CARTESIAN DIFFERENCE CATEGORIES

MARIO ALVAREZ-PICALLO [a] AND JEAN-SIMON PACAUD LEMAY [b]

[a] Department of Computer Science, University of Oxford
Current address: Edinburgh Research Center, Huawei Technologies Research & Development
e-mail address: mario.alvarez-picallo@cs.ox.ac.uk

[b] Department of Computer Science, University of Oxford
e-mail address: jean-simon.lemay@cs.ox.ac.uk

ABSTRACT. Cartesian differential categories are categories equipped with a differential combinator which axiomatizes the directional derivative. Important models of Cartesian differential categories include classical differential calculus of smooth functions and categorical models of the differential λ-calculus. However, Cartesian differential categories cannot account for other interesting notions of differentiation of a more discrete nature such as the calculus of finite differences. On the other hand, change action models have been shown to capture these examples as well as more "exotic" examples of differentiation. But change action models are very general and do not share the nice properties of Cartesian differential categories. In this paper, we introduce Cartesian difference categories as a bridge between Cartesian differential categories and change action models. We show that every Cartesian differential category is a Cartesian difference category, and how certain well-behaved change action models are Cartesian difference categories. In particular, Cartesian difference categories model both the differential calculus of smooth functions and the calculus of finite differences. Furthermore, every Cartesian difference category comes equipped with a tangent bundle monad whose Kleisli category is again a Cartesian difference category.

ACKNOWLEDGMENT

1. Introduction

In the early 2000's, Ehrhard and Regnier introduced the differential λ-calculus [ER03], an extension of the λ-calculus equipped with a differential combinator capable of taking the derivative of arbitrary higher-order functions. This development, based on models of linear logic equipped with a natural notion of "derivative" [Ehr18], sparked a wave of research into categorical models of differentiation. One of the most notable developments in the area are Cartesian differential categories [BCS09], introduced by Blute, Cockett and Seely, which provide an abstract categorical axiomatization of the directional derivative from differential calculus. The relevance of Cartesian differential categories lies in their ability to model both "classical" differential calculus (with the canonical example being the category of Euclidean spaces and smooth functions between) and the differential λ-calculus (as every categorical model for it gives rise to a Cartesian differential category [Man12a]). However, while Cartesian differential categories have proven to be an immensely successful formalism, they have, by design, some limitations. Firstly, they cannot account for certain "exotic" notions of derivative, such as the difference operator from the calculus of finite differences [Ric54]. This is because the axioms of a Cartesian differential category stipulate that derivatives should be linear in their second argument (in the same way that the directional derivative is), whereas these aforementioned discrete sorts of derivatives need not be. Additionally, every Cartesian differential category is equipped with a tangent bundle monad [CC14, Man12b] whose Kleisli category can be intuitively understood as a category of generalized vector fields.

More recently, discrete derivatives have been suggested as a semantic framework for understanding incremental computation. This led to the development of change structures [CGRO14] and change actions [APO19]. Change action models have been successfully used to provide a model for incrementalizing Datalog programs [APETJO19], and have also been shown to model the calculus of finite differences. Change action models, however, are very general, lacking many of the nice properties of Cartesian differential categories (for example, addition in a change action model is not required to be commutative), even though they are verified in most change action models. As a consequence of this generality, the tangent bundle endofunctor in a change action model can fail to be a monad.

In this work, we introduce Cartesian difference categories (Section 4.2), whose key ingredients are an infinitesimal extension operator and a difference combinator, whose axioms are a generalization of the differential combinator axioms of a Cartesian differential category. In Section 4.3, we show that every Cartesian differential category is a Cartesian difference category whose infinitesimal extension operator is zero, and conversely how every Cartesian difference category admits a full subcategory which is a Cartesian differential category. In Section 4.4, we show that every Cartesian difference category is a change action model, and conversely how a full subcategory of suitably well-behaved objects and maps of a change action model is a Cartesian difference category. In Section 5 we provide some examples of Cartesian difference categories; notably, the calculus of finite differences and the stream calculus. In Section 6, we show that every Cartesian difference category comes equipped with a monad whose Kleisli category (Section 6.3) and a certain full sub-category of its Eilenberg-Moore category (Section 6.2) are again both Cartesian difference categories. Lastly, in Section 7, we briefly discuss difference λ-categories, which are Cartesian difference categories that are Cartesian closed. We conclude with a discussion on future work regarding Cartesian difference categories.

This paper is an extended version of a conference paper [APL20] for Foundations of Software Science and Computation Structures: 23rd International Conference (FOSSACS 2020). While the overall structure and story of both versions are very similar, this version includes additional details, proofs, and results, as well as some important corrections. We highlight the main differences between this paper and the conference paper. Probably the most important change to point out is the correction made in Section 6.3. In this version, we correct the Cartesian difference structure of the Kleisli category of the tangent bundle monad. The proposed infinitesimal extension and difference combinator for said Kleisli category in the conference paper [APL20] was based on a result from another paper [APO19]. Unfortunately, we have found that the result in said other paper is incorrect, and therefore so were the proposed infinitesimal extension and difference combinator for the Kleisli category in the conference paper. Section 2 and Section 3, the background sections, are mostly the same with some added background details. In Section 4, more detailed proofs and computations are provided throughout, and both Lemma 4.8 and Lemma 4.9 are new results. In Section 4.3, we give a specific name to objects whose infinitesimal extensions is zero, which in Definition 4.11 we call ε-vanishing. In Section 4.4 we introduce flat objects, whose definition (Definition 4.15) is slightly more general than the original one. Section 5, the example section, is mostly unchanged with some added details and proofs. In Section 6, we provide full detailed computations for all the necessary structure relating to the tangent bundle monad. Section 6.2 and Section 7 are both new sections.

Conventions:

In an arbitrary category, we use the classical notation for composition as opposed to diagrammatic order which was used in [BCS09]: this means that the composite map $g \circ f : A \to C$ is the map which first does $f : A \to B$ then $g : B \to C$. By a Cartesian category we mean a category $\mathbb{X}$ with chosen finite products where we denote the binary product of objects A and B by $A \times B$ with projection maps $\pi_0 : A \times B \to A$ and $\pi_1 : A \times B \to B$ and pairing operation $\langle -, - \rangle$, so the product of maps is $f \times g = \langle f \circ \pi_0, g \circ \pi_1 \rangle$, and the chosen terminal object as $\top$ with unique terminal maps $!_A : A \to \top$.

2. Cartesian Differential Categories

In this section, we very briefly review Cartesian differential categories, so that the reader may compare Cartesian differential categories with the new notion of Cartesian *difference* categories which we introduce in Section 4.2. For a full detailed introduction on Cartesian differential categories, we refer the reader to the original paper [BCS09].

2.1. **Cartesian Left Additive Categories.** Here we recall the definition of Cartesian left additive categories [BCS09], which is the underlying structure of both Cartesian differential categories and Cartesian difference categories. Here "left additive" is meant being skew enriched [Cam18] over commutative monoids, so in particular, this means that we do not assume the existence of additive inverses, i.e., "negative elements". This left additive structure allows one to sum maps and have zero maps while allowing for maps that do not preserve the sum or zero. Maps that do preserve the additive structure are called additive maps.

Definition 2.1 [BCS09, Definition 1.1.1]. A **left additive category** is a category $\mathbb{X}$ such that each hom-set $\mathbb{X}(A, B)$ is a commutative monoid with addition operation $+ :$ $\mathbb{X}(A, B) \times \mathbb{X}(A, B) \to \mathbb{X}(A, B)$ and zero element (called the zero map) $0 \in \mathbb{X}(A, B)$, such that pre-composition preserves the additive structure:

$$(f + g) \circ h = f \circ h + g \circ h \qquad\qquad 0 \circ f = 0$$

A map k is **additive** if post-composition by k preserves the additive structure:

$$k \circ (f + g) = k \circ f + k \circ g \qquad\qquad k \circ 0 = 0$$

Definition 2.2 [Lem18, Definition 2.3]. A **Cartesian left additive category** is a Cartesian category $\mathbb{X}$ which is also a left additive category such all projection maps $\pi_0 : A \times B \to A$ and $\pi_1 : A \times B \to B$ are additive.

Examples of Cartesian left additive categories can be found in Section 5. We note that the definition given here of a Cartesian left additive category is slightly different from the one found in [BCS09, Definition 1.2.1], which also assumed that the pairing of additive maps be additive, but it is indeed equivalent as explained in [Lem18]. Furthermore, by [BCS09, Proposition 1.2.2], an equivalent axiomatization of a Cartesian left additive category is that of a Cartesian category where every object comes equipped with a commutative monoid structure such that the projection maps are monoid morphisms, which we discuss further in Lemma 4.4. Here are some useful identities on how the additive and product structures are compatible:

Lemma 2.3 [Lem18, Lemma 2.4]. *In a Cartesian left additive category $\mathbb{X}$:*

(i) *$\langle f, g \rangle + \langle h, k \rangle = \langle f + h, g + k \rangle$ and $\langle 0, 0 \rangle = 0$;*
(i) *If $f : C \to A$ and $g : C \to B$ are additive then $\langle f, g \rangle : C \to A \times B$ is additive;*
(i) *If $h : A \to B$ and $k : C \to D$ are additive then $h \times k : A \times C \to B \times D$ is additive;*
(i) *For any object A, the unique map to the terminal object $\top$ is the zero map, $!_A = 0$.*

2.2. **Cartesian Differential Categories.** Cartesian differential categories are Cartesian left additive categories that come equipped with a differential combinator, which in turn is axiomatized by the basic properties of the directional derivative from multivariable differential calculus. In the following definition, it is important to note that here, following the more recent work on Cartesian differential categories, we've flipped the convention found in [BCS09], so that the linear argument is in the second argument rather than in the first argument.

Definition 2.4 [BCS09, Definition 2.1.1]. A **Cartesian differential category** is a Cartesian left additive category equipped with a **differential combinator** D of the form

$$\frac{f : A \to B}{\mathsf{D}[f] : A \times A \to B}$$

verifying the following coherence conditions:

[**CD.1**] $\mathsf{D}[f + g] = \mathsf{D}[f] + \mathsf{D}[g]$ and $\mathsf{D}[0] = 0$
[**CD.2**] $\mathsf{D}[f] \circ \langle x, y + z \rangle = \mathsf{D}[f] \circ \langle x, y \rangle + \mathsf{D}[f] \circ \langle x, z \rangle$ and $\mathsf{D}[f] \circ \langle x, 0 \rangle = 0$
[**CD.3**] $\mathsf{D}[1_A] = \pi_1$ and $\mathsf{D}[\pi_0] = \pi_0 \circ \pi_1$ and $\mathsf{D}[\pi_1] = \pi_1 \circ \pi_1$
[**CD.4**] $\mathsf{D}[\langle f, g \rangle] = \langle \mathsf{D}[f], \mathsf{D}[g] \rangle$ and $\mathsf{D}[!_A] = !_{A \times A}$
[**CD.5**] $\mathsf{D}[g \circ f] = \mathsf{D}[g] \circ \langle f \circ \pi_0, \mathsf{D}[f] \rangle$

[CD.6] $\mathsf{D}\left[\mathsf{D}[f]\right] \circ \langle\langle x, y\rangle, \langle 0, z\rangle\rangle = \mathsf{D}[f] \circ \langle x, z\rangle$
[CD.7] $\mathsf{D}\left[\mathsf{D}[f]\right] \circ \langle\langle x, y\rangle, \langle z, 0\rangle\rangle = \mathsf{D}\left[\mathsf{D}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle$

As we will see in Section 4.3, a Cartesian difference category is a generalization of a Cartesian differential category. Therefore, we leave the discussion of the intuition of these axioms for later in Section 4.2 below. We also refer to [BCS09, Section 4] for a term calculus which may help better understand the axioms of a Cartesian differential category. We highlight that the last two axioms **[CD.6]** and **[CD.7]** can equivalently be expressed as follows:

Lemma 2.5 [CC14, Proposition 4.2]. *In the presence of the other axioms,* **[CD.6]** *and* **[CD.7]** *are equivalent to the following two axioms:*

[CD.6.a] $\mathsf{D}\left[\mathsf{D}[f]\right] \circ \langle\langle x, 0\rangle, \langle 0, y\rangle\rangle = \mathsf{D}[f] \circ \langle x, y\rangle$
[CD.7.a] $\mathsf{D}\left[\mathsf{D}[f]\right] \circ \langle\langle x, y\rangle, \langle z, w\rangle\rangle = \mathsf{D}\left[\mathsf{D}[f]\right] \circ \langle\langle x, z\rangle, \langle y, w\rangle\rangle$

The canonical example of a Cartesian differential category is the category of real smooth functions, which we will discuss in Section 5.1. Other interesting examples of Cartesian differential categories can be found throughout the literature such as categorical models of the differential λ-calculus [ER03, Man12a], the subcategory of differential objects of a tangent category [CC14], and the coKleisli category of a differential category [BCS06, BCS09].

3. Change Action Models

Change actions [APETJO19, APO19] have recently been proposed as a setting for reasoning about higher-order incremental computation, based on a discrete notion of differentiation. Together with Cartesian differential categories, they provide the core ideas behind Cartesian difference categories. In this section, we quickly review change actions and change action models, in particular, to highlight where some of the axioms of a Cartesian difference category come from. For more details on change actions, we invite readers to see the original paper [APO19].

3.1. **Change Actions.** The basic intuition for a change action is an object A equipped with a monoid ΔA which represents the possible "changes" or "updates" that might be applied to A, with the monoid structure on ΔA representing the capability to compose updates.

Definition 3.1 [APO19, Section 2]. A **change action** $\overline{A}$ in a Cartesian category $\mathbb{X}$ is a quintuple:

$$\overline{A} \equiv (A, \Delta A, \oplus_A, +_A, 0_A)$$

consisting of two objects A and ΔA, and three maps:

$$\oplus_A : A \times \Delta A \to A \qquad +_A : \Delta A \times \Delta A \to \Delta A \qquad 0_A : \top \to \Delta A$$

such that:

[CA.1] $(\Delta A, +_A, 0_A)$ is a monoid, that is, the following equalities hold:

$$+_A \circ \langle 1_{\Delta A}, 0_{\Delta A} \circ !_{\Delta A}\rangle = 1_{\Delta A} = +_A \circ \langle 0_{\Delta A} \circ !_{\Delta_A}, 1_{\Delta A}\rangle$$

$$+_A \circ (1_{\Delta A} \times +_A) = +_A \circ \langle +_A \circ (1_{\Delta A} \times \pi_0), \pi_1 \circ \pi_1\rangle$$

[**CA.2**] $\oplus_A : A \times \Delta A \to A$ is an action of ΔA on A, that is, the following equalities hold:

$$\oplus_A \circ \langle 1_A, 0_A \circ !_A \rangle = 1_A$$

$$\oplus_A \circ (1_A \times +_A) = \oplus_A \circ \langle \oplus_A \circ (1_A \times \pi_0), \pi_1 \circ \pi_1 \rangle$$

In [**CA.1**], the first line says that 0_A is a unit for $+_A$ while the second line is the associativity of $+_A$. In [**CA.2**], the first line is the unit law of the module action, while the second line is the associativity law of the module action. These are highlighted in Lemma 3.2 below. If $\overline{A} \equiv (A, \Delta A, \oplus_A, +_A, 0_A)$ is a change action, then for a pair of parallel maps $h : C \to \Delta A$ and $k : C \to \Delta A$, we define their ΔA-sum $h +_{\overline{A}} k : C \to \Delta A$ as follows:

$$h +_{\overline{A}} k = +_A \circ \langle h, k \rangle$$

Similarly, for a pair of maps $f : C \to A$ and $g : C \to \Delta A$, we define $f \oplus_{\overline{A}} g : C \to A$ as:

$$f \oplus_{\overline{A}} g = \oplus_A \circ \langle f, g \rangle$$

Here, $f \oplus_{\overline{A}} g$ should be thought of as g acting on f.

Lemma 3.2. *In a Cartesian category* $\mathbb{X}$, *let* $\overline{A} \equiv (A, \Delta A, \oplus_A, +_A, 0_A)$ *be a change action. Then for any suitable maps:*

   (i) $f +_{\overline{A}} (0_A \circ !_C) = f = (0_A \circ !_C) +_{\overline{A}} f$
   (ii) $f +_{\overline{A}} (g +_{\overline{A}} h) = (f +_{\overline{A}} g) +_{\overline{A}} h$
   (iii) $f \oplus_{\overline{A}} (0_A \circ !_C) = f$
   (iv) $f \oplus_{\overline{A}} (g +_{\overline{A}} h) = (f \oplus_{\overline{A}} g) \oplus_{\overline{A}} h$
   (v) $(f \oplus_{\overline{A}} g) \circ h = (f \circ h) \oplus_{\overline{A}} (g \circ h)$

Proof. We leave these as a simple exercises for the reader. Briefly, (i) and (ii) will follow from [**CA.1**], while (iii) and (iv) will follow from [**CA.2**], and (v) follows directly from the definition. $\qquad\square$

Change actions give rise to a notion of derivative, with a distinctly "discrete" flavour. Given change actions structure on A and B, a map $f : A \to B$ can be said to be differentiable when changes to the input (in the sense of elements of ΔA) are mapped to changes to the output (that is, elements of ΔB). In the setting of incremental computation [APETJO19, CGRO14, KPS16], this is precisely what it means for f to be incrementalizable, with the derivative of f corresponding to an incremental version of f.

Definition 3.3 [APO19, Definition 1]**.** Let

$$\overline{A} \equiv (A, \Delta A, \oplus_A, +_A, 0_A) \qquad \text{and} \qquad \overline{B} \equiv (B, \Delta B, \oplus_B, +_B, 0_B)$$

be change actions in a Cartesian category $\mathbb{X}$. For a map $f : A \to B$, a map of type $\partial[f] : A \times \Delta A \to \Delta B$ is a **derivative** of f whenever the following equalities hold:

[**CAD.1**] $f \circ (x \oplus_{\overline{A}} y) = (f \circ x) \oplus_{\overline{B}} (\partial[f] \circ \langle x, y \rangle)$
[**CAD.2**] $\partial[f] \circ \langle x, y +_{\overline{A}} z \rangle = (\partial[f] \circ \langle x, y \rangle) +_{\overline{B}} (\partial[f] \circ \langle x \oplus_{\overline{A}} y, z \rangle)$ and
           $\partial[f] \circ \langle x, 0_B \circ !_B \rangle = 0_B \circ !_{A \times \Delta A}$

The second axiom [**CAD.2**] is also known as regularity [APO19, Definition 2]. The intuition for these axioms will be explained in more detail in Section 4.2 when we explain the axioms of a Cartesian difference category. It is important to note that there is nothing in the definition that says that derivatives are necessarily unique, and therefore a map f could have multiple possible derivatives.

3.2. **Change Action Models.** The chain rule for derivatives still holds for change actions, or in other words, differentiation is compositional. Indeed by [APO19, Lemma 1], whenever $\boldsymbol{\partial}[f]$ and $\boldsymbol{\partial}[g]$ are derivatives for composable maps f and g respectively, then $\boldsymbol{\partial}[g] \circ \langle f \circ \pi_0, \boldsymbol{\partial}[f]\rangle$ is a derivative for $g \circ f$. As a corollary, change actions in $\mathbb{X}$ form a category where the maps are pairs consisting a map f and a derivative of f. For more details on this category of change actions, see [APO19, Section 3].

Definition 3.4 [APO19, Section 3]**.** Given a Cartesian category $\mathbb{X}$, define its change actions category $\mathsf{CAct}(\mathbb{X})$ as the category whose objects are change actions $\overline{A}$ in $\mathbb{X}$ and whose maps $\overline{f} : \overline{A} \to \overline{B}$ are the pairs $(f, \boldsymbol{\partial}[f])$, where $f : A \to B$ is a map in $\mathbb{X}$ and $\boldsymbol{\partial}[f] : A \times \Delta A \to \Delta B$ is a derivative for f. The identity is $(1_A, \pi_1)$, while composition of $(f, \boldsymbol{\partial}[f])$ and $(g, \boldsymbol{\partial}[g])$ is $(g \circ f, \boldsymbol{\partial}[g] \circ \langle f \circ \pi_0, \boldsymbol{\partial}[f]\rangle)$.

For a Cartesian category $\mathbb{X}$, it is straightforward to see that $\mathsf{CAct}(\mathbb{X})$ is also a Cartesian category where the terminal object and the product of objects is the same as in $\mathbb{X}$, where the projection maps are the pairs $(\pi_i, \pi_i \circ \pi_1)$, and where the pairing of maps is given point-wise $\langle (f, \boldsymbol{\partial}[f]), (g, \boldsymbol{\partial}[g])\rangle = (\langle f, g\rangle, \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g]\rangle)$. There is an obvious product-preserving forgetful functor $\mathcal{E} : \mathsf{CAct}(\mathbb{X}) \to \mathbb{X}$ sending every change action $(A, \Delta A, \oplus, +, 0)$ to its base object A and every map $(f, \boldsymbol{\partial}[f])$ to the underlying map f. Here is a useful lemma which describes the compatibility between change action structure and the product structure.

Lemma 3.5. *Let* $\overline{A} \equiv (A, \Delta A, \oplus_A, +_A, 0_A)$ *and* $\overline{B} \equiv (B, \Delta B, \oplus_B, +_B, 0_B)$ *be change actions in a Cartesian category* $\mathbb{X}$*. Then for any suitable maps:*

   (i) $\langle f, g\rangle \oplus_{\overline{A \times B}} \langle h, k\rangle = \langle f \oplus_{\overline{A}} h, g \oplus_{\overline{B}} k\rangle$
  (ii) $\langle f, g\rangle +_{\overline{A \times B}} \langle h, k\rangle = \langle f +_{\overline{A}} h, g +_{\overline{B}} k\rangle$
 (iii) $\langle 0_A, 0_B\rangle = 0_{A \times B}$

Proof. These are immediate from the definition of product structure in $\mathsf{CAct}(\mathbb{X})$, and so we leave this as excercise for the reader. $\qquad\square$

As a setting for studying differentiation, the category $\mathsf{CAct}(\mathbb{X})$ is rather lacklustre, since there is no notion of higher derivatives. Instead, one works with change action models. Informally, a change action model consists of a rule which for every object A of $\mathbb{X}$ associates a change action over it, and for every map a choice of a derivative.

Definition 3.6 [APO19, Definition 5]**.** A **change action model** is a Cartesian category $\mathbb{X}$ with a product-preserving functor $\alpha : \mathbb{X} \to \mathsf{CAct}(\mathbb{X})$ that is a section of the forgetful functor $\mathcal{E}$, that is, $\mathcal{E} \circ \alpha = 1_{\mathbb{X}}$.

For brevity, when A is an object of a change action model, we will simply write its associate change action as $\alpha(A) = (A, \Delta A, \oplus_A, +_A)$. Examples of change action models can be found in [APO19]. In particular, as was shown in [APO19, Theorem 6], every Cartesian differential category provides a change model action. We will generalize this result, and show in Section 4.4 that a Cartesian difference category also always provides a change action model.

4. Cartesian Difference Categories

In this section, we introduce *Cartesian difference categories*, which are generalizations of Cartesian differential categories. Examples of Cartesian difference categories can be found in Section 5.

4.1. **Infinitesimal Extensions in Left Additive Categories.** We first introduce infinitesimal extensions, which is an operator that turns a map into an "infinitesimal" version of itself, in the sense that every map coincides with its Taylor approximation on infinitesimal elements.

Definition 4.1. A Cartesian left additive category $\mathbb{X}$ is said to have an **infinitesimal extension** ε if every pair of objects A and B, there is a function $\varepsilon : \mathbb{X}(A, B) \to \mathbb{X}(A, B)$ such that:

[E.1] ε is a monoid morphism, that is, $\varepsilon(f + g) = \varepsilon(f) + \varepsilon(g)$ and $\varepsilon(0) = 0$
[E.2] $\varepsilon(g \circ f) = \varepsilon(g) \circ f$
[E.3] $\varepsilon(\pi_0) = \pi_0 \circ \varepsilon(1_{A \times B})$ and $\varepsilon(\pi_1) = \pi_1 \circ \varepsilon(1_{A \times B})$

By **[E.1]**, $\varepsilon(1_A)$ is an additive map, while by **[E.2]**, it follows that $\varepsilon(f) = \varepsilon(1_B) \circ f$. In light of this, it turns out that infinitesimal extensions can equivalently be described as a class of additive maps $\varepsilon_A : A \to A$. The equivalence is given by setting $\varepsilon(f) = \varepsilon_B \circ f$ and $\varepsilon_A = \varepsilon(1_A)$.

Lemma 4.2. *For a Cartesian left additive category $\mathbb{X}$, the following are in bijective correspondence:*

(i.) *An infinitesimal extension ε on $\mathbb{X}$;*
(ii.) *A family of maps $\varepsilon_A : A \to A$ indexed by objects A, such that ε_A is additive and $\varepsilon_{A \times B} = \varepsilon_A \times \varepsilon_B$.*

Therefore, for any map $f : A \to B$, $\varepsilon(f) = \varepsilon_A \circ f$.

Proof. Let ε be an infinitesimal extension. Then for each object A, define $\varepsilon_A : A \to A$ as $\varepsilon_A = \varepsilon(1_A)$. Since ε preserves the additive structure and $\varepsilon(g \circ f) = \varepsilon(g) \circ f$, it follows that:

$$
\begin{aligned}
\varepsilon_A \circ (f + g) =\ & \varepsilon(1_A) \circ (f + g) & \\
=\ & \varepsilon\left(1_A \circ (f + g)\right) & \textbf{[E.2]} \\
=\ & \varepsilon(f + g) & \\
=\ & \varepsilon(f) + \varepsilon(g) & \textbf{[E.1]} \\
=\ & \varepsilon\left(1_A \circ f\right) + \varepsilon\left(1_A \circ g\right) & \\
=\ & (\varepsilon(1_A) \circ f) \circ (\varepsilon(1_A) \circ g) & \textbf{[E.2]} \\
=\ & (\varepsilon_A \circ f) + (\varepsilon_A \circ g) &
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon_A \circ 0 =\ & \varepsilon(1_A) \circ 0 & \\
=\ & \varepsilon\left(1_A \circ 0\right) & \textbf{[E.2]} \\
=\ & \varepsilon(0) & \\
=\ & 0 & \textbf{[E.1]}
\end{aligned}
$$

So therefore, ε_A is additive. Now for a pair of objects A and B, we have the following:

$$
\begin{aligned}
\pi_0 \circ \varepsilon_{A \times B} =\ & \pi_0 \circ \varepsilon(1_{A \times B}) & \\
=\ & \varepsilon(\pi_0) & \textbf{[E.3]} \\
=\ & \varepsilon(1_A \circ \pi_0) & \\
=\ & \varepsilon(1_A) \circ \pi_0 & \textbf{[E.2]}
\end{aligned}
$$

$$= \ \varepsilon_A \circ \pi_0$$

So $\pi_0 \circ \varepsilon_{A \times B} = \varepsilon_A \circ \pi_0$, and similarly, $\pi_1 \circ \varepsilon_{A \times B} = \varepsilon_B \circ \pi_1$. Then by the universal property of the product, it follows that $\varepsilon_{A \times B} = \varepsilon_A \times \varepsilon_B$.

Conversely, suppose that each object A comes equipped with an additive map $\varepsilon_A : A \to A$ such that $\varepsilon_{A \times B} = \varepsilon_A \times \varepsilon_B$. Define $\varepsilon : \mathbb{X}(A, B) \to \mathbb{X}(A, B)$ as $\varepsilon(f) = \varepsilon_B \circ f$. Since ε_B is additive, it follows that:

$$\varepsilon(f + g) = \ \varepsilon_B \circ (f + g) = \ (\varepsilon_B \circ f) + (\varepsilon_B \circ g) = \ \varepsilon(f) + \varepsilon(g)$$

$$\varepsilon(0) = \ \varepsilon_B \circ 0 = \ 0$$

So ε is a monoid morphism, and so [**E.1**] holds. Next, it is straightforward by definition that $\varepsilon(g \circ f) = \varepsilon(g) \circ f$, so [**E.2**] holds. So it remains to show that ε is compatible with the projections. Note that $\varepsilon_{A \times B} = \varepsilon_A \times \varepsilon_B$ implies that $\pi_0 \circ \varepsilon_{A \times B} = \varepsilon_A \circ \pi_0$ and $\pi_1 \circ \varepsilon_{A \times B} = \varepsilon_B \circ \pi_1$. Therefore, we have that:

$$\varepsilon(\pi_0) = \ \varepsilon_A \circ \pi_0 = \ \pi_0 \circ \varepsilon_{A \times B}$$
$$\varepsilon(\pi_1) = \ \varepsilon_B \circ \pi_1 = \ \pi_1 \circ \varepsilon_{A \times B}$$

So we have that [**E.3**] holds. Therefore, ε is an infinitesimal extension.

Lastly, we need to show that these constructions are inverses of each other. Starting with an infinitesimal extension ε, we have that:

$$\varepsilon(f) = \ \varepsilon(1_B \circ f) = \ \varepsilon(1_B) \circ f = \ \varepsilon_B \circ f$$

While in the other direction, it is automatic that $\varepsilon(1_A) = \varepsilon_A$. So we conclude that infinitesimal extensions are bijective correspondence with a family of additive maps $\varepsilon_A : A \to A$ such that $\varepsilon_{A \times B} = \varepsilon_A \times \varepsilon_B$. $\qquad\square$

As an immediate consequence of the previous lemma, it follows that infinitesimal extensions are compatible with the product structure.

Lemma 4.3. *Let $\mathbb{X}$ be a Cartesian left additive category with an infinitesimal extension ε. Then $\varepsilon(\langle f, g \rangle) = \langle \varepsilon(f), \varepsilon(g) \rangle$ and $\varepsilon(h \times k) = \varepsilon(h) \times \varepsilon(k)$.*

Proof. We compute that:

$$\begin{aligned}
\varepsilon(\langle f, g \rangle) &= \ \varepsilon_{A \times B} \circ \langle f, g \rangle && \text{(Lemma 4.2)} \\
&= \ (\varepsilon_A \times \varepsilon_B) \circ \langle f, g \rangle && \text{(Lemma 4.2)} \\
&= \ \langle \varepsilon_A \circ f, \varepsilon_B \circ g \rangle && \\
&= \ \langle \varepsilon(f), \varepsilon(g) \rangle && \text{(Lemma 4.2)}
\end{aligned}$$

So $\varepsilon(\langle f, g \rangle) = \langle \varepsilon(f), \varepsilon(g) \rangle$. By similar calculations, we also have that:

$$\begin{aligned}
\varepsilon(h \times k) &= \ \varepsilon_{A \times B} \circ (h \times k) && \text{(Lemma 4.2)} \\
&= \ (\varepsilon_A \times \varepsilon_B) \circ (h \times k) && \text{(Lemma 4.2)} \\
&= \ (\varepsilon_A \circ h) \times (\varepsilon_B \circ k) && \\
&= \ \varepsilon(h) \times \varepsilon(k) && \text{(Lemma 4.2)}
\end{aligned}$$

So $\varepsilon(h \times k) = \varepsilon(h) \times \varepsilon(k)$. $\qquad\square$

Infinitesimal extensions equip each object with a canonical change action structure:

Lemma 4.4. *Let $\mathbb{X}$ be a Cartesian left additive category with infinitesimal extension ε. For every object A, define the three maps $\oplus_A : A \times A \to A$, $+_A : A \times A \to A$, and $0_A : \top \to A$ respectively as follows:*

$$\oplus_A = \pi_0 + \varepsilon(\pi_1) \qquad\qquad +_A = \pi_0 + \pi_1 \qquad\qquad 0_A = 0$$

Then $(A, A, \oplus_A, +_A, 0_A)$ is a change action in $\mathbb{X}$.

Proof. That $(A, +_A, 0_A)$ is a commutative monoid was shown in [BCS09, Proposition 1.2.2]. Thus **[CA.1]** holds. It remains to show that $\oplus_A$ is an action, which follows directly from the fact that ε preserves the additive structure. So we first that compute:

$$
\begin{aligned}
\oplus_A \circ \langle 1_A, 0_A \circ !_A \rangle &= (\pi_0 + \varepsilon(\pi_1)) \circ \langle 1_A, 0 \rangle \\
&= \pi_0 \circ \langle 1_A, 0_A \rangle + \varepsilon(\pi_1) \circ \langle 1_A, 0 \rangle \\
&= \pi_0 \circ \langle 1_A, 0_A \rangle + \varepsilon\left(\pi_1 \circ \langle 1_A, 0 \rangle\right) &&\textbf{[E.2]} \\
&= 1_A + \varepsilon(0) \\
&= 1_A + 0 &&\textbf{[E.1]} \\
&= 1_A
\end{aligned}
$$

So $\oplus_A \circ \langle 1_A, 0_A \circ !_A \rangle = 1_A$. Next we compute that:

$$
\begin{aligned}
\oplus_A \circ \langle \oplus_A \circ (1_A \times \pi_0), \pi_1 \circ \pi_1 \rangle &= (\pi_0 + \varepsilon(\pi_1)) \circ \langle \oplus_A \circ (1_A \times \pi_0), \pi_1 \circ \pi_1 \rangle \\
&= \pi_0 \circ \langle \oplus_A \circ (1_A \times \pi_0), \pi_1 \circ \pi_1 \rangle + \varepsilon(\pi_1) \circ \langle \oplus_A \circ (1_A \times \pi_0), \pi_1 \circ \pi_1 \rangle \\
&= \pi_0 \circ \langle \oplus_A \circ (1_A \times \pi_0), \pi_1 \circ \pi_1 \rangle + \varepsilon\left(\pi_1 \circ \langle \oplus_A \circ (1_A \times \pi_0), \pi_1 \circ \pi_1 \rangle\right) &&\textbf{[E.2]} \\
&= \oplus_A \circ (1_A \times \pi_0) + \varepsilon(\pi_1 \circ \pi_1) \\
&= (\pi_0 + \varepsilon(\pi_1)) \circ (1_A \times \pi_0) + \varepsilon(\pi_1 \circ \pi_1) \\
&= \pi_0 \circ (1_A \times \pi_0) + \varepsilon(\pi_1) \circ (1_A \times \pi_0) + \varepsilon(\pi_1 \circ \pi_1) \\
&= \pi_0 \circ (1_A \times \pi_0) + \varepsilon\left(\pi_1 \circ (1_A \times \pi_0)\right) + \varepsilon(\pi_1 \circ \pi_1) &&\textbf{[E.2]} \\
&= \pi_0 + \varepsilon(\pi_1 \circ \pi_0) + \varepsilon(\pi_1 \circ \pi_1) \\
&= \pi_0 + \varepsilon(\pi_1 \circ \pi_0 + \pi_1 \circ \pi_1) &&\textbf{[E.1]} \\
&= \pi_0 + \varepsilon(\pi_1 \circ (\pi_0 + \pi_1)) &&(\pi_1 \text{ is additive}) \\
&= \pi_0 + \varepsilon(\pi_1 \circ +_A) \\
&= \pi_0 \circ (1_A \times +_A) + \varepsilon(\pi_1 \circ (1_A \times +_A)) \\
&= \pi_0 \circ (1_A \times +_A) + \varepsilon(\pi_1) \circ (1_A \times +_A) &&\textbf{[E.2]} \\
&= (\pi_0 + \varepsilon(\pi_1)) \circ (1_A \times +_A) \\
&= \oplus_A \circ (1_A \times +_A)
\end{aligned}
$$

So $\oplus_A \circ (1_A \times +_A) = \oplus_A \circ \langle \oplus_A \circ (1_A \times \pi_0), \pi_1 \circ \pi_1 \rangle$. Thus **[CA.2]** holds. So we conclude that $(A, A, \oplus_A, +_A, 0_A)$ is a change action. $\qquad\square$

It is important to note that in an arbitrary Cartesian left additive category, $\oplus_A$, $+_A$, and 0_A are not necessarily natural transformations. That said, $+_A$ and 0_A are natural with respect to additive maps, that is, if $f : A \to B$ is additive then $f \circ 0_A = 0_B$ and $f \circ +_A = +_B \circ (f \times f)$. Setting $\overline{A} \equiv (A, A, \oplus_A, +_A, 0_A)$, we note that $f \oplus_{\overline{A}} g = f + \varepsilon(g)$ and $f +_{\overline{A}} g = f + g$, and so in particular $+_{\overline{A}} = +$. Therefore, from now on we will omit the subscripts and simply write $\oplus$ and $+$.

For every Cartesian left additive category, there are always at least two possible infinitesimal extensions: one given by setting the infinitesimal extension of a map to be zero and another given by setting the infinitesimal extension of a map to be itself.

Lemma 4.5. *For any Cartesian left additive category* $\mathbb{X}$,

(i). *Setting* $\varepsilon(f) = 0$ *defines an infinitesimal extension on* $\mathbb{X}$ *and therefore in this case,* $\oplus_A = \pi_0$ *and* $f \oplus g = f$.

(ii). *Setting* $\varepsilon(f) = f$ *defines an infinitesimal extension on* $\mathbb{X}$ *and therefore in this case,* $\oplus_A = +_A$ *and* $f \oplus g = f + g$.

Proof. These are both straightforward to check and we leave it as an exercise. $\qquad\square$

We note that while these examples of infinitesimal extensions may seem trivial, they are both very important as they will give rise to key examples of Cartesian difference categories.

4.2. **Cartesian Difference Categories.** Here we introduce Cartesian difference categories, the main novel contribution of this paper.

Definition 4.6. A **Cartesian difference category** is a Cartesian left additive category with an infinitesimal extension ε which is equipped with a **difference combinator** ∂ of the form:

$$\frac{f : A \to B}{\partial[f] : A \times A \to B}$$

verifying the following coherence conditions:

[C∂.0] $f \circ (x + \varepsilon(y)) = f \circ x + \varepsilon\left(\partial[f] \circ \langle x, y \rangle\right)$

[C∂.1] $\partial[f + g] = \partial[f] + \partial[g]$, $\partial[0] = 0$, and $\partial[\varepsilon(f)] = \varepsilon(\partial[f])$

[C∂.2] $\partial[f] \circ \langle x, y + z \rangle = \partial[f] \circ \langle x, y \rangle + \partial[f] \circ \langle x + \varepsilon(y), z \rangle$ and $\partial[f] \circ \langle x, 0 \rangle = 0$

[C∂.3] $\partial[1_A] = \pi_1$ and $\partial[\pi_0] = \pi_0 \circ \pi_1$ and $\partial[\pi_1] = \pi_1 \circ \pi_1$

[C∂.4] $\partial[\langle f, g \rangle] = \langle \partial[f], \partial[g] \rangle$

[C∂.5] $\partial[g \circ f] = \partial[g] \circ \langle f \circ \pi_0, \partial[f] \rangle$

[C∂.6] $\partial\left[\partial[f]\right] \circ \langle \langle x, y \rangle, \langle 0, z \rangle \rangle = \partial[f] \circ \langle x + \varepsilon(y), z \rangle$

[C∂.7] $\partial\left[\partial[f]\right] \circ \langle \langle x, y \rangle, \langle z, 0 \rangle \rangle = \partial\left[\partial[f]\right] \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle$

We say that $\partial[f]$ is the derivative of f.

Before giving some intuition on the axioms **[C∂.0]** to **[C∂.7]**, we first observe that one could have used change action notation $\oplus$ to express **[C∂.0]**, **[C∂.2]**, and **[C∂.6]**:

[C∂.0]: $f \circ (x \oplus y) = f \circ x \oplus \partial[f] \circ \langle x, y \rangle$

[C∂.2]: $\partial[f] \circ \langle x, y + z \rangle = \partial[f] \circ \langle x, y \rangle + \partial[f] \circ \langle x \oplus y, z \rangle$ and $\partial[f] \circ \langle x, 0 \rangle = 0$

[C∂.6]: $\partial\left[\partial[f]\right] \circ \langle \langle x, y \rangle, \langle 0, z \rangle \rangle = \partial[f] \circ \langle x \oplus y, z \rangle$

And also, just like Cartesian differential categories, **[C∂.6]** and **[C∂.7]** have alternative equivalent expressions.

Lemma 4.7. *In the presence of the other axioms,* **[C∂.6]** *and* **[C∂.7]** *are equivalent to:*

[C∂.6.a] $\partial\left[\partial[f]\right] \circ \langle \langle x, 0 \rangle, \langle 0, y \rangle \rangle = \partial[f] \circ \langle x, y \rangle$

[C∂.7.a] $\partial\left[\partial[f]\right] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle = \partial\left[\partial[f]\right] \circ \langle \langle x, z \rangle, \langle y, w \rangle \rangle$

Proof. The proof is essentially the same as [CC14, Proposition 4.2]. Assume that the other axioms **[C∂.0]** to **[C∂.5]** hold. Suppose that **[C∂.6]** and **[C∂.7]** also hold. We first compute **[C∂.6.a]** using **[C∂.6]**:

$$\partial\left[\partial[f]\right] \circ \langle \langle x, 0 \rangle, \langle 0, y \rangle \rangle = \partial[f] \circ \langle x + \varepsilon(0), y \rangle \qquad\qquad \textbf{[C}\partial\textbf{.6]}$$

$$= \boldsymbol{\partial}[f] \circ \langle x + 0, y \rangle \qquad \textbf{[E.1]}$$
$$= \boldsymbol{\partial}[f] \circ \langle x, y \rangle$$

Next we compute **[C∂.7.a]** using **[C∂.2]**, **[C∂.6]**, and **[C∂.7]**:

$$\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle, \langle z, w\rangle\rangle = \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle, \langle z, 0\rangle + \langle 0, w\rangle\rangle$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle, \langle z, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle + \varepsilon(\langle z, 0\rangle), \langle 0, w\rangle\rangle \qquad \textbf{[C∂.2]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle + \langle\varepsilon(z), \varepsilon(0)\rangle, \langle 0, w\rangle\rangle \qquad \text{(Lemma 4.3)}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle + \langle\varepsilon(z), 0\rangle, \langle 0, w\rangle\rangle \qquad \textbf{[E.1]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x + \varepsilon(z), y\rangle, \langle 0, w\rangle\rangle$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}[f] \circ \langle x + \varepsilon(z) + \varepsilon(y), w\rangle \qquad \textbf{[C∂.7]} + \textbf{[C∂.6]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}[f] \circ \langle x + \varepsilon(y) + \varepsilon(z), w\rangle$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x + \varepsilon(y), z\rangle, \langle 0, w\rangle\rangle \qquad \textbf{[C∂.6]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x + \varepsilon(y), z + 0\rangle, \langle 0, w\rangle\rangle$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x + \varepsilon(y), z + \varepsilon(0)\rangle, \langle 0, w\rangle\rangle \qquad \textbf{[E.1]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle + \langle\varepsilon(y), \varepsilon(0)\rangle, \langle 0, w\rangle\rangle$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle + \varepsilon(\langle y, 0\rangle), \langle 0, w\rangle\rangle \qquad \text{(Lemma 4.3)}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle + \langle 0, w\rangle\rangle \qquad \textbf{[C∂.2]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle, \langle z, w\rangle\rangle \qquad \textbf{[C∂.7]}$$

On the other, suppose instead that **[C∂.6.a]** and **[C∂.7.a]** hold. Then **[C∂.7]** follows immediately from **[C∂.7.a]** by setting the last term $w = 0$.

$$\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle, \langle z, 0\rangle\rangle = \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, z\rangle, \langle y, 0\rangle\rangle \qquad \textbf{[C∂.7.a]}$$

Next we compute **[C∂.6]** using **[C∂.2]**, **[C∂.6.a]**, and **[C∂.7.a]**.

$$\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle, \langle 0, z\rangle\rangle = \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle, \langle y, z\rangle\rangle \qquad \textbf{[C∂.7.a]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle, \langle y, 0\rangle + \langle 0, z\rangle\rangle$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle, \langle y, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle + \varepsilon(\langle y, 0\rangle), \langle 0, z\rangle\rangle \qquad \textbf{[C∂.2]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, y\rangle, \langle 0, 0\rangle\rangle + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle + \varepsilon(\langle y, 0\rangle), \langle 0, z\rangle\rangle \qquad \textbf{[C∂.7.a]}$$
$$= 0 + \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle + \varepsilon(\langle y, 0\rangle), \langle 0, z\rangle\rangle \qquad \textbf{[C∂.2]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle + \langle\varepsilon(y), \varepsilon(0)\rangle, \langle 0, z\rangle\rangle \qquad \text{(Lemma 4.3)}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x, 0\rangle + \langle\varepsilon(y), 0\rangle, \langle 0, z\rangle\rangle \qquad \textbf{[E.1]}$$
$$= \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle\langle x + \varepsilon(y), 0\rangle, \langle 0, z\rangle\rangle$$
$$= \boldsymbol{\partial}[f] \circ \langle x + \varepsilon(y), z\rangle \qquad \textbf{[C∂.6.a]}$$

So we conclude that **[C∂.6]** and **[C∂.7]** are equivalent to **[C∂.6.a]** and **[C∂.7.a]**. $\qquad \square$

The keen eyed reader will notice that the axioms of a Cartesian difference category are very similar to the axioms of a Cartesian differential category. Indeed, **[C∂.1]**, **[C∂.3]**, **[C∂.4]**, **[C∂.5]**, and **[C∂.7]** are the same as their Cartesian differential category counterpart. The axioms which are different are **[C∂.2]** and **[C∂.6]** where the infinitesimal extension ε is now included, and also there is the new extra axiom **[C∂.0]**. On the other hand, interestingly enough, **[CD.6.a]** is the same as **[C∂.6.a]**. We also point out that writing out **[C∂.0]**

and **[C∂.2]** using change action notion, we see that these axioms are precisely **[CAD.1]** and **[CAD.2]** respectively. To better understand **[C∂.0]** to **[C∂.7]** it may be useful to write them out using element-like notation. In element-like notation, **[C∂.0]** is written as:

$$f(x + \varepsilon(y)) = f(x) + \varepsilon\left(\boldsymbol{\partial}[f](x, y)\right)$$

This condition can be read as a generalization of the Kock-Lawvere axiom that characterizes the derivative from synthetic differential geometry [Koc06]. Broadly speaking, the Kock-Lawvere axiom states that, for any map $f : \mathcal{R} \to \mathcal{R}$ and any $x \in \mathcal{R}$ and $d \in \mathcal{D}$, there exists a unique $f'(x) \in \mathcal{R}$ verifying $f(x + d) = f(x) + d \cdot f'(x)$, where $\mathcal{D}$ is the subset of $\mathcal{R}$ consisting of infinitesimal elements. It is by analogy with the Kock-Lawvere axiom that we refer to ε as an "infinitesimal extension" as it can be thought of as embedding the space A into a subspace $\varepsilon(A)$ of infinitesimal elements.

    [C∂.1] states that the differential of a sum of maps is the sum of differentials, and similarly for zero maps and the infinitesimal extension of a map. **[C∂.2]** is the first crucial difference between a Cartesian difference category and a Cartesian differential category. In a Cartesian differential category, the differential of a map is assumed to be additive in its second argument. In a Cartesian difference category, just as derivatives for change actions, while the differential is still required to preserve zeros in its second argument, it is only additive "up to a small perturbation", that is:

$$\boldsymbol{\partial}[f](x, y + z) = \boldsymbol{\partial}[f](x, y) + \boldsymbol{\partial}[f](x + \varepsilon(y), z)$$

[C∂.3] tells us what the differential of the identity and projection maps are, while **[C∂.4]** says that the differential of a pairing of maps is the pairing of their differentials. **[C∂.5]** is the chain rule which expresses what the differential of a composition of maps is:

$$\boldsymbol{\partial}[g \circ f](x, y) = \boldsymbol{\partial}[g](f(x), \boldsymbol{\partial}[f](x, y))$$

The last two axioms, **[C∂.6]** and **[C∂.7]** tell us how to work with second order differentials. **[C∂.6]** is expressed as follows:

$$\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right](x, y, 0, z) = \boldsymbol{\partial}[f](x + \varepsilon(y), z)$$

and finally **[C∂.7]** is expressed as:

$$\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right](x, y, z, 0) = \boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right](x, z, y, 0)$$

It is interesting to note that while **[C∂.6]** is different from **[CD.6]**, its alternative version **[C∂.6.a]** is the same as **[CD.6.a]**.

$$\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right]((x, 0), (0, y)) = \boldsymbol{\partial}[f](x, z)$$

The interplay between **[C∂.0]**, **[C∂.2]**, and **[C∂.6]** gives rise to some remarkable and counter-intuitive consequences.

Lemma 4.8. *In a Cartesian difference category* $\mathbb{X}$*, for any map* $f : A \to B$*, the following equalities hold for any suitable maps:*

  (i). $\boldsymbol{\partial}[f] \circ \langle x, \varepsilon(y) \rangle = \varepsilon(\boldsymbol{\partial}[f]) \circ \langle x, y \rangle$
  (ii). $\varepsilon(\boldsymbol{\partial}[f]) \circ \langle x + \varepsilon(y), z \rangle = \varepsilon(\boldsymbol{\partial}[f]) \circ \langle x + \varepsilon^2(y), z \rangle$
  (iii). $\varepsilon^2\left(\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right]\right) \circ \langle\langle x, y \rangle, \langle z, 0 \rangle\rangle = \varepsilon^3\left(\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right]\right) \circ \langle\langle x, y \rangle, \langle z, 0 \rangle\rangle$

Proof. These are mostly straightforward calculations. We start by computing (i):

$$\begin{aligned}
\boldsymbol{\partial}[f] \circ \langle x, \varepsilon(y) \rangle &= \boldsymbol{\partial}[f] \circ (\langle x, 0 \rangle + \langle 0, \varepsilon(y) \rangle) \\
&= \boldsymbol{\partial}[f] \circ (\langle x, 0 \rangle + \langle \varepsilon(0), \varepsilon(y) \rangle) \qquad\qquad \textbf{[E.1]}
\end{aligned}$$

$$\begin{aligned}
&= \ \boldsymbol{\partial}[f] \circ (\langle x, 0 \rangle + \varepsilon(\langle 0, y \rangle)) && \text{(Lemma 4.3)} \\
&= \ \boldsymbol{\partial}[f] \circ \langle x, 0 \rangle + \varepsilon \left(\boldsymbol{\partial}^2[f] \circ \langle \langle x, 0 \rangle, \langle 0, y \rangle \rangle \right) && \textbf{[C}\boldsymbol{\partial}\textbf{.0]} \\
&= \ 0 + \varepsilon(\boldsymbol{\partial}[f] \circ \langle x, y \rangle) && \textbf{[C}\boldsymbol{\partial}\textbf{.2]} + \textbf{[C}\boldsymbol{\partial}\textbf{.6.a]} \\
&= \ \varepsilon(\boldsymbol{\partial}[f]) \circ \langle x, y \rangle && \textbf{[E.2]}
\end{aligned}$$

Next, we use (i) to compute (ii):

$$\begin{aligned}
\varepsilon(\boldsymbol{\partial}[f]) \circ \langle x + \varepsilon(y), z \rangle &= \ \varepsilon \left(\boldsymbol{\partial}[f] \circ \langle x + \varepsilon(y), z \rangle \right) && \textbf{[E.2]} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \circ \langle \langle x, y \rangle, \langle 0, z \rangle \rangle \right) && \textbf{[C}\boldsymbol{\partial}\textbf{.6]} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \circ \langle \langle x, 0 \rangle, \langle y, z \rangle \rangle \right) && \textbf{[C}\boldsymbol{\partial}\textbf{.7.a]} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, 0 \rangle, \langle y, z \rangle \rangle && \textbf{[E.2]} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \circ \langle \langle x, 0 \rangle, \varepsilon(\langle y, z \rangle) \rangle && \text{(Lemma 4.8.i)} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \circ \langle \langle x, 0 \rangle, \langle \varepsilon(y), \varepsilon(z) \rangle \rangle && \text{(Lemma 4.3)} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \circ \langle \langle x, \varepsilon(y) \rangle, \langle 0, \varepsilon(z) \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.7.a]} \\
&= \ \boldsymbol{\partial}[f] \circ \langle x + \varepsilon^2(y), \varepsilon(z) \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.6]} \\
&= \ \varepsilon(\boldsymbol{\partial}[f]) \circ \langle x + \varepsilon^2(y), z \rangle && \text{(Lemma 4.8.i)}
\end{aligned}$$

Lastly for (iii), we use (i) and (ii):

$$\begin{aligned}
\varepsilon^2 \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, y \rangle, \langle z, 0 \rangle \rangle &= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon^2(f)] \right] \circ \langle \langle x, y \rangle, \langle z, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.1]} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon^2(f)] \right] \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.7.a]} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \right) \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.1]} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, z \rangle, \varepsilon(\langle y, 0 \rangle) \rangle && \text{(Lemma 4.8.i)} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, z \rangle, \langle \varepsilon(y), \varepsilon(0) \rangle \rangle && \text{(Lemma 4.3)} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, z \rangle, \langle \varepsilon(y), 0 \rangle \rangle && \textbf{[E.2]} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, \varepsilon(y) \rangle, \langle z, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.7.a]} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, \varepsilon(y) \rangle, \langle z, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.1]} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, 0 \rangle + \langle 0, \varepsilon(y) \rangle, \langle z, 0 \rangle \rangle \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, 0 \rangle + \langle \varepsilon(0), \varepsilon(y) \rangle, \langle z, 0 \rangle \rangle && \textbf{[E.1]} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, 0 \rangle + \varepsilon(\langle 0, y \rangle), \langle z, 0 \rangle \rangle && \text{(Lemma 4.3)} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, 0 \rangle + \varepsilon^2(\langle 0, y \rangle), \langle z, 0 \rangle \rangle && \text{(Lemma 4.8.ii)} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, 0 \rangle + \langle \varepsilon^2(0), \varepsilon^2(y) \rangle, \langle z, 0 \rangle \rangle && \text{(Lemma 4.3)} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, 0 \rangle + \langle 0, \varepsilon^2(y) \rangle, \langle z, 0 \rangle \rangle && \textbf{[E.1]} \\
&= \ \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, \varepsilon^2(y) \rangle, \langle z, 0 \rangle \rangle \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, \varepsilon^2(y) \rangle, \langle z, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.1]} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, z \rangle, \langle \varepsilon^2(y), 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.7.a]} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, z \rangle, \langle \varepsilon^2(y), \varepsilon^2(0) \rangle \rangle && \textbf{[E.1]} \\
&= \ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, z \rangle, \varepsilon^2(\langle y, 0 \rangle) \rangle && \text{(Lemma 4.3)}
\end{aligned}$$

$$\begin{aligned}
&= \; \varepsilon^2 \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \right) \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle && \text{(Lemma 4.8.i)} \\
&= \; \varepsilon^2 \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle \right) && \textbf{[E.2]} \\
&= \; \varepsilon^2 \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \circ \langle \langle x, y \rangle, \langle z, 0 \rangle \rangle \right) && \textbf{[C}\boldsymbol{\partial}\textbf{.7.a]} \\
&= \; \varepsilon^2 \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon(f)] \right] \right) \circ \langle \langle x, y \rangle, \langle z, 0 \rangle \rangle && \textbf{[E.2]} \\
&= \; \varepsilon^3 \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, y \rangle, \langle z, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.1]} \quad \square
\end{aligned}$$

A desirable identity we would like to hold is a slightly stronger version of Lemma 4.8.iii:

$$\varepsilon(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right]) \circ \langle \langle x, y \rangle, \langle z, 0 \rangle \rangle = \varepsilon^2(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right]) \circ \langle \langle x, y \rangle, \langle z, 0 \rangle \rangle \tag{4.1}$$

Note that if ε is injective (i.e $\varepsilon(f) = \varepsilon(g)$ implies that $f = g$) or if ε is idempotent (i.e. $\varepsilon^2(f) = \varepsilon(f)$), then (4.1) holds. In fact, in all of our examples of Cartesian difference categories in Section 5, (4.1) holds (even for those where ε is neither injective or idempotent). It is currently unclear if (4.1) holds in any Cartesian difference category, as we do not have a counterexample. If (4.1) holds and the infinitesimal extension is nilpotent, then it turns out that **[C$\boldsymbol{\partial}$.2]** is the same as **[CD.2]**.

Lemma 4.9. *Let $\mathbb{X}$ be a Cartesian difference category with a nilpotent infinitesimal extension, that is, for every map $f : A \to B$ there is some $k \in \mathbb{N}$ such that $\varepsilon^k(f) = 0$, and suppose that (4.1) holds. Then* **[CD.2]** *holds, or in other words, every derivative $\boldsymbol{\partial}[f]$ is additive in its second argument, that is, $\boldsymbol{\partial}[f] \circ \langle x, y + z \rangle = \boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle$ and $\boldsymbol{\partial}[f] \circ \langle x, 0 \rangle = 0$.*

Proof. By **[C$\boldsymbol{\partial}$.2]**, it already holds that $\boldsymbol{\partial}[f] \circ \langle x, 0 \rangle = 0$. Therefore it remains to show that $\boldsymbol{\partial}[f] \circ \langle x, y + z \rangle = \boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle$. So suppose that $\varepsilon^k(f) = 0$ for some $k \in \mathbb{N}$. Then using Lemma 4.1, we compute that:

$$\begin{aligned}
\boldsymbol{\partial}[f] \circ \langle x, y + z \rangle = \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x + \varepsilon(y), z \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.2]} \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ (\langle x, z \rangle + \langle \varepsilon(y), 0 \rangle) \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ (\langle x, z \rangle + \langle \varepsilon(y), \varepsilon(0) \rangle) && \textbf{[E.1]} \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ (\langle x, z \rangle + \varepsilon(\langle y, 0 \rangle)) && \text{(Lemma 4.3)} \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle + \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle \right) && \textbf{[C}\boldsymbol{\partial}\textbf{.0]} \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle + \varepsilon \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle && \textbf{[E.2]} \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle + \varepsilon^k \left(\boldsymbol{\partial} \left[\boldsymbol{\partial}[f] \right] \right) \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle && \text{(Iterating (4.1))} \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle + \boldsymbol{\partial} \left[\boldsymbol{\partial}[\varepsilon^k(f)] \right] \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.1]} \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle + \boldsymbol{\partial} \left[\boldsymbol{\partial}[0] \right] \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle && (\varepsilon \text{ is nilpotent}) \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle + 0 \circ \langle \langle x, z \rangle, \langle y, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.1]} \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle + 0 \\
= \; &\boldsymbol{\partial}[f] \circ \langle x, y \rangle + \boldsymbol{\partial}[f] \circ \langle x, z \rangle
\end{aligned}$$

We conclude that $\boldsymbol{\partial}[f]$ is additive in its second argument, and so **[CD.2]** holds. $\qquad\square$

4.3. **Another look at Cartesian Differential Categories.** Here we explain how every Cartesian differential category is a Cartesian difference category where the infinitesimal extension is given by zero.

Proposition 4.10. *Every Cartesian differential category $\mathbb{X}$ with differential combinator D is a Cartesian difference category where the infinitesimal extension is defined as $\varepsilon(f) = 0$ and the difference combinator is defined to be the differential combinator, $\boldsymbol{\partial} = \mathsf{D}$.*

Proof. As noted before, the first two parts of the [**C$\boldsymbol{\partial}$.1**], the second part of [**C$\boldsymbol{\partial}$.2**], [**C$\boldsymbol{\partial}$.3**], [**C$\boldsymbol{\partial}$.4**], [**C$\boldsymbol{\partial}$.5**], and [**C$\boldsymbol{\partial}$.7**] are precisely the same as their Cartesian differential axiom counterparts. On the other hand, since $\varepsilon(f) = 0$, [**C$\boldsymbol{\partial}$.0**] and the third part of [**C$\boldsymbol{\partial}$.1**] trivially state that $0 = 0$, while the first part of [**C$\boldsymbol{\partial}$.2**] and [**C$\boldsymbol{\partial}$.6**] end up being precisely the first part of [**CD.2**] and [**CD.6**]. Therefore, the differential combinator satisfies the Cartesian difference axioms and we conclude that a Cartesian differential category is a Cartesian difference category. $\qquad\square$

Conversely, one can always build a Cartesian differential category from a Cartesian difference category by considering the objects for which the infinitesimal extension of the identity map is the zero map. We call such objects ε-vanishing.

Definition 4.11. In a Cartesian left additive category $\mathbb{X}$ with infinitesimal extension ε, an object A is said to be ε-**vanishing** if $\varepsilon_A = \varepsilon(1_A) = 0$. We denote $\mathbb{X}_{\varepsilon\text{-van}}$ to be the full subcategory of ε-vanishing objects of $\mathbb{X}$.

The full subcategory of ε-vanishing objects always forms a Cartesian left additive category.

Lemma 4.12. *In a Cartesian left additive category $\mathbb{X}$ with infinitesimal extension ε,*

  (i). *The terminal object $\top$ is ε-vanishing;*
  (ii). *If A and B are ε-vanishing then their product $A \times B$ is ε-vanishing;*

Therefore, $\mathbb{X}_{\varepsilon\text{-van}}$ is a Cartesian left additive category with the same Cartesian left additive structure as $\mathbb{X}$. Furthermore:

  (iii). *If B is ε-vanishing, then for any map $f : A \to B$, $\varepsilon(f) = 0$.*

Proof. For (i), recall that for the terminal object, $1_\top = 0$. Therefore by [**E.1**] we easily see that $\varepsilon(1_\top) = \varepsilon(0) = 0$. So the terminal object $\top$ is ε-vanishing. For (ii), first recall that by Lemma 4.2, for any pair of objects A and B, $\varepsilon_{A \times B} = \varepsilon_A \times \varepsilon_B$. Therefore, if both A and B are ε-vanishing, we have that $\varepsilon_{A \times B} = \varepsilon_A \times \varepsilon_B = 0 \times 0 = 0$. So there product of ε-vanishing objects is again ε-vanishing. Thus $\mathbb{X}_{\varepsilon\text{-van}}$ is closed under finite products and therefore we conclude that $\mathbb{X}_{\varepsilon\text{-van}}$ is also a Cartesian left additive category. For (iii), suppose that B is ε-vanishing. Then by [**E.2**], we have that $\varepsilon(f) = \varepsilon(1_B) \circ f = 0 \circ f = 0$. $\qquad\square$

For a Cartesian difference category, its subcategory of ε-vanishing objects is a Cartesian differential category.

Proposition 4.13. *For a Cartesian difference category $\mathbb{X}$ with infinitesimal extension ε and difference combinator $\boldsymbol{\partial}$, then $\mathbb{X}_{\varepsilon\text{-van}}$ is a Cartesian differential category where the differential combinator is defined to be the difference combinator, $\mathsf{D} = \boldsymbol{\partial}$.*

Proof. In Lemma 4.12 we already explained why $\mathbb{X}_{\varepsilon\text{-van}}$ is a Cartesian left additive category. So it remains to explain why $\boldsymbol{\partial}$ is a differential combinator. However, by Lemma 4.12.iii, every map in $\mathbb{X}_{\varepsilon\text{-van}}$ satisfies $\varepsilon(f) = 0$. Therefore, for maps between objects in $\mathbb{X}_{\varepsilon\text{-van}}$, the

Cartesian difference axioms are precisely the Cartesian differential axioms. So we conclude that the difference combinator is a differential combinator for this subcategory, and thus $\mathbb{X}_{\varepsilon\text{-van}}$ is a Cartesian differential category. □

Note that by Lemma 4.12.i, for any Cartesian difference category $\mathbb{X}$, the terminal object $\top$ is always ε-vanishing, and so therefore, $\mathbb{X}_{\varepsilon\text{-van}}$ is never empty. On the other hand, applying Proposition 4.13 to a Cartesian differential category results in the entire category, since every object is ε-vanishing by definition. It is also important to note that the above two propositions do not imply that if a difference combinator is a differential combinator then the infinitesimal extension must be zero. In Section 5.3, we provide such an example of a Cartesian differential category that comes equipped with a non-zero infinitesimal extension such that the differential combinator is a difference combinator with respect to this non-zero infinitesimal extension.

4.4. Cartesian Difference Categories as Change Action Models. In this section, we show how every Cartesian difference category is a particularly well-behaved change action model, and conversely how every change action model contains a Cartesian difference category.

Proposition 4.14. *Let $\mathbb{X}$ be a Cartesian difference category with infinitesimal extension ε and difference combinator $\boldsymbol{\partial}$. Define the functor $\alpha : \mathbb{X} \to \mathsf{CAct}(\mathbb{X})$ on objects as $\alpha(A) = (A, A, \oplus_A, +_A, 0_A)$ (as defined in Lemma 4.4) and on maps as $\alpha(f) = (f, \boldsymbol{\partial}[f])$. Then $(\mathbb{X}, \alpha : \mathbb{X} \to \mathsf{CAct}(\mathbb{X}))$ is a change action model.*

Proof. By Lemma 4.4, $(A, A, \oplus_A, +_A, 0_A)$ is a change action and so α is well-defined on objects. While for a map f, $\boldsymbol{\partial}[f]$ is a derivative of f in the change action sense since **[C$\boldsymbol{\partial}$.0]** and **[C$\boldsymbol{\partial}$.2]** are precisely **[CAD.1]** and **[CAD.2]**, and so α is well-defined on maps. Next we show that α preserves identities and composition, which follows from **[C$\boldsymbol{\partial}$.3]** and **[C$\boldsymbol{\partial}$.5]** respectively:

$$\begin{aligned}
\alpha(1_A) =\ & (1_A, \boldsymbol{\partial}[1_A]) \\
=\ & (1_A, \pi_1) && \textbf{[C$\boldsymbol{\partial}$.3]}
\end{aligned}$$

$$\begin{aligned}
\alpha(g \circ f) =\ & (g \circ f, \boldsymbol{\partial}[g \circ f]) \\
=\ & (g \circ f, \boldsymbol{\partial}[g] \circ \langle f \circ \pi_0, \boldsymbol{\partial}[f]\rangle) && \textbf{[C$\boldsymbol{\partial}$.5]} \\
=\ & (g, \boldsymbol{\partial}[g]) \circ (f, \boldsymbol{\partial}[g]) && \text{(by def. of comp. in } \mathsf{CAct}(\mathbb{X})) \\
=\ & \alpha(g) \circ \alpha(f)
\end{aligned}$$

So α is a functor. Next we check that α preserves projections and pairings, which will follow from **[C$\boldsymbol{\partial}$.3]** and **[C$\boldsymbol{\partial}$.4]**:

$$\begin{aligned}
\alpha(\pi_i) =\ & (\pi_i, \boldsymbol{\partial}[\pi_i]) \\
=\ & (\pi_i, \pi_i \circ \pi_1) && \textbf{[C$\boldsymbol{\partial}$.3]}
\end{aligned}$$

$$\begin{aligned}
\alpha(\langle f, g\rangle) =\ & (\langle f, g\rangle, \boldsymbol{\partial}[\langle f, g\rangle]) \\
=\ & (\langle f, g\rangle, \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g]\rangle) && \textbf{[C$\boldsymbol{\partial}$.4]} \\
=\ & \langle (f, \boldsymbol{\partial}[f]), (g, \boldsymbol{\partial}[g])\rangle
\end{aligned}$$

$$= \langle \alpha(f), \alpha(g) \rangle$$

So α preserves finite products. Lastly, it is clear that α is a section of the forgetful functor, and therefore we conclude that $(\mathbb{X}, \alpha)$ is a change action model. $\qquad \square$

Not every change action model is a Cartesian difference category. For example, change action models do not require the addition to be commutative. On the other hand, it can be shown that every change action model contains a Cartesian difference category as a full subcategory. We call the objects and maps in this subcategory *flat*. It is important to note that the definition here is a slight generalisation of the one in [APL20]. The original suggested definition implied that $\oplus = +$ and as a result $\varepsilon(f) = f$. While this indeed resulted in a Cartesian difference category (specifically one that generalized the Abelian group example from Section 5.2), we found that it was a bit too restrictive of a construction. We correct this by generalizing the definition slightly, while still keeping the construction's overall core ideas the same.

Definition 4.15. Let $(\mathbb{X}, \alpha : \mathbb{X} \to \mathsf{CAct}(\mathbb{X}))$ be a change action model. An object A is **flat** whenever the following equations hold:

[F.1] $\Delta A = A$
[F.2] $\alpha(\oplus_A) = (\oplus_A, \oplus_A \circ \pi_1)$ and $\alpha(+_A) = (+_A, +_A \circ \pi_1)$
[F.3] $\oplus_A$ is right-injective, that is, if $\oplus_A \circ \langle f, g \rangle = \oplus_A \circ \langle f, h \rangle$ then $g = h$.

Let $f : A \to B$ be a map between flat objects in the above change action model, with $\alpha(f) = (f, \partial[f])$. Then f is **flat** whenever the following condition holds:

[F.4] $0_B \oplus_B \partial[f] \circ \langle x, y \rangle = \partial[f] \circ \langle x, 0_A \oplus_A y \rangle$

We define Flat_α to be the subcategory of $\mathbb{X}$ containing all flat objects and all flat morphisms between them.

Intuitively, the condition **[F.1]** simply states that a flat object is similar to an Euclidean space in the sense that its "tangent space" is equal to itself. **[F.2]** can be read as stating that the action $\oplus_A$ and the addition of changes $+_A$ are linear. In this sense, flat objects behave like Euclidean spaces. **[F.3]** tells us that we may cancel the left argument of $\oplus_A$, which will imply that the result infinitesimal extension is injective. This cancellation ability that **[F.3]** provides will be used throughout most of the proofs in this section. Lastly, as we will see below, **[F.4]** is a generalization of Lemma 4.8.i.

We would like to show that for any change action model $(\mathbb{X}, \alpha)$, Flat_α is a Cartesian difference category. We first explain the finite product structure of Flat_α.

Lemma 4.16. *Let $(\mathbb{X}, \alpha : \mathbb{X} \to \mathsf{CAct}(\mathbb{X}))$ be a change action model. Then:*

(i). *The terminal object $\top$ is flat;*
(ii). *If A and B are flat then their product $A \times B$ is flat;*
(iii). *If A is a flat object, then the identity $1_A : A \to A$ is flat;*
(iv). *If A, B, and C are flat objects, and $f : A \to B$ and $g : B \to C$ are flat maps, then their composite $g \circ f : A \to C$ is flat;*
(v). *If A and B are flat objects, then the projection maps $\pi_0 : A \times B \to A$ and $\pi_1 : A \times B \to B$ are flat;*
(vi). *If A, B, and C are flat objects, and $f : C \to A$ and $g : C \to B$ are flat maps, then their pairing $\langle f, g \rangle : C \to A \times B$ is flat.*

Furthermore, Flat_α is a Cartesian category with the same Cartesian structure as $\mathbb{X}$.

Proof. These are straightforward and we leave it as an exercise for the reader to check for themselves. (i), (ii), (v), and (vi) follows from the fact that α preserves finite products, while (iii) and (v) follows from the functoriality of α. $\qquad\square$

As an immediate consequence, we note that for any change action model $(\mathbb{X}, \alpha)$, since the terminal object is always flat, Flat_α is never empty. Next we discuss the additive structure of Flat_α. The sum of maps $f : A \to B$ and $g : A \to B$ in Flat_α is defined using the change action structure $f + g := f +_B g$, while the zero map $0 : A \to B$ is $0 := 0_B \circ !_A$. And so we obtain that:

Lemma 4.17. *Let* $(\mathbb{X}, \alpha : \mathbb{X} \to \mathsf{CAct}(\mathbb{X}))$ *be a change action model. Then:*
  (i). *If* A *and* B *are flat objects, then the zero map* $0 : A \to B$ *is flat;*
  (ii). *If* A *and* B *are flat objects, and* $f : A \to B$ *and* $g : A \to B$ *are flat morphisms, then their sum* $f + g : A \to B$ *is flat.*

Furthermore, Flat_α *is a Cartesian left additive category.*

Proof. Most of the Cartesian left additive structure is straightforward. However, since the addition is not required to be commutative for arbitrary change actions, we will show that the addition is commutative for flat objects. Using that $\oplus_B$ is an action as in Lemma 3.2, that by [**F.2**] we have that $\oplus_B \circ \pi_1$ is a derivative for $\oplus_B$, and [**CAD.1**], we obtain that:

$$
\begin{aligned}
0_B \oplus_B (f + g) &= (0_B \oplus_B f) \oplus_B g && \text{(Lemma 3.2.iv)} \\
&= (0_B \oplus_B f) \oplus_B (g \oplus_B 0) && \text{(Lemma 3.2.iii)} \\
&= \oplus_B \circ \langle 0_B \oplus_B f, g \oplus_B 0 \rangle && \\
&= \oplus_B \circ (\langle 0_B, g \rangle \oplus_B \langle f, 0 \rangle) && \text{(Lemma 3.5.i)} \\
&= (\oplus_B \circ \langle 0_B, g \rangle) \oplus_B (\boldsymbol{\partial}[\oplus_B] \langle \langle 0_B, g \rangle, \langle f, 0 \rangle \rangle) && \textbf{[CAD.1]} \\
&= (0_B \oplus_B g) \oplus_B \boldsymbol{\partial}[\oplus_B] \langle \langle 0_B, g \rangle, \langle f, 0 \rangle \rangle && \\
&= (0_B \oplus_B g) \oplus_B (\oplus_B \circ \pi_1 \circ \langle \langle 0_B, g \rangle, \langle f, 0 \rangle \rangle) && \textbf{[F.2]} \\
&= (0_B \oplus_B g) \oplus_B (\oplus_B \circ \langle f, 0 \rangle) && \\
&= (0_B \oplus_B g) \oplus_B (f \oplus_B 0) && \\
&= (0_B \oplus_B g) \oplus_B f && \text{(Lemma 3.2.iii)} \\
&= 0_B \oplus_B (g + f) && \text{(Lemma 3.2.iv)}
\end{aligned}
$$

By [**F.3**], $\oplus_B$ is right-injective and we conclude that $f + g = g + f$. $\qquad\square$

We use the action of the change action structure to define the infinitesimal extension. So for a map $f : A \to B$ in Flat_α, define $\varepsilon(f) : A \to B$ as follows:

$$\varepsilon(f) = \oplus_B \circ \langle 0_B \circ !_A, f \rangle = 0 \oplus_B f$$

As such, we may rewrite [**F.4**] as follows:

$$\boldsymbol{\partial}[f] \circ \langle x, \varepsilon(y) \rangle = \varepsilon(\boldsymbol{\partial}[f]) \circ \langle x, y \rangle$$

which is Lemma 4.8.i.

Lemma 4.18. ε *is an infinitesimal extension for* Flat_α.

Proof. We show that ε satisfies [**E.1**], [**E.2**], [**E.3**]. Starting with [**E.1**], we compute that:

$$\varepsilon(0) = \ 0 \oplus_B 0$$

$$= 0 \qquad \text{(Lemma 3.2.iii)}$$

So $\varepsilon(0) = 0$. Next following the same idea as in the proof of Lemma 4.17, we obtain the following:

$$
\begin{aligned}
0_B \oplus_B \varepsilon(f+g) &= 0_B \oplus_B (0 \oplus_B (f+g)) \\
&= 0_B \oplus_B ((0 \oplus_B f) \oplus_B g) && \text{(Lemma 3.2.iv)} \\
&= 0_B \oplus_B (\varepsilon(f) \oplus_B g) && \text{(Lemma 3.2.iv)} \\
&= (0_B \oplus_B 0) \oplus_B (\varepsilon(f) \oplus_B g) && \text{(Lemma 3.2.iii)} \\
&= \oplus_B \circ \langle 0_B \oplus_B 0, \varepsilon(f) \oplus_B g \rangle \\
&= \oplus_B \circ (\langle 0_B, \varepsilon(f) \rangle \oplus_B \langle 0, g \rangle) && \text{(Lemma 3.5.i)} \\
&= (\oplus_B \circ \langle 0_B, \varepsilon(f) \rangle) \oplus_B (\partial[\oplus_B] \circ \langle \langle 0_B, \varepsilon(f) \rangle, \langle 0, g \rangle \rangle) && [\mathbf{CAD.1}] \\
&= (0_B \oplus \varepsilon(f)) \oplus_B (\partial[\oplus_B] \circ \langle \langle 0_B, \varepsilon(f) \rangle, \langle 0, g \rangle \rangle) \\
&= (0_B \oplus \varepsilon(f)) \oplus_B (\oplus_B \circ \pi_1 \circ \langle \langle 0_B, \varepsilon(f) \rangle, \langle 0, g \rangle \rangle) && [\mathbf{F.2}] \\
&= (0_B \oplus \varepsilon(f)) \oplus_B (\oplus_B \circ \langle 0, g \rangle) && [\mathbf{F.2}] \\
&= (0_B \oplus \varepsilon(f)) \oplus_B (0 \oplus_B g) \\
&= (0_B \oplus_B \varepsilon(f)) \oplus_B \varepsilon(g) \\
&= 0_B \oplus_B (\varepsilon(f) +_B \varepsilon(g)) && \text{(Lemma 3.2.iv)}
\end{aligned}
$$

Then by [**F.3**], it follows that $\varepsilon(f+g) = \varepsilon(f) + \varepsilon(g)$. So [**E.1**] holds. Next, it is easy to show that ε is compatible with composition:

$$
\begin{aligned}
\varepsilon(g \circ f) &= 0 \oplus_C (g \circ f) \\
&= (0 \circ f) \oplus_C (g \circ f) \\
&= (0 \oplus_C g) \circ f && \text{(Lemma 3.2.v)} \\
&= \varepsilon(g) \circ f
\end{aligned}
$$

So [**E.2**] holds. Finally, for the projections we compute:

$$
\begin{aligned}
\varepsilon(\pi_i) &= 0 \oplus \pi_i \\
&= \oplus_A \circ \langle 0_A \circ !_{A \times B}, \pi_i \rangle \\
&= \oplus_A \circ \langle \pi_i \circ 0_{A \times B} \circ !_{A \times B}, \pi_i \rangle && \text{(Lemma 3.5.iii)} \\
&= \oplus_A \circ (\pi_i \times \pi_i) \circ \langle 0_{A \times B} \circ !_{A \times B}, 1_{A \times B} \rangle \\
&= \pi_i \circ \oplus_{A \times B} \circ \langle 0_{A \times B} \circ !_{A \times B}, 1_{A \times B} \rangle \\
&= \pi_i \circ (0 \oplus 1_{A \times B}) \\
&= \pi_i \circ \varepsilon(1_{A \times B})
\end{aligned}
$$

So [**E.3**] holds. Thus we conclude that ε is an infinitesimal extension. $\qquad\square$

Next, we observe that for a flat object, the action can be expressed in terms of the infinitesimal extension.

Lemma 4.19. *For any maps $f, g : A \to B$ in Flat_α, the following equality holds:*

$$f \oplus_B g = f + \varepsilon(g)$$

Proof. The proof is a straightforward consequence of linearity of addition:

$$
\begin{aligned}
f + \varepsilon(g) &= f + (0 \oplus_B g) \\
&= (f \oplus_B 0) + (0 \oplus_B g) && \text{(Lemma 3.2.iii)} \\
&= +_B \circ (\langle f \oplus_B 0, 0 \oplus_B g \rangle) \\
&= +_B \circ (\langle f, 0 \rangle \oplus_{B \times B} \langle 0, g \rangle) && \text{(Lemma 3.5.i)} \\
&= (+_B \circ \langle f, 0 \rangle) \oplus_B (\boldsymbol{\partial}[+_B] \circ \langle \langle f, 0 \rangle, \langle 0, g \rangle \rangle) && \textbf{[CAD.1]} \\
&= (f + 0) \oplus_B (\boldsymbol{\partial}[+_B] \circ \langle \langle f, 0 \rangle, \langle 0, g \rangle \rangle) \\
&= f \oplus_B (\boldsymbol{\partial}[+_B] \circ \langle \langle f, 0 \rangle, \langle 0, g \rangle \rangle) \\
&= f \oplus_B (+_B \circ \pi_1 \circ \langle \langle f, 0 \rangle, \langle 0, g \rangle \rangle) && \textbf{[F.2]} \\
&= f \oplus_B (+_B \circ \langle 0, g \rangle) \\
&= f \oplus_B (0 + g) \\
&= f \oplus_B g
\end{aligned}
$$

Thus we conclude that the desired equality holds. □

The difference combinator for Flat_α is defined in the obvious way, that is, $\boldsymbol{\partial}[f]$ is defined as the second component of $\alpha(f)$.

Proposition 4.20. *Let $(\mathbb{X}, \alpha : \mathbb{X} \to \mathsf{CAct}(\mathbb{X}))$ be a change action model. Then Flat_α is a Cartesian difference category.*

Proof. **[C∂.0]** and **[C∂.2]** are simply a restatement of **[CAD.1]** and **[CAD.2]**. **[C∂.3]** and **[C∂.4]** follow immediately from the fact that α preserves finite products and from the structure of products in $\mathsf{CAct}(\mathbb{X})$, while **[C∂.5]** follows from the definition of composition in $\mathsf{CAct}(\mathbb{X})$. So it remains to show that **[C∂.1]**, **[C∂.6]**, and **[C∂.7]** hold. We start by proving **[C∂.1]**, for which it suffices to calculate and apply **[F.2]**:

$$
\begin{aligned}
\boldsymbol{\partial}[f + g] &= \boldsymbol{\partial}[+ \circ \langle f, g \rangle] \\
&= \boldsymbol{\partial}[+_B] \circ \langle \langle f, g \rangle \circ \pi_0, \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \rangle && \textbf{[C∂.5]} \\
&= + \circ \pi_1 \circ \langle \langle f, g \rangle \circ \pi_0, \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \rangle && \textbf{[C∂.2]} \\
&= + \circ \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \\
&= \boldsymbol{\partial}[f] + \boldsymbol{\partial}[g]
\end{aligned}
$$

It is not hard to show that $\boldsymbol{\partial}[0] = 0$ and $\boldsymbol{\partial}[\varepsilon(f)] = \varepsilon(\boldsymbol{\partial}[f])$. For the first property, on one hand we have:

$$
0 \circ \oplus \circ \langle x, y \rangle = 0 = \oplus \circ \langle 0 \circ x, 0 \circ \langle x, y \rangle \rangle
$$

On the other hand, applying **[CAD.1]**:

$$
0 \circ \oplus \circ \langle x, y \rangle = \oplus \circ \langle 0 \circ x, \boldsymbol{\partial}[0] \circ \langle x, y \rangle \rangle
$$

Hence by **[F.3]** we have $\boldsymbol{\partial}[0] \circ \langle x, y \rangle = 0 \circ \langle x, y \rangle$ for any choice of x and y. In particular, $\boldsymbol{\partial}[0] = \boldsymbol{\partial}[0] \circ \langle \pi_0, \pi_1 \rangle = 0 \circ \langle \pi_0, \pi_1 \rangle = 0$ as desired. For the infinitesimal extension, we simply apply the chain rule and the equation $\boldsymbol{\partial}[0] = 0$:

$$
\begin{aligned}
\boldsymbol{\partial}[\varepsilon(f)] &= \boldsymbol{\partial}[\oplus_B \circ \langle 0, f \rangle] \\
&= \boldsymbol{\partial}[\oplus_B] \circ \langle \langle 0, f \rangle \circ \pi_0, \boldsymbol{\partial}[\langle 0, f \rangle] \rangle && \textbf{[C∂.5]}
\end{aligned}
$$

$$
\begin{aligned}
&= \oplus_B \circ \boldsymbol{\partial}[\langle 0, f \rangle] && \textbf{[F.2]} \\
&= \oplus_B \circ \langle \boldsymbol{\partial}[0], \boldsymbol{\partial}[f] \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.4]} \\
&= \oplus_B \circ \langle 0, \boldsymbol{\partial}[f] \rangle && (\boldsymbol{\partial}[0] = 0) \\
&= \varepsilon(\boldsymbol{\partial}[f])
\end{aligned}
$$

So we conclude that **[C$\boldsymbol{\partial}$.1]** holds.

We proceed to prove axioms **[C$\boldsymbol{\partial}$.6.a]** and **[C$\boldsymbol{\partial}$.7.a]** — which, per Lemma 4.7, is equivalent to, and easier to prove than **[C$\boldsymbol{\partial}$.6]** and **[C$\boldsymbol{\partial}$.7]**. Starting with **[C$\boldsymbol{\partial}$.6.a]**, as before, we pick arbitrary $x, y : A \to B$ and calculate:

$$
\begin{aligned}
0 \oplus \boldsymbol{\partial}^2[f] \circ \langle \langle x, 0 \rangle, \langle 0, y \rangle \rangle &= \left(\boldsymbol{\partial}[f] \circ \langle x, 0 \rangle \right) \oplus \left(\boldsymbol{\partial}^2[f] \circ \langle \langle x, 0 \rangle, \langle 0, y \rangle \rangle \right) && \textbf{[C}\boldsymbol{\partial}\textbf{.2]} \\
&= \boldsymbol{\partial}[f] \circ \left(\langle x, 0 \rangle \oplus \langle 0, y \rangle \right) && \textbf{[CAD.1]} \\
&= \boldsymbol{\partial}[f] \circ \left(\langle x \oplus 0, 0 \oplus y \rangle \right) && \text{(Lemma 3.5.i)} \\
&= \boldsymbol{\partial}[f] \circ \langle x, 0 \oplus y \rangle && \text{(Lemma 3.2.iii)} \\
&= 0 \oplus \boldsymbol{\partial}[f] \circ \langle x, y \rangle && \textbf{[F.4]}
\end{aligned}
$$

Hence by **[F.3]**, $\boldsymbol{\partial}^2[f] \circ \langle \langle x, 0 \rangle, \langle 0, y \rangle \rangle = \boldsymbol{\partial}[f] \circ \langle x, y \rangle$ as desired.

Finally, for **[C$\boldsymbol{\partial}$.7.a]**, first observe that by Lemma 4.19 we can compute the following equality for any suitable maps:

$$
\begin{aligned}
(f \oplus g) \oplus (h \oplus k) &= (f + \varepsilon(g)) \oplus (h + \varepsilon(k)) && \text{(Lemma 4.19)} \\
&= f + \varepsilon(g) + \varepsilon(h + \varepsilon(k)) && \text{(Lemma 4.19)} \\
&= f + \varepsilon(g) + \varepsilon(h) + \varepsilon^2(k) && \textbf{[E.1]} \\
&= f + \varepsilon(h) + \varepsilon(g) + \varepsilon^2(k) && \text{(by commutativity of +)} \\
&= f + \varepsilon(h) + \varepsilon(g + \varepsilon(k)) && \textbf{[E.1]} \\
&= (f + \varepsilon(h)) \oplus (g + \varepsilon(k)) && \text{(Lemma 4.19)} \\
&= (f \oplus h) \oplus (g \oplus k)
\end{aligned}
$$

Thus we have the following identity:

$$
(f \oplus g) \oplus (h \oplus k) = (f \oplus h) \oplus (g \oplus k) \tag{4.2}
$$

Alternatively, we could have computed the following for any suitable maps:

$$
\begin{aligned}
(f \oplus g) \oplus (h \oplus k) &= (f + \varepsilon(g)) \oplus (h + \varepsilon(k)) && \text{(Lemma 4.19)} \\
&= f + \varepsilon(g) + \varepsilon(h + \varepsilon(k)) && \text{(Lemma 4.19)} \\
&= f + \varepsilon(g) + \varepsilon(h) + \varepsilon^2(k) && \textbf{[E.1]} \\
&= f + \varepsilon(g + h + \varepsilon(k)) && \textbf{[E.1]} \\
&= f \oplus (g + h + \varepsilon(k)) && \text{(Lemma 4.19)} \\
&= f \oplus ((g + h) \oplus k) && \text{(Lemma 4.19)}
\end{aligned}
$$

Thus we also have the following identity:

$$
(f \oplus g) \oplus (h \oplus k) = f \oplus ((g + h) \oplus k) \tag{4.3}
$$

Then we compute the following for suitable maps:

$$
f \circ ((x \oplus y) \oplus (z \oplus w)) = (f \circ (x \oplus y)) \oplus (\boldsymbol{\partial}[f] \circ \langle x \oplus y, z \oplus w \rangle) \qquad \textbf{[CAD.1]}
$$

$$\begin{aligned}
&= ((f \circ x) \oplus (\boldsymbol{\partial}[f] \circ \langle x, y \rangle)) \oplus (\boldsymbol{\partial}[f] \circ \langle x \oplus y, z \oplus w \rangle) && [\textbf{CAD.1}] \\
&= ((f \circ x) \oplus (\boldsymbol{\partial}[f] \circ \langle x, y \rangle)) \oplus (\boldsymbol{\partial}[f] \circ (\langle x, z \rangle \oplus \langle y, w \rangle)) \\
&= ((f \circ x) \oplus (\boldsymbol{\partial}[f] \circ \langle x, y \rangle)) \oplus ((\boldsymbol{\partial}[f] \circ \langle x, z \rangle) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, z \rangle, \langle y, w \rangle \rangle)) && [\textbf{CAD.1}] \\
&= (f \circ x) \oplus (((\boldsymbol{\partial}[f] \circ \langle x, y \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, z \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, z \rangle, \langle y, w \rangle \rangle)) && (4.3)
\end{aligned}$$

So we have the following identity:

$$f \circ ((x \oplus y) \oplus (z \oplus w)) =$$
$$(f \circ x) \oplus (((\boldsymbol{\partial}[f] \circ \langle x, y \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, z \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, z \rangle, \langle y, w \rangle \rangle)) \qquad (4.4)$$

However, by swapping the middle two arguments on the left hand side using (4.2) we compute:

$$\begin{aligned}
&(f \circ x) \oplus (((\boldsymbol{\partial}[f] \circ \langle x, y \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, z \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, z \rangle, \langle y, w \rangle \rangle)) \\
&= f \circ ((x \oplus y) \oplus (z \oplus w)) && (4.4) \\
&= f \circ ((x \oplus z) \oplus (y \oplus w)) && (4.2) \\
&= (f \circ x) \oplus (((\boldsymbol{\partial}[f] \circ \langle x, z \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, y \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle)) && (4.4) \\
&= (f \circ x) \oplus (((\boldsymbol{\partial}[f] \circ \langle x, y \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, z \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle)) && \text{(by com. of +)}
\end{aligned}$$

So we have that:

$$\begin{aligned}
&(f \circ x) \oplus (((\boldsymbol{\partial}[f] \circ \langle x, y \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, z \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, z \rangle, \langle y, w \rangle \rangle)) \\
&= (f \circ x) \oplus (((\boldsymbol{\partial}[f] \circ \langle x, y \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, z \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle))
\end{aligned}$$

By applying [**F.3**], we obtain that:

$$\begin{aligned}
&((\boldsymbol{\partial}[f] \circ \langle x, y \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, z \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, z \rangle, \langle y, w \rangle \rangle) \\
&= ((\boldsymbol{\partial}[f] \circ \langle x, y \rangle) + (\boldsymbol{\partial}[f] \circ \langle x, z \rangle)) \oplus (\boldsymbol{\partial}^2[f] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle)
\end{aligned}$$

By applying [**F.3**] again we finally obtain $\boldsymbol{\partial}^2[f] \langle \langle x, y \rangle, \langle z, w \rangle \rangle = \boldsymbol{\partial}^2[f] \langle \langle x, z \rangle, \langle y, w \rangle \rangle$ as desired. So we conclude that Flat_α is a Cartesian difference category. $\qquad \square$

4.5. **Linear Maps and ε-Linear Maps.** An important subclass of maps in a Cartesian differential category is the subclass of *linear maps* [BCS09, Definition 2.2.1]. One can also define linear maps in a Cartesian difference category by using the same definition.

Definition 4.21. In a Cartesian difference category, a map f is **linear** if the following equality holds: $\boldsymbol{\partial}[f] = f \circ \pi_1$.

Using element-like notation, a map f is linear if $\boldsymbol{\partial}[f](x, y) = f(y)$. Linear maps in a Cartesian difference category satisfy many of the same properties found in [BCS09, Lemma 2.2.2].

Lemma 4.22. *In a Cartesian difference category,*
  (i) *If $f : A \to B$ is linear then $\varepsilon(f) = f \circ \varepsilon(1_A)$.*
  (ii) *If $f : A \to B$ is linear, then f is additive.*
  (iii) *Identity maps, projection maps, and zero maps are linear.*
  (iv) *The composite, sum, pairing, and product of linear maps are linear.*

(v) *If $f : A \to B$ and $k : C \to D$ are linear, then for any map $g : B \to C$:*
$$\boldsymbol{\partial}[k \circ g \circ f] = k \circ \boldsymbol{\partial}[g] \circ (f \times f)$$

(vi) *If an isomorphism is linear, then its inverse is linear.*

(vii) *For any object A, $\oplus_A$ and $+_A$ are linear.*

Proof. Most of the above results are either immediate or admit a similar proof to the ones in [BCS09, Lemma 2.2.2]. We will prove i, as it differs from the differential setting:

$$
\begin{aligned}
f \circ \varepsilon(1_A) =\ & f \circ (0 + \varepsilon(1_A)) \\
=\ & f \circ 0 + \varepsilon(\boldsymbol{\partial}[f] \circ \langle 0, 1_A \rangle) & [\mathbf{C\partial.0}] \\
=\ & f \circ \pi_1 \circ \langle 0, 0 \rangle + \varepsilon(\boldsymbol{\partial}[f] \circ \langle 0, 1_A \rangle) \\
=\ & \boldsymbol{\partial}[f] \circ \langle 0, 0 \rangle + \varepsilon(f \circ \pi_1 \circ \langle 0, 1_A \rangle) & (f \text{ is linear}) \\
=\ & 0 + \varepsilon(f \circ 1_A) & [\mathbf{C\partial.2}] \\
=\ & \varepsilon(f)
\end{aligned}
$$

So $\varepsilon(f) = f \circ \varepsilon(1_A)$. $\qquad\square$

Using element-like notation, the first point of the above lemma says that if f is linear then $f(\varepsilon(x)) = \varepsilon(f(x))$. It is important to note that while all linear maps are additive, the converse is not necessarily true, see [BCS09, Corollary 2.3.4]. That said, an immediate consequence of the above lemma is that the subcategory of linear maps of a Cartesian difference category has finite biproducts.

Another interesting subclass of maps is the subclass of ε-linear maps, which are maps whose infinitesimal extension is linear.

Definition 4.23. In a Cartesian difference category, a map f is ε-**linear** if $\varepsilon(f)$ is linear.

Lemma 4.24. *In a Cartesian difference category,*

(i) *If $f : A \to B$ is ε-linear then $f \circ (x + \varepsilon(y)) = f \circ x + \varepsilon(f) \circ y$;*

(ii) *Every linear map is ε-linear;*

(iii) *The composite, sum, and pairing of ε-linear maps is ε-linear;*

(iv) *If an isomorphism is ε-linear, then its inverse is again ε-linear.*

(v) *If ε is nilpotent (.i.e for every f there exists a $k \in \mathbb{N}$ such that $\varepsilon^k(f) = 0$), then for every map f, $\boldsymbol{\partial}[f] \circ \langle 0, 1 \rangle$ is linear.*

Proof. The first four (i) through (iv) are straightforward and so we leave them as an exercise for the reader. For (v), suppose that ε is nilpotent. We need to show that $\boldsymbol{\partial}[\varepsilon(\boldsymbol{\partial}[f] \circ \langle 0, 1 \rangle)] = \varepsilon(\boldsymbol{\partial}[f] \circ \langle 0, 1 \rangle) \circ \pi_1$. So we compute:

$$
\begin{aligned}
\boldsymbol{\partial}[\varepsilon(\boldsymbol{\partial}[f] \circ \langle 0, 1 \rangle)] =\ & \varepsilon\left(\boldsymbol{\partial}\left[\boldsymbol{\partial}[f] \circ \langle 0, 1 \rangle\right]\right) & [\mathbf{C\partial.1}] \\
=\ & \varepsilon\left(\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ (\langle 0, 1 \rangle \times \langle 0, 1 \rangle)\right) & (\text{Lemma 4.22.v}) \\
=\ & \varepsilon\left(\boldsymbol{\partial}\left[\boldsymbol{\partial}[f]\right] \circ \langle \langle 0, \pi_0 \rangle, \langle 0, \pi_1 \rangle \rangle\right) \\
=\ & \varepsilon\left(\boldsymbol{\partial}[f] \circ \langle 0 + \varepsilon(\pi_0), \pi_1 \rangle\right) & [\mathbf{C\partial.6}] \\
=\ & \varepsilon\left(\boldsymbol{\partial}[f] \circ \langle \varepsilon(\pi_0), \pi_1 \rangle\right) \\
=\ & \varepsilon\left(\boldsymbol{\partial}[f] \circ (\langle 0, \pi_1 \rangle + \langle \varepsilon(\pi_0), 0 \rangle)\right) \\
=\ & \varepsilon\left(\boldsymbol{\partial}[f] \circ (\langle 0, \pi_1 \rangle + \langle \varepsilon(\pi_0), \varepsilon(0) \rangle)\right) & [\mathbf{E.1}] \\
=\ & \varepsilon\left(\boldsymbol{\partial}[f] \circ (\langle 0, \pi_1 \rangle + \varepsilon(\langle \pi_0, 0 \rangle))\right) & (\text{Lemma 4.3})
\end{aligned}
$$

$$
\begin{aligned}
&= \ \varepsilon \left(\boldsymbol{\partial}[f] \circ \langle 0, \pi_1 \rangle + \varepsilon \left(\boldsymbol{\partial}[\boldsymbol{\partial}[f]] \circ \langle \langle 0, \pi_1 \rangle, \langle \pi_0, 0 \rangle \rangle \right)\right) && \textbf{[C}\boldsymbol{\partial}\textbf{.0]} \\
&= \ \varepsilon \left(\boldsymbol{\partial}[f] \circ \langle 0, \pi_1 \rangle\right) + \varepsilon^2 \left(\boldsymbol{\partial}[\boldsymbol{\partial}[f]] \circ \langle \langle 0, \pi_1 \rangle, \langle \pi_0, 0 \rangle \rangle \right) && \textbf{[E.1]} \\
&= \ \varepsilon \left(\boldsymbol{\partial}[f]\right) \circ \langle 0, \pi_1 \rangle + \varepsilon^2 \left(\boldsymbol{\partial}[\boldsymbol{\partial}[f]]\right) \circ \langle \langle 0, \pi_1 \rangle, \langle \pi_0, 0 \rangle \rangle && \textbf{[E.2]} \\
&= \ \varepsilon \left(\boldsymbol{\partial}[f]\right) \circ \langle 0, \pi_1 \rangle + \varepsilon^k \left(\boldsymbol{\partial}[\boldsymbol{\partial}[f]]\right) \circ \langle \langle 0, \pi_1 \rangle, \langle \pi_0, 0 \rangle \rangle && \\
&&& \text{(Iterating Lemma 4.8.iii)} \\
&= \ \varepsilon \left(\boldsymbol{\partial}[f]\right) \circ \langle 0, \pi_1 \rangle + \boldsymbol{\partial}[\boldsymbol{\partial}[\varepsilon^k \left(f\right)]] \circ \langle \langle 0, \pi_1 \rangle, \langle \pi_0, 0 \rangle \rangle && \textbf{([C}\boldsymbol{\partial}\textbf{.1])} \\
&= \ \varepsilon \left(\boldsymbol{\partial}[f]\right) \circ \langle 0, \pi_1 \rangle + \boldsymbol{\partial}[\boldsymbol{\partial}[0]] \circ \langle \langle 0, \pi_1 \rangle, \langle \pi_0, 0 \rangle \rangle && (\varepsilon \text{ is nilpotent}) \\
&= \ \varepsilon \left(\boldsymbol{\partial}[f]\right) \circ \langle 0, \pi_1 \rangle + 0 \circ \langle \langle 0, \pi_1 \rangle, \langle \pi_0, 0 \rangle \rangle && \textbf{[C}\boldsymbol{\partial}\textbf{.1]} \\
&= \ \varepsilon \left(\boldsymbol{\partial}[f]\right) \circ \langle 0, \pi_1 \rangle + 0 && \\
&= \ \varepsilon(\boldsymbol{\partial}[f]) \circ \langle 0, 1 \rangle \circ \pi_1 && \\
&= \ \varepsilon(\boldsymbol{\partial}[f] \circ \langle 0, 1 \rangle) \circ \pi_1 && \textbf{[E.2]}
\end{aligned}
$$

So we conclude that $\varepsilon(\boldsymbol{\partial}[f]) \circ \langle 0, 1 \rangle$ is linear and hence $\boldsymbol{\partial}[f] \circ \langle 0, 1 \rangle$ is ε-linear.   □

Using element-like notation, the first point of the above lemma says that if f is ε-linear then $f(x + \varepsilon(y)) = f(x) + \varepsilon(f(y))$. So ε-linear maps are additive on "infinitesimal" elements (i.e. those of the form $\varepsilon(y)$). For a Cartesian differential category, linear maps in the Cartesian difference category sense are precisely the same as the Cartesian differential category sense [BCS09, Definition 2.2.1], while every map is ε-linear since $\varepsilon = 0$.

5. Examples of Cartesian Difference Categories

5.1. **Smooth Functions.** As we have shown in Proposition 4.10, every Cartesian differential category is a Cartesian difference category where the infinitesimal extension is zero. As a particular example, we consider the category of real smooth functions which, as mentioned above, can be considered to be the canonical (and motivating) example of a Cartesian differential category.

Definition 5.1. Let $\mathbb{R}$ denote the set of real numbers. The category SMOOTH is the category whose objects are Euclidean spaces $\mathbb{R}^n$ (including the point $\mathbb{R}^0 = \{*\}$), and whose maps are smooth functions $F : \mathbb{R}^n \to \mathbb{R}^m$.

SMOOTH is a Cartesian left additive category where the product structure is given by the standard Cartesian product of Euclidean spaces and where the additive structure is defined by point-wise addition, $(F + G)(\vec{x}) = F(\vec{x}) + G(\vec{x})$ and $0(\vec{x}) = (0, \ldots, 0)$, where $\vec{x} \in \mathbb{R}^n$. SMOOTH is a Cartesian differential category where the differential combinator is defined by the directional derivative of smooth functions. Explicitly, for a smooth function $F : \mathbb{R}^n \to \mathbb{R}^m$, which is in fact a tuple of smooth functions $F = (f_1, \ldots, f_n)$ where $f_i : \mathbb{R}^n \to \mathbb{R}$, $\mathsf{D}[F] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ is defined as follows:

$$
\mathsf{D}[F](\vec{x}, \vec{y}) := \left(\sum_{i=1}^n \frac{\partial f_1}{\partial u_i}(\vec{x}) y_i, \ldots, \sum_{i=1}^n \frac{\partial f_n}{\partial u_i}(\vec{x}) y_i \right)
$$

where $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Alternatively, $\mathsf{D}[F]$ can also be defined in terms of the Jacobian matrix of F. Therefore SMOOTH is a Cartesian difference category with infinitesimal extesion $\varepsilon = 0$ and with difference combinator D. Since $\varepsilon = 0$, the induced

action is simply $\vec{x} \oplus \vec{y} = \vec{x}$. A smooth function is linear in the Cartesian difference category sense precisely if it is $\mathbb{R}$-linear in the classical sense, and every smooth function is ε-linear since $\varepsilon(f) = 0$ is linear.

5.2. **Calculus of Finite Differences.** The calculus of finite differences [JJ65, Gle05] is a field which aims to apply methods from differential calculus to discrete settings. It does so by introducing a "discrete derivative" or "finite difference" operator Δ which, when applied to an integer-valued function $f : \mathbb{Z} \to \mathbb{Z}$, gives the function $\Delta(f)$ defined by

$$\Delta(f)(x) = f(x+1) - f(x)$$

Here we show that a generalization of the finite difference operator gives an example of a Cartesian difference category (but *not* a Cartesian differential category). This example was the main motivating example for developing Cartesian difference categories. The behaviour of the calculus of finite differences, as well as some of its generalizations (notably, the Boolean differential calculus [Tha81, SP09]) is captured by the category of abelian groups and arbitrary set functions between them, equipped with a suitable Cartesian difference structure.

Definition 5.2. $\overline{\mathsf{Ab}}$ is the category whose objects are abelian groups G (where we use additive notation for group structure) and where a map $f : G \to H$ is simply an arbitrary function between them (and therefore does not necessarily preserve the group structure).

Proposition 5.3. $\overline{\mathsf{Ab}}$ *is a Cartesian left additive category where the product structure is given by the standard Cartesian product of sets and where the additive structure is again given by point-wise addition, $(f + g)(x) = f(x) + g(x)$ and $0(x) = 0$. Furthermore, $\overline{\mathsf{Ab}}$ is a Cartesian difference category where the infinitesimal extension is given by the identity, that is, $\varepsilon(f) = f$, and and where the difference combinator $\boldsymbol{\partial}$ is defined as follows for a map $f : G \to H$:*

$$\boldsymbol{\partial}[f](x, y) = f(x + y) - f(x)$$

On the other hand, note that $\boldsymbol{\partial}$ is not a differential combinator for $\overline{\mathsf{Ab}}$ since it does not satisfy [**CD.6**] and part of [**CD.2**]. Indeed, since f is not necessarily a group homomorphism, [**CD.2**] fails to hold as:

$$\boldsymbol{\partial}[f](x, y + z) = f(x + y + z) - f(x)$$

is not necessarily equal to:

$$\boldsymbol{\partial}[f](x, y) + \boldsymbol{\partial}[f](x, z) = f(x + y) - f(x) + f(x + z) - f(x)$$

$\boldsymbol{\partial}$ does satisfy [**C$\boldsymbol{\partial}$.2**] and [**C$\boldsymbol{\partial}$.6**], as well as [**C$\boldsymbol{\partial}$.0**]. which in this case are respectively:

$$\boldsymbol{\partial}[f](x, y + z) = \boldsymbol{\partial}[f](x, y) + \boldsymbol{\partial}[f](x + y, z)$$

$$\boldsymbol{\partial}[\boldsymbol{\partial}[f]]((x, y), (0, z)) = \boldsymbol{\partial}[f](x + y, z)$$

$$f(x + y) = f(x) + \boldsymbol{\partial}[f](x, y)$$

However, as noted in [BLR18], it is interesting to note that $\boldsymbol{\partial}$ does satisfy [**CD.1**], the second part of [**CD.2**], [**CD.3**], [**CD.4**], [**CD.5**], [**CD.7**], and [**C$\boldsymbol{\partial}$.6.a**]. It is worth emphasizing that in [BLR18], the goal was to drop the addition and develop a "non-additive" version of Cartesian differential categories, whereas the current presentation keeps the additive structure while suitably relaxing the differential combinator to a difference combinator.

In $\overline{\mathsf{Ab}}$, since the infinitesimal operator is given by the identity, the induced action is simply the addition, $x \oplus y = x + y$. On the other hand, the linear maps in $\overline{\mathsf{Ab}}$ are precisely the group homomorphisms. Indeed, f is linear if $\boldsymbol{\partial}[f](x, y) = f(y)$. But by $[\mathbf{C}\boldsymbol{\partial}.\mathbf{0}]$ and $[\mathbf{C}\boldsymbol{\partial}.\mathbf{2}]$, we get that:

$$f(x + y) = f(x) + \boldsymbol{\partial}[f](x, y) = f(x) + f(y) \qquad f(0) = \boldsymbol{\partial}[f](x, 0) = 0$$

So f is a group homomorphism. Conversely, whenever f is a group homomorphism:

$$\boldsymbol{\partial}[f](x, y) = f(x + y) - f(x) = f(x) + f(y) - f(x) = f(y)$$

So f is linear. Since $\varepsilon(f) = f$, the ε-linear maps are precisely the linear maps.

5.3. **Module Morphisms.** Here we provide a simple example of a Cartesian difference category whose difference combinator is also a differential combinator, but where the infinitesimal extension is neither zero nor the identity.

Definition 5.4. Let R be a commutative semiring. We define the category MOD_R as the category whose objects are all R-modules and whose maps are all the R-linear maps between them.

MOD_R has finite biproducts and is, therefore, a Cartesian additive (in particular, left additive) category. Every $r \in R$ induces an infinitesimal extension ε^r defined by scalar multiplication, $\varepsilon^r(f)(m) = rf(m)$. For any choice of r, the category MOD_R is a Cartesian difference category with the infinitesimal extension ε^r and its difference combinator $\boldsymbol{\partial}$ defined as:

$$\boldsymbol{\partial}[f](m, n) = f(n)$$

R-linearity of f assures that $[\mathbf{C}\boldsymbol{\partial}.\mathbf{0}]$ holds, while the remaining Cartesian difference axioms hold trivially. In fact, $\boldsymbol{\partial}$ is also a differential combinator and therefore MOD_R is also a Cartesian differential category, but note that the Cartesian difference structure of MOD_R is (whenever R is non-zero) different than the Cartesian difference structure that corresponds to this differential structure. The induced action is given by $m \oplus n = m + rn$. By definition of $\boldsymbol{\partial}$, every map in MOD_R is linear, and by definition of ε^r and R-linearity, every map is also ε-linear.

5.4. **Stream calculus.** It is common knowledge that streams, i.e. infinite sequences of values, can be studied using methods from differential calculus. For example, in [Rut05], a notion of *stream derivative* operator is introduced, and streams are characterized as the solutions of stream differential equations involving stream derivatives. More recent work in the setting of causal functions between streams of real numbers [SJ19, SK19] has focused on extending the "classical" notion of the derivative of a real-valued function to stream-valued functions.

We seek now to show that causal functions between streams are indeed endowed with the structure of a Cartesian difference category, the corresponding difference combinator capturing the change of the stream over time. For this we will introduce an idempotent infinitesimal extension on streams that plays a similar role in this setting as Rutten's stream derivative operator [Rut05], which is given by discarding the head of the stream. On the other hand, our work is more closely related to Sprunger et al.'s work [SJ19, SK19] as it focuses on the differentiation of functions between streams, rather than the description of single streams in terms of differential equations.

For a set A, let A^ω denote the set of infinite sequences of elements of A. We write $[a_i]$ for the infinite sequence $[a_i] = (a_0, a_1, a_2, \ldots)$ and $a_{i:\omega}$ for the (infinite) subsequence $(a_i, a_{i+1}, \ldots)$. A function $f : A^\omega \to B^\omega$ is **causal** whenever the n-th element $f([a_i])_n$ of the output sequence only depends on the first n elements of $[a_i]$. More formally, the function f is causal if and only if, whenever $a_{0:n} = b_{0:n}$, then $f([a_i])_{0:n} = f([b_i])_{0:n}$.

We will restrict ourselves to considering streams over abelian groups[1], so let $\overline{\mathsf{Ab}}^\omega$ be the category whose objects are all the abelian groups and where a morphism from G to H in $\overline{\mathsf{Ab}}^\omega$ is a causal map from G^ω to H^ω. $\overline{\mathsf{Ab}}^\omega$ is a Cartesian left-additive category, where the product is given by the standard product of abelian groups and where the additive structure is lifted point-wise from the structure of $\overline{\mathsf{Ab}}$.

More concretely, whenever G, H are abelian groups, the hom-set $\overline{\mathsf{Ab}}^\omega(G, H)$ is endowed with the additive structure that comes from setting

$$(f + g)([a_i])_n = f([a_i])_n + g([a_i])_n \qquad\qquad 0([a_i])_n = 0$$

At this point, one might think that the "natural" choice of an infinitesimal extension in this setting would be something akin to Rutten's stream derivative operator, which is given by dropping the first element of the stream, that is, $([a_i])' = ([a_i])_{1:\omega}$. This, however, is not a causal function and so it does not exist in the category $\overline{\mathsf{Ab}}^\omega$. That said, it is possible to construct a larger category where this operator is used for differentiation, it however fails to satisfy [**C∂.6**]. We instead an define infinitesimal extension for $\overline{\mathsf{Ab}}^\omega$ as the *truncation* operator $\mathbf{z}$.

Definition 5.5. For an abelian group A, define the **truncation operator** $\mathbf{z}_A : A^\omega \to A^\omega$ as follows:

$$(\mathbf{z}[a_i])_0 = 0 \qquad\qquad (\mathbf{z}[a_i])_{j+1} = a_{j+1}$$

Note that $\mathbf{z}_A$ is a monoid homomorphism according to the pointwise monoid structure on A^ω. Thus it is straightforward to see that we obtain an infinitesimal extension $\mathbf{z}$ for $\overline{\mathsf{Ab}}^\omega$.

Theorem 5.6. *The category* $\overline{\mathsf{Ab}}^\omega$ *is a Cartesian difference category, with the infinitesimal extension* $\varepsilon(f) = \mathbf{z} \circ f$ *and a difference operator defined as:*

$$\boldsymbol{\partial}[f]([a_i], [b_i])_0 = f([a_i] + [b_i])_0 - f([a_i])_0$$
$$\boldsymbol{\partial}[f]([a_i], [b_i])_{n+1} = f([a_i] + \mathbf{z}([b_i]))_{n+1} - f([a_i])_{n+1}$$

Proof. We leave it to the reader to check for themselves that $\mathbf{z}$ is an infinitesimal extension. [**C∂.0**] is satisfied because of causality. Indeed, since f is causal, $f([a_i] + \mathbf{z}([b_i]))$ depends only on $([a_i] + \mathbf{z}([b_i]))_0$, but note that, by definition of $\mathbf{z}$, this term is precisely equal to a_0. Hence we obtain the following:

$$(f([a_i]) + \mathbf{z}(\boldsymbol{\partial}[f]([a_i], [b_i])))_0 = f([a_i])_0 = f([a_i] + \mathbf{z}([b_i]))_0$$

For any other index, we have that:

$$\begin{aligned}
(f([a_i]) + \mathbf{z}(\boldsymbol{\partial}[f]([a_i], [b_i])))_{i+1} &= f([a_i])_{i+1} + \mathbf{z}(\boldsymbol{\partial}[f]([a_i], [b_i]))_{i+1} \\
&= f([a_i])_{i+1} + \boldsymbol{\partial}[f]([a_i], [b_i])_{i+1} \\
&= f([a_i])_{i+1} + f([a_i] + \mathbf{z}([b_i]))_{i+1} - f([a_i])_{i+1}
\end{aligned}$$

---

[1]A similar approach to the one in [SK19] is possible where we consider streams on arbitrary difference categories, and lift the difference operator of the underlying category to its category of streams, although it would complicate the presentation of this section without gaining clarity.

$$= f\left([a_i] + \mathbf{z}([b_i])\right)_{i+1}$$

We also explicitly prove the part of [**C∂.1**] concerning the infinitesimal extension **z**:

$$\boldsymbol{\partial}[\mathbf{z}]\left([a_i], [b_i]\right)_0 = \mathbf{z}\left([a_i] + [b_i]\right)_0 - \mathbf{z}([a_i])_0 = 0 = \mathbf{z}([b_i])_0$$
$$\boldsymbol{\partial}[\mathbf{z}]\left([a_i], [b_i]\right)_{n+1} = \mathbf{z}\left([a_i] + \mathbf{z}([b_i])\right)_{n+1} - \mathbf{z}([a_i])_{n+1} = b_{n+1} = \mathbf{z}([b_i])_{n+1}$$

Finally, we show that [**C∂.3**] holds for the identity map, which is a matter of simple calculation as well (linearity of the projection maps follows by an almost identical argument). However, as we remark later, this property would fail to hold if we had chosen a more "natural" candidate for the difference operator.

$$\boldsymbol{\partial}[1]\left([a_i], [b_i]\right)_0 = a_0 + b_0 - a_0 = b_0$$
$$\boldsymbol{\partial}[1]\left([a_i], [b_i]\right)_{n+1} = \left([a_i] + \mathbf{z}([b_i])\right)_{n+1} - a_{n+1} = b_{n+1}$$

The remaining axioms can be shown to hold by similar pointwise reasoning; the corresponding calculations are very similar to the case of $\overline{\mathsf{Ab}}$. □

Remark 5.7. One might expect the difference operator in $\overline{\mathsf{Ab}}^\omega$ to be given by the simpler expression

$$\boldsymbol{\partial}[f]([a_i], [b_i]) = f([a_i] + \mathbf{z}([b_i])) - f([a_i])$$

While this satisfies some of the Cartesian difference axioms (notably [**C∂.0**]) it does not satisfy all of them: for example, [**C∂.3**] fails to hold since $\boldsymbol{\partial}[1]\left([a_i], [b_i]\right) = \mathbf{z}([b_i]) \neq [b_i]$.

Note the similarities between the difference combinator on $\overline{\mathsf{Ab}}$ and that on $\overline{\mathsf{Ab}}^\omega$. The induced action can be computed out to be:

$$([a_i] \oplus [b_i])_0 = a_0 \qquad\qquad ([a_i] \oplus [b_i])_{n+1} = a_{n+1} + b_{n+1} \ (\equiv\ a_{n+1} \oplus b_{n+1})$$

The linear maps (in the Cartesian difference category sense) in $\overline{\mathsf{Ab}}^\omega$ are precisely those maps f that are group homomorphisms (when the set of streams G^ω is equipped with the structure lifted pointwise from the group G) satisfying the additional property that whenever $[a_i]_{1:\omega} = [b_i]_{1:\omega}$ then $f([a_i])_{n+1} = f([b_i])_{n+1}$, but this is far from evident. The reader can easily verify that any such homomorphism is linear, we prove here the converse.

Proposition 5.8. *Any linear map f in $\overline{\mathsf{Ab}}^\omega$ is a group homomorphism. Furthermore, whenever $[a_i]_{1:\omega} = [b_i]_{1:\omega}$, the map f satisfies $f([a_i])_{n+1} = f([b_i])_{n+1}$.*

Proof. The first part of the proposition is simply a corollary of Lemma 4.22.ii, according to which every linear map in $\overline{\mathsf{Ab}}^\omega$ is additive and, therefore, a group homomorphism. For the second property, since f is linear, we have:

$$f\left([b_i]\right)_{n+1} = \boldsymbol{\partial}[f]\left([a_i], [b_i]\right)_{n+1} = f\left([a_i] + \mathbf{z}([b_i])\right)_{n+1} - f([a_i])_{n+1}$$

Therefore $f\left([a_i] + \mathbf{z}([b_i])\right)_{n+1} = f\left([a_i]\right)_{n+1} + f\left([b_i]\right)_{n+1}$. By setting $[a_i] = [0]$ in the above equation, and since f preserves the identity element, we establish $f\left(\mathbf{z}([b_i])\right)_{n+1} = f\left([b_i]\right)_{n+1}$, from which the second part of the desired property follows as an immediate corollary. □

On the other hand, since every map of the form $f \circ \mathbf{z}$ verifies this second property of "insensitivity" to the initial element of the stream, it follows that for a map f to be ε-linear it is sufficient (and necessary) that $f \circ \varepsilon$ be a group homomorphism.

6. Tangent Bundles in Cartesian Difference Categories

In this section, we show that the difference combinator of a Cartesian difference category induces a monad, called the *tangent bundle monad*. This construction is a generalization of the tangent bundle monad for Cartesian differential categories [CC14, Man12b]. Furthermore, we show that a full subcategory of the Eilenberg-Moore category and the Kleisli category of the tangent bundle monad is again a Cartesian difference category. The general intuition of these categories for the tangent bundle monad are more or less the same as explained in [CC14, Section 3.2]. The maps of the Kleisli category are generalized vector fields, while the considered full subcategory of the Eilenberg-Moore category consists of objects equipped with a linear map which associates tangent vectors to points.

6.1. **The Tangent Bundle Monad.** If only to introduce notation, recall that a monad on a category $\mathbb{X}$ is a triple (T, μ, η) consisting of an endofunctor $\mathsf{T} : \mathbb{X} \to \mathbb{X}$, and two natural transformations $\mu : \mathsf{T}^2 \Rightarrow \mathsf{T}$ and $\eta : 1_{\mathbb{X}} \Rightarrow \mathsf{T}$ (where $1_{\mathbb{X}} : \mathbb{X} \to \mathbb{X}$ is the identity functor), such that the following equalities hold for all objects $A \in \mathbb{X}$:

$$\mu_A \circ \eta_{\mathsf{T}(A)} = 1_{\mathsf{T}(A)} = \mu_A \circ \mathsf{T}(\eta_A) \qquad \mu_A \circ \mu_{\mathsf{T}(A)} = \mu_A \circ \mathsf{T}(\mu_A) \tag{6.1}$$

Now let $\mathbb{X}$ be a Cartesian difference category with infinitesimal extension ε and difference combinator $\boldsymbol{\partial}$. Define the functor $\mathsf{T} : \mathbb{X} \to \mathbb{X}$ as follows:

$$\mathsf{T}(A) = A \times A \qquad \mathsf{T}(f) = \langle f \circ \pi_0, \boldsymbol{\partial}[f] \rangle$$

and define the natural transformations $\eta : 1_{\mathbb{X}} \Rightarrow \mathsf{T}$ and $\mu : \mathsf{T}^2 \Rightarrow \mathsf{T}$ as follows:

$$\eta_A := \langle 1_A, 0 \rangle \qquad \mu_A := \langle \pi_{00}, \pi_{10} + \pi_{01} + \varepsilon(\pi_{11}) \rangle$$

where $\pi_{ij} = \pi_i \circ \pi_j$. Before providing the proof that (T, μ, η) is indeed a monad for any Cartesian difference category, let us provide some examples for intuition.

Example 6.1. For a Cartesian differential category, since $\varepsilon = 0$, the induced monad is precisely the monad induced by its tangent category structure [CC14, Man12b]. For example, in the Cartesian differential category SMOOTH (as defined in Section 5.1), one has that $\mathsf{T}(\mathbb{R}^n) = \mathbb{R}^n \times \mathbb{R}^n$, which is in fact the classical tangent bundle over $\mathbb{R}^n$ from differential geometry, and also that $\mathsf{T}(F)(\vec{x}, \vec{y}) = (F(\vec{x}), \mathsf{D}[F](\vec{x}, \vec{y}))$, $\eta_{\mathbb{R}^n}(\vec{x}) = (\vec{x}, \vec{0})$, and $\mu_{\mathbb{R}^n}((\vec{x}, \vec{y}), (\vec{z}, \vec{w})) = (\vec{x}, \vec{y} + \vec{z})$.

Example 6.2. In the Cartesian difference category $\overline{\mathsf{Ab}}$ (as defined in Section 5.2), the monad is given by $\mathsf{T}(G) = G \times G$ $\mathsf{T}(f)(x, y) = (f(x), f(x + y) - f(x))$, $\eta_G(x) = (x, 0)$, and $\mu_G((x, y), (z, w)) = (x, y + z + w)$.

We now show that the above construction is indeed a monad.

Proposition 6.3. *Let $\mathbb{X}$ be a Cartesian difference category with infinitesimal extension ε and difference combinator $\boldsymbol{\partial}$. Then (T, μ, η), as defined above, is a monad on $\mathbb{X}$.*

Proof. That T is a functor follows from the fact that the change action model $\alpha : \mathbb{X} \to \mathsf{CAct}(\mathbb{X})$ from Proposition 4.14 is a functor. Specifically, T is the second component of α. Next we show the naturality of η and μ. For η we have:

$$\begin{aligned} \mathsf{T}(f) \circ \eta_A = & \ \langle f \circ \pi_0, \boldsymbol{\partial}[f] \rangle \circ \eta_A \\ = & \ \langle f \circ \pi_0 \circ \eta_A, \boldsymbol{\partial}[f] \circ \eta_A \rangle \\ = & \ \langle f \circ \pi_0 \circ \langle 1_A, 0 \rangle, \boldsymbol{\partial}[f] \circ \langle 1_A, 0 \rangle \rangle \end{aligned}$$

$$\begin{aligned}
&= \langle f, 0 \rangle && [\mathbf{C}\partial.\mathbf{2}] \\
&= \langle 1_B, 0 \rangle \circ f \\
&= \eta_B \circ f
\end{aligned}$$

For the naturality of μ, first it is straightforward to check that:

$$\mathsf{T}^2(f) = \left\langle \langle f \circ \pi_{00}, \boldsymbol{\partial}[f] \circ \pi_0 \rangle, \langle \boldsymbol{\partial}[f] \circ (\pi_0 \times \pi_0), \boldsymbol{\partial}^2[f] \rangle \right\rangle$$

and therefore we compute:

$$\begin{aligned}
\mu_B \circ \mathsf{T}^2(f) &= \langle \pi_{00}, \pi_{10} + \pi_{01} + \varepsilon(\pi_{11}) \rangle \circ \mathsf{T}^2(f) \\
&= \langle \pi_{00} \circ \mathsf{T}^2(f), (\pi_{10} + \pi_{01} + \varepsilon(\pi_{11})) \circ \mathsf{T}^2(f) \rangle \\
&= \langle \pi_{00} \circ \mathsf{T}^2(f), \pi_{10} \circ \mathsf{T}^2(f) + \pi_{01} \circ \mathsf{T}^2(f) + \varepsilon(\pi_{11}) \circ \mathsf{T}^2(f) \rangle \\
&= \langle \pi_{00} \circ \mathsf{T}^2(f), \pi_{10} \circ \mathsf{T}^2(f) + \pi_{01} \circ \mathsf{T}^2(f) + \varepsilon(\pi_{11} \circ \mathsf{T}^2(f)) \rangle && [\mathbf{E.2}] \\
&= \langle f \circ \pi_{00}, \boldsymbol{\partial}[f] \circ \pi_0 + \boldsymbol{\partial}[f] \circ (\pi_0 \times \pi_0) + \varepsilon(\boldsymbol{\partial}^2[f]) \rangle \\
&= \langle f \circ \pi_{00}, \boldsymbol{\partial}[f] \circ (\pi_0 \times \pi_0) + \left(\boldsymbol{\partial}[f] \circ \pi_0 + \varepsilon(\boldsymbol{\partial}^2[f]) \right) \rangle \\
&= \langle f \circ \pi_{00}, \boldsymbol{\partial}[f] \circ \langle \pi_{00}, \pi_{01} \rangle + \left(\boldsymbol{\partial}[f] \circ \langle \pi_{00}, \pi_{10} \rangle + \varepsilon \left(\boldsymbol{\partial}^2[f] \circ \langle \langle \pi_{00}, \pi_{10} \rangle, \langle \pi_{01}, \pi_{11} \rangle \rangle \right) \right) \rangle \\
&= \langle f \circ \pi_{00}, \boldsymbol{\partial}[f] \circ \langle \pi_{00}, \pi_{01} \rangle + \boldsymbol{\partial}[f] \circ \langle \pi_{00} + \varepsilon(\pi_{01}), \pi_{10} + \varepsilon(\pi_{11}) \rangle \rangle && [\mathbf{C}\partial.\mathbf{0}] \\
&= \langle f \circ \pi_{00}, \boldsymbol{\partial}[f] \circ \langle \pi_{00}, \pi_{01} + \pi_{10} + \varepsilon(\pi_{11}) \rangle \rangle && [\mathbf{C}\partial.\mathbf{2}] \\
&= \langle f \circ \pi_0, \boldsymbol{\partial}[f] \rangle \circ \langle \pi_{00}, \pi_{01} + \pi_{10} + \varepsilon(\pi_{11}) \rangle \\
&= \mathsf{T}(f) \circ \mu_A
\end{aligned}$$

So η and μ are natural transformations. Now we show that μ and η satisfy the monad identities (6.1). Starting with $\mu_A \circ \eta_A = 1_{\mathsf{T}(A)}$:

$$\begin{aligned}
\mu_A \circ \eta_{\mathsf{T}(A)} &= \langle \pi_{00}, \pi_{10} + \pi_{01} + \varepsilon(\pi_{11}) \rangle \circ \eta_{\mathsf{T}(A)} \\
&= \langle \pi_{00} \circ \eta_{\mathsf{T}(A)}, (\pi_{10} + \pi_{01} + \varepsilon(\pi_{11})) \circ \eta_{\mathsf{T}(A)} \rangle \\
&= \langle \pi_{00} \circ \eta_{\mathsf{T}(A)}, \pi_{10} \circ \eta_{\mathsf{T}(A)} + \pi_{01} \circ \eta_{\mathsf{T}(A)} + \varepsilon(\pi_{11}) \circ \eta_{\mathsf{T}(A)} \rangle \\
&= \langle \pi_{00} \circ \eta_{\mathsf{T}(A)}, \pi_{10} \circ \eta_{\mathsf{T}(A)} + \pi_{01} \circ \eta_{\mathsf{T}(A)} + \varepsilon(\pi_{11} \circ \eta_{\mathsf{T}(A)}) \rangle && [\mathbf{E.2}] \\
&= \langle \pi_{00} \circ \langle 1_{\mathsf{T}(A)}, 0 \rangle, \pi_{10} \circ \langle 1_{\mathsf{T}(A)}, 0 \rangle + \pi_{01} \circ \langle 1_{\mathsf{T}(A)}, 0 \rangle + \varepsilon \left(\pi_{11} \circ \langle 1_{\mathsf{T}(A)}, 0 \rangle \right) \rangle \\
&= \langle \pi_0, \pi_1 + 0 + \varepsilon(0) \rangle \\
&= \langle \pi_0, \pi_1 + 0 \rangle && [\mathbf{E.1}] \\
&= \langle \pi_0, \pi_1 \rangle \\
&= 1_{\mathsf{T}(A)}
\end{aligned}$$

Next we check that $\mu_A \circ \mathsf{T}(\eta_A) = 1_{\mathsf{T}(A)}$. First note that $\mathsf{T}(\eta_A) = \langle \langle \pi_0, 0 \rangle, \langle \pi_1, 0 \rangle \rangle$. Then we compute:

$$\begin{aligned}
\mu_A \circ \mathsf{T}(\eta_A) &= \langle \pi_{00}, \pi_{01} + \pi_{10} + \varepsilon(\pi_{11}) \rangle \circ \mathsf{T}(\eta_A) \\
&= \langle \pi_{00} \circ \mathsf{T}(\eta_A), (\pi_{01} + \pi_{10} + \varepsilon(\pi_{11})) \circ \mathsf{T}(\eta_A) \rangle \\
&= \langle \pi_{00} \circ \mathsf{T}(\eta_A), \pi_{01} \circ \mathsf{T}(\eta_A) + \pi_{10} \circ \mathsf{T}(\eta_A) + \varepsilon(\pi_{11}) \circ \mathsf{T}(\eta_A) \rangle \\
&= \langle \pi_{00} \circ \mathsf{T}(\eta_A), \pi_{01} \circ \mathsf{T}(\eta_A) + \pi_{10} \circ \mathsf{T}(\eta_A) + \varepsilon \left(\pi_{11} \circ \mathsf{T}(\eta_A) \right) \rangle && [\mathbf{E.2}] \\
&= \langle \pi_0, \pi_1 + 0 + \varepsilon(0) \rangle \\
&= \langle \pi_0, \pi_1 + 0 \rangle && [\mathbf{E.1}]
\end{aligned}$$

$$= \langle \pi_0, \pi_1 \rangle$$
$$= 1_{\mathsf{T}(A)}$$

For the last of the monad laws, we first note that, since μ is linear, it follows that $\mathsf{T}(\mu) = \mu \times \mu$. Then it suffices to compute:

$$\mu_A \circ \mathsf{T}(\mu_A) = \mu_A \circ (\mu_A \times \mu_A)$$
$$= \langle \pi_{00}, \pi_{10} + \pi_{01} + \varepsilon(\pi_{11}) \rangle \circ (\mu_A \times \mu_A)$$
$$= \langle \pi_{00} \circ (\mu_A \times \mu_A), (\pi_{10} + \pi_{01} + \varepsilon(\pi_{11})) \circ (\mu_A \times \mu_A) \rangle$$
$$= \langle \pi_{00} \circ (\mu_A \times \mu_A), \pi_{10} \circ (\mu_A \times \mu_A) + \pi_{01} \circ (\mu_A \times \mu_A) + \varepsilon(\pi_{11}) \circ (\mu_A \times \mu_A) \rangle$$
$$= \langle \pi_{00} \circ (\mu_A \times \mu_A), \pi_{10} \circ (\mu_A \times \mu_A) + \pi_{01} \circ (\mu_A \times \mu_A) + \varepsilon(\pi_{11} \circ (\mu_A \times \mu_A)) \rangle \qquad \textbf{[E.2]}$$
$$= \langle \pi_0 \circ \mu_A \circ \pi_0, \pi_1 \circ \mu_A \circ \pi_0 + \pi_0 \circ \mu_A \circ \pi_1 + \varepsilon(\pi_1 \circ \mu_A \circ \pi_1) \rangle$$
$$= \langle \pi_{000}, \pi_{100} + \pi_{010} + \varepsilon(\pi_{110}) + \pi_{001} + \varepsilon(\pi_{101} + \pi_{011} + \varepsilon(\pi_{111})) \rangle$$
$$= \langle \pi_0 \circ \pi_{00}, \pi_1 \circ \pi_{00} + \pi_0 \circ (\pi_{10} + \pi_{01} + \varepsilon(\pi_{11})) + \varepsilon(\pi_1 \circ (\pi_{10} + \pi_{01} + \varepsilon(\pi_{11}))) \rangle$$
$$= \langle \pi_{00}, \pi_{10} + \pi_{01} + \varepsilon(\pi_{11}) \rangle \circ \langle \pi_{00}, \pi_{10} + \pi_{01} + \varepsilon(\pi_{11}) \rangle$$
$$= \mu_A \circ \mu_{\mathsf{T}(A)}$$

So we conclude that (T, μ, η) is a monad. $\qquad \square$

Those familiar with monads may wonder if the Kleisli triple approach might have simplified the above proof. Recall that for a category $\mathbb{X}$, a Kleisli triple is a triple $(\mathsf{T}, \eta, (\_)^{\sharp})$ consisting of a function on object T, $A \mapsto \mathsf{T}(A)$, a family of maps indexed by the objects of $\mathbb{X}$, $\eta = \{\eta_A : A \to \mathsf{T}(A)| \ A \in \mathbb{X}\}$, and an operator $(\_)^{\sharp}$ which for any map $f : A \to \mathsf{T}(B)$ results in a map $f^{\sharp} : \mathsf{T}(A) \to \mathsf{T}(B)$, and such that the following equalities hold:

$$f^{\sharp} \circ \eta_A = f \qquad \eta_A^{\sharp} = 1_{\mathsf{T}(A)} \qquad (g^{\sharp} \circ f)^{\sharp} = g^{\sharp} \circ f^{\sharp}$$

There is a bijective correspondence between monads and Kleisli triples. So for the tangent bundle monad, we could have instead defined its associated Kleisli triple and prove 3 equations instead of the functoriality of T, the naturality of μ and η, and the 3 monad identities. However, we find the operator $(\_)^{\sharp}$ slightly more complicated to work with. As such, we elected to work out the monad identities directly, since while there are more identities to prove, we find the computations simpler and the proof easier to follow. That said, it is still interesting to work out the $(\_)^{\sharp}$ for the tangent bundle monad. In general, for a map $f : A \to \mathsf{T}(B)$, $f^{\sharp} : \mathsf{T}(A) \to \mathsf{T}(B)$ is equal to the composite $\mathsf{T}(f) = \mu_B \circ \mathsf{T}(f)$. In the case of the tangent bundle monad, note that a $f : A \to \mathsf{T}(B)$ would be a pair $f = \langle f_0, f_1 \rangle$, and so we obtain that (which we leave an excercise for the reader to compute for themselves):

$$f^{\sharp} = \langle f_0 \circ \pi_0, f_1 \circ \pi_0 + \partial[f_0] + \varepsilon(\partial[f_1]) \rangle$$

We next observe that the tangent bundle functor T preserves finite products up to isomorphism. Indeed, note that $\mathsf{T}(A \times B) \cong \mathsf{T}(A) \times \mathsf{T}(B)$ via the canonical natural isomorphism:

$$\phi_{A,B} : \mathsf{T}(A \times B) \to \mathsf{T}(A) \times \mathsf{T}(B) \qquad \phi_{A,B} := \langle T(\pi_0), T(\pi_1) \rangle \qquad (6.2)$$

By $[\textbf{C}\partial\textbf{.3}]$, it follows that $\phi_{A,B} = \langle \pi_0 \times \pi_0, \pi_1 \times \pi_1 \rangle$ and its inverse $\phi_{A,B}^{-1} : \mathsf{T}(A) \times \mathsf{T}(B) \to \mathsf{T}(A \times B)$ is defined in the same way, that is, $\phi_{A,B}^{-1} = \langle \pi_0 \times \pi_0, \pi_1 \times \pi_1 \rangle$. Expanding this out, we can see that $\phi_{A,B} : (A \times B) \times (A \times B) \to (A \times A) \times (B \times B)$ swaps the middle two

arguments, and so does $\phi_{A,B}^{-1}$. The following lemma will be extremely useful in many proofs throughout the remainder of this section.

Lemma 6.4. *In a Cartesian difference category:*

   (i) μ, η, *and* ϕ *are linear;*
   (ii) $\mathsf{T}(f + g) = \mathsf{T}(f) + \mathsf{T}(g)$ *and* $\mathsf{T}(0) = 0$;
  (iii) *If* f *is linear then* $\mathsf{T}(f)$ *is linear and* $\mathsf{T}(f) = f \times f$;
  (iv) $\mathsf{T}(\pi_i) = \pi_i \circ \phi$ *and* $\phi \circ \mathsf{T}(\langle f, g \rangle) = \langle \mathsf{T}(f), \mathsf{T}(g) \rangle$ *and* $\phi \circ \mathsf{T}(h \times k) = (\mathsf{T}(h) \times \mathsf{T}(k)) \circ \phi$;
   (v) $\mathsf{T}(\boldsymbol{\partial}[f]) = \boldsymbol{\partial}[\mathsf{T}(f)] \circ \phi$ *and* $\mathsf{T}(\varepsilon(f)) = \varepsilon(\mathsf{T}(f))$;
  (vi) $\varepsilon(\mu_A) = \mu_A \circ \varepsilon(1_{\mathsf{T}(A)})$, $\varepsilon(\eta_A) = \eta_A \circ \varepsilon(1_A)$, *and* $\varepsilon(\phi_{A,B}) = \phi_{A,B} \circ \varepsilon(1_{\mathsf{T}(A \times B)})$

Proof. The first three are mostly straightforward. Indeed, (i) follows immediately from Lemma 4.22 and by construction of μ, η, and ϕ, while (ii) follows from [**C$\boldsymbol{\partial}$.1**], and (iii) follows from the definition of T and linear maps. Next, we compute the identities of (iv), the first two follow from [**C$\boldsymbol{\partial}$.3**] and [**C$\boldsymbol{\partial}$.4**]. We first show $\mathsf{T}(\pi_i) = \pi_i \circ \phi$:

$$
\begin{aligned}
\mathsf{T}(\pi_i) =\ & \langle \pi_i \circ \pi_0, \boldsymbol{\partial}[\pi_i] \rangle \\
=\ & \langle \pi_i \circ \pi_0, \pi_i \circ \pi_1 \rangle && [\mathbf{C}\boldsymbol{\partial}.\mathbf{3}] \\
=\ & \pi_i \times \pi_i \\
=\ & \pi_i \circ \phi
\end{aligned}
$$

Next we compute that $\phi \circ \mathsf{T}(\langle f, g \rangle) = \langle \mathsf{T}(f), \mathsf{T}(g) \rangle$:

$$
\begin{aligned}
\phi \circ \mathsf{T}(\langle f, g \rangle) =\ & \phi \circ \langle \langle f, g \rangle \circ \pi_0, \boldsymbol{\partial}[\langle f, g \rangle] \rangle \\
=\ & \phi \circ \langle \langle f \circ \pi_0, g \circ \pi_0 \rangle, \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \rangle && [\mathbf{C}\boldsymbol{\partial}.\mathbf{4}] \\
=\ & \langle \pi_0 \times \pi_0, \pi_1 \times \pi_1 \rangle \circ \langle \langle f \circ \pi_0, g \circ \pi_0 \rangle, \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \rangle \\
=\ & \langle (\pi_0 \times \pi_0) \circ \langle \langle f \circ \pi_0, g \circ \pi_0 \rangle, \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \rangle, (\pi_1 \times \pi_1) \circ \langle \langle f \circ \pi_0, g \circ \pi_0 \rangle, \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \rangle \rangle \\
=\ & \langle \langle \pi_0 \circ \langle f \circ \pi_0, g \circ \pi_0 \rangle, \pi_0 \circ \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \rangle, \langle \pi_1 \circ \langle f \circ \pi_0, g \circ \pi_0 \rangle, \pi_1 \circ \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g] \rangle \rangle \rangle \\
=\ & \langle \langle f \circ \pi_0, \boldsymbol{\partial}[f] \rangle, \langle g \circ \pi_0, \boldsymbol{\partial}[g] \rangle \rangle \\
=\ & \langle \mathsf{T}(f), \mathsf{T}(g) \rangle
\end{aligned}
$$

For the remaining identity, $\phi \circ \mathsf{T}(h \times k) = (\mathsf{T}(h) \times \mathsf{T}(k)) \circ \phi$, we use the previous one:

$$
\begin{aligned}
\phi \circ \mathsf{T}(h \times k) =\ & \phi \circ \mathsf{T}(\langle h \circ \pi_0, k \circ \pi_1 \rangle) \\
=\ & \langle \mathsf{T}(h \circ \pi_0), \mathsf{T}(k \circ \pi_1) \rangle && \text{(Lem.6.4.iv)} \\
=\ & \langle \mathsf{T}(h) \circ \mathsf{T}(\pi_0), \mathsf{T}(k) \circ \mathsf{T}(\pi_1) \rangle && (\mathsf{T} \text{ is a functor}) \\
=\ & \langle \mathsf{T}(h) \circ (\pi_0 \times \pi_0), \mathsf{T}(k) \circ (\pi_1 \times \pi_1) \rangle && \text{(Lem.6.4.iii)} \\
=\ & \langle \mathsf{T}(h) \circ \pi_0 \circ \phi, \mathsf{T}(k) \circ \pi_1 \circ \phi \rangle \\
=\ & \langle \mathsf{T}(h) \circ \pi_0, \mathsf{T}(k) \circ \pi_1 \rangle \circ \phi \\
=\ & (\mathsf{T}(h) \times \mathsf{T}(k)) \circ \phi
\end{aligned}
$$

Next we compute the two identities of (v). First we show $\mathsf{T}(\boldsymbol{\partial}[f]) = \boldsymbol{\partial}[\mathsf{T}(f)] \circ \phi$:

$$
\begin{aligned}
\boldsymbol{\partial}[\mathsf{T}(f)] \circ \phi =\ & \boldsymbol{\partial}[\langle f \circ \pi_0, \boldsymbol{\partial}[f] \rangle] \circ \phi \\
=\ & \langle \boldsymbol{\partial}[f \circ \pi_0], \boldsymbol{\partial}^2[f] \rangle \circ \phi && [\mathbf{C}\boldsymbol{\partial}.\mathbf{4}] \\
=\ & \langle \boldsymbol{\partial}[f] \circ (\pi_0 \times \pi_0), \boldsymbol{\partial}^2[f] \rangle \circ \phi && \text{(Lem.4.22)} \\
=\ & \langle \boldsymbol{\partial}[f] \circ (\pi_0 \times \pi_0) \circ \phi, \boldsymbol{\partial}^2[f] \circ \phi \rangle
\end{aligned}
$$

$$
\begin{aligned}
&= \left\langle \partial[f] \circ (\pi_0 \times \pi_0) \circ \langle \pi_0 \times \pi_0, \pi_1 \times \pi_1 \rangle, \partial^2[f] \circ \langle \pi_0 \times \pi_0, \pi_1 \times \pi_1 \rangle \right\rangle \\
&= \left\langle \partial[f] \circ \langle \pi_0(\pi_0 \times \pi_0), \pi_0 \circ (\pi_1 \times \pi_1) \rangle, \partial^2[f] \circ \langle \langle \pi_0 \circ \pi_0, \pi_0 \circ \pi_1 \rangle, \langle \pi_1 \circ \pi_0, \pi_1 \circ \pi_1 \rangle \rangle \right\rangle \\
&= \left\langle \partial[f] \circ \langle \pi_1 \circ \pi_0, \pi_1 \circ \pi_0 \rangle, \partial^2[f] \circ \langle \langle \pi_0 \circ \pi_0, \pi_0 \circ \pi_1 \rangle, \langle \pi_1 \circ \pi_0, \pi_1 \circ \pi_1 \rangle \rangle \right\rangle \\
&= \left\langle \partial[f] \circ \langle \pi_0 \circ \pi_0, \pi_1 \circ \pi_0 \rangle, \partial^2[f] \circ \langle \langle \pi_0 \circ \pi_0, \pi_1 \circ \pi_0 \rangle, \langle \pi_0 \circ \pi_1, \pi_1 \circ \pi_1 \rangle \rangle \right\rangle \qquad \textbf{[C}\partial\textbf{.7.a]} \\
&= \left\langle \partial[f] \circ \pi_0, \partial^2[f] \right\rangle \\
&= \mathsf{T}(\partial[f])
\end{aligned}
$$

Next we show $\mathsf{T}(\varepsilon(f)) = \varepsilon(\mathsf{T}(f))$:

$$
\begin{aligned}
\mathsf{T}(\varepsilon(f)) &= \left\langle \varepsilon(f) \circ \pi_0, \partial[\varepsilon(f)] \right\rangle \\
&= \left\langle \varepsilon(f \circ \pi_0), \varepsilon(\partial[f]) \right\rangle && \textbf{[E.2]} + \textbf{[C}\partial\textbf{.1]} \\
&= \varepsilon(\langle f \circ \pi_0, \partial[f] \rangle) && \text{(Lem.4.3)} \\
&= \varepsilon(\mathsf{T}(f))
\end{aligned}
$$

Lastly (vi) follows from the fact we have shown that μ, η, and ϕ are linear and so by Lemma 4.22.i, the desired equalities hold. $\qquad\square$

6.2. The Eilenberg-Moore Category of T.

Recall that a T-algebra of the monad (T, μ, η) is a pair (A, ν) consisting of an object A and a map $\nu : \mathsf{T}(A) \to A$ such that $\nu \circ \eta_A = 1_A$ and $\nu \circ \mathsf{T}(\nu) = \nu \circ \mu_A$, and that a T-algebra morphism $f : (A, \nu) \to (B, \omega)$ is a map $f : A \to B$ such that $\omega \circ \mathsf{T}(f) = f \circ \nu$. The Eilenberg-Moore category of (T, μ, η) is the category of T-algebras, that is, the category $\mathbb{X}^{\mathsf{T}}$ whose objects are T-algebras and whose maps are T-algebra morphisms. It is well known that for any monad on a category with finite products, the Eilenberg-Moore category also has finite products. Indeed, in this case, for T-algebras (A, ν) and (B, ω), their product is defined as:

$$
(A, \nu) \times (B, \omega) := (A \times B, (\nu \times \omega) \circ \phi_{A,B})
$$

while the projection maps and the pairing of maps are the same as in the base category. The terminal object in $\mathbb{X}^{\mathsf{T}}$ is defined as $(\top, 0)$.

Therefore, it may be tempting to think that the Eilenberg-Moore category of the tangent bundle monad is also a Cartesian difference category. Unfortunately, there are two problems with this: the sum of a T-algebra morphism is not necessarily a T-algebra morphism and the differential of a T-algebra may not necessarily be a T-algebra. Indeed, let $f, g : (A, \nu) \to (B, \omega)$ be T-algebra morphism. On the one hand, since $\mathsf{T}(f + g) = \mathsf{T}(f) + \mathsf{T}(g)$ we have that:

$$
\omega \circ \mathsf{T}(f + g) = \omega \circ (\mathsf{T}(f) + \mathsf{T}(g)) = \omega \circ \mathsf{T}(f) + \omega \circ \mathsf{T}(g) = f \circ \nu + g \circ \nu
$$

However, if ν is not additive then $f \circ \nu + g \circ \nu$ may not be equal to $(f + g) \circ \nu$. As such, one solution could be to consider the subcategory of additive T-algebras. While this full subcategory of additive T-algebras will be a Cartesian left additive category, this does not solve the problem of the differential of a T-algebra morphism. For a T-algebra morphism $f : (A, \nu) \to (B, \omega)$, its derivative $\partial[f] : A \times A \to B$ should be a T-algebra morphism of type $(A, \nu) \times (A, \nu) \to (B, \omega)$. However one cannot get very far in trying to show that $\partial[f]$ is a T-algebra morphism. The solution to this problem is instead to consider the full subcategory of *linear* T-algebras.

Define $\mathbb{X}_{lin}^{\mathsf{T}}$ as the category of *linear* T-algebra, that is, the category whose objects are T-algebras (A, ν) such that ν is linear and whose maps are arbitrary T-algebra morphisms between them. In particular, note that a map in $\mathbb{X}_{lin}^{\mathsf{T}}$ need not be linear. Like the Eilenberg-Moore category, there is an obvious forgetful functor $\mathsf{U} : \mathbb{X}_{lin}^{\mathsf{T}} \to \mathbb{X}$ defined as $\mathsf{U}(A, \nu) = A$ and $\mathsf{U}(f) = f$. A linear T-algebra structure should be interpreted as a map which linearly modifies a point by a tangent vector. Unfortunately, as mentioned in [CC14], the Eilenberg-Moore category of the tangent bundle monad has yet to be studied in full, even in classical differential geometry.

Example 6.5. For the tangent bundle monad on the Cartesian differential category SMOOTH (Example 6.1), a linear T-algebra is a pair $(\mathbb{R}^n, \nu)$ such that $\nu(\vec{x}, \vec{y}) = \vec{x} + t\vec{y}$ for some fixed $t \in \mathbb{R}$. Therefore, $\mathsf{SMOOTH}_{lin}^{\mathsf{T}}$ is equivalent to the category whose objects are pairs $(\mathbb{R}^n, t)$, with $t \in \mathbb{R}$, where a map $F : (\mathbb{R}^n, t) \to (\mathbb{R}^m, s)$ is a smooth function $F : \mathbb{R}^n \to \mathbb{R}^m$ such that $F(\vec{x} + t\vec{y}) = F(\vec{x}) + s\mathsf{D}[F](\vec{x}, \vec{y})$.

Example 6.6. For the tangent bundle monad on the Cartesian differential category $\overline{\mathsf{Ab}}$ (Example 6.2), a linear T-algebra is a pair (G, ν) such that $\nu(x, y) = x + e(y)$ for some group endomorphism $e : G \to G$. Therefore, $\overline{\mathsf{Ab}}$ is equivalent to the category whose objects are pairs (G, e), with G an abelian group and $e : G \to G$ a group endomorphism, where a map $f : (G, e) \to (H, e')$ is a function $f : G \to H$ such that

$$f(x + e(y)) = f(x) + e'\left(f(x + y)\right) - e'(f(x)).$$

We will now explain how $\mathbb{X}_{lin}^{\mathsf{T}}$ is a Cartesian difference category. Simply put, the Cartesian difference structure of $\mathbb{X}_{lin}^{\mathsf{T}}$ is the same as $\mathbb{X}$, and therefore we say that the forgetful functor preserves the Cartesian difference structure strictly. Starting with the Cartesian left additive structure: the finite product structure of $\mathbb{X}_{lin}^{\mathsf{T}}$ is defined in the same way as $\mathbb{X}^{\mathsf{T}}$, while the sum of T-algebra morphisms and zero T-algebra morphisms are defined as the sum and zero maps in $\mathbb{X}$. Of course, we must check that this is well-defined.

Lemma 6.7. *$\mathbb{X}_{lin}^{\mathsf{T}}$ is a Cartesian left additive category.*

Proof. We must first check that if (A, ν) and (B, ω) are linear T-algebras, then so is their product $(A, \nu) \times (B, \omega)$, that is, we must show that $(\nu \times \omega) \circ \phi_{A,B}$ is linear. However, since ν, ω, and $\phi_{A,B}$ are linear, it follows from Lemma 4.22.iv that $(\nu \times \omega) \circ \phi_{A,B}$ is also linear. Therefore, the product of linear T-algebras is also a linear T-algebra. Next, since zero maps are linear, $(\top, 0)$ is also a linear T-algebra. Therefore $\mathbb{X}_{lin}^{\mathsf{T}}$ has finite products since it is a full subcategory of $\mathbb{X}^{\mathsf{T}}$. Next, we show that if $f : (A, \nu) \to (B, \omega)$ and $g : (A, \nu) \to (B, \omega)$ are T-algebra morphisms between linear T-algebras, then their sum $f + g : (A, \nu) \to (B, \omega)$ is also a T-algebra morphism. This follows from the fact that linear maps are additive (Lemma 4.22.ii):

$$
\begin{aligned}
\omega \circ \mathsf{T}(f + g) &= \omega \circ (\mathsf{T}(f) + \mathsf{T}(g)) && \text{(by Prop.6.4.ii)} \\
&= \omega \circ \mathsf{T}(f) + \omega \circ \mathsf{T}(g) && \\
&= f \circ \nu + g \circ \nu && \text{(by T-alg. morph. def.)} \\
&= (f + g) \circ \nu && \text{(ν is additive)}
\end{aligned}
$$

So we have that $f + g$ is a T-algebra morphism. Similarly, we must show that zero maps between linear T-algebras (A, ν) and (B, ω) are T-algebra morphisms:

$$\omega \circ \mathsf{T}(0) = \omega \circ 0 \qquad\qquad \text{(by Prop.6.4.ii)}$$

$$= 0$$
$$= 0 \circ \nu \qquad\qquad (\nu \text{ is additive})$$

Therefore $0 : (A, \nu) \to (B, \omega)$ is a T-algebra morphism. And clearly, the remaining Cartesian left additive structure axioms hold since composition in $\mathbb{X}_{lin}^{\mathsf{T}}$ is the same as in $\mathbb{X}$. So we conclude that $\mathbb{X}_{lin}^{\mathsf{T}}$ is a Cartesian left additive category. $\qquad\square$

Similarly, the infinitesimal extension and the difference combinator of $\mathbb{X}_{lin}^{\mathsf{T}}$ are defined in the same way as in $\mathbb{X}$. Once again, while the infinitesimal extension and difference combinator axioms will automatically hold, we will have to check that this is all well-defined.

Proposition 6.8. *For a Cartesian difference category $\mathbb{X}$, the category of linear T-algebras $\mathbb{X}_{lin}^{\mathsf{T}}$ is a Cartesian difference category such that the obvious forgetful functor $\mathsf{U} : \mathbb{X}_{lin}^{\mathsf{T}} \to \mathbb{X}$ preserves the Cartesian difference structure strictly.*

Proof. We first show that if $f : (A, \nu) \to (B, \omega)$ is a T-algebra morphism between linear T-algebras, then so is $\varepsilon(f)$:

$$
\begin{aligned}
\omega \circ \mathsf{T}(\varepsilon(f)) &= \omega \circ \varepsilon(\mathsf{T}(f)) && \text{(by Prop.6.4.v)} \\
&= \varepsilon(\omega \circ \mathsf{T}(f)) && \text{(by Lem.4.22.i)} \\
&= \varepsilon(f \circ \nu) && \text{(by T-alg. morph. def.)} \\
&= \varepsilon(f) \circ \nu && \text{(by ε def.)}
\end{aligned}
$$

So we have that $\varepsilon(f) : (A, \nu) \to (B, \omega)$ is a T-algebra morphism. Next, we check that the derivative $\boldsymbol{\partial}[f] : A \times A \to B$ is a T-algebra morphism of type $(A, \nu) \times (A, \nu) \to (B, \omega)$.

$$
\begin{aligned}
\omega \circ \mathsf{T}(\boldsymbol{\partial}[f]) &= \omega \circ \boldsymbol{\partial}[\mathsf{T}(f)] \circ \phi_{A,A} && \text{(by Prop.6.4.v)} \\
&= \boldsymbol{\partial}[\omega \circ \mathsf{T}(f)] \circ \phi_{A,A} && \text{(by Lem.4.22.v since ω is linear)} \\
&= \boldsymbol{\partial}[f \circ \nu] \circ \phi_{A,A} && \text{(by T-alg. morph. def.)} \\
&= \boldsymbol{\partial}[f] \circ (\nu \times \nu) \circ \phi_{A,A} && \text{(by Lem.4.22.v since ν is linear)}
\end{aligned}
$$

So we have that $\boldsymbol{\partial}[f] : (A, \nu) \times (A, \nu) \to (B, \omega)$ is a T-algebra morphism. Since composition and the Cartesian left additive structure of $\mathbb{X}_{lin}^{\mathsf{T}}$ is the same as $\mathbb{X}$, it automatically follows that ε is an infinitesimal extension and $\boldsymbol{\partial}$ is a difference combinator on $\mathbb{X}_{lin}^{\mathsf{T}}$. Therefore, we conclude that $\mathbb{X}_{lin}^{\mathsf{T}}$ is a Cartesian difference category, and clearly the forgetful functor preserves the Cartesian difference structure strictly. $\qquad\square$

6.3. **The Kleisli Category of T.** The construction found here is different from the one found in the conference paper [APL20]. Indeed, the proposed infinitesimal extension and difference combinator in [APL20] were based on the ones that appeared in [APO19]. Unfortunately, we have found that said proposed infinitesimal extension and difference combinator failed to satisfy **[C$\boldsymbol{\partial}$.2]** and therefore both of the aforementioned results in [APL20, APO19] are incorrect. We rectify this mistake here by changing the infinitesimal extension to the correct one, while keeping the difference combinator the same.

Recall that the Kleisli category of the monad (T, μ, η) is defined as the category $\mathbb{X}_{\mathsf{T}}$ whose objects are the objects of $\mathbb{X}$, and where a map $A \to B$ in $\mathbb{X}_{\mathsf{T}}$ is a map $f : A \to \mathsf{T}(B)$ in $\mathbb{X}$, which would be a pair $f = \langle f_0, f_1 \rangle$ where $f_j : A \to B$. The identity map in $\mathbb{X}_{\mathsf{T}}$ is the monad unit $\eta_A : A \to \mathsf{T}(A)$, while composition of Kleisli maps $f : A \to \mathsf{T}(B)$ and

$g : B \to \mathsf{T}(C)$ is defined as the composite $\mu_C \circ \mathsf{T}(g) \circ f$. To distinguish between composition in $\mathbb{X}$ and $\mathbb{X}_{\mathsf{T}}$, we denote the Kleisli composition as follows:

$$g \circ^{\mathsf{T}} f = \mu_C \circ \mathsf{T}(g) \circ f \tag{6.3}$$

If $f = \langle f_0, f_1 \rangle$ and $g = \langle g_0, g_1 \rangle$, then their Kleisli composition can be worked out to be:

$$g \circ^{\mathsf{T}} f = \langle g_0, g_1 \rangle \circ^{\mathsf{T}} \langle f_0, f_1 \rangle = \langle g_0 \circ f_0, \boldsymbol{\partial}[g_0] \circ \langle f_0, f_1 \rangle + g_1 \circ (f_0 \oplus f_1) \rangle \tag{6.4}$$

Kleisli maps can be understood as "generalized" vector fields. Indeed, $\mathsf{T}(A)$ should be thought of as the tangent bundle over A, and therefore a vector field would be a map $\langle 1, f \rangle : A \to \mathsf{T}(A)$, which is of course also a Kleisli map. For more details on the intuition behind this Kleisli category see [CC14]. Furthermore, in general, the Kleisli category of a monad is equivalently to a full subcategory of the Eilenberg-Moore consisting of the free algebras. In this case, note that every free T-algebra $(\mathsf{T}(A), \mu_A)$ is also an object in $\mathbb{X}_{lin}^{\mathsf{T}}$, since μ_A is linear. Therefore, $\mathbb{X}_{\mathsf{T}}$ is equivalent to the full subcategory of free T-algebras of $\mathbb{X}_{lin}^{\mathsf{T}}$.

Example 6.9. For the tangent bundle monad on the Cartesian differential category SMOOTH (Example 6.1), a Kleisli map is a smooth function $F : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^m$, which is interpreted as a pair of smooth functions $F = \langle F_0, F_1 \rangle$ where $F_i : \mathbb{R}^n \to \mathbb{R}^m$. The composition of Kleisli maps $F = \langle F_0, F_1 \rangle$ and $G = \langle G_0, G_1 \rangle$ is computed out to be $(G \circ^{\mathsf{T}} F)(\vec{x}) = (G_0(F_0(\vec{x})), \mathsf{D}[G_0](F_0(\vec{x}), F_1(\vec{x})) + G_1(F_0(\vec{x}))$. Among these Kleisli maps are the vector fields on $\mathbb{R}^n$, which are precisely the Kleisli maps of the form $F = \langle 1_{\mathbb{R}^n}, f \rangle$, for some smooth function $f : \mathbb{R}^n \to \mathbb{R}^n$.

Example 6.10. For the tangent bundle monad on the Cartesian differential category $\overline{\mathsf{Ab}}$ (Example 6.2), a Kleisli map is a function $f : G \to H \times H$, which is interpreted as a pair of functions $f = \langle f_0, f_1 \rangle$ where $f_i : G \to H$. In this case, the composition of Kleisli maps $f = \langle f_0, f_1 \rangle$ and $g = \langle g_0, g_1 \rangle$ is $(g \circ^{\mathsf{T}} g)(x) = (g_0(g_0(x)), g_0(f_0(x) + f_1(x)) - g_0(f_0(x)) + g_1(f_0(x))$.

We now wish to explain how the Kleisli category $\mathbb{X}_{\mathsf{T}}$ is again a Cartesian difference category. We begin by exhibiting the Cartesian left additive structure of the Kleisli category $\mathbb{X}_{\mathsf{T}}$. Generally, the Kleisli category does not automatically inherit the product structure of the base category, even if the Eilenberg-Moore category does. However, since T preserves finite products up to isomorphism, it follows that its Kleisli category has finite products. As such, the product of objects in $\mathbb{X}_{\mathsf{T}}$ is defined as $A \times B$ with projections $\pi_0^{\mathsf{T}} : A \times B \to \mathsf{T}(A)$ and $\pi_1^{\mathsf{T}} : A \times B \to \mathsf{T}(B)$ defined respectively as $\pi_0^{\mathsf{T}} = \langle \pi_0, 0 \rangle$ and $\pi_1^{\mathsf{T}} = \langle \pi_1, 0 \rangle$, and the pairing of Kleisli maps $f = \langle f_0, f_1 \rangle$ and $g = \langle g_0, g_1 \rangle$ is defined as:

$$\langle f, g \rangle^{\mathsf{T}} = \phi^{-1} \circ \langle f, g \rangle = \langle \langle f_0, g_0 \rangle, \langle f_1, g_1 \rangle \rangle \tag{6.5}$$

where recall $\phi^{-1} : \mathsf{T}(A) \times \mathsf{T}(B) \to \mathsf{T}(A \times B)$ is the inverse of ϕ as defined in (6.2). The terminal object is again $\top$ and where the unique map to the terminal object is $!_A^{\mathsf{T}} = 0$. The sum of Kleisli maps $f = \langle f_0, f_1 \rangle$ and $g = \langle g_0, g_1 \rangle$ is defined as:

$$f +^{\mathsf{T}} g = f + g = \langle f_0 + g_0, f_1 + g_1 \rangle$$

and the zero Kleisli maps is simply $0^{\mathsf{T}} = 0 = \langle 0, 0 \rangle$. Therefore we conclude that the Kleisli category of the tangent monad is a Cartesian left additive category.

Lemma 6.11. $\mathbb{X}_{\mathsf{T}}$ *is a Cartesian left additive category.*

Proof. As explained above, by Lemma 6.4, T preserves the finite product structure, and thus it follows that the Kleisli category $\mathbb{X}_\mathsf{T}$ is a Cartesian category. So it remains to show that $\mathbb{X}_\mathsf{T}$ is a left additive category and that the projection maps are additive. We start by showing that the proposed additive structure is compatible with Kleisli composition. This follows from the fact that T preserves the additive structure by Lemma 6.4.(ii) and Lemma 4.22.ii, that μ is linear and therefore also additive.

$$
\begin{aligned}
(f +^\mathsf{T} g) \circ^\mathsf{T} x =\ & (f + g) \circ^\mathsf{T} x \\
=\ & \mu \circ \mathsf{T}(f + g) \circ x \\
=\ & \mu \circ (\mathsf{T}(f) + \mathsf{T}(g)) \circ x && \text{(Lem6.4.ii)} \\
=\ & \mu \circ (\mathsf{T}(f) \circ x + \mathsf{T}(g) \circ x) \\
=\ & \mu \circ \mathsf{T}(f) \circ x + \mu \circ \mathsf{T}(g) \circ x && (\mu \text{ is additive)} \\
=\ & f \circ^\mathsf{T} x + g \circ^\mathsf{T} x \\
=\ & f \circ^\mathsf{T} x +^\mathsf{T} g \circ^\mathsf{T} x
\end{aligned}
$$

$$
\begin{aligned}
0^\mathsf{T} \circ x =\ & 0 \circ^\mathsf{T} x \\
=\ & \mu \circ \mathsf{T}(0) \circ x \\
=\ & \mu \circ 0 \circ x && \text{(Lem6.4.ii)} \\
=\ & 0 \circ x && (\mu \text{ is additive)} \\
=\ & 0
\end{aligned}
$$

So we have that $\mathbb{X}_\mathsf{T}$ is a left additive category. Next we show that the projection maps are additive. Note that $\pi_i^\mathsf{T} = \eta \circ \pi_i$ and that by Lemma6.4.(iii), $\mathsf{T}(\pi_i) = \pi_i \times \pi_i$ is linear and therefore also additive. So we have that:

$$
\begin{aligned}
\pi_i^\mathsf{T} \circ^\mathsf{T} (x + y) =\ & (\eta \circ \pi_i) \circ^\mathsf{T} (x + y) \\
=\ & \mu \circ \mathsf{T}(\eta \circ \pi_i) \circ (x + y) \\
=\ & \mu \circ \mathsf{T}(\eta) \circ \mathsf{T}(\pi_i) \circ (x + y) && (\mathsf{T} \text{ is a functor)} \\
=\ & \mathsf{T}(\pi_i) \circ (x + y) && \text{(Monad identities)} \\
=\ & \mathsf{T}(\pi_i) \circ x + \mathsf{T}(\pi_i) \circ y && (\mathsf{T}(\pi_i) \text{ is additive)} \\
=\ & \mu \circ \mathsf{T}(\eta) \circ \mathsf{T}(\pi_i) \circ x + \mu \circ \mathsf{T}(\eta) \circ \mathsf{T}(\pi_i) \circ y && \text{(Monad identities)} \\
=\ & \mu \circ \mathsf{T}(\eta \circ \pi_i) \circ x + \mu \circ \mathsf{T}(\eta \circ \pi_i) \circ y && (\mathsf{T} \text{ is a functor)} \\
=\ & (\eta \circ \pi_i) \circ^\mathsf{T} x + (\eta \circ \pi_i) \circ^\mathsf{T} y \\
=\ & \pi_i^\mathsf{T} \circ^\mathsf{T} x + \pi_i^\mathsf{T} \circ^\mathsf{T} y
\end{aligned}
$$

$$
\begin{aligned}
\pi_i^\mathsf{T} \circ^\mathsf{T} 0^\mathsf{T} =\ & (\eta \circ \pi_i) \circ^\mathsf{T} 0 \\
=\ & \mu \circ \mathsf{T}(\eta \circ \pi_i) \circ 0 \\
=\ & \mu \circ \mathsf{T}(\eta) \circ \mathsf{T}(\pi_i) \circ 0 && (\mathsf{T} \text{ is a functor)} \\
=\ & \mathsf{T}(\pi_i) \circ 0 && \text{(Monad identities)} \\
=\ & 0 && (\mathsf{T}(\pi_i) \text{ is additive)}
\end{aligned}
$$

So we conclude that $\mathbb{X}_{\mathsf{T}}$ is a Cartesian left additive category. $\square$

The infinitesimal extension ε^{T} for the Kleisli category is the same as the infinitesimal extension of the base category, that is, for a Kleisli map $f = \langle f_0, f_1 \rangle$:

$$\varepsilon^{\mathsf{T}}(f) = \varepsilon(f) = \varepsilon(\langle f_0, f_1 \rangle) = \langle \varepsilon(f_0), \varepsilon(f_1) \rangle$$

where the last equality is due to Lemma 4.3. It is important to note that this is the major correction made from the conference paper version. Indeed, the infinitesimal extension suggested in [APL20] was based on the change action structure suggested in [APO19]. However, it unfortunately turns out that a result in [APO19] was incorrect, since [**CAD.2**] actually fails, and therefore the infinitesimal extension and difference combinator suggested in [APL20] do not satisfy [**C∂.2**]. Luckily, the infinitesimal extension provide in this paper does work, as we carefully prove below.

Lemma 6.12. ε^{T} *is an infinitesimal extension on* $\mathbb{X}_{\mathsf{T}}$.

Proof. We check that ε^{T} satisfies [**E.1**], [**E.2**], and [**E.3**]. For [**E.1**], we compute:

$$
\begin{aligned}
\varepsilon^{\mathsf{T}}(f +^{\mathsf{T}} g) &= \varepsilon^{\mathsf{T}}(f + g) \\
&= \varepsilon(f + g) \\
&= \varepsilon(f) + \varepsilon(g) \qquad\qquad\qquad\qquad\text{[\textbf{E.1}]} \\
&= \varepsilon(f) +^{\mathsf{T}} \varepsilon(g) \\
&= \varepsilon^{\mathsf{T}}(f) +^{\mathsf{T}} \varepsilon^{\mathsf{T}}(g)
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon^{\mathsf{T}}(0^{\mathsf{T}}) &= \varepsilon(0^{\mathsf{T}}) \\
&= \varepsilon(0) \\
&= 0 \qquad\qquad\qquad\qquad\qquad\qquad\text{[\textbf{E.1}]} \\
&= 0^{\mathsf{T}}
\end{aligned}
$$

Next for [**E.2**], we compute:

$$
\begin{aligned}
\varepsilon^{\mathsf{T}}(g \circ^{\mathsf{T}} f) &= \varepsilon(g \circ^{\mathsf{T}} f) \\
&= \varepsilon(\mu \circ \mathsf{T}(g) \circ f) \\
&= \varepsilon(\mu) \circ \mathsf{T}(g) \circ f \qquad\qquad\qquad\text{[\textbf{E.2}]} \\
&= \mu \circ \varepsilon(1) \circ \mathsf{T}(g) \circ f \qquad\qquad\text{(Lem.6.4.(vi))} \\
&= \mu \circ \varepsilon(\mathsf{T}(g)) \circ f \\
&= \mu \circ \mathsf{T}(\varepsilon(g)) \circ f \qquad\qquad\qquad\text{(Lem.6.4.(v))} \\
&= \varepsilon(g) \circ^{\mathsf{T}} f \\
&= \varepsilon^{\mathsf{T}}(g) \circ^{\mathsf{T}} f
\end{aligned}
$$

Lastly for [**E.3**], recall again that $\pi_i^{\mathsf{T}} = \eta \circ \pi_i$, so we can compute:

$$
\begin{aligned}
\varepsilon^{\mathsf{T}}(\pi_i^{\mathsf{T}}) &= \varepsilon(\pi_i^{\mathsf{T}}) \\
&= \varepsilon(\eta \circ \pi) \\
&= \varepsilon(\mathsf{T}(\pi_i) \circ \eta) \qquad\qquad\qquad\qquad\text{(Nat. of }\eta\text{)}
\end{aligned}
$$

$$\begin{aligned}
&= \; \varepsilon(\mathsf{T}(\pi_i)) \circ \eta && \textbf{[E.2]} \\
&= \; \mathsf{T}(\varepsilon(\pi_i)) \circ \eta && (\text{Lem.6.4.(v)}) \\
&= \; \mathsf{T}(\pi_i \circ \varepsilon(1)) \circ \eta && \textbf{[E.3]} \\
&= \; \mathsf{T}(\pi_i) \circ \mathsf{T}(\varepsilon(1)) \circ \eta && (\mathsf{T} \text{ is a functor}) \\
&= \; \mathsf{T}(\pi_i) \circ \eta \circ \varepsilon(1) && (\text{Nat. of } \eta) \\
&= \; \mathsf{T}(\pi_i) \circ \varepsilon(\eta) && (\text{Lem.6.4.(vi)}) \\
&= \; \mu \circ \mathsf{T}(\eta) \circ \mathsf{T}(\pi_i) \circ \varepsilon(\eta) && (\text{Monad identities}) \\
&= \; \mu \circ \mathsf{T}(\eta \circ \pi_i) \circ \varepsilon(\eta) && (\mathsf{T} \text{ is a functor}) \\
&= \; (\eta \circ \pi_i) \circ^{\mathsf{T}} \varepsilon(\eta) && \\
&= \; (\eta \circ \pi_i) \circ^{\mathsf{T}} \varepsilon^{\mathsf{T}}(\eta) && \\
&= \; \pi_i^{\mathsf{T}} \circ^{\mathsf{T}} \varepsilon^{\mathsf{T}}(\eta) &&
\end{aligned}$$

So we conclude that ε^{T} is an infinitesimal extension on $\mathbb{X}_{\mathsf{T}}$. $\qquad\square$

To define the difference combinator for the Kleisli category, first note that difference combinators by definition do not change the codomain. That is, if $f : A \to \mathsf{T}(B)$ is a Kleisli map, then the type of its derivative should be $A \times A \to \mathsf{T}(B)$, which coincides with the type of its derivative in $\mathbb{X}$. Therefore, the difference combinator $\boldsymbol{\partial}^{\mathsf{T}}$ for the Kleisli category is defined to be the difference combinator of the base category, that is, for a Kleisli map $f = \langle f_0, f_1 \rangle$, its derivative is defined as:

$$\boldsymbol{\partial}^{\mathsf{T}}[f] = \boldsymbol{\partial}[f] = \langle \boldsymbol{\partial}[f_0], \boldsymbol{\partial}[f_1] \rangle \tag{6.6}$$

where the last equality is due to $[\textbf{C}\boldsymbol{\partial}\textbf{.4}]$. We note that this difference combinator was the one suggested in [APL20] and is the derivative underlying the change action model structure suggested in [APO19]. However, as mentioned above, the difference combinator with the infinitesimal extension or change action suggested in those papers does not satisfy $[\textbf{C}\boldsymbol{\partial}\textbf{.2}]$ or $[\textbf{CAD.2}]$. Luckily, we are able to correct this and still obtain a positive result by carefully proving that the proposed infinitesimal extension and difference combinator in this paper does provide a Cartesian difference structure on the Kleisli category.

Proposition 6.13. *For a Cartesian difference category $\mathbb{X}$, the Kleisli category $\mathbb{X}_{\mathsf{T}}$ is a Cartesian difference category with infinitesimal extension ε^{T} and difference combinator $\boldsymbol{\partial}^{\mathsf{T}}$.*

Proof. We first note that we can easily compute the following:

$$\begin{aligned}
\boldsymbol{\partial}^{T}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}} &= \; \mu \circ \mathsf{T}(\boldsymbol{\partial}^{T}[f]) \circ \langle x, y \rangle^{\mathsf{T}} \\
&= \; \mu \circ \mathsf{T}(\boldsymbol{\partial}[f]) \circ \phi^{-1} \circ \langle x, y \rangle \\
&= \; \mu \circ \boldsymbol{\partial}[\mathsf{T}(f)] \circ \phi \circ \phi^{-1} \circ \langle x, y \rangle && (\text{by Prop.6.4.v}) \\
&= \; \mu \circ \boldsymbol{\partial}[\mathsf{T}(f)] \circ \langle x, y \rangle
\end{aligned}$$

So we have that $\boldsymbol{\partial}^{T}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}} = \mu \circ \boldsymbol{\partial}[\mathsf{T}(f)] \circ \langle x, y \rangle$. This will help simplify many of the calculations to follow, since $\mathsf{T}(\boldsymbol{\partial}[f])$ appears everywhere due to the definition of Kleisli composition. We now prove the Cartesian difference category axioms.

$[\textbf{C}\boldsymbol{\partial}\textbf{.0}]$ $f \circ^{\mathsf{T}} (x +^{\mathsf{T}} \varepsilon^{\mathsf{T}}(y)) = f \circ^{\mathsf{T}} x +^{\mathsf{T}} \varepsilon^{\mathsf{T}} \left(\boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}} \right)$

First note that μ is linear and therefore additive. Then we compute that:

$$f \circ^{\mathsf{T}} x +^{\mathsf{T}} \varepsilon^{\mathsf{T}} \left(\boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}} \right) = f \circ^{\mathsf{T}} x + \varepsilon \left(\boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}} \right)$$

$$= \mu \circ \mathsf{T}(f) \circ x + \varepsilon \left(\mu \circ \boldsymbol{\partial}[\mathsf{T}(f)] \circ \langle x, y \rangle \right)$$

$$= \mu \circ \mathsf{T}(f) \circ x + \varepsilon \left(\mu \right) \circ \boldsymbol{\partial}[\mathsf{T}(f)] \circ \langle x, y \rangle \qquad \textbf{[E.2]}$$

$$= \mu \circ \mathsf{T}(f) \circ x + \mu \circ \varepsilon(1_{\mathsf{T}(A)}) \circ \boldsymbol{\partial}[\mathsf{T}(f)] \circ \langle x, y \rangle \qquad \text{(Lemma 6.4.vi)}$$

$$= \mu \circ \mathsf{T}(f) \circ x + \mu \circ \varepsilon \left(\boldsymbol{\partial}[\mathsf{T}(f)] \circ \langle x, y \rangle \right) \qquad \text{(Lemma 4.2)}$$

$$= \mu \circ \left(\mathsf{T}(f) \circ x + \varepsilon \left(\boldsymbol{\partial}[\mathsf{T}(f)] \circ \langle x, y \rangle \right) \right) \qquad (\mu \text{ is additive})$$

$$= \mu \circ \mathsf{T}(f) \circ (x + \varepsilon(y)) \qquad \textbf{[C}\boldsymbol{\partial}\textbf{.0]}$$

$$= f \circ^{\mathsf{T}} (x + \varepsilon(y))$$

$$= f \circ^{\mathsf{T}} (x +^{\mathsf{T}} \varepsilon^{\mathsf{T}}(y))$$

[C$\boldsymbol{\partial}$.1] $\boldsymbol{\partial}^{\mathsf{T}}[f +^{\mathsf{T}} g] = \boldsymbol{\partial}^{\mathsf{T}}[f] +^{\mathsf{T}} \boldsymbol{\partial}^{\mathsf{T}}[g]$, $[\boldsymbol{\partial}^{\mathsf{T}}[0^{\mathsf{T}}] = \boldsymbol{\partial}[0]$, and $\boldsymbol{\partial}^{\mathsf{T}}[\varepsilon^{\mathsf{T}}(f)] = \varepsilon^{\mathsf{T}} \left(\boldsymbol{\partial}^{\mathsf{T}}[f] \right)$

Since both the sum, zero maps, infinitesimal extension, and differential combinator in the Kleisli category are the same as in the base category, by **[C$\boldsymbol{\partial}$.1]** it easily follows that:

$$\boldsymbol{\partial}^{\mathsf{T}}[f +^{\mathsf{T}} g] = \boldsymbol{\partial}[f + g] = \boldsymbol{\partial}[f] + \boldsymbol{\partial}[g] = \boldsymbol{\partial}^{\mathsf{T}}[f] +^{\mathsf{T}} \boldsymbol{\partial}^{\mathsf{T}}[g]$$

$$\boldsymbol{\partial}^{\mathsf{T}}[0^{\mathsf{T}}] = \boldsymbol{\partial}[0] = 0 = 0^{\mathsf{T}}$$

$$\boldsymbol{\partial}^{\mathsf{T}}[\varepsilon^{\mathsf{T}}(f)] = \boldsymbol{\partial}[\varepsilon(f)] = \varepsilon \left(\boldsymbol{\partial}[f] \right) = \varepsilon^{\mathsf{T}} \left(\boldsymbol{\partial}^{\mathsf{T}}[f] \right)$$

[C$\boldsymbol{\partial}$.2] $\boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, y +^{\mathsf{T}} z \rangle^{\mathsf{T}} = \boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}} + \boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x + \varepsilon^{\mathsf{T}}(y), z \rangle^{\mathsf{T}}$ and $\boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, 0^{\mathsf{T}} \rangle^{\mathsf{T}} = 0^{\mathsf{T}}$

$$\boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \left\langle x, y +^{\mathsf{T}} z \right\rangle^{\mathsf{T}} = \mu \circ \boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \left\langle x, y +^{\mathsf{T}} z \right\rangle$$

$$= \mu \circ \boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \langle x, y + z \rangle$$

$$= \mu \circ \left(\boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \langle x, y \rangle + \boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \langle x + \varepsilon(y), z \rangle \right) \qquad \textbf{[C}\boldsymbol{\partial}\textbf{.2]}$$

$$= \mu \circ \boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \langle x, y \rangle + \mu \circ \boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \langle x + \varepsilon(y), z \rangle \quad (\mu \text{ is additive})$$

$$= \mu \circ \boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \langle x, y \rangle + \mu \circ \boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \left\langle x +^{\mathsf{T}} \varepsilon^{\mathsf{T}}(y), z \right\rangle$$

$$= \boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}} + \boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x + \varepsilon^{\mathsf{T}}(y), z \rangle^{\mathsf{T}}$$

$$\boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \left\langle x, 0^{\mathsf{T}} \right\rangle^{\mathsf{T}} = \boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, 0 \rangle^{\mathsf{T}}$$

$$= \mu \circ \boldsymbol{\partial} \left[\mathsf{T}(f) \right] \circ \langle x, 0 \rangle$$

$$= 0 \qquad \textbf{[C}\boldsymbol{\partial}\textbf{.2]}$$

$$= 0^{\mathsf{T}}$$

[C$\boldsymbol{\partial}$.3] $\boldsymbol{\partial}^{\mathsf{T}}[\eta] = \pi_1^{\mathsf{T}}$ and $\boldsymbol{\partial}^{\mathsf{T}}[\pi_i^{\mathsf{T}}] = \pi_i^{\mathsf{T}} \circ^{\mathsf{T}} \pi_1^{\mathsf{T}}$

Recall that $\pi_i^\mathsf{T} = \eta \circ \pi_i$ and so $\pi_i^\mathsf{T} \circ^\mathsf{T} \pi_1^\mathsf{T} = \eta \circ \pi_i \circ \pi_1$. Then since η is linear, we have that:

$$
\begin{aligned}
\boldsymbol{\partial}^\mathsf{T}[\eta] &= \boldsymbol{\partial}[\eta] \\
&= \eta \circ \pi_1 && (\eta \text{ is linear}) \\
&= \pi_1^\mathsf{T}
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\partial}^\mathsf{T}[\pi_i^\mathsf{T}] &= \boldsymbol{\partial}[\pi_i^\mathsf{T}] \\
&= \boldsymbol{\partial}[\eta \circ \pi_i] \\
&= \eta \circ \boldsymbol{\partial}[\pi_i] && (\eta \text{ linear and Lem.4.22.v}) \\
&= \eta \circ \pi_i \circ \pi_1 && [\mathbf{C}\boldsymbol{\partial}.\mathbf{3}]
\end{aligned}
$$

$[\mathbf{C}\boldsymbol{\partial}.\mathbf{4}]$ $\boldsymbol{\partial}^\mathsf{T}[\langle f, g\rangle^\mathsf{T}] = \langle \boldsymbol{\partial}^\mathsf{T}[f], \boldsymbol{\partial}^\mathsf{T}[g]\rangle^\mathsf{T}$

First note that since ϕ is linear, so it its inverse ϕ^{-1}. Then we have that:

$$
\begin{aligned}
\boldsymbol{\partial}^\mathsf{T}[\langle f, g\rangle^\mathsf{T}] &= \boldsymbol{\partial}[\langle f, g\rangle^\mathsf{T}] \\
&= \boldsymbol{\partial}\left[\phi^{-1} \circ \langle f, g\rangle\right] \\
&= \phi^{-1} \circ \boldsymbol{\partial}\left[\langle f, g\rangle\right] && (\phi^{-1} \text{ linear and Lem.4.22.v}) \\
&= \phi^{-1} \circ \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g]\rangle && [\mathbf{C}\boldsymbol{\partial}.\mathbf{4}] \\
&= \langle \boldsymbol{\partial}[f], \boldsymbol{\partial}[g]\rangle^\mathsf{T} \\
&= \langle \boldsymbol{\partial}^\mathsf{T}[f], \boldsymbol{\partial}^\mathsf{T}[g]\rangle^\mathsf{T}
\end{aligned}
$$

$[\mathbf{C}\boldsymbol{\partial}.\mathbf{5}]$ $\boldsymbol{\partial}^\mathsf{T}[g \circ^\mathsf{T} f] = \boldsymbol{\partial}^\mathsf{T}[g] \circ^\mathsf{T} \langle f \circ^\mathsf{T} \pi_0^\mathsf{T}, \boldsymbol{\partial}^\mathsf{T}[f]\rangle^\mathsf{T}$

First note that since $\pi_0^\mathsf{T} = \eta \circ \pi_0$, it easily follows that $f \circ^\mathsf{T} \pi_0^\mathsf{T} = f \circ \pi_0$ (using the monad identities and the naturality of η). Therefore, we compute that:

$$
\begin{aligned}
\boldsymbol{\partial}^\mathsf{T}[g] \circ^\mathsf{T} \langle f \circ^\mathsf{T} \pi_0^\mathsf{T}, \boldsymbol{\partial}^\mathsf{T}[f]\rangle^\mathsf{T} &= \mu \circ \boldsymbol{\partial}\left[\mathsf{T}(g)\right] \circ \langle f \circ^\mathsf{T} \pi_0^\mathsf{T}, \boldsymbol{\partial}^\mathsf{T}[f]\rangle \\
&= \mu \circ \boldsymbol{\partial}\left[\mathsf{T}(g)\right] \circ \langle f \circ \pi_0, \boldsymbol{\partial}[f]\rangle \\
&= \mu \circ \boldsymbol{\partial}[\mathsf{T}(g) \circ f] && [\mathbf{C}\boldsymbol{\partial}.\mathbf{5}] \\
&= \boldsymbol{\partial}\left[\mu \circ \mathsf{T}(g) \circ f\right] && (\mu \text{ linear and Lem.4.22.v}) \\
&= \boldsymbol{\partial}[g \circ^\mathsf{T} f] \\
&= \boldsymbol{\partial}^\mathsf{T}[g \circ^\mathsf{T} f]
\end{aligned}
$$

For the remaining two axioms, we will instead prove $[\mathbf{C}\boldsymbol{\partial}.\mathbf{6.a}]$ and $[\mathbf{C}\boldsymbol{\partial}.\mathbf{7.a}]$. Before we do so, we first compute the following:

$$
\begin{aligned}
\boldsymbol{\partial}^\mathsf{T}\left[\boldsymbol{\partial}^\mathsf{T}[f]\right] \circ^\mathsf{T} \left\langle \langle x, y\rangle^\mathsf{T}, \langle z, w\rangle^\mathsf{T}\right\rangle^\mathsf{T} &= \mu \circ \boldsymbol{\partial}\left[\mathsf{T}(\boldsymbol{\partial}^\mathsf{T}[f])\right] \circ \left\langle \langle x, y\rangle^\mathsf{T}, \langle z, w\rangle^\mathsf{T}\right\rangle \\
&= \mu \circ \boldsymbol{\partial}\left[\mathsf{T}(\boldsymbol{\partial}[f])\right] \circ \left\langle \phi^{-1} \circ \langle x, y\rangle, \phi^{-1} \circ \langle z, w\rangle\right\rangle \\
&= \mu \circ \boldsymbol{\partial}\left[\mathsf{T}(\boldsymbol{\partial}[f])\right] \circ (\phi^{-1} \times \phi^{-1}) \circ \langle\langle x, y\rangle, \langle z, w\rangle\rangle
\end{aligned}
$$

$$= \mu \circ \boldsymbol{\partial} \left[\mathsf{T}(\boldsymbol{\partial}[f]) \circ \phi^{-1} \right] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle \qquad (\phi^{-1} \text{ linear and Lem.4.22.v})$$
$$= \mu \circ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\mathsf{T}(f)] \circ \phi \circ \phi^{-1} \right] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle \qquad (\text{by Prop.6.4.v})$$
$$= \mu \circ \boldsymbol{\partial}^2 \left[\mathsf{T}(f) \right] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle$$

Therefore we have that $\boldsymbol{\partial}^{\mathsf{T}} \left[\boldsymbol{\partial}^{\mathsf{T}}[f] \right] \circ^{\mathsf{T}} \left\langle \langle x, y \rangle^{\mathsf{T}}, \langle z, w \rangle^{\mathsf{T}} \right\rangle^{\mathsf{T}} = \mu \circ \boldsymbol{\partial}^2 \left[\mathsf{T}(f) \right] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle$.

[C$\boldsymbol{\partial}$.6.a] $\boldsymbol{\partial}^{\mathsf{T}} \left[\boldsymbol{\partial}^{\mathsf{T}}[f] \right] \circ^{\mathsf{T}} \left\langle \langle x, 0^{\mathsf{T}} \rangle^{\mathsf{T}}, \langle 0^{\mathsf{T}}, y \rangle^{\mathsf{T}} \right\rangle^{\mathsf{T}} = \boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}}$

$$\boldsymbol{\partial}^{\mathsf{T}} \left[\boldsymbol{\partial}^{\mathsf{T}}[f] \right] \circ^{\mathsf{T}} \left\langle \langle x, 0^{\mathsf{T}} \rangle^{\mathsf{T}}, \langle 0^{\mathsf{T}}, y \rangle^{\mathsf{T}} \right\rangle^{\mathsf{T}} = \mu \circ \boldsymbol{\partial}^2 \left[\mathsf{T}(f) \right] \circ \left\langle \langle x, 0^{\mathsf{T}} \rangle, \langle 0^{\mathsf{T}}, y \rangle \right\rangle$$
$$= \mu \circ \boldsymbol{\partial}^2 \left[\mathsf{T}(f) \right] \circ \langle \langle x, 0 \rangle, \langle 0, y \rangle \rangle$$
$$= \mu \circ \boldsymbol{\partial}[\mathsf{T}(f)] \circ \langle x, y \rangle \qquad \text{[C}\boldsymbol{\partial}\text{.6.a]}$$
$$= \boldsymbol{\partial}^{\mathsf{T}}[f] \circ^{\mathsf{T}} \langle x, y \rangle^{\mathsf{T}}$$

[C$\boldsymbol{\partial}$.7.a] $\boldsymbol{\partial}^{\mathsf{T}} \left[\boldsymbol{\partial}^{\mathsf{T}}[f] \right] \circ^{\mathsf{T}} \left\langle \langle x, y \rangle^{\mathsf{T}}, \langle z, w \rangle^{\mathsf{T}} \right\rangle^{\mathsf{T}} = \boldsymbol{\partial}^{\mathsf{T}} \left[\boldsymbol{\partial}^{\mathsf{T}}[f] \right] \circ \left\langle \langle x, z \rangle^{\mathsf{T}}, \langle y, w \rangle^{\mathsf{T}} \right\rangle^{\mathsf{T}}$

$$\boldsymbol{\partial}^{\mathsf{T}} \left[\boldsymbol{\partial}^{\mathsf{T}}[f] \right] \circ^{\mathsf{T}} \left\langle \langle x, y \rangle^{\mathsf{T}}, \langle z, w \rangle^{\mathsf{T}} \right\rangle^{\mathsf{T}} = \mu \circ \boldsymbol{\partial}^2 \left[\mathsf{T}(f) \right] \circ \langle \langle x, y \rangle, \langle z, w \rangle \rangle$$
$$= \mu \circ \boldsymbol{\partial} \left[\boldsymbol{\partial}[\mathsf{T}(f)] \right] \circ \langle \langle x, z \rangle, \langle y, w \rangle \rangle \qquad \text{[C}\boldsymbol{\partial}\text{.7.a]}$$
$$= \boldsymbol{\partial}^{\mathsf{T}} \left[\boldsymbol{\partial}^{\mathsf{T}}[f] \right] \circ \left\langle \langle x, z \rangle^{\mathsf{T}}, \langle y, w \rangle^{\mathsf{T}} \right\rangle^{\mathsf{T}}$$

So we conclude that the Kleisli category is a Cartesian difference category. $\qquad \square$

We also point that in the case of a Cartesian differential category, since $\varepsilon = 0$, it follows that $\varepsilon^T = 0$. Therefore we have that the Kleisli category of the tangent bundle monad of a Cartesian differential category is again a Cartesian differential category. To the knowledge of the authors, this is a novel observation.

Corollary 6.14. *For a Cartesian differential category* $\mathbb{X}$*, the Kleisli category* $\mathbb{X}_{\mathsf{T}}$ *is a Cartesian differential category with differential combinator* $\mathsf{D}^{\mathsf{T}} = \mathsf{D}$*.*

We conclude this section by briefly taking a look at the linear maps and the ε^{T}-linear maps in the Kleisli category. A Kleisli map $f = \langle f_0, f_1 \rangle$ is linear in the Kleisli category if $\boldsymbol{\partial}^{\mathsf{T}}[f] = f \circ^{\mathsf{T}} \pi_1^{\mathsf{T}}$, which amounts to requiring that:

$$\langle \boldsymbol{\partial}[f_0], \boldsymbol{\partial}[f_1] \rangle = \langle f_0 \circ \pi_1, f_1 \circ \pi_1 \rangle$$

Therefore a Kleisli map is linear in the Kleisli category if and only if it is the pairing of maps which are linear in the base category. Similarly, a Kleisli map is ε^{T}-linear if and only if is the pairing of ε-linear maps.

7. Difference λ-Categories

Categorical models of the differential λ-calculus [ER03] are known as differential λ-categories [BEM10]. However, a differential λ-category is not simply a Cartesian differential category which is Cartesian closed. In a differential λ-category, both the additive structure and the differential structure must be compatible with the curry operator. The same is true of Cartesian difference categories. In this section, we introduce difference λ-categories.

Briefly, a difference λ-category is a Cartesian difference category which is Cartesian closed and such that the Cartesian difference structure and the curry operations are compatible.

For a Cartesian closed category $\mathbb{X}$, we denote the exponential as $A \Rightarrow B$, the evaluation map as $\mathbf{ev} : (A \Rightarrow B) \times A \to B$, and the curry of a map $f : A \times B \to C$ as $\Lambda(f) : A \to (B \Rightarrow C)$, that is, $\Lambda(f)$ is the unique map such that $\mathbf{ev} \circ (\Lambda(f) \times 1_B) = f$. Conversely, given a map of type $g : A \to (B \Rightarrow C)$, define $\Lambda^{-1}(g) : A \times B \to C$ as $\Lambda^{-1}(g) = \mathbf{ev} \circ (g \times 1_B)$. Of course, Λ and Λ^{-1} are inverses of each other, that is, $\Lambda\left(\Lambda^{-1}(g)\right) = g$ and $\Lambda^{-1}\left(\Lambda(f)\right) = f$. Another useful map will be the canonical natural isomorphism $\mathbf{sw} : (A \times B) \times C \to (A \times C) \times B$ which swaps the last arguments, that is, $\mathbf{sw}$ is defined as follows:

$$\mathbf{sw} := \langle\langle \pi_{00}, \pi_1 \rangle, \pi_{10} \rangle$$

where recall that $\pi_{ij} = \pi_i \circ \pi_j$. Note that $\mathbf{sw}$ is its own inverse, that is, $\mathbf{sw} \circ \mathbf{sw} = 1$.

Definition 7.1. A **Cartesian closed left additive category** is a Cartesian left additive category $\mathbb{X}$ such that $\mathbb{X}$ is Cartesian closed and such that the curry operator preserves the additive structure, that is, $\Lambda(f + g) = \Lambda(f) + \Lambda(g)$ and $\Lambda(0) = 0$.

Definition 7.2. A **difference λ-category** is a Cartesian difference category $\mathbb{X}$, with difference combinator $\boldsymbol{\partial}$ and infinitesimal extension ε, such that $\mathbb{X}$ is a Cartesian closed left additive category and the following additional axioms hold:

[C$\boldsymbol{\partial}\lambda$.1] $\boldsymbol{\partial}[\Lambda(f)] = \Lambda\left(\boldsymbol{\partial}[f] \circ \langle (\pi_0 \times 1), (\pi_1 \times 0) \rangle\right)$
[C$\boldsymbol{\partial}\lambda$.2] $\Lambda(\varepsilon(f)) = \varepsilon\left(\Lambda(f)\right)$

The first axiom **[C$\boldsymbol{\partial}\lambda$.1]** is identical to its differential combinator analogue in a differential λ-category. As such, it follows the same broad intuition. First note that **[C$\boldsymbol{\partial}\lambda$.1]** can also equivalently written in terms of $\mathbf{sw}$ as follows:

$$\boldsymbol{\partial}[\Lambda(f)] = \Lambda\left(\boldsymbol{\partial}[f] \circ ((1 \times 1) \times \langle 1, 0 \rangle) \circ \mathbf{sw}\right)$$

which will be useful for many calculations in this section. Now given a map $f : A \times B \to C$, we usually understand the composite:

$$\boldsymbol{\partial}[f] \circ (1_{A \times B} \times \langle 1_A, 0 \rangle) \circ \mathbf{sw} : (A \times A) \times B \to C$$

as the partial derivative of f with respect to its first argument A. Hence, **[C$\boldsymbol{\partial}\lambda$.1]** states that the derivative of a curried function is precisely the curry of the partial derivative of the function with respect to its first argument. On the other hand, **[C$\boldsymbol{\partial}\lambda$.2]** simply says that the curry of an infinitesimal extension of a function is the infinitesimal extension of the curry of the function. Using λ-calculus notation, this implies that $\lambda x.\varepsilon(f(t,y)) = \varepsilon(\lambda x.f(t,y))$.

Example 7.3. Every differential λ-category [BEM10] is precisely a difference λ-category such that $\varepsilon = 0$.

Example 7.4. The category $\overline{\mathsf{Ab}}$ (as defined in Section 5.2) is a difference λ-category. Given abelian groups G, H, the exponential $G \Rightarrow H$ is defined as the abelian group of all functions between G and H, where the group structure is defined point-wise. The remaining Cartesian closed structure is defined in the standard way, that is, for an arbitrary function $f : G \times H \to K$, $\Lambda(f)(x)(y) = f(x,y)$. We leave it to the reader to check for themselves that $\overline{\mathsf{Ab}}$ is indeed a Cartesian closed left additive category. Since $\varepsilon = 1$, clearly **[C$\boldsymbol{\partial}\lambda$.2]** holds automatically. So it remains to verify that the other axiom **[C$\boldsymbol{\partial}\lambda$.1]** also holds:

$$\begin{aligned}
\boldsymbol{\partial}[\Lambda(f)](x,u)(y) &= \Lambda(f)(x+u)(y) - \Lambda(f)(x)(y) \\
&= f(x+u, y) - f(x, y)
\end{aligned}$$

$$
\begin{aligned}
&= \; f\left((x,y) + (u,0)\right) - f(x,y) \\
&= \; \boldsymbol{\partial}[f](x,y,u,0) \\
&= \; (\boldsymbol{\partial}[f] \circ ((1 \times 1) \times \langle 1,0 \rangle) \circ \mathbf{sw})(x,u,y) \\
&= \; \Lambda(\boldsymbol{\partial}[f] \circ ((1 \times 1) \times \langle 1,0 \rangle) \circ \mathbf{sw})(x,u)(y)
\end{aligned}
$$

Therefore, $\overline{\mathsf{Ab}}$ is indeed a difference λ-category.

A central property of differential λ-categories is a deep correspondence between differentiation and the evaluation map. As one would expect, the partial derivative of the evaluation map gives a first-class derivative operator, see for example [BEM10, Lemma 4.5], which provides an interpretation for the differential substitution operator in the differential λ-calculus. This property still holds in difference λ-categories, although its formulation is somewhat more involved.

Lemma 7.5. *In a difference λ-category, for maps $g : A \times B \to C$ and $f : A \to B$, the following identities hold:*

(i) $\boldsymbol{\partial}[\mathbf{ev} \circ \langle \Lambda(g), f \rangle] = \mathbf{ev} \circ \langle \boldsymbol{\partial}[\Lambda(g)], f \circ \pi_0 \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0 + \varepsilon(\pi_1), f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle$

(ii) $\boldsymbol{\partial}[\mathbf{ev} \circ \langle \Lambda(g), f \rangle] = \mathbf{ev} \circ \langle \boldsymbol{\partial}[\Lambda(g)], f \circ \pi_0 + \varepsilon(\boldsymbol{\partial}[f]) \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle$

Proof. For (i), we compute:

$$
\begin{aligned}
\boldsymbol{\partial}[\mathbf{ev} \circ \langle \Lambda(g), f \rangle] &= \; \boldsymbol{\partial}[g \circ \langle 1_A, f \rangle] \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle 1_A, f \rangle \circ \pi_0, \boldsymbol{\partial}[\langle 1_A, f \rangle] \rangle && [\mathbf{C\partial.5}] \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle \pi_1, \boldsymbol{\partial}[f] \rangle \rangle && [\mathbf{C\partial.3}] + [\mathbf{C\partial.4}] \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle \pi_1, 0 \rangle + \langle 0, \boldsymbol{\partial}[f] \rangle \rangle \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle \pi_1, 0 \rangle \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle + \varepsilon(\langle \pi_1, 0 \rangle), \langle 0, \boldsymbol{\partial}[f] \rangle \rangle && [\mathbf{C\partial.2}] \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle \pi_1, 0 \rangle \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle + \langle \varepsilon(\pi_1), 0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle \\
&&& (\text{Lem } 4.3 + [\mathbf{E.1}]) \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle \pi_1, 0 \rangle \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0 + \varepsilon(\pi_1), f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle \\
&= \; \boldsymbol{\partial}[g] \circ \langle \pi_0 \times 1_B, \pi_1 \times 0 \rangle \circ \langle 1, f \circ \pi_0 \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0 + \varepsilon(\pi_1), f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle \\
&= \; \Lambda^{-1}\left(\Lambda\left(\boldsymbol{\partial}[g] \circ \langle \pi_0 \times 1_B, \pi_1 \times 0 \rangle\right)\right) \circ \langle 1, f \circ \pi_0 \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0 + \varepsilon(\pi_1), f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle \\
&= \; \Lambda^{-1}\left(\boldsymbol{\partial}[\Lambda[g]]\right) \circ \langle 1, f \circ \pi_0 \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0 + \varepsilon(\pi_1), f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle && (\text{by } [\mathbf{C\partial\lambda.1}]) \\
&= \; \mathbf{ev} \circ (\boldsymbol{\partial}[\Lambda[g]] \times 1_B) \circ \langle 1, f \circ \pi_0 \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0 + \varepsilon(\pi_1), f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle \\
&= \; \mathbf{ev} \circ \langle \boldsymbol{\partial}[\Lambda[g]], f \circ \pi_0 \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0 + \varepsilon(\pi_1), f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle
\end{aligned}
$$

On the other hand, for (ii), we compute:

$$
\begin{aligned}
\boldsymbol{\partial}[\mathbf{ev} \circ \langle \Lambda(g), f \rangle] &= \; \boldsymbol{\partial}[g \circ \langle 1_A, f \rangle] \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle 1_A, f \rangle \circ \pi_0, \boldsymbol{\partial}[\langle 1_A, f \rangle] \rangle && [\mathbf{C\partial.5}] \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle \pi_1, \boldsymbol{\partial}[f] \rangle \rangle && [\mathbf{C\partial.3}] + [\mathbf{C\partial.4}] \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle + \langle \pi_1, 0 \rangle \rangle \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle + \varepsilon(\langle 0, \boldsymbol{\partial}[f] \rangle), \langle \pi_1, 0 \rangle \rangle && [\mathbf{C\partial.2}] \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle + \langle 0, \varepsilon(\boldsymbol{\partial}[f]) \rangle, \langle \pi_1, 0 \rangle \rangle \\
&&& (\text{Lem } 4.3 + [\mathbf{E.1}]) \\
&= \; \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle \rangle + \boldsymbol{\partial}[g] \circ \langle \langle \pi_0, f \circ \pi_0 + \varepsilon(\boldsymbol{\partial}[f]) \rangle, \langle \pi_1, 0 \rangle \rangle
\end{aligned}
$$

$$= \quad \boldsymbol{\partial}[g] \circ \langle\langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle\rangle + \boldsymbol{\partial}[g] \circ \langle \pi_0 \times 1_B, \pi_1 \times 0 \rangle \circ \langle 1, f \circ \pi_0 + \varepsilon(\boldsymbol{\partial}[f]) \rangle$$

$$= \quad \boldsymbol{\partial}[g] \circ \langle\langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle\rangle + \Lambda^{-1}\left(\Lambda\left(\boldsymbol{\partial}[g] \circ \langle \pi_0 \times 1_B, \pi_1 \times 0 \rangle \right) \right) \circ \langle 1, f \circ \pi_0 + \varepsilon(\boldsymbol{\partial}[f]) \rangle$$

$$= \quad \boldsymbol{\partial}[g] \circ \langle\langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle\rangle + \Lambda^{-1}\left(\boldsymbol{\partial}[\Lambda[g]] \right) \circ \langle 1, f \circ \pi_0 + \varepsilon(\boldsymbol{\partial}[f]) \rangle \qquad \text{(by } [\mathbf{C}\boldsymbol{\partial}\boldsymbol{\lambda}.\mathbf{1}])$$

$$= \quad \boldsymbol{\partial}[g] \circ \langle\langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle\rangle + \mathbf{ev} \circ \left(\boldsymbol{\partial}[\Lambda[g]] \times 1_B \right) \circ \langle 1, f \circ \pi_0 + \varepsilon(\boldsymbol{\partial}[f]) \rangle$$

$$= \quad \boldsymbol{\partial}[g] \circ \langle\langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle\rangle + \mathbf{ev} \circ \langle \boldsymbol{\partial}[\Lambda[g]], f \circ \pi_0 + \varepsilon(\boldsymbol{\partial}[f]) \rangle$$

$$= \quad \mathbf{ev} \circ \langle \boldsymbol{\partial}[\Lambda[g]], f \circ \pi_0 + \varepsilon(\boldsymbol{\partial}[f]) \rangle + \boldsymbol{\partial}[g] \circ \langle\langle \pi_0, f \circ \pi_0 \rangle, \langle 0, \boldsymbol{\partial}[f] \rangle\rangle$$

Thus the desired identities hold. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

8. Conclusions and Future Work

We have presented Cartesian difference categories, which generalize Cartesian differential categories to account for more discrete definitions of derivatives while providing additional structure that is absent in change action models. We have also exhibited important examples and shown that Cartesian difference categories arise quite naturally from considering tangent bundles in any Cartesian differential category. We claim that Cartesian difference categories can facilitate the exploration of differentiation in discrete spaces, by generalizing techniques and ideas from the study of their differential counterparts. For example, Cartesian differential categories can be extended to allow objects whose tangent space is not necessarily isomorphic to the object itself [Cru17]. The same generalization could be applied to Cartesian difference categories — with some caveats: for example, the equation defining a linear map (Definition 4.21) becomes ill-typed, but the notion of ε-linear map remains meaningful.

Perhaps the most important topic for further research is identifying and cataloguing more instances of Cartesian difference categories. There is a number of very natural candidates for such; among those, we would like to single out synthetic differential geometry [Koc06] and non-standard analysis [Kei13], both of which feature a notion of "infinitesimal elements". In either case, multiplying by a suitably chosen infinitesimal ε would be an obvious choice of an infinitesimal extension. As suggested by one of the anonymous reviewers, the case of non-standard analysis based on hyperreal numbers might be of particular interest: since they contain invertible infinitesimals (as opposed to the nilpotent infinitesimals from SDG), one could define the derivative of a function f as $\boldsymbol{\partial}[f](x,y) = \varepsilon^{-1}\left(f(x + \varepsilon y) - f(x) \right)$ (assuming a choice of an "infinitesimal unit" ε). With an infinitesimal extension being defined by multiplication by the above infinitesimal unit ε, most of the axioms of a Cartesian difference category can be proven by a suitable generalisation of the proofs for the calculus of finite differences (Section 5.2), although more work would be needed to prove $[\mathbf{C}\boldsymbol{\partial}.\mathbf{7}]$. Lie groups, and Lie algebras, have also been suggested by one of the anonymous reviewers as a potential source of models, but work in this direction is at a less mature stage.

Another relevant path to consider is developing the analogue of the "tensor" story for Cartesian difference categories. Indeed, an important source of examples of Cartesian differential categories are the coKleisli categories of a tensor differential category [BCS06, BCS09]. A similar result likely holds for a hypothetical "tensor difference category", but it is not immediately clear how these should be defined: $[\mathbf{C}\boldsymbol{\partial}.\mathbf{2}]$ implies that derivatives in the difference sense are non-linear and therefore their interplay with the tensor structure will be much different. As suggested by one of the anonymous reviewers, a potential leading example of a "tensor difference category" is Ab, the category of abelian groups and group

homomorphisms between them, which is a model of linear logic via a linear-non-linear adjunction between Ab and SET, the category of sets and arbitrary set functions between them [Ben94]. The induced comonad (which models the exponential modality of linear logic) on Ab is given by mapping an abelian group G to the free abelian group over the underlying set of G. Even more interesting is the fact that the coKleisli category of this comonad is isomorphic to $\overline{\mathsf{Ab}}$, as defined in Section 5.2, which is one of our main examples of a Cartesian difference category. As such, this is a pretty convincing argument that Ab will indeed be a prime candidate for an example of a "tensor difference category" and help guide us in constructing the correct definition.

A further generalization of Cartesian differential categories, categories with tangent structure [CC14] are defined directly in terms of a tangent bundle functor rather than requiring that every tangent bundle be trivial (that is, in a tangent category it may not be the case that $\mathsf{T}A = A \times A$). Some preliminary research on change actions has already shown that, when generalized in this way, change actions are precisely internal categories, but the consequences of this for change action models (and, *a fortiori*, Cartesian difference categories) are not understood. More recently, some work has emerged about differential equations using the language of tangent categories [CC17]. We believe similar techniques can be applied in a straightforward way to Cartesian difference categories, where they might be of use to give an abstract formalization of discrete dynamical systems and difference equations.

References

[APETJO19] Mario Alvarez-Picallo, Alex Eyers-Taylor, Michael Peyton Jones, and C-H Luke Ong. Fixing incremental computation. In *European Symposium on Programming*, pages 525–552. Springer, 2019.

[APL20] Mario Alvarez-Picallo and Jean-Simon Pacaud Lemay. Cartesian difference categories. In *International Conference on Foundations of Software Science and Computation Structures*, pages 57–76. Springer, 2020.

[APO19] Mario Alvarez-Picallo and C-H Luke Ong. Change actions: models of generalised differentiation. In *International Conference on Foundations of Software Science and Computation Structures*, pages 45–61. Springer, 2019.

[BCS06] Richard F Blute, J Robin B Cockett, and Robert A G Seely. Differential categories. *Mathematical structures in computer science*, 16(06):1049–1083, 2006.

[BCS09] Richard F Blute, J Robin B Cockett, and Robert A G Seely. Cartesian differential categories. *Theory and Applications of Categories*, 22(23):622–672, 2009.

[BEM10] Antonio Bucciarelli, Thomas Ehrhard, and Giulio Manzonetto. Categorical models for simply typed resource calculi. *Electronic Notes in Theoretical Computer Science*, 265:213–230, 2010.

[Ben94] P Nick Benton. A mixed linear and non-linear logic: Proofs, terms and models. In *International Workshop on Computer Science Logic*, pages 121–135. Springer, 1994.

[BLR18] Jonathan Bradet-Legris and Hayley Reid. Differential forms in non-linear cartesian differential categories. Foundational Methods in Computer Science, 2018.

[Cam18] Alexander Campbell. Skew-enriched categories. *Applied Categorical Structures*, 26(3):597–615, Jun 2018.

[CC14] J Robin B Cockett and Geoff S H Cruttwell. Differential structure, tangent structure, and sdg. *Applied Categorical Structures*, 22(2):331–417, 2014.

[CC17] JRB Cockett and GSH Cruttwell. Connections in tangent categories. *Theory and Applications of Categories*, 32(26):835–888, 2017.

[CGRO14] Yufei Cai, Paolo G Giarrusso, Tillmann Rendel, and Klaus Ostermann. A theory of changes for higher-order languages: Incrementalizing λ-calculi by static differentiation. In *ACM SIGPLAN Notices*, volume 49, pages 145–155. ACM, 2014.

[Cru17]   Geoff SH Cruttwell. Cartesian differential categories revisited. *Mathematical Structures in Computer Science*, 27(1):70–91, 2017.

[Ehr18]   Thomas Ehrhard. An introduction to differential linear logic: proof-nets, models and antiderivatives. *Mathematical Structures in Computer Science*, 28(7):995–1060, 2018.

[ER03]   T Ehrhard and L Regnier. The differential lambda-calculus. *Theoretical Computer Science*, 309(1):1–41, 2003.

[Gle05]   David Gleich. Finite calculus: A tutorial for solving nasty sums. *Stanford University*, 2005.

[JJ65]   Charles Jordan and Károly Jordán. *Calculus of finite differences*, volume 33. American Mathematical Soc., 1965.

[Kei13]   H Jerome Keisler. *Elementary calculus: An infinitesimal approach*. Courier Corporation, 2013.

[Koc06]   Anders Kock. *Synthetic differential geometry*, volume 333. Cambridge University Press, 2006.

[KPS16]   Robert Kelly, Barak A Pearlmutter, and Jeffrey Mark Siskind. Evolving the incremental λ calculus into a model of forward automatic differentiation (ad). *arXiv preprint arXiv:1611.03429*, 2016.

[Lem18]   J.-S. P. Lemay. A tangent category alternative to the faa di bruno construction. *Theory and Applications of Categories*, 33(35):1072–1110, 2018.

[Man12a]   Giulio Manzonetto. What is a categorical model of the differential and the resource λ-calculi? *Mathematical Structures in Computer Science*, 22(3):451–520, 2012.

[Man12b]   Oleksandr Manzyuk. Tangent bundles in differential lambda-categories. *arXiv preprint arXiv:1202.0411*, 2012.

[Ric54]   Clarence Hudson Richardson. *An introduction to the calculus of finite differences*. Van Nostrand, 1954.

[Rut05]   Jan JMM Rutten. A coinductive calculus of streams. *Mathematical Structures in Computer Science*, 15(1):93, 2005.

[SJ19]   David Sprunger and Bart Jacobs. The differential calculus of causal functions. *arXiv preprint arXiv:1904.10611*, 2019.

[SK19]   David Sprunger and Shin-ya Katsumata. Differentiable causal computations via delayed trace. In *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–12. IEEE, 2019.

[SP09]   Bernd Steinbach and Christian Posthoff. Boolean differential calculus. In *Logic Functions and Equations*, pages 75–103. Springer, 2009.

[Tha81]   André Thayse. *Boolean calculus of differences*. Number 101. Springer Science & Business Media, 1981.