Math. Struct. in Comp. Science (2012), *vol.* 22, *pp.* 451–520. © Cambridge University Press 2012
doi:10.1017/S0960129511000594 First published online 27 February 2012

What is a categorical model of the differential and the resource λ-calculi?

GIULIO MANZONETTO[†]

Univ. Paris 13, Sorbonne Paris Cité, LIPN, UMR 7030, CNRS, F-93430 Villetaneuse, France
Email: giulio.manzonetto@lipn.univ-paris13.fr

Received 16 June 2001; revised 8 October 2002

The *differential λ-calculus* is a paradigmatic functional programming language endowed with a syntactical differentiation operator that allows the application of a program to an argument in a linear way. One of the main features of this language is that it is *resource conscious* and gives the programmer suitable primitives to handle explicitly the resources used by a program during its execution. The differential operator also allows us to write the full Taylor expansion of a program. Through this expansion, every program can be decomposed into an infinite sum (representing non-deterministic choice) of 'simpler' programs that are strictly linear.

The aim of this paper is to develop an abstract 'model theory' for the untyped differential λ-calculus. In particular, we investigate what form a general categorical definition of a denotational model for this calculus should take. Starting from the work of Blute, Cockett and Seely on differential categories, we develop the notion of a *Cartesian closed differential category* and prove that *linear reflexive objects* living in such categories constitute sound and complete models of the untyped differential λ-calculus. We also give sufficient conditions for Cartesian closed differential categories to model the Taylor expansion. This requires that every model living in such categories equates all programs having the same full Taylor expansion.

We then provide a concrete example of a Cartesian closed differential category modelling the Taylor expansion, namely the category **MRel** of sets and relations from finite multisets to sets. We prove that the extensional model $\mathscr{D}$ of λ-calculus we have recently built in **MRel** is linear, and is thus also an extensional model of the untyped differential λ-calculus. In the same category, we build a non-extensional model $\mathscr{E}$ and prove that it is, nevertheless, extensional on its differential part.

Finally, we study the relationship between the differential λ-calculus and the *resource calculus*, which is a functional programming language combining the ideas behind the differential λ-calculus with those behind Boudol's λ-calculus with multiplicities. We define two translation maps between these two calculi and study the properties of these translations. In particular, this analysis shows that the two calculi share the same notion of a model, and thus that the resource calculus can be interpreted by translation into every linear reflexive object living in a Cartesian closed differential category.

[†] This work was partly funded by the NWO Project 612.000.936 CALMOC (CAtegorical and ALgebraic Models of Computation) and by Digiteo/Île-de-France Project 2009-28HD COLLODI (Complexity and concurrency through ludics and differential linear logic).

Contents

1. Introduction

Among the many computational formalisms that have been studied in the literature, the λ-calculus (Barendregt 1984) plays an important role as a bridge between logic and computer science. The λ-calculus was originally introduced by Church (Church 1932; 1941) as a

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

foundation for mathematics, where functions, and not sets, were primitive. This system turned out to be both consistent and successful as a tool for formalising all computable functions. However, the λ-calculus is not resource sensitive since a λ-term can erase its arguments or duplicate them an arbitrarily large number of times. This becomes problematic when one wants to deal with programs that are executed in environments with bounded resources (like PDA's) or in the presence of depletable arguments (like quantum data that cannot be duplicated for physical reasons). In these contexts, we want to be able to express the fact that a program *actually consumes* its arguments. Such an idea of 'resource consumption' is central to Girard's quantitative semantics (Girard 1988). This semantics establishes an analogy between linearity in the sense of computer science (programs using arguments exactly once) and algebraic linearity (the commutation of sums and products with scalars), giving a new mathematically very appealing interpretation of resource consumption. Drawing on these insights, Ehrhard and Regnier designed a resource sensitive paradigmatic programming language called *the differential λ-calculus* (Ehrhard and Regnier 2003).

The differential λ-calculus. The differential λ-calculus is a conservative (see Ehrhard and Regnier (2003, Proposition 19)) extension of the untyped λ-calculus with differential and linear constructions. In this language, two different operators can be used to apply a program to its argument: the usual application and a *linear application*. The latter defines a syntactic derivative operator $\mathsf{D}s \cdot t$, which is an excellent candidate for increasing control over programs executed in environments with bounded resources. Indeed, the evaluation of $\mathsf{D}s \cdot t$ (the derivative of the program s on the argument t) has a precise operational meaning: it captures the fact that t is available for s 'exactly once'. The corresponding meta-operation of substitution, which replaces exactly one (linear) occurrence of x in s by t, is called 'differential substitution' and is denoted by

$$\frac{\partial s}{\partial x} \cdot t.$$

It is worth noting that when s contains several occurrences of x, one has to choose which occurrence should be replaced, and there are several possible choices. When s does not contain any occurrence of x, then the differential substitution cannot be performed and the result is 0 (corresponding to an empty program). Thus, the differential substitution forces non-determinism in the system, which is represented by a formal sum having 0 as neutral element. In this way, the differential λ-calculus constitutes a useful framework for studying the notions of linearity and non-determinism, and the relation between them.

Taylor expansion. As expected, iterated differentiation yields a natural notion of linear approximation of the ordinary application of a program to its argument. Indeed, the syntactic derivative operator allows us to write all the derivatives of a λ-term M, so it also allows us (in the presence of countable sums) to define its *full Taylor expansion* M^*. In general, M^* will be an infinite formal linear combination of simple terms (with coefficients

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

in a field), and should satisfy, when M is the usual application NQ:

$$(NQ)^* = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\mathsf{D}^n N \cdot \left(\underbrace{Q, \ldots, Q}_{n \text{ times}} \right) \right) 0$$

where $1/n!$ is a numerical coefficient and $\mathsf{D}^n N \cdot (Q, \ldots, Q)$ stands for iterated linear application of N to n copies of Q. The precise operational meaning of the Taylor expansion has been extensively studied in Ehrhard and Regnier (2003; 2006a; 2008). The crucial property of such an expansion is that it gives a *quantitative* account to the β-reduction of λ-calculus (in the sense of Böhm tree computation). Formal connections between Taylor expansions and Böhm trees of ordinary λ-terms were presented in Ehrhard and Regnier (2006a) using a decorated version of Krivine's machine.

The resource calculus. The resource calculus is essentially a reworking of Boudol's λ-calculus with multiplicities (Boudol 1993; Boudol and Curien and Lavatelli) and provides an alternative approach to the problem of modelling resource consumption within a functional programming language. In this calculus, there is only one application operator, while the arguments can be either linear or reusable and come in finite multisets called 'bags'. Linear arguments must be used exactly once, while reusable ones can be used *ad libitum*. Also, the evaluation of a function applied to a bag of arguments in this setting may give rise to different possible choices, which correspond to the different possibilities of distributing the arguments between the occurrences of the formal parameter.

The main differences between Boudol's calculus and the resource calculus are that the former is affine, is equipped with explicit substitution and has a lazy operational semantics, while the latter is linear and is a true extension of the classical λ-calculus. The current formalisation of resource calculus was proposed in Tranquilli (2009) with the aim of defining a Curry–Howard correspondence with differential nets (Ehrhard and Regnier 2006b).

The resource calculus has been recently studied from a syntactic point of view by Pagani and Tranquilli (Pagani and Tranquilli 2009) for confluence results, by Pagani and the current author (Manzonetto and Pagani 2011) for separability results, and by Pagani and Ronchi della Rocca for results on may and must solvability (Pagani and Ronchi Della Rocca 2010). Algebraic notions of models for the *strictly linear fragment* of the resource calculus have been proposed by Carraro, Ehrhard and Salibra in Carraro *et al.* (2010). Our main focus in the current paper is a study of the differential λ-calculus, but we will also draw conclusions for the resource calculus.

Denotational semantics. Although the differential λ-calculus was born out of semantical considerations (that is, the deep analysis of coherent spaces performed by Ehrhard and Regnier), the analysis of its denotational semantics are just beginning. It is known that finiteness spaces (Ehrhard 2005) and the relational semantics of linear logic (Girard 1988) are examples of models of the *simply typed* differential λ-calculus, and thus have very limited expressive power. When it comes to the *untyped* differential λ-calculus, it is *folklore* that the relational model $\mathscr{D}$ introduced in Bucciarelli *et al.* (2007) in the relational semantics constitutes a concrete example of a model[†]. This picture is reminiscent of the beginnings of the denotational semantics of λ-calculus, when Scott's $\mathscr{D}_\infty$ was the only concrete example of a model of λ-calculus, but no general definition of a model was known. Only when an abstract model theory for this calculus had been developed were researchers able to provide rich semantics (such as the continuous (Scott 1972), stable (Berry 1978) and strongly stable semantics (Bucciarelli and Ehrhard 1991)), and general methods for building huge classes of models in these semantics.

Categorical notion of a model

The aim of the current paper is to provide a general categorical notion of a model of the untyped differential λ-calculus. Our starting point will be the work of Blute, Cockett and Seely on (Cartesian) differential categories (Blute *et al.* 2006; 2009). In these categories, a derivative operator $D(-)$ on morphisms is equationally axiomatised; the derivative of a morphism $f : A \to B$ will be a morphism $D(f) : A \times A \to B$ that is linear in its first component. Blute *et al.* then proved that these categories are sound and complete for modelling suitable term calculi. However, it turns out that the properties of differential categories are too weak for modelling the full differential λ-calculus. For this reason, we will introduce the more powerful notion of a *Cartesian closed differential category*. In such categories, we can define an operator

$$\frac{f : C \times A \to B \quad g : C \to A}{f \star g : C \times A \to B} \ (\star)$$

that can be seen as a categorical counterpart to differential substitution. Intuitively, the morphism $f \star g$ is obtained by force-feeding the second argument A of f with *one copy* of the result of g. However, the type is not modified because $f \star g$ may still depend on A.

The operator $\star$ allows us to interpret the differential λ-calculus in every *linear reflexive object* $\mathscr{U}$ living in a Cartesian closed differential category $\mathbf{C}$. For a reflexive object $\mathscr{U} = (U, \mathcal{A}, \lambda)$ 'to be linear' amounts to requiring that the morphisms $\mathcal{A}$ and λ performing the retraction $(U \Rightarrow U) \lhd U$ are linear. We will prove that this categorical notion of a model is *sound*, which means that the induced equational theory $\mathrm{Th}(\mathscr{U})$ is actually a *differential λ-theory*. We will also investigate what conditions the category $\mathbf{C}$ should satisfy in order to *model the Taylor expansion*. This means that all differential programs having the same Taylor expansion are equated in every model living in $\mathbf{C}$.

A question that arises naturally when a notion of a model of a certain calculus is introduced is whether it is *equationally complete*, that is, whether all equational theories of that calculus can be represented. For instance, in the case of the untyped λ-calculus, Scott and Koymans proved that for every λ-theory $\mathcal{T}$ there is a reflexive object $\mathscr{U}$ in a Cartesian closed category $\mathbf{C}$ such that $\mathrm{Th}(\mathscr{U}) = \mathcal{T}$. We will prove that the notion of a linear reflexive object in a Cartesian closed differential category is equationally complete for the

[†] This follows from Ehrhard and Regnier (2006b), which showed that the differential λ-calculus can be translated into differential proofnets, plus Vaux (2007b), which proved that $\mathscr{D}$ is a model of such proofnets.

differential λ-calculus, provided we only consider theories satisfying suitable properties. The first property is that in these theories the sum is considered to be idempotent, which amounts to saying that we only know whether a term appears in a result, not how many times it appears; the second is that these theories are 'extensional on linear applications', which means that $\mathsf{D}s \cdot t$ must have a functional behaviour. It turns out that these properties are quite natural in the sense that they are satisfied by all models that have arisen so far.

Relational semantics

In Bucciarelli *et al.* (2007), the current author built, in collaboration with Bucciarelli and Ehrhard, an extensional model $\mathscr{D}$ of λ-calculus living in the category **MRel** of sets and 'relations from finite multisets to sets'. This model can be viewed as a relational analogue of Scott's $\mathscr{D}_\infty$ (Ehrhard 2009). By virtue of its logical nature, $\mathscr{D}$ can be used to model several systems beyond the untyped λ-calculus. For instance, Bucciarelli *et al.* (2009) proved that it constitutes an adequate model of a λ-calculus extended with non-deterministic choice and parallel composition, while Vaux (2007b) showed that it is a model of differential proof-nets.

In the current paper, we study $\mathscr{D}$ as a model of the untyped differential λ-calculus. Indeed (as expected), the category **MRel** turns out to be an instance of the definition of a Cartesian closed differential category, and the relational model $\mathscr{D}$ is easily checked to be linear. We will then study the equational theory induced by $\mathscr{D}$ and prove that it equates all terms having the same Taylor expansion. This property follows from the fact that **MRel** models the Taylor expansion. As a simple consequence, we get that the relational semantics is *hugely incomplete* – there is a continuum of equational theories that are not representable by models living in **MRel**.

In the same category, we will also build a model $\mathscr{E}$ that can be seen as a relational analogue of Engeler's graph model (Engeler 1981). The model $\mathscr{E}$ provides an example of a non-extensional model, which is, however, extensional on linear applications.

Translations

Finally, we study the inter-relationships existing between the differential λ-calculus and the resource calculus. In fact, it is commonly believed in the scientific community that the two calculi are essentially the same, and the choice of studying one language rather than the other is more a matter of taste than substance. We will give a formal meaning to this belief by defining a translation map $(\cdot)^r$ from the differential λ-calculus to the resource calculus, and another map $(\cdot)^d$ in the opposite direction. We will prove that these translations are 'faithful' in the sense that equivalent programs of differential λ-calculus are mapped into equivalent resource programs, and *vice versa*. This shows that the two calculi share the same notion of denotational model: in particular, the resource calculus can be interpreted by translation in every linear reflexive object living in a Cartesian closed differential category.

Outline

Section 2 presents some preliminary notions and notation used in the rest of the paper. In Section 3, we present the syntax and axioms of the differential λ-calculus, and define the associated equational theories. In Section 4, we introduce the notion of a Cartesian closed differential category. In section 5, we show that linear reflexive objects in such categories are sound and complete models of the differential λ-calculus. In Section 6, we build two relational models $\mathscr{D}$ and $\mathscr{E}$ and provide a partial characterisation of their equational theories. In Section 7, we define the resource calculus and study its relationship to the differential λ-calculus. Finally, in Section 8, we discuss related work, present our conclusions and propose some further lines of research.

2. Preliminaries

To make this article more self-contained, this section summarises some definitions and results we will use later in the paper. Our main reference for category theory is Asperti and Longo (1991).

2.1. *Sets and multisets*

We use $\mathbf{N}$ to denote the set of natural numbers. Given $n \in \mathbf{N}$, we write $\mathfrak{S}_n$ for the set of all permutations (bijective maps) of the set $\{1, \dots, n\}$.

Let A be a set. We use $\mathcal{P}(A)$ to denote the powerset of A. A *multiset* m over A can be defined as an unordered list $m = [a_1, a_2, \dots]$ with repetitions such that $a_i \in S$ for all indices i. A multiset m is said to be *finite* if it is a finite list; we use $[]$ to denote the empty multiset. Given two multisets $m_1 = [a_1, a_2, \dots]$ and $m_2 = [b_1, b_2, \dots]$, the *multi-union* of m_1, m_2 is defined by $m_1 \uplus m_2 = [a_1, b_1, a_2, b_2, \dots]$.

Finally, we write $\mathcal{M}_f(A)$ for the set of all finite multisets over A.

2.2. *Cartesian (closed) categories*

Let $\mathbf{C}$ be a *Cartesian category* and A, B, C be arbitrary objects of $\mathbf{C}$. We write $\mathbf{C}(A, B)$ for the homset of morphisms from A to B; when there is no risk of confusion, we write $f : A \to B$ instead of $f \in \mathbf{C}(A, B)$. We usually write $A \times B$ to denote the *categorical product* of A and B, and $\pi_1 : A \times B \to A$, $\pi_2 : A \times B \to B$ for the associated *projections*, and given a pair of arrows $f : C \to A$ and $g : C \to B$, we use $\langle f, g \rangle : C \to A \times B$ to denote the unique arrow such that $\pi_1 \circ \langle f, g \rangle = f$ and $\pi_2 \circ \langle f, g \rangle = g$. We write $f \times g$ for the *product map of f and g*, which is defined by $f \times g = \langle f \circ \pi_1, g \circ \pi_2 \rangle$.

If the category $\mathbf{C}$ is *Cartesian closed*, we write $A \Rightarrow B$ for the *exponential object* and $\mathrm{ev}_{AB} : (A \Rightarrow B) \times A \to B$ for the *evaluation morphism*. Moreover, for any object C and arrow $f : C \times A \to B$, we write $\Lambda(f) : C \to (A \Rightarrow B)$ for the (unique) morphism such that $\mathrm{ev}_{AB} \circ (\Lambda(f) \times \mathrm{Id}_A) = f$. Finally, $\mathbb{1}$ denotes the terminal object and $!_A$ the only morphism in $\mathbf{C}(A, \mathbb{1})$.

Recall that the following equalities hold in every Cartesian closed category:

(pair) $\langle f, g \rangle \circ h = \langle f \circ h, g \circ h \rangle$

(beta-cat) $\mathrm{ev} \circ \langle \Lambda(f), g \rangle = f \circ \langle \mathrm{Id}, g \rangle$

(Curry) $\Lambda(f) \circ g = \Lambda(f \circ (g \times \mathrm{Id}))$

(Id-Curry) $\Lambda(\mathrm{ev}) = \mathrm{Id}.$

Moreover, we can define the *uncurry* operator

$$\Lambda^{-}(-) = \mathrm{ev} \circ (- \times \mathrm{Id}).$$

From (beta-cat), (Curry) and (Id-Curry) it follows that

$$\Lambda(\Lambda^{-}(f)) = f$$
$$\Lambda^{-}(\Lambda(g)) = g.$$

3. The differential lambda calculus

In this section we recall the definition of the *differential λ-calculus* (Ehrhard and Regnier 2003), together with some standard properties of the language. We also define the associated equational theories, namely, the *differential λ-theories*. The syntax we use in the present paper is freely adapted from Vaux (2007a).

3.1. *Differential lambda terms*

The set Λ^d of *differential λ-terms* and the set Λ^s of *simple terms* are defined by mutual induction as follows:

$$\begin{array}{llcl} \Lambda^d : & S, T, U, V & ::= & s \mid 0 \mid s + T \\ \Lambda^s : & s, t, u, v & ::= & x \mid \lambda x.s \mid sT \mid \mathsf{D}s \cdot t. \end{array}$$

The differential λ-term $\mathsf{D}s \cdot t$ represents the *linear application* of s to t. Intuitively, this means that s is provided with exactly one copy of t. Notice that sums may also appear in simple terms as the right-hand components of ordinary applications. Although the rule $s + t = s$ will not be valid in our axiomatisation, the sum should still be thought of as a version of non-deterministic choice where all actual choice operations are postponed.

Convention 3.1. We consider differential λ-terms up to α-conversion, and up to associativity and commutativity of the sum. The term 0 is the neutral element of the sum, so we also add the equation $S + 0 = S$.

As notation, we will write

$$\lambda x_1 \ldots x_n.s \;\; \text{for} \;\; \lambda x_1.(\cdots(\lambda x_n.s)\cdots)$$
$$sT_1 \cdots T_k \;\; \text{for} \;\; (\cdots(sT_1)\cdots)T_k.$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

We also set

$$\mathsf{D}^1 s \cdot (t_1) = \mathsf{D} s \cdot t_1$$
$$\mathsf{D}^{n+1} s \cdot (t, t_1, \dots, t_n) = \mathsf{D}^n (\mathsf{D} s \cdot t) \cdot (t_1, \dots, t_n).$$

When writing $\mathsf{D}^n s \cdot (t_1, \dots, t_n)$, we assume $n > 0$.

Definition 3.2. The *permutative equality* on differential λ-terms requires

$$\mathsf{D}^n s \cdot (t_1, \dots, t_n) = \mathsf{D}^n s \cdot \left(t_{\sigma(1)}, \dots, t_{\sigma(n)}\right)$$

for all permutations $\sigma \in \mathfrak{S}_{\mathfrak{n}}$.

From now on, differential λ-terms will also be considered up to permutative equality. This is needed, for instance, to prove the Schwarz Theorem (see Section 3.2), which allows us to speak of a differential operator. For specific λ-terms, we define

$$\begin{aligned}
\mathbf{I} &\equiv \lambda x.x \\
\mathbf{1} &\equiv \lambda xy.xy \\
\Delta &\equiv \lambda x.xx \\
\Omega &\equiv \Delta\Delta \\
\mathbf{Y} &\equiv \lambda f.(\lambda x.f(xx))(\lambda x.f(xx)) \\
\mathbf{s} &\equiv \lambda nxy.nx(xy) \\
\underline{n} &\equiv \lambda sx.s^n(x), \text{ for every natural number } n \in \mathbf{N},
\end{aligned}$$

where $\equiv$ stands for syntactical equality up to the above mentioned equivalences on differential λ-terms. Note that $\mathbf{I}$ is the identity, $\mathbf{Y}$ is Curry's fixpoint combinator, $\underline{n}$ is the nth Church numeral and $\mathbf{s}$ implements the successor function on Church numerals. The term Ω denotes the usual paradigmatic unsolvable λ-term.

Definition 3.3. Let S be a differential λ-term. The set $\mathrm{FV}(S)$ of *free variables* of S is defined inductively as follows:

$$\begin{aligned}
\mathrm{FV}(x) &= \{x\} \\
\mathrm{FV}(\lambda x.s) &= \mathrm{FV}(s) - \{x\} \\
\mathrm{FV}(sT) &= \mathrm{FV}(s) \cup \mathrm{FV}(T) \\
\mathrm{FV}(\mathsf{D} s \cdot t) &= \mathrm{FV}(s) \cup \mathrm{FV}(t) \\
\mathrm{FV}(0) &= \varnothing \\
\mathrm{FV}(s + S) &= \mathrm{FV}(s) \cup \mathrm{FV}(S).
\end{aligned}$$

Given differential λ-terms $S_1, \dots, S_k$, we set

$$\mathrm{FV}(S_1, \dots, S_k) = \mathrm{FV}(S_1) \cup \cdots \cup \mathrm{FV}(S_k).$$

We will now introduce some notation for differential λ-terms that will be particularly useful when we define the substitution operators in the next section.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Notation 3.4. We will often use the following abbreviations (note that these are just syntactic sugar, not real terms):

$$\lambda x.\left(\sum_{i=1}^{k} s_i\right) = \sum_{i=1}^{k} \lambda x.s_i$$
$$\left(\sum_{i=1}^{k} s_i\right) T = \sum_{i=1}^{k} s_i T$$
$$\mathsf{D}\left(\sum_{i=1}^{k} s_i\right) \cdot \left(\sum_{j=1}^{n} t_j\right) = \sum_{i,j} \mathsf{D} s_i \cdot t_j.$$

Intuitively, these equalities make sense since the lambda abstraction is linear, the usual application is linear in its left-hand component, and the linear application is a bilinear operator. Notice, however, that

$$S\left(\sum_{i=1}^{k} t_i\right) \neq \sum_{i=1}^{k} S t_i.$$

Note that in the particular case of empty sums, we get

$$\lambda x.0 = 0$$
$$0T = 0$$
$$\mathsf{D}0 \cdot t = 0$$
$$\mathsf{D}s \cdot 0 = 0.$$

Thus 0 annihilates any term, except when it occurs on the right of an ordinary application.

3.2. *Two kinds of substitution*

We now introduce two kinds of meta-operation for substitution in differential λ-terms: the usual capture-free substitution and differential substitution. Both definitions make free use of the abbreviations introduced in Notation 3.4.

Definition 3.5. Let S, T be differential λ-terms and x be a variable. The *capture-free substitution* of T for x in S, denoted by $S\{T/x\}$, is defined by induction on S as follows:

$$y\{T/x\} = \begin{cases} T & \text{if } x = y, \\ y & \text{otherwise,} \end{cases}$$
$$(\lambda y.s)\{T/x\} = \lambda y.s\{T/x\} \qquad \text{(where we suppose by } \alpha\text{-conversion that } x \neq y \text{ and } y \notin \mathrm{FV}(T))$$
$$(sU)\{T/x\} = (s\{T/x\})(U\{T/x\}),$$
$$(\mathsf{D}^n s \cdot (u_1, \ldots, u_n))\{T/x\} = \mathsf{D}^n(s\{T/x\}) \cdot \left(u_1\{T/x\}, \ldots, u_n\{T/x\}\right),$$
$$0\{T/x\} = 0$$
$$(s + S)\{T/x\} = s\{T/x\} + S\{T/x\}.$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Thus, $S\{T/x\}$ is the result of substituting T for all free occurrences of x in S, subject to the usual proviso about renaming bound variables in S to avoid capture of free variables in T. On the other hand, the differential substitution

$$\frac{\partial S}{\partial x}\cdot T$$

defined below denotes the result of substituting T (still avoiding capture of variables) for *exactly one* – non-deterministically chosen – linear occurrence of x in S. If there is no such occurrence in S, then the result will be 0.

Definition 3.6. Let S, T be differential λ-terms and x be a variable. The *differential substitution* of T for x in S is denoted by

$$\frac{\partial S}{\partial x}\cdot T$$

and defined by induction on S as follows:

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

$$\frac{\partial}{\partial x}(sU)\cdot T = \left(\frac{\partial s}{\partial x}\cdot T\right)U + \left(\mathsf{D}s\cdot\left(\frac{\partial U}{\partial x}\cdot T\right)\right)U$$

$$\frac{\partial}{\partial x}(\lambda y.s)\cdot T = \lambda y.\frac{\partial s}{\partial x}\cdot T \qquad \text{(where we suppose by } \alpha\text{-conversion that } x \neq y \text{ and } y \notin \mathrm{FV}(T))$$

$$\frac{\partial}{\partial x}(\mathsf{D}^n s\cdot(u_1,\ldots,u_n))\cdot T = \mathsf{D}^n\left(\frac{\partial s}{\partial x}\cdot T\right)\cdot(u_1,\ldots,u_n) + \sum_{i=1}^{n}\mathsf{D}^n s\cdot\left(u_1,\ldots,\frac{\partial u_i}{\partial x}\cdot T,\ldots,u_n\right)$$

$$\frac{\partial 0}{\partial x}\cdot T = 0$$

$$\frac{\partial}{\partial x}(s+U)\cdot T = \frac{\partial s}{\partial x}\cdot T + \frac{\partial U}{\partial x}\cdot T.$$

The definition states that the differential substitution distributes over linear constructions. We will now look briefly at the case of the standard application sU since it is the most complex one. The result of

$$\frac{\partial(sU)}{\partial x}\cdot T$$

is the sum of two terms since the differential substitution can non-deterministically be applied to either s or to U. In the first case, we can safely apply it to s since the standard application is linear in its left-hand argument, so we obtain

$$\left(\frac{\partial s}{\partial x}\cdot T\right)U.$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

However, in the other case, we cannot apply it directly to U because the standard application is *not* linear in its right-hand argument, so we follow two steps:

(i) we replace sU by $(\mathsf{D}s \cdot U)U$;
(ii) we apply the differential substitution to the linear copy of U.

Intuitively, this works because U is essentially available infinitely many times in sU, so when the differential substitution goes on U we 'extract' a linear copy of U, which receives the substitution, and we keep the other infinitely many unchanged. This will be much more evident in the definition of the analogous operation for the resource calculus (*cf.* Definition 7.3).

Example 3.7. Recall that the simple terms Δ and $\mathbf{I}$ were defined following Definition 3.2. We have

(1) $$\frac{\partial \Delta}{\partial x} \cdot \mathbf{I} = 0 \qquad \text{(since } x \text{ does not occur free in } \Delta\text{)}$$

(2) $$\frac{\partial x}{\partial x} \cdot \mathbf{I} = \mathbf{I}$$

(3) $$\frac{\partial (xx)}{\partial x} \cdot \mathbf{I} = \mathbf{I}x + (\mathsf{D}x \cdot \mathbf{I})x$$

(4) $$\frac{\partial}{\partial x}\left(\frac{\partial (xx)}{\partial x} \cdot \mathbf{I}\right) \cdot \Delta = (\mathsf{D}\mathbf{I} \cdot \Delta)x + (\mathsf{D}\Delta \cdot \mathbf{I})x + (\mathsf{D}(\mathsf{D}x \cdot \mathbf{I}) \cdot \Delta)x$$

(5) $$((\mathsf{D}x \cdot x)x)\,\{\mathbf{I}/x\} = (\mathsf{D}\mathbf{I} \cdot \mathbf{I})\mathbf{I}.$$

The differential substitution

$$\frac{\partial S}{\partial x} \cdot T$$

can be thought of as the differential of S with respect to the variable x, linearly applied to T. This may be inferred from the rule for linear application, which relates to the rule for composition of the differential. Moreover, it is easy to check that if $x \notin \mathrm{FV}(S)$ (that is, S is constant with respect to x), then

$$\frac{\partial S}{\partial x} \cdot T = 0.$$

This intuition is also reinforced by the validity of the Schwarz Theorem.

Theorem 3.8 (Schwarz Theorem). Let S, T, U be differential λ-terms. Let x and y be variables such that $x \notin \mathrm{FV}(U)$. Then

$$\frac{\partial}{\partial y}\Big(\frac{\partial S}{\partial x} \cdot T\Big) \cdot U = \frac{\partial}{\partial x}\Big(\frac{\partial S}{\partial y} \cdot U\Big) \cdot T + \frac{\partial S}{\partial x} \cdot \Big(\frac{\partial T}{\partial y} \cdot U\Big).$$

In particular, when $y \notin \mathrm{FV}(T)$, the second summand is 0 and the two differential substitutions commute.

Proof. The proof is by structural induction on S. Here we will just check the case $S \equiv vV$.

$$\begin{aligned}
\frac{\partial}{\partial y}\left(\frac{\partial vV}{\partial x}\cdot T\right)\cdot U =& \frac{\partial}{\partial y}\left(\left(\frac{\partial v}{\partial x}\cdot T\right)V + \left(\mathsf{D}v\cdot\left(\frac{\partial V}{\partial x}\cdot T\right)\right)V\right)\cdot U \\
=& \left(\frac{\partial}{\partial y}\left(\frac{\partial v}{\partial x}\cdot T\right)\cdot U\right)V \\
&+ \left(\mathsf{D}\left(\frac{\partial v}{\partial x}\cdot T\right)\cdot\left(\frac{\partial V}{\partial y}\cdot U\right)\right)V \\
&+ \left(\mathsf{D}\left(\frac{\partial v}{\partial y}\cdot U\right)\cdot\left(\frac{\partial V}{\partial x}\cdot T\right)\right)V \\
&+ \left(\mathsf{D}v\cdot\left(\frac{\partial}{\partial y}\left(\frac{\partial V}{\partial x}\cdot T\right)\cdot U\right)\right)V \\
&+ \left(\mathsf{D}\left(\mathsf{D}v\cdot\left(\frac{\partial V}{\partial x}\cdot T\right)\cdot\left(\frac{\partial V}{\partial y}\cdot U\right)\right)\right)V
\end{aligned}$$

Applying the induction hypothesis (and the permutative equality), we then get

$$\begin{aligned}
\frac{\partial}{\partial y}\left(\frac{\partial vV}{\partial x}\cdot T\right)\cdot U =& \left(\frac{\partial}{\partial x}\left(\frac{\partial v}{\partial y}\cdot U\right)\cdot T\right)V \\
&+ \left(\mathsf{D}\left(\frac{\partial v}{\partial y}\cdot U\right)\cdot\left(\frac{\partial V}{\partial x}\cdot T\right)\right)V \\
&+ \left(\mathsf{D}\left(\frac{\partial v}{\partial x}\cdot T\right)\cdot\left(\frac{\partial V}{\partial y}\cdot U\right)\right)V \\
&+ \left(\mathsf{D}v\cdot\left(\frac{\partial}{\partial x}\left(\frac{\partial V}{\partial y}\cdot U\right)\cdot T\right)\right)V \\
&+ \left(\mathsf{D}\left(\mathsf{D}v\cdot\left(\frac{\partial V}{\partial y}\cdot U\right)\right)\cdot\left(\frac{\partial V}{\partial x}\cdot T\right)\right)V \\
&+ \left(\frac{\partial v}{\partial x}\cdot\left(\frac{\partial T}{\partial y}\cdot U\right)\right)V \\
&+ \left(\mathsf{D}v\cdot\left(\frac{\partial V}{\partial x}\cdot\left(\frac{\partial T}{\partial y}\cdot U\right)\right)\right)V \\
=& \frac{\partial}{\partial x}\left(\left(\frac{\partial v}{\partial y}\cdot U\right)V + \left(\mathsf{D}v\cdot\left(\frac{\partial V}{\partial y}\cdot U\right)\right)V\right)\cdot T \\
&+ \left(\frac{\partial v}{\partial x}\cdot\left(\frac{\partial T}{\partial y}\cdot U\right)\right)V + \left(\mathsf{D}v\cdot\left(\frac{\partial V}{\partial x}\cdot\left(\frac{\partial T}{\partial y}\cdot U\right)\right)\right)V \\
=& \frac{\partial}{\partial x}\left(\frac{\partial vV}{\partial y}\cdot U\right)\cdot T + \frac{\partial vV}{\partial x}\cdot\left(\frac{\partial T}{\partial y}\cdot U\right).
\end{aligned}$$

□

For readability, we will sometimes adopt the following notation for multiple differential substitutions.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Notation 3.9. We set

$$\frac{\partial^n S}{\partial x_1,\dots,x_n}\cdot(t_1,\dots,t_n)=\frac{\partial}{\partial x_n}\left(\cdots\frac{\partial S}{\partial x_1}\cdot t_1\cdots\right)\cdot t_n$$

where $x_i \notin \mathrm{FV}(t_1,\dots,t_n)$ for all $1\leqslant i\leqslant n$.

Remark 3.10. From Theorem 3.8 we have

$$\frac{\partial^n S}{\partial x_1,\dots,x_n}\cdot(t_1,\dots,t_n)=\frac{\partial^n S}{\partial x_{\sigma(1)},\dots,x_{\sigma(n)}}\cdot(t_{\sigma(1)},\dots,t_{\sigma(n)}),\ \text{for all } \sigma\in\mathfrak{S}_n.$$

3.3. *Differential lambda theories*

In this section we introduce the axioms associated with the differential λ-calculus and define the equational theories of this calculus, namely, the *differential λ-theories*.

The axioms of the *differential λ-calculus* are (for all $s,t\in\Lambda^s$ and $T\in\Lambda^d$):

$$(\beta)\qquad (\lambda x.s)T = s\{T/x\}$$

$$(\beta_D)\qquad \mathsf{D}(\lambda x.s)\cdot t=\lambda x.\frac{\partial s}{\partial x}\cdot t.$$

Once oriented from left to right, the (β)-conversion shows how to calculate a function $\lambda x.s$ classically applied to an argument T, while the (β_D)-conversion shows how to evaluate a function $\lambda x.s$ *linearly* applied to a simple argument t.

Note that the λx does not disappear in the result of a linear application. This is because the simple term s may still contain free occurrences of x. The only way to get rid of the outer lambda abstraction in the term $\lambda x.s$ is to apply it classically to a term T, and then use the (β)-rule; when $x\notin\mathrm{FV}(s)$, a standard choice for T is 0.

The differential λ-calculus is an intensional language – there are syntactically different programs having the same extensional behaviour. We will sometimes be interested in the extensional version of this calculus, which is obtained by adding the following axiom (for every $s\in\Lambda^s$):

$$(\eta)\qquad \lambda x.sx=s,\ \text{where } x\notin\mathrm{FV}(s).$$

In the differential λ-calculus, we have another extensionality axiom, which is strictly weaker than (η), that can be safely added to the system, namely (for every $s,t\in\Lambda^s$):

$$(\eta_\partial)\qquad \lambda x.(\mathsf{D}s\cdot t)x=\mathsf{D}s\cdot t,\ \text{where } x\notin\mathrm{FV}(s,t).$$

The axiom (η_∂) states that the calculus is only extensional in its differential part, that is, in the presence of the linear application. Intuitively, this means that $\mathsf{D}s\cdot t$ must have a functional behaviour, which is always true in a simply typed setting where $s:A\to B$, $t:A$ and $\mathsf{D}s\cdot t:A\to B$. Interestingly enough, there are some very natural models of untyped differential λ-calculus that satisfy (η_∂) but do not satisfy (η) – see Section 6.3.1 for an example.

A *λ^d-relation* $\mathcal{T}$ is any set of equations between differential λ-terms (which can be thought of as a relation on $\Lambda^d\times\Lambda^d$).

A λ^d-relation $\mathcal{T}$ is said to be:

— an *equivalence* if it is closed under the following rules (for all $S, T, U \in \Lambda^d$):

$$\frac{}{S = S}\ \text{(reflexivity)}$$

$$\frac{T = S}{S = T}\ \text{(symmetry)}$$

$$\frac{S = T \quad T = U}{S = U}\ \text{(transitivity)}$$

— *compatible* if it is closed under the following rules (for all $S, T, U, V, S_i, T_i \in \Lambda^d$):

$$\frac{S = T}{\lambda x.S = \lambda x.T}\ \text{(lambda)}$$

$$\frac{S = T \quad U = V}{ST = UV}\ \text{(app)}$$

$$\frac{S = T \quad U = V}{\mathsf{D}S \cdot U = \mathsf{D}T \cdot V}\ \text{(Lapp)}$$

$$\frac{S_i = T_i \quad \text{for all } 1 \leqslant i \leqslant n}{\sum_{i=1}^{n} S_i = \sum_{i=1}^{n} T_i}\ \text{(sum).}$$

As notation, we will write $\mathcal{T} \vdash S = T$ or $S =_{\mathcal{T}} T$ for $S = T \in \mathcal{T}$.

Definition 3.11. A *differential λ-theory* is any compatible λ^d-relation $\mathcal{T}$ that is an equivalence relation and includes (β) and (β_D). A differential λ-theory $\mathcal{T}$ is said to be *differentially extensional* if it contains (η_∂) and *extensional* if it also contains (η). We say that $\mathcal{T}$ *satisfies sum idempotency* whenever $\mathcal{T} \vdash s + s = s$.

The differential λ-theories are naturally ordered by set-theoretical inclusion. We use $\lambda\beta^d$ to denote the minimum differential λ-theory, and $\lambda\beta\eta_\partial^d$ to denote the minimum differentially extensional differential λ-theory, and $\lambda\beta\eta^d$ to denote the minimum extensional differential λ-theory.

We will now give easy examples of equalities between differential λ-terms in $\lambda\beta^d, \lambda\beta\eta_\partial^d$ and $\lambda\beta\eta^d$ to assist in gaining familiarity with the operations in the calculus.

Example 3.12. Recall that $\Delta \equiv \lambda x.xx$. Then

(1) $\lambda\beta^d \vdash \quad (\mathsf{D}\Delta \cdot y)z = yz + (\mathsf{D}z \cdot y)z$

(2) $\lambda\beta^d \vdash \quad (\mathsf{D}^2\Delta \cdot (x, y))0 = (\mathsf{D}x \cdot y)0 + (\mathsf{D}y \cdot x)0$

(3) $\lambda\beta^d \vdash \quad \mathsf{D}^3\Delta \cdot (x, y, z) = \lambda r.(\mathsf{D}^2 x \cdot (y, z) + \mathsf{D}^2 y \cdot (x, z) + \mathsf{D}^2 z \cdot (x, y) + \mathsf{D}^3 r \cdot (x, y, z))r$

(4) $\lambda\beta\eta_\partial^d \vdash \quad \mathsf{D}(\lambda z.xz) \cdot y = \lambda z.(\mathsf{D}x \cdot y)z = \mathsf{D}x \cdot y$

(5) $\lambda\beta\eta^d \vdash \quad \mathsf{D}\Delta \cdot z = \lambda x.zx + \lambda x.(\mathsf{D}x \cdot z)x = z + \lambda x.(\mathsf{D}x \cdot z)x.$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Note that in this calculus (as in standard λ-calculus extended with non-deterministic choice (Dezani-Ciancaglini *et al.* 1996)), a single simple term can generate an infinite sum of terms, as in the following example.

Example 3.13. Recall (from the definitions following Definition 3.2) that $\mathbf{Y}$ is Curry's fixpoint combinator, $\underline{n}$ is the n-th Church numeral and $\mathbf{s}$ denotes the successor. Then:

$$\begin{array}{lll}(1) & \lambda\beta^d \vdash & \mathbf{Y}(x+y) = x(\mathbf{Y}(x+y)) + y(\mathbf{Y}(x+y)) \qquad \text{(for all variables } x, y) \\ (2) & \lambda\beta^d \vdash & \mathbf{Y}((\lambda z.\underline{0}) + \mathbf{s}) = \underline{0} \ + \ \mathbf{s}(\mathbf{Y}((\lambda z.\underline{0}) + \mathbf{s})) \\ & & \qquad = \underline{0} \ + \ \underline{1} \ + \ \mathbf{s}(\mathbf{s}(\mathbf{Y}((\lambda z.\underline{0}) + \mathbf{s}))) \\ & & \qquad = \cdots \end{array}$$

3.4. *A theory of Taylor expansion*

One of the most interesting consequences of adding a syntactical differential operator to the λ-calculus is that, in the presence of infinite sums, we can define the Taylor expansion of a program. Such an expansion is classically defined in the literature only for ordinary λ-terms (Ehrhard and Regnier 2003; Ehrhard and Regnier 2006a; Ehrhard and Regnier 2008). In this section we will generalise this notion to differential λ-terms. To avoid the annoying problem of handling coefficients, we will consider an idempotent sum.

Definition 3.14. Given a differential λ-term S, we define its (full) *Taylor expansion* S^* by induction on S as follows:

$$\begin{aligned} x^* &= x \\ (\lambda x.s)^* &= \lambda x.s^* \\ (\mathsf{D}^k s \cdot (t_1, \ldots, t_k))^* &= \mathsf{D}^k s^* \cdot \left(t_1^*, \ldots, t_k^*\right) \\ (sT)^* &= \Sigma_{k \in \mathbf{N}} (\mathsf{D}^k s^* \cdot \left(T^*, \ldots, T^*\right))0 \\ (s+T)^* &= s^* + T^*. \end{aligned}$$

Thus, the 'target language' of the Taylor expansion is much simpler than the full differential λ-calculus. For instance, the general application of the λ-calculus is no longer required, and we will only need iterated linear applications and ordinary applications to 0. We will, however, need countable sums, which are not, in general, present in the differential λ-calculus. From now on, the target calculus of the Taylor expansion will be denoted by Λ^d_∞.

We will write $\vec{S}$ to denote sequences of differential λ-terms $S_1, \ldots, S_k$ (with $k \geqslant 0$).

Remark 3.15. Every term $S \in \Lambda^d_\infty$ can be written as a (possibly infinite) sum of terms of the form

$$\lambda\vec{y}.(\mathsf{D}^{n_1}(\cdots(\mathsf{D}^{n_k} s \cdot (\vec{t}_k))\vec{0}) \cdots \cdot (\vec{t}_1))\vec{0}$$

where $\vec{t}_i$ is a sequence of simple terms of length $n_i \in \mathbf{N}$ (for $1 \leqslant i \leqslant k$) and the simple term s is either a variable or a lambda abstraction.

We will now try to clarify what it means for two differential λ-terms S and T to 'have the same Taylor expansion'. Indeed, we may have that $S^* = \Sigma_{i\in I} s_i$ and $T^* = \Sigma_{j\in J} t_j$ where I, J are countable sets. In this case one might be tempted to define $S^* = T^*$ by asking for the existence of a bijective correspondence between I and J such that each s_i is $\lambda\beta^d$-equivalent to some t_j. However, in the general case, this definition does not capture the equivalence between infinite sums that we have in mind. For instance, $S^* = T^*$ might hold because there are partitions $\{I_k\}_{k\in K}$ and $\{J_k\}_{k\in K}$ of I and J, respectively, such that for every $k \in K$, the sets I_k, J_k are finite and $\Sigma_{i\in I_k} s_i =_{\lambda\beta^d} \Sigma_{i\in J_k} s_j$. The naive definition works well when all summands of the two sums we are equating are 'in normal form'. Since the Λ^d_∞ calculus (essentially) enjoys strong normalisation, we can define the normal form of every $S \in \Lambda^d_\infty$ as follows.

Definition 3.16. Given $S \in \Lambda^d_\infty$, we define the *normal form of* S as follows.

— If $S \equiv \sum_{i\in I} s_i$, we set

$$\mathrm{NF}(S) = \sum_{i\in I} \mathrm{NF}(s_i).$$

— If $S \equiv \lambda\vec{y}.(\mathsf{D}^{n_1}(\cdots(\mathsf{D}^{n_k} x \cdot (\vec{t}_k))\vec{0})\cdots\cdot(\vec{t}_1))\vec{0}$, we set

$$\mathrm{NF}(S) = \lambda\vec{y}.(\mathsf{D}^{n_1}(\cdots(\mathsf{D}^{n_k} x \cdot (\mathrm{NF}(\vec{t}_k)))\vec{0})\cdots\cdot(\mathrm{NF}(\vec{t}_1)))\vec{0}.$$

— If $S \equiv \lambda\vec{y}.(\mathsf{D}^{n_1}(\cdots(\mathsf{D}^{n_k}(\lambda x.s)\cdot(\vec{t}_k))\vec{0})\cdots\cdot(\vec{t}_1))\vec{0}$ with $n_k > 0$, we set

$$\mathrm{NF}(S) =$$
$$\mathrm{NF}\left(\lambda\vec{y}.\left(\mathsf{D}^{n_1}\left(\cdots\left(\mathsf{D}^{n_{k-1}}\left(\left(\lambda x.\frac{\partial^{n_k} s}{\partial x,\ldots,x}\cdot(\vec{t}_k)\right)\vec{0}\right)\cdot(\vec{t}_{k-1})\right)\vec{0}\right)\cdots\cdot(\vec{t}_1)\right)\vec{0}\right).$$

— If $S \equiv \lambda\vec{y}.(\mathsf{D}^{n_1}(\cdots(\mathsf{D}^{n_k}((\lambda x.s)0\vec{0})\cdot(\vec{t}_k))\vec{0})\cdots\cdot(\vec{t}_1))\vec{0}$, we set

$$\mathrm{NF}(S) = \mathrm{NF}\big(\lambda\vec{y}.(\mathsf{D}^{n_1}(\cdots(\mathsf{D}^{n_k}((s\{0/x\})\vec{0})\cdot(\vec{t}_k))\vec{0})\cdots\cdot(\vec{t}_1))\vec{0}\big).$$

By Remark 3.15, this definition covers all possible cases.

We are now able to define the differential λ-theory generated by equating all differential λ-terms having the same Taylor expansion.

Definition 3.17. Given $S, T \in \Lambda^d$, we say that $\mathrm{NF}(S^*) = \mathrm{NF}(T^*)$ whenever

$$\mathrm{NF}(S^*) = \sum_{i\in I} s_i$$
$$\mathrm{NF}(T^*) = \sum_{j\in J} t_j$$

and there is an isomorphism $\imath : I \to J$ such that $\lambda\beta^d \vdash s_i = t_{\imath(i)}$. We set

$$\mathcal{E} = \{(S,T) \in \Lambda^d \times \Lambda^d \mid \mathrm{NF}(S^*) = \mathrm{NF}(T^*)\}.$$

It is not difficult to check that $\mathcal{E}$ is actually a differential λ-theory.

Two standard λ-terms s, t have the same Böhm tree (Barendregt 1984, Chapter 10) if and only if $\mathcal{E} \vdash s = t$ holds. The 'if' part of this equivalence is fairly straightforward, and the ‘only if’ part is proved in Ehrhard and Regnier (2006a). Thus, the theory $\mathcal{E}$ can be seen as an extension of the theory of Böhm trees in the context of differential λ-calculus.

4. A differential model theory

In this section we will provide the categorical framework in which the models of the differential λ-calculus live, namely, the *Cartesian closed differential categories*†. The material presented in Section 4.1 is mainly borrowed from (Blute *et al.* 2009).

4.1. *Cartesian differential categories*

Differential λ-terms will be interpreted as morphisms in a suitable category $\mathbf{C}$. Since in the syntax we have sums of terms, we need a sum on the morphisms of $\mathbf{C}$ satisfying the equations introduced in Notation 3.4. For this reason, we will focus our attention on left-additive categories.

A category $\mathbf{C}$ is *left-additive* whenever each homset has the structure of a commutative monoid $(\mathbf{C}(A,B), +_{AB}, 0_{AB})$ and $(g+h)\circ f = (g\circ f)+(h\circ f)$ and $0\circ f = 0$.

Definition 4.1. A morphism f in $\mathbf{C}$ is said to be *additive* if, in addition, it satisfies

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

A category is *Cartesian left-additive* if it is a left-additive category with products such that all projections and pairings of additive maps are additive.

Definition 4.2. A *Cartesian differential category* is a Cartesian left-additive category having an operator $D(-)$ that maps every morphism $f : A\to B$ into a morphism $D(f) : A\times A\to B$ and satisfies the following axioms:

(D1) $$D(f+g) = D(f)+D(g)$$
$$D(0) = 0$$

(D2) $$D(f)\circ\langle h+k, v\rangle = D(f)\circ\langle h, v\rangle + D(f)\circ\langle k, v\rangle$$
$$D(f)\circ\langle 0, v\rangle = 0$$

(D3) $$D(\mathrm{Id}) = \pi_1$$
$$D(\pi_1) = \pi_1\circ\pi_1$$
$$D(\pi_2) = \pi_2\circ\pi_1$$

(D4) $$D(\langle f, g\rangle) = \langle D(f), D(g)\rangle$$

(D5) $$D(f\circ g) = D(f)\circ\langle D(g), g\circ\pi_2\rangle$$

† These categories were introduced in Bucciarelli *et al.* (2010), where they were called *differential λ-categories* and proposed as models of the simply typed differential λ-calculus and simply typed resource calculus.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

(D6) $D(D(f))\circ\langle\langle g,0\rangle,\langle h,k\rangle\rangle = D(f)\circ\langle g,k\rangle$

(D7) $D(D(f))\circ\langle\langle 0,h\rangle,\langle g,k\rangle\rangle = D(D(f))\circ\langle\langle 0,g\rangle,\langle h,k\rangle\rangle\,.$

We will now provide some intuitions for these axioms:

(D1) says that the operator $D(-)$ is linear;
(D2) says that $D(-)$ is additive in its first coordinate;
(D3) and (D4) say that $D(-)$ behaves coherently with the product structure;
(D5) is the usual chain rule;
(D6) requires that $D(f)$ is linear in its first component.
(D7) states the independence of the order of 'partial differentiation'.

Remark 4.3. In a Cartesian differential category we obtain partial derivatives from the full ones by 'zeroing out' the components for which the differentiation is not required. For example, if we want to define the partial derivative $D_1(f)$ of $f : C\times A\to B$ on its first component, it is sufficient to set

$$D_1(f) = D(f)\circ(\langle \mathrm{Id}_C, 0_A\rangle\times \mathrm{Id}_{C\times A}) : C\times(C\times A)\to B.$$

Similarly, we define

$$D_2(f) = D(f)\circ(\langle 0_C, \mathrm{Id}_A\rangle\times \mathrm{Id}_{C\times A}) : A\times(C\times A)\to B,$$

as the partial derivative of f with respect to its second component.

This remark follows from the fact that every differential $D(f)$ can be reconstructed from its partial derivatives as follows:

$$\begin{aligned}
D(f) &= D(f)\circ\langle\langle \pi_1\circ\pi_1, \pi_2\circ\pi_1\rangle,\pi_2\rangle\\
&= D(f)\circ\langle\langle \pi_1\circ\pi_1, 0\rangle,\pi_2\rangle + D(f)\circ\langle\langle 0,\pi_2\circ\pi_1\rangle,\pi_2\rangle\\
&= D(f)\circ(\langle \mathrm{Id},0\rangle\times\mathrm{Id})\circ(\pi_1\times\mathrm{Id}) + D(f)\circ(\langle 0,\mathrm{Id}\rangle\times\mathrm{Id})\circ(\pi_2\times\mathrm{Id})\\
&= D_1(f)\circ(\pi_1\times\mathrm{Id}) + D_2(f)\circ(\pi_2\times\mathrm{Id}).
\end{aligned}$$

4.2. *Linear morphisms*

In Cartesian differential categories we are able to express the fact that a morphism is 'linear' by requiring that its differential is constant.

Definition 4.4. In a Cartesian differential category, a morphism $f : A\to B$ is said to be *linear* if $D(f) = f\circ\pi_1$.

Lemma 4.1. Every linear morphism $f : A\to B$ is additive.

Proof. By the definition of a linear morphism, we have $D(f) = f\circ\pi_1$. For all $g, h : C \to A$ we have

$$\begin{aligned} f \circ (g + h) &= f \circ \pi_1 \circ \langle g + h, g\rangle \\ &= D(f) \circ \langle g + h, g\rangle \\ &= D(f) \circ \langle g, g\rangle + D(f) \circ \langle h, g\rangle \\ &= f \circ \pi_1 \circ \langle g, g\rangle + f \circ \pi_1 \circ \langle h, g\rangle \\ &= f \circ g + f \circ h. \end{aligned}$$

Moreover,

$$\begin{aligned} f \circ 0 &= f \circ \pi_1 \circ \langle 0, 0\rangle \\ &= D(f) \circ \langle 0, 0\rangle \\ &= 0, \end{aligned}$$

so we can conclude that f is additive. □

Lemma 4.2. The composition of two linear morphisms is linear.

Proof. Let f, g be two linear maps. We have to prove that $D(f \circ g) = f \circ g \circ \pi_1$. By *(D5)* we have $D(f \circ g) = D(f) \circ \langle D(g), g \circ \pi_2\rangle$. Since f, g are linear we have

$$\begin{aligned} D(f) \circ \langle D(g), g \circ \pi_2\rangle &= f \circ \pi_1 \circ \langle g \circ \pi_1, g \circ \pi_2\rangle \\ &= f \circ g \circ \pi_1. \end{aligned}$$ □

Thus, in fact, every Cartesian differential category has a subcategory of linear maps.

4.3. *Cartesian closed differential categories*

Cartesian differential categories are not enough to interpret the differential λ-calculus since the differential operator does not automatically behave well with respect to the Cartesian closed structure. For this reason, we will now introduce the notion of a *Cartesian closed differential category*.

Definition 4.5. A category is *Cartesian closed left-additive* if it is a Cartesian left-additive category that is Cartesian closed and satisfies

(+-curry) $\qquad \Lambda(f + g) = \Lambda(f) + \Lambda(g)$

(0-curry) $\qquad \Lambda(0) = 0.$

Using these properties of $\Lambda(-)$, it is easy to prove that the evaluation morphism is additive in its left component.

Lemma 4.3. In every Cartesian closed left-additive category the following axioms hold (for all $f, g : C \to (A \Rightarrow B)$ and $h : C \to A$):

(+-eval) $\qquad \mathrm{ev} \circ \langle f + g, h\rangle = \mathrm{ev} \circ \langle f, h\rangle + \mathrm{ev} \circ \langle g, h\rangle$

(0-eval) $\qquad \mathrm{ev} \circ \langle 0, h\rangle = 0.$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Proof. Let $f' = \Lambda^-(f)$ and $g' = \Lambda^-(g)$. Then

$$
\begin{aligned}
\mathrm{ev}\circ\langle f+g,h\rangle &= \mathrm{ev}\circ((\Lambda(f')+\Lambda(g'))\times \mathrm{Id})\circ\langle \mathrm{Id},h\rangle && \text{(definition of } f',g'\text{)}\\
&= \Lambda^-((\Lambda(f')+\Lambda(g'))\circ\langle \mathrm{Id},h\rangle && \text{(definition of } \Lambda^-\text{)}\\
&= \Lambda^-(\Lambda(f'+g'))\circ\langle \mathrm{Id},h\rangle && \text{(+-curry)}\\
&= (f'+g')\circ\langle \mathrm{Id},h\rangle && \text{(definition of } \Lambda^-\text{)}\\
&= f'\circ\langle \mathrm{Id},h\rangle + g'\circ\langle \mathrm{Id},h\rangle && \text{(left-additivity)}\\
&= \Lambda^-(f)\circ\langle \mathrm{Id},h\rangle + \Lambda^-(g)\circ\langle \mathrm{Id},h\rangle && \text{(definition of } f',g'\text{)}\\
&= \mathrm{ev}\circ(f\times \mathrm{Id})\circ\langle \mathrm{Id},h\rangle + \mathrm{ev}\circ(g\times \mathrm{Id})\circ\langle \mathrm{Id},h\rangle && \text{(definition of } \Lambda^-\text{)}\\
&= \mathrm{ev}\circ\langle f,h\rangle + \mathrm{ev}\circ\langle g,h\rangle\,.
\end{aligned}
$$

Moreover, $\mathrm{ev}\circ\langle 0,g\rangle = \mathrm{ev}\circ\langle \Lambda(0),g\rangle = 0\circ\langle \mathrm{Id},g\rangle = 0$. □

Definition 4.6. A *Cartesian closed differential category* is a Cartesian differential category that is Cartesian closed left-additive and such that, for all $f : C\times A\to B$:

(D-curry) $$D(\Lambda(f)) = \Lambda(D(f)\circ\langle \pi_1\times 0_A, \pi_2\times \mathrm{Id}_A\rangle).$$

Indeed, in a Cartesian closed differential category there are two ways to differentiate $f : C\times A\to B$ with respect to its first component: we can use the trick of Remark 4.3; or we can 'hide' the component A by currying f and then differentiate $\Lambda(f)$. Intuitively, (D-curry) requires that these two methods are equivalent.

Lemma 4.4. In every Cartesian closed differential category, the following axiom holds (for all $h : C\to(A\Rightarrow B)$ and $g : C\to A$):

(D-eval) $$D(\mathrm{ev}\circ\langle h,g\rangle) = \mathrm{ev}\circ\langle D(h), g\circ\pi_2\rangle + D(\Lambda^-(h))\circ\langle\langle 0_C, D(g)\rangle, \langle\pi_2, g\circ\pi_2\rangle\rangle\,.$$

Proof. Let $h' = \Lambda^-(h) : C\times A\to B$. Then

$$
\begin{aligned}
D(\mathrm{ev}\circ\langle h,g\rangle) &= D(\mathrm{ev}\circ\langle \Lambda(h'),g\rangle) && \text{(definition of } h'\text{)}\\
&= D(h'\circ\langle \mathrm{Id}_C,g\rangle) && \text{(beta-cat)}\\
&= D(h')\circ\langle D(\langle \mathrm{Id}_C,g\rangle), \langle \mathrm{Id}_C,g\rangle\circ\pi_2\rangle && \text{(D5)}\\
&= D(h')\circ\langle\langle \pi_1, D(g)\rangle, \langle\pi_2, g\circ\pi_2\rangle\rangle && \text{(D4+D3)}\\
&= D(h')\circ\langle\langle \pi_1, 0_A\rangle + \langle 0_C, D(g)\rangle, \langle\pi_2, g\circ\pi_2\rangle\rangle && \text{(pairing is additive)}\\
&= D(h')\circ\langle\langle \pi_1, 0_A\rangle, \langle\pi_2, g\circ\pi_2\rangle\rangle\\
&\quad + D(h')\circ\langle\langle 0_C, D(g)\rangle, \langle\pi_2, g\circ\pi_2\rangle\rangle && \text{(D2)}\\
&= D(h')\circ\langle \pi_1\times 0_A, \pi_2\times \mathrm{Id}_A\rangle\circ\langle \mathrm{Id}_{C\times C}, g\circ\pi_2\rangle\\
&\quad + D(h')\circ\langle\langle 0_C, D(g)\rangle, \langle\pi_2, g\circ\pi_2\rangle\rangle\\
&= \mathrm{ev}\circ\langle \Lambda(D(h')\circ\langle \pi_1\times 0_A, \pi_2\times \mathrm{Id}_A\rangle), g\circ\pi_2\rangle\\
&\quad + D(h')\circ\langle\langle 0_C, D(g)\rangle, \langle\pi_2, g\circ\pi_2\rangle\rangle && \text{(beta-cat)}
\end{aligned}
$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

$$
\begin{aligned}
&= \mathrm{ev} \circ \langle D(\Lambda(h')), g \circ \pi_2 \rangle \\
&\quad + D(\Lambda^-(\Lambda(h'))) \circ \langle \langle 0_C, D(g) \rangle, \langle \pi_2, g \circ \pi_2 \rangle \rangle && \text{(D-curry)} \\
&= \mathrm{ev} \circ \langle D(h), g \circ \pi_2 \rangle \\
&\quad + D(\Lambda^-(h)) \circ \langle \langle 0_C, D(g) \rangle, \langle \pi_2, g \circ \pi_2 \rangle \rangle \,. && \text{(definition of } h')
\end{aligned}
$$

□

The axiom (D-eval) can be viewed as a chain rule for denotations of differential λ-terms (*cf.* Lemma 4.8(i)).

In Cartesian closed differential categories, we are able to define a binary operator $\star$ on morphisms, which can be viewed as the semantic counterpart of differential substitution. The idea behind $f \star g$ is to derive the map $f : A \to B$ and then apply the argument $g : A$ in its linear component. However, differential λ-terms are interpreted *in a certain context*, so we need to handle the context C and consider maps $f : C \times A \to A$ and $g : C \to A$.

Definition 4.7. The operator

$$
\frac{f : C \times A \to B \quad g : C \to A}{f \star g : C \times A \to B} \; (\star)
$$

is defined by $f \star g = D(f) \circ \big\langle \langle 0_C^{C \times A}, g \circ \pi_1 \rangle, \mathrm{Id}_{C \times A} \big\rangle$.

The morphism $f \star g$ is obtained by differentiating f in its second component (partial differentiation), and applying g in that component. The precise correspondence between $\star$ and the differential substitution is given in Theorem 5.7.

Remark 4.8. In fact, the operators $D(-)$ and $\star$ are mutually definable. To define $D(-)$ in terms of $\star$, we just set $D(f) = (f \circ \pi_2) \star \mathrm{Id}$. To check that this definition is meaningful, we show that it holds in every Cartesian differential category: indeed, by Definition 4.7,

$$
\begin{aligned}
(f \circ \pi_2) \star \mathrm{Id} &= D(f \circ \pi_2) \circ \langle \langle 0, \pi_1 \rangle, \mathrm{Id} \rangle \\
&= D(f) \circ \langle \pi_2 \circ \pi_1, \pi_2 \circ \pi_2 \rangle \circ \langle \langle 0, \pi_1 \rangle, \mathrm{Id} \rangle \\
&= D(f).
\end{aligned}
$$

Thus, we could formulate the whole theory of Cartesian closed differential categories by axiomatising the behaviour of $\star$ instead of that of $D(-)$. In the current work we prefer to use $D(-)$ because it is a more basic operation, which has already been studied in the literature, and the complexities of the two approaches are comparable.

Linear morphisms can be characterised in terms of the operator $\star$ as follows.

Lemma 4.5. A morphism $f : A \to B$ is linear if and only if for all $g : C \to A$, we have

$$
(f \circ \pi_2) \star g = (f \circ g) \circ \pi_1 \, : C \times A \to B.
$$

Proof.

($\Rightarrow$) Suppose f is linear. By the definition of $\star$, we have

$$
(f \circ \pi_2) \star g = D(f \circ \pi_2) \circ \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{C \times A} \rangle \,.
$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

By applying (D5) and (D3), this is equal to

$$D(f) \circ \langle \pi_2 \circ \pi_1, \pi_2 \circ \pi_2 \rangle \circ \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{C \times A} \rangle = D(f) \circ \langle g \circ \pi_1, \pi_2 \rangle .$$

Since f is linear, we have $D(f) = f \circ \pi_1$, so

$$D(f) \circ \langle g \circ \pi_1, \pi_2 \rangle = f \circ g \circ \pi_1.$$

($\Leftarrow$) Suppose $(f \circ \pi_2) \star g = (f \circ g) \circ \pi_1$ for all $g : C \to A$. In particular, this is true for $C = A$ and $g = \mathrm{Id}_A$. Thus we have $(f \circ \pi_2) \star \mathrm{Id}_A = f \circ \pi_1$. We can now conclude since

$$\begin{aligned}
(f \circ \pi_2) \star \mathrm{Id}_A &= D(f \circ \pi_2) \circ \langle \langle 0_A, \pi_1 \rangle, \mathrm{Id}_{A \times A} \rangle && \text{(definition of } \star) \\
&= D(f) \circ \langle \pi_2 \circ \pi_1, \pi_2 \circ \pi_2 \rangle \circ \langle \langle 0_A, \pi_1 \rangle, \mathrm{Id}_{A \times A} \rangle && \text{(D5+D3)} \\
&= D(f) \circ \langle \pi_1, \pi_2 \rangle \\
&= D(f).
\end{aligned}$$

□

The operator $\star$ enjoys the following commutation property.

Lemma 4.6. Let $f : C \times A \to B$ and $g, h : C \to A$. Then $(f \star g) \star h = (f \star h) \star g$.

Proof. We set $\varphi_g = \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{C \times A} \rangle$ and $\varphi_h = \langle \langle 0_C, h \circ \pi_1 \rangle, \mathrm{Id}_{C \times A} \rangle$. We have

$$\begin{aligned}
&(f \star g) \star h \\
&\quad = D(D(f) \circ \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{C \times A} \rangle) \circ \varphi_h \\
&\quad = D(D(f)) \circ \langle D(\langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id} \rangle), \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id} \rangle \circ \pi_2 \rangle \circ \varphi_h && \text{(D5)} \\
&\quad = D(D(f)) \circ \langle \langle \langle 0_C, D(g \circ \pi_1) \rangle, \pi_1 \rangle, \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id} \rangle \circ \pi_2 \rangle \circ \varphi_h && \text{(D4)} \\
&\quad = D(D(f)) \circ \langle \langle \langle 0_C, D(g) \circ \langle \pi_1 \circ \pi_1, \pi_1 \circ \pi_2 \rangle \rangle, \pi_1 \rangle, \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id} \rangle \circ \pi_2 \rangle \circ \varphi_h && \text{(D5)} \\
&\quad = D(D(f)) \circ \langle \langle \langle 0_C, D(g) \circ \langle 0_C, \pi_1 \rangle \rangle, \langle 0_C, h \circ \pi_1 \rangle \rangle, \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id} \rangle \rangle \\
&\quad = D(D(f)) \circ \langle \langle 0_{C \times A}, \langle 0_C, h \circ \pi_1 \rangle \rangle, \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id} \rangle \rangle && \text{(D2)} \\
&\quad = D(D(f)) \circ \langle \langle 0_{C \times A}, \langle 0_C, g \circ \pi_1 \rangle \rangle, \langle \langle 0_C, h \circ \pi_1 \rangle, \mathrm{Id} \rangle \rangle && \text{(D7)} \\
&\quad = D(D(f)) \circ \langle \langle \langle 0_C, D(h) \circ \langle 0_C, \pi_1 \rangle \rangle, \langle 0_C, g \circ \pi_1 \rangle \rangle, \langle \langle 0_C, h \circ \pi_1 \rangle, \mathrm{Id} \rangle \rangle && \text{(D2)} \\
&\quad = D(D(f)) \circ \langle \langle \langle 0_C, D(h) \circ \langle \pi_1 \circ \pi_1, \pi_1 \circ \pi_2 \rangle \rangle, \pi_1 \rangle, \langle \langle 0_C, h \circ \pi_1 \rangle, \mathrm{Id} \rangle \circ \pi_2 \rangle \circ \varphi_g \\
&\quad = D(D(f)) \circ \langle \langle \langle 0_C, D(h \circ \pi_1) \rangle, \pi_1 \rangle, \langle \langle 0_C, h \circ \pi_1 \rangle, \mathrm{Id} \rangle \circ \pi_2 \rangle \circ \varphi_g && \text{(D5)} \\
&\quad = D(D(f)) \circ \langle D(\langle \langle 0_C, h \circ \pi_1 \rangle, \mathrm{Id}_{CA} \rangle), \langle \langle 0_C, h \circ \pi_1 \rangle, \mathrm{Id} \rangle \circ \pi_2 \rangle \circ \varphi_g && \text{(D4)} \\
&\quad = D(D(f) \circ \langle \langle 0_C, h \circ \pi_1 \rangle, \mathrm{Id} \rangle) \circ \varphi_g && \text{(D5)} \\
&\quad = (f \star h) \star g.
\end{aligned}$$

□

Definition 4.9. Let $\mathrm{sw}_{ABC} = \langle \langle \pi_1 \circ \pi_1, \pi_2 \rangle, \pi_2 \circ \pi_1 \rangle : (A \times B) \times C \to (A \times C) \times B$.

Remark 4.10. We have

$$\begin{aligned}
\mathrm{sw} \circ \mathrm{sw} &= \mathrm{Id}_{(A \times B) \times C} \\
\mathrm{sw} \circ \langle \langle f, g \rangle, h \rangle &= \langle \langle f, h \rangle, g \rangle \\
D(\mathrm{sw}) &= \mathrm{sw} \circ \pi_1.
\end{aligned}$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

The following two technical lemmas will be used in Section 5.3 to show the soundness of the categorical models of the differential λ-calculus – complete proofs are given in Appendix A.

Lemma 4.7. Let $f : (C \times A) \times D \to B$ and $g : C \to A$, $h : C \to B'$. Then:

(i) $\pi_2 \star g = g \circ \pi_1$.
(ii) $(h \circ \pi_1) \star g = 0$.
(iii) $\Lambda(f) \star g = \Lambda(((f \circ \mathrm{sw}) \star (g \circ \pi_1)) \circ \mathrm{sw})$.

Outline of proof.

(i) This part follows by applying (D3).
(ii) This part follows by applying (D2), (D3) and (D5).
(iii) This part follows from (Curry), (D-curry) and (D2), (D3) and (D5). □

Lemma 4.8. Let $f : C \times A \to (D \Rightarrow B)$ and $g : C \to A$, $h : C \times A \to D$. Then:

(i) $(\mathrm{ev} \circ \langle f, h \rangle) \star g = \mathrm{ev} \circ \langle f \star g + \Lambda(\Lambda^-(f) \star (h \star g)), h \rangle$.
(ii) $\Lambda(\Lambda^-(f) \star h) \star g = \Lambda(\Lambda^-(f \star g) \star h) + \Lambda(\Lambda^-(f) \star (h \star g))$.
(iii) $\Lambda(\Lambda^-(f) \star h) \circ \langle \mathrm{Id}_C, g \rangle = \Lambda(\Lambda^-(f \circ \langle \mathrm{Id}_C, g \rangle) \star (h \circ \langle \mathrm{Id}_C, g \rangle))$.

Outline of proof.

(i) This part follows by applying (D-eval) and (beta-cat).
(ii) This equation can be simplified by using the axioms of Cartesian closed left-additive categories. Indeed, the right-hand side can be written as

$$\Lambda((\Lambda^-(f \star g) \star h) + \Lambda^-(f) \star (h \star g)).$$

By taking a morphism f' such that $f = \Lambda(f')$ and applying Lemma 4.7 (iii), part (ii) becomes equivalent to

$$((f' \star h) \circ \mathrm{sw}) \star (g \circ \pi_1) \circ \mathrm{sw} = (((f' \circ \mathrm{sw}) \star (g \circ \pi_1)) \circ \mathrm{sw}) \star h + f' \star (h \star g),$$

which follows by (Curry) and (D2–7).
(iii) This part follows by (Curry) and (D2–5). □

5. Categorical models of the differential lambda calculus

We proved in Bucciarelli *et al.* (2010) that Cartesian closed differential categories constitute *sound* models of the simply typed differential λ-calculus. In this section we will show that all reflexive objects living in these categories and satisfying a linearity condition are sound models of the *untyped* version of this calculus.

5.1. *Linear reflexive objects in Cartesian closed differential categories*

In a category **C**, an object A is a *retract* of an object B, written $A \lhd B$, if there are morphisms $f : A \to B$ and $g : B \to A$ such that $g \circ f = \mathrm{Id}_A$. When $f \circ g = \mathrm{Id}_B$ also holds, we say that A and B are *isomorphic*, written $A \cong B$, and that f, g are *isomorphisms*.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

In a Cartesian closed category **C**, we expect a *reflexive object* $\mathscr{U}$ to be a triple $(U, \mathcal{A}, \lambda)$ where U is an object of **C** and $\mathcal{A} : U \to (U \Rightarrow U)$ and $\lambda : (U \Rightarrow U) \to U$ are two morphisms performing the retraction $(U \Rightarrow U) \lhd U$. When $(U \Rightarrow U) \cong U$ we say that $\mathscr{U}$ is *extensional*.

Definition 5.1. A reflexive object $\mathscr{U} = (U, \mathcal{A}, \lambda)$ in a Cartesian closed differential category is *linear* if both $\mathcal{A}$ and λ are linear morphisms.

We are now able to provide our definition of a model of the untyped differential λ-calculus.

Definition 5.2. A *categorical model* $\mathscr{U}$ *of the differential λ-calculus* is a linear reflexive object in a Cartesian closed differential category. The model $\mathscr{U}$ is said to be *differentially extensional* (respectively, *extensional*) if its equational theory is.

It is routine to check that $\mathscr{U}$ is an extensional model (that is, $\mathrm{Th}(\mathscr{U}) \supseteq \lambda\beta\eta^d$) if and only if it is extensional as a reflexive object (that is, $(U \Rightarrow U) \cong U$).

The following lemma is useful for proving that a reflexive object in a Cartesian closed differential category is linear.

Lemma 5.1. Let $\mathscr{U}$ be a reflexive object.

(i) If $\mathcal{A}$ and $\lambda \circ \mathcal{A}$ are linear, then $\mathscr{U}$ is linear.
(ii) If $\mathscr{U}$ is extensional and either $\mathcal{A}$ or λ is linear, then $\mathscr{U}$ is linear.

Proof.

(i) Suppose $\mathcal{A}$ and $\lambda \circ \mathcal{A}$ are linear morphisms. We will show that λ is linear too. Indeed, we have

$$\begin{aligned}
D(\lambda) &= D(\lambda) \circ (\mathcal{A} \times \mathcal{A}) \circ (\lambda \times \lambda) \\
&= D(\lambda) \circ \langle \mathcal{A} \circ \pi_1, \mathcal{A} \circ \pi_2 \rangle \circ (\lambda \times \lambda) \\
&= D(\lambda) \circ \langle D(\mathcal{A}), \mathcal{A} \circ \pi_2 \rangle \circ (\lambda \times \lambda) && (\mathcal{A} \text{ linear}) \\
&= D(\lambda \circ \mathcal{A}) \circ (\lambda \times \lambda) \\
&= \lambda \circ \mathcal{A} \circ \pi_1 \circ \langle \lambda \circ \pi_1, \lambda \circ \pi_2 \rangle && (\lambda \circ \mathcal{A} \text{ linear}) \\
&= \lambda \circ \mathcal{A} \circ \lambda \circ \pi_1 \\
&= \lambda \circ \pi_1.
\end{aligned}$$

(ii) If $\mathcal{A}$ is linear, part (ii) follows directly from part (i) since $\lambda \circ \mathcal{A} = \mathrm{Id}_U$ and the identity is linear. If λ is linear, calculations analogous to those made for Part (i) show that $\mathcal{A}$ is too. □

Notice that, in general, there may be extensional reflexive objects that are not linear. However, in the concrete example of Cartesian closed differential category we will describe in Section 6, every extensional reflexive object will be linear (see Corollary 6.5).

Lemma 5.2. Let $\mathscr{U}$ be a linear reflexive object and let

$$\begin{aligned} f &: U^{n+1} \to (U \Rightarrow U) \\ h &: U^{n+1} \to U \\ g &: U^n \to U. \end{aligned}$$

Then:

(i) $\lambda \circ (f \star g) = (\lambda \circ f) \star g$.
(ii) $\mathcal{A} \circ (h \star g) = (\mathcal{A} \circ h) \star g$.

Proof.

(i) By the definition of $\star$, we have

$$(\lambda \circ f) \star g = D(\lambda \circ f) \circ \langle\langle 0_{U^n}, g \circ \pi_1\rangle, \mathrm{Id}_{U^{n+1}}\rangle\,.$$

By (D5), we have

$$D(\lambda \circ f) = D(\lambda) \circ \langle D(f), f \circ \pi_2\rangle\,.$$

Since λ is linear, we have $D(\lambda) = \lambda \circ \pi_1$, so

$$\begin{aligned} D(\lambda) \circ \langle D(f), f \circ \pi_2\rangle &= \lambda \circ \pi_1 \circ \langle D(f), f \circ \pi_2\rangle \\ &= \lambda \circ D(f). \end{aligned}$$

Hence,

$$\begin{aligned} D(\lambda \circ f) \circ \langle\langle 0_{U^n}, g \circ \pi_1\rangle, \mathrm{Id}_{U^{n+1}}\rangle &= \lambda \circ D(f) \circ \langle\langle 0_{U^n}, g \circ \pi_1\rangle, \mathrm{Id}_{U^{n+1}}\rangle \\ &= \lambda \circ (f \star g). \end{aligned}$$

(ii) The proof for this part is analogous to the proof of part (i). □

5.2. *Defining the interpretation*

Let $\vec{x} = x_1, \ldots, x_n$ be an ordered sequence of variables without repetitions. We say that $\vec{x}$ is *adequate* for $S_1, \ldots, S_k \in \Lambda^d$ if $\mathrm{FV}(S_1, \ldots, S_k) \subseteq \{x_1, \ldots, x_n\}$. Given an object U, we write $U^{\vec{x}}$ for the $\{x_1, \ldots, x_n\}$-indexed categorical product of n copies of U (when $n = 0$, we consider $U^{\vec{x}} = \mathbb{1}$). Moreover, we define the ith projection $\pi_i^{\vec{x}} : U^{\vec{x}} \to U$ by

$$\pi_i^{\vec{x}} = \begin{cases} \pi_2 & \text{if } i = n \\ \pi_i^{x_1,\ldots,x_{n-1}} \circ \pi_1 & \text{otherwise.} \end{cases}$$

Definition 5.3. Let $\mathscr{U}$ be a categorical model, S be a differential λ-term and $\vec{x} = x_1, \ldots, x_n$ be adequate for S. The *interpretation of S in $\mathscr{U}$ (with respect to $\vec{x}$)* will be a morphism

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

$[\![S]\!]_{\vec{x}} : U^{\vec{x}} \to U$ defined by induction as follows:

$$
\begin{aligned}
[\![x_i]\!]_{\vec{x}} &= \pi_i^{\vec{x}} \\
[\![sT]\!]_{\vec{x}} &= \mathrm{ev} \circ \langle \mathcal{A} \circ [\![s]\!]_{\vec{x}}, [\![T]\!]_{\vec{x}} \rangle \\
[\![\lambda z.s]\!]_{\vec{x}} &= \lambda \circ \Lambda([\![s]\!]_{\vec{x},z}) \qquad \text{(where by } \alpha\text{-conversion we assume that } z \text{ does not occur in } \vec{x}) \\
[\![\mathsf{D}^1 s \cdot (t)\,]\!]_{\vec{x}} &= \lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![s]\!]_{\vec{x}}) \star [\![t]\!]_{\vec{x}}) \\
[\![\mathsf{D}^{n+1} s \cdot (t_1, \dots, t_n, t_{n+1})\,]\!]_{\vec{x}} &= \lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![\mathsf{D}^n s \cdot (t_1, \dots, t_n)\,]\!]_{\vec{x}}) \star [\![t_{n+1}]\!]_{\vec{x}}) \\
[\![0]\!]_{\vec{x}} &= 0_U^{U^{\vec{x}}} \\
[\![s + S]\!]_{\vec{x}} &= [\![s]\!]_{\vec{x}} + [\![S]\!]_{\vec{x}}.
\end{aligned}
$$

Remark 5.4. Easy calculations give

$$[\![\mathsf{D}^n s \cdot (t_1, \dots, t_n)\,]\!]_{\vec{x}} = \lambda \circ \Lambda((\cdots (\Lambda^-(\mathcal{A} \circ [\![s]\!]_{\vec{x}}) \star [\![t_1]\!]_{\vec{x}}) \cdots) \star [\![t_n]\!]_{\vec{x}}).$$

Lemma 4.6 means that this interpretation does not depend on the chosen representative of the permutative equivalence class. In other words, we have

$$[\![\mathsf{D}^n s \cdot (t_1, \dots, t_n)\,]\!]_{\vec{x}} = [\![\mathsf{D}^n s \cdot \left(t_{\sigma(1)}, \dots, t_{\sigma(n)}\right)\,]\!]_{\vec{x}}$$

for every permutation $\sigma \in \mathfrak{S}_{\mathfrak{n}}$.

5.3. *Soundness*

Given a categorical model $\mathscr{U}$, we can define the *equational theory of* $\mathscr{U}$ as follows:

$$\mathrm{Th}(\mathscr{U}) = \{S = T \mid [\![S]\!]_{\vec{x}} = [\![T]\!]_{\vec{x}} \text{ for some } \vec{x} \text{ adequate for } S, T\}.$$

The aim of this section is to prove that the interpretation we have defined is *sound*, that is, that $\mathrm{Th}(\mathscr{U})$ is a differential λ-theory for every model $\mathscr{U}$.

The following convention allows us to simplify the statements of our theorems.

Convention 5.5. For the rest of this section we consider a fixed (but arbitrary) linear reflexive object $\mathscr{U}$ living in a Cartesian closed differential category **C**. Moreover, whenever we write $[\![S]\!]_{\vec{x}}$, we suppose that $\vec{x}$ is an adequate sequence for S.

The proof of the next lemma is easy, and is left as an exercise. Recall that the morphism sw was introduced in Definition 4.9.

Lemma 5.3. Let $S \in \Lambda^d$.

(i) If $z \notin \mathrm{FV}(S)$, then $[\![S]\!]_{\vec{x};z} = [\![S]\!]_{\vec{x}} \circ \pi_1$, where z does not occur in $\vec{x}$.
(ii) $[\![S]\!]_{\vec{x};y;z} = [\![S]\!]_{\vec{x};z;y} \circ \mathrm{sw}$, where z and y do not occur in $\vec{x}$.

Theorem 5.6 (Classic Substitution Theorem). Let $S, T \in \Lambda^d$, $\vec{x} = x_1, \dots, x_n$ and y not occurring in $\vec{x}$. Then

$$[\![S\,\{T/y\}\,]\!]_{\vec{x}} = [\![S]\!]_{\vec{x};y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle\,.$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Proof. The proof is by induction on S. The only interesting case is

$$S \equiv \mathsf{D}^n s \cdot (u_1, \ldots, u_n),$$

which we treat by cases on n:

— Case $n = 1$:
By the definition of substitution, we have

$$[\![(\mathsf{D} s \cdot u_1)\{T/y\}]\!]_{\vec{x}} = [\![\mathsf{D} s\{T/y\} \cdot u_1\{T/y\}]\!]_{\vec{x}}.$$

By the definition of $[\![-]\!]$, this is equal to

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![s\{T/y\}]\!]_{\vec{x}}) \star [\![u_1\{T/y\}]\!]_{\vec{x}}).$$

By the induction hypothesis, we then get

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![s]\!]_{\vec{x};y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle) \star ([\![u_1]\!]_{\vec{x};y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle)).$$

Applying Lemma 4.8 (iii), this is equal to

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![s]\!]_{\vec{x};y}) \star [\![u_1]\!]_{\vec{x};y}) \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle = [\![\mathsf{D} s \cdot u_1]\!]_{\vec{x};y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle.$$

— Case $n > 1$:
By the definition of substitution, we have

$$[\![(\mathsf{D}^n s \cdot (u_1, \ldots, u_n))\{T/y\}]\!]_{\vec{x}} = [\![(\mathsf{D}^n s\{T/y\} \cdot \big(u_1\{T/y\}, \ldots, u_n\{T/y\}\big))]\!]_{\vec{x}}.$$

Applying the definition of $[\![-]\!]$, this is equal to

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![\mathsf{D}^{n-1} s\{T/y\} \cdot \big(u_1\{T/y\}, \ldots, u_{n-1}\{T/y\}\big)]\!]_{\vec{x}}) \star [\![u_n\{T/y\}]\!]_{\vec{x}}).$$

By the definition of substitution, this is

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![(\mathsf{D}^{n-1} s \cdot (u_1, \ldots, u_{n-1}))\{T/y\}]\!]_{\vec{x}}) \star [\![u_n\{T/y\}]\!]_{\vec{x}}).$$

Applying the induction hypothesis twice, we get

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![(\mathsf{D}^{n-1} s \cdot (u_1, \ldots, u_{n-1})]\!]_{\vec{x},y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle) \star ([\![u_n]\!]_{\vec{x},y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle)).$$

By Lemma 4.8 (iii), this is equal to

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![\mathsf{D}^{n-1} s \cdot (u_1, \ldots, u_{n-1})]\!]_{\vec{x},y}) \star [\![u_n]\!]_{\vec{x},y}) \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle,$$

which equals

$$[\![(\mathsf{D}^n s \cdot (u_1, \ldots, u_n))]\!]_{\vec{x};y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle. \qquad \square$$

Theorem 5.7 (Differential Substitution Theorem). Let $S, T \in \Lambda^d$, $\vec{x} = x_1, \ldots, x_n$ and y not occurring in $\vec{x}$. Then

$$\left[\!\!\left[\frac{\partial S}{\partial y} \cdot T \right]\!\!\right]_{\vec{x};y} = [\![S]\!]_{\vec{x};y} \star [\![T]\!]_{\vec{x}}.$$

Proof. We use structural induction on S.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

— Case $S \equiv y$:
We have

$$\left[\!\!\left[\frac{\partial y}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},y} = [\![T]\!]_{\vec{x},y} = [\![T]\!]_{\vec{x}} \circ \pi_1 = \pi_2 \star [\![T]\!]_{\vec{x}} = [\![y]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}$$

by Lemma 4.7 (i).

— Case $S \equiv x_i \neq y$:
We have

$$\left[\!\!\left[\frac{\partial x_i}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},y} = [\![0]\!]_{\vec{x},y} = 0.$$

By Lemma 4.7 (ii), we have

$$0 = ([\![x_i]\!]_{\vec{x}} \circ \pi_1) \star [\![T]\!]_{\vec{x}} = [\![x_i]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}.$$

— Case $S \equiv \lambda z.v$:
By the definition of differential substitution, we have

$$\left[\!\!\left[\frac{\partial \lambda z.v}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},y} = \left[\!\!\left[\lambda z.\frac{\partial v}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},y} = \lambda \circ \Lambda\left(\left[\!\!\left[\frac{\partial v}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},y,z}\right).$$

Applying Lemma 5.3 (ii), this is equal to

$$\lambda \circ \Lambda\left(\left[\!\!\left[\frac{\partial v}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},z,y} \circ \mathrm{sw}\right).$$

By the induction hypothesis, we then get

$$\lambda \circ \Lambda(([\![v]\!]_{\vec{x},z,y} \star [\![T]\!]_{\vec{x},z}) \circ \mathrm{sw}).$$

Supposing, without loss of generality, that $z \notin \mathrm{FV}(T)$, we have, by Lemma 5.3 (i), $[\![T]\!]_{\vec{x},z} = [\![T]\!]_{\vec{x}} \circ \pi_1$. Thus, applying Lemma 4.7 (iii), we have

$$\lambda \circ \Lambda(([\![v]\!]_{\vec{x},z,y} \star ([\![T]\!]_{\vec{x}} \circ \pi_1)) \circ \mathrm{sw}) = \lambda \circ (\Lambda([\![v]\!]_{\vec{x},z,y} \circ \mathrm{sw}) \star [\![T]\!]_{\vec{x}}),$$

which is equal to $\lambda \circ (\Lambda([\![v]\!]_{\vec{x},y,z}) \star [\![T]\!]_{\vec{x}})$ by Lemma 5.3 (ii). Since $\mathscr{U}$ is linear, we can apply Lemma 5.2 (i) and get

$$\begin{aligned}\lambda \circ (\Lambda([\![v]\!]_{\vec{x},y,z}) \star [\![T]\!]_{\vec{x}}) &= (\lambda \circ \Lambda([\![v]\!]_{\vec{x},y,z})) \star [\![T]\!]_{\vec{x}} \\ &= [\![\lambda z.v]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}.\end{aligned}$$

— Case $S \equiv sU$:
By the definition of differential substitution, we have

$$\left[\!\!\left[\frac{\partial sU}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},y} = \left[\!\!\left[\left(\frac{\partial s}{\partial y} \cdot T\right) U\right]\!\!\right]_{\vec{x},y} + \left[\!\!\left[\left(\mathsf{D}s \cdot \left(\frac{\partial U}{\partial y} \cdot T\right)\right) U\right]\!\!\right]_{\vec{x},y}.$$

We consider the two summands separately. For the first we have

$$\left[\!\!\left[\left(\frac{\partial s}{\partial y} \cdot T\right) U\right]\!\!\right]_{\vec{x},y} = \mathrm{ev} \circ \left\langle \mathcal{A} \circ \left[\!\!\left[\frac{\partial s}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},y}, [\![U]\!]_{\vec{x},y} \right\rangle,$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

which is equal by the induction hypothesis to

$$\mathrm{ev} \circ \left\langle \mathcal{A} \circ ([\![s]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}), [\![U]\!]_{\vec{x},y} \right\rangle.$$

By Lemma 5.2 (ii), this is equal to $\mathrm{ev} \circ \langle (\mathcal{A} \circ [\![s]\!]_{\vec{x},y}) \star [\![T]\!]_{\vec{x}}, [\![U]\!]_{\vec{x},y} \rangle$.
For the second summand we have, using $\mathcal{A} \circ \lambda = \mathrm{Id}_{U \Rightarrow U}$,

$$\left[\!\!\left[\left(\mathsf{D}s \cdot \left(\frac{\partial U}{\partial y} \cdot T \right) \right) U \right]\!\!\right]_{\vec{x},y} = \mathrm{ev} \circ \left\langle \Lambda \left(\Lambda^-(\mathcal{A} \circ [\![s]\!]_{\vec{x},y}) \star \left[\!\!\left[\frac{\partial U}{\partial y} \cdot T \right]\!\!\right]_{\vec{x},y} \right), [\![U]\!]_{\vec{x},y} \right\rangle.$$

By the induction hypothesis, this is equal to

$$\mathrm{ev} \circ \left\langle \Lambda(\Lambda^-(\mathcal{A} \circ [\![s]\!]_{\vec{x},y}) \star ([\![U]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}})), [\![T]\!]_{\vec{x},y} \right\rangle.$$

Applying Lemma 4.3, we can rewrite the sum of this two summands as follows:

$$\mathrm{ev} \circ \left\langle (\mathcal{A} \circ [\![s]\!]_{\vec{x},y}) \star [\![T]\!]_{\vec{x}} + \Lambda(\Lambda^-(\mathcal{A} \circ [\![s]\!]_{\vec{x},y}) \star ([\![U]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}})), [\![U]\!]_{\vec{x},y} \right\rangle.$$

By Lemma 4.8 (i), this is

$$\left(\mathrm{ev} \circ \left\langle \mathcal{A} \circ [\![s]\!]_{\vec{x},y}, [\![U]\!]_{\vec{x},y} \right\rangle \right) \star [\![T]\!]_{\vec{x}} = [\![sU]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}.$$

— Case $S \equiv \mathsf{D}^n v \cdot (u_1, \ldots, u_n)$:
We consider subcases on n:

– Subcase $n = 1$:
By the definition of differential substitution, we have

$$\left[\!\!\left[\frac{\partial}{\partial y} (\mathsf{D}v \cdot u_1) \cdot T \right]\!\!\right]_{\vec{x},y} = \left[\!\!\left[\mathsf{D} \left(\frac{\partial v}{\partial y} \cdot T \right) \cdot u_1 \right]\!\!\right]_{\vec{x},y} + \left[\!\!\left[\mathsf{D}v \cdot \left(\frac{\partial u_1}{\partial y} \cdot T \right) \right]\!\!\right]_{\vec{x},y}.$$

We consider the two summands separately. For the first, we have

$$\left[\!\!\left[\mathsf{D} \left(\frac{\partial v}{\partial y} \cdot T \right) \cdot u_1 \right]\!\!\right]_{\vec{x},y} = \lambda \circ \Lambda \left(\Lambda^- \left(\mathcal{A} \circ \left[\!\!\left[\frac{\partial v}{\partial y} \cdot T \right]\!\!\right]_{\vec{x},y} \right) \star [\![u_1]\!]_{\vec{x},y} \right).$$

By the induction hypothesis, this is equal to

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ ([\![v]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}})) \star [\![u_1]\!]_{\vec{x},y}),$$

which is equal to

$$\lambda \circ \Lambda(\Lambda^-((\mathcal{A} \circ [\![v]\!]_{\vec{x},y}) \star [\![T]\!]_{\vec{x}}) \star [\![u_1]\!]_{\vec{x},y})$$

by Lemma 5.2 (ii).
For the second summand, we have

$$\left[\!\!\left[\mathsf{D}v \cdot \left(\frac{\partial u_1}{\partial y} \cdot T \right) \right]\!\!\right]_{\vec{x},y} = \lambda \circ \Lambda \left(\Lambda^-(\mathcal{A} \circ [\![v]\!]_{\vec{x},y}) \star \left[\!\!\left[\frac{\partial u_1}{\partial y} \cdot T \right]\!\!\right]_{\vec{x},y} \right).$$

By the induction hypothesis, this is

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![v]\!]_{\vec{x},y}) \star ([\![u_1]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}})).$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Since λ is linear, we can apply Lemma 4.5 and write the sum of the two morphisms as

$$\lambda \circ \big(\Lambda(\Lambda^-((\mathcal{A} \circ [\![v]\!]_{\vec{x},y}) \star [\![T]\!]_{\vec{x}}) \star [\![u_1]\!]_{\vec{x},y}) + \Lambda(\Lambda^-(\mathcal{A} \circ [\![v]\!]_{\vec{x},y}) \star ([\![u_1]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}))\big).$$

Applying Lemma 4.8 (ii), we obtain

$$\lambda \circ (\Lambda(\Lambda^-(\mathcal{A} \circ [\![v]\!]_{\vec{x},y}) \star [\![u_1]\!]_{\vec{x},y}) \star [\![T]\!]_{\vec{x}}),$$

which is equal to $[\![\mathsf{D}v \cdot u_1]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}$.

– Subcase $n > 1$:
Easy calculations give

$$\left[\!\!\left[\frac{\partial}{\partial y}(\mathsf{D}^n v \cdot (u_1, \ldots, u_n)) \cdot T\right]\!\!\right]_{\vec{x};y} =$$

(1) $$\left[\!\!\left[\mathsf{D}\left(\frac{\partial}{\partial y}(\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})) \cdot T\right) \cdot u_n\right]\!\!\right]_{\vec{x};y} +$$

(2) $$\left[\!\!\left[\mathsf{D}\left(\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})\right) \cdot \left(\frac{\partial u_n}{\partial y} \cdot T\right)\right]\!\!\right]_{\vec{x};y}.$$

We consider the two summands separately:

- Summand(1):

  By the induction hypothesis,

$$\lambda \circ \Lambda\left(\Lambda^-\left(\mathcal{A} \circ \left[\!\!\left[\frac{\partial}{\partial y}(\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})) \cdot T\right]\!\!\right]_{\vec{x},y}\right) \star [\![u_n]\!]_{\vec{x},y}\right)$$

  equals

$$\lambda \circ \Lambda\left(\Lambda^-\left(\mathcal{A} \circ \left([\![\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}\right)\right) \star [\![u_n]\!]_{\vec{x},y}\right),$$

  which, by Lemma 5.2 (ii), equals

$$\lambda \circ \Lambda\left(\Lambda^-\left(\left(\mathcal{A} \circ [\![\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})]\!]_{\vec{x},y}\right) \star [\![T]\!]_{\vec{x}}\right) \star [\![u_n]\!]_{\vec{x},y}\right).$$

- Summand(2):

  By the induction hypothesis,

$$\lambda \circ \Lambda\left(\Lambda^-\left(\mathcal{A} \circ [\![\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})]\!]_{\vec{x},y}\right) \star \left[\!\!\left[\frac{\partial u_n}{\partial y} \cdot T\right]\!\!\right]_{\vec{x},y}\right)$$
$$= \lambda \circ \Lambda\left(\Lambda^-\left(\mathcal{A} \circ [\![\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})]\!]_{\vec{x},y}\right) \star \left([\![u_n]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}\right)\right).$$

Since λ is linear, we have (1) + (2) is equal to

$$\lambda \circ \big(\Lambda(\Lambda^-((\mathcal{A} \circ [\![\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})]\!]_{\vec{x},y}) \star [\![T]\!]_{\vec{x}}) \star [\![u_n]\!]_{\vec{x},y}) +$$
$$\Lambda(\Lambda^-(\mathcal{A} \circ [\![\mathsf{D}^{n-1} v \cdot (u_1, \ldots, u_{n-1})]\!]_{\vec{x},y}) \star ([\![u_n]\!]_{\vec{x},y} \star [\![T]\!]_{\vec{x}}))\big)$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

By Lemma 4.8 (ii), we then get

$$\lambda \circ (\Lambda(\Lambda^-(\mathcal{A} \circ [\![\mathsf{D}^{n-1} v \cdot (u_1, \dots, u_{n-1})]\!]_{\vec{x},y}) \star [\![u_n]\!]_{\vec{x}}) \star [\![T]\!]_{\vec{x}}),$$

which, by Lemma 5.2 (i), is equal to

$$\lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ [\![\mathsf{D}^{n-1} v \cdot (u_1, \dots, u_{n-1})]\!]_{\vec{x},y}) \star [\![u_n]\!]_{\vec{x}}) \star [\![T]\!]_{\vec{x}},$$

in other words,

$$[\![\mathsf{D}^n v \cdot (u_1, \dots, u_n)]\!]_{\vec{x};y} \star [\![T]\!]_{\vec{x}}.$$

— Case $S \equiv 0$:
This case is straightforward.

— Case $S \equiv s + U$:
This case is straightforward. □

We are now able to state the main result of this section.

Theorem 5.8 (Soundness). Every linear reflexive object $\mathscr{U}$ in a Cartesian closed differential category **C** is a sound model of the differential λ-calculus.

Proof. It is easy to check that the categorical interpretation is contextual. We now prove that $\mathrm{Th}(\mathscr{U})$ is closed under the rules (β) and (β_D):

— Rule (β):
Let

$$[\![(\lambda y.s)T]\!]_{\vec{x}} = \mathrm{ev} \circ \langle \mathcal{A} \circ \lambda \circ \Lambda([\![s]\!]_{\vec{x},y}), [\![T]\!]_{\vec{x}} \rangle.$$

Since $\mathcal{A} \circ \lambda = \mathrm{Id}$, this is equal to

$$\mathrm{ev} \circ \langle \Lambda([\![s]\!]_{\vec{x},y}), [\![T]\!]_{\vec{x}} \rangle.$$

On the other side, we have

$$[\![s\{T/y\}]\!]_{\vec{x}} = [\![s]\!]_{\vec{x},y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle$$

by Theorem 5.6, and by (beta-cat),

$$[\![s]\!]_{\vec{x},y} \circ \langle \mathrm{Id}, [\![T]\!]_{\vec{x}} \rangle = \mathrm{ev} \circ \left\langle \Lambda([\![s]\!]_{\vec{x},y}), [\![T]\!]_{\vec{x}} \right\rangle.$$

— Rule (β_D):
Let

$$[\![\mathsf{D}(\lambda y.s) \cdot t]\!]_{\vec{x}} = \lambda \circ \Lambda(\Lambda^-(\mathcal{A} \circ \lambda \circ \Lambda([\![s]\!]_{\vec{x},y})) \star [\![t]\!]_{\vec{x}}).$$

Since $\mathcal{A} \circ \lambda = \mathrm{Id}$, this is equal to

$$\lambda \circ \Lambda(\Lambda^-(\Lambda([\![s]\!]_{\vec{x},y})) \star [\![t]\!]_{\vec{x}}) = \lambda \circ \Lambda([\![s]\!]_{\vec{x},y} \star [\![t]\!]_{\vec{x}}).$$

Applying Theorem 5.7, this is equal to

$$\lambda \circ \Lambda\left(\left[\!\!\left[\frac{\partial s}{\partial y} \cdot t \right]\!\!\right]_{\vec{x},y} \right) = \left[\!\!\left[\lambda y. \frac{\partial s}{\partial y} \cdot t \right]\!\!\right]_{\vec{x}}.$$

We can then conclude that $\mathrm{Th}(\mathscr{U})$ is a differential λ-theory. □

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

The above theorem shows that linear reflexive objects in Cartesian closed differential categories are sound models of the untyped differential λ-calculus.

Proposition 5.9. If $\mathscr{U}$ is extensional, then $\mathrm{Th}(\mathscr{U})$ is extensional.

Proof. As in the case of ordinary λ-calculus, easy calculations show that

$$[\![\lambda x.sx]\!]_{\vec{x}} = \lambda \circ \Lambda(\mathrm{ev}) \circ \mathcal{A} \circ [\![s]\!]_{\vec{x}},$$

which is equal to $[\![s]\!]_{\vec{x}}$ since $\Lambda(\mathrm{ev}) = \mathrm{Id}$ and $\lambda \circ \mathcal{A} = \mathrm{Id}$. □

5.4. *Equational completeness*

An important result in the ordinary λ-calculus is the *equational completeness theorem* proved in Scott (1980) and subsequently refined in Koymans (1982). This theorem states that every λ-theory is the theory of a reflexive object in a Cartesian closed category. In this section we discuss whether the categorical notion of a model of the differential λ-calculus presented in Section 5 is also complete. In other words, we investigate the question of whether for every differential λ-theory $\mathcal{T}$ there is a linear reflexive object $\mathscr{U}_{\mathcal{T}}$ living in a suitable Cartesian closed differential category $\mathbf{C}_{\mathcal{T}}$ such that $\mathrm{Th}(\mathscr{U}_{\mathcal{T}}) = \mathcal{T}$. We will be able to answer yes to this question, provided $\mathcal{T}$ is differentially extensional and satisfies sum idempotency. This restriction is quite reasonable since all known models that have arisen so far do satisfy these properties (see Sections 6.1.1 and 6.3.1). However, these conditions arise from some technical choices we have to make, and it is not yet known whether other choices might lead to a more general theorem.

Before going further, we will outline the proof of the classic Scott–Koymans' result, which is achieved in two steps:

(i) Given a λ-theory $\mathcal{T}$, one proves that the set of λ-terms modulo $\mathcal{T}$ together with the application operator defined between equivalence classes constitutes a *λ-model*† $\mathscr{M}_{\mathcal{T}}$ (called *the term model of* $\mathcal{T}$) having as theory exactly $\mathcal{T}$.

(ii) By applying a construction called the *Karoubi envelope* (Karoubi 1978) to $\mathscr{M}_{\mathcal{T}}$, one builds a (very syntactic) Cartesian closed category $\mathbf{C}_{\mathcal{T}}$ in which the identity $\mathbf{I}$ is a reflexive object such that $\mathrm{Th}(\mathbf{I}) = \mathcal{T}$.

Summing up, the idea of the proof is to find suitable λ-terms to encode the structure of the category (pairing, currying, evaluation, and the like), and then prove that they actually define a category with such a structure.

In our context, the categorical operator $D(-)$ can be easily defined in terms of the linear application. Intuitively, the term representing $D(f)$ takes in input a pair and applies the first component linearly and the second in the usual way, in accordance with the categorical axiomatisation of $D(-)$. The main problem we need to solve is that the encoding of the categorical pairing $\langle f, g\rangle$ used by Scott is not additive. Indeed, such a pairing is

† A 'λ-model' is a combinatory algebra satisfying the five axioms of Curry and the Meyer–Scott axiom – see Barendregt (1984, Chapter 5) for more details.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

defined from Church's encoding of the pair in λ-calculus given by $\langle\langle f, g\rangle\rangle \equiv \lambda x.xfg$ with projections

$$p_1 = \lambda z.z(\lambda xy.x)$$
$$p_2 = \lambda z.z(\lambda xy.y).$$

Obviously, with this definition we have

$$\langle\langle f_1 + f_2, g_1 + g_2\rangle\rangle \neq \langle\langle f_1, g_1\rangle\rangle + \langle\langle f_2, g_2\rangle\rangle$$

since the sums do not occur in linear position. We will see that the encoding of an additive pairing can be obtained using the linear application and the sum in the differential λ-calculus.

Notation 5.10. Given a differential λ-theory $\mathcal{T}$, we write $\Lambda^d_{\mathcal{T}}$ for $\Lambda^d/\mathcal{T}$.

From now until the end of the section, we set $A \circ B \equiv \lambda x.A(Bx)$. We say that $A \in \Lambda^d_{\mathcal{T}}$ is *idempotent* if $A \circ A = A$ and *additive* if $A(x + y) = Ax + Ay$.

Definition 5.11. Let $\mathcal{T}$ be a differential λ-theory. The category $\mathbf{C}_{\mathcal{T}}$ associated with $\mathcal{T}$ is defined as follows:

Objects:	$\{A \in \Lambda^d_{\mathcal{T}} \mid A$ is idempotent and additive$\}$
Arrows:	$\mathbf{C}_{\mathcal{T}}(A, B) = \{f \in \Lambda^d_{\mathcal{T}} \mid B \circ f \circ A = f\}$
Identities:	$\mathrm{Id}_A = A$
Composition:	$f \circ g$.

It is easy to verify that $\mathbf{C}_{\mathcal{T}}$ is indeed a category.

We now encode the ordered pair $\langle\langle S, T\rangle\rangle$ in the differential λ-calculus as follows – we will use this notion in the definition of categorical pairing.

Definition 5.12. The encoding of the pair into the differential λ-calculus is given by

$$\langle\langle S, T\rangle\rangle \equiv \lambda y.(S + \mathsf{D}y \cdot T), \text{ for some } y \notin \mathrm{FV}(S, T)$$

with projections

$$p_1 \equiv \lambda x.x0$$
$$p_2 \equiv \lambda x.(\mathsf{D}x \cdot \mathbf{I})00.$$

It is immediate that

$$p_i\langle\langle S_1, S_2\rangle\rangle = S_i \qquad (\text{for } i = 1, 2)$$

and that

$$\langle\langle S_1 + S_2, T_1 + T_2\rangle\rangle = \langle\langle S_1, T_1\rangle\rangle + \langle\langle S_2, T_2\rangle\rangle.$$

This encoding is inspired by the set-theoretical definition of the ordered pair: the pair of S, T is essentially the set containing S, T (the sum being the union) slightly modified to make them distinguishable. Such a distinction consists of the number of linear resources they can receive (zero for the first component; one for the second).

Given this encoding, we can endow $\mathbf{C}_{\mathcal{T}}$ with the structure of a differential Cartesian closed category, under the assumption that the sum is idempotent (like set-theoretical union).

Theorem 5.13. For all differential λ-theories $\mathcal{T}$ satisfying sum idempotency:

(i) $\mathbf{C}_{\mathcal{T}}$ is differential Cartesian closed.
(ii) The triple $(\mathbf{I}, \mathbf{1}, \mathbf{1})$ is a linear reflexive object.

Proof.

(i) We have

— *Terminal object*:
This is $\mathbb{1} \equiv \lambda xy.y$. Note that $f : A \to \mathbb{1}$ if and only if $f \equiv !_A \equiv \lambda xy.y$.

— *Products*:
Given two objects A_1, A_2, the object

$$A_1 \times A_2 \equiv \lambda z.\langle\!\langle A_1(p_1 z), A_2(p_2 z)\rangle\!\rangle$$

is the Cartesian product of A_1 and A_2:

– Projections:
We have

$$\pi_1 : A_1 \times A_2 \to A_1, \qquad \pi_1^{A_1,A_2} \equiv A_1 \circ p_1$$
$$\pi_2 : A_1 \times A_2 \to A_2, \qquad \pi_2^{A_1,A_2} \equiv A_2 \circ p_2.$$

– Pairing:
Let $f : A \to B$ and $g : A \to C$. Then

$$\langle f, g\rangle \equiv \lambda z.\langle\!\langle fz, gz\rangle\!\rangle.$$

— *Exponents*:
Given two objects A, B, the object $A \Rightarrow B \equiv \lambda z.B \circ z \circ A$ is the exponential object internalising $\mathbf{C}_{\mathcal{T}}(A, B)$. The evaluation morphism $\mathrm{ev} : (A \Rightarrow B) \times A \to B$ is defined by

$$\mathrm{ev} \equiv \lambda z.B((p_1 z)(A(p_2 z)))$$

and the curry of a morphism $f : A \times B \to C$ is given by

$$\Lambda(f) \equiv \lambda xy.f\langle\!\langle x, y\rangle\!\rangle.$$

— *Differential operator*:
Given a morphism $f : A \to B$, we define

$$D(f) \equiv \lambda z.B\big((\mathsf{D}f \cdot (A(p_1 z)))(A(p_2 z))\big)$$

— *Left-additive structure*:
We interpret the sum in the category as the sum on $\Lambda^d_{\mathcal{T}}$.
The calculations showing that everything works are straightforward but *very* lengthy. As a simple example, we will just prove that categorical pairing is indeed

additive:

$$\begin{aligned}\langle f_1 + f_2, g_1 + g_2\rangle &= \lambda z.\langle\!\langle f_1 z + f_2 z, g_1 z + g_2 z\rangle\!\rangle \\ &= \lambda y.(f_1 z + f_2 z) + \lambda y.\mathsf{D} y \cdot (g_1 z + g_2 z) \\ &= \lambda y.f_1 z + \lambda y.f_2 z + \lambda y.\mathsf{D} y \cdot (g_1 z) + \lambda y.\mathsf{D} y \cdot (g_2 z) \\ &= \langle f_1, g_1\rangle + \langle f_2, g_2\rangle .\end{aligned}$$

(ii) Note that $(\mathbf{I} \Rightarrow \mathbf{I}) = \mathbf{1}$. Then $\mathbf{I}$ is a reflexive object since $\mathbf{1} : \mathbf{1} \to \mathbf{I}$, $\mathbf{1} : \mathbf{I} \to \mathbf{1}$ and $\mathbf{1} \circ \mathbf{1} = \mathrm{Id}_{\mathbf{1}}$. Moreover, $\mathbf{I}$ is linear as a reflexive object:

$$\begin{aligned}D(\mathbf{1}) &= \lambda z.(\mathsf{D}\mathbf{1} \cdot (p_1 z))(p_2 z) \\ &= \lambda z.(\lambda xy.p_1 zy)(p_2 z) \\ &= \lambda zy.p_1 zy \\ &= \mathbf{1} \circ \pi_1 .\end{aligned}$$

□

In the above proof, the idempotency of the sum is required, for instance, to prove the axiom (D-curry). The task of finding an encoding of the additive pairing that does not require the idempotency of the sum is left for future work.

In order to provide a characterisation of the interpretation of a differential λ-term S, we need the following definition.

Definition 5.14. The *full η_∂-expansion* $\widehat{S}$ of a differential λ-term $S \in \Lambda^d$ is defined by induction (where y is a fresh variable):

$$\begin{aligned}\widehat{x} &\equiv x \\ \widehat{\lambda x.s} &\equiv \lambda x.\widehat{s} \\ \widehat{sT} &\equiv \widehat{s}\widehat{T} \\ \widehat{\mathsf{D}s \cdot t} &\equiv \lambda y.(\mathsf{D}\widehat{s} \cdot \widehat{t})y \\ \widehat{\textstyle\sum_i s_i} &\equiv \textstyle\sum_i \widehat{s_i}.\end{aligned}$$

Roughly speaking, the term $\widehat{S}$ is obtained from S by performing one η_∂-expansion in all its subterms of shape $\mathsf{D}s \cdot t$. The adjective *full* refers to the fact that the η_∂-expansion is carried out inductively on the structure of S.

Remark 5.15. It is obvious that if $\mathcal{T}$ is differentially extensional, then $\mathcal{T} \vdash S = \widehat{S}$ for all $S \in \Lambda^d$.

Proposition 5.16. In the model $\mathbf{I}$ living in $\mathbf{C}_{\mathcal{T}}$, the following holds (for some $z \notin \mathrm{FV}(S)$):

$$[\![S]\!]_{\vec{x}} = \lambda z.\widehat{S}\left\{\pi^{\vec{x}}_{x_1} z/x_1\right\}\cdots\left\{\pi^{\vec{x}}_{x_n} z/x_n\right\} : \mathbf{I}^{\vec{x}} \to \mathbf{I}.$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Proof. The proof is by induction on the structure of S. There are only two non-trivial cases:

— Case $S \equiv \lambda y.T$:
We have

$$
\begin{aligned}
[\![\lambda x_{n+1}.T]\!]_{\vec{x}} &= \mathbf{1} \circ \Lambda([\![T]\!]_{\vec{x},x_{n+1}}) && \text{(definition of } [\![\cdot]\!]_{\vec{x}}) \\
&= \mathbf{1} \circ (\lambda y_1 y_2.[\![T]\!]_{\vec{x},x_{n+1}} \langle\!\langle y_1, y_2 \rangle\!\rangle) && \text{(definition of } \Lambda(\cdot)) \\
&= \lambda y_1 y_2. \left(\lambda z.\widehat{T}\left\{\pi^{\vec{x},x_{n+1}}_{x_1} z/x_1\right\}\cdots\left\{\pi^{\vec{x},x_{n+1}}_{x_{n+1}} z/x_{n+1}\right\}\right)\langle\!\langle y_1, y_2\rangle\!\rangle \\
& && \text{(induction hypothesis)} \\
&= \lambda y_1 y_2.\widehat{T}\left\{\pi^{\vec{x}}_{x_1} y_1/x_1\right\}\cdots\left\{\pi^{\vec{x}}_{x_n} y_1/x_n\right\}\{y_2/x_{n+1}\} && (\beta\text{-reduction}) \\
&= \lambda z.(\lambda x_{n+1}.\widehat{T})\left\{\pi^{\vec{x}}_{x_1} z/x_1\right\}\cdots\left\{\pi^{\vec{x}}_{x_n} z/x_n\right\}. && (\alpha\text{-conversion})
\end{aligned}
$$

— Case $S \equiv \mathsf{D}T \cdot U$.

$$
\begin{aligned}
&[\![\mathsf{D}T\cdot U]\!]_{\vec{x}} \\
&\quad = \lambda zy.(\mathsf{D}([\![T]\!]_{\vec{x}}z)\cdot([\![U]\!]_{\vec{x}}z))y \\
&\quad = \lambda zy.\left(\mathsf{D}\left(\left(\lambda z.T\left\{\pi^{\vec{x}}_{x_1}z/x_1\right\}\cdots\left\{\pi^{\vec{x}}_{x_n}z/x_n\right\}\right)z\right)\cdot\right. \\
&\qquad\qquad \left.\left(\left(\lambda z.S\left\{\pi^{\vec{x}}_{x_1}z/x_1\right\}\cdots\left\{\pi^{\vec{x}}_{x_n}z/x_n\right\}\right)z\right)\right)y \\
&\quad = \lambda zy.(\mathsf{D}(T\left\{\pi^{\vec{x}}_{x_1}z/x_1\right\}\cdots\left\{\pi^{\vec{x}}_{x_n}z/x_n\right\})\cdot(S\left\{\pi^{\vec{x}}_{x_1}/x_1\right\}\cdots\left\{\pi^{\vec{x}}_{x_n}z/x_n\right\}))y.
\end{aligned}
$$
□

Therefore, in the theory of the model **I**, equations of the form

$$\mathsf{D}s\cdot t = \lambda y.(\mathsf{D}s\cdot t)y$$

might be added. No equation can be added when the theory $\mathcal{T}$ is already differentially extensional.

Theorem 5.17 (Equational Completeness). Every differentially extensional differential λ-theory $\mathcal{T}$ satisfying sum idempotency is the theory of a linear reflexive object in a differential Cartesian closed category.

Proof. For all closed terms $S, T \in \Lambda^d$, we have, by Proposition 5.16, that $[\![S]\!]_{\vec{x}} = [\![T]\!]_{\vec{x}}$ entails $\mathcal{T} \vdash \lambda z.\widehat{S} = \lambda z.\widehat{T}$ and, by Remark 5.15, that $\mathcal{T} \vdash \lambda z.S = \lambda z.T$. Since $\mathcal{T}$ is a differential λ-theory, we also have $\mathcal{T} \vdash (\lambda z.S)0 = (\lambda z.T)0$. Since $z \notin \mathrm{FV}(S,T)$ and $\lambda\beta^d \subseteq \mathcal{T}$, we get $\mathcal{T} \vdash S = T$, so $\mathrm{Th}(\mathbf{I}) = \mathcal{T}$. □

We will show in Section 7.3.1 that an analogous theorem holds for the resource calculus, but without the restriction to theories that are differentially extensional.

Remark 5.18. Theorem 5.17 does not mean that the theory of every model in a differential Cartesian closed category must satisfy sum-idempotency and the differential extensional axiom. These are just technical conditions arising from the specific proof of the completeness theorem given here. However, no more general proof is currently known.

The completeness theorem constitutes an important result and suggests that the notion of a model we have chosen for the differential λ-calculus is actually correct.

On the other hand, denotational models are usually introduced because they allow us to study a calculus by means of more abstract mathematical structures for which a broader range of tools and proof techniques are available. In this respect, the categorical models living in $\mathbf{C}_{\mathcal{T}}$ are unsatisfactory because they are very syntactical, and proving operational properties of differential λ-terms using these models does not make it any easier than working directly with the syntax.

For this reason, it would be interesting to find meaningful classes of models (semantics) that are *complete* in the sense that they allow us to represent all differential λ-theories. However, even in the case of the ordinary λ-calculus, it is well known that the main examples of semantics, that is, the continuous, the stable and the strongly stable semantics, are all *hugely incomplete* – there is a continuum of λ-theories that cannot be represented by models living in such semantics (Salibra 2001). We will show in Section 6.3 that a similar result holds for the relational semantics of the differential λ-calculus (Corollary 6.10).

The problem of finding a complete semantics of the (differential) λ-calculus is open and quite difficult.

5.5. *Comparison with the categorical models of the untyped lambda calculus*

The definition of a categorical model of the differential λ-calculus proposed in this paper seems to be a straightforward generalisation of the classical definition of a model of the λ-calculus, that is, the notion of a reflexive object in a Cartesian closed category. However, while this notion is by far the best-known categorical definition of a model of λ-calculus, it is not the most general one. Indeed, as pointed out in Martini (1992), there is one axiom of Cartesian closed categories that is never used in the proof of soundness for categorical models (Barendregt 1984, Proposition 5.5.5), namely, the axiom (Id-Curry), which is equivalent to requiring the unicity of the operator $\Lambda(-)$ in the category (and this entails $\Lambda(\Lambda^-(f)) = f$).

For this reason, Martini proposed reflexive objects living in *weak* Cartesian closed categories as a more general notion of a model of λ-calculus. In these categories, there is just a retraction (not an isomorphism) between the homsets $\mathbf{C}(C \times A, B) \lhd \mathbf{C}(C, A \Rightarrow B)$. So $A \Rightarrow B$ is no longer an object representing $\mathbf{C}(A, B)$ *exactly* – there are other objects that can equally well accomplish the job. Recently, de Carvalho (2007) successfully used this notion to build concrete models living in very natural weak Cartesian closed categories inspired by the semantics of linear logic.

However, this generalisation cannot be applied in our differential framework because the proof of soundness relies on the fact that $\Lambda(\Lambda^-(f)) = f$. This is actually required in order to give a meaningful interpretation of the linear application $\mathsf{D}s \cdot t$. Hence, the definition of a categorical model of the differential λ-calculus we have presented here is more different from the corresponding one for the ordinary λ-calculus than one might think at first glance.

5.6. *Modelling the Taylor expansion*

In this section we provide sufficient conditions for models living in Cartesian closed differential categories to equate all differential λ-terms having the same Taylor expansion.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

As an interesting fact, this happens to be a property of the category rather than of the reflexive objects. Therefore, all models living in a category 'modelling the Taylor expansion' have an equational theory including $\mathcal{E}$.

Since the definition of the Taylor expansion requires infinite sums, we need to consider Cartesian closed differential categories $\mathbf{C}$ where it is possible to sum infinitely many morphisms. Formally, we require that for every countable set I and every family $\{f_i\}_{i\in I}$ of morphisms $f_i : A \to B$, we have $\sum_{i\in I} f_i \in \mathbf{C}(A,B)$. In this case we say that $\mathbf{C}$ *has countable sums*. To avoid the tedious problem of handling coefficients, we assume that the sum on the morphisms is idempotent.

Definition 5.19. A Cartesian closed differential category *models the Taylor Expansion* if it has countable sums and the following axiom holds (for every $f : C \times A \to B$ and $g : C \to A$):

$$\text{(Taylor)} \qquad \mathrm{ev} \circ \langle f, g\rangle = \sum_{k\in\mathbf{N}} ((\underbrace{\cdots(\Lambda^-(f) \star g)\cdots) \star g}_{k \text{ times}}) \circ \langle \mathrm{Id}, 0\rangle\,.$$

Recall that the Taylor expansion S^* of a differential λ-term S was defined in Section 3.4. Given a model $\mathscr{U}$ of the differential λ-calculus living in a Cartesian closed differential category having countable sums, we can extend the interpretation given in Definition 5.3 to terms in Λ^d_∞ by setting $[\![\Sigma_{i\in I} s_i]\!]_{\vec{x}} = \sum_{i\in I} [\![s_i]\!]_{\vec{x}}$, for every countable set I.

Theorem 5.20. Let S be a differential λ-term and $\mathscr{U}$ be a model living in a Cartesian closed differential category having countable sums and modelling the Taylor Expansion. Then:

$$[\![S]\!]_{\vec{x}} = [\![S^*]\!]_{\vec{x}}.$$

Proof. We use structural induction on S – the only interesting case is $S \equiv sT$:

$$\begin{aligned}
[\![sT]\!]_{\vec{x}} &= \mathrm{ev} \circ \langle \mathcal{A} \circ [\![s]\!]_{\vec{x}}, [\![T]\!]_{\vec{x}}\rangle && \text{(definition of } [\![-]\!]_{\vec{x}}) \\
&= \sum_{k\in\mathbf{N}} ((\underbrace{\cdots(\Lambda^-(\mathcal{A} \circ [\![s]\!]_{\vec{x}}) \star [\![T]\!]_{\vec{x}})\cdots) \star [\![T]\!]_{\vec{x}}}_{k \text{ times}}) \circ \langle \mathrm{Id}, 0\rangle && \text{(Taylor)} \\
&= \sum_{k\in\mathbf{N}} \mathrm{ev} \circ \left\langle \Lambda((\underbrace{\cdots(\Lambda^-([\![s]\!]_{\vec{x}} \star [\![T]\!]_{\vec{x}})\cdots) \star [\![T]\!]_{\vec{x}}}_{k \text{ times}}), 0 \right\rangle && \text{(beta-cat)} \\
&= \sum_{k\in\mathbf{N}} \mathrm{ev} \circ \left\langle \mathcal{A} \circ \lambda \circ \Lambda((\underbrace{\cdots(\Lambda^-([\![s]\!]_{\vec{x}} \star [\![T]\!]_{\vec{x}})\cdots) \star [\![T]\!]_{\vec{x}}}_{k \text{ times}}), 0 \right\rangle && (\mathcal{A} \circ \lambda = \mathrm{Id}) \\
&= \sum_{k\in\mathbf{N}} \mathrm{ev} \circ \langle \mathcal{A} \circ [\![\mathsf{D}^k s \cdot (T, \ldots, T)]\!]_{\vec{x}}, 0\rangle && \text{(definition of } [\![-]\!]_{\vec{x}}) \\
&= [\![\Sigma_{k\in\mathbf{N}} (\mathsf{D}^k s \cdot (T, \ldots, T))0]\!]_{\vec{x}} && \text{(definition of } [\![-]\!]_{\vec{x}}) \\
&= [\![(sT)^*]\!]_{\vec{x}} && \text{(definition of } (\cdot)^*) \quad \square
\end{aligned}$$

By adapting the proof of Theorem 5.8, we can then prove that $[\![S^*]\!]_{\vec{x}} = [\![\mathrm{NF}(S^*)]\!]_{\vec{x}}$ for every differential λ-term S. From this fact and Theorem 5.20, we get the following result.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Corollary 5.21. Every model $\mathscr{U}$ living in a Cartesian closed differential category that models the Taylor expansion satisfies $\mathcal{E} \subseteq \mathrm{Th}(\mathscr{U})$.

6. A relational model of the differential lambda calculus

In this section we discuss the main example of a Cartesian closed differential category known in the literature; *viz.* the category **MRel** (Girard 1988; Bucciarelli *et al.* 2007), which is the co-Kleisli category of the functor $\mathcal{M}_f(-)$ over the $\star$-autonomous category **Rel** of sets and relations. We will also show that the reflexive object $\mathscr{D}$ living in **MRel** that was built in Bucciarelli *et al.* (2007) to model the ordinary λ-calculus is linear, and then that it constitutes a model of the untyped differential λ-calculus. We will then provide a partial characterisation of its equational theory showing that it contains $\lambda\beta\eta^d$ and $\mathcal{E}$ (this follows from the fact that **MRel** models the Taylor expansion).

Remark 6.1. Bucciarelli *et al.* (2010) also provided another example of a Cartesian closed differential category: the category **MFin**, which is the co-Kleisli of the functor $\mathcal{M}_f(-)$ over the $\star$-autonomous category of finiteness spaces and finitary relations (Ehrhard 2005). However, we will not discuss the category **MFin** in the current paper since it does not contain any reflexive object (Ehrhard 2005; Vaux 2009) and thus cannot be used as a semantics of the untyped differential λ-calculus. Other examples of semantics useful for modelling the untyped differential λ-calculus (including semantics that *do not* model the Taylor expansion) will be discussed in Section 8.2.

6.1. *Relational semantics*

Recall that we introduced the definitions and notation for multisets in Section 2.1. We will now provide a direct definition of the category **MRel**:

— The objects of **MRel** are all the sets.

— A morphism from A to B is a relation from $\mathcal{M}_f(A)$ to B; in other words,

$$\mathbf{MRel}(A,B) = \mathcal{P}(\mathcal{M}_f(A) \times B).$$

— The identity of A is the relation

$$\mathrm{Id}_A = \{([\alpha],\alpha) \mid \alpha \in A\} \in \mathbf{MRel}(A,A).$$

— The composition of $s \in \mathbf{MRel}(A,B)$ and $t \in \mathbf{MRel}(B,C)$ is defined by

$$\begin{aligned} t \circ s = \{(m,\gamma) \quad | \quad & \exists k \in \mathbf{N},\ \exists (m_1,\beta_1),\dots,(m_k,\beta_k) \in s \text{ such that} \\ & m = m_1 \uplus \dots \uplus m_k \text{ and } ([\beta_1,\dots,\beta_k],\gamma) \in t\ \}. \end{aligned}$$

Given two sets A_1, A_2, we use $A_1 \& A_2$ to denote their disjoint union

$$(\{1\} \times A_1) \cup (\{2\} \times A_2).$$

From now on, we will adopt the following convention.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Convention 6.2. We consider the canonical bijection between $\mathcal{M}_f(A_1) \times \mathcal{M}_f(A_2)$ and $\mathcal{M}_f(A_1 \& A_2)$ as an equality. Therefore, we will continue to use (m_1, m_2) to denote the corresponding element of $\mathcal{M}_f(A_1 \& A_2)$.

Theorem 6.3. The category **MRel** is a Cartesian closed category.

Proof. The terminal object $\mathbb{1}$ is the empty set $\emptyset$, and the unique element of $\mathbf{MRel}(A, \emptyset)$ is the empty relation.

Given two sets A_1 and A_2, their categorical product in **MRel** is their disjoint union $A_1 \& A_2$ and the projections π_1, π_2 are given by

$$\pi_i = \{([(i,a)], a) \mid a \in A_i\} \in \mathbf{MRel}(A_1 \& A_2, A_i), \text{ for } i = 1, 2.$$

It is easy to check that this is actually the categorical product of A_1 and A_2 in **MRel**; given $s \in \mathbf{MRel}(B, A_1)$ and $t \in \mathbf{MRel}(B, A_2)$, the corresponding morphism $\langle s, t\rangle \in \mathbf{MRel}(B, A_1 \& A_2)$ is given by

$$\langle s, t\rangle = \{(m, (1, a)) \mid (m, a) \in s\} \cup \{(m, (2, b)) \mid (m, b) \in t\}.$$

Given two objects A and B, the exponential object $A \Rightarrow B$ is $\mathcal{M}_f(A) \times B$ and the evaluation morphism is given by

$$\mathrm{ev}_{AB} = \{(([(m, b)], m), b) \mid m \in \mathcal{M}_f(A) \text{ and } b \in B\} \in \mathbf{MRel}((A \Rightarrow B) \& A, B).$$

It is again easy to check that we have defined an exponentiation in this way. Indeed, given any set C and any morphism $s \in \mathbf{MRel}(C \& A, B)$, there is exactly one morphism $\Lambda(s) \in \mathbf{MRel}(C, A \Rightarrow B)$ such that

$$\mathrm{ev}_{AB} \circ (\Lambda(s) \times \mathrm{Id}_S) = s,$$

and this morphism is $\Lambda(s) = \{(p, (m, b)) \mid ((p, m), b) \in s\}$. □

Theorem 6.4. The category **MRel** is a Cartesian closed differential category.

Proof. By Theorem 6.3, **MRel** is Cartesian closed. Moreover, it is Cartesian closed left-additive since every homset $\mathbf{MRel}(A, B)$ can be endowed with the additive structure $(\mathbf{MRel}(A, B), \cup, \emptyset)$.

Finally, given $f \in \mathbf{MRel}(A, B)$, we can define its derivative as follows:

$$D(f) = \{(([\alpha], m), \beta) \mid (m \uplus [\alpha], \beta) \in f\} \in \mathbf{MRel}(A \& A, B).$$

It is not difficult to check that $D(-)$ satisfies (D1–7). We will now show that (D-Curry) also holds. Let $f \subseteq (\mathcal{M}_f(C) \times \mathcal{M}_f(A)) \times B$. On one side we have

$$D(\Lambda(f)) = \{(([\gamma], m_1), (m_2, \beta)) \mid ((m_1 \uplus [\gamma], m_2), \beta) \in f\}.$$

And on the other side we have $D(f) = f_1 \cup f_2$, where

$$f_1 = \{((([\gamma], []), (m_1, m_2)), \beta) \mid ((m_1 \uplus [\gamma], m_2), \beta) \in f\}$$
$$f_2 = \{((([], [\alpha]), (m_1, m_2)), \beta) \mid ((m_1, m_2 \uplus [\alpha]), \beta) \in f\}.$$

Since **MRel** is left-additive, we have

$$(f_1 \cup f_2) \circ \langle \pi_1 \times 0, \pi_2 \times \mathrm{Id}\rangle = (f_1 \circ \langle \pi_1 \times 0, \pi_2 \times \mathrm{Id}\rangle) \cup (f_2 \circ \langle \pi_1 \times 0, \pi_2 \times \mathrm{Id}\rangle).$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Easy calculations then give

$$f_1 \circ \langle \pi_1 \times 0, \pi_2 \times \mathrm{Id} \rangle = \{(((([\gamma], m_1), m_2), \beta)) \mid ((m_1 \uplus [\gamma], m_2), \beta) \in f\}$$
$$f_2 \circ \langle \pi_1 \times 0, \pi_2 \times \mathrm{Id} \rangle = \varnothing.$$

We then get

$$\Lambda(D(f) \circ \langle \pi_1 \times 0, \pi_2 \times \mathrm{Id} \rangle) = \Lambda(f_1 \circ \langle \pi_1 \times 0, \pi_2 \times \mathrm{Id} \rangle)$$
$$= D(\Lambda(f)).$$

□

The operator $\star$ can be defined directly in **MRel** as follows:

$$f \star g = \{((m_1 \uplus m_2, m), \beta) \mid (m_1, \alpha) \in g,\ ((m_2, m \uplus [\alpha]), \beta) \in f\} \in \mathbf{MRel}(C \& A, B).$$

We now provide a characterisation of the linear morphisms of **MRel**.

Lemma 6.1. A morphism $f \in \mathbf{MRel}(A, B)$ is linear if and only if for all $(m, \beta) \in f$, we have that m is a singleton.

Proof. Easy calculations give $f \circ \pi_1 = \{((m, []), \beta) \mid (m, \beta) \in f\}$. This is equal to $D(f)$ if and only if m is a singleton. □

Corollary 6.5. In **MRel** every isomorphism is linear.

Proof. Let $f \in \mathbf{MRel}(B, A)$ and $g \in \mathbf{MRel}(A, B)$ be such that $f \circ g = \mathrm{Id}_A$ and $g \circ f = \mathrm{Id}_B$. Notice that f does not contain any pair $([], \alpha)$ since otherwise such a pair would also appear in $f \circ g$, and this is impossible since $f \circ g = \mathrm{Id}$. Similarly, g cannot contain any pair $([], \beta)$. Hence,

$$f \circ g = \{([\alpha], \alpha) \mid \exists \beta \in B\ ([\alpha], \beta) \in g \text{ and } ([\beta], \alpha) \in f\}.$$

Since by hypothesis $f \circ g = \{([\alpha], \alpha) \mid \alpha \in A\}$, we have that for all $\alpha \in A$ there is a $\beta \in B$ such that $([\beta], \alpha) \in f$. Suppose now, in order to show a contradiction, that there is a $([\alpha_1, \ldots, \alpha_k], \beta) \in g$ such that $k > 1$. From the property above there are $\beta_1, \ldots, \beta_k \in B$ such that $([\beta_i], \alpha_i) \in f$ for $1 \leqslant i \leqslant k$, so we would have $([\beta_1, \ldots, \beta_k], \beta) \in f \circ g = \mathrm{Id}_B$, which is impossible. We can then conclude by Lemma 6.1 that g is linear. Analogous considerations show that f is linear too. □

6.1.1. *An extensional relational model.* In this section we build a reflexive object $\mathscr{D}$ in **MRel** that is extensional by construction, and hence linear by Corollary 6.5. We first give some preliminary definitions.

Recall that $\mathbf{N}$ denotes the set of natural numbers. An $\mathbf{N}$-indexed sequence $\sigma = (m_1, m_2, \ldots)$ of multisets is *quasi-finite* if $m_i = []$ for all but a finite number of indices i. If A is a set, we use $\mathcal{M}_f(A)^{(\omega)}$ to denote the set of all quasi-finite $\mathbf{N}$-indexed sequences of finite multisets over A. Notice that the only inhabitant of $\mathcal{M}_f(\varnothing)^{(\omega)}$ is the sequence $([], [], [], \ldots)$.

We now define a family of sets $\{D_n\}_{n\in\mathbf{N}}$ as follows:

$$D_0 = \varnothing$$
$$D_{n+1} = \mathcal{M}_f(D_n)^{(\omega)}$$

Since the operation $A \mapsto \mathcal{M}_f(A)^{(\omega)}$ is monotonic on sets, and since $D_0 \subseteq D_1$, we have $D_n \subseteq D_{n+1}$ for all $n \in \mathbf{N}$. Finally, we set $D = \cup_{n\in\mathbf{N}} D_n$.

So we have $D_0 = \varnothing$ and $D_1 = \{([], [], \ldots)\}$. The elements of D_2 are quasi-finite sequences of multisets over a singleton, that is, quasi-finite sequences of natural numbers. More generally, an element of D can be represented as a finite tree that alternates two kinds of layers:

— ordered nodes (the quasi-finite sequences), where immediate subtrees are indexed by distinct natural numbers;

— unordered nodes where subtrees are organised in a *non-empty* multiset.

In order to define an isomorphism in **MRel** between D and $(D \Rightarrow D) = \mathcal{M}_f(D) \times D$, it is enough to note that every element $\sigma = (\sigma_0, \sigma_1, \sigma_2 \ldots) \in D$ is canonically associated with the pair $(\sigma_0, (\sigma_1, \sigma_2, \ldots))$ and *vice versa*. Given $\sigma \in D$ and $m \in \mathcal{M}_f(D)$, we write $m :: \sigma$ for the element $\tau = (\tau_1, \tau_2, \ldots) \in D$ such that $\tau_1 = m$ and $\tau_{i+1} = \sigma_i$. This defines a bijection between $\mathcal{M}_f(D) \times D$ and D, and hence an isomorphism in **MRel** as follows.

Proposition 6.6. The triple $\mathscr{D} = (D, \mathcal{A}, \lambda)$ where:

$$\lambda = \{([(m, \sigma)], m :: \sigma) \mid m \in \mathcal{M}_f(D), \sigma \in D\} \in \mathbf{MRel}(D \Rightarrow D, D)$$
$$\mathcal{A} = \{([m :: \sigma], (m, \sigma)) \mid m \in \mathcal{M}_f(D), \sigma \in D\} \in \mathbf{MRel}(D, D \Rightarrow D)$$

is an extensional categorical model of differential λ-calculus.

Proof. It is trivial that $\lambda \circ \mathcal{A} = \mathrm{Id}_D$ and $\mathcal{A} \circ \lambda = \mathrm{Id}_{D\Rightarrow D}$. We can then conclude by Corollary 6.5. □

6.2. *Interpreting the differential lambda calculus in $\mathscr{D}$*

In Section 5, we defined the interpretation of a differential λ-term in any linear reflexive object of a Cartesian closed differential category. We will now give the result of the corresponding computation when it is performed in $\mathscr{D}$.

Given a differential λ-term S and a sequence $\vec{x} = x_1, \ldots, x_n$ adequate for S, the interpretation $[\![S]\!]_{\vec{x}}$ is an element of $\mathbf{MRel}(D^{\vec{x}}, D)$, that is, $[\![S]\!]_{\vec{x}} \subseteq \mathcal{M}_f(D)^n \times D$. The

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

interpretation $[\![S]\!]_{\vec{x}}$ is defined by structural induction on S as follows:

$$[\![x_i]\!]_{\vec{x}} = \{(([],\ldots,[],[\sigma],[],\ldots,[]),\sigma) \mid \sigma \in D\} \qquad \text{(where the only non-empty multiset occurs in the } i\text{th position)}$$

$$\begin{aligned}[\![sT]\!]_{\vec{x}} = \{((m_1,\ldots,m_n),\sigma) \mid\ & \exists k \in \mathbf{N},\\ & \exists(m_1^j,\ldots,m_n^j) \in \mathcal{M}_f(D)^n, \text{ for } j=0,\ldots,k,\\ & \exists \sigma_1,\ldots,\sigma_k \in D\\ & \text{such that}\\ & m_i = m_i^0 \uplus \ldots \uplus m_i^k, \text{ for } i=1,\ldots,n,\\ & ((m_1^0,\ldots,m_n^0),[\sigma_1,\ldots,\sigma_k] :: \sigma) \in [\![s]\!]_{\vec{x}},\\ & ((m_1^j,\ldots,m_n^j),\sigma_j) \in [\![T]\!]_{\vec{x}}, \text{ for } j=1,\ldots,k\}\end{aligned}$$

$$[\![\lambda z.s]\!]_{\vec{x}} = \{((m_1,\ldots,m_n), m :: \sigma) \mid ((m_1,\ldots,m_n,m),\sigma) \in [\![s]\!]_{\vec{x},z}\}$$

(where we assume that z does not occur in $\vec{x}$)

$$\begin{aligned}[\![\mathsf{D}^1 s \cdot (t)]\!]_{\vec{x}} = \{((m_1 \uplus m_1',\ldots,m_n \uplus m_n'), m :: \beta) \mid\ & \exists \alpha \in D\ ((m_1,\ldots,m_n),\alpha) \in [\![t]\!]_{\vec{x}}\\ & \text{and}\\ & ((m_1',\ldots,m_n'), m \uplus [\alpha] :: \beta) \in [\![s]\!]_{\vec{x}}\},\end{aligned}$$

$$\begin{aligned}[\![\mathsf{D}^{n+1} s \cdot (t_1,\ldots,t_{n+1})]\!]_{\vec{x}} = \{((m_1 \uplus m_1',\ldots,m_n \uplus m_n'), m :: \beta) \mid\ & \exists \alpha \in D\ ((m_1,\ldots,m_n),\alpha) \in [\![t_{n+1}]\!]_{\vec{x}}\\ & \text{and}\\ & ((m_1',\ldots,m_n'), m \uplus [\alpha] :: \beta) \in [\![\mathsf{D}^n s \cdot (t_1,\ldots,t_n)]\!]_{\vec{x}}\}\end{aligned}$$

$$[\![0]\!]_{\vec{x}} = \varnothing$$

$$[\![s + S]\!]_{\vec{x}} = [\![s]\!]_{\vec{x}} \cup [\![S]\!]_{\vec{x}}.$$

Note that if S is a *closed* differential λ-term, then $[\![S]\!] \subseteq D$. Moreover, it is easy to check that $[\![\Omega]\!] = \varnothing$ (in fact, we know from Manzonetto (2009) that the interpretation of all unsolvable standard λ-terms is empty). In the next section we will prove some general properties of $\mathrm{Th}(\mathscr{D})$.

6.3. *An extensional model of Taylor expansion*

Manzonetto (2009) characterised the equational theory of $\mathscr{D}$, viewed as a model of the untyped λ-calculus. More precisely, we proved that $\mathrm{Th}(\mathscr{D}) = \mathcal{H}^\star$, the theory equating two λ-terms M, N whenever they behave in the same way in every context. This is not surprising since Ehrhard (2009) proved that the continuous semantics (Scott 1972) can be seen as the extensional collapse of the category **MRel** and that $\mathscr{D}$ corresponds to Scott's $\mathscr{D}_\infty$ under this collapse.

In this section we give a partial characterisation of the theory of $\mathscr{D}$ viewed as a model of the differential λ-calculus.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Remark 6.7. Given an arbitrary set I and an I-indexed family of relations $\{f_i\}_{i\in I}$ from $\mathcal{M}_f(A)$ to B, we have $\cup_{i\in I} f_i \subseteq \mathcal{M}_f(A) \times B$. In particular, **MRel** has countable sums.

Proposition 6.8. **MRel** models the Taylor expansion.

Proof. Let $f \subseteq \mathcal{M}_f(C) \times (\mathcal{M}_f(A) \times B)$ and $g \subseteq \mathcal{M}_f(C) \times A$. Easy calculations give

$$\begin{aligned}
\mathrm{ev} \circ \langle f, g\rangle = \ & \{(m,\gamma) \mid \exists k \in \mathbf{N}, \\
& \qquad \exists m_j \in \mathcal{M}_f(C),\ \text{for } j = 0,\dots,k, \\
& \qquad \exists \alpha_1,\dots,\alpha_k \in A \\
& \qquad \text{such that} \\
& \qquad m = m_0 \uplus \dots \uplus m_k,\ \text{for } i = 1,\dots,n, \\
& \qquad (m_0, ([\alpha_1,\dots,\alpha_k],\gamma)) \in f, \\
& \qquad (m_j,\alpha_j) \in g,\ \text{for } j = 1,\dots,k\} \\
= \ & \textstyle\bigcup_{k\in\mathbf{N}} \{(m,\gamma) \mid \exists m_j \in \mathcal{M}_f(C),\ \text{for } j = 0,\dots,k, \\
& \qquad \exists \alpha_1,\dots,\alpha_k \in A \\
& \qquad \text{such that} \\
& \qquad m = m_0 \uplus \dots \uplus m_k,\ \text{for } i = 1,\dots,n, \\
& \qquad (m_0, ([\alpha_1,\dots,\alpha_k],\gamma)) \in f, \\
& \qquad (m_j,\alpha_j) \in g,\ \text{for } j = 1,\dots,k\} \\
= \ & \textstyle\sum_{k\in\mathbf{N}} ((\cdots(\Lambda^-(f) \underbrace{\star g)\cdots) \star g)}_{k \text{ times}} \circ \langle \mathrm{Id}_A, \varnothing\rangle\,.
\end{aligned}$$

□

Corollary 6.9. Every categorical model $\mathscr{U}$ of the differential λ-calculus living in **MRel** satisfies $\mathcal{E} \subseteq \mathrm{Th}(\mathscr{U})$.

Another easy corollary is that the relational semantics is *incomplete*. Recall that a semantics **C** is said to be *complete* if for all differential λ-theories $\mathcal{T}$ there is a model $\mathscr{U}$ living in **C** such that $\mathrm{Th}(\mathscr{U}) = \mathcal{T}$. As we know that in **MRel** only theories including $\mathcal{E}$ are representable, it follows that no non-trivial recursively enumerable† differential λ-theory is representable in **MRel**, and since there exists a continuum of recursively enumerable differential λ-theories, we get the following result.

Corollary 6.10. The relational semantics is hugely incomplete: there are $2^{\aleph_0}$ differential λ-theories that are not representable in **MRel**.

From Corollary 6.9 we get the following (partial) characterisation of $\mathrm{Th}(\mathscr{D})$.

Corollary 6.11. The theory of $\mathscr{D}$ includes both $\lambda\beta\eta^d$ and $\mathcal{E}$.

These preliminary results and the work in Bucciarelli *et al.* (2011) lead us to the following conjecture.

† A differential λ-theory $\mathcal{T}$ is *recursively enumerable* if the $\mathcal{T}$-equivalence class of every differential λ-term is; it is said to be *trivial* if it equates all differential λ-terms.

Conjecture 1. We conjecture that

$$\mathrm{Th}(\mathscr{D}) = \{(S,T) \in \Lambda^d \times \Lambda^d \mid \text{for all contexts } C(\cdot),\ C(S) \text{ is solvable iff } C(T) \text{ is solvable}\},$$

where a context is a differential λ-term with a hole denoted by $(\cdot)$, and $C(S)$ denotes the result of substituting S (possibly with capture of variables) for the hole in C. 'Solvable' here is to be understood as *may-solvable*[†] (that is, a sum of terms converges if at least one of its components converges).

A complete syntactical characterisation of the theory of $\mathscr{D}$ is difficult to provide, and is reserved for future work.

6.3.1. *A differentially extensional but non-extensional relational model.* In this section we briefly present an example of a model $\mathscr{E}$ in the category **MRel** satisfying the axiom (η_∂) but not the axiom (η). This model, whose construction is similar to that of $\mathscr{D}$, was first introduced in Hyland *et al.* (2006) and was studied by de Carvalho in his Ph.D. thesis, where it was presented as a type system called *System R* (de Carvalho 2007, §6.3.3).

We fix a non-empty set A of 'atoms' that does not contain any pairs, and define a family of sets $\{E_n\}_{n\in\mathbf{N}}$ as follows:

$$E_0 = \varnothing$$
$$E_{n+1} = (\mathcal{M}_f(E_n) \times E_n) \cup A.$$

Finally, we set $E = \cup_{n\in\mathbf{N}} E_n$ and $\mathscr{E} = (E, \mathcal{A}, \lambda)$ where $\mathcal{A}, \lambda$ are the obvious morphisms performing the retraction $(E \Rightarrow E) \lhd E$.

Remark 6.12. It is easy to verify that $\mathscr{E}$ is linear, and thus a model of the differential λ-calculus, and it is non-extensional because the atoms in $A \subseteq E$ cannot be sent injectively into $\mathcal{M}_f(E) \times E$.

As remarked in Hyland *et al.* (2006), the model $\mathscr{E}$ is a relational analogue of Engeler's graph-model (Engeler 1981) in much the same way as $\mathscr{D}$ is the analogue of Scott's $\mathscr{D}_\infty$. The interpretation of a differential λ-term S in $\mathscr{E}$ is defined as usual and gives, up to isomorphism, a subset $[\![S]\!]_{\vec{x}} \subseteq \mathcal{M}_f(E)^n \times E$.

Lemma 6.2. The model $\mathscr{E}$ is differentially extensional.

Proof. In $\mathscr{E}$ the interpretation of the linear application does not contain any atoms, in the sense that $(\vec{m}, \alpha) \in [\![\mathsf{D}S \cdot T]\!]_{\vec{x}}$ entails $\alpha = (m', \beta)$. Presenting the model as a type system, (m', β) would be an arrow type $m' \to \beta$. This guarantees that the η_∂-expansion does not modify the interpretation. □

[†] May and must solvability were investigated in Pagani and Ronchi Della Rocca (2010) in the context of the resource calculus.

7. The resource calculus

In this section we present the resource calculus (Boudol 1993; Boudol and Curien and Lavatelli), using the formalisation *à la Tranquilli* given in Pagani and Tranquilli (2009), and show that every model of the differential λ-calculus is also a model of the resource calculus. We then discuss the (tight) relationship existing between the differential λ-calculus and the resource calculus.

7.1. *Syntax*

The resource calculus has three syntactical categories: *resource λ-terms* (Λ^r), which are in functional position; *bags* (Λ^b), which are in argument position and represent multisets of resources; and *sums*, which represent the possible results of a computation. A *resource* ($\Lambda^{(!)}$) can be linear or reusable, and in the latter case it is written with a ! superscript. An *expression* (Λ^e) is either a term or a bag.

Formally, we have the following grammar:

$$\begin{array}{llcll} \Lambda^r : & M,N,L & ::= & x \mid \lambda x.M \mid MP & \text{(resource } \lambda\text{-terms)} \\ \Lambda^{(!)} : & M^{(!)},N^{(!)} & ::= & M \mid M^! & \text{(resources)} \\ \Lambda^b : & P,Q,R & ::= & [M_1^{(!)},\dots,M_n^{(!)}] & \text{(bags)} \\ \Lambda^e : & A,B & ::= & M \mid P & \text{(expressions)} \end{array}$$

From now on, resource λ-terms are considered up to α-conversion and permutation of the resources in the bags. Intuitively, linear resources are available exactly once, while reusable resources can be used zero or many times.

Definition 7.1. Given an expression $A \in \Lambda^e$, the set $\mathrm{FV}(A)$ of *free variables of A* is defined by induction on A as follows:

$$\begin{aligned} \mathrm{FV}(x) &= \{x\} \\ \mathrm{FV}(\lambda x.M) &= \mathrm{FV}(M) - \{x\} \\ \mathrm{FV}(MP) &= \mathrm{FV}(M) \cup \mathrm{FV}(P) \\ \mathrm{FV}([]) &= \varnothing \\ \mathrm{FV}([M^{(!)}] \uplus P) &= \mathrm{FV}(M) \cup \mathrm{FV}(P). \end{aligned}$$

Given expressions $A_1,\dots,A_k$, we set

$$\mathrm{FV}(A_1,\dots,A_k) = \mathrm{FV}(A_1) \cup \dots \cup \mathrm{FV}(A_k).$$

For sums, we use $\mathbf{N}\langle\Lambda^r\rangle$ (respectively, $\mathbf{N}\langle\Lambda^b\rangle$) to denote the set of finite formal sums of terms (respectively, bags). As usual, we suppose that the sum is commutative and associative, and that 0 is its neutral element.

$$\mathbb{M},\mathbb{N} \in \mathbf{N}\langle\Lambda^r\rangle \qquad \mathbb{P},\mathbb{Q} \in \mathbf{N}\langle\Lambda^b\rangle \qquad \mathbb{A},\mathbb{B},\mathbb{C} \in \mathbf{N}\langle\Lambda^e\rangle = \mathbf{N}\langle\Lambda^r\rangle \cup \mathbf{N}\langle\Lambda^b\rangle \qquad \text{(sums)}$$

Note that in writing $\mathbf{N}\langle\Lambda^e\rangle$ we are abusing the notation as it does not denote the $\mathbf{N}$-module generated over $\Lambda^e = \Lambda^r \cup \Lambda^b$ but rather the union of the two $\mathbf{N}$-modules. In other words, sums must only be taken in the same sort.

The definition of $\mathrm{FV}(\cdot)$ is extended to elements of $\mathbf{N}\langle\Lambda^e\rangle$ in the obvious way.

Sums do not appear in the grammar for resource λ-terms, bags and expressions – indeed they may only arise on the 'surface' in this calculus (while in the differential λ-calculus, sums may appear in the right-hand argument of an application). Nevertheless, as a syntactic sugar and not as actual syntax, we extend all the constructors to sums as follows.

Notation 7.2. We use the following abbreviations with $\mathbf{N}\langle\Lambda^e\rangle$:

$$\lambda x.\sum_{i=1}^{k} M_i = \sum_{i=1}^{k} \lambda x.M_i$$

$$\left(\sum_{i=1}^{k} M_i\right)\left(\sum_{j=1}^{n} P_j\right) = \left(\sum_{i,j} M_i P_j\right)$$

$$\left[\left(\sum_{i=1}^{k} M_i\right)\right] \uplus P = \sum_{i=1}^{k} [M_i] \uplus P$$

$$\left[\left(\sum_{i=1}^{k} M_i\right)^!\right] \uplus P = [M_1^!, \ldots, M_k^!] \uplus P.$$

These equalities make sense since all constructors, apart from $(\cdot)^!$, are linear. Note the difference between these rules and the analogous ones for the differential λ-calculus introduced in Notation 3.4. In the differential λ-calculus, the application operator is only linear in its left component, while here it is bilinear.

The 0-ary version of the above equalities give us

$$\begin{aligned} \lambda x.0 &= 0 \\ \mathbb{M}0 &= 0 \\ 0\mathbb{P} &= 0 \\ [0] \uplus P &= 0 \\ [0^!] \uplus P &= P \\ 0 \uplus P &= 0, \end{aligned}$$

so 0 annihilates everything except when it lies under a $(\cdot)^!$.

Definition 7.3. Let A be an expression and N be a resource λ-term.

— $A\{N/x\}$ is the usual substitution of N for x in A. It is extended to sums as in $\mathbb{A}\{\mathbb{N}/x\}$ by linearity† in $\mathbb{A}$, and using Notation 7.2 for $\mathbb{N}$.

† A unary operator $F(\cdot)$ is extended by linearity by setting $F(\Sigma_i A_i) = \Sigma_i F(A_i)$.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

— $A\langle N/x\rangle$ is the *linear substitution* defined inductively as follows:

$$y\langle N/x\rangle = \begin{cases} N & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$
$$[M]\langle N/x\rangle = [M\langle N/x\rangle]$$
$$[M^!]\langle N/x\rangle = [M\langle N/x\rangle, M^!]$$
$$(\lambda y.M)\langle N/x\rangle = \lambda y.M\langle N/x\rangle$$
$$(MP)\langle N/x\rangle = M\langle N/x\rangle P + M(P\langle N/x\rangle)$$
$$[]\langle N/x\rangle = 0$$
$$(P \uplus R)\langle N/x\rangle = P\langle N/x\rangle \uplus R + P \uplus R\langle N/x\rangle$$

This is extended to $\mathbb{A}\langle \mathbb{N}/x\rangle$ by bilinearity† in both $\mathbb{A}$ and $\mathbb{N}$.

The operation $M\langle N/x\rangle$ on resource λ-terms is roughly equivalent to the operation

$$\frac{\partial S}{\partial x} \cdot T$$

on differential λ-terms (*cf.* Lemma 7.1). Notice that in defining $[M^!]\langle N/x\rangle$, we essentially extract a linear copy of M from the infinitely many represented by $M^!$, which then receives the substitution, and leave the others unchanged.

Example 7.4.

(1)
$$x\langle M/x\rangle = M$$
$$y\langle M/x\rangle = 0.$$

(2)
$$(x[x])\langle M+N/x\rangle = (M+N)[x] + x[M+N] = M[x] + N[x] + x[M] + x[N].$$

(3)
$$(x[x^!])\langle M+N/x\rangle = (M+N)[x^!] + x[(M+N), x^!] = M[x^!] + N[x^!] + x[M, x^!] + x[N, x^!].$$

(4)
$$(x[x^!])\{M+N/x\} = (M+N)[(M+N)^!] = M[M^!, N^!] + N[M^!, N^!].$$

As notation, we will write $\vec{L}$ for $L_1, \ldots, L_k$ and $\vec{N}^!$ for $N_1^!, \ldots, N_n^!$. We will also abbreviate $M\langle L_1/x\rangle \cdots \langle L_k/x\rangle$ as $M\left\langle \vec{L}/x\right\rangle$. Moreover, given a sequence $\vec{L}$ and an index $1 \leqslant i \leqslant k$, we will write $\vec{L}_{-i}$ for $L_1, \ldots, L_{i-1}, L_{i+1}, \ldots, L_k$.

Remark 7.5. Every applicative resource λ-term MP can be written in a unique way as $M[\vec{L}, \vec{N}^!]$.

† A binary operator $F(\cdot,\cdot)$ is extended by bilinearity by setting $F(\Sigma_i A_i, \Sigma_j B_j) = \Sigma_{i,j} F(A_i, B_j)$.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

7.2. *Resource lambda theories*

We now define the equational theories of the resource calculus, namely the *resource λ-theories*. To begin, we will present the main axiom associated with this calculus (for $\vec{L} = L_1, \ldots, L_k$ and $\vec{N} = N_1, \ldots, N_n$):

$$(\beta^r) \qquad (\lambda x.M)[\vec{L}, \vec{N}^!] = M \left\langle \vec{L}/x \right\rangle \{\Sigma_{i=1}^n N_i/x\}$$

Note that when $n = 0$, this rule becomes $(\lambda x.M)[\vec{L}] = M \left\langle \vec{L}/x \right\rangle \{0/x\}$. Once oriented from left to right, the (β^r)-conversion expresses how to calculate a function $\lambda x.M$ applied to a bag containing linear resources $\vec{L}$ and reusable resources $\vec{N}$.

Remark 7.6. The left-to-right oriented version of (β^r) corresponds to *giant-step* reduction in the terminology of Pagani and Tranquilli (2009). In the same paper, the authors also consider a *baby-step* reduction rule. They prove that both reductions are confluent and that every giant-step can be emulated by several baby-steps. For our purposes we can consider the rule (β^r) without loss of generality because both reductions generate the same equational theory.

In the resource calculus, the axiom equating all resource λ-terms having the same extensional behaviour has the form

$$(\eta^r) \qquad \lambda x.M[x^!] = M, \text{ where } x \notin \mathrm{FV}(M).$$

In this context, the axiom (η_∂) of the differential λ-calculus has no analogue since the application of a resource λ-term to a bag essentially corresponds to a sequence of linear applications always followed by a classic application (see Definition 7.10, below). Therefore, the linear application where (η_∂) should act is hidden.

The resource calculus can be seen as a proper extension of the classical λ-calculus.

Remark 7.7. The classical λ-calculus can be easily injected within the resource calculus. Indeed, given an ordinary λ-term M, it is sufficient to translate every subterm of M of the form PQ into $P[Q^!]$. In this restricted system, the rules (β^r) and (η^r) are completely equivalent to the classic (β) and (η)-conversions, respectively.

We can now define the equational theories associated with this calculus, namely, the *resource λ-theories*.

A *λ^r-relation* $\mathcal{R}$ is any set of equations between sums of resource λ-terms (or bags). Thus $\mathcal{R}$ can be thought of as a binary relation on $\mathbf{N}\langle\Lambda^e\rangle$.

A λ^r-relation $\mathcal{R}$ is said to be:

— an *equivalence* if it is closed under the following rules (for all $\mathbb{A}, \mathbb{B}, \mathbb{C} \in \mathbf{N}\langle\Lambda^e\rangle$):

$$\frac{}{\mathbb{A} = \mathbb{A}} \text{ (reflexivity)}$$

$$\frac{\mathbb{B} = \mathbb{A}}{\mathbb{A} = \mathbb{B}} \text{ (symmetry)}$$

$$\frac{\mathbb{A} = \mathbb{B} \quad \mathbb{B} = \mathbb{C}}{\mathbb{A} = \mathbb{C}} \text{ (transitivity)}$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

— *compatible* if it is closed under the following structural rules (for all $\mathbb{M}, \mathbb{N}, \mathbb{M}_i, \mathbb{N}_i \in \mathbf{N}\langle\Lambda^r\rangle$ and $\mathbb{P}, \mathbb{Q} \in \mathbf{N}\langle\Lambda^b\rangle$):

$$\frac{\mathbb{M} = \mathbb{N}}{\lambda x.\mathbb{M} = \lambda x.\mathbb{N}} \text{ (lambda)}$$

$$\frac{\mathbb{M} = \mathbb{N} \quad \mathbb{Q} = \mathbb{P}}{\mathbb{M}\mathbb{P} = \mathbb{N}\mathbb{Q}} \text{ (app)}$$

$$\frac{\mathbb{M} = \mathbb{N} \quad \mathbb{P} = \mathbb{Q}}{[\mathbb{M}^{(!)}] \uplus \mathbb{P} = [\mathbb{N}^{(!)}] \uplus \mathbb{Q}} \text{ (bag)}$$

$$\frac{\mathbb{M}_i = \mathbb{N}_i \quad \text{for all } 1 \leqslant i \leqslant n}{\sum_{i=1}^n \mathbb{M}_i = \sum_{i=1}^n \mathbb{N}_i} \text{ (sum)}$$

As notation, we will write $\mathcal{R} \vdash \mathbb{M} = \mathbb{N}$ or $\mathbb{M} =_{\mathcal{R}} \mathbb{N}$ for $\mathbb{M} = \mathbb{N} \in \mathcal{R}$.

Definition 7.8. A *resource λ-theory* is any compatible λ^r-relation $\mathcal{R}$ that is an equivalence relation and includes (β^r). $\mathcal{R}$ is said to be *extensional* if it also contains (η^r). We say that $\mathcal{R}$ *satisfies sum idempotency* whenever $\mathcal{R} \vdash M + M = M$.

We use $\lambda\beta^r$ (respectively, $\lambda\beta\eta^r$) to denote the minimum resource λ-theory (respectively, the minimum extensional resource λ-theory).

Example 7.9.

(1)
$$\begin{aligned} \lambda\beta^r &\vdash (\lambda x.x[x])[\mathbf{I}] = 0 \\ \lambda\beta^r &\vdash (\lambda x.x[x])[\mathbf{I},\mathbf{I}] = \mathbf{I} \\ \lambda\beta^r &\vdash (\lambda x.x[x])[\mathbf{I},\mathbf{I},\mathbf{I}] = 0. \end{aligned}$$

(2)
$$\lambda\beta^r \vdash (\lambda x.x[x])[M,N] = M[N] + N[M].$$

(3)
$$\begin{aligned} \lambda\beta^r \vdash (\lambda x.x[x,x])[(\lambda y.y[y^!])^!] &= (\lambda x.x[x^!])[\lambda y.y[y^!], \lambda z.z[z^!]] \\ &= 2(\lambda y.y[y^!])[(\lambda z.z[z^!])^!]. \end{aligned}$$

(4)
$$\begin{aligned} \lambda\beta\eta^r \vdash (\lambda xz.y[y][z^!])[] &= \lambda z.y[y][z^!] \\ &= y[y]. \end{aligned}$$

7.3. *From the resource to the differential lambda calculus…*

In this section we show that every linear reflexive object living in a Cartesian closed differential category is also a sound model of the untyped resource calculus. This result is achieved by first translating the resource calculus in the differential λ-calculus, and then applying the machinery of Section 5.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Definition 7.10. The resource calculus can be easily translated into the differential λ-calculus as follows:

$$x^d = x$$
$$(\lambda x.M)^d = \lambda x.M^d$$
$$\left(M\left[L_1,\ldots,L_k,N_1^!,\ldots,N_n^!\right]\right)^d = \left(\mathsf{D}^k M^d\cdot(L_1^d,\ldots,L_k^d)\right)\left(\sum_{i=1}^n N_i^d\right).$$

The translation is then extended to elements in $\mathbf{N}\langle\Lambda^r\rangle$ by setting

$$\left(\sum_{i=1}^n M_i\right)^d = \sum_{i=1}^n M_i^d.$$

The next lemma shows that this translation behaves well with respect to the differential and the usual substitution.

Lemma 7.1. Let $M, N \in \Lambda^r$ and x be a variable. Then:

(i) $$(M\langle N/x\rangle)^d = \frac{\partial M^d}{\partial x}\cdot N^d.$$

(ii) $$(M\{N/x\})^d = M^d\left\{N^d/x\right\}.$$

Proof.

(i) We use structural induction on M. The only difficult case is $M \equiv M'\left[\vec{L},\vec{N}^!\right]$. By the definitions of $(-)^d$ and linear substitution, we have

$$\left(\left(M'\left[\vec{L},\vec{N}^!\right]\right)\langle N/x\rangle\right)^d = \left(M'\langle N/x\rangle\left[\vec{L},\vec{N}^!\right]\right)^d + \left(M'\left(\left[\vec{L},\vec{N}^!\right]\langle N/x\rangle\right)\right)^d$$

(a) $$= \left(M'\langle N/x\rangle\left[\vec{L},\vec{N}^!\right]\right)^d +$$

(b) $$\left(\sum_{j=1}^k M'\left[L_j\langle N/x\rangle,\vec{L}_{-j},\vec{N}^!\right]\right)^d +$$

(c) $$\left(\sum_{i=1}^n M'\left[N_i\langle N/x\rangle,\vec{L},\vec{N}^!\right]\right)^d.$$

We consider the three summands separately.

(a) By the definition of $(-)^d$, we have

$$\left(M'\langle N/x\rangle\left[\vec{L},\vec{N}^!\right]\right)^d = \left(\mathsf{D}^k(M'\langle N/x\rangle)^d\cdot\left(\vec{L}^d\right)\right)\left(\sum_{i=1}^n N_i^d\right).$$

Applying the induction hypothesis, this is equal to

$$\left(\mathsf{D}^k\left(\frac{\partial(M')^d}{\partial x}\cdot N^d\right)\cdot\left(\vec{L}^d\right)\right)\left(\sum_{i=1}^n N_i^d\right).$$

(b) By the definition of the translation map $(-)^d$, we have

$$\left(\sum_{j=1}^k M'\left[L_j\langle N/x\rangle,\vec{L}_{-j},\vec{N}^!\right]\right)^d$$
$$=\sum_{j=1}^k\left(\mathsf{D}^{k-1}\left(\mathsf{D}\,(M')^d\cdot(L_j\langle N/x\rangle)^d\right)\cdot\left(\vec{L}^d_{-j}\right)\right)\left(\sum_{i=1}^n N_i^d\right).$$

Applying the induction hypothesis, this is equal to

$$\sum_{j=1}^k\left(\mathsf{D}^{k-1}\left(\mathsf{D}\,(M')^d\cdot\left(\frac{\partial L_j^d}{\partial x}\cdot N^d\right)\right)\cdot\left(\vec{L}^d_{-j}\right)\right)\left(\sum_{i=1}^n N_i^d\right).$$

(c) By the definition of $(-)^d$, we have

$$\left(\sum_{j=1}^n M'\left[N_j\langle N/x\rangle,\vec{L},\vec{N}^!\right]\right)^d$$
$$=\sum_{j=1}^n\left(M'\left[N_j\langle N/x\rangle,\vec{L},\vec{N}^!\right]\right)^d$$
$$=\sum_{j=1}^n\left(\mathsf{D}^k\left(\mathsf{D}\,(M')^d\cdot(N_j\langle N/x\rangle)^d\right)\cdot\left(\vec{L}^d\right)\right)\left(\sum_{i=1}^n N_i^d\right).$$

Applying the induction hypothesis, this is equal to

$$\sum_{j=1}^n\left(\mathsf{D}^k\left(\mathsf{D}\,(M')^d\cdot\left(\frac{\partial N_j^d}{\partial x}\cdot N^d\right)\right)\cdot\left(\vec{L}^d\right)\right)\left(\sum_{i=1}^n N_i^d\right),$$

and by permutative equality, this is equal to

$$\sum_{j=1}^n\left(\mathsf{D}\left(\mathsf{D}^k\,(M')^d\cdot\left(\vec{L}^d\right)\right)\cdot\left(\frac{\partial N_j^d}{\partial x}\cdot N^d\right)\right)\left(\sum_{i=1}^n N_i^d\right).$$

To conclude the proof it is sufficient to verify that

$$\frac{\partial}{\partial x}\left(\left(\mathsf{D}^k\,(M')^d\cdot\left(\vec{L}^d\right)\right)\left(\sum_{i=1}^n N_i^d\right)\right)\cdot N^d$$

is equal to the sum of (a), (b) and (c).

(ii) This part follows by a straightforward induction on M. □

The translation $(\cdot)^d$ is 'faithful' in the sense expressed by the next proposition.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Proposition 7.11. For all $M \in \Lambda^r$, we have $\lambda\beta^r \vdash M = N$ implies $\lambda\beta^d \vdash M^d = N^d$.

Proof. It is easy to check that the proposition holds for the contextual rules. Suppose then that $\lambda\beta^r \vdash M = N$ because

$$M \equiv (\lambda x.M') \left[\vec{L}, \vec{N}^!\right]$$
$$N \equiv M' \left\langle \vec{L}/x \right\rangle \left\{ \sum_{i=1}^{n} N_i/x \right\}.$$

By definition of the map $(-)^d$, we have

$$\left((\lambda x.M') \left[\vec{L}, \vec{N}^!\right]\right)^d = \left(\mathsf{D}^k \left(\lambda x.(M')^d\right) \cdot \left(\vec{L}^d\right)\right) \left(\sum_{i=1}^{n} N_i^d\right)$$
$$=_{\lambda\beta^d} \left(\lambda x. \frac{\partial^k \left(M'\right)^d}{\partial x, \ldots, x} \cdot (\vec{L}^d)\right) \left(\sum_{i=1}^{n} N_i^d\right)$$
$$=_{\lambda\beta^d} \left(\frac{\partial^k (M')^d}{\partial x, \ldots, x} \cdot (\vec{L}^d)\right) \left\{\sum_{i=1}^{n} N_i^d/x\right\},$$

which is equal to N^d by Lemma 7.1. □

Remark 7.12. The two results above generalise straightforwardly to sums of resource λ-terms (that is, to elements $\mathbb{M} \in \mathbf{N}\langle\Lambda^r\rangle$).

7.3.1. *Interpreting the resource calculus by translation.* Given a linear reflexive object $\mathscr{U}$ living in a Cartesian closed differential category **C**, it is possible to interpret resource λ-terms trough their translation $(-)^d$. Indeed, it is sufficient to set

$$[\![M]\!]_{\vec{x}} = [\![M^d]\!]_{\vec{x}} : U^n \to U.$$

From this fact, and Proposition 7.11 and Remark 7.12, it follows that $\mathscr{U}$ is a sound model of the untyped resource calculus.

Remark 7.13. If $\mathscr{U}$ is an extensional model of the differential λ-calculus, then it is also an extensional model of the resource calculus. Indeed

$$[\![(\lambda x.M[x^!])^d]\!]_{\vec{x}} = [\![\lambda x.M^d x]\!]_{\vec{x}} = [\![M^d]\!]_{\vec{x}}.$$

We can prove a completeness result for the resource calculus that is stronger than the one for the differential λ-calculus. More precisely, we can get rid of the hypothesis that the theory is differentially extensional. Indeed, for every resource λ-theory $\mathcal{R}$, the differential λ-theory $\mathcal{T}$ generated† by

$$\{S = T \mid \exists \mathbb{M}, \mathbb{N} \in \mathbf{N}\langle\Lambda^r\rangle\ S = M^d, T = N^d, \mathcal{R} \vdash M = N\}$$

† The differential λ-theory generated by a set E of equations is the smallest differential λ-theory including E.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

is such that $\mathcal{R} \vdash \mathbb{M} = \mathbb{N}$ if and only if $\mathcal{T} \vdash \mathbb{M}^d = \mathbb{N}^d$. When $\mathcal{R}$ satisfies sum idempotency, $\mathcal{T}$ also does, so we can apply the construction described in Section 5.4 and get a Cartesian closed differential category $\mathbf{C}_{\mathcal{T}}$ where $\mathbf{I}$ is a linear reflexive object. Then one can prove the following lemma, which is similar to Proposition 5.16 except that the axiom (η_∂) no longer plays a role since in the translation of the resource calculus, the linear application is always followed by a regular application, so the η_∂-expansion disappears by (β^r)-conversion.

Lemma 7.2. For every $\mathbb{M} \in \mathbf{N}\langle\Lambda^r\rangle$, we have (for some $z \notin \mathrm{FV}(\mathbb{M})$)

$$[\![\mathbb{M}]\!]_{\vec{x}} = \lambda z.\mathbb{M}^d \left\{\pi^{\vec{x}}_{x_1} z/x_1\right\} \cdots \left\{\pi^{\vec{x}}_{x_n} z/x_n\right\} : \mathbf{I}^{\vec{x}} \to \mathbf{I}.$$

Proof. The only interesting case is $\mathbb{M} = M[\vec{L}, \vec{N}^!]$. In the following, we use σ_r to denote the sequence of substitutions $\{\pi^{\vec{x}}_{x_1} r/x_1\} \cdots \{\pi^{\vec{x}}_{x_n} r/x_n\}$. We have

$$\begin{aligned}
[\![M[\vec{L}, \vec{N}^!]]\!]_{\vec{x}} &= \left[\!\!\left[\left(\mathsf{D}^k M^d \cdot \left(\vec{L}^d\right)\right) \left(\Sigma_i N_i^d\right) \right]\!\!\right]_{\vec{x}} && \text{(definition of } (\cdot)^d)\text{)} \\
&= \mathrm{ev} \circ \left\langle \mathbf{1} \circ [\![\left(\mathsf{D}^k M^d \cdot \left(\vec{L}^d\right)\right.]\!]_{\vec{x}} \right), [\![\Sigma_i N_i^d]\!]_{\vec{x}} \right\rangle && \text{(definition of } [\![\cdot]\!]_{\vec{x}}) \\
&= \mathrm{ev} \circ \left\langle \lambda ry. \left(\mathsf{D}^k \left([\![M^d]\!]_{\vec{x}} r\right) \cdot \left([\![\vec{L}]\!]_{\vec{x}} r\right)\right) y, \Sigma_i [\![N_i^d]\!]_{\vec{x}} \right\rangle && \text{(calculations)} \\
&= \mathrm{ev} \circ \left\langle \lambda ry. \left(\mathsf{D}^k \left(M^d \Sigma_r\right) \cdot \left(\vec{L}^d \Sigma_r\right)\right) y, \Sigma_i \lambda r. N_i^d \Sigma_r \right\rangle && \text{(induction hypothesis)} \\
&= \lambda z. \left(\left(\lambda ry. \left(\mathsf{D}^k \left(M^d \Sigma_r\right) \cdot \left(\vec{L}^d \Sigma_r\right)\right) y\right) z\right) \left(\Sigma_i \left(\lambda r. N_i^d \Sigma_r\right) z\right) && \text{(calculations)} \\
&= \lambda z. \left(\lambda y. \left(\mathsf{D}^k \left(M^d \Sigma_z\right) \cdot \left(\vec{L}^d \Sigma_z\right)\right) y\right) \left(\Sigma_i N_i^d \Sigma_z\right) && ((\beta^r)\text{-conversion}) \\
&= \lambda z. \left(\mathsf{D}^k \left(M^d \Sigma_z\right) \cdot \left(\vec{L}^d \Sigma_z\right)\right) \left(\Sigma_i N_i^d \Sigma_z\right) && ((\beta^r)\text{-conversion}) \\
&= \lambda z. \left(M[\vec{L}, \vec{N}^!]\right)^d \left\{\pi^{\vec{x}}_{x_1} z/x_1\right\} \cdots \left\{\pi^{\vec{x}}_{x_n} z/x_n\right\}. && \text{(definition of } (\cdot)^d)
\end{aligned}$$

□

As a corollary, we get equational completeness for the resource calculus.

Corollary 7.14 (Equational Completeness). Every resource λ-theory $\mathcal{R}$ satisfying sum idempotency is the theory of a linear reflexive object in a differential Cartesian closed category.

7.4. *And back…*

In this section we define a translation from the differential to the resource calculus. This translation is more tricky because in the differential λ-calculus the result of the linear application $\mathsf{D}(\lambda x.s) \cdot t$ mantains the lambda abstraction (since it waits for other arguments that may substitute the remaining occurrences of x in s), but the naively corresponding resource λ-term $(\lambda x.M)[N]$ erases it (since all other free occurrences of x in M are substituted by 0).

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Definition 7.15. The differential λ-calculus can be translated into the resource calculus as follows:

$$\begin{aligned} x^r &= x \\ (\lambda x.s)^r &= \lambda x.s^r \\ (sT)^r &= s^r[(T^r)^!] \\ (\mathsf{D}^k s\cdot(t_1,\dots,t_k))^r &= \lambda y.s^r[t_1^r,\dots,t_k^r,y^!] \qquad \text{(where } y \text{ is a fresh variable)} \\ (s+S)^r &= s^r + S^r. \end{aligned}$$

Notice that while the shape of the term $\lambda y.s^r[t_1^r,\dots,t_k^r,y^!]$ looks similar to an (η^r)-expansion of $s^r[t_1^r,\dots,t_k^r]$, it is not[†]! Indeed, in the (η^r)-axiom, $y^!$ is supposed to be in a singleton bag.

Lemma 7.3. Let $S,T\in\Lambda^d$ and x be a variable. Then:

(i) $$\left(\frac{\partial S}{\partial x}\cdot T\right)^r = S^r\langle T^r/x\rangle\,.$$

(ii) $$(S\{T/x\})^r = S^r\{T^r/x\}\,.$$

Proof.

(i) We use structural induction on S. If S is a variable, a lambda abstraction or a sum, the lemma follows directly from the induction hypothesis.

— Case $S\equiv \mathsf{D}^k s\cdot(t_1,\dots,t_k)$:
We have

$$\begin{aligned} &\left(\frac{\partial}{\partial x}(\mathsf{D}^k s\cdot(t_1,\dots,t_k))\cdot T\right)^r \\ &= \sum_{i=1}^k\left(\left(\mathsf{D}^k s\cdot\left(t_1,\dots,\frac{\partial t_i}{\partial x}\cdot T,\dots,t_k\right)\right)\right)^r \\ &\quad + \left(\left(\mathsf{D}^k\left(\frac{\partial s}{\partial x}\cdot T\right)\cdot(t_1,\dots,t_k)\right)\right)^r && \left(\text{definition of } \frac{\partial(\cdot)}{\partial x}\cdot T\right) \\ &= \sum_{i=1}^k \lambda y.s^r\left[t_1^r,\dots,\left(\frac{\partial t_i}{\partial x}\cdot T\right)^r,\dots,t_k^r,y^!\right] \\ &\quad + \lambda y.\left(\frac{\partial s}{\partial x}\cdot T\right)^r[t_1^r,\dots,t_k^r,y^!] && (\text{definition of } (\cdot)^r) \\ &= \sum_{i=1}^k \lambda y.s^r[t_1^r,\dots,t_i^r\langle T^r/x\rangle,\dots,t_k^r,y^!] \\ &\quad + \lambda y.(s^r\langle T^r/x\rangle)[t_1^r,\dots,t_k^r,y^!] && (\text{induction hypothesis}) \\ &= (\lambda y.s^r[t_1^r,\dots,t_k^r,y^!])\langle T^r/x\rangle && (\text{definition of } \langle T^r/x\rangle) \\ &= (\mathsf{D}^k s\cdot(t_1,\dots,t_k))^r\langle T^r/x\rangle\,. && (\text{definition of } (\cdot)^r) \end{aligned}$$

† However, a connection with (η_∂)-conversion can be found in Proposition 7.17(iii).

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

— Case $S \equiv sU$:
By definition, we have

$$\left(\frac{\partial(sU)}{\partial x}\cdot T\right)^r = \left(\left(\frac{\partial s}{\partial x}\cdot T\right)U + \left(\mathsf{D}s\cdot\left(\frac{\partial U}{\partial x}\cdot T\right)\right)U\right)^r$$
$$= \left(\left(\frac{\partial s}{\partial x}\cdot T\right)U\right)^r + \left(\left(\mathsf{D}s\cdot\left(\frac{\partial U}{\partial x}\cdot T\right)\right)U\right)^r$$
$$= \left(\frac{\partial s}{\partial x}\cdot T\right)^r[(U^r)^!] + \left(\lambda y.s^r\left[\left(\frac{\partial U}{\partial x}\cdot T\right)^r, y^!\right]\right)[(U^r)^!].$$

By the induction hypothesis, this is equal to

$$(s^r\langle T^r/x\rangle)[(U^r)^!] + (\lambda y.s^r[U^r\langle T^r/x\rangle, y^!])[(U^r)^!].$$

By β-conversion, this is equal to

$$(s^r\langle T^r/x\rangle)[(U^r)^!] + s^r[U^r\langle T^r/x\rangle, (U^r)^!].$$

Then, by the definition of linear substitution, this is

$$(s^r[(U^r)^!])\langle T^r/x\rangle = (sU)^r\langle T^r/x\rangle.$$

(ii) This part follows from a straightforward induction on S. □

The next proposition shows that the translation $(\cdot)^r$ is faithful too.

Proposition 7.16. For all $S, T \in \Lambda^d$, we have $\lambda\beta^d \vdash S = T$ implies $\lambda\beta^r \vdash S^r = T^r$.

Proof. It is easy to check that the proposition holds for the contextual rules.
Suppose $\lambda\beta^d \vdash S = T$ holds because

$$S \equiv \mathsf{D}^k(\lambda x.s)\cdot(u_1,\ldots,u_k)$$
$$T \equiv \lambda x.\frac{\partial^k s}{\partial x,\ldots,x}\cdot(u_1,\ldots,u_k).$$

Then we have

$$\begin{array}{rcll} S^r & = & \lambda y.(\lambda x.s^r)[u_1^r,\ldots,u_k^r,y^!] & \text{(definition of } (\cdot)^r) \\ & =_{\lambda\beta^r} & \lambda y.s^r\langle u_1^r/x\rangle\cdots\langle u_k^r/x\rangle\{y/x\} & (\beta^r\text{-conversion}) \\ & \equiv & \lambda x.s^r\langle u_1^r/x\rangle\cdots\langle u_k^r/x\rangle & (\alpha\text{-conversion}) \\ & = & \lambda x.\left(\dfrac{\partial^k s}{\partial x,\ldots,x}\cdot(u_1,\ldots,u_k)\right)^r & \text{(Lemma 7.3(i))} \\ & = & T^r. & \text{(definition of } (\cdot)^r) \end{array}$$ □

The two translations $(\cdot)^d$ and $(\cdot)^r$ are not exactly the inverses of each other. The next proposition presents the properties that they do satisfy, which are summarised in Figure 1 in terms of retractions and isomorphisms between the two calculi.

Proposition 7.17. The translations $(\cdot)^d$ and $(\cdot)^r$ enjoy the following properties:

(i) For all ordinary λ-terms s,

$$(s^r)^d \equiv s.$$

The differential λ-calculus		**The resource calculus**
with $=_{\lambda\beta\eta^d_\partial}$	$\lhd$	with $=_{\lambda\beta^r}$
with $=_{\lambda\beta^d}$	$\rhd$	with $=_{\lambda\beta^r}$
with $=_{\lambda\beta\eta^d}$	$\cong$	with $=_{\lambda\beta\eta^r}$

Fig. 1. Relationships between the differential and resource calculus.

(ii) For some $S \in \Lambda^d$ and $\mathbb{M} \in \mathbf{N}\langle\Lambda^r\rangle$,

$$(S^r)^d \not\equiv S$$
$$(\mathbb{M}^d)^r \not\equiv \mathbb{M}.$$

(iii) For all $S \in \Lambda^d$,

$$\lambda\beta\eta^d_\partial \vdash (S^r)^d = S.$$

(iv) For all $\mathbb{M} \in \mathbf{N}\langle\Lambda^r\rangle$,

$$\lambda\beta^r \vdash (\mathbb{M}^d)^r = \mathbb{M}.$$

Proof.

(i) This part follows from a straightforward induction on the structure of s.

(ii) For instance

$$\begin{aligned}((\mathsf{D}x \cdot x)^r)^d &= (\lambda y.x[x, y^!])^d \\ &= \lambda y.(\mathsf{D}x \cdot x)y \\ &\not\equiv \mathsf{D}x \cdot x.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}((x[L])^d)^r &= ((\mathsf{D}x \cdot y)0)^r \\ &= (\lambda z.x[y, z^!])0 \not\equiv x[L].\end{aligned}$$

(iii) We use induction on the structure of S:

— Case $S \equiv \mathsf{D}^k s \cdot (t_1, \ldots, t_k)$:
By the definition of $(\cdot)^r$, we have

$$\begin{aligned}((\mathsf{D}^k s \cdot (t_1, \ldots, t_k))^r)^d &= (\lambda y.s^r[t_1^r, \cdots, t_k^r, y^!])^d \\ &= \lambda y.(\mathsf{D}^k (s^r)^d \cdot \big((t_1^r)^d, \cdots, (t_k^r)^d\big))y.\end{aligned}$$

By the induction hypothesis, we have

$$\begin{aligned}(s^r)^d &=_{\lambda\beta\eta^d_\partial} s \\ (t_i^r)^d &=_{\lambda\beta\eta^d_\partial} t_i\end{aligned}$$

for all $1 \leqslant i \leqslant k$. Therefore, we get

$$\begin{aligned}\lambda y.(\mathsf{D}^k (s^r)^d \cdot \big((t_1^r)^d, \cdots, (t_k^r)^d\big))y &=_{\lambda\beta\eta^d_\partial} \lambda y.(\mathsf{D}^k s \cdot (t_1, \ldots, t_k))y \\ &=_{\lambda\beta\eta^d_\partial} \mathsf{D}^k s \cdot (t_1, \ldots, t_k)\,.\end{aligned}$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

— Case $S \equiv sT$:
We have

$$((sT)^r)^d = (s^r[(T^r)^!])^d = (s^r)^d(T^r)^d.$$

By the induction hypothesis, we know that

$$(s^r)^d =_{\lambda\beta\eta^d_\partial} s$$
$$(T^r)^d =_{\lambda\beta\eta^d_\partial} T,$$

so we can conclude

$$(s^r)^d(T^r)^d =_{\lambda\beta\eta^d_\partial} sT.$$

— All other cases are trivial.

(iv) We use induction on the structure of $\mathbb{M}$ – the only interesting case is $\mathbb{M} \equiv M[\vec{L}, \vec{N}^!]$. We have

$$((M[\vec{L}, \vec{N}^!])^d)^r = \left(\left(\mathsf{D}^k M^d \cdot \left(\vec{L}^d\right)\right)\left(\sum_{i=1}^n N_i^d\right)\right)^r$$
$$= (\lambda y.(M^d)^r[(\vec{L}^d)^r, y^!])[((\vec{N}^d)^r)^!].$$

By the induction hypothesis, we know that

$$(M^d)^r =_{\lambda\beta^r} M$$
$$(L_j^d)^r =_{\lambda\beta^r} L_j$$
$$(N_i^d)^r =_{\lambda\beta^r} N_i,$$

so

$$(\lambda y.(M^d)^r[(\vec{L}^d)^r, y^!])[((\vec{N}^d)^r)^!] =_{\lambda\beta^r} (\lambda y.M[\vec{L}, y^!])[(\vec{N})^!].$$

Then, since $y \notin \mathrm{FV}(M, \vec{L})$, we have

$$(\lambda y.M[\vec{L}, y^!])[(\vec{N})^!] =_{\lambda\beta^r} M[\vec{L}, \vec{N}^!].$$ □

8. Discussion, and related and further work

In this paper we have proposed a general categorical definition of models of the untyped differential λ-calculus, namely the notion of a linear reflexive object living in a Cartesian closed differential category. We have proved that this notion of a model is:

(i) Sound – in other words, the equational theory induced by a model is actually a differential λ-theory;
(ii) Inhabited – indeed we have given concrete examples of such a definition in the form of the models $\mathscr{D}$ and $\mathscr{E}$ living in **MRel** and all the syntactic models built trough the revised Scott–Koymans' construction;
(iii) Equationally complete – provided we restrict consideration to differentially extensional differential λ-theories satisfying sum idempotency.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Finally, we have shown that the equational theories of the differential λ-calculus and of the resource calculus are tightly connected. Formally, we have provided faithful translations between the two calculi, thus showing that they share the same notion of a model. In particular, this shows that linear reflexive objects in Cartesian closed differential categories are also sound models of the untyped resource calculus. For the resource calculus, we have been able to prove an even stronger equational completeness theorem, in the sense that it holds for all resource λ-theories satisfying sum idempotency.

8.1. *Related work*

This paper is in spirit a continuation of the work on (Cartesian) differential categories done in Blute *et al.* (2006; 2009) and can be considered as a long version of Bucciarelli *et al.* (2010). Note, however, that all the calculi considered in those papers were simply typed. Moreover, our aim here has been to find a suitable notion of semantics for Ehrhard and Regnier's differential λ-calculus (so we have assumed the calculus as given), while in Blute *et al.* (2006; 2009) the goal was to provide a categorical axiomatisation of a differential operator and then find a calculus (namely, the term logic) that suits the categories under consideration. In particular, the differential calculus presented in Blute *et al.* (2009) was slightly different from Ehrhard and Regnier's differential λ-calculus in some key ways.

On the one hand, the calculus defined in Blute *et al.* (2009) has no λ-abstraction, so it is not an extension of λ-calculus, but, on the other hand, it does have explicit substitutions and constructors for the pairing, the projections and every n-ary function. Also, the treatment of differentiation is different – in the Leibniz-style approach of Blute *et al.* (2009), the notation for differentiation becomes

$$\frac{\Gamma, x : A \vdash t : B \quad \Gamma \vdash s : A \quad \Gamma \vdash u : A}{\Gamma \vdash \frac{\mathsf{d}t}{\mathsf{d}x}(s) \cdot u : B}\ (\partial)$$

where the variable x is bound in t. Hence, differential terms are built using the binder

$$\frac{\mathsf{d}(\cdot)}{\mathsf{d}x}.$$

Intuitively,

$$\frac{\mathsf{d}t}{\mathsf{d}x}(s)$$

denotes the derivative of t at s† and determines a linear transformation, so that it could be typed as

$$\frac{\mathsf{d}t}{\mathsf{d}x}(s) : A \multimap B,$$

while u is the point where the derivative is calculated. The lack of λ-abstraction in this system is not a true difference because λ-terms could be added without problems. In

† In other words, the Jacobian matrix.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

addition to the usual equations from λ-calculus, one should just add

$$\lambda x.(s+t) = \lambda x.s + \lambda x.t$$
$$\frac{\mathsf{d}(\lambda y.s)}{\mathsf{d}x}(t)\cdot u = \lambda y.\frac{\mathsf{d}s}{\mathsf{d}x}(t)\cdot u$$

and the resulting system is conjectured to have linear reflexive objects living in Cartesian closed differential categories as sound and complete models. To understand this better, we can sketch the translation from the differential λ-calculus into this system as follows:

$$(\mathsf{D}s\cdot t)^\circ = \lambda x_0.\left(\frac{\mathsf{d}\,(s^\circ x)}{\mathsf{d}x}(x_0)\cdot t^\circ\right),$$
$$\left(\frac{\partial s}{\partial x}\cdot t\right)^\circ = \frac{\mathsf{d}s^\circ}{\mathsf{d}x}(x)\cdot t^\circ.$$

where x_0 is some fresh variable. This calculus is certainly more standard from a mathematical point of view, while we think the differential and resource calculi are more standard from a computer scientist point of view. We believe that the language in Blute *et al.* (2009) would be a more promising option for getting a completeness result in the simply typed setting; on the other hand, in the untyped case, the language would suffer the same problems we encountered in Theorem 5.17, namely, the completeness only applies for differentially extensional theories satisfying sum idempotency. Finally, once the calculus is stripped of types and constructors (since in this paper we are interested in the pure untyped setting), it becomes quite similar to the differential λ-calculus. For all these reasons, we decided not to analyse the calculus of Blute *et al.* (2009) any further.

8.2. *Other examples of Cartesian closed differential categories*

In Section 6 we presented **MRel** (and mentioned **MFin** in Remark 6.1) as an instance of the definition of a Cartesian closed differential category. We will briefly discuss here some other examples of such categories that have been recently defined in the literature.

In fact, *game semantics* is an inexhaustible source of differential categories, indeed, resource usage is represented rather explicitly in games and strategies. In collaboration with Laird and McCusker, we showed in Laird *et al.* (2011) that the games model $\mathbf{G}^\otimes$ of Idealised Algol with non-determinism introduced in Harmer and McCusker (1999) contains a (definable) differential operator giving it the structure of a Cartesian closed differential category. The category $\mathbf{G}^\otimes$ is cpo-enriched, has arenas as objects and suitable non-deterministic strategies as morphisms. Intuitively, this category is additive since non-deterministic strategies are closed under union and the linearity of a strategy on a certain component is captured by the fact that the strategy plays *exactly once* in that component.

Moreover, Laird *et al.* (2011) provided a general categorical construction for building differential categories. Its key step takes a symmetric monoidal category with countable biproducts, embeds it in its *Karoubi envelope* and then constructs the *cofree cocommutative comonoid* on this category (following the recipe in Melliès *et al.* (2009)) and a differential operator on the Kleisli category of the corresponding comonad. Since biproducts may be added to any category by free constructions, this gives a way to embed any symmetric

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

monoidal (closed) category in a Cartesian (closed) differential category. This construction allows us to recover both the category **MRel**, starting from the terminal symmetric monoidal closed category (one object, one morphism), and $\mathbf{G}^{\otimes}$, starting from a symmetric monoidal category of *exhausting games*.

The category $\mathbf{G}^{\otimes}$, just like **MRel**, models the Taylor expansion. Natural examples of differential Cartesian closed categories that *do not* model the Taylor expansion have been recently defined in Carraro *et al.* (2010) by introducing new exponential operations on **Rel**. The intuition behind this construction is rather simple: the authors replace the set of natural numbers (which are used for counting multiplicities of elements in multisets) with more general semi-rings containing elements ω such that $\omega + 1 = \omega$ (that is, elements that are essentially infinite). In these models with infinite multiplicities, all differential constructions are available, but the Taylor formula does not hold. Indeed, in these categories it is possible to find a morphism $f \neq 0$ such that, for all $n \in \mathbf{N}$, the nth derivative of f evaluated on 0 is equal to 0: the Taylor expansion of such an f is the 0 map, and hence the morphism is different from its Taylor expansion. In particular, the authors exhibit models where the interpretation of Ω is different from 0.

8.3. *Algebraic approach*

Another interesting line of research would be to provide an algebraic definition of a model of the differential λ-calculus. In other words, we would like to introduce a class of algebras modelling the differential λ-calculus in the same way as combinatory algebras model the regular one. This would open the way to generalising the powerful techniques developed in Lusin and Salibra (2004), Manzonetto and Salibra (2010) and Salibra (2000) for analysing combinatory algebras. For instance, we have proved in collaboration with Salibra that combinatory algebras satisfy good algebraic properties, such as a Stone representation theorem stating that every combinatory algebra is decomposable in a weak Boolean product of non-decomposable algebras (Manzonetto and Salibra 2010). This allowed us, among other things, to give a uniform proof of incompleteness for the main semantics of λ-calculus (that is, the continuous, stable and strongly stable semantics).

A first attempt at providing algebraic models of the resource calculus was recently presented in Carraro *et al.* (2010), where the authors introduced the notion of 'resource λ-models' and showed that they are suitable for modelling the *finite* resource calculus (that is, the promotion-free fragment). However, at least at the moment, a generalisation allowing to model the full fragment of resource calculus (or, equivalently, the differential λ-calculus) does not seem easy, and is reserved for future work.

Appendix A. Technical appendix

This technical appendix gives the full proofs of the two main lemmas in Section 4.3 (Lemmas 4.7 and 4.8). The proofs are not particularly difficult, but quite long and require some preliminary notation.

Notation A.1. We will adopt the following notation:

— Given a sequence of indices $\vec{i} = i_1, \ldots, i_k$ with $i_j \in \{1, 2\}$, we write $\pi_{\vec{i}}$ for $\pi_{i_1} \circ \cdots \circ \pi_{i_k}$. Thus $\pi_{1,2} = \pi_1 \circ \pi_2$.

— For brevity, when writing a Cartesian product of objects as subscript of 0 or Id, we will replace the operator $\times$ by simple juxtaposition. For instance, the morphism $\mathrm{Id}_{(A\times B)\times(C\times D)}$ will be written $\mathrm{Id}_{(AB)(CD)}$.

We will write '(proj)' to refer to the rules

$$\pi_1 \circ \langle f, g \rangle = f$$
$$\pi_2 \circ \langle f, g \rangle = g$$

that hold in every Cartesian category. Recall that

$$\mathrm{sw}_{ABC} = \langle\langle \pi_{1,1}, \pi_2 \rangle, \pi_{2,1} \rangle : (A \times B) \times C \to (A \times C) \times B.$$

Lemma 4.7. Let $f : (C \times A) \times D \to B$, $g : C \to A$, $h : C \to B'$.

(i) $\pi_2 \star g = g \circ \pi_1$.
(ii) $(h \circ \pi_1) \star g = 0$.
(iii) $\Lambda(f) \star g = \Lambda(((f \circ \mathrm{sw}) \star (g \circ \pi_1)) \circ \mathrm{sw})$.

Proof.

(i) We have

$$\begin{aligned} \pi_2 \star g &= D(\pi_2) \circ \langle\langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{CA} \rangle && \text{(definition of } \star) \\ &= \pi_2 \circ \pi_1 \circ \langle\langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{CA} \rangle && \text{(D3)} \\ &= \pi_2 \circ \langle 0_C, g \circ \pi_1 \rangle && \text{(proj)} \\ &= g \circ \pi_1. && \text{(proj)} \end{aligned}$$

(ii) We have

$$\begin{aligned} (h \circ \pi_1) \star g &= D(h \circ \pi_1) \circ \langle\langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{CA} \rangle && \text{(definition of } \star) \\ &= D(h) \circ \langle D(\pi_1), \pi_{1,2} \rangle \, \langle\langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{CA} \rangle && \text{(D5)} \\ &= D(h) \circ \langle \pi_1 \circ \pi_1, \pi_{1,2} \rangle \circ \langle\langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{C\times A} \rangle && \text{(D3)} \\ &= D(h) \circ \langle 0_C, \pi_1 \rangle \, . && \text{(proj)} \\ &= 0. && \text{(D2)} \end{aligned}$$

(iii) We first prove the following claim.

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Claim A.2. Let $g : C \to A$. Then the following diagram commutes:

$$\begin{array}{ccc}
(C \times A) \times D & \xrightarrow{\langle \pi_1, \mathrm{Id}_{C \times A} \rangle \times \mathrm{Id}_D} & (C \times (C \times A)) \times D \\
\downarrow{\scriptstyle \langle \pi_1 \times \mathrm{Id}_D, \mathrm{sw} \rangle} & & \downarrow{\scriptstyle (\langle 0_C, g \rangle \times \mathrm{Id}_{C \times A}) \times \mathrm{Id}_D} \\
(C \times D) \times ((C \times D) \times A) & & ((C \times A) \times (C \times A)) \times D \\
\downarrow{\scriptstyle \langle 0_{C \times D}, g \circ \pi_1 \rangle \times \mathrm{Id}_{(C \times D) \times A}} & & \downarrow{\scriptstyle \langle \pi_1 \times 0_D, \pi_2 \times \mathrm{Id}_D \rangle} \\
((C \times D) \times A) \times ((C \times D) \times A) & \xrightarrow{\langle D(\mathrm{sw}), \mathrm{sw} \circ \pi_2 \rangle} & ((C \times A) \times D) \times ((C \times A)) \times D)
\end{array}$$

Proof of Claim A.2. We have

$$\begin{aligned}
&\langle \pi_1 \times 0_D, \pi_2 \times \mathrm{Id}_D \rangle \circ ((\langle 0_C, g \rangle \times \mathrm{Id}_{CA}) \times \mathrm{Id}_D) \circ (\langle \pi_1, \mathrm{Id}_{CA} \rangle \times \mathrm{Id}_D) \\
&\quad = \langle \langle \langle 0_C, g \circ \pi_{1,1} \rangle, 0_D \rangle, \langle \pi_{2,1}, \pi_2 \rangle \rangle \circ \langle \langle \pi_1, \langle \pi_1, \pi_2 \rangle \rangle \circ \pi_1, \pi_2 \rangle \\
&\quad = \langle \langle \langle 0_C, g \circ \pi_{1,1} \rangle, 0_D \rangle, \langle \pi_{2,1}, \pi_2 \rangle \rangle \circ \langle \langle \pi_{1,1}, \langle \pi_{1,1}, \pi_{2,1} \rangle \rangle, \pi_2 \rangle \\
&\quad = \langle \langle \langle 0_C, g \circ \pi_{1,1} \rangle, 0_D \rangle, \langle \langle \pi_{1,1}, \pi_{2,1} \rangle, \pi_2 \rangle \rangle \\
&\quad = \langle \langle \langle 0_C, g \circ \pi_{1,1} \rangle, 0_D \rangle, \langle \langle \pi_{1,1,2}, \pi_{2,2} \rangle, \pi_{2,1,2} \rangle \rangle \circ \langle \pi_1 \times \mathrm{Id}_D, \mathrm{sw} \rangle \\
&\quad = \langle \langle \langle \pi_{1,1,1}, \pi_{2,1} \rangle, \pi_{2,1,1} \rangle, \langle \langle \pi_{1,1,2}, \pi_{2,2} \rangle, \pi_{2,1,2} \rangle \rangle \circ \\
&\qquad \langle \langle 0_{CD}, g \circ \pi_{1,1} \rangle, \pi_2 \rangle \circ \langle \pi_1 \times \mathrm{Id}_D, \mathrm{sw} \rangle \\
&\quad = \langle D(\mathrm{sw}), \mathrm{sw} \circ \pi_2 \rangle \circ (\langle 0_{CD}, g \circ \pi_1 \rangle \times \mathrm{Id}_{(CD)A}) \circ \langle \pi_1 \times \mathrm{Id}_D, \mathrm{sw} \rangle . \qquad \square
\end{aligned}$$

We can now conclude the proof of part (iii) of Lemma 4.7 as follows:

$$\begin{aligned}
\Lambda(f) \star g &= D(\Lambda(f)) \circ \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{CA} \rangle && \text{(definition of } \star\text{)} \\
&= \Lambda(D(f) \circ \langle \pi_1 \times 0_D, \pi_2 \times \mathrm{Id}_D \rangle) \circ \langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{CA} \rangle && \text{(D-curry)} \\
&= \Lambda(D(f) \circ \langle \pi_1 \times 0_D, \pi_2 \times \mathrm{Id}_D \rangle \circ ((\langle \langle 0_C, g \circ \pi_1 \rangle, \mathrm{Id}_{CA} \rangle) \times \mathrm{Id}_D)) && \text{(Curry)} \\
&= \Lambda(D(f) \circ \langle D(\mathrm{sw}), \mathrm{sw} \circ \pi_2 \rangle \\
&\qquad \circ (\langle 0_{CD}, g \circ \pi_1 \rangle \times \mathrm{Id}_{(CD)A}) \circ \langle \pi_1 \times \mathrm{Id}_D, \mathrm{sw} \rangle) && \text{(Claim A.2)} \\
&= \Lambda(D(f \circ \mathrm{sw}) \circ (\langle 0_{CD}, g \circ \pi_1 \rangle \times \mathrm{Id}_{(CD)A}) \circ \langle \pi_1, \mathrm{Id} \rangle \circ \mathrm{sw}) && \text{(D5)} \\
&= \Lambda(((f \circ \mathrm{sw}) \star (g \circ \pi_1)) \circ \mathrm{sw}). && \text{(definition of } \star\text{)}
\end{aligned}$$

$\square$

Lemma 4.8. Let

$$\begin{aligned}
&f : C \times A \to (D \Rightarrow B) \\
&g : C \to A \\
&h : C \times A \to D.
\end{aligned}$$

Then:

(i) $(\mathrm{ev} \circ \langle f, h \rangle) \star g = \mathrm{ev} \circ \langle f \star g + \Lambda(\Lambda^-(f) \star (h \star g)), h \rangle$
(ii) $\Lambda(\Lambda^-(f) \star h) \star g = \Lambda(\Lambda^-(f \star g) \star h) + \Lambda(\Lambda^-(f) \star (h \star g))$
(iii) $\Lambda(\Lambda^-(f) \star h) \circ \langle \mathrm{Id}_C, g \rangle = \Lambda(\Lambda^-(f \circ \langle \mathrm{Id}_C, g \rangle) \star (h \circ \langle \mathrm{Id}_C, g \rangle))$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Proof.

(i) Let $\varphi \equiv \langle\langle 0_C, g\circ\pi_1\rangle, \mathrm{Id}_{CA}\rangle$. Then

$$
\begin{aligned}
&(\mathrm{ev}\circ\langle f,h\rangle)\star g\\
&\quad= D(\mathrm{ev}\circ\langle f,h\rangle)\circ\varphi && \text{(definition of } \star)\\
&\quad= (\mathrm{ev}\circ\langle D(f),h\circ\pi_2\rangle + D(\Lambda^-(f))\circ\langle\langle 0_{CA}, D(h)\rangle,\langle\pi_2,h\circ\pi_2\rangle\rangle)\circ\varphi && \text{(D-eval)}\\
&\quad= \mathrm{ev}\circ\langle D(f),h\circ\pi_2\rangle\circ\varphi\\
&\qquad + D(\Lambda^-(f))\circ\langle\langle 0_{CA}, D(h)\circ\varphi\rangle,\langle \mathrm{Id}_{CA},h\rangle\rangle && \text{(Definition 4.2)}\\
&\quad= \mathrm{ev}\circ\langle D(f)\circ\varphi,h\rangle\\
&\qquad + D(\Lambda^-(f))\circ\langle\langle 0_{CA},(h\star g)\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle\circ\langle \mathrm{Id}_{CA},h\rangle && \text{(definition of } \star)\\
&\quad= \mathrm{ev}\circ\langle f\star g,h\rangle + (\Lambda^-(f)\star(h\star g))\circ\langle \mathrm{Id},h\rangle && \text{(definition of } \star)\\
&\quad= \mathrm{ev}\circ\langle f\star g,h\rangle + \mathrm{ev}\circ\langle\Lambda(\Lambda^-(f)\star(h\star g)),h\rangle && \text{(beta-cat)}\\
&\quad= \mathrm{ev}\circ\langle f\star g + \Lambda(\Lambda^-(f)\star(h\star g)),h\rangle && \text{(Lemma 4.3)}
\end{aligned}
$$

(ii) We first simplify the equation

$$\Lambda(\Lambda^-(f)\star h)\star g = \Lambda(\Lambda^-(f\star g)\star h) + \Lambda(\Lambda^-(f)\star(h\star g))$$

to get rid of the Cartesian closed structure. The right-hand side can be rewritten as

$$\Lambda((\Lambda^-(f\star g)\star h) + \Lambda^-(f)\star(h\star g)).$$

Taking a morphism $f' : (C\times A)\times D \to B$ such that $f = \Lambda(f')$ and applying Lemma 4.7 (iii), we then discover that proving part (ii) is equivalent to showing that:

$$((f'\star h)\circ \mathrm{sw})\star(g\circ\pi_1)\circ \mathrm{sw} = (((f'\circ \mathrm{sw})\star(g\circ\pi_1))\circ \mathrm{sw})\star h + f'\star(h\star g).$$

By the definition of $\star$, we have

$$((f'\star h)\circ \mathrm{sw})\star(g\circ\pi_1)\circ \mathrm{sw} = D(D(f')\circ\langle\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle,\mathrm{sw}\rangle)\circ\langle\langle 0_{CD}, g\circ\pi_{1,1}\rangle,\mathrm{sw}\rangle\,.$$

Writing $\varphi \equiv \langle\langle 0_{CD}, g\circ\pi_{1,1}\rangle,\mathrm{sw}\rangle$ and $D^2(f)$ for $D(D(f))$. Then

$$
\begin{aligned}
&D^2(f')\circ\langle\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle,\mathrm{sw}\rangle)\circ\varphi\\
&\quad= D^2(f')\circ\langle D(\langle\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle,\mathrm{sw}\rangle),\langle\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle,\mathrm{sw}\rangle\circ\pi_2\rangle\circ\varphi && \text{(D5)}\\
&\quad= D^2(f')\circ\langle D(\langle\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle,\mathrm{sw}\rangle)\circ\varphi,\\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle,\mathrm{sw}\rangle\circ\pi_2\circ\varphi\rangle && \text{(pair)}\\
&\quad= D^2(f')\circ\langle\langle D(\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle)\circ\varphi, D(\mathrm{sw})\circ\varphi\rangle,\\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle,\mathrm{sw}\rangle\circ \mathrm{sw}\rangle && \text{(D4)}\\
&\quad= D^2(f')\circ\langle\langle D(\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle)\circ\varphi, D(\mathrm{sw})\circ\varphi\rangle,\\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle. && \text{(Remark 4.10)}
\end{aligned}
$$

Since

$$
\begin{aligned}
&\langle D(\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle)\circ\varphi, D(\mathrm{sw})\circ\varphi\rangle\\
&\qquad\qquad = \langle 0, D(\mathrm{sw})\circ\varphi\rangle + \langle D(\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle)\circ\varphi, 0\rangle,
\end{aligned}
$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

we can apply D2 and rewrite the expression above as a sum of two morphisms:

(1) $D^2(f')\circ\langle\langle 0_{(CA)D}, D(\mathrm{sw})\circ\varphi\rangle,\langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle$ +

(2) $D^2(f')\circ\langle\langle D(\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle)\circ\varphi, 0_{(CA)D}\rangle,\langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle$

We now show that (1) $= (((f'\circ\mathrm{sw})\star(g\circ\pi_1))\circ\mathrm{sw})\star h$. Indeed, we have

$$
\begin{aligned}
& D^2(f')\circ\langle\langle 0_{(CA)D}, D(\mathrm{sw})\circ\varphi\rangle,\langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle \\
&\quad = D^2(f')\circ\langle\langle 0_{(CA)D}, \mathrm{sw}\circ\pi_1\circ\varphi\rangle,\langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle && \text{(Remark 4.10)} \\
&\quad = D^2(f')\circ\langle\langle 0_{(CA)D}, \mathrm{sw}\circ\langle 0_{CD}, g\circ\pi_{1,1}\rangle\rangle,\langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle && \text{(proj)} \\
&\quad = D^2(f')\circ\langle\langle 0_{(CA)D}, \langle\langle 0_C, g\circ\pi_{1,1}\rangle, 0_D\rangle\rangle, \\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle && \text{(Remark 4.10)} \\
&\quad = D^2(f')\circ\langle\langle\langle\langle 0_C, 0_A\rangle, 0_D\rangle, \langle 0_{CA}, h\circ\pi_1\rangle\rangle, \\
&\qquad\qquad \langle\langle\langle 0_C, g\circ\pi_{1,1}\rangle, 0_D\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle && \text{(D7)} \\
&\quad = D^2(f')\circ\langle\langle\langle\langle 0_C, D(g)\circ\langle 0_C,\pi_{1,1}\rangle\rangle, 0_D\rangle, \langle 0_{CA}, h\circ\pi_1\rangle\rangle \\
&\qquad\qquad \langle\langle\langle 0_C, g\circ\pi_{1,1}\rangle, 0_D\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle. && \text{(D2)}
\end{aligned}
$$

Letting $\psi \equiv \langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle$, we then have

$$
\begin{aligned}
& D(D(f))\circ\langle\langle\langle\langle 0_C, D(g)\circ\langle 0_C,\pi_{1,1}\rangle\rangle, 0_D\rangle, \langle 0_{CA}, h\circ\pi_1\rangle\rangle, \langle\langle\langle 0_C, g\circ\pi_{1,1}\rangle, 0_D\rangle,\mathrm{Id}_{(CA)D}\rangle\rangle \\
&\quad = D(D(f))\circ\langle\langle\langle\langle 0_C, D(g)\circ\langle\pi_{1,1,1},\pi_{1,1,2}\rangle\rangle, 0_D\rangle,\pi_1\rangle, \\
&\qquad\qquad \langle\langle\langle 0_C, g\circ\pi_{1,1,2}\rangle, 0_D\rangle,\pi_2\rangle\rangle\circ\psi && \text{(proj)} \\
&\quad = D(D(f))\circ\langle\langle\langle\langle 0_C, D(g)\circ\langle D(\pi_{1,1}),\pi_{1,1,2}\rangle\rangle, 0_D\rangle,\pi_1\rangle, \\
&\qquad\qquad \langle\langle\langle 0_C, g\circ\pi_{1,1,2}\rangle, 0_D\rangle,\pi_2\rangle\rangle\circ\psi && \text{(D3)} \\
&\quad = D(D(f))\circ\langle\langle\langle\langle 0_C, D(g\circ\pi_{1,1})\rangle, 0_D\rangle,\pi_1\rangle, \langle\langle\langle 0_C, g\circ\pi_{1,1,2}\rangle, 0_D\rangle,\pi_2\rangle\rangle\circ\psi && \text{(D5)} \\
&\quad = D(D(f))\circ\langle\langle\langle\langle D(0_C), D(g\circ\pi_{1,1})\rangle, D(0_D)\rangle, D(\mathrm{Id}_{(CA)D})\rangle, \\
&\qquad\qquad \langle\langle\langle 0_C, g\circ\pi_{1,1,2}\rangle, 0_D\rangle,\pi_2\rangle\rangle\circ\psi && \text{(D1)} \\
&\quad = D(D(f))\circ\langle D(\langle\langle\langle 0_C, g\circ\pi_{1,1}\rangle, 0_D\rangle,\mathrm{Id}_{(CA)D}\rangle), \\
&\qquad\qquad \langle\langle\langle 0_C, g\circ\pi_{1,1}\rangle, 0_D\rangle,\mathrm{Id}_{(CA)D}\rangle\circ\pi_2\rangle\circ\psi && \text{(D4)} \\
&\quad = D(D(f)\circ\langle\langle\langle 0_C, g\circ\pi_{1,1}\rangle, 0_D\rangle,\mathrm{Id}_{(CA)D}\rangle)\circ\psi && \text{(D5)} \\
&\quad = D(D(f)\circ\langle\mathrm{sw}\circ\langle 0_{CD}, g\circ\pi_{1,1}\rangle, \mathrm{sw}\circ\mathrm{sw}\rangle)\circ\psi && \text{(Remark 4.10)} \\
&\quad = D(D(f)\circ\langle\mathrm{sw}\circ\pi_1, \mathrm{sw}\circ\pi_2\rangle\circ\langle\langle 0_{CD}, g\circ\pi_{1,1}\rangle, \mathrm{sw}\rangle)\circ\psi && \text{(proj)} \\
&\quad = D(D(f)\circ\langle D(\mathrm{sw}), \mathrm{sw}\circ\pi_2\rangle\circ\langle\langle 0_{CD}, g\circ\pi_{1,1}\rangle, \mathrm{sw}\rangle)\circ\psi && \text{(Remark 4.10)} \\
&\quad = D(D(f\circ\mathrm{sw})\circ\langle\langle 0_{CD}, g\circ\pi_{1,1}\rangle,\mathrm{Id}_{(CD)A}\rangle\circ\mathrm{sw})\circ\langle\langle 0_{CA}, h\circ\pi_1\rangle,\mathrm{Id}_{(CA)D}\rangle && \text{(D5)} \\
&\quad = (((f\circ\mathrm{sw})\star(g\circ\pi_1))\circ\mathrm{sw})\star h. && \text{(definition of } \star)
\end{aligned}
$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

We will now show that (2) $= f \star (h \star g)$, which will conclude the proof of part (ii).

$$
\begin{aligned}
& D^2(f)\circ\langle\langle D(\langle 0_{CA}, h\circ\langle\pi_{1,1},\pi_2\rangle\rangle)\circ\varphi, 0_{(CA)D}\rangle, \langle\langle 0_{CA}, h\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle\rangle \\
&\quad = D^2(f)\circ\langle\langle\langle 0_{CA}, D(h\circ\langle\pi_{1,1},\pi_2\rangle)\rangle\circ\varphi, 0_{(CA)D}\rangle, \\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle\rangle && \text{(D1+D4)} \\
&\quad = D^2(f)\circ\langle\langle\langle 0_{CA}, D(h)\circ\langle D(\langle\pi_{1,1},\pi_2\rangle), \langle\pi_{1,1,2},\pi_{2,2}\rangle\rangle\rangle\circ\varphi, 0_{(CA)D}\rangle, \\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle\rangle && \text{(D5)} \\
&\quad = D^2(f)\circ\langle\langle\langle 0_{CA}, D(h)\circ\langle\langle D(\pi_{1,1}), D(\pi_2)\rangle, \langle\pi_{1,1,2},\pi_{2,2}\rangle\rangle\rangle\circ\varphi, 0_{(CA)D}\rangle, \\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle\rangle && \text{(D4+D3)} \\
&\quad = D^2(f)\circ\langle\langle\langle 0_{CA}, D(h)\circ\langle\langle\pi_{1,1,1},\pi_{2,1}\rangle, \langle\pi_{1,1,2},\pi_{2,2}\rangle\rangle\rangle\circ\varphi, 0_{(CA)D}\rangle, \\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle\rangle && \text{(D5+D3)} \\
&\quad = D^2(f)\circ\langle\langle\langle 0_{CA}, D(h)\circ\langle\langle 0_C, g\circ\pi_{1,1}\rangle, \pi_1\rangle\rangle, 0_{(CA)D}\rangle, \\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle\rangle && \text{(proj)} \\
&\quad = D(f)\circ\langle\langle 0_{CA}, D(h)\circ\langle\langle 0_C, g\circ\pi_{1,1}\rangle, \pi_1\rangle\rangle, \mathrm{Id}_{(CA)D}\rangle && \text{(D6)} \\
&\quad = D(f)\circ\langle\langle 0_{CA}, D(h)\circ\langle\langle 0_C, g\circ\pi_1\rangle, \mathrm{Id}_{CA}\rangle\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle && \text{(proj)} \\
&\quad = f\star(h\star g) && \text{(definition of } \star)
\end{aligned}
$$

(iii) By (Curry), we have

$$\Lambda(\Lambda^-(f)\star h)\circ\langle\mathrm{Id}_C, g\rangle = \Lambda((\Lambda^-(f)\star h)\circ(\langle\mathrm{Id}_C, g\rangle\times\mathrm{Id}_D)),$$

so if we can show that

$$(\Lambda^-(f)\star h)\circ(\langle\mathrm{Id}_C, g\rangle\times\mathrm{Id}_D) = \Lambda^-(f\circ\langle\mathrm{Id}_C, g\rangle)\star(h\circ\langle\mathrm{Id}_C, g\rangle),$$

we are done.

So we proceed as follows:

$$
\begin{aligned}
& (\Lambda^-(f)\star h)\circ(\langle\mathrm{Id}_C, g\rangle\times\mathrm{Id}_D) \\
&\quad = D(\Lambda^-(f))\circ\langle\langle 0_{CA}, h\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle\circ(\langle\mathrm{Id}_C, g\rangle\times\mathrm{Id}_D) && \text{(definition of } \star) \\
&\quad = D(\mathrm{ev}\circ\langle f\circ\pi_1,\pi_2\rangle)\circ \\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\pi_1\rangle, \mathrm{Id}_{(CA)D}\rangle\circ(\langle\mathrm{Id}_C, g\rangle\times\mathrm{Id}_D) && \text{(definition of } \Lambda^-) \\
&\quad = D(\mathrm{ev})\circ\langle\langle D(f\circ\pi_1), D(\pi_2)\rangle, \langle f\circ\pi_{1,2},\pi_{2,2}\rangle\rangle\circ \\
&\qquad\qquad \langle\langle 0_{CA}, h\rangle\circ\langle\pi_1, g\circ\pi_1\rangle, \langle\mathrm{Id}_C, g\rangle\times\mathrm{Id}_D\rangle && \text{(D5+D4)} \\
&\quad = D(\mathrm{ev})\circ\langle\langle D(f)\circ\langle\pi_{1,1},\pi_{1,2}\rangle, \pi_{2,1}\rangle, \langle f\circ\pi_{1,2},\pi_{2,2}\rangle\rangle\circ \\
&\qquad\qquad \langle\langle 0_{CA}, h\circ\langle\pi_1, g\circ\pi_1\rangle\rangle, \langle\mathrm{Id}_C, g\rangle\times\mathrm{Id}_D\rangle && \text{(D5+D3)} \\
&\quad = D(\mathrm{ev})\circ\langle\langle D(f)\circ\langle 0_{CA}, \langle\pi_1, g\circ\pi_1\rangle\rangle, h\circ\langle\pi_1, g\circ\pi_1\rangle\rangle, \langle f\circ\langle\pi_1, g\circ\pi_1\rangle, \pi_2\rangle\rangle && \text{(proj)} \\
&\quad = D(\mathrm{ev})\circ\langle\langle D(f)\circ\langle\langle 0_C, D(g)\circ\langle 0_C, \mathrm{Id}_C\rangle\rangle, \langle\mathrm{Id}_C, g\rangle\rangle, h\circ\langle\pi_1, g\circ\pi_1\rangle\rangle, \\
&\qquad\qquad \langle f\circ\langle\mathrm{Id}_C, g\rangle, \mathrm{Id}_D\rangle\rangle && \text{(D2)}
\end{aligned}
$$

https://doi.org/10.1017/S0960129511000594 Published online by Cambridge University Press

Setting $\varphi = \langle\langle 0_C, h\circ\langle \pi_1, g\circ\pi_1\rangle\rangle, \mathrm{Id}_{CD}\rangle$, this equals

$$
\begin{aligned}
&D(\mathrm{ev})\circ\langle\langle D(f)\circ\langle\langle \pi_{1,1}, D(g)\circ\langle \pi_{1,1},\pi_{1,2}\rangle\rangle, \langle \pi_{1,2}, g\circ\pi_{1,2}\rangle\rangle, \pi_{2,1}\rangle,\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \langle f\circ\langle \pi_{1,2}, g\circ\pi_{1,2}\rangle, \pi_{2,2}\rangle\circ\varphi\rangle\\
&\quad = D(\mathrm{ev})\circ\langle\langle D(f\circ\langle \pi_1, g\circ\pi_1\rangle), D(\pi_2)\rangle, \langle f\circ\langle \pi_{1,2}, g\circ\pi_{1,2}\rangle, \pi_{2,2}\rangle\rangle\circ\varphi && \text{(D5)}\\
&\quad = D(\mathrm{ev})\circ\langle D(\langle f\circ\langle \pi_1, g\circ\pi_1\rangle, \pi_2\rangle), \langle f\circ\langle \pi_{1,2}, g\circ\pi_{1,2}\rangle, \pi_{2,2}\rangle\rangle\circ\varphi && \text{(D4)}\\
&\quad = D(\mathrm{ev}\circ\langle f\circ\langle \pi_1, g\circ\pi_1\rangle, \pi_2\rangle)\circ\varphi && \text{(D5)}\\
&\quad = D(\Lambda^-(f\circ\langle \mathrm{Id}_C, g\rangle))\circ\varphi && \text{(definition of } \Lambda^-)\\
&\quad = \Lambda^-(f\circ\langle \mathrm{Id}_C, g\rangle)\star(h\circ\langle \mathrm{Id}_C, g\rangle). && \text{(definition of } \star)
\end{aligned}
$$

□

Acknowledgements

We are grateful to Antonio Bucciarelli, Thomas Ehrhard and Guy McCusker for interesting discussions. Many thanks are also due to Michele Pagani, Paolo Tranquilli and the anonymous reviewers for helpful comments and suggestions.

References

Asperti, A. and Longo, G. (1991) *Categories, types and structures. Category theory for the working computer scientist*, M.I.T. Press.

Barendregt, H. P. (1984) *The Lambda calculus: Its syntax and semantics*, North-Holland.

Berry, G. (1978) Stable models of typed lambda-calculi. In: Ausiello, G. and Böhm, C. (eds.) ICALP. *Springer-Verlag Lecture Notes in Computer Science* **62** 72–89.

Blute, R. F., Cockett, J. R. B. and Seely, R. A. G. (2006) Differential categories. *Mathematical Structures in Computer Science* **16** (6) 1049–1083.

Blute, R. F., Cockett, J. R. B. and Seely, R. A. G. (2009) Cartesian differential categories. *Theory and Applications of Categories* **22** (23) 622–672.

Boudol, G. (1993) The lambda-calculus with multiplicities. In: *CONCUR'93* 1–6.

Boudol, G., Curien, P.-L. and Lavatelli, C. (1999) A semantics for lambda calculi with resources. *Mathematical Structures in Computer Science* **9** (4) 437–482.

Bucciarelli, A., Carraro, A., Ehrhard, T. and Manzonetto, G. (2011) Full abstraction for resource calculus with tests. In: Bezem, M. (ed.) Computer Science Logic (CSL'11) – 25th International Workshop/20th Annual Conference of the EACSL. *Leibniz International Proceedings in Informatics (LIPIcs)* **12**, Schloss Dagstuhl–Leibniz-Zentrum für Informatik 97–111.

Bucciarelli, A. and Ehrhard, T. (1991) Sequentiality and strong stability. In: *Sixth Annual IEEE Symposium on Logic in Computer Science*, IEEE Computer Society Press 138–145.

Bucciarelli, A., Ehrhard, T. and Manzonetto, G. (2007) Not enough points is enough. In: Proc. of Comp. Sci. Logic 2007. *Springer-Verlag Lecture Notes in Computer Science* **4646** 298–312.

Bucciarelli, A., Ehrhard, T. and Manzonetto, G. (2009) A relational model of a parallel and non-deterministic lambda-calculus. In: International Symposium on Logical Foundations of Computer Science (LFCS 2009). *Springer-Verlag Lecture Notes in Computer Science* **5407** 107–121.

Bucciarelli, A., Ehrhard, T. and Manzonetto, G. (2010) Categorical models for simply typed resource calculi. *Electronic Notes in Theoretical Computer Science* **265** 213–230.

Carraro, A., Ehrhard, T. and Salibra, A. (2010) Exponentials with infinite multiplicities. In: Computer Science Logic 2010. *Springer-Verlag Lecture Notes in Computer Science* **6247** 170–184.

Carraro, A., Ehrhard, T. and Salibra, A. (2010) Resource combinatory algebras. In: Mathematical Foundations of Computer Science 2010. *Springer-Verlag Lecture Notes in Computer Science* **6281** 233–245.

Church, A. (1932) A set of postulates for the foundation of logic. *Annals of Mathematics* **33** 346–366.

Church, A. (1941) *The Calculi of Lambda-Conversion*, Princeton University Press.

de Carvalho, D. (2007) *Sémantiques de la logique linéaire et temps de calcul*, Ph.D. thesis, Université de la Méditerranée, Aix-Marseille II.

Dezani-Ciancaglini, M., de'Liguoro, U. and Piperno, A. (1996) Filter models for conjunctive–disjunctive lambda-calculi. *Theoretical Computer Science* **170** (1–2) 83–128.

Ehrhard, T. (2005) Finiteness spaces. *Mathematical Structures in Computer Science* **15** (4) 615–646.

Ehrhard, T. (2009) The Scott model of linear logic is the extensional collapse of its relational model.

Ehrhard, T. and Regnier, L. (2003) The differential lambda-calculus. *Theoretical Computer Science* **309** (1–3) 1–41.

Ehrhard, T. and Regnier, L. (2006a) Böhm trees, Krivine's machine and the Taylor expansion of lambda-terms. In: CiE. *Springer-Verlag Lecture Notes in Computer Science* **3988** 186–197.

Ehrhard, T. and Regnier, L. (2006b) Differential interaction nets. *Theoretical Computer Science* **364** (2) 166–195.

Ehrhard, T. and Regnier, L. (2008) Uniformity and the Taylor expansion of ordinary lambda-terms. *Theoretical Computer Science* **403** (2–3) 347–372.

Engeler, E. (1981) Algebras and combinators. *Algebra Universalis* **13** (3) 389–392.

Girard, J.-Y. (1988) Normal functors, power series and λ-calculus. *Annals of Pure and Applied Logic* **37** (2) 129–177.

Harmer, R. and McCusker, G. (1999) A fully abstract game semantics for finite nondeterminism. In: *Proceedings, Fourteenth Annual IEEE Symposium on Logic in Computer Science*, IEEE Computer Society Press 422–430.

Hyland, J. M. E., Nagayama, M., Power, J. and Rosolini, G. (2006) A category theoretic formulation for Engeler-style models of the untyped λ-calculus. *Electronic Notes in Theoretical Computer Science* **161** 43–57.

Karoubi, M. (1978) *K-theory*, Springer-Verlag.

Koymans, C. P. J. (1982) Models of the lambda calculus. *Information and Control* **52** (3) 306–332.

Laird, J., Manzonetto, G. and McCusker, G. (2011) Constructing differential categories and deconstructing categories of games. In: Aceto, L., Henzinger, M. and Sgall, J. (eds.) ICALP (2). *Springer-Verlag Lecture Notes in Computer Science* **6756** 186–197.

Lusin, S. and Salibra, A. (2004) The lattice of lambda theories. *Journal of Logic and Computation* **14** 373–394.

Manzonetto, G. (2009) A general class of models of $\mathcal{H}^\star$. In: Mathematical Foundations of Computer Science 2009 (MFCS'09). *Springer-Verlag Lecture Notes in Computer Science* **5734** 574–586.

Manzonetto, G. and Pagani, M. (2011) Böhm theorem for resource lambda calculus through Taylor expansion. In: Typed Lambda Calculi and Applications (TLCA'11). *Springer-Verlag Lecture Notes in Computer Science* **6690** 153–168.

Manzonetto, G. and Salibra, A. (2010) Applying universal algebra to lambda calculus. *Journal of Logic and Computation* **20** (4) 877–915.

Martini, S. (1992) Categorical models for non-extensional λ-calculi and combinatory logic. *Mathematical Structures in Computer Science* **2** (3) 327–357.

Melliès, P.-A., Tabareau, N. and Tasson, C. (2009) An explicit formula for the free exponential modality of linear logic. In: Albers, S., Marchetti-Spaccamela, A., Matias, Y., Nikoletseas, S. and Thomas, W. (eds.) Automata, Languages and Programming. *Springer-Verlag Lecture Notes in Computer Science* **5556** 247–260.

Pagani, M. and Ronchi Della Rocca, S. (2010) Solvability in resource lambda-calculus. In: Foundations of Software Science and Computational Structures (FOSSACS'10). *Springer-Verlag Lecture Notes in Computer Science* **6014** 358–373.

Pagani, M. and Tranquilli, P. (2009) Parallel reduction in resource λ-calculus. In: APLAS. *Springer-Verlag Lecture Notes in Computer Science* **5904** 226–242.

Salibra, A. (2000) On the algebraic models of lambda calculus. *Theoretical Computer Science* **249** 197–240.

Salibra, A. (2001) A continuum of theories of lambda calculus without semantics. In: *16th Annual IEEE Symposium on Logic in Computer Science*, IEEE Computer Society Press 334–343.

Scott, D. (1980) Relating theories of the lambda-calculus. In: Hindley, J. R. and Seldin, J. P. (eds.) *Essays on Combinatory Logic, Lambda-Calculus, and Formalism*, Academic Press 589–606.

Scott, D. S. (1972) Continuous lattices. In: *Toposes, algebraic geometry and logic*, Springer-Verlag.

Tranquilli, P. (2009) *Nets between determinism and nondeterminism*, Ph.D. thesis, Univ. of Paris 7 and Univ. of Roma 3.

Vaux, L. (2007a) The differential $\lambda\mu$-calculus. *Theoretical Computer Science* **379** (1–2) 166–209.

Vaux, L. (2007b) *λ-calcul différentiel et logique classique: interactions calculatoires*, Ph.D. thesis, Univ. de la Méditerranée.

Vaux, L. (2009) Primitive recursion in finiteness spaces. Technical report, Laboratoire de Mathématiques de l'Université de Savoie.