



HAL
open science

Cartesian Coherent Differential Categories

Thomas Ehrhard, Aymeric Walch

► **To cite this version:**

Thomas Ehrhard, Aymeric Walch. Cartesian Coherent Differential Categories. Thirty-Eighth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), Jun 2023, Boston (MA), United States. hal-04015264v2

HAL Id: hal-04015264

<https://hal.science/hal-04015264v2>

Submitted on 6 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cartesian Coherent Differential Categories

Thomas Ehrhard

Université Paris Cité, CNRS, Inria, IRIF, F-75013, Paris, France

Aymeric Walch

Université Paris Cité, CNRS, IRIF, F-75013, Paris, France

June 6, 2023

Abstract

We extend to general cartesian categories the idea of Coherent Differentiation recently introduced by Ehrhard in the setting of categorical models of Linear Logic. The first ingredient is a summability structure which induces a partial left-additive structure on the category. Additional functoriality and naturality assumptions on this summability structure implement a differential calculus which can also be presented in a formalism close to Blute, Cockett and Seely's cartesian differential categories. We show that a simple term language equipped with a natural notion of differentiation can easily be interpreted in such a category.

Contents

Introduction	2
1 Left summability structure	4
1.1 Left pre-summability structures	4
1.2 Left summability structures	6
1.3 Comparison with summability structures	7
2 Differential	8
2.1 Differential Structure	8
2.2 Linearity	9
2.3 The Differentiation Monad	11
3 Interpreting the axioms as properties of the derivative	12
4 Compatibility with the cartesian product	15
4.1 Cartesian product and summability structure	16
4.2 Cartesian product and differential structure	17
4.3 Partial derivatives	18
4.4 Generalization to arbitrary finite products	19
4.5 Multilinear morphism	20
5 Kleisli category of the exponential of a model of LL	24
5.1 Coherent differentiation in a linear setting	24
5.2 The example of probabilistic coherence spaces	27

6	Link with cartesian differential categories	29
6.1	Cartesian left additive categories	29
6.2	Cartesian differential categories	29
7	A first order coherent differential language	30
7.1	Terms	31
7.2	Semantics	33
7.3	Reduction	34

Introduction

This article is a long version of a paper, with the same title and by the same authors, accepted at the *ACM/IEEE Symposium on Logic in Computer Science 2023*. In particular, all the proofs which are missing in the conference version are provided in the present article.

Linear Logic (LL) and its models [1] strongly suggest that differentiation of proofs should be a natural operation extracting their best “local” linear approximation. Remember that for any E, F Banach spaces, $f : E \rightarrow F$ is differentiable at $x \in E$ if there is a neighborhood U of 0 in E and a linear and continuous function $\varphi : E \rightarrow F$ such that, for all $u \in U$

$$f(x + u) = f(x) + \varphi(u) + o(\|u\|). \quad (1)$$

When φ exists, it is unique and is denoted as $f'(x)$. When $f'(x)$ exists for all $x \in E$, the function $f' : E \rightarrow \mathcal{L}(E, F)$, where $\mathcal{L}(E, F)$ is the Banach space of linear and continuous functions $E \rightarrow F$, is called the *differential* of f . This function can itself admit a differential and so on. When all these iterated differentials exist one says that f is *smooth* and the n th derivative of f is a function $f^{(n)} : E \rightarrow \mathcal{L}_n(E, F)$ where $\mathcal{L}_n(E, F)$ is the space of n -linear symmetric functions $E^n \rightarrow F$. It can even happen that f is locally (or even globally) expressed using its iterated derivatives by means of the *Taylor Formula* $f(x + u) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x)(u, \dots, u)$; when this holds locally at any point x , f is said to be *analytic*.

Based on categorical models of LL where morphisms are analytic functions, the differential λ -calculus and differential LL provide a logical and syntactical account of differentiation. A program of type $A \Rightarrow B$ can be turned into a program of type $A \Rightarrow (A \multimap B)$. This provides a new approach of finite approximations of functions by a syntactical version of the Taylor Formula which has shown relevance in the study of the λ -calculus and of LL. Differentiation is deeply connected with *addition*, as it can already be seen in its definition Eq. (1). This connection also appears when writing the differential of $f : \mathbb{R}^n \rightarrow \mathbb{R}$ as a sum of *partial derivatives*:

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

and, of course, in the Taylor formula itself. For this reason, until recently, all categorical models of the differential λ -calculus and of differential LL [2, 3] were using categories where hom-sets have a structure of commutative monoid and both formalisms feature a formal and unrestricted addition operation on terms or proofs of the same type. The only available operational interpretation of such a sum being erratic choice, these formalisms are inherently non-deterministic.

Recently, the first author observed [4] that, in a setting where all coefficients are non-negative, differentiation survives to strong restrictions on the use of addition. Consider for instance a function $[0, 1] \rightarrow [0, 1]$ which is smooth on $[0, 1)$ and all of whose iterated

derivatives are everywhere ≥ 0 ¹. If $x, u \in [0, 1]$ are such that $x + u \in [0, 1]$ then $f(x) + f'(x)u \leq f(x + u) \in [0, 1]$ (this makes sense even if $f'(1) = \infty$, which can happen: take $f(x) = 1 - \sqrt{1 - x}$). So if S is the set of all such pairs (x, u) that we call *summable*, we can consider the function $\tilde{D}(f) : (x, u) \mapsto (f(x), f'(x)u)$ as a map $S \rightarrow S$. This basic observation is generalized in [5] to a wide range of categorical models \mathcal{L} of LL including coherence spaces, probabilistic coherence spaces *etc.* where hom-sets have only a *partially defined* addition. In these *summable categories*, S becomes an endofunctor $\mathcal{L} \rightarrow \mathcal{L}$ equipped with an additional structure which allows to define summability and (partial) sums in a very general way and turns out to induce a monad. Differentiation is then axiomatized as a distributive law between this monad (similar to the tangent bundle monad of a tangent category [6]) and the resource comonad $!_-$ of the LL structure of the category² \mathcal{L} . Indeed, this distributive law allows one to extend S to $\mathcal{L}_!$ the Kleisli category of $!_-$ and this extension $\tilde{D} : \mathcal{L}_! \rightarrow \mathcal{L}_!$ turns out to be a monad which has all the required properties of differentiation.

It is well known that $\mathcal{L}_!$ is a cartesian closed category, and it can be interesting to drift away from the LL structure of \mathcal{L} by only looking at the structure of its Kleisli category. This is what happened with differentiation. It was first axiomatized in a typical LL setting with additive categories and differential categories [2]. It was then carried to the setting of cartesian categories with left-additive categories and cartesian differential categories (CDC) [3]. Unsurprisingly, the Kleisli categories of the former provide instances of the latter, but cartesian differential categories cover a wider range of models. As mentioned in [5], differential categories can be seen as a special instance of summable categories equipped with differentiation (we will call those *coherent differential categories*) in which addition is unrestricted. Naturally, we can wonder if there is a notion of *cartesian coherent differential categories*, that arise as the Kleisli categories of coherent differential categories, and that generalize CDC to a partial setting.

We provide a positive answer to this question. We define coherent differentiation in an *arbitrary category*, whose morphisms are intuitively considered as smooth. So we start from a category \mathcal{C} equipped with a map³ $\tilde{D} : \text{Ob}(\mathcal{C}) \rightarrow \text{Ob}(\mathcal{C})$ given together with morphisms $\pi_{0,X}, \pi_{1,X}, \sigma_X \in \mathcal{C}(\tilde{D}(X), X)$ (for each $X \in \text{Obj}(\mathcal{C})$). The intuition is that $\tilde{D}(X)$ is the object of summable pairs of elements of X , that π_i are the obvious projections and that σ computes the sums. We assume π_0, π_1 to be jointly monic and this is sufficient to say when $f_0, f_1 \in \mathcal{C}(X, Y)$ are summable: this is when there is a necessarily unique $h \in \mathcal{C}(X, \tilde{D}(Y))$ such that $\pi_i \circ h = f_i$ and when this holds we set $f_0 + f_1 = \sigma \circ h$. Under suitable assumptions this very light structure suffices to equip hom-sets of \mathcal{C} with a structure of partial commutative monoid which is compatible with composition on the left⁴.

This structure is a convenient setting for differentiation: it suffices to furthermore equip \tilde{D} with a functorial action on morphisms wrt. which some morphisms (definable in terms of π_0, π_1, σ) are natural. This is the notion of *coherent differential category* whose axioms are in one-to-one correspondence with those of a CDC. Just as in tangent categories [6, 7], our functor \tilde{D} can be equipped with a monad structure. Contrarily to the additive framework of CDC our differentiation functor \tilde{D} is not defined in terms of the cartesian product so it is important to understand how it interacts with the cartesian product when available: this is formalized by the concept of *cartesian coherent differential category* (CCDC). This compatibility can be expressed in terms of a strength with which \tilde{D} can be equipped, turning it into a commutative monad. This induces a satisfactory theory of partial derivatives. We

¹This actually implies that f is analytic.

²Which by the way needs not be a fully-fledged LL model.

³It will become a functor and even a monad later.

⁴And not on the right in general since, intuitively, the morphisms of \mathcal{C} are not assumed to be linear.

provide a concrete example of such a category based on probabilistic coherence spaces and illustrate our formalism by interpreting a simple term language equipped with a notion of differentiation in a CCDC.

1 Left summability structure

We introduce in this section the notion of *left summability structure* in order to generalize the notion of summability structure introduced in [5] to a setting where morphisms are not necessarily additive.

1.1 Left pre-summability structures

Let \mathcal{C} be a category with objects $\mathbf{Obj}(\mathcal{C})$ and hom-set $\mathcal{C}(X, Y)$ for any $X, Y \in \mathbf{Obj}(\mathcal{C})$. We assume that any hom-set $\mathcal{C}(X, Y)$ contains a distinguished morphism $0^{X, Y}$ (usually X and Y are kept implicit) such that for any $f \in \mathcal{C}(Z, X)$, $0^{X, Y} \circ f = 0^{Z, Y}$.

Definition 1. A *summable pairing structure* on a category \mathcal{C} is a tuple $(\tilde{\mathbf{D}}, \pi_0, \pi_1, \sigma)$ where:

- $\tilde{\mathbf{D}} : \mathbf{Obj}(\mathcal{C}) \rightarrow \mathbf{Obj}(\mathcal{C})$ is a map (a functional class) on objects;
- $(\pi_{0, X})_{X \in \mathbf{Obj}(\mathcal{C})}$, $(\pi_{1, X})_{X \in \mathbf{Obj}(\mathcal{C})}$ and $(\sigma_X)_{X \in \mathbf{Obj}(\mathcal{C})}$ are families of morphisms in $\mathcal{C}(\tilde{\mathbf{D}}X, X)$. The object X will usually be kept implicit;
- π_0 and π_1 are jointly monic: for any $f, g \in \mathcal{C}(Y, \tilde{\mathbf{D}}X)$, if $\pi_0 \circ f = \pi_0 \circ g$ and $\pi_1 \circ f = \pi_1 \circ g$ then $f = g$.

We assume in what follows that \mathcal{C} is equipped with a summable pairing structure $(\tilde{\mathbf{D}}, \pi_0, \pi_1, \sigma)$.

Definition 2. Two morphisms $f_0, f_1 \in \mathcal{C}(X, Y)$ are said to be *summable* if there exists $h \in \mathcal{C}(X, \tilde{\mathbf{D}}Y)$ such that $\pi_i \circ h = f_i$. The joint monicity of the π_i 's ensures that when h exists, it is unique. We set $\langle\langle f_0, f_1 \rangle\rangle := h$, and we call it the *witness* of the sum. By definition, $\pi_i \circ \langle\langle f_0, f_1 \rangle\rangle = f_i$. Then we set $f_0 + f_1 := \sigma \circ \langle\langle f_1, f_2 \rangle\rangle$.

Remark 1. A more standard approach to notations would be to write π_1 and π_2 instead of π_0 and π_1 . The reason we proceed that way is that [Equation \(1\)](#) will be formalized in our setting with the use of a pair $\langle\langle f(x), f'(x) \cdot u \rangle\rangle$. That is, the left element of this pair is of order 0, and the right element is of order 1.

Notations 1. We write $f_0 \boxplus f_1$ for the property that f_0 and f_1 are summable. We say that an algebraic expression containing binary sums is *well defined* if each pair of morphisms involved in these sums are summable. For example, $(f_0 + f_1) + f_2$ is well defined if $f_0 \boxplus f_1$ and $(f_0 + f_1) \boxplus f_2$.

Proposition 1. *The morphism π_0 and π_1 are summable with witness $\langle\langle \pi_0, \pi_1 \rangle\rangle = \text{id}$ and sum $\pi_0 + \pi_1 = \sigma$.*

Proof. $\pi_i \circ \text{id} = \pi_i$ so by definition, $\pi_0 \boxplus \pi_1$ with witness id and sum $\sigma \circ \text{id} = \sigma$. □

Proposition 2 (Left compatibility of sum). *For any $f_0, f_1 \in \mathcal{C}(Y, Z)$ and $g \in \mathcal{C}(X, Y)$, if $f_0 \boxplus f_1$, then $(f_0 \circ g) \boxplus (f_1 \circ g)$ with witness $\langle\langle f_0 \circ g, f_1 \circ g \rangle\rangle = \langle\langle f_0, f_1 \rangle\rangle \circ g$. Moreover, $(f_0 \circ g) + (f_1 \circ g) = (f_0 + f_1) \circ g$.*

Proof. Let $w = \langle\langle f_0, f_1 \rangle\rangle \circ g$. Then $\pi_i \circ w = f_i \circ g$ so w is a witness for the summability of $f_0 \circ g$ and $f_1 \circ g$. And $f_0 \circ g + f_1 \circ g := \sigma \circ w = (f_0 + f_1) \circ g$. \square

An important class of morphisms is that of additive morphisms, for which addition is compatible with composition on the right.

Definition 3. A morphism $h \in \mathcal{C}(Y, Z)$ is *additive* if $h \circ 0 = 0$ and if for any $f_0, f_1 \in \mathcal{C}(X, Y)$, if $f_0 \boxplus f_1$ then $h \circ f_0 \boxplus h \circ f_1$ and $h \circ (f_0 + f_1) = h \circ f_0 + h \circ f_1$. Note that id is additive and that the composition of two additive morphisms is an additive morphism.

Proposition 3. A morphism h such that $h \circ 0 = 0$ is additive if and only if $h \circ \pi_0 \boxplus h \circ \pi_1$ with sum $h \circ \sigma$.

Proof. For the forward implication, recall that $\pi_0 \boxplus \pi_1$ with sum σ . Thus by additivity of h , $h \circ \pi_0 \boxplus h \circ \pi_1$ with sum $h \circ \sigma$. For the reverse implication, assume that $f_0 \boxplus f_1$. Since $h \circ \pi_0 \boxplus h \circ \pi_1$, **Proposition 2** ensures that $h \circ f_0 = h \circ \pi_0 \circ \langle\langle f_0, f_1 \rangle\rangle$ and $h \circ f_1 = h \circ \pi_1 \circ \langle\langle f_0, f_1 \rangle\rangle$ are summable, with sum $(h \circ \pi_0 + h \circ \pi_1) \circ \langle\langle f_0, f_1 \rangle\rangle = h \circ \sigma \circ \langle\langle f_0, f_1 \rangle\rangle = h \circ (f_0 + f_1)$. \square

Definition 4. The summable pairing structure $(\tilde{D}, \pi_0, \pi_1, \sigma)$ is a *left pre-summability structure* if π_0, π_1 and σ are additive.

The additivity of the projections implies that the sum behaves well with respect to the operation $\langle\langle _, _ \rangle\rangle$ itself.

Proposition 4. Assume that π_0 and π_1 are additive. Then for any $f_0, f_1, g_0, g_1 \in \mathcal{C}(X, Y)$, if $f_0 \boxplus f_1, g_0 \boxplus g_1$ and $\langle\langle f_0, f_1 \rangle\rangle \boxplus \langle\langle g_0, g_1 \rangle\rangle$, then $f_0 \boxplus g_0, f_1 \boxplus g_1, (f_0 + g_0) \boxplus (f_1 + g_1)$ and $\langle\langle f_0, f_1 \rangle\rangle + \langle\langle g_0, g_1 \rangle\rangle = \langle\langle f_0 + g_0, f_1 + g_1 \rangle\rangle$.

Proof. By additivity of π_i , $\pi_i \circ \langle\langle f_0, f_1 \rangle\rangle = f_i$ and $\pi_i \circ \langle\langle g_0, g_1 \rangle\rangle = g_i$ are summable with sum $f_i + g_i = \pi_i \circ (\langle\langle f_0, f_1 \rangle\rangle + \langle\langle g_0, g_1 \rangle\rangle)$. Since $\pi_0 \boxplus \pi_1$ this entails by **Proposition 2** that $f_0 + g_0, f_1 + g_1$ are summable with witness $\langle\langle f_0, f_1 \rangle\rangle + \langle\langle g_0, g_1 \rangle\rangle$. \square

The additivity of σ implies that whenever $\langle\langle f_0, f_1 \rangle\rangle \boxplus \langle\langle g_0, g_1 \rangle\rangle$, one has $\sigma \circ \langle\langle f_0, f_1 \rangle\rangle \boxplus \sigma \circ \langle\langle g_0, g_1 \rangle\rangle$ and

$$\sigma \circ (\langle\langle f_0, f_1 \rangle\rangle + \langle\langle g_0, g_1 \rangle\rangle) = (\sigma \circ \langle\langle f_0, f_1 \rangle\rangle) + (\sigma \circ \langle\langle g_0, g_1 \rangle\rangle)$$

Assuming the additivity of the projections, the additivity of σ implies that whenever $\langle\langle f_0, f_1 \rangle\rangle, \langle\langle g_0, g_1 \rangle\rangle$ exists, the two sums below are well defined (see **Notations 1**) and

$$(f_0 + g_0) + (f_1 + g_1) = (f_0 + f_1) + (g_0 + g_1). \quad (2)$$

Proposition 5. The morphisms 0 and 0 are summable with witness 0 and sum 0 . In particular, 0 is additive.

Proof. On the one hand, $\pi_i \circ 0 = 0$ by additivity of π_i , so $0 \boxplus 0$ with witness 0 . On the other hand, $\sigma \circ 0 = 0$ by additivity of σ so $0 + 0 = 0$. In particular, 0 is additive thanks to **Proposition 3** because $0 \circ \pi_0 = 0$ and $0 \circ \pi_1 = 0$ are summable with witness 0 and sum $0 = 0 \circ \sigma$. \square

1.2 Left summability structures

We consider a category \mathcal{C} equipped with a left pre-summability structure $(\tilde{\mathcal{D}}, \pi_0, \pi_1, \sigma)$. The goal of this section is to make $(\mathcal{C}(X, Y), +, 0)$ a *partial* commutative monoid. Similar structures appear in [8] or more recently in [9], in a setting where sums can be infinitary. Our partial monoids have only finite sums⁵. More crucially, the categorical notion of summability defined above is essential for us whereas it is not categorically formalized in these works.

Definition 5. The left pre-summability structure is *commutative* if for any object X , $\pi_1, \pi_0 \in \mathcal{C}(\tilde{\mathcal{D}}X, X)$ are summable with sum σ . Then we set $\gamma = \langle\langle \pi_1, \pi_0 \rangle\rangle \in \mathcal{C}(\tilde{\mathcal{D}}X, \tilde{\mathcal{D}}X)$ so that $\pi_i \circ \gamma = \pi_{1-i}$. This property is called ($\tilde{\mathcal{D}}$ -com).

Proposition 6 (Commutativity). *The left pre-summability structure is commutative if and only if for any $f_0, f_1 \in \mathcal{C}(X, Y)$, if $f_0 \boxplus f_1$ then $f_1 \boxplus f_0$ and $f_0 + f_1 = f_1 + f_0$.*

Proof. For the direct implication, assume that $f_0 \boxplus f_1$. Then $\pi_i \circ \gamma \circ \langle\langle f_0, f_1 \rangle\rangle = \pi_{1-i} \circ \langle\langle f_0, f_1 \rangle\rangle = f_{1-i}$ so $f_1 \boxplus f_0$ with witness $\gamma \circ \langle\langle f_0, f_1 \rangle\rangle$. Furthermore, $f_1 + f_0 = \sigma \circ \gamma \circ \langle\langle f_0, f_1 \rangle\rangle = \sigma \circ \langle\langle f_0, f_1 \rangle\rangle = f_0 + f_1$. Conversely, $\pi_0 \boxplus \pi_1$ so by commutativity $\pi_1 \boxplus \pi_0$ and $\pi_1 + \pi_0 = \pi_0 + \pi_1 = \sigma$. \square

Definition 6. The left pre-summability structure *has 0 as a neutral element* if for any object X , $\text{id}_X \boxplus 0$ and $0 \boxplus \text{id}_X$ with sums equal to id_X . We call this property ($\tilde{\mathcal{D}}$ -zero). We define $\iota_0, \iota_1 \in \mathcal{C}(X, \tilde{\mathcal{D}}X)$ as $\iota_0 := \langle\langle \text{id}_X, 0 \rangle\rangle$ and $\iota_1 := \langle\langle 0, \text{id}_X \rangle\rangle$.

Proposition 7 (Neutrality of 0). *The left pre-summability structure has 0 as a neutral element if and only if for any morphism $f \in \mathcal{C}(X, Y)$, $0 \boxplus f$, $f \boxplus 0$ and $f + 0 = 0 + f = f$.*

Proof. By definition of ι_0 , $\pi_0 \circ \iota_0 \circ f = \text{id} \circ f = f$ and $\pi_1 \circ \iota_0 \circ f = 0 \circ f = 0$. So $f \boxplus 0$ with witness $\iota_0 \circ f$ and $f + 0 = \sigma \circ \iota_0 \circ f = \text{id} \circ f = f$. We do the same for $0 + f$ with ι_1 . Conversely, we apply the neutrality of 0 on id to get that $\text{id} \boxplus 0$ and $0 \boxplus \text{id}$, with sum id . \square

Associativity is not that straightforward, as there are two possible notions. The situation is similar in the infinitary setting of [9] with the distinction between Weak Partition Associativity and Partition Associativity.

Definition 7 (Weak Associativity). The operation $+$ is called *weakly associative* if whenever $(f_0 + f_1) + f_2$ and $f_0 + (f_1 + f_2)$ are well defined (recall **Notations 1**), we have $(f_0 + f_1) + f_2 = f_0 + (f_1 + f_2)$.

Definition 8 (Associativity). The operation $+$ is called *associative* if whenever $(f_0 + f_1) + f_2$ or $f_0 + (f_1 + f_2)$ is well defined, the other expression is also well defined and $(f_0 + f_1) + f_2 = f_0 + (f_1 + f_2)$.

We need to work in a partial setting in which addition is associative: this is required for instance in **Section 2.1** to define $\theta = \langle\langle \pi_0 \circ \pi_0, \pi_0 \circ \pi_1 + \pi_1 \circ \pi_0 \rangle\rangle$. This associativity seems related to a kind of positivity of morphisms.

Example 1. Let $x, y \in [-1, 1]$ be summable when $|x| + |y| \leq 1$, with $x + y$ as sum. Then $+$ is weakly associative, but is not associative. Indeed, take $x_0 = -\frac{1}{2}, x_1 = \frac{1}{2}, y_1 = 1$. Then $(x_0 + x_1) + y_1$ is defined, but $x_0 + (x_1 + y_1)$ is not since $|x_1| + |y_1| = \frac{3}{2} > 1$. However, the same definition on $[0, 1]$ yields an associative operation.

⁵Although the extension of the finite sum to an infinitary operations will have to be considered when dealing with fixpoints.

Recall from [Equation \(2\)](#) that whenever $\langle\langle f_0, f_1 \rangle\rangle, \langle\langle g_0, g_1 \rangle\rangle$ exists, the expressions $(f_0 + g_0) + (f_1 + g_1)$ and $(f_0 + f_1) + (g_0 + g_1)$ are well defined and equal. Taking $g_0 = 0$ and assuming [\(\$\tilde{D}\$ -zero\)](#), this means that whenever $\langle\langle f_0, f_1 \rangle\rangle, \langle\langle 0, g_1 \rangle\rangle$ exists, $(f_0 + f_1) + g_1$ and $f_0 + (f_1 + g_1)$ are well defined and equal. Taking $f_1 = 0$ and assuming [\(\$\tilde{D}\$ -zero\)](#), this means that whenever $\langle\langle f_0, 0 \rangle\rangle, \langle\langle g_0, g_1 \rangle\rangle$ exist, $f_0 + (g_0 + g_1)$ and $(f_0 + g_0) + g_1$ are well defined and equal. Thus associativity holds if [\(\$\tilde{D}\$ -zero\)](#) holds and if whenever $(f_0 + f_1) + g_1$ is defined (respectively $f_0 + (g_0 + g_1)$ is defined), then $\langle\langle f_0, f_1 \rangle\rangle, \langle\langle 0, g_1 \rangle\rangle$ exists (respectively $\langle\langle f_0, 0 \rangle\rangle, \langle\langle g_0, g_1 \rangle\rangle$ exists). This shows that associativity follows from the following axiom.

Definition 9. The left pre-summability structure *admits witnesses* if for any $f, g \in \mathcal{C}(Y, \tilde{D}X)$, if $\sigma \circ f \boxplus \sigma \circ g$ then $f \boxplus g$. We call this property [\(\$\tilde{D}\$ -witness\)](#).

Theorem 1. *The properties [\(\$\tilde{D}\$ -zero\)](#), [\(\$\tilde{D}\$ -com\)](#) and [\(\$\tilde{D}\$ -witness\)](#) give to $\mathcal{C}(X, Y)$ the structure of a partial commutative monoid for any objects X, Y . That is, for any $f, f_0, f_1, f_2 \in \mathcal{C}(X, Y)$:*

- $f \boxplus 0, 0 \boxplus f$ and $0 + f = f + 0 = f$;
- If $f_0 \boxplus f_1$ then $f_1 \boxplus f_0$ and $f_0 + f_1 = f_1 + f_0$;
- If $(f_0 + f_1) + f_2$ or $f_0 + (f_1 + f_2)$ is defined, then both are defined and $(f_0 + f_1) + f_2 = f_0 + (f_1 + f_2)$.

One can define inductively from this binary sum a notion of arbitrary finite sum. The empty family is always summable with sum 0. The family $(f_i)_{i \in I}$ for $I \neq \emptyset$ is summable if $\exists i_0 \in I$ such that $(f_i)_{i \in I/\{i_0\}}$ is summable and if $(\sum_{i \in I/\{i_0\}} f_i) \boxplus f_{i_0}$. Then we set $\sum_{i \in I} f_i := \sum_{i \in I/\{i_0\}} f_i + f_{i_0}$. [Theorem 2](#) shown in [5] ensures that the choice of order for the sum is irrelevant.

Theorem 2. *A family $(f_i)_{i \in I}$ is summable if and only if for all partition⁶ I_1, \dots, I_n of I , we have that for all $j \in \llbracket 1, n \rrbracket := \{1, \dots, n\}$, $(f_i)_{i \in I_j}$ is summable and $(\sum_{i \in I_j} f_i)_{j \in \llbracket 1, n \rrbracket}$ is summable. Moreover, $\sum_{i \in I} f_i = \sum_{j \in \llbracket 1, n \rrbracket} \sum_{i \in I_j} f_i$.*

Definition 10. A *left summability structure* is a left pre-summability structure $(\tilde{D}, \pi_0, \pi_1, \sigma)$ such that [\(\$\tilde{D}\$ -zero\)](#), [\(\$\tilde{D}\$ -com\)](#) and [\(\$\tilde{D}\$ -witness\)](#) hold.

1.3 Comparison with summability structures

In the LL setting of [5], the first author introduced a notion of pre-summability structure $(S, \pi_0, \pi_1, \sigma)$ as a summable pairing structure (recall [Definition 1](#)) where S is a functor for which π_0, π_1, σ are natural transformations.

Theorem 3. *The following are equivalent*

- $(S, \pi_0, \pi_1, \sigma)$ is a left pre-summability structure and every morphism is additive;
- $(S, \pi_0, \pi_1, \sigma)$ is a pre-summability structure [5].

Remember that in [5], the underlying category \mathcal{C} is assumed to be enriched over the monoidal category of pointed sets, the distinguished element of $\mathcal{C}(X, Y)$ being denoted 0. In particular $f \circ 0 = 0$ always holds.

⁶Where we admit that some I_j s can be empty.

Proof. Let $(\mathbb{S}, \pi_0, \pi_1, \sigma)$ be a left pre-summability structure in which every morphism is additive. By [Proposition 3](#), for any $f \in \mathcal{C}(X, Y)$ we can define $\mathbb{S}f := \langle\langle f \circ \pi_0, f \circ \pi_1 \rangle\rangle$ and the following equations hold: $\pi_i \circ \mathbb{S}f = f \circ \pi_i$, $\sigma \circ \mathbb{S}f = f \circ \sigma$. Furthermore, \mathbb{S} is a functor: $\pi_i \circ \text{Sid} = \text{id} \circ \pi_i = \pi_i \circ \text{id}$ and $\pi_i \circ \mathbb{S}f \circ \mathbb{S}g = f \circ \pi_i \circ \mathbb{S}g = f \circ g \circ \pi_i = \pi_i \circ \mathbb{S}(f \circ g)$. Thus, by joint monicity of the π_i , $\text{Sid} = \text{id}$ and $\mathbb{S}(f \circ g) = \mathbb{S}f \circ \mathbb{S}g$. Then the equations $\pi_i \circ \mathbb{S}f = f \circ \pi_i$ and $\sigma \circ \mathbb{S}f = f \circ \sigma$ introduced above correspond to the naturality of π_0, π_1 and σ .

Conversely, let $(\mathbb{S}, \pi_0, \pi_1, \sigma)$ be a pre-summability structure in the sense of [\[5\]](#). The naturality of π_0 and π_1 ensures that for any f , $f \circ \pi_0 \boxplus f \circ \pi_1$ with witness $\mathbb{S}f$. The naturality of σ ensures that the sum of those two morphisms is $\sigma \circ \mathbb{S}f = f \circ \sigma$. Finally, $f \circ 0 = 0$ by assumption. So every morphism is additive by [Proposition 3](#). In particular, π_0, π_1 and σ are additive, so $(\mathbb{S}, \pi_0, \pi_1, \sigma)$ is a left pre-summability structure. \square

Corollary 1. *The summability structures of [\[5\]](#) are the left summability structures where all morphisms are additive.*

2 Differential

2.1 Differential Structure

Recall from [Equation \(1\)](#) the main idea of the differential calculus. We generalize it to a partial additive setting: f is differentiable at x if for any u , if $x \boxplus u$ then $f'(x) \cdot u$ is defined, $f(x) \boxplus f'(x) \cdot u$ and, intuitively, $f(x + u) \simeq f(x) + f'(x) \cdot u$. Hence the differential of f can be seen as a function $\tilde{\mathbb{D}}f$ that maps a pair of two summable elements $\langle\langle x, u \rangle\rangle$ to a pair of two summable elements $\tilde{\mathbb{D}}f(x, u) = \langle\langle f(x), f'(x) \cdot u \rangle\rangle$.

Definition 11. A *pre-differential structure* is a left summability structure $(\tilde{\mathbb{D}}, \pi_0, \pi_1, \sigma)$ together with, for each $X, Y \in \text{Obj}(\mathcal{C})$, an operator $\mathcal{C}(X, Y) \rightarrow \mathcal{C}(\tilde{\mathbb{D}}X, \tilde{\mathbb{D}}Y)$, also denoted as $\tilde{\mathbb{D}}$, and such that $\pi_0 \circ \tilde{\mathbb{D}}f = f \circ \pi_0$. We define the *differential* of f as $\mathbb{D}f := \pi_1 \circ \tilde{\mathbb{D}}f \in \mathcal{C}(\tilde{\mathbb{D}}X, Y)$. By our assumptions $\tilde{\mathbb{D}}f = \langle\langle f \circ \pi_0, \mathbb{D}f \rangle\rangle$.

At this point we do not assume $\tilde{\mathbb{D}}$ to be a functor, this will be the Chain Rule. Then the equation $\pi_0 \circ \tilde{\mathbb{D}}f = f \circ \pi_0$ will be the naturality of π_0 . We can also introduce three families of morphisms θ , \mathbb{l} and \mathbb{c} whose naturality will correspond to some axioms of differentiation. This is very similar to what happens in tangent categories [\[7\]](#), the difference being the structure of the functor $\tilde{\mathbb{D}}$ itself⁷.

The additivity of σ ensures that $\sigma \circ \pi_0 \boxplus \sigma \circ \pi_1$. That is, $(\pi_0 \circ \pi_0 + \pi_1 \circ \pi_0) \boxplus (\pi_0 \circ \pi_1 + \pi_1 \circ \pi_1)$. By associativity, this implies that $((\pi_0 \circ \pi_0 + \pi_1 \circ \pi_0) + \pi_0 \circ \pi_1) + \pi_1 \circ \pi_1$ is well defined, so $(\pi_0 \circ \pi_0 + \pi_1 \circ \pi_0) + \pi_0 \circ \pi_1$ is well defined. By associativity again, $\pi_0 \circ \pi_0 + (\pi_1 \circ \pi_0 + \pi_0 \circ \pi_1)$ is well defined, so [Definition 12](#) below makes sense.

Definition 12. For any object X , define $\theta \in \mathcal{C}(\tilde{\mathbb{D}}^2 X, \tilde{\mathbb{D}}X)$ as $\theta := \langle\langle \pi_0 \circ \pi_0, \pi_1 \circ \pi_0 + \pi_0 \circ \pi_1 \rangle\rangle$.

By ($\tilde{\mathbb{D}}$ -zero), $(\pi_0 + 0) \boxplus (0 + \pi_1)$ so by ($\tilde{\mathbb{D}}$ -witness) $\langle\langle \pi_0, 0 \rangle\rangle \boxplus \langle\langle 0, \pi_1 \rangle\rangle$. So [Definition 13](#) below makes sense.

Definition 13. For any object X , define $\mathbb{l} \in \mathcal{C}(\tilde{\mathbb{D}}X, \tilde{\mathbb{D}}^2 X)$ as $\mathbb{l} := \langle\langle \langle\langle \pi_0, 0 \rangle\rangle, \langle\langle 0, \pi_1 \rangle\rangle \rangle\rangle$.

⁷There might be a way to combine tangent categories and coherent differentiation in one notion allowing to axiomatize objects similar to manifolds where the tangent spaces have an addition of vectors which is only partially defined. The first step should be to develop convincing concrete examples of such objects, which might be related to the semantics of Type Theory.

By [Proposition 2](#) (left compatibility) $\pi_0 \circ (\pi_0 + \pi_1) \boxplus \pi_1 \circ (\pi_0 + \pi_1)$. By additivity of π_0 and π_1 , it means that $(\pi_0 \circ \pi_0 + \pi_0 \circ \pi_1) \boxplus (\pi_1 \circ \pi_0 + \pi_1 \circ \pi_1)$. So by ([D-tilde-witness](#)), $\langle\langle \pi_0 \circ \pi_0, \pi_0 \circ \pi_1 \rangle\rangle \boxplus \langle\langle \pi_1 \circ \pi_0, \pi_1 \circ \pi_1 \rangle\rangle$ and [Definition 14](#) below makes sense.

Definition 14. For any object X , we can define $c \in \mathcal{C}(\tilde{D}^2 X, \tilde{D}^2 X)$ as $c := \langle\langle \pi_0 \circ \pi_0, \pi_0 \circ \pi_1 \rangle\rangle, \langle\langle \pi_1 \circ \pi_0, \pi_1 \circ \pi_1 \rangle\rangle$.

It is probably easier to understand those morphisms by how they operate on witnesses. This corresponds to [Proposition 8](#) below. The proof is a straightforward computation using the joint monicity of π_0 and π_1 .

Proposition 8. For any $x, u, v, w \in \mathcal{C}(U, X)$ such that $\langle\langle x, u \rangle\rangle, \langle\langle v, w \rangle\rangle$ is defined,

$$\begin{aligned} \theta \circ \langle\langle x, u \rangle\rangle, \langle\langle v, w \rangle\rangle &= \langle\langle x, u + v \rangle\rangle \\ c \circ \langle\langle x, u \rangle\rangle, \langle\langle v, w \rangle\rangle &= \langle\langle x, v \rangle\rangle, \langle\langle u, w \rangle\rangle \\ l \circ \langle\langle x, u \rangle\rangle &= \langle\langle x, 0 \rangle\rangle, \langle\langle 0, u \rangle\rangle \end{aligned}$$

Definition 15. A *differential structure* is a pre-differential structure $(\tilde{D}, \pi_0, \pi_1, \sigma)$ where the following axioms hold, using the associated notation Df introduced in [Definition 11](#):

- (1) ([Dproj-lin](#)) $D\pi_0 = \pi_0 \circ \pi_1, D\pi_1 = \pi_1 \circ \pi_1$;
- (2) ([Dsum-lin](#)) $D\sigma = \sigma \circ \pi_1, D0 = 0$;
- (3) ([D-chain](#)) \tilde{D} is a functor (Chain Rule);
- (4) ([D-add](#)) ι_0, θ are natural transformations (additivity of the derivative);
- (5) ([D-lin](#)) l is a natural transformation (linearity of the derivatives);
- (6) ([D-Schwarz](#)) c is a natural transformation (Schwarz Rule).

A *coherent differential category* is a category \mathcal{C} equipped with a differential structure.

The axiom ([Dproj-lin](#)) corresponds to an important structural property of \tilde{D} with regard to $\langle\langle _, _ \rangle\rangle$. The axiom ([Dsum-lin](#)) corresponds to the additivity of the derivative operator, that is, $(f + g)' = f' + g'$. The axiom ([D-chain](#)) corresponds to the Chain Rule of the differential calculus. The axiom ([D-add](#)) says that $u \mapsto f'(x) \cdot u$ is additive. The axiom ([D-lin](#)) says that $u \mapsto f'(x) \cdot u$ is not only additive, but also equal to its own derivative in 0. It is shown in Prop. 4.2 of [7] (in the left-additive setting of cartesian differential categories) that it implies that $u \mapsto f'(x) \cdot u$ is equal to its own derivative in any points. The same reasoning can be generalized to our setting, but it would require too much technical development to be developed in this paper. Finally, the axiom ([D-Schwarz](#)) corresponds to the Schwarz Rule, that is, the second derivative $f''(x)$ (a bilinear map) is symmetric. An account of these axioms as properties of D can be found in [Section 3](#) and might help the reader understand the ideas mentioned above.

2.2 Linearity

For the rest of this section, \mathcal{C} is only assumed to be equipped with a pre-differential structure. Any use of an axiom of coherent differential categories will be made explicit.

Definition 16 ([D-tilde-linearity](#)). A morphism $f \in \mathcal{C}(X, Y)$ is *D-tilde-linear* if the following diagrams commute.

$$\begin{array}{ccc}
\tilde{D}X & \xrightarrow{\tilde{D}f} & \tilde{D}Y & & \tilde{D}X & \xrightarrow{\tilde{D}f} & \tilde{D}Y & & X & & \\
\pi_1 \downarrow & & \downarrow \pi_1 & & \sigma \downarrow & & \downarrow \sigma & & 0 \downarrow & \searrow 0 & \\
X & \xrightarrow{f} & Y & & X & \xrightarrow{f} & Y & & X & \xrightarrow{f} & Y
\end{array}$$

Remark 2. The first diagram can also be written as $D(f) = f \circ \pi_1$ and means that $\tilde{D}f = \langle\langle f \circ \pi_0, f \circ \pi_1 \rangle\rangle$.

Proposition 9. *A morphism f is \tilde{D} -linear if and only if it is additive and $Df = f \circ \pi_1$ (that is, $\tilde{D}f = \langle\langle f \circ \pi_0, f \circ \pi_1 \rangle\rangle$).*

Proof. Assume that f is \tilde{D} -linear. Then $f \circ 0 = 0$ and, by [Remark 2](#), $f \circ \pi_0 \boxplus f \circ \pi_1$ of witness $\tilde{D}f$. Thus $f \circ \pi_0 + f \circ \pi_1 := \sigma \circ \tilde{D}f = f \circ \sigma$ by assumption. So f is additive by [Proposition 3](#), and $Df = f \circ \pi_1$ by assumption. Conversely, only the second diagram is not part of the assumptions.

$$\begin{aligned}
\sigma \circ \tilde{D}f &= (\pi_0 + \pi_1) \circ \tilde{D}f \\
&= \pi_0 \circ \tilde{D}f + \pi_1 \circ \tilde{D}f \quad \text{by [Proposition 2](#)} \\
&= f \circ \pi_0 + f \circ \pi_1 \quad \text{by assumption} \\
&= f \circ (\pi_0 + \pi_1) = f \circ \sigma \quad \text{by additivity of } f
\end{aligned}$$

Thus f is \tilde{D} -linear. □

Thus \tilde{D} -linear morphisms are in particular additive. As we will see, our notion of additive and \tilde{D} -linear morphisms ultimately coincides with that of [\[3\]](#), so this distinction between additivity and linearity is as relevant as it is in their setting.

Corollary 2. *([Dproj-lin](#)) is equivalent to the linearity of π_0 and π_1 . ([Dsum-lin](#)) is equivalent to the linearity of σ and 0 .*

Thus \tilde{D} -linear morphisms are special instances of additive ones. Our notion of additive and \tilde{D} -linear morphisms ultimately coincides with the one of [\[3\]](#) thanks to [Proposition 10](#) below, so this distinction between additivity and linearity is as relevant as it is in their setting.

Proposition 10. *Assuming ([Dproj-lin](#)), ([D-chain](#)) and ([D-add](#)), any morphism $h \in \mathcal{C}(X, Y)$ such that $Dh = h \circ \pi_1$ is additive, hence \tilde{D} -linear.*

Proof. The proof relies on [Propositions 18 and 19](#) of [Section 3](#). If $h = Dh \circ \pi_1$, then for any $g \in \mathcal{C}(Z, X)$, $h \circ g = h \circ \pi_1 \circ \langle\langle 0, g \rangle\rangle = Dh \circ \langle\langle 0, g \rangle\rangle$. Thus, $h \circ 0 = Dh \circ \langle\langle 0, 0 \rangle\rangle = 0$ by [Proposition 18](#), and $h \circ (f_0 + f_1) = Dh \circ \langle\langle 0, f_0 + f_1 \rangle\rangle = Dh \circ \langle\langle 0, f_0 \rangle\rangle + Dh \circ \langle\langle 0, f_1 \rangle\rangle = h \circ f_0 + h \circ f_1$ by [Proposition 19](#) again. Thus, h is additive, so h is \tilde{D} -linear by [Proposition 9](#). □

Thanks to ([D-chain](#)), ([Dproj-lin](#)) and ([Dsum-lin](#)), we can show that linear morphisms are closed under composition, witnesses and sum.

Proposition 11. *Assuming ([D-chain](#)), \tilde{D} -linear morphisms are closed under composition and inverses.*

Proof. Easy verification using the functoriality of \tilde{D} . □

Proposition 12 (\tilde{D} -linearity and pairing). *Assume ([D-chain](#)) and ([Dproj-lin](#)). Assume that $h_0, h_1 \in \mathcal{C}(X, Y)$ are summable and both \tilde{D} -linear. Then $\langle\langle h_0, h_1 \rangle\rangle$ is \tilde{D} -linear.*

Proof. Let us do the diagram involving σ , the other two being very similar. By joint monicity of the π_i 's, it suffices to solve the diagram chase below for $i = 0, 1$.

$$\begin{array}{ccccc}
\tilde{D}X & \xrightarrow{\tilde{D}\langle\langle h_0, h_1 \rangle\rangle} & \tilde{D}^2Y & \xrightarrow{\sigma} & \tilde{D}Y \\
\downarrow \sigma & \searrow (a) \tilde{D}h_i & \tilde{D}\pi_i \downarrow & (b) & \downarrow \pi_i \\
& & \tilde{D}Y & \xrightarrow{\sigma} & Y \\
& & \uparrow h_i & & \uparrow \pi_i \\
X & \xrightarrow{\langle\langle h_0, h_1 \rangle\rangle} & \tilde{D}Y & &
\end{array}$$

(a) is a consequence of **(D-chain)**, (b) is a consequence of **(Dproj-lin)** and (c) is the \tilde{D} -linearity of h_i . \square

Proposition 13. *Assuming **(D-chain)** and **(Dproj-lin)**, σ is \tilde{D} -linear if and only if for all $h_0, h_1 \in \mathcal{C}(X, Y)$ summable and both \tilde{D} -linear, $h_0 + h_1$ is \tilde{D} -linear.*

Proof. Assume that h_0, h_1 are \tilde{D} -linear. By **Proposition 12**, $\langle\langle h_0, h_1 \rangle\rangle$ is \tilde{D} -linear so $h_0 + h_1 = \sigma \circ \langle\langle h_0, h_1 \rangle\rangle$ is \tilde{D} -linear (\tilde{D} -linearity is closed under composition). Conversely, $\sigma = \pi_0 + \pi_1$ and π_0, π_1 are \tilde{D} -linear so σ is \tilde{D} -linear. \square

Corollary 3. *Assuming **(Dproj-lin)**, **(Dsum-lin)** and **(D-chain)**, $\iota_0, \iota_1, c, l, \theta$ are all \tilde{D} -linear.*

Proof. All these morphisms are obtained through pairing, sums and composition of \tilde{D} -linear maps. \square

On a side note, by **Remark 2** the \tilde{D} -linearity of π_i means that $\tilde{D}\pi_i = \langle\langle \pi_i \circ \pi_0, \pi_i \circ \pi_1 \rangle\rangle$. In particular, it implies that $c = \langle\langle \tilde{D}\pi_0, \tilde{D}\pi_1 \rangle\rangle$. This is very useful because the differential of a pair can then be obtained from the pair of the differentials.

Proposition 14. *Assume **(Dproj-lin)**, **(D-chain)**. Let $f_0, f_1 \in \mathcal{C}(X, Y)$ such that $f_0 \boxplus f_1$. Then $\tilde{D}f_0 \boxplus \tilde{D}f_1$ and $\langle\langle \tilde{D}f_0, \tilde{D}f_1 \rangle\rangle = c \circ \tilde{D}\langle\langle f_0, f_1 \rangle\rangle$.*

Proof. $\pi_i \circ c \circ \tilde{D}\langle\langle f_0, f_1 \rangle\rangle = \tilde{D}\pi_i \circ \tilde{D}\langle\langle f_0, f_1 \rangle\rangle = \tilde{D}f_i$. \square

2.3 The Differentiation Monad

Proposition 15. *Assuming **(D-chain)**, **(Dproj-lin)** and **(Dsum-lin)**, the following diagrams commute.*

$$\begin{array}{ccc}
\tilde{D}X & \xrightarrow{\tilde{D}\iota_0} & \tilde{D}^2X & \xleftarrow{\iota_0} & \tilde{D}X \\
& \searrow \text{id}_{\tilde{D}X} & \downarrow \theta & \swarrow \text{id}_{\tilde{D}X} & \\
& & \tilde{D}X & &
\end{array}
\quad
\begin{array}{ccc}
\tilde{D}^3X & \xrightarrow{\tilde{D}\theta_X} & \tilde{D}^2X \\
\theta_{\tilde{D}X} \downarrow & & \downarrow \theta_X \\
\tilde{D}^2X & \xrightarrow{\theta_X} & \tilde{D}X
\end{array}$$

Proof. By **Corollary 3**, ι_0 is \tilde{D} -linear. Thus by **Remark 2**, $\tilde{D}\iota_0 = \langle\langle \iota_0 \circ \pi_0, \iota_0 \circ \pi_1 \rangle\rangle = \langle\langle \langle\langle \pi_0, 0 \rangle\rangle, \langle\langle \pi_1, 0 \rangle\rangle \rangle$. Hence $\theta \circ \tilde{D}\iota_0 = \langle\langle \pi_0, 0 + \pi_1 \rangle\rangle = \langle\langle \pi_0, \pi_1 \rangle\rangle = \text{id}_{\tilde{D}X}$ by **Proposition 8**. Next $\iota_0^{\tilde{D}X} = \langle\langle \langle\langle \pi_0, \pi_1 \rangle\rangle, \langle\langle 0, 0 \rangle\rangle \rangle$ since $\langle\langle \pi_0, \pi_1 \rangle\rangle = \text{id}$ and $\langle\langle 0^{X,X}, 0^{X,X} \rangle\rangle = 0^{\tilde{D}X, \tilde{D}X}$. By **Proposition 8** again, $\theta \circ \iota_0 = \langle\langle \pi_0, \pi_1 + 0 \rangle\rangle = \langle\langle \pi_0, \pi_1 \rangle\rangle = \text{id}_{\tilde{D}X}$ so the triangles commute.

The square is a direct computation. We use simple juxtaposition for the composition of projections for the sake of readability. The bottom path can be reduced using left compatibility of addition ([Proposition 2](#)) and additivity of the projections:

$$\begin{aligned}\theta \circ \theta &= \langle\langle \pi_0 \circ \pi_0 \circ \theta, \pi_1 \circ \pi_0 \circ \theta + \pi_0 \circ \pi_1 \circ \theta \rangle\rangle \\ &= \langle\langle \pi_0 \circ \pi_0 \circ \pi_0, \pi_1 \circ \pi_0 \circ \pi_0 + \pi_0 \circ (\pi_1 \circ \pi_0 + \pi_0 \circ \pi_1) \rangle\rangle \\ &= \langle\langle \pi_0 \circ \pi_0 \circ \pi_0, \pi_1 \circ \pi_0 \circ \pi_0 + (\pi_0 \circ \pi_1 \circ \pi_0 + \pi_0 \circ \pi_0 \circ \pi_1) \rangle\rangle.\end{aligned}$$

The upper path can be reduced by \tilde{D} -linearity of θ and left compatibility of sum ([Proposition 2](#)):

$$\begin{aligned}\theta \circ \tilde{D}\theta &= \langle\langle \pi_0 \circ \pi_0 \circ \tilde{D}\theta, \pi_1 \circ \pi_0 \circ \tilde{D}\theta + \pi_0 \pi_1 \circ \tilde{D}\theta \rangle\rangle \\ &= \langle\langle \pi_0 \circ \theta \circ \pi_0, \pi_1 \circ \theta \circ \pi_0 + \pi_0 \circ \theta \circ \pi_1 \rangle\rangle \\ &= \langle\langle \pi_0 \circ \pi_0 \circ \pi_0, (\pi_1 \circ \pi_0 + \pi_0 \circ \pi_1) \circ \pi_0 + \pi_0 \circ \pi_0 \circ \pi_1 \rangle\rangle \\ &= \langle\langle \pi_0 \circ \pi_0 \circ \pi_0, (\pi_1 \circ \pi_0 \circ \pi_0 + \pi_0 \circ \pi_1 \circ \pi_0) + \pi_0 \circ \pi_0 \circ \pi_1 \rangle\rangle.\end{aligned}$$

We conclude that those two morphisms are equal, using the associativity of the partial sum. \square

Corollary 4. (*Dproj-lin*), (*Dsum-lin*), (*D-chain*) and (*D-add*) imply that $(\tilde{D}, \iota_0, \theta)$ is a monad.

3 Interpreting the axioms as properties of the derivative

In this section, \mathcal{C} is only assumed to be a category equipped with a pre-differential structure ([Definition 11](#)). We show that the various axioms of a coherent differential category correspond to standard rules of the differential calculus, written as properties about $D(f)$. The results of this section are only necessary for [Section 6](#) but they also provide some intuitions on the axioms of coherent differentiation.

All the proofs are similar, and consist in using the joint monicity of π_0 and π_1 to reduce the axioms to a set of equations, then show that only one of those equations is non trivial. In what follows, “linear” always means \tilde{D} -linear.

Proposition 16. \tilde{D} is a functor if and only if $D(\text{id}) = \pi_1$ and $D(g \circ f) = D(g) \circ \langle\langle f \circ \pi_0, D(f) \rangle\rangle$.

Proof. \tilde{D} is a functor if and only if $\tilde{D}\text{id}_X = \text{id}_{\tilde{D}X}$ and for any g, f , $\tilde{D}(g \circ f) = \tilde{D}g \circ \tilde{D}f$. By joint monicity of the π_i , $\tilde{D}\text{id} = \text{id}$ if and only if $\pi_i \circ \tilde{D}\text{id} = \pi_i \circ \text{id} = \pi_i$. But $\pi_0 \circ \tilde{D}\text{id} = \text{id} \circ \pi_0 = \pi_0$ by assumptions on Pre-Differential Structures. So $\tilde{D}\text{id} = \text{id}$ if and only if $\pi_1 \circ \tilde{D}\text{id} = \pi_1$, that is, if and only if $D(\text{id}) = \pi_1$.

Similarly, $\pi_0 \circ \tilde{D}g \circ \tilde{D}f = g \circ \pi_0 \circ \tilde{D}f = g \circ f \circ \pi_0 = \pi_0 \circ \tilde{D}(g \circ f)$ by assumption on pre-differential-structures. So by joint monicity of the π_i , $\tilde{D}(g \circ f) = \tilde{D}g \circ \tilde{D}f$ if and only if $\pi_1 \circ \tilde{D}(g \circ f) = \pi_1 \circ \tilde{D}g \circ \tilde{D}f$. By definition of D , this corresponds exactly to the equation $D(g \circ f) = D(g) \circ \tilde{D}f = D(g) \circ \langle\langle f \circ \pi_0, D(f) \rangle\rangle$ \square

Proposition 17. Assuming (*Dproj-lin*), σ is linear if and only if $\tilde{D}\sigma = \tilde{D}\pi_0 + \tilde{D}\pi_1$. Assuming (*Dproj-lin*) and (*D-chain*), σ is linear if and only if for any f_0, f_1 that are summable, $\tilde{D}(f_0 + f_1) = \tilde{D}f_0 + \tilde{D}f_1$ (recall that $\tilde{D}f_0 \boxplus \tilde{D}f_1$ by [Proposition 14](#)).

Proof. By linearity of π_i , $\tilde{D}\pi_i = \langle\langle \pi_i \circ \pi_0, \pi_i \circ \pi_1 \rangle\rangle$ so by **Proposition 4**, $\tilde{D}\pi_0 + \tilde{D}\pi_1 = \langle\langle \pi_0 \circ \pi_0 + \pi_1 \circ \pi_0, \pi_0 \circ \pi_1 + \pi_1 \circ \pi_1 \rangle\rangle = \langle\langle (\pi_0 + \pi_1) \circ \pi_0, (\pi_0 + \pi_1) \circ \pi_1 \rangle\rangle = \langle\langle \sigma \circ \pi_0, \sigma \circ \pi_1 \rangle\rangle$. But σ is linear if and only if $\tilde{D}\sigma = \langle\langle \sigma \circ \pi_0, \sigma \circ \pi_1 \rangle\rangle$ by **Proposition 9**, that is, if and only if $\tilde{D}\sigma = \tilde{D}\pi_0 + \tilde{D}\pi_1$.

For the second part of the lemma, notice that the right statement for $f_0 = \pi_0$ and $f_1 = \pi_1$ is exactly $\tilde{D}\sigma = \tilde{D}\pi_0 + \tilde{D}\pi_1$, so the converse direction holds. For the forward direction, notice that

$$\begin{aligned} \tilde{D}(f_0 + f_1) &= \tilde{D}(\sigma \circ \langle\langle f_0, f_1 \rangle\rangle) \\ &= \tilde{D}\sigma \circ \tilde{D}\langle\langle f_0, f_1 \rangle\rangle \quad \text{by (D-chain)} \\ &= (\tilde{D}\pi_0 + \tilde{D}\pi_1) \circ \tilde{D}\langle\langle f_0, f_1 \rangle\rangle \quad \text{by assumptions} \\ &= \tilde{D}\pi_0 \circ \tilde{D}\langle\langle f_0, f_1 \rangle\rangle + \tilde{D}\pi_1 \circ \tilde{D}\langle\langle f_0, f_1 \rangle\rangle \\ &= \tilde{D}f_0 + \tilde{D}f_1 \quad \text{by (D-chain)} \end{aligned}$$

□

Corollary 5. *Assuming (Dproj-lin) and (D-chain), σ is linear if and only if for any f_0, f_1 that are summable, $D(f_0 + f_1) = D(f_0) + D(f_1)$*

Proof. The linearity of σ is equivalent to $\tilde{D}(f_0 + f_1) = \tilde{D}f_0 + \tilde{D}f_1$ for any f_0, f_1 summable. By **Proposition 4**, this is equivalent to $\langle\langle (f_0 + f_1) \circ \pi_0, D(f_0 + f_1) \rangle\rangle = \langle\langle f_0 \circ \pi_0 + f_1 \circ \pi_0, D(f_0) + D(f_1) \rangle\rangle$. The left compatibility of addition (**Proposition 2**) ensures that the first coordinates are always equal. So σ is linear if and only if for all $f_0 \boxplus f_1$, $D(f_0 + f_1) = D(f_0) + D(f_1)$. □

Proposition 18. *The following assertions are equivalent:*

- (1) ι_0 is natural;
- (2) For any $f \in \mathcal{C}(X, Y)$, $Df \circ \iota_0 = 0$;
- (3) For any $f \in \mathcal{C}(X, Y)$, any object U and $x \in \mathcal{C}(U, X)$, $Df \circ \langle\langle x, 0 \rangle\rangle = 0$.

Proof. (1) \Leftrightarrow (2). By joint monicity of the π_i , for any $f \in \mathcal{C}(X, Y)$, $\tilde{D}f \circ \iota_0 = \iota_0 \circ f$ if and only if $\pi_0 \circ \tilde{D}f \circ \iota_0 = \pi_0 \circ \iota_0 \circ f = f$ and $\pi_1 \circ \tilde{D}f \circ \iota_0 = \pi_1 \circ \iota_0 \circ f = 0$. The first condition always holds by naturality of π_0 and definition of ι_0 . So ι_0 is natural if and only if the second identity holds. This equation is precisely (2).

(2) \Leftrightarrow (3). The forward direction is directly obtained by composing the identity of (2) by x on the right. The reverse is directly obtained by applying the equation of (3) to $x = \text{id}_X$. □

Proposition 19. *Assuming (Dproj-lin) and (D-chain), the following assertions are equivalent:*

- (1) θ is natural;
- (2) for any $f \in \mathcal{C}(X, Y)$, $Df \circ \tilde{D}\pi_0 \boxplus Df \circ \pi_0$ and $Df \circ \theta = Df \circ \tilde{D}\pi_0 + Df \circ \pi_0$;
- (3) for any $f \in \mathcal{C}(X, Y)$, any object U and any $x, u, v \in \mathcal{C}(U, X)$ that are summable, $Df \circ \langle\langle x, u \rangle\rangle \boxplus Df \circ \langle\langle x, v \rangle\rangle$ and

$$Df \circ \langle\langle x, u + v \rangle\rangle = Df \circ \langle\langle x, u \rangle\rangle + Df \circ \langle\langle x, v \rangle\rangle.$$

Proof. (1) \Leftrightarrow (2). By joint monicity of the π_i , for any $f \in \mathcal{C}(X, Y)$, $\tilde{D}f \circ \theta = \theta \circ \tilde{D}^2f$ if and only if $\pi_0 \circ \tilde{D}f \circ \theta = \pi_0 \circ \theta \circ \tilde{D}^2f$ and $\pi_1 \circ \tilde{D}f \circ \theta = \pi_1 \circ \theta \circ \tilde{D}^2f$. The equation $\pi_0 \circ \tilde{D}f \circ \theta = \pi_0 \circ \theta \circ \tilde{D}^2f$ always holds. Indeed

$$\begin{aligned} \pi_0 \circ \tilde{D}f \circ \theta &= f \circ \pi_0 \circ \theta \quad \text{by naturality of } \pi_0 \\ &= f \circ \pi_0 \circ \pi_0 \quad \text{by definition of } \theta \\ \pi_0 \circ \theta \circ \tilde{D}^2f &= \pi_0 \circ \pi_0 \circ \tilde{D}^2f \quad \text{by definition of } \theta \\ &= f \circ \pi_0 \circ \pi_0 \quad \text{by naturality of } \pi_0. \end{aligned}$$

The left hand side of the equation $\pi_1 \circ \tilde{D}f \circ \theta = \pi_1 \circ \theta \circ \tilde{D}^2f$ is $D(f) \circ \theta$ by definition. The right hand side rewrites as follows.

$$\begin{aligned} \pi_1 \circ \theta \circ \tilde{D}^2f &= (\pi_0 \circ \pi_1 + \pi_1 \circ \pi_0) \circ \tilde{D}^2f \\ &= \pi_0 \circ \pi_1 \circ \tilde{D}^2f + \pi_1 \circ \pi_0 \circ \tilde{D}^2f \quad \text{by Proposition 2} \\ &= \pi_1 \circ \tilde{D}\pi_0 \circ \tilde{D}^2f + \pi_1 \circ \pi_0 \circ \tilde{D}^2f \quad \text{by } \tilde{D}\text{-linearity of } \pi_0 \\ &= \pi_1 \circ \tilde{D}(\pi_0 \circ \tilde{D}f) + \pi_1 \circ \pi_0 \circ \tilde{D}^2f \quad \text{by (D-chain)} \\ &= \pi_1 \circ \tilde{D}(f \circ \pi_0) + \pi_1 \circ \tilde{D}f \circ \pi_0 \quad \text{by naturality of } \pi_0 \\ &= \pi_1 \circ \tilde{D}f \circ \tilde{D}\pi_0 + \pi_1 \circ \tilde{D}f \circ \pi_0 \quad \text{by (D-chain)} \\ &= Df \circ \tilde{D}\pi_0 + Df \circ \pi_0 \quad \text{by definition} \end{aligned}$$

So this second equation under consideration is equivalent to the equation of (2).

(2) \Leftrightarrow (3). Recall that $\tilde{D}\pi_0 = \langle\langle \pi_0 \circ \pi_0, \pi_0 \circ \pi_1 \rangle\rangle$ by linearity of π_0 . Then the forward direction is directly obtained by composing the equation of (2) with $\langle\langle x, v \rangle\rangle, \langle\langle u, 0 \rangle\rangle$ on the right. The converse is directly obtained by applying the equation of (3) to $x = \pi_0 \circ \pi_0$, $u = \pi_1 \circ \pi_0$ and $v = \pi_0 \circ \pi_1$. \square

Remark 3. Notice that $D(D(f)) = \pi_1 \circ \tilde{D}(\pi_1 \circ \tilde{D}f) = \pi_1 \circ \tilde{D}\pi_1 \circ \tilde{D}^2f = \pi_1 \circ \pi_1 \circ \tilde{D}^2f$ assuming (D-chain) and (Dproj-lin). Thus, $D(D(f))$ is nothing more than the rightmost coordinate of \tilde{D}^2f . This will be useful for what follows in this part.

Proposition 20. *Assuming (Dproj-lin), (D-chain) and the naturality of ι_0 , the following assertions are equivalent:*

- (1) \lrcorner is natural;
- (2) for all morphism $f \in \mathcal{C}(X, Y)$, $D(D(f)) \circ \lrcorner = D(f)$;
- (3) for all morphism $f \in \mathcal{C}(X, Y)$, for all morphisms $x, u \in \mathcal{C}(U, X)$ summable,

$$D(D(f)) \circ \langle\langle x, 0 \rangle\rangle, \langle\langle 0, u \rangle\rangle = D(f) \circ \langle\langle x, u \rangle\rangle.$$

Proof. By joint monicity of the π_i , \lrcorner is natural if and only if for all f and for all $i, j \in \{0, 1\}$, $\pi_i \circ \pi_j \circ \tilde{D}^2f \circ \lrcorner = \pi_i \circ \pi_j \circ \lrcorner \circ \tilde{D}f$. By Remark 3 (and because $\pi_1 \circ \pi_1 \circ \lrcorner = \pi_1$), the equation for $i = j = 1$ corresponds exactly to the equation $D(D(f)) \circ \lrcorner = D(f)$. Thus, it suffices to show that $\pi_i \circ \pi_j \circ \tilde{D}^2f \circ \lrcorner = \pi_i \circ \pi_j \circ \lrcorner \circ \tilde{D}f$ always holds when $(i, j) \neq (1, 1)$ to conclude that (1) is equivalent to (2).

- Case $i = 0, j = 0$: $\pi_0 \circ \pi_0 \circ \lrcorner \circ \tilde{D}f = \pi_0 \circ \tilde{D}f = f \circ \pi_0$ and $\pi_0 \circ \pi_0 \circ \tilde{D}^2f \circ \lrcorner = f \circ \pi_0 \circ \pi_0 \circ \lrcorner = f \circ \pi_0$;

- Case $i = 1, j = 0$: $\pi_0 \circ \pi_1 \circ l \circ \tilde{D}f = 0 \circ \tilde{D}f = 0$ and $\pi_1 \circ \pi_0 \circ \tilde{D}^2 f \circ l = \pi_1 \circ \tilde{D}f \circ \pi_0 \circ l = \pi_1 \circ \tilde{D}f \circ \iota_0 \circ \pi_0 = \pi_1 \circ \iota_0 \circ f \circ \pi_0 = 0$ thanks to the naturality of ι_0 ;
- Case $i = 0, j = 1$: $\pi_0 \circ \pi_1 \circ l \circ \tilde{D}f = 0 \circ \tilde{D}f = 0$ and $\pi_0 \circ \pi_1 \circ \tilde{D}^2 f \circ l = \pi_1 \circ \tilde{D}\pi_0 \circ \tilde{D}^2 f \circ l = \pi_1 \circ \tilde{D}f \circ \tilde{D}\pi_0 \circ l = \pi_1 \circ \tilde{D}f \circ \iota_0 \circ \pi_0 = \pi_1 \circ \iota_0 \circ f \circ \pi_0 = 0$ thanks to the naturality of ι_0 .

Next (2) is a particular case of (3) for $x = \pi_0$ and $u = \pi_1$. Conversely, assuming (2) we have that $D(D(f)) \circ \langle\langle x, 0 \rangle\rangle, \langle\langle 0, u \rangle\rangle = D(D(f)) \circ l \circ \langle\langle x, u \rangle\rangle = D(f) \circ \langle\langle x, u \rangle\rangle$. \square

Proposition 21. *Assuming (*Dproj-lin*) and (*D-chain*), the following assertions are equivalent:*

- (1) c is natural;
- (2) for all morphism $f \in \mathcal{C}(X, Y)$, $D(D(f)) \circ c = D(D(f))$;
- (3) for all morphism $f \in \mathcal{C}(X, Y)$ and $x, u, v, w \in \mathcal{C}(U, X)$ that are summable,

$$D(D(f)) \circ \langle\langle x, u \rangle\rangle, \langle\langle v, w \rangle\rangle = D(D(f)) \circ \langle\langle x, v \rangle\rangle, \langle\langle u, w \rangle\rangle$$

Proof. By joint monicity of the π_i , c is natural if and only if for all f and for all $i, j \in \{0, 1\}$, $\pi_i \circ \pi_j \circ \tilde{D}^2 f \circ c = \pi_i \circ \pi_j \circ c \circ \tilde{D}^2 f$. But $\pi_i \circ \pi_j \circ c \circ \tilde{D}^2 f = \pi_j \circ \pi_i \circ \tilde{D}^2 f$. Then, by **Remark 3**, the equation for $i = j = 1$ corresponds exactly to the equation $D(D(f)) \circ c = D(D(f))$. Thus, it suffices to show that $\pi_i \circ \pi_j \circ \tilde{D}^2 f \circ c = \pi_j \circ \pi_i \circ \tilde{D}^2 f$ when $(i, j) \neq (1, 1)$ to conclude that (1) is equivalent to (2).

- $i = 0, j = 0$: The equation holds by reflexivity of equality.
- $i = 1, j = 0$: $\pi_0 \circ \pi_1 \circ \tilde{D}^2 f = \pi_1 \circ \tilde{D}\pi_0 \circ \tilde{D}^2 f = \pi_1 \circ \tilde{D}f \circ \tilde{D}\pi_0$ and $\pi_1 \circ \pi_0 \circ \tilde{D}^2 f \circ c = \pi_1 \circ \tilde{D}f \circ \pi_0 \circ c = \pi_1 \circ \tilde{D}f \circ \tilde{D}\pi_0$ so both sides are equal.
- $i = 0, j = 1$: $\pi_0 \circ \pi_1 \circ \tilde{D}^2 f \circ c = \pi_1 \circ \pi_0 \circ \tilde{D}^2 f$ if and only if $\pi_0 \circ \pi_1 \circ \tilde{D}^2 f = \pi_1 \circ \pi_0 \circ \tilde{D}^2 f \circ c$ because c is involutive. But this equation holds, as seen above.

Next, (2) is a particular case of (3) for $x = \pi_0 \circ \pi_0$, $u = \pi_0 \circ \pi_1$, $v = \pi_1 \circ \pi_0$ and $w = \pi_1 \circ \pi_1$. Conversely, if (3) holds then $D(D(f)) \circ \langle\langle x, u \rangle\rangle, \langle\langle v, w \rangle\rangle = D(D(f)) \circ c \circ \langle\langle x, v \rangle\rangle, \langle\langle u, w \rangle\rangle = D(D(f)) \circ \langle\langle x, v \rangle\rangle, \langle\langle u, w \rangle\rangle$. \square

4 Compatibility with the cartesian product

We assume in this section that \mathcal{C} is cartesian and is equipped with a left summability structure $(\tilde{D}, \pi_0, \pi_1, \sigma)$.

Notations 2. We use $\&$ for the cartesian product, following the notations of LL. For any objects Y_0, Y_1 , the projection will be written as $\mathbf{p}_i \in \mathcal{C}(Y_0 \& Y_1, Y_i)$ and the pairing of $f_0 \in \mathcal{C}(X, Y_0)$ and $f_1 \in \mathcal{C}(X, Y_1)$ as $\langle f_0, f_1 \rangle$. Finally, the terminal object will be written \top . Note that the uniqueness of the pairing in the universal property of the cartesian product can be understood as the joint monicity of the \mathbf{p}_i .

4.1 Cartesian product and summability structure

Definition 17. The summability structure $(\tilde{D}, \pi_0, \pi_1, \sigma)$ is *compatible with the cartesian product* if $\langle 0, 0 \rangle = 0$ and, for all $f_0, g_0 \in \mathcal{C}(X, Y_0)$ and $f_1, g_1 \in \mathcal{C}(X, Y_1)$:

- $\langle f_0, f_1 \rangle \boxplus \langle g_0, g_1 \rangle$ if and only if $f_0 \boxplus g_0$ and $f_1 \boxplus g_1$
- and then $\langle f_0, f_1 \rangle + \langle g_0, g_1 \rangle = \langle f_0 + f_1, g_0 + g_1 \rangle$.

That is, sums are computed componentwise. Let us break down this definition in more details.

Proposition 22. *The following are equivalent*

- $\mathfrak{p}_0, \mathfrak{p}_1$ are additive;
- $\langle 0, 0 \rangle = 0$ and for all $f_0, g_0 \in \mathcal{C}(X, Y_0)$ and $f_1, g_1 \in \mathcal{C}(X, Y_1)$, if $\langle f_0, f_1 \rangle \boxplus \langle g_0, g_1 \rangle$ then $f_0 \boxplus g_0, f_1 \boxplus g_1$ and $\langle f_0, f_1 \rangle + \langle g_0, g_1 \rangle = \langle f_0 + f_1, g_0 + g_1 \rangle$.

Proof. Assume that $\mathfrak{p}_0, \mathfrak{p}_1$ are additive. Then $\mathfrak{p}_i \circ 0 = 0 = \mathfrak{p}_i \circ \langle 0, 0 \rangle$. Thus by joint monicity, $0 = \langle 0, 0 \rangle$. Furthermore, assume that $\langle f_0, f_1 \rangle \boxplus \langle g_0, g_1 \rangle$. Then by additivity of \mathfrak{p}_i , $\mathfrak{p}_i \circ \langle f_0, f_1 \rangle = f_i$ and $\mathfrak{p}_i \circ \langle g_0, g_1 \rangle = g_i$ are summable and $f_i + g_i = \mathfrak{p}_i \circ (\langle f_0, f_1 \rangle + \langle g_0, g_1 \rangle)$. So the joint monicity of the \mathfrak{p}_i implies that $\langle f_0, f_1 \rangle + \langle g_0, g_1 \rangle = \langle f_0 + f_1, g_0 + g_1 \rangle$.

Conversely, since $\langle 0, 0 \rangle = 0$ we have $\mathfrak{p}_i \circ 0 = \mathfrak{p}_i \circ \langle 0, 0 \rangle = 0$. Let $f, g \in \mathcal{C}(X, Y_0 \& Y_1)$ be summable. One can write $f = \langle \mathfrak{p}_0 \circ f, \mathfrak{p}_1 \circ f \rangle$ and $g = \langle \mathfrak{p}_0 \circ g, \mathfrak{p}_1 \circ g \rangle$. Since $f \boxplus g$ we have $\mathfrak{p}_i \circ f \boxplus \mathfrak{p}_i \circ g$ and $f + g = \langle \mathfrak{p}_0 \circ f + \mathfrak{p}_0 \circ g, \mathfrak{p}_1 \circ f + \mathfrak{p}_1 \circ g \rangle$. Applying \mathfrak{p}_i on this equation yields that $\mathfrak{p}_i \circ (f + g) = \mathfrak{p}_i \circ f + \mathfrak{p}_i \circ g$ so \mathfrak{p}_i is additive. \square

Corollary 6. *If \mathfrak{p}_0 and \mathfrak{p}_1 are additive, then $0 \& 0 = 0$ and for all $f_0, g_0 \in \mathcal{C}(X_0, Y_0)$ and $f_1, g_1 \in \mathcal{C}(X_1, Y_1)$, if $f_0 \& f_1 \boxplus g_0 \& g_1$ then $f_0 \boxplus g_0, f_1 \boxplus g_1$ and $f_0 \& f_1 + g_0 \& g_1 = (f_0 + g_0) \& (f_1 + g_1)$.*

Proof. We simply use the fact that $f \& g = \langle f \circ \mathfrak{p}_0, g \circ \mathfrak{p}_1 \rangle$ and **Proposition 22** together with the left compatibility of sum with regard to composition (**Proposition 2**). \square

We now assume that the projections \mathfrak{p}_0 and \mathfrak{p}_1 are additive. This allows us to define a morphism $c_{\&} \in \mathcal{C}(\tilde{D}(X_0 \& X_1), \tilde{D}X_0 \& \tilde{D}X_1)$ for any objects X_0, X_1 as $c_{\&} := \langle \langle \mathfrak{p}_0 \circ \pi_0, \mathfrak{p}_0 \circ \pi_1 \rangle, \langle \mathfrak{p}_1 \circ \pi_0, \mathfrak{p}_1 \circ \pi_1 \rangle \rangle$. In other words, $\pi_i \circ \mathfrak{p}_j \circ c_{\&} = \mathfrak{p}_j \circ \pi_i$, that is

$$c_{\&} \circ \langle \langle f_0, f_1 \rangle, \langle g_0, g_1 \rangle \rangle = \langle \langle f_0, g_0 \rangle, \langle f_1, g_1 \rangle \rangle.$$

This is very reminiscent of the flip c (it swaps the two middle coordinates), except that there are no summability conditions associated with the $\langle _, _ \rangle$ pairing.

Theorem 4. *The following assertions are equivalent*

- (1) $c_{\&}$ is an isomorphism;
- (2) $\pi_0 \& \pi_0 \boxplus \pi_1 \& \pi_1$;
- (3) for any $f_0, g_0 \in \mathcal{C}(X, Y_0)$, $f_1, g_1 \in \mathcal{C}(X, Y_1)$, if $f_0 \boxplus g_0$ and $f_1 \boxplus g_1$ then $f_0 \& f_1 \boxplus g_0 \& g_1$;
- (4) for any $f_0, g_0 \in \mathcal{C}(X, Y_0)$, $f_1, g_1 \in \mathcal{C}(X, Y_1)$, if $f_0 \boxplus g_0$ and $f_1 \boxplus g_1$ then $\langle f_0, f_1 \rangle \boxplus \langle g_0, g_1 \rangle$

and then $\langle \pi_0 \& \pi_0, \pi_1 \& \pi_1 \rangle = c_{\&}^{-1}$.

Proof. (1) \Rightarrow (2): Assume that $c_{\&}$ is an isomorphism with inverse w . Then $\pi_i \circ \mathbf{p}_j = \pi_i \circ \mathbf{p}_j \circ c_{\&} \circ w = \mathbf{p}_j \circ \pi_i \circ w$. But $\pi_i \circ \mathbf{p}_j = \mathbf{p}_j \circ (\pi_i \& \pi_i)$ by naturality of \mathbf{p}_j so $\mathbf{p}_j \circ \pi_i \circ w = \mathbf{p}_j \circ (\pi_i \& \pi_i)$. By joint monicity of the \mathbf{p}_j 's we have $\pi_i \circ w = (\pi_i \& \pi_i)$. That is $w = \langle \pi_0 \& \pi_0, \pi_1 \& \pi_1 \rangle$.

(2) \Rightarrow (1): Assume that $\pi_0 \& \pi_0 \boxplus \pi_1 \& \pi_1$, of witness w . Then, $\mathbf{p}_j \circ \pi_i \circ w = \mathbf{p}_j \circ (\pi_i \& \pi_i) = \pi_i \circ \mathbf{p}_j$. Hence

$$\begin{aligned} \mathbf{p}_j \circ \pi_i \circ w \circ c_{\&} &= \pi_i \circ \mathbf{p}_j \circ c_{\&} = \mathbf{p}_j \circ \pi_i \\ \pi_i \circ \mathbf{p}_j \circ c_{\&} \circ w &= \mathbf{p}_j \circ \pi_i \circ w = \pi_i \circ \mathbf{p}_j \end{aligned}$$

By joint monicity of the \mathbf{p}_j 's and of the π_i 's we get $w \circ c_{\&} = \text{id}_{\tilde{\mathcal{D}}(X_0 \& X_1)}$ and $c_{\&} \circ w = \text{id}_{\tilde{\mathcal{D}}X_0 \& \tilde{\mathcal{D}}X_1}$.

(2) \Rightarrow (3): We have $\langle f_0, g_0 \rangle \in \mathcal{C}(X, \tilde{\mathcal{D}}Y_0)$ and $\langle f_1, g_1 \rangle \in \mathcal{C}(X, \tilde{\mathcal{D}}Y_1)$. Let $w = \langle \pi_0 \& \pi_0, \pi_1 \& \pi_1 \rangle \circ (\langle f_0, g_0 \rangle \& \langle f_1, g_1 \rangle)$. We have $\pi_0 \circ w = f_0 \& f_1$ and $\pi_1 \circ w = g_0 \& g_1$ so that $f_0 \& f_1 \boxplus g_0 \& g_1$.

(3) \Rightarrow (2): (2) is a particular case of case (3).

(3) \Rightarrow (4): Assume that $f_0 \boxplus g_0$ and $f_1 \boxplus g_1$. Then by assumption, $f_0 \& f_1 \boxplus g_0 \& g_1$. Let $w = \langle f_0 \& f_1, g_0 \& g_1 \rangle \circ \langle \text{id}, \text{id} \rangle$. Then $\pi_0 \circ w = \langle f_0, f_1 \rangle$ and $\pi_1 \circ w = \langle g_0, g_1 \rangle$ so that $\langle f_0, f_1 \rangle \boxplus \langle g_0, g_1 \rangle$.

(4) \Rightarrow (3): Assume that $f_0 \boxplus g_0$ and $f_1 \boxplus g_1$. Then $f_0 \circ \mathbf{p}_0 \boxplus g_0 \circ \mathbf{p}_0$ and $f_1 \circ \mathbf{p}_1 \boxplus g_1 \circ \mathbf{p}_1$ by left compatibility wrt. composition ([Proposition 2](#)). Hence, by assumption, $\langle f_0 \circ \mathbf{p}_0, f_1 \circ \mathbf{p}_1 \rangle \boxplus \langle g_0 \circ \mathbf{p}_0, g_1 \circ \mathbf{p}_1 \rangle$. That is $f_0 \& f_1 \boxplus g_0 \& g_1$. \square

Corollary 7. *A summability structure is compatible with the cartesian product if and only if $\mathbf{p}_0, \mathbf{p}_1$ are additive and $c_{\&}$ is an isomorphism.*

4.2 Cartesian product and differential structure

We now assume that \mathcal{C} is a cartesian category with a pre-differential structure $(\tilde{\mathcal{D}}, \pi_0, \pi_1, \sigma)$.

Definition 18. The (pre-)differential structure $(\tilde{\mathcal{D}}, \pi_0, \pi_1, \sigma)$ is *compatible with the cartesian product* if the underlying summability structure is compatible with the cartesian product, and if $\mathbf{p}_0, \mathbf{p}_1$ are $\tilde{\mathcal{D}}$ -linear. A *cartesian coherent differential category* (CCDC) is a coherent differential category whose cartesian product is compatible with the differential structure.

We assume that \mathcal{C} is a CCDC. By $\tilde{\mathcal{D}}$ -linearity of \mathbf{p}_0 and \mathbf{p}_1 , all constructions involving only the cartesian product are $\tilde{\mathcal{D}}$ -linear.

Proposition 23. *If $h_0 \in \mathcal{C}(X, Y_0)$ and $h_1 \in \mathcal{C}(X, Y_1)$ are $\tilde{\mathcal{D}}$ -linear, then $\langle h_0, h_1 \rangle$ is $\tilde{\mathcal{D}}$ -linear. If $f_0 \in \mathcal{C}(X_0, Y_0)$ and $f_1 \in \mathcal{C}(X_1, Y_1)$ are $\tilde{\mathcal{D}}$ -linear, then $f_0 \& f_1$ is $\tilde{\mathcal{D}}$ -linear.*

Proof. For the first statement we proceed as for [Proposition 12](#) except that the pairing as a summable pair is replaced by the pairing of the cartesian product. The second statement follows from the first one, because $f_0 \& f_1 = \langle f_0 \circ \mathbf{p}_0, f_1 \circ \mathbf{p}_1 \rangle$, the projections are $\tilde{\mathcal{D}}$ -linear, and $\tilde{\mathcal{D}}$ -linearity is closed under composition. \square

For any objects X_0, X_1 , there is a natural transformation $\langle \tilde{\mathcal{D}}\mathbf{p}_0, \tilde{\mathcal{D}}\mathbf{p}_1 \rangle \in \mathcal{C}(\tilde{\mathcal{D}}(X_0 \& X_1), \tilde{\mathcal{D}}X_0 \& \tilde{\mathcal{D}}X_1)$. By $\tilde{\mathcal{D}}$ -linearity of \mathbf{p}_0 and \mathbf{p}_1 this natural transformation is equal to $\langle \langle \mathbf{p}_0 \circ \pi_0, \mathbf{p}_0 \circ \pi_1 \rangle, \langle \mathbf{p}_1 \circ \pi_0, \mathbf{p}_1 \circ \pi_1 \rangle \rangle = c_{\&}$. Whence a result similar to [Proposition 14](#).

Proposition 24. *For any $f_0 \in \mathcal{C}(X, Y_0)$ and $f_1 \in \mathcal{C}(X, Y_1)$, $\langle \tilde{\mathcal{D}}f_0, \tilde{\mathcal{D}}f_1 \rangle = c_{\&} \circ \tilde{\mathcal{D}}\langle f_0, f_1 \rangle$*

Proof. $\mathbf{p}_i \circ c_{\&} \circ \tilde{\mathcal{D}}\langle f_0, f_1 \rangle = \tilde{\mathcal{D}}\mathbf{p}_i \circ \tilde{\mathcal{D}}\langle f_0, f_1 \rangle = \tilde{\mathcal{D}}f_i$. \square

4.3 Partial derivatives

Using $c_{\&}^{-1}$, we define two natural transformations

$$\begin{aligned}\Phi^0 &= (c_{\&})^{-1} \circ (\text{id}_{\tilde{D}X_0} \& \iota_0) \in \mathcal{C}(\tilde{D}X_0 \& X_1, \tilde{D}(X_0 \& X_1)) \\ \Phi^1 &= (c_{\&})^{-1} \circ (\iota_0 \& \text{id}_{\tilde{D}X_1}) \in \mathcal{C}(X_0 \& \tilde{D}X_1, \tilde{D}(X_0 \& X_1))\end{aligned}$$

Note that $c_{\&}$, $(c_{\&})^{-1}$, Φ^0 and Φ^1 are all \tilde{D} -linear, thanks to [Propositions 11](#) and [23](#) and [Corollary 3](#).

Proposition 25. $\Phi^0 = \langle\langle \pi_0 \& \text{id}_{X_1}, \pi_1 \& 0 \rangle\rangle$ and $\Phi^1 = \langle\langle \text{id}_{X_0} \& \pi_0, 0 \& \pi_1 \rangle\rangle$

Proof. By [Theorem 4](#), $(c_{\&})^{-1} = \langle\langle \pi_0 \& \pi_0, \pi_1 \& \pi_1 \rangle\rangle$ and the result follows by a straightforward computation. \square

Definition 19 (Partial derivative). If $f \in \mathcal{C}(X_0 \& X_1, Y)$ one can define $\tilde{D}_0 f := \tilde{D}f \circ \Phi^0 \in \mathcal{C}(\tilde{D}X_0 \& X_1, \tilde{D}Y)$ and $\tilde{D}_1 f := \tilde{D}f \circ \Phi^1 \in \mathcal{C}(X_0 \& \tilde{D}X_1, \tilde{D}Y)$, the *partial derivatives* of f .

Proposition 26. For any $f \in \mathcal{C}(X_0 \& X_1, Y)$, $\pi_0 \circ \tilde{D}_0 f = f \circ (\pi_0 \& \text{id})$ and $\pi_0 \circ \tilde{D}_1 f = f \circ (\text{id} \& \pi_0)$.

Proof. $\pi_0 \circ \tilde{D}_0 f = \pi_0 \circ \tilde{D}f \circ \Phi^0 = f \circ \pi_0 \circ \Phi^0 = f \circ (\pi_0 \& \text{id})$ by [Proposition 25](#). The proof for Φ^1 is similar. \square

Proposition 27. The following diagram commutes.

$$\begin{array}{ccc}\tilde{D}(X_0 \& \tilde{D}X_1) & \xleftarrow{\Phi^0} & \tilde{D}X_0 \& \tilde{D}X_1 & \xrightarrow{\Phi^1} & \tilde{D}(\tilde{D}X_0 \& X_1) \\ \tilde{D}\Phi^1 \downarrow & & & & & \downarrow \tilde{D}\Phi^0 \\ \tilde{D}^2(X_0 \& X_1) & \xrightarrow{\quad c \quad} & & & & \tilde{D}^2(X_0 \& X_1)\end{array}$$

Proof. We use [Proposition 25](#) to compute $\tilde{D}\Phi^1 \circ \Phi^0$ and $\tilde{D}\Phi^0 \circ \Phi^1$. Since Φ^0 is \tilde{D} -linear, $\tilde{D}\Phi^0 = \langle\langle \Phi^0 \circ \pi_0, \Phi^0 \circ \pi_1 \rangle\rangle$ by [Remark 2](#). Thus

$$\begin{aligned}\tilde{D}\Phi^0 \circ \Phi^1 &= \tilde{D}\Phi^0 \circ \langle\langle \text{id}_{X_0} \& \pi_0, 0 \& \pi_1 \rangle\rangle \\ &= \langle\langle \Phi^0 \circ (\text{id}_{X_0} \& \pi_0), \Phi^0 \circ (0 \& \pi_1) \rangle\rangle \\ &= \langle\langle \langle\langle \pi_0 \& \pi_0, \pi_1 \& 0 \rangle\rangle, \langle\langle 0 \& \pi_1, 0 \& 0 \rangle\rangle \rangle\end{aligned}$$

Similarly, $\tilde{D}\Phi^1 \circ \Phi^0 = \langle\langle \langle\langle \pi_0 \& \pi_0, 0 \& \pi_1 \rangle\rangle, \langle\langle \pi_1 \& 0, 0 \& 0 \rangle\rangle \rangle$. The commutation results from [Proposition 8](#). \square

Proposition 28. The following diagram commutes

$$\begin{array}{ccc}\tilde{D}(X_0 \& \tilde{D}X_1) & \xleftarrow{\Phi^0} & \tilde{D}X_0 \& \tilde{D}X_1 & \xrightarrow{\Phi^1} & \tilde{D}(\tilde{D}X_0 \& X_1) \\ \tilde{D}\Phi^1 \downarrow & & \downarrow c_{\&}^{-1} & & \downarrow \tilde{D}\Phi^0 \\ \tilde{D}^2(X_0 \& X_1) & \xrightarrow{\theta} & \tilde{D}(X_0 \& X_1) & \xleftarrow{\theta} & \tilde{D}^2(X_0 \& X_1)\end{array}$$

Proof. Thanks to the computation of $\tilde{D}\Phi^0 \circ \Phi^1$ in the proof of [Proposition 27](#), we know that $\theta \circ \tilde{D}\Phi^0 \circ \Phi^1 = \langle\langle \pi_0 \& \pi_0, \pi_1 \& 0 + 0 \& \pi_1 \rangle\rangle = \langle\langle \pi_0 \& \pi_0, \pi_1 \& \pi_1 \rangle\rangle$ by [Corollary 6](#). So $\theta \circ \tilde{D}\Phi^0 \circ \Phi^1 = (c_{\&})^{-1}$ by [Theorem 4](#). A similar computation yields the result for $\theta \circ \tilde{D}\Phi^1 \circ \Phi^0$. \square

Remark 4. We can check that the natural morphisms Φ^0, Φ^1 are *strenghts* [10, 11] for the monad $(\tilde{D}, \iota_0, \theta)$. Then the diagram of [Proposition 28](#) means that this strong monad is a *commutative monad*. The diagrams can be checked by hand, but are also a consequence of very generic properties about strong monads on cartesian categories.

As mentioned in [12] in paragraph 2.3, any monad (M, η, μ) on a cartesian category can be endowed with the structure of a *colax symmetric monoidal monad*⁸ taking

- n^0 is the unique element of $\mathcal{C}(M\top, \top)$
- $n^2_{X_1, X_2} := \langle Mp_1, Mp_2 \rangle \in \mathcal{C}(M(X_1 \& X_2), MX_1 \& MX_2)$

If n^2 and n^0 are isos, M becomes a (*strong*) *symmetric monoidal monad*. This is what happens here for $M = \tilde{D}$, because $n^2 = c_{\&}$ and we can show that n^0 is an isomorphism with inverse ι_0 using the join monicity of the π_i . But symmetric monoidal monad are the same as commutative monads as shown in [10, 14], and it turns out that the strenghts induced from the symmetric monoidal structure are exactly Φ^0 and Φ^1 .

The axioms ([D-Schwarz](#)) and ([D-add](#)) carry to the setting of partial derivatives very naturally thanks to [Propositions 27](#) and [28](#) respectively, giving the full fledged Schwarz and Leibniz rules. The fact that the Leibniz rule is a consequence of the additivity of the derivative is not surprising, as it is also the case in the usual differential calculus: $f'(x, y) \cdot (u, v) = f'(x, y) \cdot (u, 0) + f'(x, y) \cdot (0, v) = \frac{\partial f}{\partial x}(x, y) \cdot u + \frac{\partial f}{\partial y}(x, y) \cdot v$.

Proposition 29 (Leibniz rule). $\tilde{D}f \circ c_{\&}^{-1} = \theta \circ \tilde{D}_0 \tilde{D}_1 f = \theta \circ \tilde{D}_1 \tilde{D}_0 f$

Proof. Let us prove that $\tilde{D}f \circ c_{\&}^{-1} = \theta \circ \tilde{D}_0 \tilde{D}_1 f$.

$$\begin{aligned} \theta \circ \tilde{D}_0 \tilde{D}_1 f &= \theta \circ \tilde{D}(\tilde{D}f \circ \Phi^1) \circ \Phi^0 && \text{by definition} \\ &= \theta \circ \tilde{D}^2 f \circ \tilde{D}\Phi^1 \circ \Phi^0 && \text{by (D-chain)} \\ &= \tilde{D}f \circ \theta \circ \tilde{D}\Phi^1 \circ \Phi^0 && \text{by (D-add)} \\ &= \tilde{D}f \circ c_{\&}^{-1} && \text{by Proposition 28} \end{aligned} \quad \square$$

The proof of $\tilde{D}f \circ c_{\&}^{-1} = \theta \circ \tilde{D}_1 \tilde{D}_0 f$ is similar.

Proposition 30 (Schwarz rule). $\tilde{D}_0 \tilde{D}_1 f = c \circ \tilde{D}_1 \tilde{D}_0 f$

Proof. Very similar to that of [Proposition 29](#), except that it uses the naturality of c of ([D-Schwarz](#)) instead of the naturality of θ . \square

4.4 Generalization to arbitrary finite products

Notations 3. Recall that the existence of arbitrary finite products is equivalent to the existence of a binary product and a terminal object. In order to stay consistent with the current notations, we write the finite products starting from 0: $X_0 \& \dots \& X_n$. We allow empty products, with the convention that taking $n = -1$ yields a product $X_0 \& \dots \& X_{-1} := \top$.

The constructions above can be extended to arbitrary finite products. One can indeed define a (symmetric monoidal) natural transformation $c_{\&}^n \in \mathcal{C}(\tilde{D}(X_0 \& \dots \& X_n), \tilde{D}X_0 \& \dots \& \tilde{D}X_n)$ inductively by $(c_{\&})_X := t_{\tilde{D}\top} \in \mathcal{C}(\tilde{D}\top, \top)$, $(c_{\&}^0)_X := \text{id}_{\tilde{D}X} \in \mathcal{C}(\tilde{D}X, \tilde{D}X)$ and $c_{\&}^{n+1} := c_{\&} \circ \langle c_{\&}^n, \text{id}_{\tilde{D}X_{n+1}} \rangle$. By associativity of the cartesian product, this definition does not depend on the actual parenthesizing of $X_0 \& \dots \& X_n$.

⁸Also called oplax symmetric monoidal monad, or symmetric comonoidal monad, or Hopf monad, see [13]

Notations 4. Let $X_0, Y_0, \dots, X_n, Y_n \in \text{Obj}(\mathcal{C})$, $i \in \llbracket 0, n \rrbracket$ and $f_k \in \mathcal{C}(X_k, Y_k)$ for each $k \neq i$. Let $g \in \mathcal{C}(X_i, Y_i)$. Define $(g; f_{-i}) := f_0 \& \dots \& f_{i-1} \& g \& f_{i+1} \& \dots \& f_n$ in which we use f_i everywhere except at position i where we use g .

Similarly to the binary case, one can then define a strength $\Phi^i \in \mathcal{C}(X_0 \& \dots \& \tilde{D}X_i \& \dots \& X_n, \tilde{D}(X_0 \& \dots \& X_n))$ as

$$\Phi^i := (\mathbf{c}_{\&}^n)^{-1} \circ (\text{id}_{\tilde{D}X_i}; (\iota_0)_{-i}).$$

Proposition 31. $\mathbf{c}_{\&}^n$ is an isomorphism and $(\mathbf{c}_{\&}^n)^{-1} = \langle\langle \pi_0 \& \dots \& \pi_0, \pi_1 \& \dots \& \pi_1 \rangle\rangle$. Hence, $\Phi^i = \langle\langle (\pi_0; \text{id}_{-i}), (\pi_1; 0_{-i}) \rangle\rangle$.

Proof. The equation on $\mathbf{c}_{\&}^n$ is obtained by unfolding the inductive definition and using [Theorem 4](#). The equations on the Φ^i 's follow from this, as in [Proposition 25](#). \square

Definition 20. For any $f \in \mathcal{C}(X_0 \& \dots \& X_n, Y)$ one can define the i -th partial derivative of f as $\tilde{D}_i f := \tilde{D}f \circ \Phi^i \in \mathcal{C}(X_0 \& \dots \& \tilde{D}X_i \& \dots \& X_n, \tilde{D}Y)$.

Proposition 32. $\pi_0 \circ \tilde{D}_i f = f \circ (\pi_0; \text{id}_{-i})$.

Proof. Same as [Proposition 26](#). \square

Definition 21. For any $X \in \text{Obj}(\mathcal{C})$ and $n \geq 0$, we can define $\theta_X^k \in \mathcal{C}(\tilde{D}^{n+1}X, \tilde{D}X)$ as the composition of k copies of θ : $\theta_X^0 = \text{id}_{\tilde{D}X}$ and $\theta_X^{k+1} = \theta_X^k \circ \theta_{\tilde{D}^k X}$. We define similarly $\pi_i^k \in \mathcal{C}(\tilde{D}^k X, X)$.

Note that $\theta^k = \langle\langle \pi_0^{k+1}, \sum_{j=0}^k \pi_0^j \circ \pi_1 \circ \pi_0^{k-j} \rangle\rangle$. In other words, the right component of θ^k sums over all of the possible combinations of k left projections and one right projection. One can prove a generalization of [Proposition 28](#) for $n \geq 1$,

$$(\mathbf{c}_{\&}^n)^{-1} = \theta^n \circ \tilde{D}^n \Phi^{\alpha(n)} \circ \dots \circ \tilde{D} \Phi^{\alpha(1)} \circ \tilde{D} \Phi^{\alpha(0)}$$

for any α permutation of $\llbracket 0, n \rrbracket$. As in [Proposition 29](#), this generalizes the Leibniz Rule to the n -ary case.

Proposition 33 (Leibniz, generalized). *For any $n \geq 1$ and for any permutation α of $\llbracket 0, n \rrbracket$,*

$$\tilde{D}f \circ (\mathbf{c}_{\&}^n)^{-1} = \theta^n \circ \tilde{D}_{\alpha(n)} \dots \tilde{D}_{\alpha(0)} f.$$

4.5 Multilinear morphism

We generalize to multivariate functions the notion of additivity and \tilde{D} -linearity.

Definition 22. A morphism $f \in \mathcal{C}(Y_0 \& \dots \& Y_n, Z)$ is additive in its i^{th} argument (for $i \in \llbracket 0, n \rrbracket$) if $f \circ (0; \text{id}_{-i}) = 0$ and if for all $h_0, h_1 \in \mathcal{C}(X, Y_i)$ such that $h_0 \boxplus h_1$, then $f \circ (h_0; \text{id}_{-i}) \boxplus f \circ (h_1; \text{id}_{-i})$ and

$$f \circ (h_0; \text{id}_{-i}) + f \circ (h_1; \text{id}_{-i}) = f \circ (h_0 + h_1; \text{id}_{-i})$$

Proposition 34. *A morphism $f \in \mathcal{C}(Y_0 \& \dots \& Y_n, Z)$ such that $f \circ (0; \text{id}_{-i}) = 0$ is additive in its i^{th} argument if and only if $f \circ (\pi_0; \text{id}_{-i}) \boxplus f \circ (\pi_1; \text{id}_{-i})$ with sum $f \circ (\sigma; \text{id}_{-i})$.*

Proof. The proof is the same as [Proposition 3](#), using the fact that for any $k \in \{0, 1\}$, $f \circ (h_k; \text{id}_{-i}) = f \circ (\pi_k; \text{id}_{-i}) \circ (\langle\langle h_0, h_1 \rangle\rangle; \text{id}_{-i})$. \square

Definition 23. A morphism $f \in \mathcal{C}(X_0 \& \cdots \& X_n, Y)$ is *linear in its i^{th} argument* if it is additive in this argument and if $\pi_1 \circ \tilde{D}_i f = f \circ (\pi_1; \text{id}_{-i})$.

As in **Proposition 10**, **(D-add)** ensures that the equation $\pi_1 \circ \tilde{D}_i f = f \circ (\pi_1; \text{id}_{-i})$ is a sufficient condition for linearity in the i^{th} argument.

Proposition 35. Assume that $\pi_1 \circ \tilde{D}_i f = f \circ (\pi_1; \text{id}_{-i})$. Then f is additive in its i^{th} argument, hence linear in that argument.

Proof. The equation allows rewriting $f \circ (h; \text{id}_{-i})$ as follows.

$$\begin{aligned} f \circ (h; \text{id}_{-i}) &= f \circ (\pi_1; \text{id}_{-i}) \circ (\langle\langle 0, h \rangle\rangle; \text{id}_{-i}) \\ &= \tilde{D}_i f \circ (\langle\langle 0, h \rangle\rangle; \text{id}_{-i}) \quad \text{by assumption} \\ &= \tilde{D}f \circ \Phi^i \circ (\langle\langle 0, h \rangle\rangle; \text{id}_{-i}) \\ &= \tilde{D}f \circ \langle\langle (0; \text{id}_{-i}), (h; 0_{-i}) \rangle\rangle \quad \text{by Proposition 31} \end{aligned}$$

In particular, $f \circ (0; \text{id}_{-i}) = \tilde{D}f \circ \langle\langle (0; \text{id}_{-i}), (0; 0_{-i}) \rangle\rangle$. But $(0; 0_{-i}) = 0$ by **Corollary 6**. So by **(D-add)** and **Proposition 18**, $f \circ (0; \text{id}_{-i}) = 0$. Similarly, if $h_0 \boxplus h_1$,

$$\begin{aligned} f \circ (h_0 + h_1; \text{id}_{-i}) &= \tilde{D}f \circ \langle\langle (0; \text{id}_{-i}), (h_0 + h_1; 0_{-i}) \rangle\rangle \\ &= \tilde{D}f \circ \langle\langle (0; \text{id}_{-i}), (h_0; 0_{-i}) + (h_1; 0_{-i}) \rangle\rangle \quad \text{by Corollary 6} \\ &= \tilde{D}f \circ \langle\langle (0; \text{id}_{-i}), (h_0; 0_{-i}) \rangle\rangle + \tilde{D}f \circ \langle\langle (0; \text{id}_{-i}), (h_1; 0_{-i}) \rangle\rangle \quad \text{by (D-add) and Proposition 19} \\ &= f \circ (h_0; \text{id}_{-i}) + f \circ (h_1; \text{id}_{-i}). \end{aligned}$$

□

Definition 24. A morphism $f \in \mathcal{C}(X_0 \& \cdots \& X_n, Y)$ is *multilinear* (and more precisely, $(n+1)$ -linear) if it is linear in all of its argument. Note that the 1-linear morphisms are exactly the \tilde{D} -linear ones.

As a sanity check of the notion, we can use the result below together with the Leibniz rule to show a result similar to the fact that in differential calculus, if Φ is a bilinear map, then $\Phi'(x, y) \cdot (u, v) = \Phi(x, v) + \Phi(u, y)$.

Lemma 1. For any $f \in \mathcal{C}(X_0 \& \cdots \& X_n, Y)$ and $i, j \in \llbracket 0, n \rrbracket$ such that $i \neq j$,

$$\tilde{D}\pi_0 \circ \tilde{D}_i \tilde{D}_j f = \tilde{D}_i f \circ (\pi_0; \text{id}_{-j})$$

Proof. This is a direct computation

$$\begin{aligned} \tilde{D}\pi_0 \circ \tilde{D}_i \tilde{D}_j f &= \tilde{D}\pi_0 \circ \tilde{D}(\tilde{D}_j f) \circ \Phi^i \\ &= \tilde{D}(\pi_0 \circ \tilde{D}_j f) \circ \Phi^i \quad \text{by (D-chain)} \\ &= \tilde{D}(f \circ (\pi_0; \text{id}_{-j})) \circ \Phi^i \quad \text{by Proposition 32} \\ &= \tilde{D}f \circ \tilde{D}(\pi_0; \text{id}_{-j}) \circ \Phi^i \quad \text{by (D-chain)} \\ &= \tilde{D}f \circ \Phi^i \circ (\pi_0; \text{id}_{-j}) \quad \text{since } \Phi^i \text{ natural and } i \neq j \\ &= \tilde{D}_i f \circ (\pi_0; \text{id}_{-j}) \end{aligned}$$

□

Theorem 5. For any $(n + 1)$ -linear morphism $f \in \mathcal{C}(X_0 \& \cdots \& X_n, Y)$

$$\begin{aligned}\pi_0 \circ \tilde{D}f \circ (\mathbf{c}\&)^{-1} &= f \circ (\pi_0 \& \cdots \& \pi_0) \\ \pi_1 \circ \tilde{D}f \circ (\mathbf{c}\&)^{-1} &= f \circ (\pi_1 \& \pi_0 \& \cdots \& \pi_0) + \cdots + f \circ (\pi_0 \& \cdots \& \pi_0 \& \pi_1)\end{aligned}$$

Proof. We will write the proof for $n = 1$. The general case relies on the same arguments. The first equation is just a direct consequence of the naturality of π_0 and [Proposition 25](#). For the second equation, Leibniz ([Proposition 29](#)) ensures that $\pi_1 \circ \tilde{D}f \circ \mathbf{c}\&^{-1} = \pi_1 \circ \theta \circ \tilde{D}_0 \tilde{D}_1 f = \pi_1 \circ \pi_0 \circ \tilde{D}_0 \tilde{D}_1 f + \pi_0 \circ \pi_1 \circ \tilde{D}_0 \tilde{D}_1 f$. We can compute those two summands separately.

$$\begin{aligned}\pi_1 \circ \pi_0 \circ \tilde{D}_0 \tilde{D}_1 f &= \pi_1 \circ \tilde{D}_1 f \circ (\pi_0 \& \text{id}) \quad \text{by [Proposition 26](#)} \\ &= f \circ (\text{id} \& \pi_1) \circ (\pi_0 \& \pi_1) \quad \text{by bilinearity of } f \\ &= f \circ (\pi_0 \& \pi_1)\end{aligned}$$

$$\begin{aligned}\pi_0 \circ \pi_1 \circ \tilde{D}_0 \tilde{D}_1 f &= \pi_1 \circ \tilde{D} \pi_0 \circ \tilde{D}_0 \tilde{D}_1 f \quad \text{by linearity of } \pi_0 \\ &= \pi_1 \circ \tilde{D}_0 f \circ (\text{id} \& \pi_0) \quad \text{by [Lemma 1](#)} \\ &= f \circ (\pi_1 \& \text{id}) \circ (\text{id} \& \pi_0) \quad \text{by bilinearity of } f \\ &= f \circ (\pi_1 \& \pi_0)\end{aligned}$$

Which concludes the proof. □

We can expand on the ideas of the proof [Lemma 1](#) to show the following result. This result is crucial, as it explains how to project on a series of partial derivatives.

Proposition 36. Let $n \geq 0$, $f \in \mathcal{C}(X_0 \& \cdots \& X_n)$, $d \geq 0$ and $i, i_1, \dots, i_d \in \llbracket 0, n \rrbracket$. Then,

$$\tilde{D}^d \pi_0 \circ \tilde{D}_{i_d} \dots \tilde{D}_{i_1} \tilde{D}_i f = \tilde{D}_{i_d} \dots \tilde{D}_{i_1} f \circ (\tilde{D}^{h_d(i)} \pi_0; \text{id}_{-i})$$

where $h_d(i) = \#\{k \in \llbracket 1, d \rrbracket \mid i_k = i\}$. Furthermore, if f is $(n + 1)$ -linear, then

$$\tilde{D}^d \pi_1 \circ \tilde{D}_{i_d} \dots \tilde{D}_{i_1} \tilde{D}_i f = \tilde{D}_{i_d} \dots \tilde{D}_{i_1} f \circ (\tilde{D}^{h_d(i)} \pi_1; \text{id}_{-i})$$

Proof. By induction on d . The case $d = 0$ is [Proposition 32](#) for π_0 , and the definition of n -linearity for π_1 . We deal with the inductive step for π_0 . The inductive step for π_1 is dealt with similarly.

$$\begin{aligned}\tilde{D}^{d+1} \pi_0 \circ \tilde{D}_{i_{d+1}} \dots \tilde{D}_{i_1} \tilde{D}_i f &= \tilde{D}(\tilde{D}^d \pi_0) \circ \tilde{D}(\tilde{D}_{i_d} \dots \tilde{D}_{i_1} \tilde{D}_i f) \circ \Phi^{i_{d+1}} \quad \text{by definition} \\ &= \tilde{D}(\tilde{D}^d \pi_0 \circ \tilde{D}_{i_d} \dots \tilde{D}_{i_1} \tilde{D}_i f) \circ \Phi^{i_{d+1}} \quad \text{by [\(D-chain\)](#)} \\ &= \tilde{D}(\tilde{D}_{i_d} \dots \tilde{D}_{i_1} \tilde{D}_i f \circ (\tilde{D}^{h_d(i)} \pi_0; \text{id}_{-i})) \circ \Phi^{i_{d+1}} \quad \text{by inductive hypothesis} \\ &= \tilde{D} \tilde{D}_{i_d} \dots \tilde{D}_{i_1} \tilde{D}_i f \circ \tilde{D}(\tilde{D}^{h_d(i)} \pi_0; \text{id}_{-i}) \circ \Phi^{i_{d+1}} \quad \text{by [\(D-chain\)](#)}$$

The next step is to use the naturality of $\Phi^{i_{d+1}}$:

$$\tilde{D}(f_0 \& \cdots \& f_n) \circ \Phi^{i_{d+1}} = (\tilde{D}f_{i_{d+1}}; f_{-i_{d+1}})$$

If $i_{d+1} = i$, then

$$\tilde{D}(\tilde{D}^{h_d(i)} \pi_0; \text{id}_{-i}) \circ \Phi^{i_{d+1}} = \Phi^{i_{d+1}} \circ (\tilde{D}^{h_d(i)+1} \pi_0; \text{id}_{-i})$$

If $i_{d+1} \neq i$ then

$$\tilde{D}(\tilde{D}^{h_d(i)}\pi_0; \text{id}_{-i}) \circ \Phi^{i_{d+1}} = \Phi^{i_{d+1}} \circ (\tilde{D}^{h_d(i)}\pi_0; \text{id}_{-i})$$

In both case,

$$\tilde{D}(\tilde{D}^{h_d(i)}\pi_0; \text{id}_{-i}) \circ \Phi^{i_{d+1}} = \Phi^{i_{d+1}} \circ (\tilde{D}^{h_{d+1}(i)}\pi_0; \text{id}_{-i})$$

Consequently:

$$\begin{aligned} \tilde{D}^{d+1}\pi_0 \circ \tilde{D}_{i_{d+1}} \dots \tilde{D}_{i_1} \tilde{D}_i f &= \tilde{D} \tilde{D}_{i_d} \dots \tilde{D}_{i_1} \tilde{D}_i f \circ \Phi^{i_{d+1}} \circ (\tilde{D}^{h_{d+1}(i)}\pi_0; \text{id}_{-i}) \\ &= \tilde{D}_{i_{d+1}} \tilde{D}_{i_d} \dots \tilde{D}_{i_1} \tilde{D}_i f \circ (\tilde{D}^{h_{d+1}(i)}\pi_0; \text{id}_{-i}) \end{aligned}$$

which concludes the proof. \square

This property instantiated in $d = 1$ gives back something similar to [Lemma 1](#).

Corollary 8. *If $f \in \mathcal{C}(X_0 \& \dots \& X_n)$ is $(n+1)$ -linear, then for any $i, j \in \llbracket 0, n \rrbracket$ such that $i \neq j$ and for any $k \in \{0, 1\}$,*

$$\tilde{D}\pi_k \circ \tilde{D}_i \tilde{D}_j f = \tilde{D}_i f \circ (\pi_k; \text{id}_{-j})$$

$$\tilde{D}\pi_k \circ \tilde{D}_i \tilde{D}_i f = \tilde{D}_i f \circ (\tilde{D}\pi_k; \text{id}_{-i})$$

We can use this corollary to show that the partial derivative of a $(n+1)$ -linear morphism is also $(n+1)$ -linear.

Theorem 6. *If $f \in \mathcal{C}(X_0 \& \dots \& X_n)$ is $(n+1)$ -linear, then for any $i \in \llbracket 0, n \rrbracket$, $\tilde{D}_i f$ is $(n+1)$ -linear.*

Proof. Let $j \in \llbracket 0, n \rrbracket$. The goal is to prove that $\pi_1 \circ \tilde{D}_j \tilde{D}_i f = \tilde{D}_i f \circ (\pi_1; \text{id}_{-j})$. By joint monicity of the π_k , it suffices to prove that $\pi_k \circ \pi_1 \circ \tilde{D}_j \tilde{D}_i f = \pi_k \circ \tilde{D}_i f \circ (\pi_1; \text{id}_{-j})$ for any $k \in \{0, 1\}$. If $i \neq j$,

$$\begin{aligned} \pi_k \circ \pi_1 \circ \tilde{D}_j \tilde{D}_i f &= \pi_1 \circ \tilde{D}\pi_k \circ \tilde{D}_j \tilde{D}_i f \quad \text{by } \tilde{D}\text{-linearity of } \pi_1 \\ &= \pi_1 \circ \tilde{D}_j f \circ (\pi_k; \text{id}_{-i}) \quad \text{by } \text{Corollary 8} \\ &= f \circ (\pi_1; \text{id}_{-j}) \circ (\pi_k; \text{id}_{-i}) \quad \text{since } f \text{ is } (n+1)\text{-linear} \\ &= f \circ (\pi_k; \text{id}_{-i}) \circ (\pi_1; \text{id}_{-j}) \quad \text{since } i \neq j \\ &= \pi_k \circ \tilde{D}_i f \circ (\pi_1; \text{id}_{-j}) \quad \text{since } f \text{ is } (n+1)\text{-linear} \end{aligned}$$

The case $i = j$ is very similar

$$\begin{aligned} \pi_k \circ \pi_1 \circ \tilde{D}_i \tilde{D}_i f &= \pi_1 \circ \tilde{D}\pi_k \circ \tilde{D}_i \tilde{D}_i f \quad \text{by } \tilde{D}\text{-linearity of } \pi_1 \\ &= \pi_1 \circ \tilde{D}_i f \circ (\tilde{D}\pi_k; \text{id}_{-i}) \quad \text{by } \text{Corollary 8} \\ &= f \circ (\pi_1; \text{id}_{-i}) \circ (\tilde{D}\pi_k; \text{id}_{-i}) \quad \text{since } f \text{ is } (n+1)\text{-linear} \\ &= f \circ (\pi_k; \text{id}_{-i}) \circ (\pi_1; \text{id}_{-i}) \quad \text{since } \pi_k \text{ is } \tilde{D}\text{-linear} \\ &= \pi_k \circ \tilde{D}_i f \circ (\pi_1; \text{id}_{-i}) \quad \text{since } f \text{ is } (n+1)\text{-linear.} \end{aligned}$$

\square

Composition with a linear morphism preserves multilinearity. Thus, the Leibniz rule ensures that if f is multilinear then $\tilde{D}f$ is also multilinear.

Proposition 37. *If $f \in \mathcal{C}(X_0 \& \cdots \& X_n, Y)$ is $(n + 1)$ -linear and $h \in \mathcal{C}(Y, Z)$ is linear, then $h \circ f$ is $(n + 1)$ -linear.*

Proof. This follows from a straightforward computation $\pi_1 \circ \tilde{D}_i(h \circ f) = \pi_1 \circ \tilde{D}(h \circ f) \circ \Phi^i = \pi_1 \circ \tilde{D}h \circ \tilde{D}f \circ \Phi^i = h \circ \pi_1 \circ \tilde{D}_i f = h \circ f \circ (\pi_1; \text{id}_{-i})$. \square

Theorem 7. *If $f \in \mathcal{C}(X_0 \& \cdots \& X_n, Y)$ is $(n + 1)$ -linear, then $\tilde{D}f \circ (c_{\&}^n)^{-1} \in \mathcal{C}(\tilde{D}X_0 \& \cdots \& \tilde{D}X_n, \tilde{D}Y)$ is also $(n + 1)$ -linear.*

Proof. By Leibniz (Proposition 33), $\tilde{D}f \circ (c_{\&}^n)^{-1} = \theta^n \circ \tilde{D}_{\alpha(n)} \cdots \tilde{D}_{\alpha(0)} f$. But the partial derivatives preserves multilinearity by Theorem 6 and composition by θ^n on the left preserves multilinearity by Proposition 37. \square

5 Kleisli category of the exponential of a model of LL

5.1 Coherent differentiation in a linear setting

Let \mathcal{L} be a symmetric monoidal closed category that is a model of LL, and more precisely a Seely category in the sense of [15]. We write the composition of $f \in \mathcal{L}(X, Y)$ with $g \in \mathcal{L}(Y, Z)$ as $g f$ to stress the intuition that the morphisms of \mathcal{L} are linear. The axioms of a Seely category include the existence of a cartesian product $\&$ and a comonad $(!, \text{der}, \text{dig})$ on \mathcal{L} , where $\text{der}_X \in \mathcal{L}(!X, X)$ and $\text{dig}_X \in \mathcal{L}(!X, !!X)$ are natural transformations. The *Kleisli category* $\mathcal{L}_!$ of this comonad is the category whose objects are the objects of \mathcal{L} and whose hom-sets are $\mathcal{L}_!(X, Y) = \mathcal{L}(!X, Y)$. Composition is defined in this category as $g \circ f = g ! f \text{ dig}$ and the identity at X is der_X , the unit of the comonad. It is well known that $\mathcal{L}_!$ is a cartesian (closed) category, with the same cartesian product $\&$ as \mathcal{L} .

The goal of this section is to show that coherent differentiation on \mathcal{L} as introduced in [5] in the setting of LL gives $\mathcal{L}_!$ a CCDC structure.

Theorem 8. *Any differential structure on a summable category \mathcal{L} (see [5]) induces a CCDC structure on $\mathcal{L}_!$.*

Let us detail first what the assumption means. The category \mathcal{L} is said to be *summable* [5] if it has a summability structure $(S, \pi_0, \pi_1, \sigma)$ in the sense of the first author. By Theorem 3, this means that $(S, \pi_0, \pi_1, \sigma)$ is a left summability structure in the sense of Definition 10 where every morphism is additive and the functorial action of S is given by $Sf := \langle\langle f \circ \pi_0, f \circ \pi_1 \rangle\rangle$. Then, we can define $\iota_i, \theta, !$ and c as usual⁹. The difference is that the additivity of every morphism ensures that those families are natural transformations for the functor S . In particular, (S, ι_0, θ) is *de facto* a monad. The category \mathcal{L} is said to be *summable as a cartesian category* if $\langle\text{Sp}_0, \text{Sp}_1\rangle = c_{\&}$ is an isomorphism¹⁰. Because every morphism of \mathcal{L} is additive, this corresponds by Corollary 7 to the fact that the cartesian product is compatible with the left summability structure as in Definition 17.

It is well known that there is a faithful functor $\text{Der} : \mathcal{L} \rightarrow \mathcal{L}_!$ which maps X to X and $f \in \mathcal{L}(X, Y)$ to $f \text{ der}_X \in \mathcal{L}_!(X, Y)$. We can show that this functor induces a left summability structure $(\tilde{D}, \text{Der}(\pi_0), \text{Der}(\pi_1), \text{Der}(\sigma))$ on $\mathcal{L}_!$ (where $\tilde{D}X := SX$) compatible with the cartesian product $\&$ of $\mathcal{L}_!$. The reason is that Der preserves monicity and additivity, thanks to the well known fact that $\text{Der}(h) \circ f = h f$. Finally, the definition of Der ensures

⁹Note that in [5], θ is called τ

¹⁰We can show that the condition required in [5] that $0 \in \mathcal{L}(S\top, \top)$ is an isomorphism always hold, using the joint monicity of the π_i

that $\text{Der}(\langle\langle f_0, f_1 \rangle\rangle) = \langle\langle \text{Der}(f_0), \text{Der}(f_1) \rangle\rangle$. In particular, the families of morphism generated by the Left Summability Structure $(\widetilde{\text{D}}, \text{Der}(\pi_0), \text{Der}(\pi_1), \text{Der}(\sigma))$ in [Definitions 6](#) and [12](#) to [14](#) are $\text{Der}(\iota_i)$, $\text{Der}(\theta)$, $\text{Der}(\text{l})$ and $\text{Der}(\text{c})$ respectively.

Then a *differential structure* on a summable category \mathcal{L} is a natural transformation $\partial_X \in \mathcal{L}(!SX, S!X)$ satisfying some equations called $(\partial\text{-chain})$, $(\partial\text{-local})$, $(\partial\text{-lin})$, $(\partial\text{-\&})$ and $(\partial\text{-Schwarz})$ (see [\[5\]](#)). The first axiom, $(\partial\text{-chain})$, is a compatibility condition of ∂ with regard to dig and der , making ∂ a *distributive law* between the functor S and the comonad $!_-$.

Definition 25. A *distributive law* between a functor $F : \mathcal{L} \rightarrow \mathcal{L}$ and the comonad $!_-$ on \mathcal{L} is a natural transformation $\lambda^F \in \mathcal{L}(!FX, F!X)$ such that the two following diagrams commute.

$$\begin{array}{ccc} !FX & \xrightarrow{\lambda_X^F} & F!X \\ \text{der}_{FX} \searrow & & \downarrow F\text{der}_X \\ & & FX \end{array} \quad \begin{array}{ccc} !FX & \xrightarrow{\lambda_X^F} & F!X \\ \text{dig}_{FX} \downarrow & & \downarrow F\text{dig}_X \\ !!FX & \xrightarrow{! \lambda_X^F} & !F!X \xrightarrow{\lambda_{!X}^F} & F!!X \end{array}$$

A definition of distributive laws can be found in [\[16\]](#), together with a proof of [Propositions 38](#) and [39](#) stated below (corollary 5.11 of [\[16\]](#))¹¹.

Proposition 38. *Let $F : \mathcal{C} \rightarrow \mathcal{C}$ be an endofunctor. There is a bijection between distributive laws $\lambda_F \in \mathcal{L}(!FX, F!X)$ and liftings¹² \hat{F} of F on $\mathcal{L}_!$. A lifting \hat{F} of F is a functor $\hat{F} : \mathcal{L}_! \rightarrow \mathcal{L}_!$ such that $\hat{F}X = FX$ and $\hat{F}(\text{Der}(h)) = \text{Der}(Fh)$.*

Proof. Given a distributive law $\lambda^F \in \mathcal{L}(!FX, F!X)$, one can define an extension mapping X to FX and $f \in \mathcal{L}_!(X, Y)$ to $F(f) \lambda_X^F \in \mathcal{L}_!(FX, FY)$. We can check that it is a functor using the diagrams of distributive laws, and a lifting of F using the naturality of λ^F . Conversely, any lifting \hat{F} of F induces a family $\lambda_X^F = \hat{F}id_X \in \mathcal{L}_!(!FX, F!X)$. The two diagrams of distributive law comes from the functoriality of \hat{F} and the naturality comes from the fact that \hat{F} is an extension of F . \square

Remark 5. Let $F, G : \mathcal{L} \rightarrow \mathcal{L}$ be two functors, with respective lifting \hat{F} and \hat{G} associated to the distributive laws $\lambda^F \in \mathcal{L}(!FX, F!X)$ and $\lambda^G \in \mathcal{L}(!GX, G!X)$. Then $\hat{G}\hat{F}$ is a lifting of GF and the distributive law associated with $\hat{G}\hat{F}$ is the following natural transformation: $\lambda_X^{GF} = G(\lambda_X^F) \lambda_{FX}^G \in \mathcal{L}(!GF X, GF!X)$.

The result below, proved in [\[16\]](#), is rather overlooked. While the proof is indeed quite simple, it provides a very interesting perspective on the idea of extending structure to a Kleisli (or similarly to an Eilenberg-Moore) category.

Proposition 39. *Let $F, G : \mathcal{L} \rightarrow \mathcal{L}$ be two endofunctors. Assume that \hat{F} and \hat{G} are lifting of F and G respectively, and let λ^F and λ^G be their respective associated distributive law. Let $\alpha_X \in \mathcal{L}(FX, GX)$ be a natural transformation. Then $\text{Der}(\alpha_X) \in \mathcal{L}_!(\hat{F}X, \hat{G}X)$ is natural if and only if the following diagram commutes.*

$$\begin{array}{ccc} !FX & \xrightarrow{\lambda^F} & F!X \\ !\alpha \downarrow & & \downarrow \alpha \\ !GX & \xrightarrow{\lambda^G} & G!X \end{array} \quad (3)$$

¹¹These observations are made in the more general setting of 2-categories

¹²The word ‘‘extension’’ is also used. We use the term lifting in order to stick to the terminology of [\[16\]](#)

Proof. straightforward computation. \square

In the case of differentiation, the axiom $(\partial\text{-chain})$ implies that $\partial \in \mathcal{C}(!SX, S!X)$ is a distributive law between the comonad $!$ and the functor S . This means that S can be lifted to an endofunctor \tilde{D} on $\mathcal{L}_!$. Besides, there is a trivial distributive law $\text{id}_{!X} \in \mathcal{L}(!X, !X)$ associated to the lifting of the identity functor on \mathcal{L} to the identity functor on $\mathcal{L}_!$. Then $(\partial\text{-local})$ is an instance of [Equation \(3\)](#) in which $F = S$, $G = \text{Id}$ and $\alpha = \pi_0$. This means that $(\partial\text{-local})$ holds if and only if $\text{Der}(\pi_0) \in \mathcal{L}_!(\tilde{D}X, X)$ is a natural transformation. Thus, $(\tilde{D}, \text{Der}(\pi_0), \text{Der}(\pi_1), \text{Der}(\sigma))$ is a pre-differential structure on $\mathcal{L}_!$ (in the sense of [Definition 11](#)) and $(D\text{-chain})$ holds.

Moreover, since \tilde{D} is a lifting of S , for any $h \in \mathcal{L}(X, Y)$, the morphism $\text{Der}(h) \in \mathcal{L}_!(X, Y)$ is \tilde{D} -linear. Indeed, $\text{Der}(\pi_0) \circ \tilde{D}(\text{Der}(h)) = \text{Der}(\pi_0) \circ \text{Der}(Sh) = \text{Der}(\pi_0 Sh) = \text{Der}(h \pi_0) = \text{Der}(h) \circ \text{Der}(\pi_0)$. As a result, $\text{Der}(\pi_i), \text{Der}(\sigma), \text{Der}(p_i)$ are all linear so $(D\text{proj-lin})$, $(D\text{sum-lin})$ hold and the pre-differential structure is compatible with the cartesian product.

Furthermore, $(\partial\text{-lin})$ consists of two instances of [Equation \(3\)](#). The first one is an instance in which $F = S$, $G = \text{Id}$ and $\alpha = \iota_0 \in \mathcal{L}(X, SX)$. The second one is an instance in which $F = S^2$, $G = S$ and $\alpha = \theta \in \mathcal{L}(S^2X, X)$. Indeed, as we saw in [Remark 5](#), there is a distributive law $S(\partial_X) \partial_{SX} \in \mathcal{C}(!S^2X, S^2!X)$ associated to \tilde{D}^2 , the lifting of S^2 to $\mathcal{L}_!$. So $(\partial\text{-lin})$ holds if and only if $\text{Der}(\iota_0) \in \mathcal{L}_!(\tilde{D}^2X, \tilde{D}X)$ and $\text{Der}(\theta) \in \mathcal{L}_!(\tilde{D}^2X, \tilde{D}X)$ are natural transformations, that is if and only if $(D\text{-add})$ hold¹³.

Finally, $(\partial\text{-Schwarz})$ consists of an instance of [Equation \(3\)](#) in which $F = S^2$, $G = S^2$ and $\alpha = c$. So $(\partial\text{-Schwarz})$ holds if and only if $\text{Der}(c) \in \mathcal{L}_!(\tilde{D}^2X, \tilde{D}^2X)$ is natural. The only lacking axiom is $(D\text{-lin})$ that corresponds to the naturality of $\text{Der}(!)$. Thanks to [Proposition 39](#), it would hold if and only if the diagram below commutes.

$$\begin{array}{ccc} !SX & \xrightarrow{\partial_X} & S!X \\ \Downarrow & & \downarrow h_X \\ !S^2X & \xrightarrow{\partial_{SX}} S!SX \xrightarrow{S\partial_X} & S^2!X \end{array} \quad (4)$$

This diagram is not mentioned in [\[5\]](#) but makes perfectly sense in the setting of coherent differentiation in LL and holds in all known LL models of coherent differentiation. The study of the consequences of this diagram is left for further work. This ends the proof of [Theorem 8](#).

Remark 6. The only remaining axiom is $(\partial\text{-\&})$ that deals with the Seely isomorphisms $m^n \in \mathcal{L}(!X_0 \otimes \dots \otimes !X_n, !(X_0 \& \dots \& X_n))$ of the Seely category \mathcal{L} . It is possible to define in LL a notion of multilinearity: given any $l \in \mathcal{L}(X_0 \otimes \dots \otimes X_n, Y)$, one can define $\tilde{l} \in \mathcal{L}_!(X_0 \& \dots \& X_n, Y)$ as $\tilde{l} = l(\text{der} \otimes \dots \otimes \text{der})(m^n)^{-1}$. Then a morphism in $\mathcal{L}_!(X_0 \& \dots \& X_n, Y)$ is $(n+1)$ -linear (in the sense of LL) if it can be written as \tilde{h} for some h . The axiom $(\partial\text{-\&})$ allows to show that any $(n+1)$ -linear morphism in the sense of LL is also $(n+1)$ -linear in the sense of [Definition 24](#). A proof of this fact can be implicitly found in Theorem 4.26 of [\[17\]](#). This is a crucial fact, because it shows that what really matters is the $(n+1)$ -linearity in terms of CCDC rather than the $(n+1)$ -linearity in terms of LL.

Many models of LL have a coherent differential structure, such as coherence spaces, non-uniform coherence spaces and probabilistic coherence spaces. Thus, their Kleisli categories

¹³As we saw, this gives to \tilde{D} the structure of a Monad on $\mathcal{L}_!$. In fact, $(\partial\text{-chain})$ and $(\partial\text{-lin})$ taken together make ∂ a distributive law between the monad \tilde{D} and the comonad $!$. There is a striking symmetry, because it also allows to lift $!$ to a comonad on \mathcal{L}_S the Kleisli category of S

are all CCDCs. This provides a rich variety of examples. We present here the example of probabilistic coherence spaces.

5.2 The example of probabilistic coherence spaces

A *probabilistic coherence space* (PCS) [18] is a pair $X = (|X|, \text{PX})$ where $|X|$ is a set and $\text{PX} \subseteq (\mathbb{R}_{\geq 0})^{|X|}$ satisfies $\text{PX} = \{x \in (\mathbb{R}_{\geq 0})^{|X|} \mid \forall x' \in \mathcal{P}' \langle x, x' \rangle := \sum_{a \in |X|} x_a x'_a \leq 1\}$ for some $\mathcal{P}' \subseteq (\mathbb{R}_{\geq 0})^{|X|}$ called a *predual* of X . To avoid ∞ coefficients it is also assumed that $\forall a \in |X| 0 < \sup_{x' \in \mathcal{P}'} x'_a < \infty$ and then it is easily checked that for all $\forall a \in |X| 0 < \sup_{x \in \text{PX}} x_a < \infty$.

A multiset of elements of a set I is a function $m : I \rightarrow \mathbb{N}$ such that the set $\text{supp}(m) = \{i \in I \mid m(i) \neq 0\}$ is finite. The set $\mathcal{M}_{\text{fin}}(I)$ of these multisets is the free commutative monoid generated by I . We use $[i_1, \dots, i_k]$ for the $m \in \mathcal{M}_{\text{fin}}(I)$ such that $m(i) = \#\{j \mid i_j = i\}$, for $i_1, \dots, i_k \in I$.

Given PCSs X and Y , a function $f : \text{PX} \rightarrow \text{PY}$ is *analytic*¹⁴ if there is a *matrix* $t \in (\mathbb{R}_{\geq 0})^{\mathcal{M}_{\text{fin}}(|X|) \times |Y|}$ such that, for all $x \in \text{PX}$ and $b \in |Y|$, one has $f(x)_b = \sum_{(m,b) \in \mathcal{M}_{\text{fin}}(|X|) \times |Y|} t_{m,b} x^m$ where $x^m = \prod_{a \in |X|} x_a^{m(a)}$. Thanks to the fact that all the coefficients in t are finite, it is not difficult to see that they can be recovered from the function f itself by means of iterated differentiation, see [18]. So an analytic function has *exactly one* associated matrix.

The identity function $\text{PX} \rightarrow \text{PX}$ is analytic (of matrix t given by $t_{m,a} = \delta_{m,[a]}$) and the composition of two analytic functions is still analytic. We use \mathbf{APcoh} for the category whose objects are PCSs and morphisms are analytic functions. For instance, if 1 is the PCS $(\{*\}, [0, 1])$ then $f_1, f_2 : [0, 1] \rightarrow [0, 1]$ given by $f_1(x) = 1 - \sqrt{1 - x^2}$ and $f_2(x) = e^{x-1}$ are in $\mathbf{APcoh}(1, 1)$, but $f_3(x) = 2x - x^2$ is not because of the negative coefficient. The (pointwise) sum of two analytic functions $\text{PX} \rightarrow \text{PY}$ is always well defined $\text{PX} \rightarrow \mathbb{R}_{\geq 0}^{|Y|}$, but is not necessarily in $\mathbf{APcoh}(X, Y)$ so \mathbf{APcoh} is not left-additive¹⁵.

If X is a PCS then $\tilde{\text{D}}X = (\{0, 1\} \times |X|, \text{P}(\tilde{\text{D}}X) = \{z \in (\mathbb{R}_{\geq 0})^{\{0,1\} \times |X|} \mid \pi_0(z) + \pi_1(z) \in \text{PX}\})$, where $\pi_i(z)_a = z_{i,a}$, is a PCS. Then $\pi_0, \pi_1 \in \mathbf{APcoh}(\tilde{\text{D}}X, X)$ and we have also $\sigma \in \mathbf{APcoh}(\tilde{\text{D}}X, X)$ given by $\sigma(z) = \pi_0(z) + \pi_1(z)$. In other words $\tilde{\text{D}}X$ is the PCS whose elements are the pairs $(x, u) \in \text{PX}^2$ such that $x + u \in \text{PX}$. In that way we have equipped \mathbf{APcoh} with a left pre-summability structure and the associated notion of summability is the obvious one: $f_0, f_1 \in \mathbf{APcoh}(X, Y)$ are summable if their pointwise sum $f_0 + f_1$ is in $\mathbf{APcoh}(X, Y)$ (the matrix of this sum is the sum of the matrices of f_0 and f_1). It is easily checked that this left pre-summability structure is a left summability structure (see Definition 10).

As explained in Section 2.1, differentiation boils down to extending the operation $\tilde{\text{D}}$ to morphisms in such a way that the conditions of Definition 15 be satisfied. Given $f \in$

¹⁴There is also a purely functional characterization of these functions as those which are totally monotone and Scott continuous, see [19]

¹⁵At least for this most natural addition.

$\mathbf{APcoh}(X, Y)$ of matrix t and $(x, u) \in \mathbf{P}(\tilde{\mathbf{D}}X)$ we have

$$\begin{aligned}
f(x+u) &= \sum_{(m,b) \in \mathcal{M}_{\text{fin}}(|X|) \times |Y|} t_{m,b} (x+u)^m \\
&= \sum_{(m,b) \in \mathcal{M}_{\text{fin}}(|X|) \times |Y|} t_{m,b} \sum_{p \leq m} \binom{m}{p} x^{m-p} u^p \\
&= f(x) + \sum_{a \in \text{supp}(m)} \binom{m}{[a]} x^{m-[a]} u_a + r(x, u) \\
&= f(x) + \sum_{a \in \text{supp}(m)} m(a) x^{m-[a]} u_a + r(x, u)
\end{aligned}$$

where $\binom{m}{p} = \prod_{a \in |X|} \binom{m(a)}{p(a)} \in \mathbb{N}$ when $p \leq m$ for the pointwise order. In these expressions the remainder $r(x, u)$ is a power series in x and u all of whose monomials have total degree > 1 in u (such as $x_a u_b u_c$ if $a, b, c \in |X|$). In particular $\|r(x, u)\| \in o(\|u\|)$ where $\|x\| = \sup\{\langle x, x' \rangle \mid x' \in \mathcal{P}'\} \in [0, 1]$ for any preduel of X (this norm does not depend on the choice of \mathcal{P}'). Using [Definition 11](#) we set

$$Df(x, u) = \sum_{a \in \text{supp}(m)} m(a) x^{m-[a]} u_a.$$

Since all coefficients of t are ≥ 0 we have $f(x) + Df(x, u) \leq f(x+u)$ for the pointwise order so that $\tilde{\mathbf{D}}f(x, u) = (f(x), Df(x, u)) \in \mathbf{P}(\tilde{\mathbf{D}}Y)$. In that way we have defined an analytic function $\tilde{\mathbf{D}}f \in \mathbf{APcoh}(\tilde{\mathbf{D}}X, \tilde{\mathbf{D}}Y)$ and it is easily checked that \mathbf{APcoh} is a coherent differential category in the sense of [Definition 15](#). For the two examples above we get $Df_2(x, u) = e^{x-1}u$ and $Df_1(x, u) = xu/\sqrt{1-x^2}$ which seems to be undefined when $x = 1$ but is not because then *we must have* $u = 0$ and so $Df_1(1, 0) = 0$.

An analytic $f \in \mathbf{APcoh}(X, Y)$ is *linear* if its matrix t satisfies that whenever $t_{m,b} \neq 0$, one has $m = [a]$ for some $a \in |X|$. This notion of linearity¹⁶ coincides with both additivity [Definition 3](#) and $\tilde{\mathbf{D}}$ -linearity [Definition 16](#).

The category \mathbf{APcoh} is cartesian, with $\top = (\emptyset, \{0\})$ and $X \& Y = (\{0\} \times |X| \cup \{1\} \times |Y|)$, $\{z \in (\mathbb{R}_{\geq 0})^{\{0\} \times |X| \cup \{1\} \times |Y|} \mid \mathbf{p}_0(z) \in \mathbf{P}X \text{ and } \mathbf{p}_1(z) \in \mathbf{P}Y\}$ which is easily seen to be a PCS (\mathbf{p}_i is defined exactly as π_i) such that $\mathbf{P}(X \& Y) = \mathbf{P}X \times \mathbf{P}Y$ up to a trivial bijection. The projections \mathbf{p}_i are additive, and $\mathbf{c}_{\&}$ (see [Section 4.1](#)) is an iso: if $((x, u), (y, v)) \in \mathbf{P}(\tilde{\mathbf{D}}X \& \tilde{\mathbf{D}}Y)$ then $((x, y), (u, v)) \in \mathbf{P}(\tilde{\mathbf{D}}(X \& Y))$ since $(x, y) + (u, v) = (x+u, y+v)$ so the summability structure is compatible with the cartesian product by [Corollary 7](#).

An $f \in \mathbf{APcoh}(X_0 \& X_1, Y)$ is *bilinear* in X_0, X_1 if it is linear (or additive) separately in both inputs, which is equivalent to saying that its matrix t satisfies that if $t_{m,b} \neq 0$ then $m = [(0, a_0), (1, a_1)]$ with $a_i \in |X_i|$ for $i = 0, 1$. Let $\mathbf{N} = (\mathbb{N}, \{x \in (\mathbb{R}_{\geq 0})^{\mathbb{N}} \mid \sum_{n \in \mathbb{N}} x_n \leq 1\})$ which represents the type of integers in \mathbf{APcoh} , then the function $h : \mathbf{APcoh}(\mathbf{N} \& \mathbf{N} \& \mathbf{N}, \mathbf{N})$ given by $h(u, x, y) = u_0 x + (\sum_{n=1}^{\infty} u_n) y$ is bilinear in $\mathbf{N}, \mathbf{N} \& \mathbf{N}$ and can be understood as an *ifzero* operator. The function $k \in \mathbf{APcoh}(\mathbf{N}, \mathbf{N})$ such that $k(x)_n = x_{n+1}$ is linear and represents the successor operation.

¹⁶Which arises from the fact that \mathbf{APcoh} is the Kleisli category of the comonad “!” on the PCS model of LL of [\[18\]](#).

6 Link with cartesian differential categories

We show in this section that CCDCs are a generalization of cartesian differential categories [3].

6.1 Cartesian left additive categories

We rely on the presentation of [20] for left additive categories, since this article uses a minimal set of assumptions.

Definition 26. A left additive category is a category such that each hom-set is a commutative monoid, with addition $+$ and zero 0 commuting with composition on the right, that is $(f + g) \circ h = f \circ h + g \circ h$ and $0 \circ f = 0$.

Definition 27. A morphism h is additive if addition is compatible with composition with h on the left, that is $h \circ (f + g) = h \circ f + h \circ g$ and $h \circ 0 = 0$. Note that the identity is additive, and additive morphisms are closed under addition and composition.

Definition 28. A cartesian left additive category is a left additive category such that the projections are additive.

Given a cartesian left additive category \mathcal{C} , one can define a summable pairing structure (Definition 1) $(\tilde{D}_{\&}, \mathfrak{p}_0, \mathfrak{p}_1, \mathfrak{p}_0 + \mathfrak{p}_1)$ with $\tilde{D}_{\&}X = X \& X$. Then one can check that all morphisms are summable (the witness of $f \boxplus g$ is $\langle f, g \rangle$). Moreover the left additivity of the category ensures that the notion of sum induced by $(\tilde{D}_{\&}, \mathfrak{p}_0, \mathfrak{p}_1, \mathfrak{p}_0 + \mathfrak{p}_1)$ coincides with the native structure of monoid on the hom-sets. In particular, a morphism is additive in the sense of Definition 3 if and only if it is additive in the sense of Definition 28. Consequently, $\mathfrak{p}_0, \mathfrak{p}_1$ and $\mathfrak{p}_0 + \mathfrak{p}_1$ are additive. Thus, $(\tilde{D}_{\&}, \mathfrak{p}_0, \mathfrak{p}_1, \mathfrak{p}_0 + \mathfrak{p}_1)$ is a left pre-summability structure. Finally, it is a left summability structure because $(\tilde{D}$ -witness) trivially holds (everything is summable), and $(\tilde{D}$ -zero), $(\tilde{D}$ -com) hold thanks to the fact that everything is summable and that $(\mathcal{C}(X, Y), +, 0)$ is a commutative monoid.

Conversely any left summability structure on \mathcal{C} of shape $(\tilde{D}_{\&}, \mathfrak{p}_0, \mathfrak{p}_1, \sigma)$ with $\tilde{D}_{\&}X = X \& X$ endows each hom-set with a commutative monoid structure and Proposition 2 ensures that the category is left additive. Then, as above, a morphism is additive in the sense of Definition 3 if and only if it is additive in the sense of Definition 28. Thus $\mathfrak{p}_0, \mathfrak{p}_1$ are additive so the category is cartesian left additive. Moreover $\sigma = \mathfrak{p}_0 + \mathfrak{p}_1$ by Proposition 1 so the left summability structure induced by the monoid on the hom-set coincides with the left summability structure we started from. We just proved Theorem 9 below.

Theorem 9. *Let \mathcal{C} be a cartesian category. Define $\tilde{D}_{\&}X = X \& X$. There is a bijection between the monoid structures on the hom-set that make \mathcal{C} a cartesian left additive category and the left summability structures $(\tilde{D}, \pi_0, \pi_1, \sigma)$ on \mathcal{C} such that $\tilde{D} = \tilde{D}_{\&}$, $\pi_0 = \mathfrak{p}_0$ and $\pi_1 = \mathfrak{p}_1$.*

Remark 7. Any left summability structure on \mathcal{C} of shape $(\tilde{D}_{\&}, \mathfrak{p}_0, \mathfrak{p}_1, \sigma)$ with $\tilde{D}_{\&}X = X \& X$ is *de facto* compatible with the cartesian product. The additivity of \mathfrak{p}_0 and \mathfrak{p}_1 is part of the axioms of summability, and $c_{\&}$ is an isomorphism thanks to Theorem 4 and the fact that everything is summable.

6.2 Cartesian differential categories

We give the axioms of a cartesian differential category following the alternative formulation of [7] for convenience.

Definition 29. A cartesian differential category is a cartesian left additive category \mathcal{C} equipped with a differential combinator D that maps each morphism $f \in \mathcal{C}(X, Y)$ to a morphism $Df \in \mathcal{C}(X \& X, Y)$ such that

- (1) $Dp_0 = p_0 \circ p_1$, $Dp_1 = p_1 \circ p_1$;
- (2) $D0 = 0$ and $D(f + g) = Df + Dg$;
- (3) $Did = \pi_1$ and $D(g \circ f) = Dg \circ \langle f \circ \pi_0, Df \rangle$;
- (4) $Df \circ \langle x, 0 \rangle = 0$ and $Df \circ \langle x, u + v \rangle = Df \circ \langle x, u \rangle + Df \circ \langle x, v \rangle$;
- (5) $DDf \circ \langle \langle x, 0 \rangle, \langle 0, u \rangle \rangle = Df \circ \langle x, u \rangle$;
- (6) $DDf \circ \langle \langle x, u \rangle, \langle v, w \rangle \rangle = DDf \circ \langle \langle x, v \rangle, \langle u, w \rangle \rangle$.

Note that the axiom $Did = p_1$ seems to be missing from the axioms given in [7], although it can be found in the original formulation in [3]. There is usually another axiom, that states that $D\langle f, g \rangle = \langle Df, Dg \rangle$. But as observed in [20], this axiom is a consequence of the linearity of the projections and of the chain rule so we discard it.

Let \mathcal{C} be a left additive category. As stated in [Theorem 9](#), the structure of monoid in the hom-set arises from a summability structure $(\tilde{D}_{\&}, p_0, p_1, p_0 + p_1)$ compatible with the cartesian product. Then, there is a bijection between pre-differential structures on top of this summability structure and differential combinators in the sense of [Definition 29](#): we can define the functorial action of $\tilde{D}_{\&}$ from D as $\tilde{D}_{\&}f := \langle f \circ p_0, Df \rangle$, and we can define D from $\tilde{D}_{\&}$ as $Df = p_1 \circ \tilde{D}_{\&}f$.

Besides, we have shown in [Section 3](#) that the axioms of coherent differentiation are equivalent to some equational properties on D . When the underlying left summability structure is $(\tilde{D}_{\&}, p_0, p_1, p_0 + p_1)$, those properties turn out to be exactly the axioms of cartesian differential categories. The axiom (1) corresponds to [\(Dproj-lin\)](#). By [Corollary 5](#), (2) corresponds to [\(Dsum-lin\)](#). By [Proposition 16](#), (3) corresponds to [\(D-chain\)](#). By [Propositions 18 and 19](#), (4) corresponds to [\(D-add\)](#). By [Proposition 20](#), (5) corresponds to [\(D-lin\)](#). By [Proposition 21](#), (6) corresponds to [\(D-Schwarz\)](#).

Finally, the differential structures on top of the left summability structure $(\tilde{D}_{\&}, p_0, p_1, \sigma)$ are *de facto* compatible with the cartesian product, because the linearity of p_0 and p_1 is included in (1).

Theorem 10. *The cartesian differential categories are exactly the cartesian coherent differential categories in which $\tilde{D}X = X \& X$, $\pi_0 = p_0$, $\pi_1 = p_1$.*

Remark 8. In [3], h is said to be linear if $D(h) = h \circ p_1$. Then [Proposition 10](#) ensures that this notion of linearity exactly corresponds through [Theorem 10](#) to our notion of \tilde{D} -linearity introduced in [Definition 16](#).

Remark 9. Every cartesian differential category is also a tangent category [7], and the tangent functor induced from D is exactly the same functor as $\tilde{D}_{\&}$. This makes sense, as coherent differentiation and tangent categories are very similar: they extend cartesian differential categories by generalizing addition in two different ways.

7 A first order coherent differential language

We introduce a first order language associated to these models. Note that a development of a whole coherent differential PCF of which our language can be roughly considered as a

fragment can already be found in [17], with a semantics based on [5]. Our main contribution here is that CCDCs provide the tools for a more principled and synthetic treatment of the semantics. This tighter connection between syntax and semantics allows for the development of new ideas, such as a more systematic treatment of multilinearity.

7.1 Terms

Definition 30. Let \mathcal{B} be a set of ground type symbols, ranged over by α, β, \dots . For any $\alpha \in \mathcal{B}$ and $h \in \mathbb{N}$, $\tilde{D}^h \alpha$ is a ground type. General types are inductively defined by

$$A, B, C := \tilde{D}^h \alpha \mid A \& B.$$

For any type A , we define the type $\tilde{D}A$ inductively on A by $\tilde{D}\tilde{D}^h \alpha = \tilde{D}^{h+1} \alpha$ and $\tilde{D}(A \& B) = \tilde{D}A \& \tilde{D}B$.

Definition 31. Let φ, ψ, \dots be function symbols. Each function symbol φ is uniquely assigned a *function type* of the form $A_0, \dots, A_n \rightarrow B$ where A_i and B are types. Then, $n + 1$ is called the arity of φ , denoted as $\text{ar}(\varphi)$.

A function symbol φ of type $A_0, \dots, A_n \rightarrow B$ will be interpreted in section [Section 7.2](#) as a $(n + 1)$ -linear morphisms $\llbracket \varphi \rrbracket \in \mathcal{C}(\llbracket A_0 \rrbracket \& \dots \& \llbracket A_n \rrbracket, \llbracket B \rrbracket)$ (recall [Definition 24](#)). Note that the types A_i can themselves be products and need not be ground types. For example, a 2-linear map in $\mathcal{C}((A \& B) \& C, D)$ can by no means be seen as a 3-linear map in $\mathcal{C}(A \& B \& C, D)$.

Definition 32. Define *functions* as

$$f, g, \dots := \varphi \mid \pi_i^A \mid \rho_i^{A,B} \mid \iota_i^A \mid \theta_n^A$$

where $i \in \{0, 1\}$, $n \geq 0$, φ are function symbols and A, B are types. Each function f has a function type: π_0^A, π_1^A have type $\tilde{D}A \rightarrow A$, ι_0^A, ι_1^A have type $A \rightarrow \tilde{D}A$, the θ_n^A have type $\tilde{D}^{n+1}A \rightarrow \tilde{D}A$ and $\rho_0^{A,B}, \rho_1^{A,B}$ have types $A \& B \rightarrow A$ and $A \& B \rightarrow B$ respectively. Notice that projections have arity 1 and not 2. The type attached to the constructors π_i, ρ_i, ι_i and θ_n will always be kept implicit in what follows.

Remark 10. Taking $n = -1$ allows to write constants.

Definition 33. Let \mathcal{V} be a set of variable symbols. The set Λ_1 of terms is defined inductively as follows

$$t, u, \dots := \langle t_0, t_1 \rangle \mid f^\zeta(t_0, \dots, t_n) \mid x$$

where $x \in \mathcal{V}$, f are function symbols of arity $n + 1$ and $\zeta \in \llbracket 0, n \rrbracket^*$, the set of finite words¹⁷ of elements of $\llbracket 0, n \rrbracket$.

Remark 11. Nothing prevents us from adding to this calculus non multilinear function symbols, assuming that the formal derivatives for the function symbols are also provided. We focus on multilinear functions though, due to the nature of the basic operations of PCF. A coherent differential PCF would contain a base type **nat**, two function symbols **pred** and **succ** of type **nat** \rightarrow **nat**, a family of function symbols **if** ^{A} of type **nat**, $A \& A \rightarrow A$ (conditional) and a family of function symbols **let** ^{A} of type **nat**, $(\mathbf{nat} \rightarrow A) \rightarrow A$ (call-by-value on the type of integers). An analysis of the semantics of these symbols in coherent differentiation in the LL setting of [17] or in the example of [Section 5.2](#) indeed shows that

¹⁷Such a word represents a successive application of partial derivatives on the multilinear symbol f , more on this in [Section 7.2](#).

$$\begin{array}{c}
\frac{x : A \in \Gamma}{\Gamma \vdash x : A} \text{ (Var)} \quad \frac{\Gamma \vdash t_0 : A \quad \Gamma \vdash t_1 : B}{\Gamma \vdash \langle t_0, t_1 \rangle : A \& B} \text{ (Pair)} \\
\\
\frac{f : A_0, \dots, A_n \rightarrow B \quad \zeta \in \llbracket 0, n \rrbracket^* \quad (\Gamma \vdash t_i : \tilde{D}^{|\zeta|_i} A_i)_{i=0}^n}{\Gamma \vdash f^\zeta(t_0, \dots, t_n) : \tilde{D}^{|\zeta|} B} \text{ (App)}
\end{array}$$

Figure 1: Typing rules

$$\begin{aligned}
\partial(x, y) &= \begin{cases} x & \text{if } y = x \\ \iota_0(y) & \text{otherwise} \end{cases} \\
\partial(x, \langle t_0, t_1 \rangle) &= \langle \partial(x, t_0), \partial(x, t_1) \rangle \\
\partial(x, f^\zeta(t_0, \dots, t_n)) &= \theta_n(f^{\zeta n \dots 10}(\partial(x, t_0), \dots, \partial(x, t_n)))
\end{aligned}$$

Figure 2: Differential of a term

`pred` and `succ` should be interpreted as linear morphisms, and that `ifA` and `letA` should be interpreted as 2-linear morphisms. Using the fact that variables can be used in a non-linear way as well as the PCF fixpoint operator, it is then possible to write terms whose interpretation is not multilinear. For instance, f_1 of [Section 5.2](#) is the semantics of a term, see [\[4\]](#).

Notations 5. For any word ζ , we write $|\zeta|$ for its length, and $|\zeta|_j$ for the number of occurrences of the letter j . We will write f for f^ϵ , where ϵ is the empty word. Notice that when $\text{ar}(f) = 0$, a word $\zeta \in \llbracket 0, 0 \rrbracket^*$ can be uniquely seen as an integer $d = |\zeta|$. We will then write $f^{(d)}$ for f^ζ .

We introduce the typing rules in [Figure 1](#). The systematic treatment of multilinear morphisms allows for a great factorization of the rules. We write $f : A_0, \dots, A_n \rightarrow B$ if f has type $A_0, \dots, A_n \rightarrow B$. Given any term t , one can define a term $\partial(x, t)$ by induction on t . The inductive steps are given in [Figure 2](#).

Proposition 40. *If $\Gamma, x : A \vdash t : B$ then $\Gamma, x : \tilde{D}A \vdash \partial(x, t) : \tilde{D}B$*

Proof. By induction on the typing derivation.

- If the last rule applied is (Var) then the first possibility is that $t = x$ and $\Gamma, x : A \vdash x : A$. But then, $\partial(x, x) = x$ and $\Gamma, x : \tilde{D}A \vdash x : \tilde{D}A$. The second possibility is that $t = y$ with $y \neq x$ and $\Gamma \vdash y : B$. But then, $\partial(x, y) = \iota_0(y)$ and $\Gamma \vdash \iota_0(y) : \tilde{D}B$. Thus, $\Gamma, x : \tilde{D}A \vdash \iota_0(y) : \tilde{D}B$ in both cases.
- If the last rule applied is (Pair), then $t = \langle t_0, t_1 \rangle$, t is of type $B_0 \& B_1$, $\Gamma, x : A \vdash t_0 : B_0$ and $\Gamma, x : A \vdash t_1 : B_1$. But $\partial(x, t) = \langle \partial(x, t_0), \partial(x, t_1) \rangle$. By induction hypothesis $\Gamma, x : \tilde{D}A \vdash \partial(x, t_0) : \tilde{D}B_0$ and $\Gamma, x : \tilde{D}A \vdash \partial(x, t_1) : \tilde{D}B_1$. Thus, by applying (Pair), $\Gamma, x : \tilde{D}A \vdash \langle \partial(x, t_0), \partial(x, t_1) \rangle : \tilde{D}B_0 \& \tilde{D}B_1$. But $\tilde{D}B_0 \& \tilde{D}B_1 = \tilde{D}(B_0 \& B_1)$ so $\Gamma, x : \tilde{D}A \vdash \partial(x, \langle t_0, t_1 \rangle) : \tilde{D}(B_0 \& B_1)$.
- If the last rule applied is (App) then $t = f^\zeta(t_0, \dots, t_n)$, f has some type $A_0, \dots, A_n \rightarrow B$, and $\Gamma, x : A \vdash t : \tilde{D}^{|\zeta|}B$. Besides, for any i , $\Gamma, x : A \vdash t_i : \tilde{D}^{|\zeta|_i}A_i$. By induction

hypothesis, $\Gamma, x : \tilde{D}A \vdash \partial(x, t_i) : \tilde{D}^{|\zeta|+1}A_i$. But $|\zeta n \dots 10|_i = |\zeta|_i + 1$ so applying the (App) rule gives a derivation for $\Gamma, x : \tilde{D}A \vdash f^{\zeta n \dots 10}(\partial(x, t_0), \dots, \partial(x, t_n)) : \tilde{D}^{|\zeta|+n+1}B$. Applying the (App) rule again for $f = \theta_n$ yields a derivation of $\Gamma, x : \tilde{D}A \vdash \theta_n(f^{\zeta n \dots 10}(\partial(x, t_0), \dots, \partial(x, t_n))) : \tilde{D}^{|\zeta|+1}B$, which concludes the proof. \square

7.2 Semantics

Let \mathcal{C} be a CCDC. For the sake of simplicity, we assume that $\tilde{D}(X \& Y) = \tilde{D}X \& \tilde{D}Y$ and $c_{\&} = \text{id}^{18}$. Assume that we are given an object $\llbracket \alpha \rrbracket$ of \mathcal{C} for any ground type symbol α . Then one can interpret any type as an object: $\llbracket \tilde{D}^h \alpha \rrbracket = \tilde{D}^h \llbracket \alpha \rrbracket$ and $\llbracket A \& B \rrbracket = \llbracket A \rrbracket \& \llbracket B \rrbracket$. It follows by a straightforward induction that $\llbracket \tilde{D}A \rrbracket = \tilde{D}\llbracket A \rrbracket$. This interpretation extends as usual to contexts, setting $\llbracket x_0 : A_0, \dots, x_n : A_n \rrbracket = \llbracket A_0 \rrbracket \& \dots \& \llbracket A_n \rrbracket$. The semantics of the empty context is \top .

Assume that we are given a $(n+1)$ -linear morphism $\llbracket \varphi \rrbracket \in \mathcal{C}(\llbracket A_0 \rrbracket \& \dots \& \llbracket A_n \rrbracket, \llbracket B \rrbracket)$ for any function symbol $\varphi : A_0, \dots, A_n \rightarrow B$. Then any function $f : A_0, \dots, A_n \rightarrow B$ can be interpreted as an $(n+1)$ -linear morphism $\llbracket f \rrbracket$ by setting $\llbracket \pi_i \rrbracket = \pi_i$, $\llbracket \iota_i \rrbracket = \iota_i$, $\llbracket \theta_n \rrbracket = \theta^n$ (as defined in [Definition 21](#)) and $\llbracket \mathbf{p}_i \rrbracket = \mathbf{p}_i$.

Remark 12. Since $c_{\&} = \text{id}$, we have $\tilde{D}\mathbf{p}_i = \tilde{D}\mathbf{p}_i \circ (c_{\&})^{-1} = \tilde{D}\mathbf{p}_i \circ \langle \pi_0 \& \pi_0, \pi_1 \& \pi_1 \rangle = \langle \mathbf{p}_i \circ (\pi_0 \& \pi_0), \mathbf{p}_i \circ (\pi_1 \& \pi_1) \rangle = \langle \pi_0 \circ \mathbf{p}_i, \pi_1 \circ \mathbf{p}_i \rangle = \mathbf{p}_i$. Notice also that $\langle \tilde{D}f_0, \tilde{D}f_1 \rangle = \tilde{D}\langle f_0, f_1 \rangle$ by [Proposition 24](#).

Theorem 11. *For any term t such that $\Gamma \vdash t : A$, we can define $\llbracket t \rrbracket_{\Gamma} \in \mathcal{C}(\llbracket \Gamma \rrbracket, \llbracket A \rrbracket)$.*

Proof. We proceed by induction on the term.

- If $t = x$ then the last typing rule must be (Var) so that $\Gamma = \Gamma_0, x : A, \Gamma_1$. Define $\llbracket x \rrbracket_{\Gamma} = \mathbf{p}_{|\Gamma_0|} \in \mathcal{C}(\llbracket \Gamma_0 \rrbracket \& \llbracket A \rrbracket \& \llbracket \Gamma_1 \rrbracket, \llbracket A \rrbracket)$.
- If $t = \langle t_0, t_1 \rangle$ then the last typing rule must be (Pair), so t is of type $A \& B$, $\Gamma \vdash t_0 : A$ and $\Gamma \vdash t_1 : B$. By induction, one can define $\llbracket t_0 \rrbracket_{\Gamma} \in \mathcal{C}(\llbracket \Gamma \rrbracket, \llbracket A \rrbracket)$ and $\llbracket t_1 \rrbracket_{\Gamma} \in \mathcal{C}(\llbracket \Gamma \rrbracket, \llbracket B \rrbracket)$. Then we define $\llbracket \langle t_0, t_1 \rangle \rrbracket_{\Gamma} = \langle \llbracket t_0 \rrbracket_{\Gamma}, \llbracket t_1 \rrbracket_{\Gamma} \rangle \in \mathcal{C}(\llbracket \Gamma \rrbracket, \llbracket A \& B \rrbracket)$.
- If $t = f^{\zeta}(t_0, \dots, t_n)$ with $f : A_0, \dots, A_n \rightarrow B$ then the last typing rule must be (App). That is, t must be of type $D^{|\zeta|}B$ for some type B and for $i = 0, \dots, n$ we have a derivation of $\Gamma \vdash t_i : \tilde{D}^{|\zeta|}A_i$. By inductive hypothesis, we can define $\llbracket t_i \rrbracket_{\Gamma} \in \mathcal{C}(\llbracket \Gamma \rrbracket, \llbracket D^{|\zeta|}A_i \rrbracket)$. But $\llbracket D^{|\zeta|}A_i \rrbracket = \tilde{D}^{|\zeta|} \llbracket A_i \rrbracket$ and $\tilde{D}_{\zeta_k} \dots \tilde{D}_{\zeta_1} \llbracket f \rrbracket \in \mathcal{C}(\tilde{D}^{|\zeta|} \llbracket A_0 \rrbracket \& \dots \& \tilde{D}^{|\zeta|} \llbracket A_n \rrbracket, \tilde{D}^{|\zeta|} \llbracket B \rrbracket)$. Thus, we can set $\llbracket f^{\zeta_1 \dots \zeta_k}(t_0, \dots, t_n) \rrbracket_{\Gamma} = \tilde{D}_{\zeta_k} \dots \tilde{D}_{\zeta_1} \llbracket f \rrbracket \circ \langle \llbracket t_0 \rrbracket_{\Gamma}, \dots, \llbracket t_n \rrbracket_{\Gamma} \rangle$.

\square

Notations 6. We use $\llbracket x \rrbracket_{\Gamma} = \mathbf{p}_x$ for the projection on $\llbracket \Gamma \rrbracket$ to the coordinate where x appears in Γ .

Remark 13. In particular, $\llbracket \pi_i^{(d)}(t) \rrbracket = \tilde{D}^d \pi_i \circ \llbracket t \rrbracket$, $\llbracket \iota_i^{(d)}(t) \rrbracket = \tilde{D}^d \iota_i \circ \llbracket t \rrbracket$, $\llbracket \theta_n^{(d)}(t) \rrbracket = \tilde{D}^d \theta^n \circ \llbracket t \rrbracket$. More importantly, $\llbracket \mathbf{p}_i^{(d)}(t) \rrbracket = \tilde{D}^d \mathbf{p}_i \circ \llbracket t \rrbracket = \mathbf{p}_i \circ \llbracket t \rrbracket$ because of our assumption that $c_{\&}$ is the identity.

¹⁸This assumption is by no mean necessary but it simplifies the notations and the results

Notations 7. For any word $\zeta = \zeta_1 \cdots \zeta_k$ in $\llbracket 0, n \rrbracket^k$, define $\tilde{D}_\zeta := \tilde{D}_{\zeta_k} \cdots \tilde{D}_{\zeta_1}$. Then for any $f \in \mathcal{C}(X_0 \& \dots \& X_n, Y)$, $\tilde{D}_\zeta f \in \mathcal{C}(\tilde{D}^{|\zeta|_0} X_0 \& \cdots \& \tilde{D}^{|\zeta|_n} X_n, \tilde{D}^{|\zeta|} Y)$. Note that $\tilde{D}_{\zeta \cdot \delta} = \tilde{D}_\delta \tilde{D}_\zeta$. Then, [Proposition 36](#) can be seen as the property that for any f $(n+1)$ -linear, for any word δ of length d , $\tilde{D}^d \pi_i \circ \tilde{D}_\delta \tilde{D}_j f = \tilde{D}_\delta f \circ (\tilde{D}^{|\delta|_j} \pi_i; \text{id}_{-j})$

The main result of this section on the calculus consists in showing that the semantics of this syntactical derivative operation corresponds to the derivative in the model.

Theorem 12. *If $\Gamma, x : A \vdash t : B$ then $\llbracket \partial(x, t) \rrbracket_{\Gamma, x : \tilde{D}A} = \tilde{D}_1 \llbracket t \rrbracket_{\Gamma, x : A}$ where $\llbracket t \rrbracket_{\Gamma, x : A}$ is seen as a morphisms of $\mathcal{C}(\llbracket \Gamma \rrbracket \& \llbracket A \rrbracket, \llbracket B \rrbracket)$.*

Proof. By induction on t .

- If $t = x$ then $\llbracket t \rrbracket_{\Gamma, x : A} = \mathbf{p}_1 \in \mathcal{C}(\llbracket \Gamma \rrbracket \& \llbracket A \rrbracket, \llbracket A \rrbracket)$. Then $\tilde{D}_1 \mathbf{p}_1 = \tilde{D} \mathbf{p}_1 \circ \Phi^1 = \tilde{D} \mathbf{p}_1 \circ \langle \text{id} \& \pi_0, 0 \& \pi_1 \rangle = \langle \mathbf{p}_1 \circ (\text{id} \& \pi_0), \mathbf{p}_1 \circ (0 \& \pi_1) \rangle = \langle \pi_0 \circ \mathbf{p}_1, \pi_1 \circ \mathbf{p}_1 \rangle = \mathbf{p}_1$ using [Proposition 25](#) and the linearity of \mathbf{p}_1 .
- If $t = y \neq x$ then $\llbracket t \rrbracket_{\Gamma, x : A} = \llbracket y \rrbracket_\Gamma \circ \mathbf{p}_0 = \mathbf{p}_y \circ \mathbf{p}_0 \in \mathcal{C}(\llbracket \Gamma \rrbracket \& \llbracket A \rrbracket, \llbracket B \rrbracket)$. Then $\tilde{D}_1 (\mathbf{p}_y \circ \mathbf{p}_0) = \tilde{D} \mathbf{p}_y \circ \tilde{D} \mathbf{p}_0 \circ \Phi^1 = \tilde{D} \mathbf{p}_y \circ \tilde{D} \mathbf{p}_0 \circ \langle \text{id} \& \pi_0, 0 \& \pi_1 \rangle = \tilde{D} \mathbf{p}_y \circ \langle \mathbf{p}_0 \circ (\text{id} \& \pi_0), \mathbf{p}_0 \circ (0 \& \pi_1) \rangle = \tilde{D} \mathbf{p}_y \circ \langle \mathbf{p}_0, 0 \rangle = \langle \mathbf{p}_y \circ \mathbf{p}_0, 0 \rangle = \llbracket \iota_0(y) \rrbracket = \llbracket \partial(x, y) \rrbracket$.
- If $t = \langle t_0, t_1 \rangle$, then $\llbracket \partial(x, t) \rrbracket = \llbracket \langle \partial(x, t_0), \partial(x, t_1) \rangle \rrbracket = \langle \llbracket \partial(x, t_0) \rrbracket, \llbracket \partial(x, t_1) \rrbracket \rangle$. By inductive hypothesis, $\llbracket \partial(x, t) \rrbracket = \langle \tilde{D}_1 \llbracket t_0 \rrbracket, \tilde{D}_1 \llbracket t_1 \rrbracket \rangle$. But $\langle \tilde{D}_1 \llbracket t_0 \rrbracket, \tilde{D}_1 \llbracket t_1 \rrbracket \rangle = \langle \tilde{D} \llbracket t_0 \rrbracket \circ \Phi^1, \tilde{D} \llbracket t_1 \rrbracket \circ \Phi^1 \rangle = \langle \tilde{D} \llbracket t_0 \rrbracket, \tilde{D} \llbracket t_1 \rrbracket \rangle \circ \Phi^1$. By [Remark 12](#), this is equal to $\tilde{D} \langle \llbracket t_0 \rrbracket, \llbracket t_1 \rrbracket \rangle \circ \Phi^1 = \tilde{D}_1 \langle \llbracket t_0 \rrbracket, \llbracket t_1 \rrbracket \rangle = \tilde{D}_1 \llbracket t \rrbracket$.
- If $t = f^\zeta(t_0, \dots, t_n)$ then by definition $\partial(x, t) = \theta_n(f^{\zeta n \cdots 10}(\partial(x, t_0), \dots, \partial(x, t_n)))$. Thus, $\llbracket \partial(x, t) \rrbracket = \theta^n \circ \tilde{D}_{n \cdots 10} \tilde{D}_\zeta f \circ \langle \llbracket \partial(x, t_0) \rrbracket, \dots, \llbracket \partial(x, t_n) \rrbracket \rangle = \theta^n \circ \tilde{D}_{n \cdots 10} \tilde{D}_\zeta f \circ \langle \tilde{D}_1 \llbracket t_0 \rrbracket, \dots, \tilde{D}_1 \llbracket t_n \rrbracket \rangle$ by inductive hypothesis. But then, the Leibniz rule ([Proposition 33](#)) states that $\theta^n \circ \tilde{D}_{n \cdots 10} \tilde{D}_\zeta f = \tilde{D} \tilde{D}_\zeta f$. Thus, $\llbracket \partial(x, t) \rrbracket = \tilde{D} \tilde{D}_\zeta f \circ \langle \tilde{D} \llbracket t_0 \rrbracket \circ \Phi^1, \dots, \tilde{D} \llbracket t_n \rrbracket \circ \Phi^1 \rangle = \tilde{D} \tilde{D}_\zeta f \circ \langle \tilde{D} \llbracket t_0 \rrbracket, \dots, \tilde{D} \llbracket t_n \rrbracket \rangle \circ \Phi^1 = \tilde{D} (\tilde{D}_\zeta f \circ \langle \llbracket t_0 \rrbracket, \dots, \llbracket t_n \rrbracket \rangle) \circ \Phi^1 = \tilde{D} \llbracket t \rrbracket \circ \Phi^1 = \tilde{D}_1 \llbracket t \rrbracket$.

□

7.3 Reduction

We introduce in this section a set of reduction rules that deals with the differential content of the terms. The set of rules is more compact than the one given in [\[17\]](#), but covers all of the rules concerning the fragment we are looking at.

Remark 14. We could have added a construct $t[u/x]$ for explicit substitutions, with the typing rule

$$\frac{\Gamma, x : A \vdash t : B \quad \Gamma \vdash u : A}{\Gamma \vdash t[u/x] : B} (\text{Cut})$$

as well as reduction rules that performs the substitution steps (for example, $x[u/x] \rightarrow_{\Lambda_1} u$). We decided not to do so because, in a higher order λ -calculus setting, such explicit substitutions are not necessary.

The main difference with the differential lambda-calculus of [\[21\]](#) is the absence of sum, because we do not want a non deterministic typing rules such as

$$\frac{\Gamma \vdash t : A \quad \Gamma \vdash u : A}{\Gamma \vdash t + u : A}$$

But the reduction of a π_1 against a θ will introduce sums. Handling sum without the typing rule above is tricky, because of subject reduction. There will be no guarantee indeed that if $\Gamma \vdash t + u : A$ and $t \rightarrow_{\Lambda_1} t'$ then $\Gamma \vdash t' + u : A$. For this reason, we chose a conservative approach, by keeping sums as a formal multiset on top of the terms.

Definition 34. A term multiset is a finite multiset of term.

See [Section 5.2](#) for the notations we use on multisets. We define a reduction \rightarrow_{Λ_1} from terms to term multisets. The reduction rules are given in [Figure 3](#). Then we define $\rightarrow_{\Lambda_1}^?$ as the “reflexive” closure of \rightarrow_{Λ_1} . That is, $t \rightarrow_{\Lambda_1}^? L$ if $t \rightarrow_{\Lambda_1} L$ or if $L = [t]$. It allows to lift \rightarrow_{Λ_1} to a reduction from a term multiset to a term multiset in a monadic fashion: if $t_1 \rightarrow_{\Lambda_1} L_1$ and for all $i \neq 1$, $t_i \rightarrow_{\Lambda_1}^? L_i$, then

$$[t_1, \dots, t_n] \rightarrow_{\mathcal{M}(\Lambda_1)} \sum_{i=1}^n L_i$$

where \sum is the multiset union, that is, the pointwise sum of the functions $L_i : \Lambda_1 \rightarrow \mathbb{N}$.

$$\begin{aligned} & \mathbf{p}_i^{(d)}(\langle t_0, t_1 \rangle) \rightarrow_{\Lambda_1} [t_i] \\ & \pi_i^{(d)}(f^{\zeta j \delta}(t_0, \dots, t_n)) \rightarrow_{\Lambda_1} [f^{\zeta \delta}(t_0, \dots, \pi_i^{(|\delta|j)}(t_j), \dots, t_n)] \quad \text{where } |\delta| = d \\ & \pi_i^{(d)}(\iota_i^{(d)}(t)) \rightarrow_{\Lambda_1} [t] \\ & \pi_i^{(d)}(\iota_{1-i}^{(d)}(t)) \rightarrow_{\Lambda_1} [] \\ & \pi_0^{(d)}(\theta_n^{(d)}(t)) \rightarrow_{\Lambda_1} [(\pi_0^{(d)})^{n+1}(t)] \\ & \pi_1^{(d)}(\theta_n^{(d)}(t)) \rightarrow_{\Lambda_1} \sum_{k=0}^n [(\pi_0^{(d)})^k \pi_1^{(d)} (\pi_0^{(d)})^{n-k}(t)]. \end{aligned}$$

Here, $(\pi_i^{(d)})^n$ is a notation for n successive applications of $\pi_i^{(d)}$.

Figure 3: Reduction rules

Definition 35. A term multiset $[t_1, \dots, t_n]$ of type A in context Γ is \mathcal{C} -summable if $\llbracket t_1 \rrbracket_{\Gamma}, \dots, \llbracket t_n \rrbracket_{\Gamma}$ are summable (in the sense of [Theorem 2](#)). Then, we define $\llbracket [t_1, \dots, t_n] \rrbracket_{\Gamma} = \llbracket t_1 \rrbracket_{\Gamma} + \dots + \llbracket t_n \rrbracket_{\Gamma}$. Note that $[]$ is always \mathcal{C} -summable, and $\llbracket [] \rrbracket = 0$.

The main point of coherent differentiation is that the reduction \rightarrow_{Λ_1} will always introduce term multisets that are \mathcal{C} -summable, for any model \mathcal{C} .

Theorem 13 (Invariance of semantics under reduction). *For any $\Gamma \vdash t : A$, if $t \rightarrow_{\Lambda_1} L$ then L is \mathcal{C} -summable and $\llbracket L \rrbracket_{\Gamma} = \llbracket t \rrbracket_{\Gamma}$.*

Proof. Let us consider every application of the rule \rightarrow_{Λ_1} . Note that when a term multiset has one element, it is always \mathcal{C} -summable and $\llbracket [t] \rrbracket = \llbracket t \rrbracket$.

$$\begin{aligned} \llbracket \mathbf{p}_i^{(d)}(\langle t_0, t_1 \rangle) \rrbracket &= \tilde{D}^d \mathbf{p}_i \circ \langle \llbracket t_0 \rrbracket, \llbracket t_1 \rrbracket \rangle \\ &= \mathbf{p}_i \circ \langle \llbracket t_0 \rrbracket, \llbracket t_1 \rrbracket \rangle \\ &= \llbracket t_i \rrbracket. \end{aligned}$$

The rule below is the one where most of the differential content appears. Recall that $\llbracket f \rrbracket$ is assumed to be multilinear, for any function f . It implies that $\tilde{D}_\zeta \llbracket f \rrbracket$ is also multilinear by [Theorem 6](#), so it is possible to apply [Proposition 36](#) on it.

$$\begin{aligned}
& \llbracket \pi_i^{(d)}(f^{\zeta j \delta}(t_0, \dots, t_n)) \rrbracket \\
&= \tilde{D}^d \pi_i \circ \tilde{D}_\delta \tilde{D}_j \tilde{D}_\zeta \llbracket f \rrbracket \circ \langle \llbracket t_0 \rrbracket, \dots, \llbracket t_n \rrbracket \rangle \\
&= \tilde{D}_\delta \tilde{D}_\zeta \llbracket f \rrbracket \circ (\tilde{D}^{|\delta|_j} \pi_i; \text{id}_{-j}) \circ \langle \llbracket t_0 \rrbracket, \dots, \llbracket t_n \rrbracket \rangle \quad \text{by [Proposition 36](#)} \\
&= \tilde{D}_{\zeta \delta} \llbracket f \rrbracket \circ \langle \llbracket t_0 \rrbracket, \dots, \tilde{D}^{|\delta|_j} \pi_i \circ \llbracket t_j \rrbracket, \dots, \llbracket t_n \rrbracket \rangle \\
&= \llbracket f^{\zeta \delta}(t_0, \dots, \pi_i^{(|\delta|_j)}(t_j), \dots, t_n) \rrbracket.
\end{aligned}$$

The three next rules are rather standard and are consequence of the definition of π_i , ι_j and θ^n .

$$\begin{aligned}
\llbracket \pi_i^{(d)}(\iota_i^{(d)}(t)) \rrbracket &= \tilde{D}^d \pi_i \circ \tilde{D}^d \iota_i \circ \llbracket t \rrbracket \\
&= \tilde{D}^d(\pi_i \circ \iota_i) \circ \llbracket t \rrbracket \quad \text{by [\(D-chain\)](#)} \\
&= \tilde{D}^d \text{id} \circ \llbracket t \rrbracket = \llbracket t \rrbracket \quad \text{by [\(D-chain\)](#)}
\end{aligned}$$

$$\begin{aligned}
\llbracket \pi_i^{(d)}(\iota_{1-i}^{(d)}(t)) \rrbracket &= \tilde{D}^d \pi_i \circ \tilde{D}^d \iota_{1-i} \circ \llbracket t \rrbracket \\
&= \tilde{D}^d(\pi_i \circ \iota_{1-i}) \circ \llbracket t \rrbracket \quad \text{by [\(D-chain\)](#)} \\
&= \tilde{D}^d 0 \circ \llbracket t \rrbracket = 0 \circ \llbracket t \rrbracket \quad \text{by [\(Dsum-lin\)](#)} \\
&= 0 = \llbracket \llbracket \rrbracket \rrbracket
\end{aligned}$$

$$\begin{aligned}
\llbracket \pi_0^{(d)}(\theta_n^{(d)}(t)) \rrbracket &= \tilde{D}^d \pi_0 \circ \tilde{D}^d \theta^n \circ \llbracket t \rrbracket \\
&= \tilde{D}^d(\pi_0 \circ \theta^n) \circ \llbracket t \rrbracket \quad \text{by [\(D-chain\)](#)} \\
&= \tilde{D}^d(\pi_0^{n+1}) \circ \llbracket t \rrbracket \\
&= (\tilde{D}^d \pi_0)^{n+1} \circ \llbracket t \rrbracket \quad \text{by [\(D-chain\)](#)} \\
&= \llbracket (\pi_0^{(d)})^{n+1}(t) \rrbracket.
\end{aligned}$$

The last rule is where finite multisets of size greater than 1 are introduced. Most lines in the following sequence of equations should be understood as follows: “the sum above is well defined, so the sum below is well defined and they are equal”.

$$\begin{aligned}
\llbracket \pi_1^{(d)}(\theta_n^{(d)}(t)) \rrbracket &= \tilde{D}^d \pi_1 \circ \tilde{D}^d \theta^n \circ \llbracket t \rrbracket \\
&= \tilde{D}^d(\pi_1 \circ \theta^n) \circ \llbracket t \rrbracket \quad \text{by [\(D-chain\)](#)} \\
&= \left(\tilde{D}^d \left(\sum_{k=0}^n \pi_0^k \circ \pi_1 \circ \pi_0^{n-k} \right) \right) \circ \llbracket t \rrbracket \\
&= \left(\sum_{k=0}^n \tilde{D}^d(\pi_0^k \circ \pi_1 \circ \pi_0^{n-k}) \right) \circ \llbracket t \rrbracket \quad \text{by [\(Dsum-lin\)](#) and [Proposition 17](#)} \\
&= \sum_{k=0}^n \tilde{D}^d(\pi_0^k \circ \pi_1 \circ \pi_0^{n-k}) \circ \llbracket t \rrbracket \quad \text{by [Proposition 2](#)}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{k=0}^n (\tilde{D}^d \pi_0)^k \circ \tilde{D}^d \pi_1 \circ (\tilde{D}^d \pi_0)^{n-k} \circ \llbracket t \rrbracket \quad \text{by (D-chain)} \\
&= \sum_{k=0}^n \llbracket (\pi_0^{(d)})^k \pi_1^{(d)} (\pi_0^{(d)})^{n-k} (t) \rrbracket.
\end{aligned}$$

Thus, $\sum_{k=0}^n \llbracket (\pi_0^{(d)})^k \pi_1^{(d)} (\pi_0^{(d)})^{n-k} (t) \rrbracket$ is \mathcal{C} -summable of semantics $\llbracket \pi_1^{(d)} (\theta_n^{(d)}(t)) \rrbracket$. \square

Corollary 9. *For any term multiset $\Gamma \vdash L : A$ that is \mathcal{C} -summable, if $L \rightarrow_{\mathcal{M}(\Lambda_1)} L'$ then L' is \mathcal{C} -summable and $\llbracket L' \rrbracket_{\Gamma} = \llbracket L \rrbracket_{\Gamma}$.*

Proof. Assume that $[t_1, \dots, t_n]$ is \mathcal{C} -summable and that $[t_1, \dots, t_n] \rightarrow_{\mathcal{M}(\Lambda_1)} L$. That is, for any i , $t_i \rightarrow_{\Lambda_1}^? [t_i^1, \dots, t_i^{k_i}]$ and $L = \sum_{i=1}^n [t_i^1, \dots, t_i^{k_i}]$. Then by [Theorem 13](#), for any i , $\llbracket t_i^1 \rrbracket, \dots, \llbracket t_i^{k_i} \rrbracket$ are summable of sum $\llbracket t_i \rrbracket$. By assumption, $\llbracket t_1 \rrbracket, \dots, \llbracket t_n \rrbracket$ are summable, that is, $\sum_{j=1}^{k_1} \llbracket t_1^j \rrbracket, \dots, \sum_{j=1}^{k_n} \llbracket t_n^j \rrbracket$ are summable. By [Theorem 2](#), it means that the family $\llbracket t_1^1 \rrbracket, \dots, \llbracket t_1^{k_1} \rrbracket, \dots, \llbracket t_n^1 \rrbracket, \dots, \llbracket t_n^{k_n} \rrbracket$ is summable of sum

$$\sum_{i=1}^n \sum_{j=1}^{k_i} \llbracket t_i^j \rrbracket = \sum_{i=1}^n \llbracket t_i \rrbracket$$

Thus L is \mathcal{C} -summable and $\llbracket L \rrbracket = \llbracket [t_1, \dots, t_n] \rrbracket$. \square

The usage of such term multisets may seem somewhat non deterministic. But any multiset generated by reductions of the calculus can be interpreted as a summable family in deterministic models such as probabilistic coherence spaces¹⁹ (see [Section 5.2](#)) or non uniform coherence spaces. This determinism of the models allows to prove in [\[17\]](#) a result that roughly state that whenever a closed term of *type integer* reduces to a term multiset $C + \llbracket \nu \rrbracket$ (where ν are the usual integer variables of PCF), then $\llbracket C \rrbracket = 0$. That is, only one of the branches of the reduction rule

$$\pi_1^{(d)} (\theta_n^{(d)}(t)) \rightarrow_{\Lambda_1} \sum_{k=0}^n \llbracket (\pi_0^{(d)})^k \pi_1^{(d)} (\pi_0^{(d)})^{n-k} (t) \rrbracket$$

produces a non empty multiset. The proof relies on the fact that any term of type integer will be interpreted in **APcoh** as a Dirac distribution δ_n on \mathbb{N} or as the zero distribution, because the calculus does not feature any form of probabilistic branching. Thus, a term multiset of type integer is **APcoh**-summable if and only if there is at most one term in the multiset whose semantic is not 0. In particular, $\llbracket \nu \rrbracket^{\mathbf{APcoh}} = \delta_\nu$ and $C + \llbracket \nu \rrbracket$ is **APcoh**-summable (by [Corollary 9](#)) so $\llbracket C \rrbracket^{\mathbf{APcoh}} = 0$. One can also use non-uniform coherence spaces for proving the same result in a similar way. This observation led to the development of a completely deterministic Krivine Machine for a coherent differential version of PCF in [\[17\]](#), extending the projections path with a writable memory structure.

Conclusion

We have introduced and studied a general categorical framework for coherent differentiation, a new approach to the differential calculus which does not require the ambient category to be (left-)additive. We have also proposed some basic syntactical constructs accounting

¹⁹Probabilistic branching is by no mean a form of non determinism

in a term language for these new categorical constructs. These are the foundations for a principled and systematic approach to the denotational semantics of functional programming languages like (probabilistic) PCF extended with coherent differentiation. As shown in [17] such an extension can perfectly feature general recursive definitions as well as deterministic or probabilistic behaviors, in sharp contrast with the Differential λ -calculus [22] which is inherently non-deterministic. Accordingly, the next step will be to specialize the present general axiomatization to the case where the category is cartesian closed.

Acknowledgment

We thank the reviewers for their careful reading and helpful comments. This work was partly supported by the ANR project *Probabilistic Programming Semantics (PPS)* ANR-19-CE48-0014.

References

- [1] J. Girard, “Linear logic,” *Theoretical Computer Science*, vol. 50, pp. 1–102, 1987. [Online]. Available: [https://doi.org/10.1016/0304-3975\(87\)90045-4](https://doi.org/10.1016/0304-3975(87)90045-4)
- [2] R. Blute, R. Cockett, and R. Seely, “Differential categories,” *Mathematical Structures in Computer Science*, vol. 16, pp. 1049 – 1083, 12 2006.
- [3] —, “Cartesian differential categories,” *Theory and Applications of Categories*, vol. 22, pp. 622–672, 01 2009.
- [4] T. Ehrhard, “Differentials and distances in probabilistic coherence spaces,” *Logical Methods in Computer Science*, vol. 18, no. 3, 2022. [Online]. Available: [https://doi.org/10.46298/lmcs-18\(3:2\)2022](https://doi.org/10.46298/lmcs-18(3:2)2022)
- [5] —, “Coherent differentiation,” *Mathematical Structures in Computer Science*, p. 1–52, 2023.
- [6] J. Rosicky, “Abstract tangent functors,” *Diagramme*, no. 12, 1984.
- [7] R. Cockett and G. Cruttwell, “Differential Structure, Tangent Structure, and SDG,” *Applied Categorical Structures*, vol. 22, 04 2014.
- [8] M. A. Arbib and E. G. Manes, “Partially additive categories and flow-diagram semantics,” *Journal of Algebra*, vol. 62, no. 1, pp. 203–227, 1980. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0021869380902124>
- [9] P. Hines, “A categorical analogue of the monoid semiring construction,” *Mathematical Structures in Computer Science*, vol. 23, no. 1, p. 55–94, 2013.
- [10] A. Kock, “Strong functors and monoidal monads,” *Archiv der Mathematik*, vol. 23, pp. 113–120, 12 1972.
- [11] E. Moggi, “Notions of computation and monads,” *Information and Computation*, vol. 93, no. 1, pp. 55–92, 1991, selections from 1989 IEEE Symposium on Logic in Computer Science. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0890540191900524>

- [12] M. Aguiar, M. Haim, and I. Franco, “Monads on higher monoidal categories,” *Applied Categorical Structures*, vol. 26, 06 2018.
- [13] I. Moerdijk, “Monads on tensor categories,” *Journal of Pure and Applied Algebra*, vol. 168, no. 2, pp. 189–208, 2002, category Theory 1999: selected papers, conference held in Coimbra in honour of the 90th birthday of Saunders Mac Lane. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0022404901000962>
- [14] A. Kock, “Monads on symmetric monoidal closed categories,” *Archiv der Mathematik*, vol. 21, pp. 1–10, 01 1970.
- [15] P.-A. Melliès, “Categorical semantics of linear logic,” *Panoramas et Synthèses*, vol. 27, pp. 1–196, 2009.
- [16] J. Power and H. Watanabe, “Combining a monad and a comonad,” *Theoretical Computer Science*, vol. 280, no. 1, pp. 137–162, 2002, coalgebraic Methods in Computer Science. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S030439750100024X>
- [17] T. Ehrhard, “A coherent differential PCF,” 2022. [Online]. Available: <https://arxiv.org/abs/2205.04109>
- [18] V. Danos and T. Ehrhard, “Probabilistic coherence spaces as a model of higher-order probabilistic computation,” *Information and Computation*, vol. 152, no. 1, pp. 111–137, 2011.
- [19] R. Crubillé, “Probabilistic Stable Functions on Discrete Cones are Power Series,” in *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018*, A. Dawar and E. Grädel, Eds. ACM, 2018, pp. 275–284. [Online]. Available: <https://doi.org/10.1145/3209108.3209198>
- [20] J.-S. P. Lemay, “A tangent category alternative to the Faa di Bruno construction,” *Theory and Applications of Categories*, vol. 33, no. 35, pp. 1072–1110, 2018. [Online]. Available: <http://www.tac.mta.ca/tac/volumes/33/35/33-35abs.html>
- [21] T. Ehrhard and L. Regnier, “The differential lambda-calculus,” *Theoretical Computer Science*, vol. 309, no. 1, pp. 1–41, 2003. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S030439750300392X>
- [22] —, “The differential lambda-calculus,” *Theoretical Computer Science*, vol. 309, no. 1-3, pp. 1–41, 2003.