{dig}_{X\&Y}} !!(X \,\&\, Y) \xrightarrow{!(\mathsf{m}^2_{X,Y})^{-1}} !(!X \otimes !Y) \xrightarrow{!(\mathsf{der}_X \otimes \mathsf{der}_Y)} !(X \otimes Y)$$

This structure is quite important in particular when considering the Eilenberg-Moore category $\mathcal{L}^!$ of the comonad !_. Remember that an object of that category is a pair $P = (\underline{P}, h_P)$ where $\underline{P}$ is an object of $\mathcal{L}$ and $h_P \in \mathcal{L}(\underline{P}, !\underline{P})$ satisfies

$$\begin{array}{ccc} \underline{P} & \xrightarrow{h_P} & !\underline{P} \\ & \searrow^{\mathsf{Id}_{\underline{P}}} & \downarrow{\scriptstyle \mathsf{der}_{\underline{P}}} \\ & & \underline{P} \end{array} \qquad \begin{array}{ccc} \underline{P} & \xrightarrow{h_P} & !\underline{P} \\ {\scriptstyle h_P}\downarrow & & \downarrow{\scriptstyle \mathsf{dig}_{\underline{P}}} \\ !\underline{P} & \xrightarrow{!h_P} & !!\underline{P} \end{array}$$

and a morphism $P \to Q$ in $\mathcal{L}^!$ is an $f \in \mathcal{L}(\underline{P}, \underline{Q})$ such that

$$\begin{array}{ccc} \underline{P} & \xrightarrow{f} & \underline{Q} \\ {\scriptstyle h_P}\downarrow & & \downarrow{\scriptstyle h_Q} \\ !\underline{P} & \xrightarrow{!f} & !\underline{Q} \end{array}$$

Equipped with μ^0, the object 1 is a !-coalgebra that we simply denote as 1 and, given two !-coalgebras P and Q, the pair $(\underline{P} \otimes \underline{Q}, h)$ where h is defined as the following composition of morphisms

$$\underline{P} \otimes \underline{Q} \xrightarrow{h_P \otimes h_Q} !\underline{P} \otimes !\underline{Q} \xrightarrow{\mu^2_{\underline{P},\underline{Q}}} !(\underline{P} \otimes \underline{Q})$$

is a !-coalgebra that we denote as $P \otimes Q$.

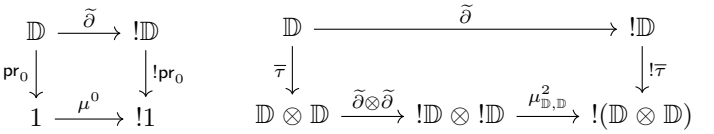
Definition 5.10. *($\widetilde{\partial}$-local) We say that $\widetilde{\partial}$ satisfies ($\widetilde{\partial}$-local) if the following diagram commutes*

$$\begin{array}{ccc} 1 & \xrightarrow{\mu^0} & !1 \\ {\scriptstyle \overline{\pi}^{\&}_0}\downarrow & & \downarrow{\scriptstyle !\overline{\pi}^{\&}_0} \\ \mathbb{D} & \xrightarrow{\widetilde{\partial}} & !\mathbb{D} \end{array}$$

In other words, $\overline{\pi}^{\&}_0$ is a !-coalgebra morphism from 1 *to* $(\mathbb{D}, \widetilde{\partial})$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Definition 5.11. *($\widetilde{\partial}$-add) We say that $\widetilde{\partial}$ satisfies ($\widetilde{\partial}$-add) if the following diagrams commute*



In other words, the unit and the comultiplication of the bimonoid $\mathbb{D}$ are !-coalgebra morphisms from $(\mathbb{D}, \widetilde{\partial})$ to 1 and to $(\mathbb{D}, \widetilde{\partial}) \otimes (\mathbb{D}, \widetilde{\partial})$, respectively.

In Ehrhard (2023b), we proved the following result.

Theorem 5.2. *There is a bijective correspondence between the natural transformations $\partial_X \in \mathcal{L}(!\mathsf{S}X, \mathsf{S}!X)$ which are differential structures on $\mathcal{L}$ and the morphisms $\widetilde{\partial} \in \mathcal{L}(\mathbb{D}, !\mathbb{D})$ which satisfy ($\widetilde{\partial}$-chain), ($\widetilde{\partial}$-local) and ($\widetilde{\partial}$-add).*

Let us just explain how the distributive law ∂ is defined when $\widetilde{\partial}$ is given: first we define a morphism ∂'_X as the following composition of morphisms

$$!(\mathbb{D} \multimap X) \otimes \mathbb{D} \xrightarrow{!(\mathbb{D}\multimap X)\otimes\widetilde{\partial}} !(\mathbb{D} \multimap X) \otimes !\mathbb{D} \xrightarrow{\mu^2} !((\mathbb{D} \multimap X) \otimes \mathbb{D}) \xrightarrow{!\mathsf{ev}} !X \tag{4}$$

and then we set $\partial_X = \mathsf{cur}\,(\,\partial'_X\,) \in \mathcal{L}(!\mathsf{S}X, \mathsf{S}!X)$.

Definition 5.12. *A differential representably summable resource category is a representably summable resource category $\mathcal{L}$ equipped with a morphism $\widetilde{\partial} \in \mathcal{L}(\mathbb{D}, !\mathbb{D})$ which satisfies ($\widetilde{\partial}$-chain), ($\widetilde{\partial}$-local) and ($\widetilde{\partial}$-add).*

The whole point of these definitions is the observation in Ehrhard (2023b) that such a differential structure on $\mathbb{D}$ is quite easy to obtain. Remember in particular that a SMC $\mathcal{L}$ is *Lafont* if !_ is the comonad associated with an adjunction between $\mathcal{L}$ and the category $\mathcal{L}^{\otimes}$ of commutative comonoids on $\mathcal{L}$. More precisely, this means that the obvious forgetful functor $\mathcal{L}^{\otimes} \to \mathcal{L}$ has a right adjoint, and !_ is the comonad on $\mathcal{L}$ induced by this adjunction.[8] It turns out that many interesting and non additive models of LL are Lafont resource categories, here are a few examples but there are many others:

- the category of coherence spaces with the multiset based exponential;
- the category of hypercoherence spaces with the multiset based exponential;[9]
- the category of nonuniform coherence spaces equipped with the Boudes exponential Bucciarelli and Ehrhard (2001); Boudes (2011);
- the category of PCSs with its unique known exponential Danos and Ehrhard (2011); Crubillé (2017).

Theorem 5.3. *If $\mathcal{L}$ is a Lafont resource category which is representably summable, then $\mathbb{D}$ has exactly one differential coalgebra structure.*

Sketch of the proof. We know that $(\mathbb{D}, \mathsf{pr}_0, \overline{\tau})$ is a commutative comonoid, see Proposition 4.1. This structure induces the announced !-coalgebra structure on $\mathbb{D}$ through the Lafont property of $\mathcal{L}$. □

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

► **Example. 5.1.** *We continue Example 4.1. The category* **Coh** *is a resource category. Its tensor product is defined by* $|E_1 \otimes E_2| = |E_1| \times |E_2|$ *and coherence given by* $(a_1, a_2) \frown_{E\otimes F} (a_1', a_2')$ *if* $(a_i \frown_{E_i} a_i')_{i=1,2}$, *which is easily seen to be a functor: given* $(s_i \in \mathbf{Coh}(E_i, F_i))_{i=1,2}$, *the set* $s_1 \otimes s_2 = \{((a_1, a_2), (b_1, b_2)) \mid ((a_i, b_i) \in s_i)_{i=1,2}\}$ *is an element of* $\mathbf{Coh}(E_1 \otimes E_2, F_1 \otimes F_2)$. *The associated unit is* $1 = (\{*\}, =)$.

The resource modality originally introduced by Girard, that we considered in Section 3.1 and that fails to provide an adequate differential setting, is defined by taking for $!E$ *the set of finite cliques of* E, *with* $x_1 \frown_{!E} x_2$ *if* $x_1 \cup x_2 \in \mathrm{Cl}(E)$. *As is well-known* **Coh** *has also a free exponential whose definition is quite similar to the original one: take for* $|!E|$ *the set of all* $m \in \mathcal{M}_{\mathrm{fin}}(|E|)$ *such that* $\mathsf{supp}\,(m) \in \mathrm{Cl}(E)$, *and coherence given by* $m_1 \frown_{!E} m_2$ *if* $m_1 + m_2 \in |!E|$.

Notice that $|!\mathbb{D}|$ *is the set of all finite multisets of elements of* $\{0, 1\}$ *since* $0 \frown_{\mathbb{D}} 1$, *and* $m \frown_{!\mathbb{D}} m'$ *holds for all* $m, m' \in |!\mathbb{D}|$. *The differential structure induced by the fact that this exponential is free is* $\widetilde{\partial} \in \mathbf{Coh}(\mathbb{D}, !\mathbb{D})$; *it is given by*

$$\widetilde{\partial} = \{(i, [i_1, \dots, i_k]) \mid k \in \mathbb{N},\ i, i_1, \dots, i_k \in \{0, 1\} \text{ and } i = i_1 + \cdots + i_k\}\,.$$

In other words $(i, m) \in \widetilde{\partial}$ *if* $i = 0$ *and all the elements of* m *are* 0, *or* $i = 1$ *and all the elements of* m *are* 0 *but exactly one, which is* 1, *that is*

$$\widetilde{\partial} = \{(0, k[0]) \mid k \in \mathbb{N}\} \cup \{(1, [1] + k[0] \mid k \in \mathbb{N})\}\,.$$

Since $0 \frown_{\mathbb{D}} 1$ *we have* $m \frown_{!\mathbb{D}} m'$ *for all* $m, m' \in |!\mathbb{D}|$. *Moreover if* $(0, m), (1, m') \in \widetilde{\partial}$, *we have* $m \neq m'$ *as required, since* $1 \in \mathsf{supp}\,(m') \setminus \mathsf{supp}\,(m)$. *Let us check for instance that the second diagram of* ($\widetilde{\partial}$*-add) commutes*[10]. *So let* $(i, p) \in |\mathbb{D}| \times |!(\mathbb{D} \otimes \mathbb{D})|$ *and let us write* $p = [(l_1, r_1), \dots, (l_k, r_k))]$. *Saying that* $(i, p) \in !\overline{\tau}\ \widetilde{\partial}$ *means that there is* $m \in |!\mathbb{D}|$ *such that* $(i, m) \in \widetilde{\partial}$ *and* $(m, p) \in !\overline{\tau}$. *That is,* $m = [i_1, \dots, i_k]$ *with* $i = i_1 + \cdots + i_k$, $(i_j = l_j + r_j)_{j=1}^k$. *To summarize,* $(i, p) \in !\overline{\tau}\ \widetilde{\partial}$ *holds iff* $i = l_1 + \cdots + l_k + r_1 + \cdots + r_k$. *Saying that* $(i, p) \in \mu^2\ (\widetilde{\partial} \otimes \widetilde{\partial})\ \overline{\tau}$ *means that there are* $l, r \in \{0, 1\}$ *such that* $l + r = i$, *and there are* $m, q \in |!\mathbb{D}|$ *such that* $(l, m), (r, q) \in \widetilde{\partial}$ *and* $((m, q), p) \in \mu^2$. *This latter condition means that* $m = [l_1, \dots, l_k]$, $q = [r_1, \dots, r_k]$ *(remember that* $p = [(l_1, r_1), \dots, (l_k, r_k)]$ *). This shows that* $(i, p) \in \mu^2\ (\widetilde{\partial} \otimes \widetilde{\partial})\ \overline{\tau}$ *iff* $i = l + r$ *where* $l = l_1 + \cdots + l_k$ *and* $r = r_1 + \cdots + r_k$, *that is, iff* $(i, p) \in !\overline{\tau}\ \widetilde{\partial}$.

It is also interesting to describe the associated distributive law $\partial_E \in \mathbf{Coh}(!\mathsf{S}E, \mathsf{S}!E)$. *The composition of morphisms described in Equation (4) gives us* $\partial'_E \in \mathbf{Coh}(!(\mathbb{D} \multimap)\ E \otimes \mathbb{D}, !E)$:

$$\begin{aligned}\partial'_E = \{&(([(i_1, a_1), \dots, (i_k, a_k)], i), [a_1, \dots, a_k]) \\ &\mid k \in \mathbb{N},\ i, i_1, \dots, i_k \in \{0, 1\},\ i = i_1 + \cdots + i_k \text{ and } \{(i_1, a_1), \dots, (i_k, a_k)\} \in \mathrm{Cl}(\mathbb{D} \multimap E)\}\end{aligned}$$

and hence

$$\begin{aligned}\partial_E = \{&([(i_1, a_1), \dots, (i_k, a_k)], (i, [a_1, \dots, a_k])) \\ &\mid k \in \mathbb{N},\ i, i_1, \dots, i_k \in \{0, 1\},\ i = i_1 + \cdots + i_k \text{ and } \{(i_1, a_1), \dots, (i_k, a_k)\} \in \mathrm{Cl}(\mathbb{D} \multimap E)\}\,.\end{aligned}$$

This means that the differential $\mathsf{T}t = (\mathsf{S}t)\ \partial_E \in \mathbf{Coh}_!(\mathsf{S}E, \mathsf{S}F)$ *of* $t \in \mathbf{Coh}_!(E, F)$ *is given by*

$$\begin{aligned}\mathsf{T}t = \{&([(i_1, a_1), \dots, (i_k, a_k)], (i, b)) \in |!\mathsf{S}E| \times |\mathsf{S}F| \\ &\mid ([a_1, \dots, a_k], b) \in t,\ i, i_1, \dots, i_k \in \{0, 1\},\ i = i_1 + \cdots + i_k \\ &\qquad \text{and } \{(i_1, a_1), \dots, (i_k, a_k)\} \in \mathrm{Cl}(\mathbb{D} \multimap E)\} \\ = \{&([(0, a_1), \dots, (0, a_k)], (0, b)) \mid ([a_1, \dots, a_k], b) \in t\} \\ &\cup \{([(0, a_1), \dots, (0, a_k), (1, a)], (1, b)) \mid ([a_1, \dots, a_k, a], b) \in t \text{ and } (a \frown_E a_j)_{j=1}^k\}\end{aligned}$$

where the condition $(a \frown_E a_j)_{j=1}^k$ *comes from the fact that we must have* $\{(0, a_1), \dots, (0, a_k), (1, a)\} \in \mathrm{Cl}(\mathbb{D} \multimap E)$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

We can define a stable[11] *function* $\hat{t} : \mathrm{Cl}(E) \to \mathrm{Cl}(F)$ *by* $\hat{t}(x) = \{b \in |Y| \mid \exists m \in |!E| \ \ \mathsf{supp}\,(m) \subseteq x \text{ and } (m, b) \in t\}$. *Then, under the identification*

$$\mathrm{Cl}(\mathsf{S}E) = \{(x, u) \in \mathrm{Cl}(E)^2 \mid x \cup u \in \mathrm{Cl}(E) \text{ and } x \cap u = \emptyset\}$$

we have

$$\widehat{\mathsf{T}t}(x, u) = (\hat{t}(x), \bigcup_{a \in u} \hat{t}(x \cup \{a\}) \setminus \hat{t}(x))$$

and we recover the initial intuition of the derivative of a stable function, see Section 3.1. This differentiation $t \mapsto \mathsf{T}t$ *is functorial, in other words, the chain rule holds, because we have moved to the free exponential, whose web uses finite multicliques instead of finite cliques.*

The great benefit of this systematic approach based on the representable differential structure of **Coh** *is that now* $\mathsf{T}t$ *is a morphism* $\mathsf{S}E \to \mathsf{S}F$ *in* $\mathbf{Coh}_!$ *and hence induces a stable function* $\mathrm{Cl}(\mathsf{S}E) \to \mathrm{Cl}(\mathsf{S}F)$*: this is a way of saying that the differential depends stably from the point where it is computed.*

▶ **Example. 5.2.** *We continue Example 4.2. The category* **Pcoh** *is also a resource category. We have seen in Example 4.2 that this category is a cartesian SMCC which is actually* $*$*-autonomous for the dualizing object* $\bot = 1$. *Then we define* $!X$ *by* $|!X| = \mathcal{M}_{\mathrm{fin}}(|X|)$ *and* $\mathsf{P}(!X) = \{x^! \mid x \in \mathsf{P}X\}^{\bot\bot}$ *where* $x^! \in (\mathbb{R}_{\geq 0})^{|!X|}$ *is defined by* $x^!_m = x^m = \prod_{a \in |X|} x_a^{m(a)}$ *(see* Remark 4.9*). Given* $t \in \mathbf{Pcoh}(X, Y)$*, it is easy to check that there is exactly one* $!t \in (\mathbb{R}_{\geq 0})^{|!X| \times |!Y|}$ *such that*

$$\forall x \in \mathsf{P}X \quad !t \cdot x^! = (t \cdot x)^! \,. \tag{5}$$

Explicitly, a simple computation using Equation (5) *shows that*

$$\forall (m, p) \in |!X| \times |!Y| \quad (!t)_{m,p} = \sum_{r \in \mathsf{L}(m,p)} \begin{bmatrix} p \\ r \end{bmatrix} t^r$$

where $\mathsf{L}(m, p)$ *is the set of all* $r \in \mathcal{M}_{\mathrm{fin}}(|X| \times |Y|)$ *such that* $\forall a \in |X| \ \ \sum_{b \in |Y|} r(a, b) = m(a)$ *and* $\forall b \in |Y| \ \ \sum_{a \in |X|} r(a, b) = p(b)$*, and* $\left[\begin{smallmatrix} p \\ r \end{smallmatrix}\right] = \prod_{b \in |Y|} \frac{p(b)!}{\prod_{a \in |X|} r(a,b)!} \in \mathbb{N}$ *is a multinomial coefficient. It can be proven that if* $s, t \in \mathbf{Pcoh}_!(!X, Y)$ *satisfy* $\forall x \in \mathsf{P}X \ s \cdot x^! = t \cdot x^!$*, then* $s = t$ *(as matrices). So the function* $\hat{t} : \mathsf{P}X \to \mathsf{P}Y$ *defined by* $\hat{t}(x) = t \cdot x^!$ *fully determines* t*; such a function* $\mathsf{P}X \to \mathsf{P}Y$ *will be called an analytic function since indeed we have*

$$\hat{t}(x) = \Big(\sum_{m \in |!X|} t_{m,b} x^m \Big)_{b \in |Y|}$$

meaning that $\hat{t}$ *is defined as a (generalized) power series with nonnegative coefficients.*

This functor is a comonad with counit $\mathsf{der}_X \in \mathbf{Pcoh}(!X, X)$ *characterized by* $\widehat{\mathsf{der}_X}(x) = x$ *and* $\mathsf{dig}_X \in \mathbf{Pcoh}(!X, !!X)$ *by* $\widehat{\mathsf{dig}_X}(x) = x^{!!}$*, that is, as matrices,* $(\,\mathsf{der}_X\,)_{m,a} = \boldsymbol{\delta}_{m,[a]}$ *(for* $(m, a) \in |!X| \times |X|$*) and* $(\,\mathsf{dig}_X\,)_{m,M} = \boldsymbol{\delta}_{m,\sum M}$*. It is easily checked to be strong monoidal from* $(\mathbf{Pcoh}, \top, \&)$ *to* $(\mathbf{Pcoh}, 1, \otimes)$*. For instance the Seely isomorphism* $\mathsf{m}^2_{X,Y} \in \mathbf{Pcoh}(!X \otimes !Y, !(X \,\&\, Y)$ *is fully characterized by* $\mathsf{m}^2_{X,Y} \cdot (x^! \otimes y^!) = (x, y)^!$ *(identifying* $\mathsf{P}(X \,\&\, Y)$ *with* $\mathsf{P}X \times \mathsf{P}Y$*).*

It was proved in Crubillé (2017) that this exponential is the free one, that is, the SMC **Pcoh** *is a Lafont category. So by* Theorem 5.3 *the object* $\mathbb{D} = 1 \,\&\, 1$ *has a structure of* $!$*-coalgebra* $\widetilde{\partial} \in \mathbf{Pcoh}(\mathbb{D}, !\mathbb{D})$*, which is characterized by*

$$\widetilde{\partial}_{i,[i_1,\ldots,i_k]} = \boldsymbol{\delta}_{i,i_1+\cdots+i_k}$$

for $i, i_1, \ldots, i_k \in \{0, 1\}$*. This structure turns* **Pcoh** *into a representable coherent differential category. So we have an induced distributive law* $\partial_X \in \mathbf{Pcoh}(!\mathsf{S}X, \mathsf{S}!X)$ *and an easy computation shows that*

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

(using the notation $i \cdot m$ introduced in Section 2)

$$(\partial_X)_{p,(i,m)} = \begin{cases} 1 & \text{if } p = 0 \cdot m \\ m(a) & \text{if } m = m_0 + [a] \text{ and } p = 0 \cdot m_0 + [(1,a)] \\ 0 & \text{otherwise.} \end{cases}$$

so that the extension T *of* S *to* $\mathbf{Pcoh}_!$ *acts as follows on morphisms. Let* $t \in \mathbf{Pcoh}(!X, Y)$ *then* $\mathsf{T}t \in \mathbf{Pcoh}(!\mathsf{S}X, \mathsf{S}Y)$ *is given by*

$$(\mathsf{T}t)_{p,(i,b)} = \begin{cases} t_{m,b} & \text{if } i = 0 \text{ and } p = 0 \cdot m \\ m(a)t_{m,b} & \text{if } i = 1, p = 0 \cdot m_0 + [(1,a)] \text{ and } m = m_0 + [a] \\ 0 & \text{otherwise.} \end{cases}$$

This means that, given an element of $\mathsf{P}(\mathsf{S}X)$ *that we consider as a pair* $(x,u) \in \mathsf{P}X^2$ *such that* $x + u \in \mathsf{P}X$*, we have*

$$\widehat{\mathsf{T}}t(x,u) = (\hat{t}(x), \hat{t}'(x) \cdot u)$$

where

$$\hat{t}'(x) \cdot u = \left(\sum_{\substack{m \in \mathcal{M}_{\mathrm{fin}}(|X|) \\ a \in |X|}} (m(a)+1) t_{m+[a],b} x^m u_a \right)_{b \in |Y|} = \left(\lim_{\varepsilon \to 0^+} \frac{\hat{t}(x+\varepsilon u)_b - \hat{t}(x)_b}{\varepsilon} \right)_{b \in |Y|}$$

is the differential of $\hat{t}$ *computed at* x *in the direction* u*.*

7. The Closed Case

So far we have not considered function space constructions, but the reader acquainted with the denotational semantics of LL probably knows that both SMCs **Coh** and **Pcoh** are symmetric monoidal closed categories and that the associated Kleisli categories $\mathbf{Coh}_!$ and $\mathbf{Pcoh}_!$ are cartesian closed.

More generally, when $\mathcal{L}$ is a resource category that is closed (as an SMC), we know that $\mathcal{L}_!$ is a CCC.[12] We use the notation $X \Rightarrow Y$ for the object of morphisms from X to Y in $\mathcal{L}_!$, which is $!X \multimap Y$ and whose associated evaluation morphism $\mathsf{Ev} \in \mathcal{L}_!((X \Rightarrow Y) \,\&\, X, Y)$ is defined as the following composition of morphisms in $\mathcal{L}$.

$$!((X \Rightarrow Y) \,\&\, X) \xrightarrow{(\mathsf{m}^2)^{-1}} !(!X \multimap Y) \otimes !X \xrightarrow{\mathsf{der}_{!X \multimap Y}} (!X \multimap Y) \otimes !X \xrightarrow{\mathsf{ev}} Y$$

Given $f \in \mathcal{L}_!(Z \,\&\, X, Y) = \mathcal{L}(!(Z \,\&\, X), Y)$, we have $f\,\mathsf{m}^2 \in \mathcal{L}(!Z \otimes !X, Y)$ and hence $\mathsf{cur}\left(f\,\mathsf{m}^2\right) \in \mathcal{L}_!(Z, X \Rightarrow Y)$, and this morphism $\mathsf{Cur}f = \mathsf{cur}\left(f\,\mathsf{m}^2\right)$ is uniquely characterized by the equation

$$\mathsf{Ev} \circ \left(\mathsf{Cur}f \,\&\, X\right) = f\,.$$

The main ingredient in the interpretation of a coherent differential λ-calculus in the CCC associated with a coherent differential resource category, which is closed (as an SMC) will be a morphism allowing us to internalize the action of the T functor as a morphism $\mathsf{T}^{\mathsf{int}} \in \mathcal{L}_!(X \Rightarrow Y, \mathsf{T}X \Rightarrow \mathsf{T}Y)$. This is not a surprise since one of the main features of a strong monad is precisely to make such internalizations possible. This morphism is defined by $\mathsf{T}^{\mathsf{int}} = \mathsf{Cur}\left(\mathsf{T}\,\mathsf{Ev} \circ \Psi^2_{X \Rightarrow Y, X}\right)$ where $\mathsf{T}\,\mathsf{Ev} \circ \Psi^2_{X \Rightarrow Y, X}$ is typed as follows in $\mathcal{L}_!$:

$$(X \Rightarrow Y) \,\&\, \mathsf{T}X \xrightarrow{\Psi^2_{X \Rightarrow Y, X}} \mathsf{T}((X \Rightarrow Y) \,\&\, X) \xrightarrow{\mathsf{T}\,\mathsf{Ev}} \mathsf{T}Y\,.$$

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Let $f \in \mathcal{L}_!(Z \,\&\, X, Y)$. There are two morphisms in $\mathcal{L}_!(Z, \mathsf{T}X \Rightarrow \mathsf{T}Y)$ that we can naturally define using f, namely

$$\mathsf{T}^{\mathsf{int}} \circ \mathsf{Cur} f \quad \text{and} \quad \mathsf{Cur}\left(\mathsf{T}_2 f\right) = \mathsf{Cur}\left(\mathsf{T}f \circ \Psi^2_{Z,X}\right)$$

Proposition 6.1. *For any $f \in \mathcal{L}_!(Z \,\&\, X, Y)$ we have $\mathsf{T}^{\mathsf{int}} \circ \mathsf{Cur} f = \mathsf{Cur}\left(\mathsf{T}f \circ \Psi^2_{Z,X}\right)$.*

The proof is easy, and the meaning of this statement is that $\mathcal{L}$ validates a form of "differential β-reduction".

Consider now $f \in \mathcal{L}_!(Z \,\&\, X, Y \Rightarrow U)$ and $g \in \mathcal{L}_!(Z \,\&\, X, Y)$; the morphism f should be seen as a function $Y \to U$ depending on two parameters in Z and X and g as an argument for that function, similarly parameterized. We can apply f to g, defining $(f)g = \mathsf{Ev} \circ \langle f, g\rangle \in \mathcal{L}_!(Z \,\&\, X, U)$ and then we can take the derivative of $(f)g$ wrt. the second parameter which is

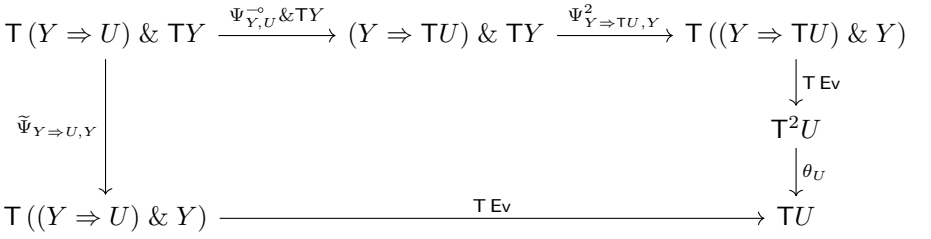
$$\mathsf{T}_2(f)g = \mathsf{T}(f)g \circ \Psi^2 \in \mathcal{L}_!(Z \,\&\, \mathsf{T}X, \mathsf{T}U)\,.$$

On the other hand, we have $\mathsf{T}_2 g \in \mathcal{L}_!(Z \,\&\, \mathsf{T}X, \mathsf{T}Y)$ and $\mathsf{T}_2 f \in \mathcal{L}_!(Z \,\&\, \mathsf{T}X), \mathsf{T}\,(Y \Rightarrow U))$ so that $\Psi^{\multimap}_{Y,U} \circ \mathsf{T}_2 f \in \mathcal{L}_!(Z \,\&\, \mathsf{T}X, Y \Rightarrow \mathsf{T}U)$ where $\Psi^{\multimap}_{U,V} = \mathsf{Der}\varphi^{\multimap}_{!Y,U} \in \mathcal{L}_!(\mathsf{T}\,(Y \Rightarrow U)\,, Y \Rightarrow \mathsf{T}U)$ is an iso by 4.13 (S-—∘). Therefore, $\mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,U} \circ \mathsf{T}_2 f \in \mathcal{L}_!(Z \,\&\, \mathsf{T}X, \mathsf{T}Y \Rightarrow \mathsf{T}^2 U)$ and hence $(\mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,U} \circ \mathsf{T}_2 f)\mathsf{T}_2 g \in \mathcal{L}_!(Z \,\&\, \mathsf{T}X, \mathsf{T}^2 U)$. Remember that $\theta_U \in \mathcal{L}_!(\mathsf{T}^2 U, \mathsf{T}U)$ is the multiplication of the monad T on $\mathcal{L}_!$.

Theorem 6.1. *If $f \in \mathcal{L}_!(Z \,\&\, X, Y \Rightarrow U)$ and $g \in \mathcal{L}_!(Z \,\&\, X, Y)$, then we have*

$$\mathsf{T}_2(f)g = \theta_U \circ (\mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,U} \circ \mathsf{T}_2 f)\mathsf{T}_2 g = ((\mathsf{T}Y \Rightarrow \theta_U) \circ \mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,U} \circ \mathsf{T}_2 f)\mathsf{T}_2 g\,.$$

Proof sketch. It suffices to establish the following commutation



The proof is not straightforward and uses crucially 5.5 (∂-&). It can be found in Ehrhard (2023a). □

This theorem shows how to compute the derivative of an application wrt. one of its parameters.

8. Fixpoints

We start with stating a few standard results about fixpoints in a CCC which is enriched in ω-cpos.

8.1 Reminder about fixpoint operators in a CCC

Definition 7.1. *An ω-cpo is a poset that has a least element 0 and where any monotone sequence has a lub. If D and D' are ω-cpos, a function $f : D \to D'$ is Scott-continuous (or simply continuous) if f is monotone and commutes with the lubs of monotone sequences.*

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Remark 7.1. In the literature, Scott continuity is usually defined as preservation of the lubs of arbitrary directed sets. However, such lubs do not always exist in the situations we are interested in. This is specifically the case in categories arising in continuous probabilistic settings: in such situations, lubs of monotone countable sequences can be assumed to exist thanks to the monotone convergence theorem though more general directed lubs do not always exist. We nevertheless use the term "Scott continuity" since the fundamental ideas of Dana Scott are also central in such models.

Proposition 7.1. *If D is an ω-cpo and $f : D \to D$ is continuous, then f has a least fixpoint, which is* $\sup_{n\in\mathbb{N}} f^n(0)$.

Definition 7.2. *A λ-category is a CCC, which is enriched in ω-cpos and in which the pairing operation $\mathcal{C}(X, Y_1) \times \mathcal{C}(X, Y_2) \to \mathcal{C}(X, Y_1 \,\&\, Y_2)$ and the currying operation $\mathcal{C}(Z \,\&\, X, Y) \to \mathcal{C}(Z, X \Rightarrow Y)$ are Scott continuous.*

Proposition 7.2. *Let $\mathcal{C}$ be a λ-category and Y be an object of $\mathcal{C}$. Then any morphism $h \in \mathcal{C}(Y, Y)$ has a least fixpoint, that is, there is a morphism $y \in \mathcal{C}(\top, Y)$ such that $h \circ y = y$ and y is minimal with this property in the poset $\mathcal{C}(\top, X)$.*

Proof. One defines a sequence $(y_n \in \mathcal{C}(\top, Y))_{n\in\mathbb{N}}$ by setting $y_0 = 0$ and $y_{n+1} = f \circ y_n$ and using the fact that composition is monotone, one checks that this sequence is monotone. Its lub y satisfies the required property by Scott continuity of composition. □

Theorem 7.1. *Let $\mathcal{C}$ be a λ-category. For any objects X of $\mathcal{C}$, there is a morphism $\mathcal{Y} \in \mathcal{C}(X \Rightarrow X, X)$ such that, for any morphism $f \in \mathcal{C}(Z, X \Rightarrow X)$, the morphism $\mathcal{Y} \circ f \in \mathcal{C}(Z, X)$ is the least morphism $g \in \mathcal{C}(Z, X)$ such that* $\mathsf{Ev} \circ \langle f, g\rangle = g$.

This is a standard result in semantics; we give the proof because we think it helps understanding Section 8.2.

Proof. Apply Proposition 7.2 with $Y = ((X \Rightarrow X) \Rightarrow X)$ and $h = \mathsf{Cur}\,(H) \in \mathcal{C}(Y, Y)$ where H is the following composition of morphisms in $\mathcal{C}$:

$$
\begin{array}{c}
Y \,\&\, (X \Rightarrow X) \\
\downarrow{\scriptstyle Y\&\langle \mathsf{Id},\mathsf{Id}\rangle} \\
Y \,\&\, (X \Rightarrow X) \,\&\, (X \Rightarrow X) \\
\downarrow{\scriptstyle \langle \mathsf{pr}_3,\mathsf{pr}_1,\mathsf{pr}_2\rangle} \\
(X \Rightarrow X) \,\&\, Y \,\&\, (X \Rightarrow X) \\
\downarrow{\scriptstyle (X\Rightarrow X)\&\mathsf{Ev}} \\
(X \Rightarrow X) \,\&\, X \\
\downarrow{\scriptstyle \mathsf{Ev}} \\
X
\end{array}
$$

which gives us $\mathcal{Y}' \in \mathcal{C}(\top, (X \Rightarrow X) \Rightarrow X)$ which is the least morphism such that $h \circ \mathcal{Y}' = \mathcal{Y}'$. Then $\mathcal{Y}$ is the following composition of morphisms in $\mathcal{C}$:

$$
X \Rightarrow X \xrightarrow{\langle 0, X\Rightarrow X\rangle} \top \,\&\, (X \Rightarrow X) \xrightarrow{\mathcal{Y}'\&(X\Rightarrow X)} ((X \Rightarrow X) \Rightarrow X) \,\&\, (X \Rightarrow X) \xrightarrow{\mathsf{Ev}} X
$$

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

By monotonicity and Scott continuity of all the CCC operations in $\mathcal{C}$, it follows that $\mathcal{Y} = \sup_{n\in\mathbb{N}} \mathcal{Y}_n$ where $(\mathcal{Y}_n \in \mathcal{C}(X \Rightarrow X, X))_{n\in\mathbb{N}}$ is the (obviously monotone) sequence of morphisms inductively defined by $\mathcal{Y}_0 = 0$ and $\mathcal{Y}_{n+1} = \mathsf{Ev} \circ \langle X \Rightarrow X, \mathcal{Y}_n\rangle$.

Therefore

$$\begin{aligned}\mathcal{Y} \circ f &= \sup_{n\in\mathbb{N}} (\mathcal{Y}_n \circ f)\\ &= \sup_{n\in\mathbb{N}} (\mathcal{Y}_{n+1} \circ f) \quad \text{since } \mathcal{Y}_0 = 0\\ &= \sup_{n\in\mathbb{N}} (\,\mathsf{Ev} \circ \langle f, \mathcal{Y}_n \circ f\rangle)\\ &= \sup_{n\in\mathbb{N}} g_n\end{aligned}$$

where $(g_n \in \mathcal{C}(Z, X))_{n\in\mathbb{N}}$ is the (obviously monotone) sequence of morphisms inductively defined by $g_0 = 0$ and $g_{n+1} = \mathsf{Ev} \circ \langle f, g_n\rangle$. □

8.2 The differential of fixpoints

Let $\mathcal{L}$ be a summable category.

Definition 7.3. *Let $f, g \in \mathcal{L}(X, Y)$, we write $f \leq g$ if there is an $h \in \mathcal{L}(X, Y)$ such that f and h are summable and $g = f + h$.*

Lemma 7.1. *The relation $\leq$ is a preorder relation on $\mathcal{L}(X, Y)$. The composition operation $\mathcal{L}(X, Y) \times \mathcal{L}(Y, Z) \to \mathcal{L}(X, Z)$ is monotone wrt. this preorder relation and if $\mathcal{L}$ is a symmetric monoidal summable category (i.e., (S-⊗) holds), then the tensor product of $\mathcal{L}$ is monotone wrt. this preorder relation. If $\mathcal{L}$ is a cartesian summable category (i.e., (S-&) holds), then the pairing operation $\mathcal{L}(X, Y_1) \times \mathcal{L}(X, Y_2) \to \mathcal{L}(X, Y_1 \,\&\, Y_2)$ is monotone.*

Proof. Monotonicity of composition results from Lemma 4.5. The two next statements are obvious consequences of 4.11 (S-⊗) and 4.14 (S-&), respectively. □

Definition 7.4. *We say that $\mathcal{L}$ is Scott if the following conditions are satisfied:*

- *in any homset, the relation $\leq$ is an order relation;*
- *for any objects X and Y, any monotone sequence $(f_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{L}(X, Y)$ has a least upper bound $\sup_{n\in\mathbb{N}} f_n \in \mathcal{L}(X, Y)$;*
- *the composition operation $\mathcal{L}(X, Y) \times \mathcal{L}(Y, Z) \to \mathcal{L}(X, Z)$ is Scott-continuous in the sense that it commutes with the lubs of monotone sequences (its domain being equipped with the product order relation);*
- *if $\mathcal{L}$ is a symmetric monoidal summable category, we also require $\otimes$ to commute with the lubs of monotone sequences;*
- *and if $\mathcal{L}$ is a summable resource category, then the functor !_ is required to be monotone and to commute with the lubs of monotone sequences.*

Lemma 7.2. *Let $\mathcal{L}$ be a summable symmetric monoidal closed category which is Scott. Then,* cur : $\mathcal{L}(Z \otimes X, Y) \to \mathcal{L}(Z, X \multimap Y)$ *is continuous.*

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Proof. Monotonicity results from Lemma 4.16. Continuity results from the fact that the inverse of the map $\mathsf{cur} : \mathcal{L}(Z \otimes X, Y) \to \mathcal{L}(Z, X \multimap Y)$ is the function $f \mapsto \mathsf{ev}\ (f \otimes X)$, which is continuous by our assumptions about $\mathcal{L}$. □

Proposition 7.3. *Let $\mathcal{L}$ be a cartesian summable category. If $\mathcal{L}$ is Scott then the pairing operation $\mathcal{L}(X, Y_1) \times \mathcal{L}(X, Y_2) \to \mathcal{L}(X, Y_1 \,\&\, Y_2)$ is continuous.*

Proof. By the universal property of the cartesian product. □

Theorem 7.2. *Let $\mathcal{L}$ be a Scott summable resource category which is closed (as an SM category). Then the cartesian closed category $\mathcal{L}_!$ is a λ-category (in the sense of Definition 7.2).*

The proof is straightforward, using Lemma 7.2.

► **Example. 7.1.** *Saying that $s, t \in \mathbf{Coh}_!(E, F)$ satisfy $s \leq t$ simply means $s \subseteq t$ and the fact that* **Coh** *is Scott comes from the fact that the set of cliques of a coherence space is closed under directed unions.*

The situation is completely similar in **Pcoh**. *Given $x, y \in \mathsf{P}X$, we have $x \leq y$ iff $x_a \leq y_a$ for all $a \in |X|$ as easily checked. Therefore, a monotone sequence $(x(n) \in \mathsf{P}X)_{n\in\mathbb{N}}$ has a lub in $\mathsf{P}X$, namely $x = (\sup_{n\in\mathbb{N}} x(n)_a)_{a\in|X|} \in \mathsf{P}X$. It follows easily that* **Pcoh** *is Scott. As a consequence, for any probabilistic coherence space X, we have $\mathcal{Y} \in \mathbf{Pcoh}_!(X \Rightarrow X, X)$, which maps $t \in \mathbf{Pcoh}_!(X, X)$ to its least fixpoint $\sup_{n\in\mathbb{N}} \hat{t}^n(0)$. The fact that this fixpoint operator is itself an analytic morphism is a remarkable property of this semantics and is deeply related to the fact that, in this semantics, the morphisms are matrices with nonnegative coefficients.*

If this were not the case, we could accept as a morphism the following $w \in \mathbf{Pcoh}_!(1 \,\&\, 1, 1)$ such that $\widehat{w}(u, v) = u + v - uv$. In other words, for $m \in |1 \,\&\, 1|$, we have $w_{m,*} = 1$ if $m \in \{[(1, *)], [(2, *)]\}$, $w_{m,*} = -1$ if $m = [(1, *), (2, *)]$ and $w_{m,*} = 0$ otherwise. If the fixpoint operator were accepted by this semantics, we would be able to define $t \in \mathbf{Pcoh}_!(1, 1)$ such that $\hat{t}(u) = \widehat{w}(u, \hat{t}(u)) = u + \hat{t}(u) - u\hat{t}(u)$, that is $u(1 - \hat{t}(u)) = 0$. So we must have $\hat{t}(u) = 1$ if $u \in (0, 1]$, and $\hat{t}(0) = 0$ because $\hat{t}(0)$ should be the least fixpoint of the function $\widehat{w}(0, _)$. So the function $\hat{t} : [0, 1] \to [0, 1]$ is not continuous, and a fortiori cannot be described as a power series (even with possibly negative coefficients).

Let $\mathcal{L}$ be a coherent differential resource category, which is closed (as an SMC) and Scott. Given $f \in \mathcal{L}_!(Z \,\&\, X, Y \Rightarrow Y)$, we can define $\mathsf{Fix}\,(f) \in \mathcal{L}_!(Z \,\&\, X, Y)$ as $\mathsf{Fix}\,(f) = \mathcal{Y} \circ f$ where $\mathcal{Y} \in \mathcal{L}_!(Y \Rightarrow Y, Y)$ comes from Theorem 7.1. Notice that

$$\mathsf{Fix}\,(f) = (f)\mathsf{Fix}\,(f)\,.$$

Remember also that the family $(\mathsf{Fix}_n\,(f) \in \mathcal{L}_!(Z \,\&\, X, Y))_{n\in\mathbb{N}}$ defined inductively by $\mathsf{Fix}_0\,(f) = 0$ and $\mathsf{Fix}_{n+1}\,(f) = (f)\mathsf{Fix}_n\,(f)$ is monotone and that

$$\mathsf{Fix}\,(f) = \sup_{n\in\mathbb{N}} \mathsf{Fix}_n\,(f)\ .$$

Theorem 7.3. *Let $\mathcal{L}$ be a coherent differential resource category, which is closed (as an SMC) and Scott. Let $f \in \mathcal{L}_!(Z \,\&\, X, Y \Rightarrow Y)$, we have*

$$\mathsf{T}_2\left(\mathsf{Fix}\,(f)\right) = \mathsf{Fix}\left((\mathsf{T}Y \Rightarrow \theta_Y) \circ \mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,Y} \circ \mathsf{T}_2 f\right)$$

Proof. It suffices to prove by induction on $n \in \mathbb{N}$ that

$$\mathsf{T}_2\left(\mathsf{Fix}_n\,(f)\right) = \mathsf{Fix}_n\left((\mathsf{T}Y \Rightarrow \theta_Y) \circ \mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,Y} \circ \mathsf{T}_2 f\right)\ .$$

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

The base case is obvious. Next we have

$$\begin{aligned}
\mathsf{T}_2\left(\mathsf{Fix}_{n+1}\left(f\right)\right) &= \mathsf{T}_2\left((f)\mathsf{Fix}_n\left(f\right)\right)\\
&= ((\mathsf{T}Y \Rightarrow \theta_Y) \circ \mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,Y} \circ \mathsf{T}_2 f)\mathsf{T}_2\left(\mathsf{Fix}_n\left(f\right)\right) \quad \text{by Theorem 6.1}\\
&= ((\mathsf{T}Y \Rightarrow \theta_Y) \circ \mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,Y} \circ \mathsf{T}_2 f)\mathsf{Fix}_n\left((\mathsf{T}Y \Rightarrow \theta_Y) \circ \mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,Y} \circ \mathsf{T}_2 f\right)\\
&\qquad\qquad \text{by inductive hypothesis}\\
&= \mathsf{Fix}_{n+1}\left((\mathsf{T}Y \Rightarrow \theta_Y) \circ \mathsf{T}^{\mathsf{int}} \circ \Psi^{\multimap}_{Y,Y} \circ \mathsf{T}_2 f\right) \quad \text{by definition.}
\end{aligned}$$

□

Remark 7.2. So the differential of a fixpoint can itself be written as a fixpoint, meaning that we can combine our coherent differential calculus with general fixpoints, which is another major difference with the differential λ-calculus and LL. These results justify the way we deal with fixpoints in $\mathsf{PCF}_{\mathsf{cd}}$ in Section 9.

8.3 Linear and multilinear morphisms

Definition 7.5. *A morphism $f \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, X_n, Y)$ is n-linear in $X_1, \ldots, X_n$ if there is $f_0 \in \mathcal{L}(X_1 \otimes \cdots \otimes X_n, Y)$ such that f coincides with the following composition of morphisms*

$$!\left(X_1 \,\&\, \cdots \,\&\, X_n\right) \xrightarrow{(\mathsf{m}^n)^{-1}} !X_1 \otimes \cdots \otimes !X_n \xrightarrow{\mathsf{der}_{X_1} \otimes \cdots \otimes \mathsf{der}_{X_n}} X_1 \otimes \cdots \otimes X_n \xrightarrow{f_0} Y$$

and in that case we write $f = \mathsf{Der}_n f_0$.

Theorem 7.4. *If $f \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, X_n, Y)$ is n-linear and $i \in \{1, \ldots, n\}$, then the ith partial differential $\mathsf{T}_i f \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, \mathsf{T}X_i \,\&\, \cdots \,\&\, X_n, \mathsf{T}Y)$ of f satisfies the following commutation in $\mathcal{L}_!$*

$$\begin{array}{ccc}
X_1 \,\&\, \cdots \,\&\, \mathsf{T}X_i \,\&\, \cdots \,\&\, X_n & \xrightarrow{\mathsf{T}_i f} & \mathsf{T}Y\\
{\scriptstyle X_1 \& \cdots \& \pi_j \& \cdots \& X_n}\downarrow & & \downarrow{\scriptstyle \pi_j}\\
X_1 \,\&\, \cdots \,\&\, X_i \,\&\, \cdots \,\&\, X_n & \xrightarrow{f} & Y
\end{array}$$

for $j = 0, 1$.

Notice that this diagram commutes for $j = 0$ for any $f \in \mathcal{L}_!(X_1 \,\&\, \cdots \,\&\, X_n, Y)$, so this result concerns only the case $j = 1$. This result is essential in the semantics of the constructions $\mathsf{if}^d(M, P, Q)$ and $\mathsf{let}^d(x, M, P)$ of the language $\mathsf{PCF}_{\mathsf{cd}}$ we describe now.

9. A Syntax for Coherent Differentiation

To conclude the paper, we describe briefly a syntax $\mathsf{PCF}_{\mathsf{cd}}$, which extends Scott-Plotkin's PCF with differentiation. The syntax is directly derived from the semantical framework described above. Its theory is developed in full detail in Ehrhard (2023a) to which we refer, so that most results in this section are provided without proofs. We start with some simple considerations about rewriting systems.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

9.1 Rewriting systems

Usually, a rewriting system is a set T of terms together with a rewriting relation $\rho \subseteq T \times T$.

In the present setting as well as in the original differential λ-calculus of Section 3.2, a term $t \in T$ (or a state of the Krivine machine that we will introduce) can reduce to several different terms, not because several redexes are available in t (as in the usual λ-calculus), but because t reduces to a "sum" of terms, since the rewriting system must somehow implement the Leibniz rule of Calculus.

So the rewriting relations that we consider have type $\rho \subseteq T \times \mathcal{M}_{\mathrm{fin}}(T)$. Such a relation can be lifted into a relation $\rho^\dagger \subseteq \mathcal{M}_{\mathrm{fin}}(T) \times \mathcal{M}_{\mathrm{fin}}(T)$ defined by $(m, m') \in \rho^\dagger$ if $m = [t] + m_1$ and $m' = m_0 + m_1$ with $(t, m_0) \in \rho$.

Lemma 8.1. *The reflexive and transitive closure $\rho^{\dagger *}$ of $\rho^\dagger$ is the least reflexive and transitive relation on $\mathcal{M}_{\mathrm{fin}}(T)$ such that*

- *if $(t, m) \in \rho$ then $([t], m) \in \rho^{\dagger *}$;*
- *and if $((m_i, m_i') \in \rho^{\dagger *})_{i=1,2}$ then $(m_1 + m_2, m_1' + m_2') \in \rho^{\dagger *}$.*

To enforce the algebraic flavor of this kind of rewriting, we adopt the following conventions:

- the singleton multiset $[t] \in \mathcal{M}_{\mathrm{fin}}(T)$ is simply written t;
- the empty multiset $[\,] \in \mathcal{M}_{\mathrm{fin}}(T)$ is simply written 0;
- and we use $t_1 + \cdots + t_k$ for $[t_1, \ldots, t_k] \in \mathcal{M}_{\mathrm{fin}}(T)$.

In other words, we identify $\mathcal{M}_{\mathrm{fin}}(T)$ with the free $\mathbb{N}$-semimodule generated by T.

9.2 Syntax

The grammar of types is inductively defined by

$$A, B, \cdots := \mathsf{D}^d\iota \mid A \Rightarrow B$$

where $d \in \mathbb{N}$ and then, given a type A, we define the type $\mathsf{D}A$ inductively by $\mathsf{D}(\mathsf{D}^d\iota) = \mathsf{D}^{d+1}\iota$ and $\mathsf{D}(A \Rightarrow B) = A \Rightarrow \mathsf{D}B$.

The syntax of terms is inductively defined as follows; we split it into three kinds of constructions:

$$\begin{aligned} M, N, P, Q, \ldots := {} & x \mid (P)N \mid \lambda x^A\, M \mid \mathsf{fix}\,(M) && \lambda\text{-calculus} \\ & \mid \underline{\nu} \mid \mathsf{succ}^d(M) \mid \mathsf{pred}^d(M) \mid \mathsf{if}^d(M, P, Q) \mid \mathsf{let}^d(x, M, N) && \text{arithmetics} \\ & \mid \mathsf{D}\,M \mid \iota_i^d(M) \mid \theta^d(M) \mid \mathsf{c}_l^d(M) \mid \pi_i^d(M) && \text{differentiation} \end{aligned}$$

where $\nu \in \mathbb{N}$, $d, l \in \mathbb{N}$, and $i \in \{0, 1\}$. The language has only one ground type ι for integers and associated successor and predecessor functions as well as a conditional that allows one to test if an integer is zero. There is also a let construct making it possible to use call-by-value on integers, which is crucial when the language is extended with probabilistic choice. The d exponents that decorate these constructs as well as the differential constructs are made necessary by the definition of the differential modification operation, see Figure 1.

Definition 8.1. *We say that a type E is sharp if it is not of the shape $E = \mathsf{D}A$ for some other type A, which simply means that $E = (A_1 \Rightarrow \cdots \Rightarrow A_n \Rightarrow \iota)$ (with the usual convention that $\Rightarrow$ associates on the right). We use letters E, F to denote types when we want to stress that they are sharp.*

Remark 8.1. For any type A there is exactly one $d \in \mathbb{N}$ and one sharp type such that $A = \mathsf{D}^d E$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

$$\partial(x,x) = x \qquad \partial(x,y) = \iota_0^0(y) \quad \text{if } y \neq x$$
$$\partial(x,(P)N) = (\theta^0(\mathsf{D}\,\partial(x,P)))\partial(x,N) \qquad \partial(x,\lambda y^B\,P) = \lambda y^B\,\partial(x,P)$$
$$\partial(x,\mathsf{fix}(P)) = \mathsf{fix}(\theta^0(\mathsf{D}\,\partial(x,P))) \qquad \partial(x,\underline{\nu}) = \iota_0^0(\underline{\nu})$$
$$\partial(x,\mathsf{succ}^d(N)) = \mathsf{succ}^{d+1}(\partial(x,N)) \qquad \partial(x,\mathsf{pred}^d(N)) = \mathsf{pred}^{d+1}(\partial(x,N))$$
$$\partial(x,\mathsf{if}^d(N,P,Q)) = \theta^0(\mathsf{c}_d^0(\mathsf{if}^{d+1}(\partial(x,N),\partial(x,P),\partial(x,Q))))$$
$$\partial(x,\mathsf{let}^d(y,N,P)) = \theta^0(\mathsf{c}_d^0(\mathsf{let}^{d+1}(y,\partial(x,N),\partial(x,P))))$$
$$\text{with } y \neq x$$
$$\partial(x,\mathsf{D}\,M) = \mathsf{c}_0^0(\mathsf{D}\,\partial(x,M)) \qquad \partial(x,\iota_i^d(M)) = \iota_i^{d+1}(\partial(x,M))$$
$$\partial(x,\theta^d(M)) = \theta^{d+1}(\partial(x,M)) \qquad \partial(x,\mathsf{c}_l^d(M)) = \mathsf{c}_l^{d+1}(\partial(x,M))$$
$$\partial(x,\pi_i^d(M)) = \pi_i^{d+1}(\partial(x,M))$$

Figure 1. Definition of the differential modification $\partial(x, M)$, by induction on M.

$$\frac{}{\Gamma, x : A \vdash x : A} \qquad \frac{\Gamma \vdash P : A \Rightarrow B \quad \Gamma \vdash N : A}{\Gamma \vdash (P)N : B} \qquad \frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda x^A\, M : A \Rightarrow B} \qquad \frac{\Gamma \vdash M : A \Rightarrow A}{\Gamma \vdash \mathsf{fix}(M) : A}$$

$$\frac{\nu \in \mathbb{N}}{\Gamma \vdash \underline{\nu} : \iota} \qquad \frac{\Gamma \vdash M : \mathsf{D}^d\iota}{\Gamma \vdash \mathsf{succ}^d(M) : \mathsf{D}^d\iota} \qquad \frac{\Gamma \vdash M : \mathsf{D}^d\iota}{\Gamma \vdash \mathsf{pred}^d(M) : \mathsf{D}^d\iota}$$

$$\frac{\Gamma \vdash M : \mathsf{D}^d\iota \quad \Gamma \vdash P : A \quad \Gamma \vdash Q : A}{\Gamma \vdash \mathsf{if}^d(M,P,Q) : \mathsf{D}^d A} \qquad \frac{\Gamma \vdash M : \mathsf{D}^d\iota \quad \Gamma, x : \iota \vdash P : A}{\Gamma \vdash \mathsf{let}^d(x,M,P) : \mathsf{D}^d A}$$

$$\frac{\Gamma \vdash M : A \Rightarrow B}{\Gamma \vdash \mathsf{D}\,M : \mathsf{D}A \Rightarrow \mathsf{D}B} \qquad \frac{\Gamma \vdash M : \mathsf{D}^d A}{\Gamma \vdash \iota_i^d(M) : \mathsf{D}^{d+1} A} \qquad \frac{\Gamma \vdash M : \mathsf{D}^{d+2} A}{\Gamma \vdash \theta^d(M) : \mathsf{D}^{d+1} A}$$

$$\frac{\Gamma \vdash M : \mathsf{D}^{d+l+2} A}{\Gamma \vdash \mathsf{c}_l^d(M) : \mathsf{D}^{d+l+2} A} \qquad \frac{\Gamma \vdash M : \mathsf{D}^{d+1} A}{\Gamma \vdash \pi_i^d(M) : \mathsf{D}^d A}$$

Figure 2. Typing rules for $\mathsf{PCF}_{\mathsf{cd}}$.

The typing rules are given in Figure 2.

As explained in Ehrhard (2023a), this syntax can be equipped with a rewriting system β_{cd}, which is inspired by the categorical setting described previously; we refer to this article for a precise definition of β_{cd}. Beyond ordinary substitution, the definition of this rewriting system requires a *differential modification* operator $\partial(x, M)$ whose definition, by induction on M, is given in Figure 1. This inductive definition uses the constructions $\iota_i^d(N)$, $\theta^d(N)$, and $\mathsf{c}_l^d(N)$, which are associated with the natural transformations similarly named in the description of summability structures in Section 5. See also Section 9.3, and more specifically Figure 3, for an operational description of these constructs.

Remark 8.2. It is important to notice that, in sharp contrast with the linear substitution $\frac{\partial M}{\partial x} \cdot N$ of the differential λ-calculus, the construction $\partial(x, M)$ does not introduce actual sums of terms, but only potential ones by inserting $\theta^d(_)$ syntactic constructs at various places.

Lemma 8.2. *If* $\Gamma, x : A \vdash M : B$, *then* $\Gamma, x : \mathsf{D}A \vdash \partial(x, M) : \mathsf{D}B$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

$$(\varepsilon\delta, \pi_i^{\mathsf{len}(\delta)}(N), s) \to (\varepsilon i\delta, N, s)$$
$$(\varepsilon i\delta, \iota_i^{\mathsf{len}(\delta)}(N), s) \to (\varepsilon\delta, N, s) \qquad (\varepsilon i\delta, \iota_{1-i}^{\mathsf{len}(\delta)}(N), s) \to 0$$
$$(\varepsilon 0\delta, \theta^{\mathsf{len}(\delta)}(N), s) \to (\varepsilon 00\delta, N, s) \qquad (\varepsilon 1\delta, \theta^{\mathsf{len}(\delta)}(N), s) \to (\varepsilon 01\delta, N, s) + (\varepsilon 10\delta, N, s)$$
$$(\varepsilon i_1 \cdots i_{l+2}\delta, \mathsf{c}_l^{\mathsf{len}(\delta)}(N), s) \to (\varepsilon i_{l+2} i_1 \cdots i_{l+1}\delta, N, s)$$

Figure 3. Transition rules for states – handling the access word.

Proof hint. Straightforward induction on M, or rather on the derivation of the typing judgment $\Gamma, x : A \vdash M : B$. As a first example, assume that $M = \mathsf{fix}\,(N)$ with $\Gamma, x : A \vdash N : B \Rightarrow B$ so that $\Gamma, x : A \vdash \mathsf{fix}\,(N) : B$. By inductive hypothesis, we have $\Gamma, x : \mathsf{D}A \vdash \partial(x, N) : B \Rightarrow \mathsf{D}B$ and hence $\Gamma, x : \mathsf{D}A \vdash \mathsf{D}\,\partial(x, N) : \mathsf{D}B \Rightarrow \mathsf{D}^2B$ so that $\Gamma, x : \mathsf{D}A \vdash \theta^0(\,\mathsf{D}\,\partial(x, N)) : \mathsf{D}B \Rightarrow \mathsf{D}B$ since $(\mathsf{D}B \Rightarrow \mathsf{D}^2B) = \mathsf{D}^2(\mathsf{D}B \Rightarrow B)$. Therefore, $\Gamma, x : \mathsf{D}A \vdash \mathsf{fix}\,(\theta^0(\,\mathsf{D}\,\partial(x, N))) : \mathsf{D}B$ as required.

As a second example, take $M = \mathsf{if}^d(N, P, Q)$ with $\Gamma, x : A \vdash N : \mathsf{D}^d\iota$, $\Gamma, x : A \vdash P : C$ and $\Gamma, x : A \vdash Q : C$ so that $B = \mathsf{D}^dC$. By inductive hypothesis, we have $\Gamma, x : \mathsf{D}A \vdash \partial(x, N) : \mathsf{D}^{d+1}\iota$, $\Gamma, x : \mathsf{D}A \vdash \partial(x, P) : \mathsf{D}C$ and $\Gamma, x : \mathsf{D}A \vdash \partial(x, Q) : \mathsf{D}C$. It follows that

$$\Gamma, x : \mathsf{D}A \vdash \mathsf{if}^{d+1}(\partial(x, N), \partial(x, P), \partial(x, Q)) : \mathsf{D}^{d+1+1}C$$

and hence

$$\Gamma, x : \mathsf{D}A \vdash \mathsf{c}_d^0(\mathsf{if}^{d+1}(\partial(x, N), \partial(x, P), \partial(x, Q))) : \mathsf{D}^{d+2}C$$

and finally

$$\Gamma, x : \mathsf{D}A \vdash \theta^0(\mathsf{c}_d^0(\mathsf{if}^{d+1}(\partial(x, N), \partial(x, P), \partial(x, Q)))) : \mathsf{D}B$$

as required. □

Remark 8.3. The main purpose of this differential modification is to make the following rewriting possible

$$\mathsf{D}\left(\lambda x^A\, M\right) \beta_{\mathsf{cd}}\, \lambda x^{\mathsf{D}A}\, \partial(x, M)\,,$$

which gives its operational meaning to the D construction of the syntax, exactly as the ordinary β-rewriting $(\lambda x^A\, M)N\ \beta_{\mathsf{cd}}\ M\,[N/x]$ gives its operational meaning to the application construct of the λ-calculus.

In contrast with the differential substitution of the differential λ-calculus, differentiation in $\mathsf{PCF}_{\mathsf{cd}}$ requires a combination of differential modification and ordinary substitution. Let N be such that $\Gamma \vdash N : DA$, so that N should be intuitively considered as the pair made of $(N_i = \pi_i^0(N))_{i=0,1}$ which satisfy $(\Gamma \vdash N_i : A)_{i=0,1}$ and are summable in the type A. Then the term

$$\pi_1^0(\partial(x, M)\,[N/x]\,)$$

of $\mathsf{PCF}_{\mathsf{cd}}$ has the same meaning as the term

$$\left(\frac{\partial M}{\partial x} \cdot \pi_1^0(N)\right)\left[\pi_0^0(N)/x\right]$$

of the differential λ-calculus. For that reason we can understand $\mathsf{PCF}_{\mathsf{cd}}$ as a sub-language of the differential λ-calculus (extended with integers, pairs and fixpoint operators).

Remark 8.4. It is important to notice that the rewriting β_{cd} we equip $\mathsf{PCF}_{\mathsf{cd}}$ with is not an ordinary rewriting relation from terms to terms, but from terms to finite multisets of terms as explained in Section 9.1. More specifically, there are exactly two rewriting rules that produce non-singleton multisets, namely $\pi_i^d(\iota_{1-i}^d(M))\ \beta_{\mathsf{cd}}\ 0$ (for $i = 0, 1$) and $\pi_1^d(\theta^d(M))\ \beta_{\mathsf{cd}}\ \pi_0^d(\pi_1^d(M)) +$

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

$$\frac{}{() : \iota \vdash \iota} \qquad \frac{s : \iota \vdash \iota}{\mathsf{succ} \cdot s : \iota \vdash \iota} \qquad \frac{s : \iota \vdash \iota}{\mathsf{pred} \cdot s : \iota \vdash \iota}$$

$$\frac{\vdash P : \mathsf{D}^d E \qquad \vdash Q : \mathsf{D}^d E \qquad s : E \vdash \iota \qquad \delta \in \{0,1\}^d}{\mathsf{if}(\delta, P, Q) \cdot s : \iota \vdash \iota} \qquad \frac{x : \iota \vdash P : \mathsf{D}^d E \qquad s : E \vdash \iota \qquad \delta \in \{0,1\}^d}{\mathsf{let}(\delta, x, P) \cdot s : \iota \vdash \iota}$$

$$\frac{\vdash P : A \qquad s : E \vdash \iota}{\mathsf{arg}(P) \cdot s : A \Rightarrow E \vdash \iota} \qquad \frac{s : \mathsf{D}A \Rightarrow E \vdash \iota \qquad i \in \{0,1\}}{\mathsf{D}(i) \cdot s : A \Rightarrow E \vdash \iota}$$

Figure 4. Typing rules for stacks.

$\pi_1^d(\pi_0^d(M))$. For this rewriting system, one can prove a form of subject reduction, which expresses that if $\Gamma \vdash : MA$ and $M\ \beta_{\mathsf{cd}}\ \sum_{i=1}^k M_i$, then we have $\Gamma \vdash M_i : A$ for $i = 1, \ldots, k$. Notice that actually $k \in \{0, 1, 2\}$.

Lemma 8.2, together with an ordinary substitution lemma (if $\Gamma, x : A \vdash M : B$ and $\Gamma \vdash N : A$ then $\Gamma \vdash M\,[N/x] : B$), allows us to prove subject reduction.

Theorem 8.1. *If* $\Gamma \vdash : MA$ *and* $M\ \beta_{\mathsf{cd}}\ \sum_{i=1}^k M_i$, *then* $(\Gamma \vdash M_i : A)_{i=1}^k$.

Corollary 8.1. *Assume that* $M_1, \ldots, M_p$ *are terms such that* $(\Gamma \vdash M_i : A)_{j=1}^p$. *If* $\sum_{j=1}^p M_j\ \beta_{\mathsf{cd}}^{\dagger *}$ $\sum_{j=1}^{p'} M'_j$, *then we have* $(\Gamma \vdash M'_j : A)_{j=1}^{p'}$.

9.3 Operational semantics

Rather than providing a complete definition of the $\mathsf{PCF}_{\mathsf{cd}}$ reduction system, which is lengthy and has already been given in Ehrhard (2023a),[13] we describe a seemingly more canonical "Krivine machine" which evaluates terms M of $\mathsf{PCF}_{\mathsf{cd}}$ such that $\vdash M : \iota$.

A state of the machine is a triple (δ, M, s) where M is a closed term, s is a stack and $\delta \in \{0,1\}^{<\omega}$. Our stacks are defined by the following grammar:

$$s, t, \ldots := () \mid \mathsf{arg}(M) \cdot s \mid \mathsf{succ} \cdot s \mid \mathsf{pred} \cdot s \mid \mathsf{if}(\delta, P, Q) \cdot s \mid \mathsf{let}(\delta, x, P) \cdot s \mid \mathsf{D}(i) \cdot s\,.$$

Stacks are typed by judgments of shape $s : E \vdash \iota$ where E is a sharp type. This typing judgment means intuitively that the stack s is a continuation which expects an argument of type E to produce a result of type ι. The typing rules for stacks are given in Figure 4.

Definition 8.2. *A state* $e = (\delta, M, s)$ *is well typed if* $\vdash M : \mathsf{D}^{\mathsf{len}(\delta)} E$ *and* $s : E \vdash \iota$ *for some sharp type* E.

The transition rules for states are given in Figures 3, Figures 5, Figures 6 where we classified them into three categories. Notice that this is a rewriting system in the sense of Section 9.1, that is, from states to finite multisets (or finite formal sums) of states. The only rules yielding actual sums are the 3rd and 5th transition rules in Figure 3.

With a stack s such that $s : E \vdash \iota$ we can associate a context $\langle s \rangle [\ ]$, that is, a closed term of type ι with one hole $[\ ]$ of type E in linear position; the definition of this context is given in Figure 7.

Lemma 8.3. *If* $s : E \vdash \iota$ *and* $\vdash M : E$, *then* $\vdash \langle s \rangle [M] : \iota$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

$$(\delta, \mathsf{succ}^{\mathsf{len}(\delta)}(M), s) \to (\delta, M, \mathsf{succ}\cdot s) \qquad (\delta, \mathsf{pred}^{\mathsf{len}(\delta)}(M), s) \to (\delta, M, \mathsf{pred}\cdot s)$$
$$(\varepsilon\delta, \mathsf{if}^{\mathsf{len}(\delta)}(M,P,Q), s) \to (\delta, M, \mathsf{if}(\varepsilon,P,Q)\cdot s) \qquad (\varepsilon\delta, \mathsf{let}^{\mathsf{len}(\delta)}(x,M,P), s) \to (\delta, M, \mathsf{let}(\varepsilon,x,P)\cdot s)$$
$$(\delta, (N)P, s) \to (\delta, N, \mathsf{arg}(P)\cdot s) \qquad (\delta, \mathsf{fix}(N), s) \to (\delta, N, \mathsf{arg}(\mathsf{fix}(N))\cdot s)$$
$$(\delta i, \mathsf{D}\,N, s) \to (\delta, N, \mathsf{D}(i)\cdot s)$$

Figure 5. Transition rules for states – pushing onto the stack.

$$(\langle\rangle, \underline{\nu}, \mathsf{succ}\cdot s) \to (\langle\rangle, \underline{\nu+1}, s) \qquad (\langle\rangle, \underline{\nu}, \mathsf{pred}\cdot s) \to (\langle\rangle, \underline{\nu \dot{-} 1}, s)$$
$$(\langle\rangle, \underline{0}, \mathsf{if}(\delta,P,Q)\cdot s) \to (\delta, P, s) \qquad (\langle\rangle, \underline{\nu+1}, \mathsf{if}(\delta,P,Q)\cdot s) \to (\delta, Q, s)$$
$$(\langle\rangle, \underline{\nu}, \mathsf{let}(\delta,x,P)\cdot s) \to (\delta, P\,[\underline{\nu}/x]\,, s)$$
$$(\delta, \lambda x^B\, N, \mathsf{arg}(P)\cdot s) \to (\delta, N\,[P/x]\,, s) \qquad (\delta, \lambda x^B\, N, \mathsf{D}(i)\cdot s) \to (\delta i, \lambda x^{\mathsf{D}B}\, \partial(x,N), s)$$

Figure 6. Transition rules for states – popping from the stack.

$$\langle()\rangle = [\,]$$
$$\langle \mathsf{succ}\cdot s\rangle = \langle s\rangle[\mathsf{succ}^0([\,])] \qquad \langle \mathsf{pred}\cdot s\rangle = \langle s\rangle[\mathsf{pred}^0([\,])]$$
$$\langle \mathsf{if}(\delta,P,Q)\cdot s\rangle = \langle s\rangle[\pi^0_\delta(\mathsf{if}^0([\,],P,Q))] \qquad \langle \mathsf{let}(\delta,x,P)\cdot s\rangle = \langle s\rangle[\pi^0_\delta(\mathsf{let}^0(x,[\,],P))]$$
$$\langle \mathsf{arg}(P)\cdot s\rangle = \langle s\rangle[([\,])P] \qquad \langle \mathsf{D}(i)\cdot s\rangle = \langle s\rangle[\pi^0_i(\mathsf{D}[\,])]$$

Figure 7. Context associated with a stack.

Proof. Straightforward induction on s. □

Definition 8.3. *Given a state e we define a term $\langle e\rangle$ by $\langle(\delta, M, s)\rangle = \langle s\rangle[\pi^0_\delta(M)]$.*

Lemma 8.4. *If e is a well-typed state, then $\vdash \langle e\rangle : \iota$.*

Proof. Direct consequence of Lemma 8.3. □

Theorem 8.2. *If $e \to \sum_{i=1}^k e_i$, then $\langle e\rangle\ \beta_{\mathsf{cd}}\ \sum_{i=1}^k \langle e_i\rangle$ (notice that $k \in \{0,1,2\}$).*

Remark 8.5. Performing transitions from the state $e = (\delta, M, s)$ amounts actually to evaluating the term $\pi^0_\delta(M)$ in the environment s in a "weak head" restriction of the $\mathsf{PCF}_{\mathsf{cd}}$ reduction system. As a whole, we could consider (δ, s) as the context $\langle s\rangle[\pi^0_\delta([\,])]$, which suggests to integrate the access path δ in the stack since the purpose of the stack is to store the current context of evaluation.

We did not do so because many rules of the rewriting system β_{cd} express some commutations between the $\pi^d_i(_)$ constructs (stored in the δ component of the state) and the other constructs of the language (stored in the s component). These commutations express that the s-context and the δ-context act in parallel on the term component of the machine, strongly suggesting to keep them separate. The benefit of this choice is that, in the transition rules of Figures 3, Figures 5, Figures 6, we do not mention these commutations explicitly: they are implemented in a purely implicit way, which is a major improvement of this machine wrt. the rewriting system β_{cd}.

For instance, the rewriting system features the reduction $\pi^0_i((N)P)\ \beta_{\mathsf{cd}}\ (\pi^0_i(N))P$ where we see that the action of the projection is transferred from $(N)P$ to N. This transfer of action of the projection is implemented implicitly in the transitions $(\delta, \pi^0_i((N)P), s) \to (\delta i, (N)P, s) \to (\delta i, N, \mathsf{arg}(P)\cdot s)$ of Figures 5 and 3. Using only a stack for storing the context, we would have obtained a sequence of reductions like $(\pi^0_i((N)P), s) \to ((N)P, \pi^0_i\cdot s) \to (N, \mathsf{arg}(P)\cdot\pi^0_i\cdot s)$. But

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

we might have $N = \theta^0(\lambda x^B\, Q)$ and then the only natural option – keeping in mind the fundamental principle that the stack should be accessed only from the top – would be to push again the θ^0 onto the stack, leading to something like $(\lambda x^B\, Q, \theta^0 \cdot \mathsf{arg}(P) \cdot \pi_i^0 \cdot s)$ but then the argument that the abstraction $\lambda x^B\, Q$ is waiting for is not available on the top of the stack. To solve this issue, we would need an equivalence relation on stacks accounting for the above-mentioned commutation reduction rules of β_{cd}. In other words, the stack should not be indexed by a finite totally ordered set (i.e., should not be a list), but rather by a tree or perhaps a more general directed acyclic graph. Our dichotomy between the stack and the access word avoids these technicalities in a very simple and, we think, natural way. Notice by the way that the access word is not dealt with as a stack since we insert and remove elements anywhere in the word, and even perform cyclic permutations of factors, see Figure 3. A simple implementation of such a data structure could use a linked list.

Definition 8.4. *A final state is a state of shape* $(\langle\rangle, \underline{\nu}, ())$ *with* $\nu \in \mathbb{N}$.

Lemma 8.5. *If e is a well-typed state and there is no transition from e, then e is a final state.*

Proof. Simple case analysis on the typing rules of terms and stacks. □

Let Θ_{cd} be the set of all well typed states.

Lemma 8.6. *If $e \in \Theta_{\mathsf{cd}}$, $e \to u$ and e' is a state such that $u_{e'} \neq 0$, then $e' \in \Theta_{\mathsf{cd}}$.*

Proof. Simple inspection of the transition rules. As an example taken from Figure 3, assume that $e = (\varepsilon 1\delta, \theta^{\mathsf{len}(\delta)}(N), s)$ and $u = (\varepsilon 01\delta, N, s) + (\varepsilon 10\delta, N, s)$. So $e' \in \{(\varepsilon 01\delta, N, s), (\varepsilon 10\delta, N, s)\}$, say $e' = (\varepsilon 10\delta, N, s)$. Let $d = \mathsf{len}(\delta)$. There must be a type A such that $\vdash N : \mathsf{D}^{d+2}A$ so that $\vdash \theta^d(N) : \mathsf{D}^{d+1}A$. There are uniquely determined sharp type E and $h \in \mathbb{N}$ such that $A = \mathsf{D}^h E$ and hence $\vdash \theta^d(N) : \mathsf{D}^{d+1+h}E$ and since e is well typed we must have $h = \mathsf{len}(\varepsilon)$ and $s : E \vdash \iota$. So $\mathsf{len}(\varepsilon 10\delta) = d + 2 + h$ and since $\vdash N : \mathsf{D}^{d+2+h}E$, the state e' is well typed.

Let us also deal with the case $e = (\delta i, \mathsf{D}\, N, s)$ and $e' = (\delta, N, \mathsf{D}(i) \cdot s)$ of Figure 5. For $\mathsf{D}\, N$ to be typed we need $\vdash N : B \Rightarrow A$ and then we have $\vdash \mathsf{D}\, N : \mathsf{D}B \Rightarrow \mathsf{D}A$. There are uniquely determined sharp type E and $d \in \mathbb{N}$ such that $A = \mathsf{D}^d E$, so that $\vdash \mathsf{D}\, N : \mathsf{D}^{d+1}(\mathsf{D}B \Rightarrow E)$. Since e is well typed, we must have $d = \mathsf{len}(\delta)$ and $s : \mathsf{D}B \Rightarrow E \vdash \iota$ so that $\mathsf{D}(i) \cdot s : B \Rightarrow E \vdash \iota$ (see Figure 4) and hence e' is well typed since we have $\vdash N : \mathsf{D}^d(B \Rightarrow E)$.

As a last example consider the case $e = (\delta, \lambda x^B\, N, \mathsf{D}(i) \cdot s)$ and $e' = (\delta i, \lambda x^{\mathsf{D}B}\, \partial(x, N), s)$ from Figure 6. We must have $x : B \vdash N : A$ for some type $A = \mathsf{D}^d E$ (with E sharp and $d \in \mathbb{N}$ uniquely defined). Accordingly $\vdash \lambda x^B\, N : \mathsf{D}^d(B \Rightarrow E)$ and since e is well typed, we must have $d = \mathsf{len}(\delta)$ and $\mathsf{D}(i) \cdot s : B \Rightarrow E \vdash \iota$ which, by the typing rules of Figure 4, entails $s : \mathsf{D}B \Rightarrow E \vdash \iota$. By Lemma 8.2, we have $x : \mathsf{D}B \vdash \partial(x, N) : \mathsf{D}^{d+1}E$ so that e' is well typed since $\mathsf{len}(\delta i) = d + 1$. □

Remark 8.6. Notice that, on one side, the transition rules of Figures 5 to 3 are deterministic in the sense that, for any $e \in \Theta_{\mathsf{cd}}$, there is at most one $u \in \mathcal{M}_{\mathsf{fin}}(\Theta_{\mathsf{cd}})$ such that $e\, \beta_{\mathsf{cd}}\, u$, and that, when there is no such transition from e, then e is final in the sense of Definition 8.4. So we can define a function $\rho_{\mathsf{cd}} : \Theta_{\mathsf{cd}} \to \mathcal{M}_{\mathsf{fin}}(\Theta_{\mathsf{cd}})$ such that $\rho_{\mathsf{cd}}(e) = u$ if $e\, \beta_{\mathsf{cd}}\, u$ and $\rho_{\mathsf{cd}}(e) = e$ if e is final.

On the other side, these transition rules contain some nondeterminism precisely in the fact that transitions are from a state e to a finite multisets of states $e_1 + \cdots + e_k$ (which can be understood as the various possible results of a transition from e) and not from states to states. One of the purposes of the next section is to show that this nondeterminism is an illusion.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

9.4 Denotational semantics

Let $\mathcal{L}$ be a coherent differential resource category, which is closed (as an SMC) and Scott and where the coproduct $\mathsf{N} = \oplus_{i\in\mathbb{N}} 1$ exists.

First, we define by induction on the type A an object $[\![A]\!]$. We take $[\![\iota]\!] = \mathsf{N}$ and more generally $[\![\mathsf{D}^d\iota]\!] = \mathsf{T}^d\mathsf{N}$. And then $[\![A \Rightarrow B]\!] = (![\![A]\!] \multimap [\![B]\!])$.

Then, given a term M of $\mathsf{PCF}_{\mathsf{cd}}$, a context $\Gamma = (x_1 : A_1, \dots, x_n : A_n)$ and a type B such that $\Gamma \vdash : MB$, one defines by induction on the typing derivation of $\Gamma \vdash : MB$ (i.e., by induction on M) an element $[\![M]\!]_\Gamma \in \mathcal{L}_!([\![\Gamma]\!], [\![B]\!])$.

We refer to Ehrhard (2023a) for the precise definition of this interpretation of terms; the syntax of $\mathsf{PCF}_{\mathsf{cd}}$ has been chosen in order to make it fairly straightforward. Concerning the "object of integers" N, we use

- the existence of a canonical isomorphism $\chi_{\mathsf{N}} \in \mathcal{L}(1 \oplus \mathsf{N}, \mathsf{N})$, which is the key ingredient for interpreting $\mathsf{succ}^d(M)$, $\mathsf{pred}^d(M)$ and $\mathsf{if}^d(M, P, Q)$;
- the existence of a canonical !-coalgebra structure on N, which is the key ingredient for interpreting $\mathsf{let}^d(x, M, P)$. This is due to the fact that 1 is a !-coalgebra (thanks to the Seely isomorphisms) and to the fact that !-coalgebras are closed under arbitrary colimits which exist in $\mathcal{L}$.

As an example, using the definition of N as a coproduct, we can define, for any object X of $\mathcal{L}$, a morphism $f \in \mathcal{L}(\mathsf{N} \otimes (X \,\&\, X)\,, X)$ uniquely characterized by

$$\begin{array}{ccc} 1 \otimes (X \,\&\, X) & \xrightarrow{\overline{0}\otimes(X\&X)} & \mathsf{N} \otimes (X \,\&\, X) \\ {\scriptstyle\lambda}\downarrow & & \downarrow{\scriptstyle f} \\ X \,\&\, X & \xrightarrow{\quad\mathsf{pr}_1\quad} & X \end{array} \qquad \begin{array}{ccc} 1 \otimes (X \,\&\, X) & \xrightarrow{\overline{\nu+1}\otimes(X\&X)} & \mathsf{N} \otimes (X \,\&\, X) \\ {\scriptstyle\lambda}\downarrow & & \downarrow{\scriptstyle f} \\ X \,\&\, X & \xrightarrow{\quad\mathsf{pr}_2\quad} & X \end{array}$$

where, for $\nu \in \mathbb{N}$, the morphism $\overline{\nu} \in \mathcal{L}(1, \mathsf{N})$ is the νth injection of 1 into the coproduct N. We set $\overline{\mathsf{if}} = \mathsf{Der}_2 f \in \mathcal{L}_!(\mathsf{N} \,\&\, (X \,\&\, X), Y)$ which is bilinear (see Section 8.3). Then using Definition 5.8 we define $\overline{\mathsf{if}}^d = \mathsf{T}_1^d\overline{\mathsf{if}} \in \mathcal{L}_!(\mathsf{T}^d\mathsf{N} \,\&\, (X \,\&\, X), \mathsf{T}^d X)$ that we use straightforwardly to interpret the $\mathsf{if}^d(M, P, Q)$ construct of $\mathsf{PCF}_{\mathsf{cd}}$.

Then one can prove a standard substitution lemma.

Lemma 8.7. *If* $\Gamma, x : A \vdash M : B$ *and* $\Gamma \vdash : NA$, *one has* $[\![M\,[N/x]]\!]_\Gamma = [\![M]\!]_{\Gamma,x:A} \circ \langle [\![\Gamma]\!], [\![N]\!]_\Gamma \rangle$ *in* $\mathcal{L}_!$.

Notice indeed that $[\![M]\!]_{\Gamma,x:A} \in \mathcal{L}_!([\![\Gamma]\!] \,\&\, [\![A]\!], [\![B]\!])$ and $\langle [\![\Gamma]\!], [\![N]\!]_\Gamma \rangle \in \mathcal{L}_!([\![\Gamma]\!], [\![\Gamma]\!] \,\&\, [\![A]\!])$.

We have an analogous lemma for the differential modification.

Lemma 8.8. *If* $\Gamma, x : A \vdash M : B$ *then* $[\![\partial(x, M)]\!]_{\Gamma,x:\mathsf{D}A} = \mathsf{T}_2\left([\![M]\!]_{\Gamma,x:A}\right)$.

Notice that $[\![M]\!]_{\Gamma,x:A} \in \mathcal{L}_!([\![\Gamma]\!] \,\&\, [\![A]\!], [\![B]\!])$ and hence $\mathsf{T}_2\left([\![M]\!]_{\Gamma,x:A}\right) \in \mathcal{L}_!([\![\Gamma]\!] \,\&\, \mathsf{T}[\![A]\!], \mathsf{T}[\![B]\!])$ so that the equation above is well typed.

Theorem 8.3. *If* $\Gamma \vdash : MA$ *and* $M\ \beta_{\mathsf{cd}}\ \sum_{i=1}^k M_i$, *then the morphisms* $[\![M_i]\!]_\Gamma \in \mathcal{L}_!([\![\Gamma]\!], [\![A]\!])$ *are summable in* $\mathcal{L}(![\![\Gamma]\!], [\![A]\!])$ *and we have* $[\![M]\!]_\Gamma = \sum_{i=1}^k [\![M_i]\!]_\Gamma$.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

This result expresses the soundness of this denotational semantics. The proof uses Lemmas 8.7 and 8.8. Using the notions introduced in Section 9.1, this generalizes easily as follows.

Corollary 8.2. *Assume that* $M_1, \dots, M_p$ *are terms such that* $(\Gamma \vdash M_i : A)_{j=1}^{p}$ *and the morphisms* $(\llbracket M_i \rrbracket_\Gamma \in \mathcal{L}(!\llbracket \Gamma \rrbracket, \llbracket A \rrbracket))_{j=1}^{p}$ *are summable. If* $\sum_{j=1}^{p} M_j \; \beta_{\mathsf{cd}}^{\dagger *} \; \sum_{j=1}^{p'} M'_j$*, then the morphisms* $(\llbracket M'_j \rrbracket_\Gamma \in \mathcal{L}(!\llbracket \Gamma \rrbracket, \llbracket A \rrbracket))_{j=1}^{p'}$ *are summable, and we have* $\sum_{j=1}^{p} \llbracket M_j \rrbracket_\Gamma = \sum_{j=1}^{p'} \llbracket M'_j \rrbracket_\Gamma$.

9.4.1 Adequacy and determinism

Now we specialize to the case where $\mathcal{L} = \mathbf{Pcoh}$, see Example 4.2 and 5.2.

If $\vdash M : A$ then we know that $\llbracket M \rrbracket \in \mathsf{P}\llbracket A \rrbracket \subseteq (\mathbb{R}_{\geq 0})^{|\llbracket A \rrbracket|}$. Moreover, a simple inspection of the definition of the semantics shows that actually $\llbracket M \rrbracket \in \mathbb{N}^{|\llbracket A \rrbracket|}$. Of course the situation would be different if the language $\mathsf{PCF}_{\mathsf{cd}}$ were extended with a probabilistic choice operator (or more simply, *e.g.*, with a "constant" rand of type ι which has probability $1/2$ to reduce to $\underline{0}$ and $1/2$ to reduce to $\underline{1}$), but this is not the case in the present paper and in Ehrhard (2023a).

If $A = \iota$, this means that $\llbracket M \rrbracket \in \mathbb{N}^{\mathbb{N}}$ and that we have

$$\sum_{\nu \in \mathbb{N}} \llbracket M \rrbracket_\nu \in [0, 1]$$

so that $\forall \nu \in \mathbb{N} \; \llbracket M \rrbracket_\nu \in \{0, 1\}$ and there is at most one $\nu \in \mathbb{N}$ such $\llbracket M \rrbracket_\nu = 1$. In other words, either $\llbracket M \rrbracket = 0$ or $\llbracket M \rrbracket = \mathsf{e}_\nu$ (for a uniquely determined $\nu \in \mathbb{N}$).

Theorem 8.4. *Let* M *be a term such that* $\vdash M : \iota$ *and let* $\nu \in \mathbb{N}$*. The two following conditions are equivalent.*

- $\llbracket M \rrbracket = \mathsf{e}_\nu$
- $(\langle\rangle, M, ()) \; \beta_{\mathsf{cd}}^{\dagger *} \; \underline{\nu}$.

The implication $\Leftarrow$ boils down to Corollary 8.2 through the translation $\langle e \rangle$ from states to terms and Theorem 8.2. The implication $\Rightarrow$ is proven using an adaptation of the reducibility method applied to an intersection typing system associated with a relational semantics of $\mathsf{PCF}_{\mathsf{cd}}$ which underlies the **Pcoh** semantics.

So the calculus $\mathsf{PCF}_{\mathsf{cd}}$, and its operational semantics formalized by our Krivine machine, is essentially deterministic in the sense that, starting from a well typed state $(\langle\rangle, M, ())$, there is at most one reduction path which leads to a final state $(\langle\rangle, M, ())$ where $\nu \in \mathbb{N}$ is uniquely determined by M (interpreting the reduction $e \to e_1 + e_2$ as a nondeterministic choice), the other ones leading to 0. The situation is not completely satisfactory yet since we do not know *a priori* which transition path is "the good one".

Another important contribution of Ehrhard (2023a) is a solution of this issue based on a simple and natural idea suggested to the author by Guillaume Geoffroy: make the access word δ of a state (δ, N, s) writable.

10. Conclusion

We have presented coherent differentiation from a semantical and syntactical point of view, explaining how this new setting combines the ideas of differential LL with determinism and with probabilistic computations.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Even if we consider this as a major improvement wrt. the earlier approaches to differential LL, the precise meaning of the resulting functional calculus is still mysterious. More recently, in a joint work with Aymeric Walch, we have extended this approach to iterated derivatives and to Taylor expansions of terms, still in a deterministic setting (Ehrhard and Walch 2025). These new results might provide the sought programming interpretation of coherent differentiation: it indeed allows one to put, within the language and in a functional way,[14] strong restrictions on the resource consumption of programs.

Acknowledgments. I would like to thank first my PhD student Aymeric Walch who made many important observations about the first presentation of coherent differentiation in Ehrhard (2023b) of which the present paper has benefited crucially, mainly in Sections 5.2, 5.3 and 6. Many thanks also to the reviewers for their insightful suggestions. This work was partly supported by the ANR project *Probabilistic Programming Semantics (PPS)* ANR-19-CE48-0014.

Notes

1 The number $M^*_s \in \mathbb{Q}$ is the coefficient of s in the Taylor expansion of M; it is > 0. More precisely $M^*_s = 1/k$ for some $k \in \mathbb{N} \setminus \{0\}$, which is a kind of "multinomial coefficient" which can be computed from s.
2 Called *elementary* situation in Ehrhard (2023a,b).
3 Because the fixpoint operators of the untyped λ-calculus are hardly compatible with differentiation, and, in the concrete models at hand, the morphisms of $\mathcal{L}_!(X, X)$ do not have fixpoints in general.
4 In our definition of $\frac{\partial M}{\partial x} \cdot N$, we are using implicitly the fact that our λ-calculus is equipped with a CBN operational semantics, translated in LL by the standard Girard translation. If our calculus were CBV, translated in LL through the "boring" Girard translation, the situation would be different.
5 See Ehrhard and Walch (2023a) where the theory is developed without this assumption.
6 In previous articles about coherent differentiation, we used the word *elementary* for this situation because of its fundamentally simple nature. As observed by a reviewer of Ehrhard and Walch (2023a), the word *representable* describes the situation in a more accurate way, so we prefer to move to this terminology.
7 It is not really necessary that all products exist, we only need 1 & 1 to exist, but this assumption is not very strong anyway.
8 More concretely, but more fuzzily also: in a Lafont resource category, any commutative comonoid is a !-coalgebra.
9 This exponential has not been formally introduced as far as we know but is easy to describe.
10 Remember that this verification is not really needed since we know that 5.9 ($\widetilde{\partial}$-chain), 5.10 ($\widetilde{\partial}$-local) and 5.11 ($\widetilde{\partial}$-add) hold by the simple fact that the exponential is free. We think it is nevertheless useful to have a better intuition on the morphism $\widetilde{\partial}$.
11 This means that $\hat{t}$ commutes with unions of directed families of cliques and with intersections of bounded nonempty finite families of cliques.
12 This is the categorical counterpart of the Girard's translation of intuitionistic logic into linear logic.
13 We hope to be able to improve and somehow simplify this system in the near future.
14 That is, without using powerful effects such as a global store.

References

Barr, M. (1979). $*$-*Autonomous Categories*, Number 752 in Lecture Notes in Mathematics. Springer-Verlag.
Blute, R. F., Cockett, J. R. B. and Seely, R. A. G. (2009). Cartesian differential categories. *Theory and Applications of Categories* **22** 622–672.
Boudes, P. (2011). Non-uniform (hyper/multi)coherence spaces. *Mathematical Structures in Computer Science* **21** (1) 1–40.
Bucciarelli, A. and Ehrhard, T. (2001). On phase semantics and denotational semantics: the exponentials. *Annals of Pure and Applied Logic* **109** (3) 205–241.
Cockett, J. R. B. and Cruttwell, G. S. H. (2014). Differential structure, tangent structure, and SDG. *Applied Categorical Structures* **22** (2) 331–417.
Crubillé, R., Ehrhard, T., Pagani, M. and Tasson, C. (2017). The free exponential modality of probabilistic coherence spaces. In: Esparza, J. and Murawski, A. S. (eds.) Foundations of Software Science and Computation Structures - 20th International Conference, FOSSACS (2017), Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2017, Uppsala, Sweden, April 22-29, 2017, Proceedings, vol. **10203** of Lecture Notes in Computer Science, 20–35.
Danos, V. and Ehrhard, T. (2011). Probabilistic coherence spaces as a model of higher-order probabilistic computation. *Information and Computation* **209** (6) 966–991.
Ehrhard, T. (2002). On Köthe sequence spaces and linear logic. *Mathematical Structures in Computer Science* **12** 579–623.
Ehrhard, T. (2005). Finiteness spaces. *Mathematical Structures in Computer Science* **15** (4) 615–646.

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press

Ehrhard, T. (2018). An introduction to differential linear logic: proof-nets, models and antiderivatives. *Mathematical Structures in Computer Science* **28** (7) 995–1060.
Ehrhard, T. (2022). Differentials and distances in probabilistic coherence spaces. *Logical Methods in Computer Science* **18** (3).
Ehrhard, T. (2023a). A coherent differential PCF. *Logical Methods in Computer Science* **19** (4).
Ehrhard, T. (2023b). Coherent differentiation. *Mathematical Structures in Computer Science* **33** 1–52.
Ehrhard, T. and Regnier, L. (2003). The differential lambda-calculus. *Theoretical Computer Science* **309** (1-3) 1–41.
Ehrhard, T. and Regnier, L. (2004). Differential interaction nets. In: Proceedings of WoLLIC'04, vol. **103** of Electronic Notes in Theoretical Computer Science, Elsevier Science, 35–74.
Ehrhard, T. and Regnier, L. (2006a). Böhm trees, Krivine machine and the Taylor expansion of ordinary lambda-terms. In: Beckmann, A., Berger, U., Löwe, B. and Tucker, J. V. (eds.) *Logical Approaches to Computational Barriers*, vol. **3988** of Lecture Notes in Computer Science, Springer-Verlag, 186–197.
Ehrhard, T. and Regnier, L. (2006b). Differential interaction nets. *Theoretical Computer Science* **364** (2) 166–195.
Ehrhard, T. and Regnier, L. (2008). Uniformity and the Taylor expansion of ordinary lambda-terms. *Theoretical Computer Science* **403** (2-3) 347–372.
Ehrhard, T., Tasson, C. and Pagani, M. (2014). Probabilistic coherence spaces are fully abstract for probabilistic PCF. In: Jagannathan, S. and Sewell, P. (ed.) *The 41st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL '14, San Diego, CA, USA, January 20-21, 2014*, San Diego, CA, USA, ACM, 309–320.
Ehrhard, T. and Walch, A. (2023a). Cartesian coherent differential categories. In: LICS, IEEE, 1–13.
Ehrhard, T. and Walch, A. (2025). Coherent taylor expansion as a bimonad. *Mathematical Structures in Computer Science* **35**. https://doi.org/10.1017/s0960129525000040
Girard, J.-Y. (1986). The system F of variable types, fifteen years later. *Theoretical Computer Science* **45** (2) 159–192.
Girard, J.-Y. (1987). Linear logic. *Theoretical Computer Science* **50** 1–102.
Girard, J.-Y. (2004). Between logic and quantic: a tract. In: Ehrhard, T., Girard, J.-Y., Ruet, P. and Scott, P. (eds.) *Linear Logic in Computer Science*, vol. **316** of London Mathematical Society Lecture Notes Series, Cambridge University Press, 346–381.
Hines, P. (2013). A categorical analogue of the monoid semiring construction. *Mathematical Structures in Computer Science* **23** (1) 55–94.
Kerjean, M. and Lemay, J.-S. P. (2023). Taylor expansion as a Monad in models of DiLL. In: LICS, IEEE, 1–13.
Kock, A. (2010). *Synthetic Geometry of Manifolds*, vol. **180** of Cambridge Tracts in Mathematics. Cambridge: Cambridge University Press.
Melliès, P.-A. (2009). Categorical semantics of linear logic. *Panoramas et Synthèses* **27** 1–196.
Plotkin, G. D. (1977). LCF considered as a programming language. *Theoretical Computer Science* **5** (3) 223–255.
Blute, R. F., Cockett, J. R. B., Lemay, J.-S. P., Seely, R. A. G. (2020). Differential categories revisited. *Applied Categorical Structures* **28** (2) 171–235.
Rosický, J. (1984). Abstract tangent functors. *Diagrammes* **12** JR1–JR11.

Cite this article: Ehrhard T (2026). From differential linear logic to coherent differentiation. *Mathematical Structures in Computer Science* **35**(e41), 1–49. https://doi.org/10.1017/S0960129525100212

https://doi.org/10.1017/S0960129525100212 Published online by Cambridge University Press