University of Oxford

Fourth Year Project

Simply-typed Differential and Resource λ-Calculus

Carol Mak

St. Catherine's College

supervised by Dr. Luke Ong

A project submitted for the degree of

MMathCompSci Mathematics and Computer Science

Trinity 2016

Abstract

The differential λ-calculus augments the λ-calculus with differential operators that mimic the rules of the standard differential calculus. The extension, and an equivalent calculus, the resource λ-calculus, give expression to resource usage of a computation. Bucciarelli *et al.* have shown that cartesian closed differential categories are models of simply-typed differential λ-theories. This project proves the converse, which is a form of *completeness*: given a typed differential λ-theory, we construct the "smallest" category in which one can soundly model the theory. Moreover, we show that, under reasonable assumptions, differential λ-theory is the internal language of cartesian closed differential category. Finally, we present the relational model as a cartesian closed differential category and show that it is incomplete.

Contents

1 Introduction

λ-calculus is introduced by Church in [Chu41] as a foundation of mathematics which, instead of focusing on sets, focuses on functions. It proves to be a consistent and elegant logical system for computation. Indeed, many functional programming languages, e.g. Haskell, OCaml and Lisp, are based heavily on λ-calculus. However, λ-calculus does not provide explicit information on how much resource is required in evaluating a program, since there is no restriction on the number of times an argument is being used in a function. In an environment where resource is limited, this is a problem. Ehrhard and Regnier in [ER03] designed the *differential λ-calculus* which is an extension of λ-calculus and is resource-sensitive in the sense that we would know exactly if arguments would actually be used during the evaluation of the program. After studying models of linear logic where the existence of differential operator naturally arises, they made a key observation that the linear notion in logic (using arguments exactly once) coincides with the linear notion in algebra, which is strongly suggestive in the terminology of linear logic suggested by Girard in [Gir87]. They drew on these insights and presented this new calculus.

Differential λ-calculus Differential λ-calculus was introduced to explore differentiation as a syntactic operation in λ-calculus by extending it with differential operators, mimicking the rules of the algebraic differential calculus. It extends λ-calculus in two directions. First they introduced a new restricted binary operation representing the derivative of a function along a point, called *differential application* and denoted as $\mathsf{D}s \cdot t$, which has the intuitive meaning that s is provided with exactly one copy of t. With this new application, the authors introduced a new corresponding *differential substitution*, denoted as $\dfrac{\partial s}{\partial x} \cdot t$, which represents the result of substituting exactly one (non-deterministically chosen) linear occurrences of x in s by t. Since differential substitution is non-deterministic, if there are multiple linear occurrences of x in s, one has a choice of which linear occurrences of x to be substituted. In the light of this, the authors extended λ-terms in another direction, introducing *sums* to the syntax. The introduction of sums allows differential λ-calculus to be non-deterministic and yet confluent (Church-Rosser property).

This new differential λ-calculus is a breakthrough since we can now apply useful theorems from differential calculus in the context of λ-calculus. For example, in [ER03; ER08], the authors studied the Taylor expansion of a normal λ-term and showed that we can "express a term as an infinite sum of purely differential terms all of which contain only (multi)linear applications and applications to 0", which is potentially useful in controlling programs in an environment with limited resources.

Resource λ-calculus Ehrhard and Regnier were not the first to suggest a variant of λ-calculus that is resource-sensitive. Early in 1993, Boudol drew inspirations from the encoding of λ-calculus into π-calculus by Milner in [Mil92], and suggested *λ-calculus with multiplicities* in [Bou93] which is resource-sensitive. Instead of having an additional application, Boudol's calculus refines λ-calculus by introducing a new type of argument in an application sT, where T is not a term but a *bag of resources*, i.e. a multiset of terms each with a multiplicity, denoted as $T \equiv (M_1^{m_1} | M_2^{m_2} | \dots | M_n^{m_n})$. The multiplicity of the term indicates the number of available copies of the term in the reduction. One can construct a normal λ-term by having a term with an infinite multiplicity. However, this refinement requires lazy β-reduction and explicit substitution in the syntax in order to perform the non-deterministic choice during reduction. Therefore, it is not an extension from λ-calculus and is different from differential λ-calculus. However after the introduction of differential λ-calculus, Pagani and Tranquilli in [PT09] presented the *resource λ-calculus* which is similar to Boudol's calculus but "enriched with the dynamics of differential λ-calculus", as stated in [Tra11], and is translatable to and from differential λ-calculus as shown in [BEM10; Man12].

Categorical model Lambek first discovered the relation between cartesian closed categories and simply-typed λ-calculus in [Lam80]. He showed that any model of a typed λ-theory is a cartesian closed category and any cartesian closed category is a model of some typed λ-theory. We describe this relationship by saying that typed λ-theory is the *internal language* of cartesian closed category. After the introduction of differential λ-calculus, it is natural to ask what kind of category has this differential λ-theory as its internal language.

Drawing on the insights of defining differentiation in λ-calculus from [ER03], Blute, Cockett and Seely in [BCS06] abstracted the notion of differentiation categorically and introduced the *differential category*. In this category, linear maps are represented as morphisms and the differentiable maps are represented as morphisms in its coKleisli category. In the subsequent paper [BCS09], the same authors directly axiomatize the differentiable maps, i.e. characterize the coKleisli structure of differential categories, and introduced the *cartesian differential category*, which is a left additive category with finite products and a cartesian differential operator that satisfies a set of more complicated conditions. In this category, differentiable maps are represented as morphisms, which matches the intuitive meaning of differential λ-terms. However, since the cartesian differential operator of cartesian differential category does not behave well with exponentials, one cannot model differential λ-calculus in it.

Bucciarelli, Ehrhard and Manzonetto added a new condition for the cartesian differential operator in [BEM10] which ensures that the operator behaves well with exponentials and proved that one can soundly model differential λ-theories in this category, which they called *cartesian closed differential category*. It is known *in the folklore* that the completeness result is true, i.e. given a differential λ-theory, one can construct a *classifying*, i.e. "smallest", cartesian closed differential category such that this theory can be modelled soundly in.

Contributions The highlight of this project is to prove this completeness theorem. We first construct a category according to the syntax of the given differential λ-theory and prove that it is indeed a cartesian closed differential category that soundly models the theory. After that, we show that it is classifying by establishing an equivalence between the category of cartesian closed differential functors and the category of differential models. Assuming that we can extend the theory with constants and function symbols, I conclude that differential λ-theory is the *internal language* of cartesian closed differential category. This result allows us to reason about cartesian closed differential category using differential λ-theory. Finally, I discuss the notion of complete categories with respect to differential λ-theories.

Outline The project is split into two parts: syntax and category. We first look at the syntax of differential λ-calculus in section 2, and the syntax of resource λ-calculus in section 3. Then, we conclude part one by looking at translation maps between the calculi in section 4. After presenting the syntax, section 5 describes the cartesian closed differential category and presents the categorical interpretation of differential λ-calculus in it. Following that, we prove the soundness and completeness theorem and show that differential λ-theory is the internal language of cartesian closed differential category and give some examples on how to prove properties of the category using the theory. Finally, we discuss the notion of complete categories and show that the relational model, which is a classic example of a carteisan closed differential category, is incomplete.

There are two appendices attached at the end of this project. Appendix A presents the relationship between simply-typed λ-calculus and cartesian closed category, proving both soundness and completeness results. This is meant for readers who are not familiar with these results. Appendix B provides the proofs of some lemmas and propositions.

2 Differential λ-Calculus

Differential λ-calculus is introduced by Ehrhard and Regnier in [ER03] as a λ-calculus with a syntactic operation "differentiation" that mimics differential calculus.

2.1 Differential λ-Terms

Let's look at the definition of differentiation in differential calculus. As defined in [Die60], a function $f : U \to V$ on vector spaces is said to be differentiable at $x \in U$ if there is a linear function $f'(x) : U \to V$ that best approximates the slope of f at x. Given $u \in U$, $f'(x) \cdot u$ can be read as the linear application of $f'(x)$ to u. So, the function $x \mapsto f'(x) \cdot u$ from U to V would be linear in u. We denote this function as $Df \cdot u : U \to V$ and call it the derivative of f along u.

Keeping this in mind, one can extend λ-terms by introducing a new construct that mimics this derivative. Given two terms s and t, $\mathsf{D}s \cdot t$ is defined to be the *differential application* of s to t, i.e. s is provided with exactly one copy of t. We can look at $\mathsf{D}s \cdot t$ as the derivative of s along t. The reason for including *sums* in this calculus will become apparent when we move onto the definition of differential substitution.

The syntax stated here is borrowed from [Vau07].

Definition 2.1 (Differential λ-Terms)**.** Assume we have an infinitely countable set of variables $\mathcal{V}$. The collection Λ^d of *differential λ-terms* and the collection Λ^s of *simple terms* are defined by mutual induction as follows:

$$\begin{aligned} \Lambda^s : \quad & s, t, u, v ::= x \mid \lambda x.s \mid s\,T \mid \mathsf{D}s \cdot t \mid \langle\rangle \mid \langle s, t\rangle \mid \mathsf{Fst}(s) \mid \mathsf{Snd}(s) \\ \Lambda^d : \quad & S, T, U, V ::= 0 \mid s \mid s + T \end{aligned}$$

where $x \in \mathcal{V}$.

We consider differential λ-terms up to α-conversion, and up to associativity and commutativity of the sum. We write $S \equiv T$ if S and T are syntactically equivalent. The term 0 is the unit element of the sum, thus $S + 0 \equiv S \equiv 0 + S$. The set $\mathsf{FV}(S)$ of free variables of S and the capture-free substitution of x by T in S, denoted by $S[T/x]$, are defined as usual.

Let us explore more on the linearity of differential λ-calculus. Adopting the notion of linear from linear logic, we say that a subterm of a term s is in a *linear position* if the subterm is used exactly once in the reduction of s. To be more precise, we say that

- x is in a linear position in variable x;
- s is in a linear position in the abstraction $\lambda x.s$;
- s is in a linear position in the application $s\,T$;
- s and t are both in a linear position in the differential application $\mathsf{D}s \cdot t$;
- s and t are both in a linear position in the pair $\langle s, t\rangle$;
- s and T are both in a linear position in the sum $s + T$.

In contrast to differential application, the normal application is linear in the function but not in the argument. This makes sense as there is no restriction on the number of times the argument is being used in the function in a normal application.

Example 2.1. Consider the subterms of the differential λ-term $S \equiv \lambda y.\big((\mathsf{D}x \cdot y)\,(x+y)\big)$.

- x and y are both in a linear position in $\mathsf{D}x \cdot y$ and $x + y$,
- $\mathsf{D}x \cdot y$ is in a linear position in the application $(\mathsf{D}x \cdot y)\,(x + y)$, but $x + y$ is not,
- $(\mathsf{D}x \cdot y)\,(x + y)$ is in a linear position in the abstraction $\lambda y.\big((\mathsf{D}x \cdot y)\,(x + y)\big)$.

So, $x + y$ is the only subterm of S that is not in a linear position in any subterm of S. However, x and y are in a linear position in $x + y$.

To clarify the coincidence of the two notions of linearity, consider the space of differentiable functions from U to V. For any differentiable functions $f, g : U \to V$, the sum of the two functions is defined to be $(f + g)(u) = f(u) + g(u)$ for any $u \in U$. Moreover, we also have $D(f + g) \cdot u = Df \cdot u + Dg \cdot u$ and $Df \cdot (u + v) = Df \cdot u + Df \cdot v$ for any $u, v \in U$. If we consider terms as differentiable functions, then we have

$$(s + t)u = su + tu \qquad \mathsf{D}(s + t) \cdot u = \mathsf{D}s \cdot u + \mathsf{D}t \cdot u$$
$$\lambda x.(s + t) = \lambda x.s + \lambda x.t \qquad \mathsf{D}s \cdot (u + v) = \mathsf{D}s \cdot u + \mathsf{D}s \cdot v$$

In the light of this, we have the following abbreviations.

Notation 2.1. Note that sums are only allowed in the argument of an application, so the definition given above does not allow the followings as terms, but we will find the following *syntactic sugars* useful.

$$\lambda x.\Big(\sum_{i=1}^{n} s_i\Big) \equiv \sum_{i=1}^{n} \lambda x.s_i \qquad \mathsf{D}\Big(\sum_{i=1}^{n} s_i\Big) \cdot \Big(\sum_{j=1}^{m} t_i\Big) \equiv \sum_{i=1,j=1}^{n,m} \mathsf{D}s_i \cdot t_j$$
$$\Big(\sum_{i=1}^{n} s_i\Big)\, T \equiv \sum_{i=1}^{n} s_i\, T \qquad \Big\langle \sum_{i=1}^{n} s_i, \sum_{j=1}^{m} t_j \Big\rangle \equiv \langle s_1, t_1\rangle + \Big\langle \sum_{i=2}^{n} s_i, \sum_{j=2}^{m} t_j \Big\rangle$$

The definition of differential λ-terms allow the argument of an application to be a sum of terms. Since it is not in a linear position, in general,

$$s\,\Big(\sum_{i=1}^{n} T_i\Big) \not\equiv \sum_{i=1}^{n} s\, T_i.$$

Note that if we sum over 0 terms, 0 annihilates all things except when it is the argument of an application.

2.2 Differential Substitution

After introducing differential application which mimics the derivative $Df \cdot u : x \mapsto f'(x) \cdot u$, the authors introduced another operation which mimics the result of applying $f'(x)$ to u linearly. This operation is similar to substitution, and is given the name *differential substitution*, denoted as $\dfrac{\partial S}{\partial x} \cdot T$. It replaces exactly one linear occurrences of x (non-deterministically chosen) by T in S.

Consider the situation where there are multiple linear occurrences of a variable in a term. To perform differential substitution, one has a choice of which linear occurrences of the variable to be substituted. So, this substitution becomes a non-deterministic operation. The authors of [ER03] use *sums* to interpret this non-deterministic property. Thus, differential substitution $\frac{\partial S}{\partial x} \cdot T$ is actually the sum of all possible terms obtained by substituting exactly one linear occurrences of the variable x by T in the term S. Furthermore, if there is no such occurrence in S, the result will be 0.

Definition 2.2 (Differential substitution)**.** Let S, T be differential λ-terms and x be a variable. The *differential substitution* of x by T in S, denoted by $\frac{\partial S}{\partial x} \cdot T$, is defined by induction on S as follows:

$$\begin{aligned}
\frac{\partial y}{\partial x} \cdot T &\equiv \begin{cases} T & \text{if } x \equiv y, \\ 0 & \text{otherwise.} \end{cases} \\
\frac{\partial}{\partial x}(\lambda y.s) \cdot T &\equiv \lambda y.\Big(\frac{\partial s}{\partial x} \cdot T\Big) \quad \text{assuming } x \not\equiv y \text{ and } y \notin \mathsf{FV}(T) \\
\frac{\partial}{\partial x}(s\,U) \cdot T &\equiv \Big(\frac{\partial s}{\partial x} \cdot T\Big)\,U + \bigg(\mathsf{D}s \cdot \Big(\frac{\partial U}{\partial x} \cdot T\Big)\bigg)\,U \\
\frac{\partial}{\partial x}(\mathsf{D}s \cdot u) \cdot T &\equiv \mathsf{D}\Big(\frac{\partial s}{\partial x} \cdot T\Big) \cdot u + \mathsf{D}s \cdot \Big(\frac{\partial u}{\partial x} \cdot T\Big) \\
\frac{\partial}{\partial x}\langle\rangle \cdot T &\equiv \langle\rangle \\
\frac{\partial}{\partial x}\langle s, u\rangle \cdot T &\equiv \Big\langle \frac{\partial s}{\partial x} \cdot T, \frac{\partial u}{\partial x} \cdot T \Big\rangle \\
\frac{\partial}{\partial x}\mathsf{Fst}(s) \cdot T &\equiv \mathsf{Fst}\Big(\frac{\partial s}{\partial x} \cdot T\Big) \\
\frac{\partial}{\partial x}\mathsf{Snd}(s) \cdot T &\equiv \mathsf{Snd}\Big(\frac{\partial s}{\partial x} \cdot T\Big) \\
\frac{\partial 0}{\partial x} \cdot T &\equiv 0 \\
\frac{\partial}{\partial x}(s + U) \cdot T &\equiv \frac{\partial s}{\partial x} \cdot T + \frac{\partial U}{\partial x} \cdot T
\end{aligned}$$

Since differential application is a bilinear operator, a linear occurrence of x in $\mathsf{D}s \cdot t$ is either a linear occurrence in s or t, so we can just sum over both possibilities and obtain

$$\mathsf{D}\big(\frac{\partial s}{\partial x} \cdot T\big) \cdot u + \mathsf{D}s \cdot \big(\frac{\partial u}{\partial x} \cdot T\big).$$

However, this is not the case for normal application $s\,U$. If the linear occurrence of x is in s, we can apply the substitution directly to s as it is in a linear position, producing the term $\big(\frac{\partial s}{\partial x} \cdot T\big)\,U$. However, if the linear occurrence of x is in U, we would need to "linearise" the application by "extracting" a linear copy of U, obtaining $(\mathsf{D}s \cdot U)\,U$, and perform the substitution to the linear copy.

$$s\,U \xrightarrow[\text{linear copy of } U]{\text{extracting a}} (\mathsf{D}s \cdot U)\,U \xrightarrow[\text{the linear copy}]{\text{substitution on}} \bigg(\mathsf{D}s \cdot \Big(\frac{\partial U}{\partial x} \cdot T\Big)\bigg)\,U$$

Again this coincides with the algebraic notion of differentiating a composition of two functions.

$$(f \circ g)'(x) = f'\big(g(x)\big) \cdot g'(x) = \big(Df \cdot g'(x)\big)\ g(x)$$

Note that "linearising" applications is not valid in this calculus. i.e. $s\,U \not\equiv (\mathsf{D}s \cdot U)\,U$. It is only used implicitly in the differential substitution of application.

Example 2.2. Let x, y and z be distinct variables. Consider the differential substitution of x by z in the term $S \equiv \lambda y.\big((\mathsf{D}x \cdot y)\ (x+y)\big)$.

$$\begin{aligned}
&\frac{\partial}{\partial x}\Big(\lambda y.\big((\mathsf{D}x \cdot y)\ (x+y)\big)\Big) \cdot z \\
&\equiv \lambda y.\Big(\frac{\partial\ (\mathsf{D}x \cdot y)\ (x+y)}{\partial x} \cdot z\Big) \\
&\equiv \lambda y.\Big(\Big(\frac{\partial\ \mathsf{D}x \cdot y}{\partial x} \cdot z\Big)\ (x+y) + \Big(\mathsf{D}(\mathsf{D}x \cdot y) \cdot \Big(\frac{\partial (x+y)}{\partial x} \cdot z\Big)\Big)\ (x+y)\Big) \\
&\equiv \lambda y.\Big(\Big(\mathsf{D}\big(\frac{\partial x}{\partial x} \cdot z\big) \cdot y + \mathsf{D}x \cdot \big(\frac{\partial y}{\partial x} \cdot z\big)\Big)\ (x+y) + \\
&\qquad \Big(\mathsf{D}(\mathsf{D}x \cdot y) \cdot \Big(\frac{\partial x}{\partial x} \cdot z + \frac{\partial y}{\partial x} \cdot z\Big)\Big)\ (x+y)\Big) \\
&\equiv \lambda y.\Big((\mathsf{D}z \cdot y + \mathsf{D}x \cdot 0)\ (x+y) + \big(\mathsf{D}(\mathsf{D}x \cdot y) \cdot (z+0)\big)\ (x+y)\Big) \\
&\equiv \lambda y.\Big((\mathsf{D}z \cdot y)\ (x+y)\Big) + \lambda y.\Big(\big(\mathsf{D}(\mathsf{D}x \cdot y) \cdot z\big)\ (x+y)\Big)
\end{aligned}$$

As discussed in example 2.1, the only subterm that is not in a linear position of any other subterm of S is $x+y$. Yet, x is a linear occurrence in $x+y$. In the differential substitution $\frac{\partial S}{\partial x} \cdot z$, we consider all possible terms obtained by substituting one linear occurrences of the variable x by z in the term S. Hence, we also substitute this linear occurrence of x in $x+y$ and obtain $\lambda y.\Big(\big(\mathsf{D}(\mathsf{D}x \cdot y) \cdot z\big)\ (x+y)\Big)$.

The following proposition tells us that if there is no free occurrences of x in a term S, the result of differential substituting x in S is 0.

Proposition 2.1. If $x \notin \mathsf{FV}(S)$, then $\frac{\partial S}{\partial x} \cdot T \equiv 0$.

Proof. Easy induction on the structure of S. □

We have established enough machinery to define differentiation in this calculus. Recall the algebraic notion of differentiation. The derivative of a function f along u is the function

$$Df \cdot u : \quad x \longmapsto f'(x) \cdot u.$$

In this calculus, we reduce the derivative of a function, i.e. an abstraction $\lambda x.s$, along a term t to another abstraction similar to $\lambda x.s$, $\lambda x.s'$, where s' is the term where exactly one linear occurrence of x (non-deterministically chosen) is substituted by t. i.e.

$$\mathsf{D}(\lambda x.s) \cdot t \longrightarrow \lambda x.\Big(\frac{\partial s}{\partial x} \cdot t\Big)$$

Example 2.3. Consider the differential application of $\lambda x.xz(yz)$ to w.

$$\begin{aligned}
\mathsf{D}\Big(\lambda x.xz(yz)\Big) \cdot w &\longrightarrow \lambda x.\Big(\frac{\partial\ xz(yz)}{\partial x} \cdot w\Big) \\
&\equiv \lambda x.\Big(\big(\frac{\partial xz}{\partial x} \cdot w\big)\ (yz) + \big(\mathsf{D}xz \cdot (\frac{\partial yz}{\partial x} \cdot w)\big)\ (yz)\Big) \\
&\equiv \lambda x.\Big(\Big(\big(\frac{\partial x}{\partial x} \cdot w\big)\ z + \big(\mathsf{D}x \cdot (\frac{\partial z}{\partial x} \cdot w)\big)\ z\Big)\ (yz) + 0\Big) \\
&\equiv \lambda x.\big((wz+0)\ (yz)\big) \\
&\equiv \lambda x.wz(yz)
\end{aligned}$$

There is only one linear occurrence of x in $\mathbf{s}$ and it is substituted by w. Note that the result of the differential substitution remains an abstraction and, in this case, that is the only difference between the reduction of the differential application $\mathsf{D}\Big(\lambda x.xz(yz)\Big) \cdot w$ and the classical application $\big(\lambda x.xz(yz)\big)\ w$, where

$$\big(\lambda x.xz(yz)\big)\ w \longrightarrow wz(yz).$$

Let us consider the differential application of $\lambda x.xx$ to y.

$$\begin{aligned}\mathsf{D}(\lambda x.xx) \cdot y \longrightarrow \lambda x.\Big(\frac{\partial xx}{\partial x} \cdot y\Big) & \\ \equiv \lambda x.\Big(\big(\frac{\partial x}{\partial x} \cdot y\big)\ x + \big(\mathsf{D}x \cdot \big(\frac{\partial x}{\partial x} \cdot y\big)\big)\ x\Big) & \\ \equiv \lambda x.\big(yx + (\mathsf{D}x \cdot y)\ x\big) &\end{aligned}$$

Note that in $\lambda x.xx$, x occurs twice and the reduction only allows exactly one linear occurrences of x being substituted.

2.3 Type System $\mathcal{D}$

Definition 2.3 (Type System $\mathcal{D}$)**.** Assume we have a collection of type variables TV. *Types* and *type contexts* are defined as follows:

Types	$\alpha, \beta ::= \mathsf{unit} \mid \gamma \mid \alpha \times \beta \mid \alpha \Rightarrow \beta$	where $\gamma \in \mathsf{TV}$
Type Contexts	$\Gamma, \Delta ::= \varnothing \mid \Gamma \cup \{x : \alpha\}$	assuming that $\{x : \alpha\} \notin \Gamma$

A differential λ-term s is a *differential typed term* if there is a type context Γ and a type α such that $\Gamma \vdash_{\mathcal{D}} s : \alpha$ is derivable in the type system $\mathcal{D}$ defined by the following rules,

$$(var)\ \frac{}{\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} x : \alpha} \qquad (abs)\ \frac{\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} s : \beta}{\Gamma \vdash_{\mathcal{D}} \lambda x.s : \alpha \Rightarrow \beta}$$

$$(app)\ \frac{\Gamma \vdash_{\mathcal{D}} s : \alpha \Rightarrow \beta \qquad \Gamma \vdash_{\mathcal{D}} t : \alpha}{\Gamma \vdash_{\mathcal{D}} s\,t : \beta}$$

$$(\mathsf{unit})\ \frac{}{\Gamma \vdash_{\mathcal{D}} \langle\rangle : \mathsf{unit}} \qquad (pair)\ \frac{\Gamma \vdash_{\mathcal{D}} s : \alpha \qquad \Gamma \vdash_{\mathcal{D}} t : \beta}{\Gamma \vdash_{\mathcal{D}} \langle s, t\rangle : \alpha \times \beta}$$

$$(\mathsf{Fst})\ \frac{\Gamma \vdash_{\mathcal{D}} p : \alpha \times \beta}{\Gamma \vdash_{\mathcal{D}} \mathsf{Fst}(p) : \alpha} \qquad (\mathsf{Snd})\ \frac{\Gamma \vdash_{\mathcal{D}} p : \alpha \times \beta}{\Gamma \vdash_{\mathcal{D}} \mathsf{Snd}(p) : \beta}$$

$$(\mathsf{D})\ \frac{\Gamma \vdash_{\mathcal{D}} s : \alpha \Rightarrow \beta \qquad \Gamma \vdash_{\mathcal{D}} t : \alpha}{\Gamma \vdash_{\mathcal{D}} \mathsf{D}s \cdot t : \alpha \Rightarrow \beta} \qquad (sum)\ \frac{\Gamma \vdash_{\mathcal{D}} s_i : \alpha \qquad \forall i \in I}{\Gamma \vdash_{\mathcal{D}} \sum_{i \in I} s_i : \alpha}$$

Remark. The differential application $\mathsf{D}s \cdot t$ has the same function type $\alpha \Rightarrow \beta$ as s as it is mimicking the derivative of s along t.

Example 2.4. Let $\Gamma = \{x:\alpha, y:\alpha, z:\alpha\}$ and $\Gamma' = \{y:\alpha, z:\alpha\}$. We show that the term $\mathsf{D}\big(\mathsf{D}(\lambda x.x)\cdot y\big)\cdot z$ is a differential typed term.

$$\dfrac{(\mathsf{D})\dfrac{(abs)\dfrac{(var)\dfrac{}{\Gamma \vdash_{\mathcal{D}} x : \alpha}}{\Gamma' \vdash_{\mathcal{D}} \lambda x.x : \alpha \Rightarrow \alpha} \quad (var)\dfrac{}{\Gamma' \vdash_{\mathcal{D}} y : \alpha}}{\Gamma' \vdash_{\mathcal{D}} \mathsf{D}(\lambda x.x)\cdot y : \alpha \Rightarrow \alpha} \quad (var)\dfrac{}{\Gamma' \vdash_{\mathcal{D}} z : \alpha}}{\Gamma' \vdash_{\mathcal{D}} \mathsf{D}\big(\mathsf{D}(\lambda x.x)\cdot y\big)\cdot z : \alpha \Rightarrow \alpha}\,(\mathsf{D})$$

Lemma 2.2 (Substitution Lemma for $\mathcal{D}$)**.** Let $\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} S : \beta$ and $\Gamma \vdash_{\mathcal{D}} T : \alpha$. We have

(i) $\Gamma \vdash_{\mathcal{D}} S[T/x] : \beta$

(ii) $\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} \dfrac{\partial S}{\partial x} \cdot T : \beta$

Proof. Easy induction on structure of S. □

2.4 Differential λ-theory

Definition 2.4. A *simply-typed theory* $\mathscr{T}$ with respect to a type system $\mathcal{T}$ is a collection of rules of the form $\Gamma \vdash_{\mathcal{T}} s = t : \alpha$, where $\Gamma \vdash_{\mathcal{T}} s : \alpha$ and $\Gamma \vdash_{\mathcal{T}} t : \alpha$ are derivable. We write $\mathscr{T} \triangleright \Gamma \vdash_{\mathcal{T}} s = t : \alpha$ to indicate that $\Gamma \vdash_{\mathcal{T}} s = t : \alpha$ is a rule in $\mathscr{T}$. The $\lambda\beta\eta_{\mathsf{D}}$-theory is the smallest theory that is closed under the following rules,

$$(app)\,\dfrac{\Gamma \vdash_{\mathcal{D}} s = s' : \alpha \Rightarrow \beta \qquad \Gamma \vdash_{\mathcal{D}} t = t' : \alpha}{\Gamma \vdash_{\mathcal{D}} st = s't' : \beta} \qquad (abs)\,\dfrac{\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} s = t : \beta}{\Gamma \vdash_{\mathcal{D}} \lambda x.s = \lambda x.t : \alpha \Rightarrow \beta}$$

$$(\beta)\,\dfrac{\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} s : \beta \qquad \Gamma \vdash_{\mathcal{D}} t : \alpha}{\Gamma \vdash_{\mathcal{D}} (\lambda x.s)\,t = s\,[t/x] : \beta} \qquad (\eta)\,\dfrac{x \notin \mathsf{FV}(s)}{\Gamma \vdash_{\mathcal{D}} \lambda x.s\,x = s : \alpha \Rightarrow \beta}$$

$$(\mathsf{unit})\,\dfrac{\Gamma \vdash_{\mathcal{D}} s : \mathsf{unit}}{\Gamma \vdash_{\mathcal{D}} s = \langle\rangle : \mathsf{unit}} \qquad (pair)\,\dfrac{\Gamma \vdash_{\mathcal{D}} p : \alpha \times \beta}{\Gamma \vdash_{\mathcal{D}} \langle \mathsf{Fst}(p), \mathsf{Snd}(p)\rangle = p : \alpha \times \beta}$$

$$(\mathsf{Fst})\,\dfrac{\Gamma \vdash_{\mathcal{D}} s : \alpha \qquad \Gamma \vdash_{\mathcal{D}} t : \beta}{\Gamma \vdash_{\mathcal{D}} \mathsf{Fst}(\langle s, t\rangle) = s : \alpha} \qquad (\mathsf{Snd})\,\dfrac{\Gamma \vdash_{\mathcal{D}} s : \alpha \qquad \Gamma \vdash_{\mathcal{D}} t : \beta}{\Gamma \vdash_{\mathcal{D}} \mathsf{Snd}(\langle s, t\rangle) = t : \beta}$$

$$(app_{\mathsf{D}})\,\dfrac{\Gamma \vdash_{\mathcal{D}} s = t : \alpha \Rightarrow \beta \qquad \Gamma \vdash_{\mathcal{D}} u = v : \alpha}{\Gamma \vdash_{\mathcal{D}} \mathsf{D}s \cdot u = \mathsf{D}t \cdot v : \alpha \Rightarrow \beta}$$

$$(\beta_{\mathsf{D}})\,\dfrac{\Gamma \vdash_{\mathcal{D}} \mathsf{D}(\lambda x.s)\cdot t : \alpha \Rightarrow \beta}{\Gamma \vdash_{\mathcal{D}} \mathsf{D}(\lambda x.s)\cdot t = \lambda x.\big(\dfrac{\partial s}{\partial x}\cdot t\big) : \alpha \Rightarrow \beta}$$

$$(sw_{\mathsf{D}})\,\dfrac{}{\Gamma \vdash_{\mathcal{D}} \mathsf{D}\;(\mathsf{D}s \cdot t)\cdot u = \mathsf{D}\;(\mathsf{D}s \cdot u)\cdot t : \alpha} \qquad (sum)\,\dfrac{\Gamma \vdash_{\mathcal{D}} s_i = t_i : \alpha \qquad \forall i \in I}{\Gamma \vdash_{\mathcal{D}} \sum_{i\in I} s_i = \sum_{i\in I} t_i : \alpha}$$

with the usual reflexive, symmetric and transitivity rules.

We say a simply-typed theory $\mathscr{T}$ is a *differential* λ*-theory* if it is closed under all rules in $\lambda\beta\eta_{\mathsf{D}}$.

Remark. Since we are only considering simply-typed theories in this project, we write theory meaning simply-typed theory.

The following lemma is a direct consequence of the (sw_{D}) rule. It shows that differential substitution of the same variable is unordered.

Lemma 2.3. If $x \notin \mathsf{FV}(T) \cup \mathsf{FV}(U)$, then $\dfrac{\partial}{\partial x}\Big(\dfrac{\partial S}{\partial x} \cdot T\Big) \cdot U \equiv \dfrac{\partial}{\partial x}\Big(\dfrac{\partial S}{\partial x} \cdot U\Big) \cdot T$.

Proof. Easy induction on the structure of S. □

Notation 2.2. Writing $\vec{t} \equiv t_1, \ldots, t_k$, we have the following syntactic sugars

$$\mathsf{D}^k s \cdot \vec{t} \equiv \mathsf{D}(\ldots(\mathsf{D}(\mathsf{D}s \cdot t_1) \cdot t_2) \ldots) \cdot t_k \qquad \frac{\partial^k s}{\partial x^k} \cdot \vec{t} \equiv \frac{\partial}{\partial x}\Big(\ldots\Big(\frac{\partial}{\partial x}\Big(\frac{\partial s}{\partial x} \cdot t_1\Big) \cdot t_2\Big) \ldots\Big) \cdot t_k.$$

Let $\hat{\vec{t}}$ be a permutation of $\vec{t}$ and $\mathscr{T}$ be a differential λ-theory. By the (sw_{D}) rule, we have

$$\mathscr{T} \rhd \quad \Gamma \vdash_{\mathcal{D}} \mathsf{D}^k s \cdot \hat{\vec{t}} = \mathsf{D}^k s \cdot \vec{t} : \alpha.$$

And, by Lemma 2.3, if $x \notin \mathsf{FV}(s) \cup \bigcup_{i=1}^k \mathsf{FV}(t_i)$, then

$$\mathscr{T} \rhd \quad \Gamma \vdash_{\mathcal{D}} \frac{\partial^k s}{\partial x^k} \cdot \hat{\vec{t}} = \frac{\partial^k s}{\partial x^k} \cdot \vec{t} : \alpha.$$

Moreover by the (β_{D}) rule, we know that

$$\mathscr{T} \rhd \quad \Gamma \vdash_{\mathcal{D}} \mathsf{D}^k(\lambda x.s) \cdot \vec{t} = \lambda x.\Big(\frac{\partial^k s}{\partial x^k} \cdot \vec{t}\Big) : \alpha.$$

Example 2.5. In example 2.4, we showed that $\Gamma' \vdash_{\mathcal{D}} \mathsf{D}\big(\mathsf{D}(\lambda x.x) \cdot y\big) \cdot z : \alpha \Rightarrow \alpha$. In any differential λ-theory $\mathscr{T}$, we can write $\mathsf{D}\big(\mathsf{D}(\lambda x.x) \cdot y\big) \cdot z$ as $\mathsf{D}^2(\lambda x.x) \cdot (y,z)$, and have

$$\mathscr{T} \rhd \quad \Gamma' \vdash_{\mathcal{D}} \mathsf{D}^2(\lambda x.x) \cdot (y,z) = \lambda x.\Big(\frac{\partial^2 x}{\partial x^2} \cdot (y,z)\Big) \equiv 0 : \alpha \Rightarrow \alpha.$$

3 Resource λ-Calculus

Resource λ-calculus is suggested by Pagani and Tranquilli in [PT09] and refined by Bucciarelli *et al.* in [BEM10], which is based on Boudol's calculus [Bou93] but "enriched with the dynamics of differential λ-calculus".

Resource λ-calculus has a *bag of resources* as the arguments in an application, where a resource is either a *reusable term*, with a ! superscript, or a *linear term*. As the names suggest, a reusable term is a term that is infinitely available, and a linear term provides exactly one copy of itself. Note that this way of indicating a reusable term comes from Girard's linear logic in [Gir87]. In the light of differential λ-calculus, *sums* are introduced to this calculus to represent all the possible results of a computation. In addition to the capture-free substitution, *resource substitution* is also added. Moreover, reductions are not lazy as in Boudol's calculus. All these make it possible to define translations between this calculus and differential λ-calculus.

3.1 Resource Terms and Bags of Resources

Recall 3.1 (Multisets)**.** A *multiset* m over a set S is an unordered list with repetitions, denoted as $m = [a_1, a_2, \dots]$, where $a_i \in S$ for all i. We say m is *finite* if it is a finite list. The *union* of multisets $m_1 = [a_1, a_2, \dots]$ and $m_2 = [b_1, b_2, \dots]$ is $m_1 \uplus m_2 = [a_1, b_1, a_2, b_2, \dots]$. We write $\mathcal{N}(S)$ as the set of all finite multisets over S. We denote $\mathbf{1}$ as the empty multiset $[\,]$.

Example 3.1. (i) $[1, 1, 2] \uplus [2, 3] = [1, 1, 2, 2, 3]$

(ii) $\mathcal{N}(\omega) = \{[n_1, n_2, \dots, n_m] : m \geq 0, n_i \in \omega\}$

Definition 3.1 (Resource Terms and Bags)**.** Assume we have an infinitely countable set of variables $\mathcal{V}$. The collection Λ^r of *resource terms*, the collection $\Lambda^{(!)}$ of *resources* and the collection Λ^b of *bags* are defined by mutual induction as follows, where $x \in \mathcal{V}$:

$$
\begin{aligned}
\Lambda^r: &\quad M, N, L ::= x \mid \lambda x.M \mid M\,P \mid \langle\rangle \mid \langle M, N\rangle \mid \mathsf{Fst}(M) \mid \mathsf{Snd}(M) && \text{(terms)}\\
\Lambda^{(!)}: &\quad M^{(!)}, N^{(!)} ::= M \mid M^! && \text{(resources)}\\
\Lambda^b: &\quad P, Q, R ::= \mathbf{1} \mid P \uplus [M^{(!)}] && \text{(bags)}\\
\Lambda^e: &\quad A, B, C ::= M \mid P && \text{(expressions)}
\end{aligned}
$$

The collection Λ^e of *expressions* is just the (disjoint) union of the collection Λ^r and Λ^b.

Definition 3.2 (Sums)**.** The collection of *sums of terms* is the set of all finite multisets over Λ^r, $\mathcal{N}(\Lambda^r)$ and the collection of *sums of bags* is the set of all finite multisets over Λ^b, $\mathcal{N}(\Lambda^b)$, with 0 referring to the empty sum.

$$
\begin{aligned}
\mathbb{M}, \mathbb{N}, \mathbb{L} &\in \mathcal{N}(\Lambda^r) && \text{(sums of terms)}\\
\mathbb{P}, \mathbb{Q}, \mathbb{R} &\in \mathcal{N}(\Lambda^b) && \text{(sums of bags)}\\
\mathbb{A}, \mathbb{B}, \mathbb{C} &\in \mathcal{N}(\Lambda^r) \cup \mathcal{N}(\Lambda^b) && \text{(sums of expressions)}
\end{aligned}
$$

We consider expressions up to α-conversion and write $A \equiv B$ if A and B are syntactically equivalent. The set $\mathsf{FV}(A)$ of free variables of A and the capture-free substitution of x by a resource term N in an expression A, denoted by $A[N/x]$, are defined as usual. It is easy to see that we can extend both the free variable and substitution to sums as in $\mathsf{FV}(\mathbb{A})$ and $\mathbb{A}[\mathbb{N}/x]$ by linearity in $\mathbb{A}$.

Example 3.2. Similar to Boudol's calculus, we can express any λ-terms using resource terms. For example,

$$\boldsymbol{I} \equiv \lambda x.x \qquad \boldsymbol{\omega} \equiv \lambda x.x\,[x^!] \qquad \boldsymbol{\Omega} \equiv \boldsymbol{\omega}\,[\boldsymbol{\omega}^!] \qquad \boldsymbol{\Delta} \equiv \lambda xy.x[(x\,[x^!, y^!])^!] \qquad \boldsymbol{\Theta} \equiv \boldsymbol{\Delta}\,[\boldsymbol{\Delta}^!]$$

Notation 3.1. The definitions for terms and bags do not include any sums. Thus, the sums are always on the "surface". Nonetheless, we would define the following *syntactic sugars* where we extend all constructors to the sums. All "positions" in resource λ-calculus are linear, except $(-)^!$. So, unlike in differential λ-calculus, application is a bilinear operator in resource λ-calculus. This makes sense since sums only arise on the "surface" in resource λ-calculus, while the argument of an application in differential λ-calculus could be a sum.

$$\lambda x.\Big(\sum_{i=1}^{n} M_i\Big) \equiv \sum_{i=1}^{n} \lambda x.M_i \qquad\qquad \Big[\sum_{i=1}^{n} M_i\Big] \uplus P \equiv \sum_{i=1}^{n} \Big([M_i] \uplus P\Big)$$

$$\Big(\sum_{i=1}^{n} M_i\Big)\Big(\sum_{j=1}^{m} P_j\Big) \equiv \sum_{i=1}^{n}\sum_{j=1}^{m}(M_i P_j) \qquad\qquad \Big[\big(\sum_{i=1}^{n} M_i\big)^!\Big] \uplus P \equiv [M_1^!, \ldots, M_n^!] \uplus P$$

Note that if we sum over 0 expressions, 0 annihilates all things except under $(-)^!$.

3.2 Resource Substitution

Definition 3.3 (Resource Substitution)**.** Let A be an expression, N be a resource term and x be a variable. The *resource substitution*, denoted as $A\langle N/x\rangle$, is defined by induction on A as follows:

$$\begin{aligned}
y\langle N/x\rangle &\equiv \begin{cases} N & \text{if } x \equiv y, \\ 0 & \text{otherwise.} \end{cases}\\
(\lambda y.M)\langle N/x\rangle &\equiv \lambda y.(M\langle N/x\rangle) \quad \text{assuming that } x \not\equiv y,\ y \notin \mathsf{FV}(N)\\
M\,P\langle N/x\rangle &\equiv M\langle N/x\rangle\,P + M\,P\langle N/x\rangle\\
\langle\rangle\langle N/x\rangle &\equiv \langle\rangle\\
\langle M, L\rangle\langle N/x\rangle &\equiv \langle M\langle N/x\rangle, L\langle N/x\rangle\rangle\\
\mathsf{Fst}(M)\langle N/x\rangle &\equiv \mathsf{Fst}(M\langle N/x\rangle)\\
\mathsf{Snd}(M)\langle N/x\rangle &\equiv \mathsf{Snd}(M\langle N/x\rangle)\\
[M]\langle N/x\rangle &\equiv [M\langle N/x\rangle]\\
[M^!]\langle N/x\rangle &\equiv [M\langle N/x\rangle, M^!]\\
\mathbf{1}\langle N/x\rangle &\equiv 0\\
(P \uplus R)\langle N/x\rangle &\equiv P\langle N/x\rangle \uplus R + P \uplus R\langle N/x\rangle
\end{aligned}$$

Similar to capture-free substitution, resource substitution can be extended to sums as in $\mathbb{A}\langle \mathbb{N}/x\rangle$ by *bilinearity* in $\mathbb{A}$ and $\mathbb{N}$.

In differential substitution, we discussed how to linearly substitute a linear occurrence of x to the argument of an application, which is not in a linear position. However, in resource λ-calculus, this is not necessary as application is bilinear. We can safely distributes the substitution over application. The only operator that is not linear in resource λ-calculus is $(-)^!$. Thus, for $[M^!]\langle N/x\rangle$, we again "extract" a linear copy from the reusable resource $M^!$, apply the resource substitution to the linear copy, leaving the reusable copy unchanged and obtain $[M\langle N/x\rangle, M^!]$. This equation

$$\begin{aligned}\big(s\,[U^!]\big)\langle T/x\rangle &\equiv \big(s\langle T/x\rangle\big)\,[U^!] + s\,\big([U^!]\langle T/x\rangle\big)\\ &\equiv \big(s\langle T/x\rangle\big)\,[U^!] + s\,[U\langle T/x\rangle, U^!]\end{aligned}$$

closely resembles the differential substitution of application.

Example 3.3. Consider the following resource substitution.

$$\begin{aligned}&\big(\lambda y.(x\,[y,(x+y)^!])\big)\langle z/x\rangle\\ &\equiv \lambda y.\big((x\,[y,(x+y)^!])\langle z/x\rangle\big)\\ &\equiv \lambda y.\big((x\langle z/x\rangle\big)\,[y,(x+y)^!] + x\,([y,(x+y)^!]\langle z/x\rangle)\big)\\ &\equiv \lambda y.\big(z\,[y,(x+y)^!] + x\,([y\langle z/x\rangle,(x+y)^!] + [y,(x+y)\langle z/x\rangle,(x+y)^!])\big)\\ &\equiv \lambda y.\big(z\,[y,(x+y)^!] + x\,([0,(x+y)^!] + [y,z+0,(x+y)^!])\big)\\ &\equiv \lambda y.\big(z\,[y,(x+y)^!]\big) + \lambda y.\big(x\,[y,z,(x+y)^!]\big)\end{aligned}$$

Note that in example 2.2, we have

$$\frac{\partial}{\partial x}\Big(\lambda y.\big((\mathsf{D}x\cdot y)\ (x+y)\big)\Big)\cdot z \equiv \lambda y.\Big((\mathsf{D}z\cdot y)\ (x+y)\Big) + \lambda y.\Big(\big(\mathsf{D}(\mathsf{D}x\cdot y)\cdot z\big)\ (x+y)\Big)$$

which closely resembles our result here.

3.3 Type System $\mathcal{R}$

Definition 3.4 (Type System $\mathcal{R}$)**.** A sum of expression $\mathbb{A}$ is a *resource typed term* if there is a type context Γ and a type α such that $\Gamma \vdash_{\mathcal{R}} \mathbb{A} : \alpha$ is derivable in the type system $\mathcal{R}$ defined by the following rules,

$$(bag)\ \frac{\Gamma \vdash_{\mathcal{R}} N : \alpha \qquad \Gamma \vdash_{\mathcal{R}} P : \alpha}{\Gamma \vdash_{\mathcal{R}} [N^{(!)}] \uplus P : \alpha} \qquad\qquad (\mathbf{1})\ \frac{}{\Gamma \vdash_{\mathcal{R}} \mathbf{1} : \alpha}$$

$$(sum)\ \frac{\Gamma \vdash_{\mathcal{R}} A : \alpha \qquad \Gamma \vdash_{\mathcal{R}} \mathbb{B} : \alpha \qquad \mathbb{B} \neq 0}{\Gamma \vdash_{\mathcal{R}} A + \mathbb{B} : \alpha}$$

and the usual (var), (abs), (app), (unit), $(pair)$, (Fst) and (Snd) rules.

3.4 Resource λ-theory

Notation 3.2. We will have the following abbreviations,

$$\begin{aligned}\vec{L} &\equiv L_1,\dots,L_n & M[\vec{N}^!/x] &\equiv M[N_1^!/x]\dots[N_k^!/x]\\ \vec{N}^! &\equiv N_1^!,\dots,N_k^! & M\langle\vec{L}/x\rangle &\equiv M\langle L_1/x\rangle\dots\langle L_n/x\rangle\end{aligned}$$

We also write $\vec{L}_{-j} \equiv L_1,\dots,L_{j-1},L_{j+1},\dots,L_n$.

Every application MP can be written uniquely as $M[\vec{L},\vec{N}^!]$, where $P \equiv [\vec{L},\vec{N}^!]$ is a multiset which is unordered.

Definition 3.5. The $\lambda\beta\eta_{\mathsf{R}}$-theory is the smallest theory that is closed under the following rules,

$$(\beta_{\mathsf{R}})\,\frac{\Gamma \cup \{x:\alpha\} \vdash_{\mathcal{R}} M:\beta \qquad \Gamma \vdash_{\mathcal{R}} [\vec{L},\vec{N}^!]:\alpha}{\Gamma \vdash_{\mathcal{R}} (\lambda x.M)\,[\vec{L},\vec{N}^!] = M\,\langle \vec{L}/x\rangle\,[\sum_i N_i/x]:\beta}$$

$$(\eta_{\mathsf{R}})\,\frac{\Gamma \vdash_{\mathcal{R}} M:\alpha \qquad x \notin \mathsf{FV}(M)}{\Gamma \vdash_{\mathcal{R}} \lambda x.M\,[x^!] = M:\alpha}$$

$$(sum)\,\frac{\Gamma \vdash_{\mathcal{R}} M_i = N_i:\alpha \qquad \forall i \in I}{\Gamma \vdash_{\mathcal{R}} \sum_{i\in I} M_i = \sum_{i\in I} N_i:\alpha} \qquad (bag)\,\frac{\Gamma \vdash_{\mathcal{R}} M = N:\alpha \qquad \Gamma \vdash_{\mathcal{R}} P = Q:\alpha}{\Gamma \vdash_{\mathcal{R}} [M^{(!)}] \uplus P = [N^{(!)}] \uplus Q:\alpha}$$

with the usual reflexive, symmetric, transitivity, (app), (abs), (unit), $(pair)$, (Fst) and (Snd) rules. We say a theory $\mathscr{T}$ is a *resource λ-theory* if it is closed under all the rules in $\lambda\beta\eta_{\mathsf{R}}$.

Remark. In the (β_{R}) rule, if we only have linear resources in the bag, we would have

$$(\beta_{\mathsf{R}})\,\frac{\Gamma \cup \{x:\alpha\} \vdash_{\mathcal{R}} M:\beta \qquad \Gamma \vdash_{\mathcal{R}} [\vec{L}]:\alpha}{\Gamma \vdash_{\mathcal{R}} (\lambda x.M)[\vec{L}] = M\langle \vec{L}/x\rangle[0/x]:\beta}$$

If there is no reusable resources in a bag, this (β_{R}) rule substitutes every free occurrences of x by 0 after the resource substitution.

Example 3.4. In resource λ-theories, $(\lambda x.x)\,[y]$ is equated to a variable

$$(\lambda x.x)\,[y] = x\,\langle y/x\rangle[0/x] \equiv y\,[0/x] \equiv y,$$

whereas in differential λ-theories, the differential application $\mathsf{D}(\lambda x.x)\cdot y$ is equated to an abstraction

$$\mathsf{D}(\lambda x.x)\cdot y = \lambda x.(\frac{\partial x}{\partial x}\cdot y) \equiv \lambda x.y.$$

4 Translation Between Differential λ-Calculus and Resource λ-Calculus

As mentioned in the previous section, resource λ-calculus suggested in [PT09] is inspired by the dynamic of differential λ-calculus. It is not surprising that these two calculi are closely related.

Along with introducing the categorical models for simply-typed differential λ-calculus, Bucciarelli *et al.* in [BEM10] defined a "faithful" translation map from resource λ-calculus to differential λ-calculus. In a subsequent paper [Man12], Manzonetto refined the translation map and provided the backward direction.

4.1 From Resource λ-Calculus to Differential λ-Calculus

Resource terms can be easily translated to differential λ-terms.

$$
\begin{aligned}
(-)^d : \qquad \Lambda^r &\longrightarrow \Lambda^d \\
x &\longmapsto x \\
\lambda x.M &\longmapsto \lambda x.M^d \\
M[\vec{L}, \vec{N}^!] &\longmapsto (\mathsf{D}^k M^d \cdot \vec{L}^d) \sum_i N_i^d \\
\langle\rangle &\longmapsto \langle\rangle \\
\langle M, N\rangle &\longmapsto \langle M^d, N^d\rangle \\
\mathsf{Fst}(M) &\longmapsto \mathsf{Fst}(M^d) \\
\mathsf{Snd}(M) &\longmapsto \mathsf{Snd}(M^d)
\end{aligned}
$$

Remark. This translation can be extended to the sums of term by $\left(\sum_i M_i\right)^d = \sum_i M_i^d$.

Note that $(-)^d$ is a partial function. Consider the equivalent resource terms $x\,[y, z]$ and $x\,[z, y]$.

$$
(x\,[y, z])^d \equiv \mathsf{D}(\mathsf{D}x \cdot y) \cdot z \qquad\qquad (x\,[z, y])^d \equiv \mathsf{D}(\mathsf{D}x \cdot z) \cdot y.
$$

In differential λ-calculus, $\mathsf{D}(\mathsf{D}x \cdot y) \cdot z$ and $\mathsf{D}(\mathsf{D}x \cdot z) \cdot y$ are *not* syntactically equivalent. However, according to the (sw_{D}) rule, they are equal under any differential λ-theory.

Example 4.1. The resource term in example 3.3 can be translated to the differential term we considered in example 2.2 via $(-)^d$.

$$
\begin{aligned}
\left(\lambda y.(x\,[y, (x+y)^!])\right)^d &\equiv \lambda y.(x\,[y, (x+y)^!])^d \\
&\equiv \lambda y.\left((\mathsf{D}x^d \cdot y^d)\ (x+y)^d\right) \\
&\equiv \lambda y.\left((\mathsf{D}x \cdot y)\ (x+y)\right)
\end{aligned}
$$

The following lemma and proposition show that $(-)^d$ behaves well with the substitutions, the typing systems and the theories. The full proofs can be found in Appendix B.

Lemma 4.1. Let $M, N \in \Lambda^r$ and x be a variable.

(i) $(M[N/x])^d \equiv M^d[N^d/x]$

(ii) $(M\langle N/x\rangle)^d \equiv \dfrac{\partial M^d}{\partial x} \cdot N^d$

Proposition 4.2. For any $M, N \in \Lambda^r$, we have

(i) $\Gamma \vdash_{\mathcal{R}} M : \alpha \quad \iff \quad \Gamma \vdash_{\mathcal{D}} M^d : \alpha$,

(ii) If M and N are provably equal in the theory $\lambda\beta\eta_{\mathsf{R}}$, then their translations are also provably equal in $\lambda\beta\eta_{\mathsf{D}}$. i.e.

$$\lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} M = N : \alpha \qquad \Longrightarrow \qquad \lambda\beta\eta_{\mathsf{D}} \rhd \Gamma \vdash_{\mathcal{D}} M^d = N^d : \alpha.$$

4.2 The Converse - from Differential λ-Calculus to Resource λ-Calculus

This direction is more tricky. As discussed in example 3.4, in resource λ-theories, if there is no reusable resources in the bag, the (β_{R}) rule substitutes all free occurrences of x by 0 after the resource substitution. Whereas in differential λ-theories, the abstraction is kept.

Consider the following translation map from differential λ-terms to resource λ-terms.

$$\begin{aligned}
(-)^r : \qquad \Lambda^d &\longrightarrow \Lambda^r \\
x &\longmapsto x \\
\lambda x.s &\longmapsto \lambda x.s^r \\
sT &\longmapsto s^r\ [(T^r)^!] \\
\mathsf{D}s \cdot t &\longmapsto \lambda y.\big(s^r\ [t^r, y^!]\big) \qquad \text{where } y \text{ is a fresh variable} \\
\langle\rangle &\longmapsto \langle\rangle \\
\langle s, t\rangle &\longmapsto \langle s^r, t^r\rangle \\
\mathsf{Fst}(s) &\longmapsto \mathsf{Fst}(s^r) \\
\mathsf{Snd}(s) &\longmapsto \mathsf{Snd}(s^r) \\
s + T &\longmapsto s^r + T^r
\end{aligned}$$

Example 4.2. In contrast to example 4.1, we show that the differential term we consider in example 2.2 can *not* be translated to the resource term in example 3.3 via $(-)^r$.

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

Similar to $(-)^d$, the translation $(-)^r$ behaves well with the substitutions, typing systems and the theories. The proofs of the following lemma and proposition can be found in Appendix B.

Lemma 4.3. Let $S, T \in \Lambda^d$ and x a variable.

(i) $\lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} \big(S[T/x]\big)^r = S^r[T^r/x] : \alpha$

(ii) $\lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} \left(\frac{\partial S}{\partial x} \cdot T\right)^r = S^r\langle T^r/x\rangle : \alpha$

Proposition 4.4. For any $s, t \in \Lambda^d$, we have

(i) $\Gamma \vdash_{\mathcal{D}} s : \alpha \quad \Longleftrightarrow \quad \Gamma \vdash_{\mathcal{R}} s^r : \alpha$,

(ii) If s and t are provably equal in the theory $\lambda\beta\eta_{\mathsf{D}}$, then their translations are also provably equal in $\lambda\beta\eta_{\mathsf{R}}$. i.e.

$$\lambda\beta\eta_{\mathsf{D}} \rhd \Gamma \vdash_{\mathcal{D}} s = t : \alpha \qquad \Longrightarrow \qquad \lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} s^r = t^r : \alpha.$$

4.3 Resource λ-Calculus and Differential λ-Calculus

As shown in example 4.2, the translations $(-)^d$ and $(-)^r$ are not inverse of each other, yet we will see that they are close enough.

Proposition 4.5. For any differential typed term $S \in \Lambda^d$ and resource typed term $\mathbb{M} \in \mathcal{N}(\Lambda^r)$,

$$\lambda\beta\eta_{\mathsf{D}} \rhd \Gamma \vdash_{\mathcal{D}} (S^r)^d = S : \alpha \qquad \text{and} \qquad \lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} (\mathbb{M}^d)^r = \mathbb{M} : \alpha.$$

Proof. We first prove that $\lambda\beta\eta_{\mathsf{D}} \rhd \Gamma \vdash_{\mathcal{D}} (S^r)^d = S : \alpha$ for all differential typed terms $S \in \Lambda^d$ by induction on the structure of S. The only interesting cases are when

- $S \equiv s\,U$ is an application,

$$\lambda\beta\eta_{\mathsf{D}} \rhd \big((s\,U)^r\big)^d \equiv (s^r\,[(U^r)^!])^d \equiv (s^r)^d\,(U^r)^d = s\,U : \alpha.$$

- $S \equiv \mathsf{D}s \cdot t$ is a differential application,

$$\begin{aligned}
\lambda\beta\eta_{\mathsf{D}} \rhd \Gamma \vdash_{\mathcal{D}} \big((\mathsf{D}s \cdot t)^r\big)^d &\equiv \big(\lambda y.(s^r\,[t^r, y^!])\big)^d \\
&\equiv \lambda y.\big((\mathsf{D}(s^r)^d \cdot (t^r)^d)\,y^d\big) \\
&\equiv \lambda y.\big((\mathsf{D}s \cdot t)\,y\big) && \text{(IH)} \\
&= \mathsf{D}s \cdot t : \alpha && (\eta)
\end{aligned}$$

Now, we consider sums of resource typed terms $\mathbb{M} \in \mathcal{N}(\Lambda^r)$. We prove that $\Gamma \vdash_{\mathcal{R}} \big((\mathbb{M})^d\big)^r = \mathbb{M}$ by induction on the structure of $\mathbb{M}$. The only interesting case is when $\mathbb{M} \equiv M\,[\vec{L}, \vec{N}^!]$ is an application.

$$\begin{aligned}
\lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} \big((M\,[\vec{L}, \vec{N}^!])^d\big)^r &\equiv \big((\mathsf{D}^k M^d \cdot \vec{L}^d) \sum N_i^d\big)^r \\
&\equiv (\mathsf{D}^k M^d \cdot \vec{L}^d)^r\,[\big(\sum (N_i^d)^r\big)^!] \\
&\equiv \big(\lambda y.((M^d)^r\,[(\vec{L}^d)^r, y^!])\big)\,[\big((\vec{N}^d)^r\big)^!] && (3.1) \\
&= \big(\lambda y.(M\,[\vec{L}, y^!])\big)\,[\vec{N}^!] && \text{(IH)} \\
&= M\,[\vec{L}, \vec{N}^!] : \alpha && (\eta_{\mathsf{R}})
\end{aligned}$$

□

This relationship between differential and resource λ-theories confirm the initial idea that differential λ-calculus is resource-sensitive.

5 Model of Differential λ-Calculus

In this section, we describe the *cartesian closed differential category* presented by Bucciarelli *et al.* in [BEM10]. It is based on the *cartesian differential category* introduced by Blute *et al.* in [BCS09]. Next, we prove that one can soundly model differential λ-theories in cartesian closed differential category as shown in [BEM10]. After that, we prove the converse by giving a construction of the "smallest" cartesian closed differential category that can soundly model any given differential λ-theory.

Furthermore, assuming that we can extend the theory with constants and function symbols, we show that differential λ-theory is the *internal language* of cartesian closed differential category and "prove" properties about the category using the theory. Finally, we discuss the notion of *complete* categories with respect to differential λ-theories.

5.1 Cartesian Differential Category

In differential category, morphisms are viewed as linear functions and coKleisli morphisms as differentiable functions. *Cartesian differential category* resembles the coKleisli category of a differential category and directly axiomatizes the differentiable functions.

Definition 5.1 (Left additive and Additive)**.** A category $\mathcal{C}$ is *left additive* if every homset $\mathcal{C}(A, B)$ is enriched with a commutative monoid $(\mathcal{C}(A, B), +_{AB}, 0_{AB})$ and, the additive structure is preserved by composition on the left. i.e.

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

We say a morphism f in a left additive category is *additive* if it also preserves the additive structure of the homset on the right. i.e.

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

Example 5.1. Consider the category FVect of finite dimensional vector spaces and differentiable functions. For any differentiable functions $g, h : U \to V$, we can define a differentiable function $g + h : U \to V$, which sends $u \mapsto g(u) + h(u)$, where for any differentiable function $f : W \to U$,

$$\forall w \in W \qquad ((g + h) \circ f)(w) = (g + h)(f(w)) = g(f(w)) + h(f(w)) = (g \circ f + h \circ f)(w)$$

Moreover, the zero function $0_{UV} : u \mapsto 0$ is preserved by composition on the left. Thus, FVect is left additive. However, not every differentiable functions in FVect is additive. For instance, $(x^2 \circ (x + 2x))(1) = x^2(1 + 2) = 3^2 = 9$, but $(x^2 \circ x + x^2 \circ (2x))(1) = x^2(1) + x^2(2) = 1^2 + 2^2 = 5$.

Definition 5.2 (Cartesian left additive category)**.** A *cartesian left additive category* is a left additive category with products where projections and pairings of additive maps are additive. i.e.

- π_1 and π_2 are additive,
- if f and g are additive, their pairing $\langle f, g\rangle$ is also additive.

Example 5.2. In FVect, it is easy to check that the projections $p_1 : (u, v) \mapsto u$ and $p_2 : (u, v) \mapsto v$ are additive. Also, given additive differentiable functions f and g, their pairing $\langle f, g\rangle : v \mapsto (f(v), g(v))$ is also additive. Thus, FVect is a cartesian left additive category.

As discussed before, the derivative of a function $f : U \to V$ along a point $u \in U$ is the function $Df \cdot u : U \to V$ which maps $x \mapsto f'(x) \cdot u$. Consider the situation where we abstract $u \in U$ as well. Then, we obtain a function from U to the space of linear functions $U \Rightarrow V$, which we denote as

$$\begin{array}{rccc} Df : & U & \longrightarrow & (U \Rightarrow V) \\ & u & \longmapsto & Df \cdot u \end{array}$$

Note that Df is also differentiable. Now, we consider such a map in a cartesian left additive category. Given a morphism $f : A \to B$, we define a cartesian differential operator $D_\times[f]$ which has the type $A \to (A \Rightarrow B)$. Since the cartesian left additive category is not necessarily closed, we "uncurry" the morphism and obtain $D_\times[f] : A \times A \to B$.

Definition 5.3 (Cartesian differential operator)**.** An operator $D_\times[-] : \mathcal{C}(A, B) \to \mathcal{C}(A \times A, B)$ of a cartesian left additive category is a *cartesian differential operator* if it satisfies the following axioms:

[CD1] $D_\times$ is linear:

$$D_\times[f + g] = D_\times[f] + D_\times[g], \qquad D_\times[0] = 0$$

[CD2] $D_\times$ is additive in its first coordinate:

$$D_\times[f] \circ \langle h + k, v\rangle = D_\times[f] \circ \langle h, v\rangle + D_\times[f] \circ \langle k, v\rangle, \qquad D_\times[f] \circ \langle 0, v\rangle = 0$$

[CD3] $D_\times$ behaves with projections:

$$D_\times[\mathsf{Id}] = \pi_1, \qquad D_\times[\pi_1] = \pi_1 \circ \pi_1, \qquad D_\times[\pi_2] = \pi_2 \circ \pi_1$$

[CD4] $D_\times$ behaves with pairings:

$$D_\times[\langle f, g\rangle] = \langle D_\times[f], D_\times[g]\rangle$$

[CD5] Chain rule:

$$D_\times[g \circ f] = D_\times[g] \circ \langle D_\times[f], f \circ \pi_2\rangle$$

[CD6] $D_\times[f]$ is linear in its first component:

$$D_\times[D_\times[f]] \circ \langle\langle g, 0\rangle, \langle h, k\rangle\rangle = D_\times[f] \circ \langle g, k\rangle$$

[CD7] Independence of order of partial differentiation:

$$D_\times[D_\times[f]] \circ \langle\langle 0, h\rangle, \langle g, k\rangle\rangle = D_\times[D_\times[f]] \circ \langle\langle 0, g\rangle, \langle h, k\rangle\rangle$$

We say $D_\times[f] : A \times A \to B$ is the *derivative* of $f : A \to B$.

Definition 5.4 (Cartesian differential category)**.** A cartesian left additive category with a cartesian differential operator is a *cartesian differential category*.

Example 5.3. We can define a cartesian differential operator in FVect as follows,

$$D_\times[\vec{f}](\vec{x},\vec{y}) = \left[\frac{\partial f_j}{\partial x_i}\bigg|_{\vec{y}} \right]_{i=1,\,j=1}^{m,\,n} \vec{x}$$

where $\vec{f} = (f_1, \ldots, f_n)$, $\vec{x} = (x_1, \ldots, x_m)$ and $\vec{y} = (y_1, \ldots, y_m)$. It is easy to check that $D_\times[-]$ satisfies [CD1-7]. Thus, FVect is a cartesian differential category.

Remark. In cartesian differential category, the partial derivative of a morphism can be obtained by zeroing out the components on which the differentiation is not required. Consider a morphism $f : A_1 \times A_2 \to B$. To obtain the partial derivative of f on the first component, we zero out the component A_2. i.e.

$$A_1 \times (A_1 \times A_2) \xrightarrow{\langle \mathsf{Id}_{A_1}, 0_{A_2}\rangle \times \mathsf{Id}_{A_1\times A_2}} (A_1 \times A_2) \times (A_1 \times A_2) \xrightarrow{D_\times[f]} B$$

Therefore, we define $D^1_\times[f] := D_\times[f] \circ \big(\langle \mathsf{Id}_{A_1}, 0_{A_2}\rangle \times \mathsf{Id}_{A_1\times A_2}\big)$ to be the derivative of f on its first component. Similarly, we define $D^2_\times[f] := D_\times[f] \circ \big(\langle 0_{A_1}, \mathsf{Id}_{A_2}\rangle \times \mathsf{Id}_{A_1\times A_2}\big)$ to be the derivative of f on its second component.

5.2 Cartesian Closed Differential Category

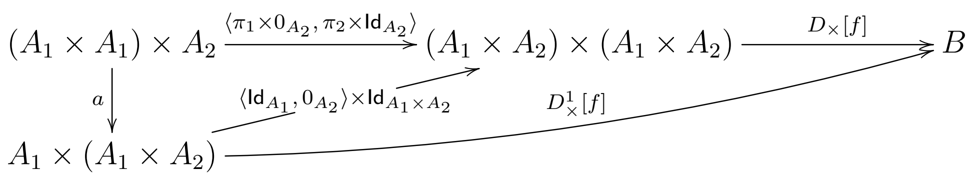
Cartesian differential category is not enough to provide categorical semantics for differential λ-calculus as the cartesian differential operator does not necessarily behave well with exponentials. Bucciarelli *et al.* added a new rule to the cartesian differential operator and proved that a cartesian differential category with such an operator can soundly model differential λ-calculus.

Definition 5.5 (Cartesian closed differential category)**.** A cartesian differential category is a *cartesian closed differential category* if

- it is cartesian closed,
- $\lambda(-)$ preserves the additive structure, i.e. $\lambda(f+g) = \lambda(f) + \lambda(g)$ and $\lambda(0) = 0$,
- $D_\times[-]$ satisfies the following (D-curry) rule. For any $f : A_1 \times A_2 \to B$,

$$D_\times[\lambda(f)] = \lambda(D_\times[f] \circ \langle \pi_1 \times 0_{A_2}, \pi_2 \times \mathsf{Id}_{A_2}\rangle)$$

It is easy to check that the following diagram commutes.



Hence, we can rewrite the rule (D-curry) as follows,

$$D_\times[\lambda(f)] = \lambda(D^1_\times[f] \circ a)$$

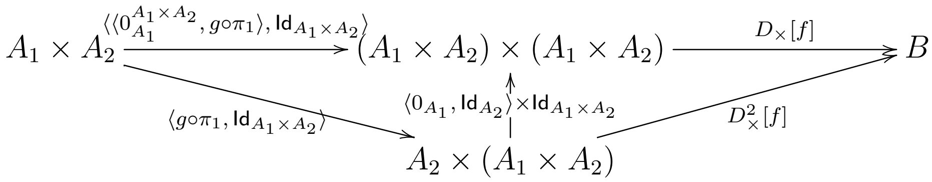
where $a : (A_1 \times A_1) \times A_2 \to A_1 \times (A_1 \times A_2)$ is the associative morphism. Intuitively, (D-curry) says that if we curry f and take its derivative, it is the same as currying the derivative of f in its first component.

Recall that in defining the categorical model for simply-typed λ-calculus, capture-free substitution is modelled by composition of morphisms. Since differential substitution is introduced for differential λ-calculus, we need another binary operator that models it.

Definition 5.6. Let $\star : \mathcal{C}(A_1 \times A_2, B) \times \mathcal{C}(A_1, A_2) \to \mathcal{C}(A_1 \times A_2, B)$ be a binary operator where

$$f \star g := D_\times[f] \circ \langle\langle 0_{A_1}^{A_1 \times A_2}, g \circ \pi_1 \rangle, \mathsf{Id}_{A_1 \times A_2} \rangle$$

The binary operator $\star$ can be viewed as the counterpart of differential substitution. It is easy to check that the following diagram commutes.



So, we can write

$$f \star g = D^2_\times[f] \circ \langle g \circ \pi_1, \mathsf{Id}_{A_1 \times A_2} \rangle.$$

Intuitively, $f \star g$ is obtained by force-feeding the second argument of f with exactly one copy of the result of g.

The following proposition tells us that the order of force-feeding f with results of morphisms does not matter. The proof can be found in Appendix B.

Proposition 5.1. Let $f : A_1 \times A_2 \to B$ and $g, h : A_1 \to A_2$ be morphisms in a cartesian closed differential category. We have

$$(f \star g) \star h = (f \star h) \star g.$$

5.3 Categorical Semantics of Differential Typed Terms

Definition 5.7 (Categorical Semantics of Differential Typed Terms)**.** Let $\mathcal{C}$ be a cartesian closed differential category. A *structrue* $\mathbb{M}$ in $\mathcal{C}$ is specified by giving each type variable $\gamma \in \mathsf{TV}$ an object $[\![\gamma]\!]_{\mathbb{M}}$ of $\mathcal{C}$. The *interpretation* of types and typed terms with respect to the structure $\mathbb{M}$ is defined by induction as follows, where $\top$ is the terminal object of $\mathcal{C}$ and τ_A is the unique morphism from A to $\top$,

- Types
  $[\![\mathsf{unit}]\!] := \top$
  $[\![\gamma]\!] := [\![\gamma]\!]_{\mathbb{M}}$, where $\gamma \in \mathsf{TV}$
  $[\![\alpha \times \beta]\!] := [\![\alpha]\!] \times [\![\beta]\!]$
  $[\![\alpha \Rightarrow \beta]\!] := [\![\alpha]\!] \Rightarrow [\![\beta]\!]$

- Type Contexts
  $[\![\varnothing]\!] := \top$
  $[\![\Gamma \cup \{x : \alpha\}]\!] := [\![\Gamma]\!] \times [\![\alpha]\!]$

- Typed Terms

$$(var)\ \frac{}{[\![\Gamma \cup \{x : \alpha\} \vdash x : \alpha]\!] := \pi_2 : [\![\Gamma]\!] \times [\![\alpha]\!] \to [\![\alpha]\!]}$$

$$(abs)\ \frac{[\![\Gamma \cup \{x : \alpha\} \vdash s : \beta]\!] = f : [\![\Gamma]\!] \times [\![\alpha]\!] \to [\![\beta]\!]}{[\![\Gamma \vdash \lambda x.s : \alpha \Rightarrow \beta]\!] := \lambda(f) : [\![\Gamma]\!] \to ([\![\alpha]\!] \Rightarrow [\![\beta]\!])}$$

$$(app)\ \frac{[\![\Gamma \vdash s : \alpha \Rightarrow \beta]\!] = S : [\![\Gamma]\!] \to ([\![\alpha]\!] \Rightarrow [\![\beta]\!]) \qquad [\![\Gamma \vdash t : \alpha]\!] = T : [\![\Gamma]\!] \to [\![\alpha]\!]}{[\![\Gamma \vdash s\, t : \beta]\!] := \mathsf{ev} \circ \langle S, T \rangle : [\![\Gamma]\!] \to ([\![\alpha]\!] \Rightarrow [\![\beta]\!]) \times [\![\alpha]\!] \to [\![\beta]\!]}$$

$$(\mathsf{unit})\ \frac{}{[\![\Gamma \vdash \langle\rangle : \mathsf{unit}]\!] := \tau_{[\![\Gamma]\!]} : [\![\Gamma]\!] \to \top}$$

$$(pair)\ \frac{[\![\Gamma \vdash s : \alpha]\!] = S \qquad [\![\Gamma \vdash t : \beta]\!] = T}{[\![\Gamma \vdash \langle s, t \rangle : \alpha \times \beta]\!] := \langle S, T \rangle : [\![\Gamma]\!] \to [\![\alpha]\!] \times [\![\beta]\!]}$$

$$(\mathsf{Fst})\ \frac{[\![\Gamma \vdash p : \alpha \times \beta]\!] = P : [\![\Gamma]\!] \to [\![\alpha]\!] \times [\![\beta]\!]}{[\![\Gamma \vdash \mathsf{Fst}(p) : \alpha]\!] := \pi_1 \circ P : [\![\Gamma]\!] \to [\![\alpha]\!]}$$

$$(\mathsf{Snd})\ \frac{[\![\Gamma \vdash p : \alpha \times \beta]\!] = P : [\![\Gamma]\!] \to [\![\alpha]\!] \times [\![\beta]\!]}{[\![\Gamma \vdash \mathsf{Snd}(p) : \beta]\!] := \pi_2 \circ P : [\![\Gamma]\!] \to [\![\beta]\!]}$$

$$(\mathsf{D})\ \frac{[\![\Gamma \vdash_{\mathcal{D}} s : \alpha \Rightarrow \beta]\!] = S : [\![\Gamma]\!] \to ([\![\alpha]\!] \Rightarrow [\![\beta]\!]) \qquad [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!] = T : [\![\Gamma]\!] \to [\![\alpha]\!]}{[\![\Gamma \vdash_{\mathcal{D}} \mathsf{D}s \cdot t : \alpha \Rightarrow \beta]\!] := \lambda\big(\lambda^-(S) \star T\big) : [\![\Gamma]\!] \to ([\![\alpha]\!] \Rightarrow [\![\beta]\!])}$$

$$(sum)\ \frac{[\![\Gamma \vdash_{\mathcal{D}} s_i : \alpha]\!] = S_i : [\![\Gamma]\!] \to [\![\alpha]\!] \qquad \forall i \in I}{[\![\Gamma \vdash_{\mathcal{D}} \sum_{i \in I} s_i : \alpha]\!] := \sum_{i \in I} S_i : [\![\Gamma]\!] \to [\![\alpha]\!]}$$

Remark. These interpretations should come as no surprise. The interpretation of the (D) rule is basically uncurrying S, force-feed one copy of the result of T to the uncurried S, and then currying the result.

Recall 5.2 (Model)**.** We say a structure $\mathbb{M}$ *satisfies* a rule $\Gamma \vdash s = t : \alpha$ if the interpretations of $\Gamma \vdash s : \alpha$ and $\Gamma \vdash t : \alpha$ with respect to $\mathbb{M}$ are the same. i.e.

$$[\![\Gamma \vdash s : \alpha]\!] = [\![\Gamma \vdash t : \alpha]\!].$$

A structure $\mathbb{M}$ is a *model* of a theory $\mathscr{T}$ if $\mathbb{M}$ satisfies all the rules in $\mathscr{T}$. i.e.

$$\mathscr{T} \rhd \Gamma \vdash s = t : \alpha \qquad \Longrightarrow \qquad [\![\Gamma \vdash s : \alpha]\!] = [\![\Gamma \vdash t : \alpha]\!].$$

5.4 Soundness and Completeness Theorem

The following substitution lemma shows that capture-free/differential substitution can be modelled soundly by composition/the binary operator $\star$. Note that the (D-curry) rule is required in proving the following lemma which is crucial in proving the (β_{D}) case in the soundness theorem.

Lemma 5.3. (Generalized Substitution Lemma) Let $\mathbb{M}$ be a structure in a cartesian closed differential category. If $\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} s : \beta$ and $\Gamma \vdash_{\mathcal{D}} t : \alpha$, then

(i) $[\![\Gamma \vdash_{\mathcal{D}} s[t/x] : \beta]\!] = [\![\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} s : \beta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!]\big\rangle$;

(ii) $[\![\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} \frac{\partial s}{\partial x} \cdot t : \beta]\!] = [\![\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} s : \beta]\!] \star [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!]$.

Remark. The proof of this generalized substitution lemma is very technical and can be found in the appendix of [Man12].

Theorem 5.4. (Soundness Theorem) Given a cartesian closed differential category $\mathcal{C}$, any structure $\mathbb{M}$ in $\mathcal{C}$ is a model of $\lambda\beta\eta_{\mathsf{D}}$.

Proof. We show that a structure $\mathbb{M}$ satisfies all rules in $\lambda\beta\eta_{\mathsf{D}}$ by induction. The cases for the rules in $\lambda\beta\eta$ can be found in th proof of Theorem A.5.

(β_{D}) Let $S = [\![\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} s : \beta]\!]$ and $T = [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!]$. We have

$$\begin{aligned}
[\![\Gamma \vdash_{\mathcal{D}} \mathsf{D}(\lambda x.s) \cdot t : \alpha \Rightarrow \beta]\!] &= \lambda\Big(\lambda^{-}\big(\lambda(S)\big) \star T\Big) \\
&= \lambda(S \star T) \\
&= \lambda\Big([\![\Gamma \cup \{x : \alpha\} \vdash_{\mathcal{D}} \frac{\partial s}{\partial x} \cdot t : \beta]\!]\Big) \qquad \text{(Lemma 5.3 (ii))} \\
&= [\![\Gamma \vdash_{\mathcal{D}} \lambda x.\big(\frac{\partial s}{\partial x} \cdot t\big) : \alpha \Rightarrow \beta]\!]
\end{aligned}$$

Thus $\mathbb{M}$ satisfies $\Gamma \vdash_{\mathcal{D}} \mathsf{D}(\lambda x.s) \cdot t = \lambda x.\big(\frac{\partial s}{\partial x} \cdot t\big) : \alpha \Rightarrow \beta$.

(sw_{D}) Let $S = [\![\Gamma \vdash_{\mathcal{D}} s : \alpha \Rightarrow \beta]\!]$, $T = [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!]$ and $U = [\![\Gamma \vdash_{\mathcal{D}} u : \alpha]\!]$.

$$\begin{aligned}
[\![\Gamma \vdash_{\mathcal{D}} \mathsf{D}(\mathsf{D}s \cdot t) \cdot u : \alpha \Rightarrow \beta]\!] &= \lambda\big(\lambda^{-}\big(\lambda(\lambda^{-}(S) \star T)\big) \star U\big) \\
&= \lambda\big((\lambda^{-}(S) \star T) \star U\big) \\
&= \lambda\big((\lambda^{-}(S) \star U) \star T\big) \qquad \text{(Prop. 5.1)} \\
&= [\![\Gamma \vdash_{\mathcal{D}} \mathsf{D}(\mathsf{D}s \cdot u) \cdot t : \alpha \Rightarrow \beta]\!]
\end{aligned}$$

Thus $\mathbb{M}$ satisfies $\Gamma \vdash_{\mathcal{D}} \mathsf{D}(\mathsf{D}s \cdot t) \cdot u = \mathsf{D}(\mathsf{D}s \cdot u) \cdot t : \alpha \Rightarrow \beta$.

The induction case for the (app_{D}) and (sum) rules are trivial. □

Soundness tells us that the categorical semantics of differential typed terms actually makes sense. Thus, we can now define a differential λ-theory based on any cartesian closed differential category.

Corollary 5.5. Every cartesian closed differential category $\mathcal{C}$ gives rise to a differential λ-theory $\mathsf{Th}(\mathcal{C})$.

Proof. Let $\mathbb{M}$ be a structure in $\mathcal{C}$. Let

$$\mathsf{Th}(\mathcal{C}) := \{\ \Gamma \vdash_{\mathcal{D}} s = t : \alpha \mid [\![\Gamma \vdash_{\mathcal{D}} s : \alpha]\!] = [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!]\ \}.$$

By soundness, $\lambda\beta\eta_{\mathsf{D}}$ is included in $\mathsf{Th}(\mathcal{C})$. So, $\mathsf{Th}(\mathcal{C})$ is indeed a differential λ-theory. $\square$

We have proved that for any cartesian closed differential category, one can define a differential λ-theory with respect to it. Now, we consider the converse. Given a differential λ-theory, we would like to construct a cartesian closed differential category in which the theory can be modelled soundly in the category. More specifically, we would like to construct the "smallest" such category. We call it the *classifying category* of the theory. The classifying category is the "smallest" in the sense that for any other categories $\mathcal{D}$, in which the theory can be modelled soundly, there is a cartesian closed differential functor F from the classifying category to $\mathcal{D}$ such that the interpretation of the theory in $\mathcal{D}$ can be expressed as the composition of the interpretation of the theory in the classifying category and F.

To have a notion of "smallest", we consider the category $\mathsf{CCDCat}_{\cong}(\mathcal{C}, \mathcal{D})$ of *cartesian closed differential functors* $\mathcal{C} \to \mathcal{D}$ and natural isomorphisms. We say a functor $F : \mathcal{C} \to \mathcal{D}$ between cartesian closed differential categories is a cartesian closed differential functor if F preserves

- the additive structure, i.e. $F(f + g) = F(f) + F(g)$ and $F(0) = 0$,
- products via the isomorphism $\Phi := \langle F\pi_1, F\pi_2\rangle$,
- exponentials via the isomorphism $\Psi := \lambda\big(F(\mathsf{ev}) \circ \Phi\big)$,
- the cartesian differential operator, i.e. $F\big(D_\times[f]\big) = D_\times\big[F(f)\big] \circ \Phi$.

We also consider the category $\mathsf{DMod}_{\cong}(\mathscr{T}, \mathcal{C})$ of models of a differential λ-theory $\mathscr{T}$ in a cartesian closed differential category $\mathcal{C}$ and *additive* model homomorphisms. A model homomorphism $h : \mathbb{M} \to \mathbb{N}$ is given by isomorphisms $h_\gamma : [\![\gamma]\!]_{\mathbb{M}} \to [\![\gamma]\!]_{\mathbb{N}}$ for each type variable γ, and

$$h_{\alpha\times\beta} := h_\alpha \times h_\beta \qquad \text{and} \qquad h_{\alpha\Rightarrow\beta} := h_\alpha^{-1} \Rightarrow h_\beta := \lambda\big(h_\beta \circ \mathsf{ev} \circ (\mathsf{Id} \times h_\alpha^{-1})\big).$$

Definition 5.8 (Classifying Category)**.** Given a differential λ-theory $\mathscr{T}$, we say a category, denoted $\mathsf{Cl}(\mathscr{T})$, is *classifying* if there is a "generic" model that soundly interprets $\mathscr{T}$ in $\mathsf{Cl}(\mathscr{T})$ and for every cartesian closed differential category $\mathcal{D}$, there is a natural equivalence

$$\mathsf{CCDCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D}) \simeq \mathsf{DMod}_{\cong}(\mathscr{T}, \mathcal{D}).$$

So, for any cartesian closed differential category $\mathcal{D}$, a model of $\mathscr{T}$ in $\mathcal{D}$ can be represented by a functor from the classifying category $\mathsf{Cl}(\mathscr{T})$ to $\mathcal{D}$. Thus, the classifying category is the "smallest". It is not difficult to see that for any differential λ-theory, its classifying category is unique up to isomorphism.

We set up the equivalence in the forward direction via the following family of modelling functors.

Definition 5.9 (Modelling functors)**.** Let $\mathcal{C}$ and $\mathcal{D}$ be cartesian closed differential categories, $\mathscr{T}$ be a differential λ-theory in $\mathcal{C}$ and $\mathbb{M}$ be a model of $\mathscr{T}$ in $\mathcal{C}$. We define a *family* *of modelling functors* $\mathsf{DAp}_{\mathbb{M}} : \mathsf{CCDCat}_{\cong}(\mathcal{C}, \mathcal{D}) \to \mathsf{DMod}_{\cong}(\mathscr{T}, \mathcal{D})$ which for any cartesian closed differential functor $F : \mathcal{C} \to \mathcal{D}$, $\mathsf{DAp}_{\mathbb{M}} F$ is a model of $\mathscr{T}$ in $\mathcal{D}$, where

$$[\![\gamma]\!]_{\mathsf{DAp}_{\mathbb{M}} F} := F([\![\gamma]\!]_{\mathbb{M}}),$$

and for any natural isomorphism $\phi : F \to G$, $\mathsf{DAp}_{\mathbb{M}} \phi : \mathsf{DAp}_{\mathbb{M}} F \to \mathsf{DAp}_{\mathbb{M}} G$ is an additive model homomorphism where

$$\big(\mathsf{DAp}_{\mathbb{M}} \phi\big)_{\gamma} := \phi_{[\![\gamma]\!]_{\mathbb{M}}}.$$

Remark. It is easy to check that $\mathsf{DAp}_{\mathbb{M}} F$ is indeed a model of $\mathscr{T}$ in $\mathcal{D}$, $\mathsf{DAp}_{\mathbb{M}} \phi$ is indeed an additive model homomorphism and $\mathsf{DAp}_{\mathbb{M}}$ is a well-defined functor from $\mathsf{CCDCat}_{\cong}(\mathcal{C}, \mathcal{D})$ to $\mathsf{DMod}_{\cong}(\mathscr{T}, \mathcal{D})$.

The following technical lemma will be used in proving completeness. Complete proofs can be found in Appendix B.

Lemma 5.6. (i) If $x \notin \mathsf{FV}(t)$, then

$$(\mathsf{D}(\lambda x.(t[t'/x'])) \cdot u)\ s = (\mathsf{D}(\lambda x'.t) \cdot ((\mathsf{D}(\lambda x.t') \cdot u)\ s))\ t'[s/x].$$

(ii) if $x \not\equiv y$, $x, y \notin \mathsf{FV}(M)$ and $x, y, z \notin \mathsf{FV}(u) \cup \mathsf{FV}(v)$,

$$\left(\frac{\partial}{\partial x}\big(M[\langle x, y\rangle/z]\big) \cdot u\right)\ [v/x] = \left(\frac{\partial M}{\partial z} \cdot \langle u, 0\rangle\right)\ [\langle v, y\rangle/z].$$

Theorem 5.7 (Completeness Theorem)**.** Every differential λ-theory $\mathscr{T}$ has a classifying cartesian closed differential category $\mathsf{Cl}(\mathscr{T})$.

Proof. We first construct a cartesian closed differential category based on the syntax of the given differential λ-theory $\mathscr{T}$ and then prove that it is classifying by presenting the "inverse" for the functor $\mathsf{DAp}_{\mathbb{G}} : \mathsf{CCDCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D}) \to \mathsf{DMod}_{\cong}(\mathscr{T}, \mathcal{D})$.

Given a differential λ-theory $\mathscr{T}$, we define a cartesian closed differential category $\mathsf{Cl}(\mathscr{T})$ where

- objects are types of $\mathscr{T}$,
- morphisms $\boldsymbol{f} : \alpha \to \beta$ are equivalence classes of typed terms $\big[\{x : \alpha\} \vdash M : \beta\big]$, where two typed terms are equivalent if they are provably equal in $\mathscr{T}$. We write $\{x : \alpha\} \vdash M : \beta$ instead of $\big[\{x : \alpha\} \vdash M : \beta\big]$,
- composition of $\boldsymbol{g} = \{y : \beta\} \vdash N : \gamma$ and $\boldsymbol{f} = \{x : \alpha\} \vdash M : \beta$ is

$$\boldsymbol{f} \circ \boldsymbol{g} = \{x : \alpha\} \vdash N[M/y] : \gamma,$$

- identity morphism of the object α is $\mathbf{Id}_{\boldsymbol{\alpha}} := \{x : \alpha\} \vdash x : \alpha$,
- product of objects α and β is $\alpha \times \beta$ with projections

$$\boldsymbol{\pi_1} = \{z : \alpha \times \beta\} \vdash \mathsf{Fst}(z) : \alpha \qquad \text{and} \qquad \boldsymbol{\pi_2} = \{z : \alpha \times \beta\} \vdash \mathsf{Snd}(z) : \beta,$$

- pairing of morphisms $\boldsymbol{f} = \{x : \gamma\} \vdash M : \alpha$ and $\boldsymbol{g} = \{x : \gamma\} \vdash N : \beta$ is

$$\langle \boldsymbol{f}, \boldsymbol{g}\rangle = \{x : \gamma\} \vdash \langle M, N\rangle : \alpha \times \beta$$

- exponential of objects β and γ is $\beta \Rightarrow \gamma$ and the evaluating morphism is

$$\mathbf{ev} = \{z : (\beta \Rightarrow \gamma) \times \beta\} \vdash \mathsf{Fst}(z)\ \mathsf{Snd}(z) : \gamma,$$

where for any morphism $\boldsymbol{f} = \{z : \alpha \times \beta\} \vdash M : \gamma$, the exponential mate of $\boldsymbol{f}$ is

$$\boldsymbol{\lambda}(\boldsymbol{f}) = \{x : \alpha\} \vdash \lambda y.(M[\langle x, y\rangle/z]) : \beta \Rightarrow \gamma,$$

where x and y are distinct and fresh variables,

- given objects α and β, $\big(\mathsf{Cl}(\mathscr{T})(\alpha, \beta), +, \mathbf{0}\big)$ is a commutative monoid where the sum of the morphisms $\boldsymbol{f} = \{x : \alpha\} \vdash M : \beta$ and $\boldsymbol{g} = \{x : \alpha\} \vdash N : \beta$ is

$$\boldsymbol{f} + \boldsymbol{g} := \{x : \alpha\} \vdash M + N : \beta,$$

and the unit is $\mathbf{0} := \{x : \alpha\} \vdash 0 : \beta$,

- given morphism $\boldsymbol{f} = \{x : \alpha\} \vdash M : \beta$, its derivative is

$$\boldsymbol{D}_{\times}[\boldsymbol{f}] := \{y : \alpha \times \alpha\} \vdash \big(\mathsf{D}(\lambda x.M) \cdot \mathsf{Fst}(y)\big)\ \mathsf{Snd}(y) : \beta,$$

where y is a fresh variable.

The proof that $\mathsf{Cl}(\mathscr{T})$ is a cartesian closed category can be found in the proof of Theorem A.7. To verify $\mathsf{Cl}(\mathscr{T})$ is a cartesian closed differential category, it is enough to prove that it is cartesian left additive, and $\boldsymbol{D}_{\times}[-]$ satisfies [CD1-7] and (D-curry).

Verifying that $\mathsf{Cl}(\mathscr{T})$ is a cartesian left additive category is easy. We would only look at the requirement that the pairing of additive morphisms is additive, since this explains the syntactic sugar on sums of pairing terms in differential λ-calculus given in 2.1.

Let $\boldsymbol{f} = \{y : \gamma\} \vdash M : \alpha$ and $\boldsymbol{g} = \{y : \gamma\} \vdash N : \beta$ be additive and $\boldsymbol{h} = \{z : \delta\} \vdash H : \gamma$ and $\boldsymbol{k} = \{z : \delta\} \vdash K : \gamma$ be morphisms.

$$\begin{aligned}
\langle \boldsymbol{f}, \boldsymbol{g}\rangle \circ (\boldsymbol{h} + \boldsymbol{k}) &= (\{y : \gamma\} \vdash \langle M, N\rangle : \alpha \times \beta) \ \circ\ (\{z : \delta\} \vdash H + K : \gamma) \\
&= \{z : \delta\} \vdash \langle M, N\rangle[H + K/y] : \alpha \times \beta \\
&= \{z : \delta\} \vdash \langle M[H + K/y], N[H + K/y]\rangle : \alpha \times \beta \\
&= \langle \boldsymbol{f} \circ (\boldsymbol{h} + \boldsymbol{k}), \boldsymbol{g} \circ (\boldsymbol{h} + \boldsymbol{k})\rangle \\
&= \langle \boldsymbol{f} \circ \boldsymbol{h} + \boldsymbol{f} \circ \boldsymbol{k}, \boldsymbol{g} \circ \boldsymbol{h} + \boldsymbol{g} \circ \boldsymbol{k}\rangle && (\boldsymbol{f} \text{ and } \boldsymbol{g} \text{ are additive}) \\
&= \langle \boldsymbol{f} \circ \boldsymbol{h}, \boldsymbol{g} \circ \boldsymbol{h}\rangle + \langle \boldsymbol{f} \circ \boldsymbol{k} + \boldsymbol{g} \circ \boldsymbol{k}\rangle \\
&= \langle \boldsymbol{f}, \boldsymbol{g}\rangle \circ \boldsymbol{h} + \langle \boldsymbol{f}, \boldsymbol{g}\rangle \circ \boldsymbol{k}
\end{aligned}$$

$$\begin{aligned}
\langle \boldsymbol{f}, \boldsymbol{g}\rangle \circ \mathbf{0} &= (\{y : \gamma\} \vdash \langle M, N\rangle : \alpha \times \beta) \ \circ\ (\{z : \delta\} \vdash 0 : \gamma) \\
&= \{z : \delta\} \vdash \langle M, N\rangle[0/y] : \alpha \times \beta \\
&= \{z : \delta\} \vdash \langle M[0/y], N[0/y]\rangle : \alpha \times \beta \\
&= \langle \boldsymbol{f} \circ \mathbf{0}, \boldsymbol{g} \circ \mathbf{0}\rangle \\
&= \langle \mathbf{0}, \mathbf{0}\rangle && (\boldsymbol{f} \text{ and } \boldsymbol{g} \text{ are additive}) \\
&= \mathbf{0}
\end{aligned}$$

We check that $\boldsymbol{D}_{\times}[-]$ is indeed a cartesian differential operator that satisfies [CD1-7] and (D-curry).

[CD1] $\boldsymbol{D}_\times[\boldsymbol{f}+\boldsymbol{g}] = \boldsymbol{D}_\times[\boldsymbol{f}] + \boldsymbol{D}_\times[\boldsymbol{g}]$ and $\boldsymbol{D}_\times[\mathbf{0}] = \mathbf{0}$.

$$\begin{aligned}
&\boldsymbol{D}_\times[\boldsymbol{f}+\boldsymbol{g}] \\
&= \boldsymbol{D}\big[\{x:\alpha\} \vdash M+N:\beta\big] \\
&= \{y:\alpha\times\alpha\} \vdash \Big(\mathsf{D}\big(\lambda x.(M+N)\big)\cdot\mathsf{Fst}(y)\Big)\ \mathsf{Snd}(y):\beta \\
&= \{y:\alpha\times\alpha\} \vdash \big(\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y)\big)\ \mathsf{Snd}(y) + \big(\mathsf{D}(\lambda x.N)\cdot\mathsf{Fst}(y)\big)\ \mathsf{Snd}(y):\beta \\
&= \boldsymbol{D}_\times[\boldsymbol{f}] + \boldsymbol{D}_\times[\boldsymbol{g}]
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{D}_\times[\mathbf{0}] &= \boldsymbol{D}\big[\{x:\alpha\}\vdash 0:\beta\big] \\
&= \{y:\alpha\times\alpha\} \vdash \big(\mathsf{D}(\lambda x.0)\cdot\mathsf{Fst}(y)\big)\ \mathsf{Snd}(y):\beta \\
&= \{y:\alpha\times\alpha\} \vdash 0:\beta = \mathbf{0}
\end{aligned}$$

[CD2] $\boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\boldsymbol{h}+\boldsymbol{k},\boldsymbol{v}\rangle = \boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\boldsymbol{h},\boldsymbol{v}\rangle + \boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\boldsymbol{k},\boldsymbol{v}\rangle$ and $\boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\mathbf{0},\boldsymbol{v}\rangle = \mathbf{0}$.

$$\begin{aligned}
&\boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\boldsymbol{h}+\boldsymbol{k},\boldsymbol{v}\rangle \\
&= \boldsymbol{D}\big[\ \{x:\alpha\}\vdash M:\beta\ \big]\ \circ\ \big\langle\ \{y:\gamma\}\vdash H+K:\alpha\ \ \{y:\gamma\}\vdash V:\alpha\ \big\rangle \\
&= \Big(\{z:\alpha\times\alpha\}\vdash\big(\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(z)\big)\ \mathsf{Snd}(z):\beta\Big)\ \circ\ \big(\{y:\gamma\}\vdash\langle H+K,V\rangle:\alpha\big) \\
&= \{y:\gamma\}\vdash\Big(\big(\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(z)\big)\ \mathsf{Snd}(z)\Big)\ [\langle H+K,V\rangle/z]:\beta \\
&= \{y:\gamma\}\vdash\big(\mathsf{D}(\lambda x.M)\cdot(H+K)\big)\ V:\beta \\
&= \{y:\gamma\}\vdash\big(\mathsf{D}(\lambda x.M)\cdot(H)\big)\ V + \big(\mathsf{D}(\lambda x.M)\cdot(K)\big)\ V:\beta \\
&= \boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\boldsymbol{h},\boldsymbol{v}\rangle + \boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\boldsymbol{k},\boldsymbol{v}\rangle
\end{aligned}$$

$$\boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\mathbf{0},\boldsymbol{v}\rangle = \{y:\gamma\}\vdash\big(\mathsf{D}(\lambda x.M)\cdot 0\big)\ V:\beta = \{y:\gamma\}\vdash 0:\beta = \mathbf{0}$$

[CD3] $\boldsymbol{D}_\times[\mathbf{Id}] = \boldsymbol{\pi_1}$, $\boldsymbol{D}_\times[\boldsymbol{\pi_1}] = \boldsymbol{\pi_1}\circ\boldsymbol{\pi_1}$ and $\boldsymbol{D}_\times[\boldsymbol{\pi_2}] = \boldsymbol{\pi_2}\circ\boldsymbol{\pi_1}$.

$$\begin{aligned}
\boldsymbol{D}_\times[\mathbf{Id}] &= \boldsymbol{D}\big[\ \{x:\alpha\}\vdash x:\alpha\ \big] \\
&= \{y:\alpha\times\alpha\}\vdash\big(\mathsf{D}(\lambda x.x)\cdot\mathsf{Fst}(y)\big)\ \mathsf{Snd}(y):\alpha \\
&= \{y:\alpha\times\alpha\}\vdash\Big(\lambda x.\big(\frac{\partial x}{\partial x}\cdot\mathsf{Fst}(y)\big)\Big)\ \mathsf{Snd}(y):\alpha && (\beta_{\mathsf{D}}) \\
&= \{y:\alpha\times\alpha\}\vdash\big(\lambda x.\mathsf{Fst}(y)\big)\ \mathsf{Snd}(y):\alpha \\
&= \{y:\alpha\times\alpha\}\vdash\mathsf{Fst}(y):\alpha && (\beta) \\
&= \boldsymbol{\pi_1}
\end{aligned}$$

$$\begin{aligned}
&\boldsymbol{D}_\times[\boldsymbol{\pi_1}] \\
&= \boldsymbol{D}\big[\ \{z:\alpha\times\beta\}\vdash\mathsf{Fst}(z):\alpha\ \big] \\
&= \{y:(\alpha\times\beta)\times(\alpha\times\beta)\}\vdash\big(\mathsf{D}(\lambda z.\mathsf{Fst}(z))\cdot\mathsf{Fst}(y)\big)\ \mathsf{Snd}(y):\alpha \\
&= \{y:(\alpha\times\beta)\times(\alpha\times\beta)\}\vdash\Big(\lambda z.\big(\frac{\partial\mathsf{Fst}(z)}{\partial z}\cdot\mathsf{Fst}(y)\big)\Big)\ \mathsf{Snd}(y):\alpha && (\beta_{\mathsf{D}}) \\
&= \{y:(\alpha\times\beta)\times(\alpha\times\beta)\}\vdash\Big(\lambda z.\mathsf{Fst}\big(\frac{\partial z}{\partial z}\cdot\mathsf{Fst}(y)\big)\Big)\ \mathsf{Snd}(y):\alpha \\
&= \{y:(\alpha\times\beta)\times(\alpha\times\beta)\}\vdash\big(\lambda z.\mathsf{Fst}(\mathsf{Fst}(y))\big)\ \mathsf{Snd}(y):\alpha \\
&= \{y:(\alpha\times\beta)\times(\alpha\times\beta)\}\vdash\mathsf{Fst}(\mathsf{Fst}(y)):\alpha && (\beta) \\
&= \boldsymbol{\pi_1}\circ\boldsymbol{\pi_1}
\end{aligned}$$

The proof for $\boldsymbol{D}_\times[\boldsymbol{\pi_2}] = \boldsymbol{\pi_2}\circ\boldsymbol{\pi_1}$ is similar.

[CD4] $\boldsymbol{D_\times[\langle f, g\rangle] = \langle D_\times[f], D_\times[g]\rangle}$

$$
\begin{aligned}
&\boldsymbol{D_\times[\langle f, g\rangle]}\\
&= \{y : \alpha\times\alpha\} \vdash \big(\mathsf{D}(\lambda x.\langle M, N\rangle)\cdot \mathsf{Fst}(y)\big)\ \mathsf{Snd}(y) : \beta\times\gamma\\
&= \{y : \alpha\times\alpha\} \vdash \Big(\lambda x.(\frac{\partial}{\partial x}\langle M, N\rangle \cdot \mathsf{Fst}(y))\Big)\ \mathsf{Snd}(y) : \beta\times\gamma && (\beta_{\mathsf{D}})\\
&= \{y : \alpha\times\alpha\} \vdash \Big(\lambda x.\langle\frac{\partial M}{\partial x}\cdot \mathsf{Fst}(y)\ \ \frac{\partial N}{\partial x}\cdot \mathsf{Fst}(y)\rangle\Big)\ \mathsf{Snd}(y) : \beta\times\gamma\\
&= \{y : \alpha\times\alpha\} \vdash \Big\langle\Big(\lambda x.(\frac{\partial M}{\partial x}\cdot \mathsf{Fst}(y))\Big)\ \mathsf{Snd}(y)\\
&\qquad\qquad\qquad\quad \Big(\lambda x.(\frac{\partial N}{\partial x}\cdot \mathsf{Fst}(y))\Big)\ \mathsf{Snd}(y)\Big\rangle : \beta\times\gamma && (\dagger)\\
&= \langle \boldsymbol{D_\times[f]}, \boldsymbol{D_\times[g]}\rangle
\end{aligned}
$$

(†) Note that

$$
\Gamma \vdash \quad (\lambda x.\langle s, t\rangle)\ u \overset{\beta}{=} \langle s, t\rangle[u/x] = \langle s[u/x], t[u/x]\rangle \overset{\beta}{=} \langle(\lambda x.s)\ u, (\lambda x.t)\ u\rangle \quad : \alpha
$$

[CD5] $\boldsymbol{D_\times[f\circ g] = D_\times[f]\circ\langle D_\times[g], g\circ\pi_2\rangle}$

$$
\begin{aligned}
&\boldsymbol{D_\times[f]\circ\langle D_\times[g], g\circ\pi_2\rangle}\\
&= \Big(\{w : \beta\times\beta\} \vdash \big(\mathsf{D}(\lambda y.N)\cdot \mathsf{Fst}(w)\big)\ \mathsf{Snd}(w) : \gamma\Big) \circ\\
&\quad \Big(\{z : \alpha\times\alpha\} \vdash \big\langle\ \big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(z)\big)\ \mathsf{Snd}(z),\ M[\mathsf{Snd}(z)/x]\ \big\rangle : \beta\times\beta\Big)\\
&= \{z : \alpha\times\alpha\} \vdash \Big(\big(\mathsf{D}(\lambda y.N)\cdot \mathsf{Fst}(w)\big)\ \mathsf{Snd}(w)\Big)\\
&\qquad\qquad\qquad \big[\big\langle\big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(z)\big)\ \mathsf{Snd}(z), M[\mathsf{Snd}(z)/x]\big\rangle/w\big] : \gamma\\
&= \{z : \alpha\times\alpha\} \vdash \Big(\mathsf{D}(\lambda y.N)\cdot \big(\big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(z)\big)\ \mathsf{Snd}(z)\big)\Big)\ M[\mathsf{Snd}(z)/x] : \gamma\\
&= \{z : \alpha\times\alpha\} \vdash \big(\mathsf{D}(\lambda x.N[M/y])\cdot \mathsf{Fst}(z)\big)\ \mathsf{Snd}(z) : \gamma && (\text{Lemma 5.6 (i)})\\
&= \boldsymbol{D}\big[\ \{x : \alpha\} \vdash N[M/y] : \gamma\ \big]\\
&= \boldsymbol{D_\times[f\circ g]}
\end{aligned}
$$

[CD6] $\boldsymbol{D_\times[D_\times[f]]\circ\langle\langle g, 0\rangle, \langle h, k\rangle\rangle = D_\times[f]\circ\langle g, k\rangle}$

$$
\begin{aligned}
&\boldsymbol{D_\times[D_\times[f]]\circ\langle\langle g, 0\rangle, \langle h, k\rangle\rangle}\\
&= \boldsymbol{D}\big[\ \{y : \alpha\times\alpha\} \vdash \big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(y)\big)\ \mathsf{Snd}(y) : \beta\ \big] \quad\circ\\
&\quad \big(\{w : \gamma\} \vdash \langle\langle G, 0\rangle, \langle H, K\rangle\rangle : (\alpha\times\alpha)\times(\alpha\times\alpha)\big)\\
&= \{w : \gamma\} \vdash \Big(\Big(\mathsf{D}\big(\lambda y.\big(\big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(y)\big)\ \mathsf{Snd}(y)\big)\big)\cdot \mathsf{Fst}(z)\Big)\ \mathsf{Snd}(z)\Big)\\
&\qquad\qquad \big[\langle\langle G, 0\rangle, \langle H, K\rangle\rangle/z\big] : \beta\\
&= \{w : \gamma\} \vdash \Big(\mathsf{D}\big(\lambda y.\big(\big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(y)\big)\ \mathsf{Snd}(y)\big)\big)\cdot\langle G, 0\rangle\Big)\ \langle H, K\rangle : \beta\\
&= \{w : \gamma\} \vdash \Big(\lambda y.\Big(\frac{\partial}{\partial y}\big(\big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(y)\big)\ \mathsf{Snd}(y)\big)\cdot\langle G, 0\rangle\Big)\Big)\ \langle H, K\rangle : \beta && (\beta_{\mathsf{D}})\\
&= \{w : \gamma\} \vdash \Big(\lambda y.\Big(\Big(\big(\frac{\partial}{\partial y}\big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(y)\big)\big)\cdot\langle G, 0\rangle\Big)\ \mathsf{Snd}(y) +\\
&\qquad\qquad \Big(\mathsf{D}\big(\mathsf{D}(\lambda x.M)\cdot \mathsf{Fst}(y)\big)\cdot\big(\frac{\partial\ \mathsf{Snd}(y)}{\partial y}\cdot\langle G, 0\rangle\big)\Big)\ \mathsf{Snd}(y)\Big)\Big)\ \langle H, K\rangle : \beta
\end{aligned}
$$

$$= \{w:\gamma\} \vdash \Big(\lambda y.\Big(\Big(\mathsf{D}\big(\lambda x.(\frac{\partial M}{\partial y}\cdot\langle G,0\rangle)\big)\cdot\mathsf{Fst}(y)\ +\ \mathsf{D}(\lambda x.M)\cdot(\frac{\partial\,\mathsf{Fst}(y)}{\partial y}\cdot\langle G,0\rangle)\Big)$$
$$\mathsf{Snd}(y)\ +\ \big(\mathsf{D}\big(\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y)\big)\cdot 0\big)\ \mathsf{Snd}(y)\Big)\Big)\ \langle H,K\rangle:\beta \qquad (\beta_{\mathsf{D}})$$
$$= \{w:\gamma\} \vdash \Big(\lambda y.\big(0+(\mathsf{D}(\lambda x.M)\cdot G)\ \mathsf{Snd}(y)+0\big)\Big)\ \langle H,K\rangle:\beta \qquad \text{(Prop. 2.1)}$$
$$= \{w:\gamma\} \vdash \big(\mathsf{D}(\lambda x.M)\cdot G\big)\ K:\beta \qquad (\beta)$$
$$= \{w:\gamma\} \vdash \big((\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y))\ \mathsf{Snd}(y)\big)\ \big[\langle G,K\rangle/y\big]:\beta$$
$$= \boldsymbol{D}_\times[\boldsymbol{f}]\circ\langle\boldsymbol{g},\boldsymbol{k}\rangle$$

[CD7] $\boldsymbol{D}_\times[\boldsymbol{D}_\times[\boldsymbol{f}]]\circ\langle\langle\boldsymbol{0},\boldsymbol{h}\rangle,\langle\boldsymbol{g},\boldsymbol{k}\rangle\rangle = \boldsymbol{D}_\times[\boldsymbol{D}_\times[\boldsymbol{f}]]\circ\langle\langle\boldsymbol{0},\boldsymbol{g}\rangle,\langle\boldsymbol{h},\boldsymbol{k}\rangle\rangle$

$$\boldsymbol{D}_\times[\boldsymbol{D}_\times[\boldsymbol{f}]]\circ\langle\langle\boldsymbol{0},\boldsymbol{h}\rangle,\langle\boldsymbol{g},\boldsymbol{k}\rangle\rangle$$
$$= \boldsymbol{D}\big[\ \{y:\alpha\times\alpha\}\vdash(\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y))\ \mathsf{Snd}(y):\beta\ \big]\ \circ\ \langle\langle\boldsymbol{0},\boldsymbol{h}\rangle,\langle\boldsymbol{g},\boldsymbol{k}\rangle\rangle$$
$$= \Big(\{z:(\alpha\times\alpha)\times(\alpha\times\alpha)\}\vdash$$
$$\Big(\mathsf{D}\big(\lambda y.((\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y))\ \mathsf{Snd}(y))\big)\cdot\mathsf{Fst}(z)\Big)\ \mathsf{Snd}(z):\beta\Big)\ \circ$$
$$\Big(\{w:\gamma\}\vdash\langle\langle 0,H\rangle,\langle G,K\rangle\rangle:(\alpha\times\alpha)\times(\alpha\times\alpha)\Big)$$
$$= \{w:\gamma\}\vdash\Big(\mathsf{D}\big(\lambda y.((\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y))\ \mathsf{Snd}(y))\big)\cdot\langle 0,H\rangle\Big)\ \langle G,K\rangle:\beta$$
$$= \{w:\gamma\}\vdash\Big(\lambda y.\Big(\frac{\partial}{\partial y}\big((\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y))\ \mathsf{Snd}(y)\big)\cdot\langle 0,H\rangle\Big)\Big)\ \langle G,K\rangle:\beta \qquad (\beta_{\mathsf{D}})$$
$$= \{w:\gamma\}\vdash\Big(\lambda y.\Big(\big(\frac{\partial}{\partial y}\big(\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y)\big)\cdot\langle 0,H\rangle\big)\ \mathsf{Snd}(y)\ +$$
$$\Big(\mathsf{D}\big(\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y)\big)\cdot\Big(\frac{\partial\ \mathsf{Snd}(y)}{\partial y}\cdot\langle 0,H\rangle\Big)\Big)\ \mathsf{Snd}(y)\Big)\Big)\ \langle G,K\rangle:\beta$$
$$= \{w:\gamma\}\vdash\Big(\lambda y.\Big(\big(\mathsf{D}0\cdot\mathsf{Fst}(y)+\mathsf{D}(\lambda x.M)\cdot 0\big)\ \mathsf{Snd}(y)\ +$$
$$\big(\mathsf{D}\big(\mathsf{D}(\lambda x.M)\cdot\mathsf{Fst}(y)\big)\cdot H\big)\ \mathsf{Snd}(y)\Big)\Big)\ \langle G,K\rangle:\beta$$
$$= \{w:\gamma\}\vdash\big(\mathsf{D}(\mathsf{D}(\lambda x.M)\cdot G)\cdot H\big)\ K:\beta \qquad (\beta)$$
$$= \{w:\gamma\}\vdash\big(\mathsf{D}(\mathsf{D}(\lambda x.M)\cdot H)\cdot G\big)\ K:\beta \qquad (sw_{\mathsf{D}})$$
$$= \boldsymbol{D}_\times[\boldsymbol{D}_\times[\boldsymbol{f}]]\circ\langle\langle\boldsymbol{0},\boldsymbol{g}\rangle,\langle\boldsymbol{h},\boldsymbol{k}\rangle\rangle$$

(D-curry) $\mathbf{D}[\boldsymbol{\lambda}(\boldsymbol{f})] = \boldsymbol{\lambda}\big(\mathbf{D}[\boldsymbol{f}]\circ\langle\boldsymbol{\pi_1}\times\boldsymbol{0},\boldsymbol{\pi_2}\times\mathbf{Id}\rangle\big)$

$$\boldsymbol{\lambda}\big(\mathbf{D}[\boldsymbol{f}]\circ\langle\boldsymbol{\pi_1}\times\boldsymbol{0},\boldsymbol{\pi_2}\times\mathbf{Id}\rangle\big)$$
$$= \boldsymbol{\lambda}\Big(\big(\{u:(\alpha\times\beta)\times(\alpha\times\beta)\}\vdash(\mathsf{D}(\lambda z.M)\cdot\mathsf{Fst}(u))\ \mathsf{Snd}(u):\gamma\big)\quad\circ$$
$$\big(\{v:(\alpha\times\alpha)\times\beta\}\vdash\Big\langle\big\langle\mathsf{Fst}(\mathsf{Fst}(v)),0\big\rangle,\big\langle\mathsf{Snd}(\mathsf{Fst}(v)),\mathsf{Snd}(v)\big\rangle\Big\rangle\big)\Big)$$
$$= \boldsymbol{\lambda}\Big(\{v:(\alpha\times\alpha)\times\beta\}\vdash\big(\mathsf{D}(\lambda z.M)\cdot\big\langle\mathsf{Fst}(\mathsf{Fst}(v)),0\big\rangle\big)\ \big\langle\mathsf{Snd}(\mathsf{Fst}(v)),\mathsf{Snd}(v)\big\rangle:\gamma\Big)$$
$$= \{w:\alpha\times\alpha\}\vdash\lambda y.\Big(\Big(\big(\mathsf{D}(\lambda z.M)\cdot\big\langle\mathsf{Fst}(\mathsf{Fst}(v)),0\big\rangle\big)$$
$$\big\langle\mathsf{Snd}(\mathsf{Fst}(v)),\mathsf{Snd}(v)\big\rangle\Big)\ [\langle w,y\rangle/v]\Big):\beta\Rightarrow\gamma$$
$$= \{w:\alpha\times\alpha\}\vdash\lambda y.\Big(\big(\mathsf{D}(\lambda z.M)\cdot\langle\mathsf{Fst}(w),0\rangle\big)\ \langle\mathsf{Snd}(w),y\rangle\Big):\beta\Rightarrow\gamma$$

$$
\begin{aligned}
&= \{w : \alpha \times \alpha\} \vdash \lambda y.\Big(\big(\lambda z.\big(\frac{\partial M}{\partial z} \cdot \langle \mathsf{Fst}(w), 0\rangle\big)\big)\ \langle \mathsf{Snd}(w), y\rangle\Big) : \beta \Rightarrow \gamma && (\beta_{\mathsf{D}})\\
&= \{w : \alpha \times \alpha\} \vdash \lambda y.\Big(\big(\frac{\partial M}{\partial z} \cdot \langle \mathsf{Fst}(w), 0\rangle\big)\ [\langle \mathsf{Snd}(w), y\rangle/z]\Big) : \beta \Rightarrow \gamma && (\beta)\\
&= \{w : \alpha \times \alpha\} \vdash \lambda y.\Big(\big(\frac{\partial\ M[\langle x, y\rangle/z]}{\partial x} \cdot \mathsf{Fst}(w)\big)\ [\mathsf{Snd}(w)/x]\Big) : \beta \Rightarrow \gamma && (5.6\text{ (ii)})\\
&= \{w : \alpha \times \alpha\} \vdash \Big(\lambda xy.\big(\frac{\partial\ M[\langle x, y\rangle/z]}{\partial x} \cdot \mathsf{Fst}(w)\big)\Big)\ \mathsf{Snd}(w) : \beta \Rightarrow \gamma && (\beta)\\
&= \{w : \alpha \times \alpha\} \vdash \Big(\lambda x.\big(\frac{\partial\ (\lambda y.(M[\langle x, y\rangle/z]))}{\partial x} \cdot \mathsf{Fst}(w)\big)\Big)\mathsf{Snd}(w) : \beta \Rightarrow \gamma\\
&= \{w : \alpha \times \alpha\} \vdash (\mathsf{D}(\lambda xy.(M[\langle x, y\rangle/z])) \cdot \mathsf{Fst}(w))\ \mathsf{Snd}(w) : \beta \Rightarrow \gamma && (\beta_{\mathsf{D}})\\
&= \mathbf{D}[\{x : \alpha\} \vdash \lambda y.(M[\langle x, y\rangle/z]) : \beta \Rightarrow \gamma]\\
&= \mathbf{D}[\boldsymbol{\lambda}(\boldsymbol{f})]
\end{aligned}
$$

Thus, $\mathsf{Cl}(\mathscr{T})$ is indeed a cartesian closed differential category. Now, we show that it is classifying by providing the "inverse" of the modelling functors.

Let $\mathbb{G}$ be a structure of $\mathscr{T}$ in $\mathsf{Cl}(\mathscr{T})$ which sets

$$[\![\gamma]\!]_{\mathbb{G}} := \gamma.$$

It is easy to check that $\mathbb{G}$ is indeed a model of $\mathscr{T}$. Let $\mathcal{D}$ be a cartesian closed differential category. We define $\mathsf{DAp}_{\mathbb{G}}^{-1} : \mathsf{DMod}_{\cong}(\mathscr{T}, \mathcal{D}) \to \mathsf{CCDCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D})$ where, for any model $\mathbb{M}$ of $\mathscr{T}$ in $\mathcal{D}$,

$$
\begin{aligned}
\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M}) : \qquad & \mathsf{Cl}(\mathscr{T}) \longrightarrow \mathcal{D}\\
& \alpha \longmapsto [\![\alpha]\!]_{\mathbb{M}}\\
& \{x : \alpha\} \vdash M : \beta \longmapsto [\![\{x : \alpha\} \vdash M : \beta]\!]_{\mathbb{M}} : [\![\alpha]\!]_{\mathbb{M}} \to [\![\beta]\!]_{\mathbb{M}}.
\end{aligned}
$$

Soundness tells us that $\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})$ is well-defined. Note that

$$\Phi := \big\langle \mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{\pi_1}), \mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{\pi_2})\big\rangle = \mathsf{Id} \quad \text{and} \quad \Psi := \lambda\big(\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\mathsf{ev}) \circ \Phi\big) = \mathsf{Id}$$

are isomorphisms. Thus, $\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})$ preserves products and exponentials. For any morphisms $\boldsymbol{g} = \{x : \alpha\} \vdash G : \beta$ and $\boldsymbol{h} = \{x : \alpha\} \vdash H : \beta$,

$$
\begin{aligned}
\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{g} + \boldsymbol{h}) &= [\![\{x : \alpha\} \vdash G + H : \beta]\!]\\
&= [\![\{x : \alpha\} \vdash G : \beta]\!] + [\![\{x : \alpha\} \vdash H : \beta]\!]\\
&= \mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{g}) + \mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{h}).
\end{aligned}
$$

$$\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(0) = [\![\{x : \alpha\} \vdash 0 : \beta]\!] = \mathbf{0}.$$

Hence, $\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})$ preserves the additive structure. Now, consider a morphism $\boldsymbol{f} = \{x : \alpha\} \vdash M : \beta$,

$$
\begin{aligned}
&\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})\big(\boldsymbol{D}_{\times}[\boldsymbol{f}]\big)\\
&= \mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})\Big(\{y : \alpha \times \alpha\} \vdash \big(\mathsf{D}(\lambda x.M) \cdot \mathsf{Fst}(y)\big)\ \mathsf{Snd}(y) : \beta\Big)\\
&= [\![\{y : \alpha \times \alpha\} \vdash \big(\mathsf{D}(\lambda x.M) \cdot \mathsf{Fst}(y)\big)\ \mathsf{Snd}(y) : \beta]\!]\\
&= \mathbf{ev} \circ \Big\langle [\![\{y : \alpha \times \alpha\} \vdash \mathsf{D}(\lambda x.M) \cdot \mathsf{Fst}(y) : \alpha \Rightarrow \beta]\!], [\![\{y : \alpha \times \alpha\} \vdash \mathsf{Snd}(y) : \alpha]\!]\Big\rangle
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(\boldsymbol{\lambda}^{-}([\![\{y:\alpha\times\alpha\} \vdash \lambda x.M : \alpha \Rightarrow \beta]\!]) \star [\![\{y:\alpha\times\alpha\} \vdash \mathsf{Fst}(y) : \alpha]\!]\big), \boldsymbol{\pi_2} \Big\rangle \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(\boldsymbol{\lambda}^{-}(\lambda([\![\{y:\alpha\times\alpha, x:\alpha\} \vdash M : \beta]\!])) \star \boldsymbol{\pi_1}\big), \boldsymbol{\pi_2} \Big\rangle \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big((\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{f}) \circ \boldsymbol{\pi_2}) \star \boldsymbol{\pi_1}\big), \boldsymbol{\pi_2} \Big\rangle \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(\boldsymbol{D}_{\times}\big[\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{f}) \circ \boldsymbol{\pi_2}\big] \circ \langle\langle \mathbf{0}, \boldsymbol{\pi_1} \circ \boldsymbol{\pi_1}\rangle, \mathbf{Id}\rangle\big), \boldsymbol{\pi_2} \Big\rangle \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(\boldsymbol{D}_{\times}\big[\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{f})\big] \circ \langle \boldsymbol{\pi_2} \circ \boldsymbol{\pi_1}, \boldsymbol{\pi_2} \circ \boldsymbol{\pi_2}\rangle \circ \langle\langle \mathbf{0}, \boldsymbol{\pi_1} \circ \boldsymbol{\pi_1}\rangle, \mathbf{Id}\rangle\big), \boldsymbol{\pi_2} \Big\rangle && \text{(CD5)} \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(\boldsymbol{D}_{\times}\big[\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{f})\big] \circ \langle \boldsymbol{\pi_1} \circ \boldsymbol{\pi_1}, \boldsymbol{\pi_2}\rangle\big), \boldsymbol{\pi_2} \Big\rangle \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(\boldsymbol{D}_{\times}\big[\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{f})\big] \circ (\boldsymbol{\pi_1} \times \mathbf{Id})\big), \boldsymbol{\pi_2} \Big\rangle \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(\boldsymbol{D}_{\times}\big[\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{f})\big]\big) \circ \boldsymbol{\pi_1}, \boldsymbol{\pi_2} \Big\rangle && \text{(Prop. A.3)} \\
&= \mathbf{ev} \circ \Big(\boldsymbol{\lambda}\big(\boldsymbol{D}_{\times}\big[\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{f})\big]\big) \times \mathsf{Id}\Big) \\
&= D_{\times}\big[\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})(\boldsymbol{f})\big]
\end{aligned}
$$

Since $\Phi = \mathsf{Id}$, $\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M})$ preserves cartesian differential operator and we conclude that it is a cartesian closed differential functor.

For any additive model homomorphism $h : \mathbb{M} \to \mathbb{N}$ in $\mathsf{DMod}_{\cong}(\mathscr{T}, \mathcal{D})$, we define the natural isomorphism $\mathsf{DAp}_{\mathbb{G}}^{-1}(h) : \mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{M}) \to \mathsf{DAp}_{\mathbb{G}}^{-1}(\mathbb{N})$ by setting

$$\big(\mathsf{DAp}_{\mathbb{G}}^{-1}(h)\big)_{\alpha} := h_{\alpha} : [\![\alpha]\!]_{\mathbb{M}} \to [\![\alpha]\!]_{\mathbb{N}}.$$

We verify that $\big(\mathsf{DAp}_{\mathbb{G}}^{-1}(h)\big)_{\alpha}$ is a natural transformation by induction on the length of the derivation of the differential typed terms $\Gamma \vdash M : \beta$ that the following diagram commutes.

$$
\begin{array}{ccc}
[\![\Gamma]\!]_{\mathbb{M}} & \xrightarrow{h_{\Gamma}} & [\![\Gamma]\!]_{\mathbb{N}} \\
{\scriptstyle [\![\Gamma \vdash M : \beta]\!]_{\mathbb{M}}} \downarrow & & \downarrow {\scriptstyle [\![\Gamma \vdash M : \beta]\!]_{\mathbb{N}}} \\
[\![\beta]\!]_{\mathbb{M}} & \xrightarrow{h_{\beta}} & [\![\beta]\!]_{\mathbb{N}}
\end{array}
$$

The cases for the rules $(var), (abs), (app), (\mathsf{unit}), (pair), (\mathsf{Fst})$ and (Snd) are already considered in the proof of Theorem A.7.

(sum) If $\Gamma \vdash s_i : \alpha$ for all $i \in I$, then $\Gamma \vdash \sum_i s_i : \alpha$ is also a typed term

$$
\begin{aligned}
[\![\Gamma \vdash \sum_i s_i : \alpha]\!]_{\mathbb{N}} \circ h_{\Gamma} &= \Big(\sum_i [\![\Gamma \vdash s_i : \alpha]\!]_{\mathbb{N}}\Big) \circ h_{\Gamma} \\
&= \sum_i \big([\![\Gamma \vdash s_i : \alpha]\!]_{\mathbb{N}} \circ h_{\Gamma}\big) && (\mathcal{D} \text{ is left additive}) \\
&= \sum_i \big(h_{\alpha} \circ [\![\Gamma \vdash s_i : \alpha]\!]_{\mathbb{M}}\big) && \text{(IH)} \\
&= h_{\alpha} \circ \sum_i [\![\Gamma \vdash s_i : \alpha]\!]_{\mathbb{M}} && (h_{\alpha} \text{ is additive}) \\
&= h_{\alpha} \circ [\![\Gamma \vdash \sum_i s_i : \alpha]\!]_{\mathbb{M}}
\end{aligned}
$$

(D) For typed terms $\Gamma \vdash s : \alpha \Rightarrow \beta$ and $\Gamma \vdash t : \alpha$, there is typed term $\Gamma \vdash \mathsf{D}s \cdot t : \alpha \Rightarrow \beta$. Let $S_M := [\![\Gamma \vdash s : \alpha \Rightarrow \beta]\!]_{\mathbb{M}}$, $S_N := [\![\Gamma \vdash s : \alpha \Rightarrow \beta]\!]_{\mathbb{N}}$ and $T_M := [\![\Gamma \vdash t : \alpha]\!]_{\mathbb{M}}$,

$T_N := \llbracket \Gamma \vdash t : \alpha \rrbracket_{\mathbb{N}}$. By inductive hypothesis, we know that $S_N \circ h_\Gamma = h_{\alpha \Rightarrow \beta} \circ S_M$ and $T_N \circ h_\Gamma = h_\alpha \circ T_M$. So, on the left hand side, we have

$$
\begin{aligned}
&\llbracket \Gamma \vdash \mathsf{D}s \cdot t : \alpha \Rightarrow \beta \rrbracket_{\mathbb{N}} \circ h_\Gamma \\
&= \lambda\big(\lambda^-(S_N) \star T_N\big) \circ h_\Gamma \\
&= \lambda\big(D_\times[\lambda^-(S_N)] \circ \langle\langle 0, T_N \circ \pi_1\rangle, \mathsf{Id}\rangle\big) \circ h_\Gamma \\
&= \lambda\big(D_\times[\mathsf{ev}_{\mathbb{N}} \circ \langle S_N \circ \pi_1, \pi_2\rangle] \circ \langle\langle 0, T_N \circ \pi_1\rangle, \mathsf{Id}\rangle \circ (h_\Gamma \times \mathsf{Id})\big) && \text{(Prop. A.3)} \\
&= \lambda\big(D_\times[\mathsf{ev}_{\mathbb{N}}] \circ \big\langle D_\times[\langle S_N \circ \pi_1, \pi_2\rangle], \langle S_N \circ \pi_1, \pi_2\rangle \circ \pi_2\big\rangle \circ \\
&\qquad \langle\langle 0, T_N \circ \pi_1\rangle, \mathsf{Id}\rangle \circ (h_\Gamma \times \mathsf{Id})\big) && \text{(CD5)} \\
&= \lambda\big(D_\times[\mathsf{ev}_{\mathbb{N}}] \circ \big\langle\langle D_\times[S_N] \circ \langle \pi_1 \circ \pi_1, \pi_1 \circ \pi_2\rangle, \pi_2 \circ \pi_1\rangle, \langle S_N \circ \pi_1, \pi_2\rangle \circ \pi_2\big\rangle \circ \\
&\qquad \big\langle\langle 0, T_N \circ h_\Gamma \circ \pi_1\rangle, \langle h_\Gamma \circ \pi_1, \pi_2\rangle\big\rangle\big) && \text{(CD4, 5)} \\
&= \lambda\big(D_\times[\mathsf{ev}_{\mathbb{N}}] \circ \big\langle\langle D_\times[S_N] \circ \langle 0, h_\Gamma \circ \pi_1\rangle, T_N \circ h_\Gamma \circ \pi_1\rangle, \langle S_N \circ h_\Gamma \circ \pi_1, \pi_2\rangle\big\rangle\big) \\
&= \lambda\big(D_\times[\mathsf{ev}_{\mathbb{N}}] \circ \big\langle\langle 0, T_N \circ h_\Gamma \circ \pi_1\rangle, \langle S_N \circ h_\Gamma \circ \pi_1, \pi_2\rangle\big\rangle\big) && \text{(CD2)} \\
&= \lambda\big(D_\times[\mathsf{ev}_{\mathbb{N}}] \circ \big\langle\langle 0, h_\alpha \circ T_M \circ \pi_1\rangle, \langle h_{\alpha \Rightarrow \beta} \circ S_M \circ \pi_1, \pi_2\rangle\big\rangle\big) && \text{(IH)}
\end{aligned}
$$

And on the right hand side,

$$
\begin{aligned}
&h_{\alpha \Rightarrow \beta} \circ \llbracket \Gamma \vdash \mathsf{D}s \cdot t : \alpha \Rightarrow \beta \rrbracket_{\mathbb{M}} \\
&= \lambda\big(h_\beta \circ \mathsf{ev} \circ (\mathsf{Id} \times h_\alpha^{-1})\big) \circ \lambda\big(\lambda^-(S_M) \star T_M\big) \\
&= \lambda\Big(h_\beta \circ \mathsf{ev} \circ (\mathsf{Id} \times h_\alpha^{-1}) \circ \Big(\lambda\big(\lambda^-(S_M) \star T_M\big) \times \mathsf{Id}\Big)\Big) && \text{(Prop. A.3)} \\
&= \lambda\Big(h_\beta \circ \big(\lambda^-(S_M) \star T_M\big) \circ (\mathsf{Id} \times h_\alpha^{-1})\Big) \\
&= \lambda\Big(h_\beta \circ D_\times\big[\lambda^-(S_M)\big] \circ \big\langle\langle 0, T_M \circ \pi_1\rangle, \mathsf{Id}\big\rangle \circ (\mathsf{Id} \times h_\alpha^{-1})\Big) \\
&= \lambda\Big(h_\beta \circ D_\times\big[\mathsf{ev}_{\mathbb{M}} \circ \langle S_M \circ \pi_1, \pi_2\rangle\big] \circ \big\langle\langle 0, T_M \circ \pi_1\rangle, \mathsf{Id}\big\rangle \circ (\mathsf{Id} \times h_\alpha^{-1})\Big) \\
&= \lambda\Big(h_\beta \circ D_\times\big[\mathsf{ev}_{\mathbb{M}}\big] \circ \big\langle\langle D_\times[S_M \circ \pi_1], D_\times[\pi_2]\rangle, \langle S_M \circ \pi_1, \pi_2\rangle \circ \pi_2\big\rangle \circ \\
&\qquad \big\langle\langle 0, T_M \circ \pi_1\rangle, \mathsf{Id}\big\rangle \circ (\mathsf{Id} \times h_\alpha^{-1})\Big) && \text{(CD5)} \\
&= \lambda\Big(h_\beta \circ D_\times\big[\mathsf{ev}_{\mathbb{M}}\big] \circ \big\langle\langle D_\times[S_M] \circ \langle \pi_1 \circ \pi_1, \pi_1 \circ \pi_2\rangle, \pi_2 \circ \pi_1\rangle, \langle S_M \circ \pi_1, \pi_2\rangle \circ \pi_2\big\rangle \circ \\
&\qquad \big\langle\langle 0, T_M \circ \pi_1\rangle, \langle \pi_1, h_\alpha^{-1} \circ \pi_2\rangle\big\rangle\Big) && \text{(CD3, 5)} \\
&= \lambda\Big(h_\beta \circ D_\times\big[\mathsf{ev}_{\mathbb{M}}\big] \circ \big\langle\langle D_\times[S_M] \circ \langle 0, \pi_1\rangle, T_M \circ \pi_1\rangle, \langle S_M \circ \pi_1, h_\alpha^{-1} \circ \pi_2\rangle\big\rangle\Big) \\
&= \lambda\Big(h_\beta \circ D_\times\big[\mathsf{ev}_{\mathbb{M}}\big] \circ \big\langle\langle 0, T_M \circ \pi_1\rangle, \langle S_M \circ \pi_1, h_\alpha^{-1} \circ \pi_2\rangle\big\rangle\Big) && \text{(CD2)} \\
&= \lambda\Big(D_\times\big[\mathsf{ev}_{\mathbb{N}}\big] \circ \big((h_{\alpha \Rightarrow \beta} \times h_\alpha) \times (h_{\alpha \Rightarrow \beta} \times h_\alpha)\big) \circ \\
&\qquad \big\langle\langle 0, T_M \circ \pi_1\rangle, \langle S_M \circ \pi_1, h_\alpha^{-1} \circ \pi_2\rangle\big\rangle\Big) && \text{(IH on } \mathsf{ev}) \\
&= \lambda\Big(D_\times\big[\mathsf{ev}_{\mathbb{N}}\big] \circ \big\langle\langle h_{\alpha \Rightarrow \beta} \circ 0, h_\alpha \circ T_M \circ \pi_1\rangle, \langle h_{\alpha \Rightarrow \beta} \circ S_M \circ \pi_1, h_\alpha \circ h_\alpha^{-1} \circ \pi_2\rangle\big\rangle\Big) \\
&= \lambda\Big(D_\times\big[\mathsf{ev}_{\mathbb{N}}\big] \circ \big\langle\langle 0, h_\alpha \circ T_M \circ \pi_1\rangle, \langle h_{\alpha \Rightarrow \beta} \circ S_M \circ \pi_1, \pi_2\rangle\big\rangle\Big) && (h_{\alpha \Rightarrow \beta} \text{ is additive}) \\
&= \llbracket \Gamma \vdash \mathsf{D}s \cdot t : \alpha \Rightarrow \beta \rrbracket_{\mathbb{N}} \circ h_\Gamma
\end{aligned}
$$

Thus, $\mathsf{DAp}_{\mathbb{G}}^{-1} : \mathsf{DMod}_{\cong}(\mathscr{T}, \mathcal{D}) \to \mathsf{CCDCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D})$ is a natural transformation. Since h_α is an isomorphism for any type α, $\mathsf{DAp}_{\mathbb{G}}^{-1}$ is a natural isomorphism.

Lastly, we check that $\mathsf{DAp}_{\mathbb{G}}$ and $\mathsf{DAp}_{\mathbb{G}}^{-1}$ are indeed equivalence by defining the following natural isomorphisms

$$\mu : \mathsf{DAp}_{\mathbb{G}}\mathsf{DAp}_{\mathbb{G}}^{-1} \cong \mathsf{Id}_{\mathsf{DMod}_{\cong}(\mathscr{T},\mathcal{D})} \qquad \text{and} \qquad \nu : \mathsf{Id}_{\mathsf{CCDCat}_{\cong}(\mathsf{Cl}(\mathscr{T}),\mathcal{D})} \cong \mathsf{DAp}_{\mathbb{G}}^{-1}\mathsf{DAp}_{\mathbb{G}}$$

such that for any model $\mathbb{M}$ of $\mathscr{T}$ in $\mathcal{D}$, $\mu_{\mathbb{M}} : \mathsf{DAp}_{\mathbb{G}}\mathsf{DAp}_{\mathbb{G}}^{-1}\mathbb{M} \to \mathbb{M}$ is defined as

$$(\mu_{\mathbb{M}})_\gamma := \mathsf{Id}_{[\![\gamma]\!]_{\mathbb{M}}} : [\![\gamma]\!]_{\mathsf{DAp}_{\mathbb{G}}\mathsf{DAp}_{\mathbb{G}}^{-1}\mathbb{M}} = [\![\gamma]\!]_{\mathbb{M}} \to [\![\gamma]\!]_{\mathbb{M}}$$

and for any cartesian closed functor $F : \mathsf{Cl}(\mathscr{T}) \to \mathcal{D}$, we define

$$(\nu_F)_\alpha := \mathsf{Id}_{F\alpha} : F\alpha \to \big(\mathsf{DAp}_{\mathbb{G}}^{-1}(\mathsf{DAp}_{\mathbb{G}}F)\big)\alpha = F([\![\alpha]\!]_{\mathbb{G}}) = F\alpha.$$

Obviously, μ and ν are natural isomorphisms. Thus,

$$\mathsf{DAp}_{\mathbb{G}} : \quad \mathsf{CCDCat}_{\cong}(\mathsf{Cl}(\mathscr{T}),\mathcal{D}) \simeq \mathsf{DMod}_{\cong}(\mathscr{T},\mathcal{D}) \quad : \mathsf{DAp}_{\mathbb{G}}^{-1}$$

and $\mathsf{Cl}(\mathscr{T})$ is indeed a classifying category with the model $\mathbb{G}$. □

5.5 Internal Language of Cartesian Closed Differential Category

In this subsection, we assume that we have an extension of differential λ-theory with *constants and function symbols*. We further assume that the soundness and completeness results are preserved in this extended theory. These assumptions are reasonable as shown in [Cro93].

With such an extended theory, given any cartesian closed differential category $\mathcal{C}$, we can define a collection of function symbols which allows us to prove $\mathsf{Cl}(\mathsf{Th}(\mathcal{C})) \simeq \mathcal{C}$. Thus, one can say that the *internal language* of a cartesian closed differential category $\mathcal{C}$ is the differential λ-theory $\mathsf{Th}(\mathcal{C})$. i.e. $\mathsf{Th}(\mathcal{C})$ precisely describes the internal structures of $\mathcal{C}$. This correspondence gives us a way to "prove" properties about the category using the theory.

Let $\mathcal{C}$ be a cartesian closed differential category. We have

- $\mathsf{TV} := Ob(\mathcal{C})$, i.e. every object of $\mathcal{C}$ is a type variable,
- every morphism of the form $k : \top \to A$ is a constant, every morphism of the form $f : A \to B$ is a function symbol and there are function symbols $I_\alpha : [\![\alpha]\!] \to \alpha$ and $J_\alpha : \alpha \to [\![\alpha]\!]$ for every type α, so the collection of function symbols is

$$\begin{aligned}\mathscr{F} := &\{k : A \mid k \in \mathcal{C}(\top, A)\} \cup \{f : A \to B \mid f \in \mathcal{C}(A,B) \text{ and } A \neq \top\} \cup \\ &\{I_\alpha : [\![\alpha]\!] \to \alpha, J_\alpha : \alpha \to [\![\alpha]\!] \mid \alpha \text{ is a type}\}.\end{aligned}$$

The canonical structure $\mathbb{M}$ sets $[\![A]\!] := A$ for every $A \in \mathsf{TV}$, $[\![k]\!] := k$, $[\![f]\!] := f$ and $I_\alpha = J_\alpha := \mathsf{Id}_{[\![\alpha]\!]}$ for every type α. By soundness,

$$\mathsf{Th}(\mathcal{C}) := \{\Gamma \vdash_{\mathcal{D}} s = t : \alpha \mid [\![\Gamma \vdash_{\mathcal{D}} s : \alpha]\!] = [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!]\}$$

is differential λ-theory, and $\mathbb{M}$ is a model of $\mathsf{Th}(\mathcal{C})$ in $\mathcal{C}$.

We now show that $\mathsf{Cl}(\mathsf{Th}(\mathcal{C}))$ and $\mathcal{C}$ are equivalent by considering the following functors.

$$\begin{aligned} Eq : \mathsf{Cl}(\mathsf{Th}(\mathcal{C})) &\longrightarrow \mathcal{C} \\ \alpha &\longmapsto [\![\alpha]\!] \\ \{x:\alpha\} \vdash_{\mathcal{D}} M : \beta &\longmapsto [\![\{x:\alpha\} \vdash_{\mathcal{D}} M : \beta]\!] \end{aligned} \qquad \begin{aligned} Eq^{-1} : \mathcal{C} &\longrightarrow \mathsf{Cl}(\mathsf{Th}(\mathcal{C})) \\ A &\longmapsto A \\ f : A \to B &\longmapsto \{x : A\} \vdash_{\mathcal{D}} f(x) : B \end{aligned}$$

It is easy to see that $Eq \circ Eq^{-1} \cong \mathsf{Id}_{\mathcal{C}}$. To show $Eq^{-1} \circ Eq \cong \mathsf{Id}_{\mathsf{Cl}(\mathsf{Th}(\mathcal{C}))}$, we define natural transformations $\mu : Eq^{-1} \circ Eq \to \mathsf{Id}_{\mathsf{Cl}(\mathsf{Th}(\mathcal{C}))}$ and $\nu : \mathsf{Id}_{\mathsf{Cl}(\mathsf{Th}(\mathcal{C}))} \to Eq^{-1} \circ Eq$ where for every type α,

$$\mu_\alpha := \{x : [\![\alpha]\!]\} \vdash_{\mathcal{D}} I_\alpha(x) : \alpha \qquad\qquad \nu_\alpha := \{x : \alpha\} \vdash_{\mathcal{D}} J_\alpha(x) : [\![\alpha]\!].$$

It is easy to show that they are indeed natural isomorphisms.

Thus, $Eq : \mathsf{Cl}(\mathsf{Th}(\mathcal{C})) \simeq \mathcal{C} : Eq^{-1}$ and the *internal language* of $\mathcal{C}$ is the differential λ-theory $\mathsf{Th}(\mathcal{C})$. We illustrate in the following example that one can use $\mathsf{Th}(\mathcal{C})$ to reason about $\mathcal{C}$.

Example 5.4. Let $\mathcal{C}$ be a cartesian closed differential category and $f : (C \times A) \times D \to B, g : C \to A$ and $h : C \to B'$ be morphisms in $\mathcal{C}$ and $sw := \langle\langle \pi_1 \circ \pi_1, \pi_2\rangle, \pi_2 \circ \pi_1\rangle : (A \times B) \times C \to (A \times C) \times B$.

(i) $\pi_2 \star g = g \circ \pi_1$

(ii) $(h \circ \pi_1) \star g = 0$

(iii) $\lambda(f) \star g = \lambda\big(((f \circ sw) \star (g \circ \pi_1)) \circ sw\big)$

We convert f, g and h to typed terms using function symbols and write $[\![-]\!]$ for the canonical model of $\mathsf{Th}(\mathcal{C})$ in $\mathcal{C}$. Since $\mathsf{Th}(\mathcal{C})$ is the internal language of $\mathcal{C}$, we can prove the results by showing that the corresponding typed terms are the same.

(i) Writing $\pi_2 = [\![\{x : C, y : A\} \vdash_{\mathcal{D}} y : A]\!]$,

$$\begin{aligned}
\pi_2 \star g &= [\![\{x : C, y : A\} \vdash_{\mathcal{D}} y : A]\!] \star [\![\{x : C\} \vdash_{\mathcal{D}} g(x) : A]\!] \\
&= [\![\{x : C, y : A\} \vdash_{\mathcal{D}} \frac{\partial y}{\partial y} \cdot g(x) : A]\!] && \text{(Lemma 5.3)} \\
&= [\![\{x : C, y : A\} \vdash_{\mathcal{D}} g(x) : A]\!] \\
&= [\![\{z : C \times A\} \vdash_{\mathcal{D}} g(\mathsf{Fst}(z)) : A]\!] \\
&= g \circ \pi_1
\end{aligned}$$

(ii) Note that the result of differential substitution in a function symbol where the variable is not free is zero.

$$\begin{aligned}
(h \circ \pi_1) \star g &= [\![\{y : C, x : A\} \vdash_{\mathcal{D}} h(y) : B']\!] \star [\![\{y : C\} \vdash_{\mathcal{D}} g(y) : A]\!] \\
&= [\![\{y : C, x : A\} \vdash_{\mathcal{D}} \frac{\partial h(y)}{\partial x} \cdot g(x) : A]\!] && \text{(Lemma 5.3)} \\
&= [\![\{y : C, x : A\} \vdash_{\mathcal{D}} 0 : A]\!] \\
&= 0
\end{aligned}$$

(iii) Note that we can write $sw_1 := [\![\{(x : C, y : D), z : A\} \vdash_{\mathcal{D}} \langle\langle x, z\rangle, y\rangle : (C \times A) \times D]\!]$ and $sw_2 := [\![\{(r : C, s : A), l : D\} \vdash_{\mathcal{D}} \langle\langle r, l\rangle, s\rangle : (C \times D) \times A]\!]$. We first look at the morphism $((f \circ sw_1) \star (g \circ \pi_1)) \circ sw_2$.

$$\begin{aligned}
&((f \circ sw_1) \star (g \circ \pi_1)) \circ sw_2 \\
&= \big([\![\{(x : C, y : D), z : A\} \vdash_{\mathcal{D}} f(\langle\langle x, z\rangle, y\rangle) : B]\!] \star [\![\{x : C, y : D\} \vdash_{\mathcal{D}} g(x) : A]\!]\big) \circ sw_2 \\
&= \big([\![\{(x : C, y : D), z : A\} \vdash_{\mathcal{D}} \frac{\partial\, f(\langle\langle x, z\rangle, y\rangle)}{\partial z} \cdot g(x) : B]\!]\big) \circ sw_2 && \text{(Lemma 5.3)} \\
&= [\![\{(r : C, s : A), l : D\} \vdash_{\mathcal{D}} \frac{\partial\, f(\langle\langle r, s\rangle, l\rangle)}{\partial s} \cdot g(r) : B]\!]
\end{aligned}$$

$$
\begin{aligned}
&\lambda\big(((f \circ sw_1) \star (g \circ \pi_1)) \circ sw_2\big) \\
&= [\![\{r : C, s : A\} \vdash_{\mathcal{D}} \lambda l.\Big(\frac{\partial\ f(\langle\langle r, s\rangle, l\rangle)}{\partial s} \cdot g(r)\Big) : B]\!] \\
&= [\![\{r : C, s : A\} \vdash_{\mathcal{D}} \frac{\partial\ \lambda l.f(\langle\langle r, s\rangle, l\rangle)}{\partial s} \cdot g(r) : B]\!] \\
&= [\![\{r : C, s : A\} \vdash_{\mathcal{D}} \lambda l.f(\langle\langle r, s\rangle, l\rangle) : B]\!] \star [\![\{z : C\} \vdash_{\mathcal{D}} g(z) : A]\!] \qquad \text{(Lemma 5.3)} \\
&= \lambda(f) \star g
\end{aligned}
$$

Note that these proofs are more straightforward than the categorical proofs given in [Man12].

We have shown that one can reason about cartesian closed differential category using differential λ-theory. Is it possible to reason about differential λ-theories using cartesian closed differential category? The main appeal of doing so is to allow the use of proof techniques available in the more abstract mathematical structure (category) in reasoning about differential λ-theories.

Following the definition given in [Sim95], we say a cartesian closed differential category is *complete* with respect to a differential λ-theory $\mathscr{T}$ if

$$
\mathscr{T} \rhd \Gamma \vdash_{\mathcal{D}} s = t : \alpha \qquad \Longleftrightarrow \qquad \text{for any structure } \mathbb{M}, \quad [\![\Gamma \vdash_{\mathcal{D}} s : \alpha]\!] = [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!]
$$

Consider the "smallest" differential λ-theory, namely $\lambda\beta\eta_{\mathsf{D}}$. It is easy to see that the classifying category $\mathsf{Cl}(\lambda\beta\eta_{\mathsf{D}})$ is complete with respect to $\lambda\beta\eta_{\mathsf{D}}$. However, the classifying category is too syntactical to be useful. In the following subsection, we consider the relational model and show that it is incomplete with respect to $\lambda\beta\eta_{\mathsf{D}}$.

5.6 Relational Model

The *relational model* MRel is the main example of cartesian closed differential category known in the literature [Gir88; BEM07; BEM10; Man12]. Other examples are the finiteness spaces semantics MFin in [BEM10] and game semantics in [LMM13; TO]. Despite being very simple, we can define a cartesian differential operator in MRel and prove that it is a cartesian closed differential category.

We describe MRel directly, where

- objects are all the sets,
- morphism $f : A \to B$ is a relation from $\mathcal{N}(A)$ to B, i.e. $\mathsf{MRel}(A, B) := \mathcal{P}(\mathcal{N}(A) \times B)$,
- identity of A is the relation $\mathsf{Id}_A := \{([a], a) : a \in A\}$,
- composition of $s \in \mathsf{MRel}(A, B)$ and $r \in \mathsf{MRel}(B, C)$ is given by

$$
r \circ s := \left\{ (m, c) : \begin{array}{l} \exists k \in \mathbb{N} \text{ and } \exists (m_1, b_1), \ldots, (m_k, b_k) \in s \text{ such that} \\ m = m_1 \uplus \cdots \uplus m_k \text{ and } ([b_1, \ldots, b_k], c) \in r\} \end{array} \right\}.
$$

It is easy to verify that MRel is indeed a category. We now give the cartesian closed structure of MRel.

- For any objects A and B, the categorical product $A \times B$ is given by the disjoint union $A \uplus B := (\{1\} \times A) \cup (\{2\} \times B)$, and the terminal object is given by $\varnothing$.

- For any morphisms $s \in \mathsf{MRel}(A, B_1)$ and $t \in \mathsf{MRel}(A, B_2)$, the pairing is given by

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

- For any objects A and B, the exponential object $A \Rightarrow B$ is given by $\mathcal{N}(A) \times B$ and

$$\mathsf{ev} := \{((([(m, b)], m), b) : m \in \mathcal{N}(A), b \in B\}.$$

  For any morphism $s \in \mathsf{MRel}(A \times B, C)$, there is a unique morphism

$$\lambda(s) := \{(p, (m, b)) : ((p, m), b) \in s\}$$

  such that $s = \mathsf{ev} \circ (\lambda(s) \times \mathsf{Id})$.

As shown in [BEM07], MRel is a cartesian closed category. It is also a cartesian closed left additive category with a cartesian differential operator.

- Each homset has a commutative monoid $(\mathsf{MRel}(A, B), \cup, \varnothing)$.
- Given $s \in \mathsf{MRel}(A, B)$, its derivative is

$$D_\times[s] = \{((([a], m), b) : (m \uplus [a], b) \in s\}$$

It is easy to see that $D_\times[-]$ satisfies the axioms [CD1-7] and (D-curry). Thus, MRel is a cartesian closed differential category.

By soundness, we have

$$\lambda\beta\eta_{\mathsf{D}} \rhd \Gamma \vdash_{\mathcal{D}} s = t : \alpha \quad \Longrightarrow \quad \text{for any structure } \mathbb{M} \text{ in } \mathsf{MRel}, [\![\Gamma \vdash_{\mathcal{D}} s : \alpha]\!] = [\![\Gamma \vdash_{\mathcal{D}} t : \alpha]\!].$$

We prove that the converse is false. Let $\mathbb{M}$ be a structure in MRel and $\Gamma \vdash_{\mathcal{D}} s : \alpha$. Say $[\![\Gamma \vdash_{\mathcal{D}} s : \alpha]\!] = f : [\![\Gamma]\!] \to [\![\alpha]\!]$. Note that in $\lambda\beta\eta_{\mathsf{D}}$, sums are *not* idempotent. i.e. $\lambda\beta\eta_{\mathsf{D}} \vdash_{\mathcal{D}} s \neq s + s : \alpha$. However,

$$\begin{aligned}
[\![\Gamma \vdash_{\mathcal{D}} s + s : \alpha]\!] &= [\![\Gamma \vdash_{\mathcal{D}} s : \alpha]\!] \cup [\![\Gamma \vdash_{\mathcal{D}} s : \alpha]\!] \\
&= f \cup f \\
&= f \\
&= [\![\Gamma \vdash_{\mathcal{D}} s : \alpha]\!]
\end{aligned}$$

Thus, MRel is not complete.

It is actually quite difficult to find a "meaningful" cartesian closed differential category that is complete with respect to $\lambda\beta\eta_{\mathsf{D}}$.

6 Conclusion

Summary At the start of this project, we followed the research of differentiation in λ-calculus. We explained how the syntax of differential λ-calculus is discovered and showed that it is resource-sensitive by providing translation maps between differential λ-calculus and resource λ-calculus. After that, we turned to research of differentiation in the categorical setting. We described the cartesian closed differential category and presented the proof that every structure in a cartesian closed differential category is a model of $\lambda\beta\eta_{\mathsf{D}}$.

Next, I have provided a completeness result for differential λ-theory, thus confirmed that cartesian closed differential categories precisely capture differential λ-theory. Moreover, assuming that we can extend differential λ-theory with constants and function symbols, I have showed that the internal language of a cartesian closed differential category $\mathcal{C}$ is exactly the differential λ-theory $\mathsf{Th}(\mathcal{C})$, and have given examples to illustrate how one can "prove" properties about the category using differential λ-theories. Finally, I have showed that the relational model is not complete with respect to $\lambda\beta\eta_{\mathsf{D}}$.

This project can be improved by including a treatment of differential λ-theory with constants and function symbols, which is not expected to pose any difficulty.

Further Work Inspired by [Sim95], it would be interesting to find the requirements for cartesian closed differential categories to be complete with respect to $\lambda\beta\eta_{\mathsf{D}}$, or an arbitrary differential λ-theory.

In a more categorical direction, Laird *et al.* gave a construction of cartesian closed differential category from symmetric monoidal closed category in [LMM13]. As suggested in [BEM10], it would then be interesting to see if one can transfer the graphical notion of symmetric monoidal closed category to cartesian closed differential category.

References

[BCS06] Richard Blute, J. Robin B. Cockett, and R. A. G. Seely. "Differential categories". In: *Mathematical Structures in Computer Science* 16.6 (2006), pp. 1049–1083.

[BCS09] Richard Blute, J. Robin B. Cockett, and R. A. G. Seely. "Cartesian differential categories". In: *Theory and Applications of Categories* 22.23 (2009), pp. 622–672.

[BEM07] Antonio Bucciarelli, Thomas Ehrhard, and Giulio Manzonetto. "Not Enough Points Is Enough". In: *Computer Science Logic, 21st International Workshop, CSL 2007, 16th Annual Conference of the EACSL, Lausanne, Switzerland, September 11-15, 2007, Proceedings*. Ed. by Jacques Duparc and Thomas A. Henzinger. Vol. 4646. Lecture Notes in Computer Science. Springer, 2007, pp. 298–312. ISBN: 978-3-540-74914-1.

[BEM10] Antonio Bucciarelli, Thomas Ehrhard, and Giulio Manzonetto. "Categorical Models for Simply Typed Resource Calculi". In: *Electr. Notes Theor. Comput. Sci.* 265 (2010), pp. 213–230.

[Bou93] Gérard Boudol. *The Lambda-Calculus with Multiplicities*. Preliminiary Report. INRIA-Sophia Antipolis, 1993.

[Chu41] Alonzo Church. *The Calculi of Lambda-Conversion*. Princeton University Press, 1941.

[Cro93] Roy L. Crole. *Categories for Types*. Cambridge University Press, 1993. ISBN: 0-521-45701-7.

[Die60] Jean Dieudonné. *Foundations of Modern Analysis*. London: Academic Press Inc., 1960. Chap. 8, pp. 141–190.

[ER03] Thomas Ehrhard and Laurent Regnier. "The differential lambda-calculus". In: *Theor. Comput. Sci.* 309.1-3 (2003), pp. 1–41.

[ER08] Thomas Ehrhard and Laurent Regnier. "Uniformity and the Taylor expansion of ordinary lambda-terms". In: *Theor. Comput. Sci.* 403.2-3 (2008), pp. 347–372.

[Gir87] Jean-Yves Girard. "Linear Logic". In: *Theor. Comput. Sci.* 50 (1987), pp. 1–102.

[Gir88] Jean-Yves Girard. "Normal functors, power series and λ-calculus". In: *Ann. Pure Appl. Logic* 37.2 (1988), pp. 129–177.

[Lam80] Joachim Lambek. "From lambda-calculus to cartesian closed categories". In: *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*. Ed. by J. P. Seldin and J. Hindley. Academic Press, 1980, pp. 376–402.

[LMM13] Jim Laird, Giulio Manzonetto, and Guy McCusker. "Constructing differential categories and deconstructing categories of games". In: *Inf. Comput.* 222 (2013), pp. 247–264.

[Man12] Giulio Manzonetto. "What is a categorical model of the differential and the resource λ-calculi?" In: *Mathematical Structures in Computer Science* 22.3 (2012), pp. 451–520.

[Mil92] Robin Milner. "Functions as Processes". In: *Mathematical Structures in Computer Science* 2.2 (1992), pp. 119–141.

[PT09] Michele Pagani and Paolo Tranquilli. "Parallel Reduction in Resource Lambda-Calculus". In: *Programming Languages and Systems, 7th Asian Symposium, APLAS 2009, Seoul, Korea, December 14-16, 2009. Proceedings*. Ed. by Zhenjiang Hu. Vol. 5904. Lecture Notes in Computer Science. Springer, 2009, pp. 226–242. ISBN: 978-3-642-10671-2.

[Sim95] Alex K. Simpson. "Categorical completeness results for the simply-typed lambda-calculus". In: *Typed Lambda Calculi and Applications, Second International Conference on Typed Lambda Calculi and Applications, TLCA '95, Edinburgh, UK, April 10-12, 1995, Proceedings*. Ed. by Mariangiola Dezani-Ciancaglini and Gordon D. Plotkin. Vol. 902. Lecture Notes in Computer Science. Springer, 1995, pp. 414–427. ISBN: 3-540-59048-X.

[TO] Takeshi Tsukada and C.-H. L. Ong. "Plays as resource terms via non-idempotent intersection types". In: *ACM/IEEE Symposium on Logic in Computer Science, LICS 2016, Proceedings*, to appear.

[Tra11] Paolo Tranquilli. "Intuitionistic differential nets and lambda-calculus". In: *Theor. Comput. Sci.* 412.20 (2011), pp. 1979–1997.

[Vau07] Lionel Vaux. "The differential lambdamu-calculus". In: *Theor. Comput. Sci.* 379.1-2 (2007), pp. 166–209.

Index

Appendices

A Simply-typed λ-Calculus and Cartesian Closed Category

This section is targeted for readers who are not familiar with the soundness and completeness theorems for simply-typed λ-calculus and cartesian closed categories. Most material in this section is borrowed from Crole's book [Cro93].

First we recall the syntax of *simply-typed λ-calculus*. Note that to prove completeness, we include projection and pairing terms in our calculus. After that, we describe the categorical semantics of the simply-typed λ-calculus in a *cartesian closed category* with respect to a *structure* $\mathbb{M}$, which gives the interpretations of all the type variables. We say a structure is a *model* of a theory if it satisfies every rule in the theory. i.e. If $s = t$ is a rule in the theory, the morphisms interpreting s and t are the same in the category. We then prove the soundness theorem stating that any structure is a model of the $\lambda\beta\eta$ theory. A corollary of soundness shows that every cartesian closed category $\mathcal{C}$ give rise to a λ-theory $\mathsf{Th}(\mathcal{C})$. After that, we prove the converse. Given a λ-theory $\mathscr{T}$, we construct a classifying cartesian closed category $\mathsf{Cl}(\mathscr{T})$ which is the "smallest" cartesian closed category that can soundly model the theory.

A.1 Syntax

Definition A.1 (λ-terms)**.** Assume we have a infinitely countable set of variables $\mathcal{V}$. The collection Λ of *λ-terms* is defined as follows:

$$\Lambda : \quad s,t,u,v ::= \; x \mid \lambda x.s \mid s\,t \mid \langle\rangle \mid \langle s,t\rangle \mid \mathsf{Fst}(s) \mid \mathsf{Snd}(s) \qquad \text{where } x \in \mathcal{V}$$

Remark. We consider λ-terms up to α-conversion, indicated by $\equiv$. The set of all free variables in a term $\mathsf{FV}(-)$ and capture-free substitution $s\,[t/x]$ are defined as usual.

Definition A.2 (Typed Terms)**.** Assume we have a collection of type variables TV. *Types* and *type contexts* are defined as follows:

Types	$\alpha,\beta ::= \mathsf{unit} \mid \gamma \mid \alpha\times\beta \mid \alpha\Rightarrow\beta$	where $\gamma \in \mathsf{TV}$
Type Contexts	$\Gamma,\Delta ::= \varnothing \mid \Gamma\cup\{x:\alpha\}$	assuming that $\{x:\alpha\}\notin\Gamma$

A λ-term s is a *typed term* if there is a type context Γ and a type α such that $\Gamma \vdash s : \alpha$ is derivable in the type system with the following rules:

$$(var)\,\frac{}{\Gamma\cup\{x:\alpha\}\vdash x:\alpha} \qquad\qquad (abs)\,\frac{\Gamma\cup\{x:\alpha\}\vdash s:\beta}{\Gamma\vdash\lambda x.s:\alpha\Rightarrow\beta}$$

$$(app)\,\frac{\Gamma\vdash s:\alpha\Rightarrow\beta \qquad \Gamma\vdash t:\alpha}{\Gamma\vdash s\,t:\beta}$$

$$(\mathsf{unit})\ \frac{}{\Gamma \vdash \langle\rangle : \mathsf{unit}} \qquad (pair)\ \frac{\Gamma \vdash s : \alpha \qquad \Gamma \vdash t : \beta}{\Gamma \vdash \langle s, t\rangle : \alpha \times \beta}$$

$$(\mathsf{Fst})\ \frac{\Gamma \vdash p : \alpha \times \beta}{\Gamma \vdash \mathsf{Fst}(p) : \alpha} \qquad (\mathsf{Snd})\ \frac{\Gamma \vdash p : \alpha \times \beta}{\Gamma \vdash \mathsf{Snd}(p) : \beta}$$

Proposition A.1 (Weakening)**.** Let $\Gamma \vdash s : \alpha$ be a typed term and $\Gamma \subseteq \Gamma'$, then $\Gamma' \vdash s : \alpha$ is derivable.

Definition A.3 (Theory)**.** A *simply-typed theory* $\mathscr{T}$ is a collection of rules of the form $\Gamma \vdash s = t : \alpha$, where $\Gamma \vdash s : \alpha$ and $\Gamma \vdash t : \alpha$ are derivable. We write $\mathscr{T} \rhd \Gamma \vdash s = t : \alpha$ to indicate that $\Gamma \vdash s = t : \alpha$ is a rule in $\mathscr{T}$. The *simply-typed $\lambda\beta\eta$-theory*, denoted by $\lambda\beta\eta$, is the smallest theory that is reflective, symmetric, transitive and closed under and the following rules,

$$(app)\ \frac{\Gamma \vdash s = s' : \alpha \Rightarrow \beta \qquad \Gamma \vdash t = t' : \alpha}{\Gamma \vdash st = s't' : \beta} \qquad (abs)\ \frac{\Gamma \cup \{x : \alpha\} \vdash s = t : \beta}{\Gamma \vdash \lambda x.s = \lambda x.t : \alpha \Rightarrow \beta}$$

$$(\beta)\ \frac{\Gamma \cup \{x : \alpha\} \vdash s : \beta \qquad \Gamma \vdash t : \alpha}{\Gamma \vdash (\lambda x.s)\, t = s\,[t/x] : \beta} \qquad (\eta)\ \frac{x \notin \mathsf{FV}(s)}{\Gamma \vdash \lambda x.s\, x = s : \alpha \Rightarrow \beta}$$

$$(\mathsf{unit})\ \frac{\Gamma \vdash s : \mathsf{unit}}{\Gamma \vdash s = \langle\rangle : \mathsf{unit}} \qquad (pair)\ \frac{\Gamma \vdash p : \alpha \times \beta}{\Gamma \vdash \langle \mathsf{Fst}(p), \mathsf{Snd}(p)\rangle = p : \alpha \times \beta}$$

$$(\mathsf{Fst})\ \frac{\Gamma \vdash s : \alpha \qquad \Gamma \vdash t : \beta}{\Gamma \vdash \mathsf{Fst}(\langle s, t\rangle) = s : \alpha} \qquad (\mathsf{Snd})\ \frac{\Gamma \vdash s : \alpha \qquad \Gamma \vdash t : \beta}{\Gamma \vdash \mathsf{Snd}(\langle s, t\rangle) = t : \beta}$$

A *simply-typed λ-theory* is a theory that is closed under all the rules in $\lambda\beta\eta$.

Remark. Since we are only considering simply-typed theories in this project, we write theory meaning simply-typed theory.

A.2 Categorical Semantics of Typed Terms

Please refer to [Cro93] for a detailed discussion on how such a categorical semantics arises. Here, we only provide the definition and point out essential lemmas in proving the soundness and completeness results.

We first provide the categorical interpretation of the syntax with respect to a structure in a cartesian closed category.

Definition A.4 (Categorical Semantics of Typed Terms)**.** Let $\mathcal{C}$ be a cartesian closed category. A *structrue* $\mathbb{M}$ in $\mathcal{C}$ is specified by giving each type variable $\gamma \in \mathsf{TV}$ an object $[\![\gamma]\!]_{\mathbb{M}}$ of $\mathcal{C}$. The *interpretation* of types and typed terms with respect to the structure $\mathbb{M}$ is defined by induction as follows, where $\top$ is the terminal object of $\mathcal{C}$ and τ_A is the unique morphism from the object A to $\top$,

- Types
  $[\![\mathsf{unit}]\!] := \top$
  $[\![\gamma]\!] := [\![\gamma]\!]_{\mathbb{M}}$, where $\gamma \in \mathsf{TV}$
  $[\![\alpha \times \beta]\!] := [\![\alpha]\!] \times [\![\beta]\!]$
  $[\![\alpha \Rightarrow \beta]\!] := [\![\alpha]\!] \Rightarrow [\![\beta]\!]$
- Type Contexts
  $[\![\varnothing]\!] := \top$
  $[\![\Gamma \cup \{x : \alpha\}]\!] := [\![\Gamma]\!] \times [\![\alpha]\!]$
- Typed Terms

$$(var)\ \frac{}{[\![\Gamma \cup \{x : \alpha\} \vdash x : \alpha]\!] := \pi_2 : [\![\Gamma]\!] \times [\![\alpha]\!] \to [\![\alpha]\!]}$$

$$(abs)\ \frac{[\![\Gamma \cup \{x : \alpha\} \vdash s : \beta]\!] = f : [\![\Gamma]\!] \times [\![\alpha]\!] \to [\![\beta]\!]}{[\![\Gamma \vdash \lambda x.s : \alpha \Rightarrow \beta]\!] := \lambda(f) : [\![\Gamma]\!] \to ([\![\alpha]\!] \Rightarrow [\![\beta]\!])}$$

$$(app)\ \frac{[\![\Gamma \vdash s : \alpha \Rightarrow \beta]\!] = S : [\![\Gamma]\!] \to ([\![\alpha]\!] \Rightarrow [\![\beta]\!]) \qquad [\![\Gamma \vdash t : \alpha]\!] = T : [\![\Gamma]\!] \to [\![\alpha]\!]}{[\![\Gamma \vdash s\, t : \beta]\!] := \mathsf{ev} \circ \langle S, T\rangle : [\![\Gamma]\!] \to ([\![\alpha]\!] \Rightarrow [\![\beta]\!]) \times [\![\alpha]\!] \to [\![\beta]\!]}$$

$$(\mathsf{unit})\ \frac{}{[\![\Gamma \vdash \langle\rangle : \mathsf{unit}]\!] := \tau_{[\![\Gamma]\!]} : [\![\Gamma]\!] \to \top}$$

$$(pair)\ \frac{[\![\Gamma \vdash s : \alpha]\!] = S \qquad [\![\Gamma \vdash t : \beta]\!] = T}{[\![\Gamma \vdash \langle s, t\rangle : \alpha \times \beta]\!] := \langle S, T\rangle : [\![\Gamma]\!] \to [\![\alpha]\!] \times [\![\beta]\!]}$$

$$(\mathsf{Fst})\ \frac{[\![\Gamma \vdash p : \alpha \times \beta]\!] = P : [\![\Gamma]\!] \to [\![\alpha]\!] \times [\![\beta]\!]}{[\![\Gamma \vdash \mathsf{Fst}(p) : \alpha]\!] := \pi_1 \circ P : [\![\Gamma]\!] \to [\![\alpha]\!]}$$

$$(\mathsf{Snd})\ \frac{[\![\Gamma \vdash p : \alpha \times \beta]\!] = P : [\![\Gamma]\!] \to [\![\alpha]\!] \times [\![\beta]\!]}{[\![\Gamma \vdash \mathsf{Snd}(p) : \beta]\!] := \pi_2 \circ P : [\![\Gamma]\!] \to [\![\beta]\!]}$$

We say a structure $\mathbb{M}$ *satisfies* a rule $\Gamma \vdash s = t : \alpha$ if the interpretations of $\Gamma \vdash s : \alpha$ and $\Gamma \vdash t : \alpha$ with respect to $\mathbb{M}$ are the same. i.e. $[\![\Gamma \vdash s : \alpha]\!] = [\![\Gamma \vdash t : \alpha]\!]$. We say a structure $\mathbb{M}$ is a *model* of a theory $\mathscr{T}$ if $\mathbb{M}$ satisfies all the rules in $\mathscr{T}$. i.e.

$$\mathscr{T} \rhd \Gamma \vdash s = t : \alpha \qquad \Longrightarrow \qquad [\![\Gamma \vdash s : \alpha]\!] = [\![\Gamma \vdash t : \alpha]\!].$$

Remark. Since any λ-theory must be closed under all the rules in $\lambda\beta\eta$, a model of any λ-theory must also be a model of $\lambda\beta\eta$.

A.3 Soundness and Completeness Theorems

Now we show that this interpretation actually makes sense. i.e. If two typed terms are equal in $\lambda\beta\eta$, their interpretations are also the same with respect to any structure in any cartesian closed category. First we show that capture-free substitution is soundly modelled by categorical composition with the help of the following propositions on cartesian closed category.

Proposition A.2. Let $\mathcal{C}$ be a cartesian closed category and $t : A \to D$ be a morphism in $\mathcal{C}$. Let $g : (C \times A) \times D \to B$ and $k : (C \times D) \times A \to B$ be morphisms such that $k = g \circ \phi$, where $\phi := \langle \pi_1 \times \mathsf{Id}_A, \pi_2 \circ \pi_1 \rangle : (C \times D) \times A \to (C \times A) \times D$. Then,

$$g \circ \langle \mathsf{Id}_C \times \mathsf{Id}_A, t \circ \pi_1 \rangle = k \circ \langle \mathsf{Id}_C, t \rangle \times \mathsf{Id}_A.$$

Proof. Consider

$$\begin{aligned} \phi \circ \langle \mathsf{Id}_C, t \rangle \times \mathsf{Id}_A &= \langle \pi_1 \times \mathsf{Id}_A, \pi_2 \circ \pi_1 \rangle \circ \big(\langle \mathsf{Id}_C, t \rangle \times \mathsf{Id}_A \big) \\ &= \big\langle \big(\pi_1 \times \mathsf{Id}_A\big) \circ \big(\langle \mathsf{Id}_C, t \rangle \times \mathsf{Id}_A \big), \pi_2 \circ \pi_1 \circ \big(\langle \mathsf{Id}_C, t \rangle \times \mathsf{Id}_A \big) \big\rangle \\ &= \big\langle \big(\pi_1 \circ \langle \mathsf{Id}_C, t \rangle\big) \times \mathsf{Id}_A, \pi_2 \circ \pi_1 \circ \big\langle \langle \mathsf{Id}_C, t \rangle \circ \pi_1, \mathsf{Id}_A \circ \pi_2 \big\rangle \big\rangle \\ &= \langle \mathsf{Id}_C \times \mathsf{Id}_A, t \circ \pi_1 \rangle. \end{aligned}$$

Thus,

$$g \circ \langle \mathsf{Id}_C \times \mathsf{Id}_A, t \circ \pi_1 \rangle = g \circ \phi \circ \langle \mathsf{Id}_C, t \rangle \times \mathsf{Id}_A = k \circ \langle \mathsf{Id}_C, t \rangle \times \mathsf{Id}_A.$$

□

Proposition A.3. Let $\mathcal{C}$ be a cartesian closed category and $f : A \times B \to C$ and $g : A' \to A$ be morphisms in $\mathcal{C}$, then

$$\lambda(f) \circ g = \lambda(f \circ (g \times \mathsf{Id}_B))$$

Proof. Note that for any $h : A' \times B \to C$, there is an *unique* morphism $\lambda(h) : A' \to (B \Rightarrow C)$ that satisfies the equation $h = \mathsf{ev} \circ \big(\lambda(h) \times \mathsf{Id}_B\big)$. So, for any $k : A' \to (B \Rightarrow C)$, we have $k = \lambda(\mathsf{ev} \circ (k \times \mathsf{Id}_B))$.

$$\begin{aligned} \lambda(f) \circ g &= \lambda\Big(\mathsf{ev} \circ \big((\lambda(f) \circ g) \times \mathsf{Id}_B\big)\Big) \\ &= \lambda\Big(\mathsf{ev} \circ \big((\lambda(f) \times \mathsf{Id}_B) \circ (g \times \mathsf{Id}_B)\big)\Big) \\ &= \lambda\Big(\big(\mathsf{ev} \circ (\lambda(f) \times \mathsf{Id}_B)\big) \circ (g \times \mathsf{Id}_B)\Big) \\ &= \lambda\Big(f \circ (g \times \mathsf{Id}_B)\Big) \end{aligned}$$

□

Lemma A.4. (Substitution Lemma) Let $\mathbb{M}$ be a structure in a cartesian closed category $\mathcal{C}$. If $\Gamma \cup \{x : \alpha\} \vdash s : \beta$ and $\Gamma \vdash t : \alpha$, then

$$[\![\Gamma \vdash s[t/x] : \beta]\!] = [\![\Gamma \cup \{x : \alpha\} \vdash s : \beta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!] \big\rangle.$$

Proof. Induction on the structure of typed term s.

(var) If $s \equiv x$ is a variable, then $\alpha = \beta$, and

$$\begin{aligned}[\![\Gamma \vdash x[t/x] : \beta]\!] = [\![\Gamma \vdash t : \beta]\!] &= \pi_2 \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \beta]\!]\big\rangle \\ &= [\![\Gamma \cup \{x : \beta\} \vdash x : \beta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \beta]\!]\big\rangle\end{aligned}$$

(abs) If $s \equiv \lambda y.u$ is an abstraction, say $\Gamma \cup \{x : \alpha\} \vdash \lambda y.u : \zeta \Rightarrow \eta$, then

$$\begin{aligned}
&[\![\Gamma \vdash (\lambda y.u)[t/x] : \zeta \Rightarrow \eta]\!] \\
&= [\![\Gamma \vdash \lambda y.u[t/x] : \zeta \Rightarrow \eta]\!] \\
&= \lambda\Big([\![\Gamma \cup \{y : \zeta\} \vdash u[t/x] : \eta]\!]\Big) \\
&= \lambda\Big([\![\Gamma \cup \{y : \zeta\} \cup \{x : \alpha\} \vdash u : \eta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!] \times [\![\zeta]\!]}, [\![\Gamma \cup \{y : \zeta\} \vdash t : \zeta]\!]\big\rangle\Big) \\
&= \lambda\Big([\![\Gamma \cup \{y : \zeta\} \cup \{x : \alpha\} \vdash u : \eta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]} \times \mathsf{Id}_{[\![\zeta]\!]}, [\![\Gamma \vdash t : \alpha]\!] \circ \pi_1\big\rangle\Big) \\
&= \lambda\Big([\![\Gamma \cup \{x : \alpha\} \cup \{y : \zeta\} \vdash u : \eta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle \times \mathsf{Id}_{[\![\zeta]\!]}\Big) \qquad \text{(Prop. A.2)} \\
&= \lambda\Big([\![\Gamma \cup \{x : \alpha\} \cup \{y : \zeta\} \vdash u : \eta]\!]\Big) \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle \qquad \text{(Prop. A.3)} \\
&= [\![\Gamma \cup \{x : \alpha\} \vdash \lambda y.u : \zeta \Rightarrow \eta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle
\end{aligned}$$

(app) If $s \equiv uv$ is an application, we have

$$\begin{aligned}
&[\![\Gamma \vdash uv[t/x] : \beta]\!] \\
&= [\![\Gamma \vdash u[t/x]v[t/x] : \beta]\!] \\
&= \mathsf{ev} \circ \big\langle [\![\Gamma \vdash u[t/x] : \zeta \Rightarrow \beta]\!], [\![\Gamma \vdash v[t/x] : \zeta]\!]\big\rangle \\
&= \mathsf{ev} \circ \Big\langle [\![\Gamma \cup \{x : \alpha\} \vdash u : \zeta \Rightarrow \beta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle, \\
&\quad [\![\Gamma \cup \{x : \alpha\} \vdash v : \zeta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle\Big\rangle \qquad \text{(IH)} \\
&= \mathsf{ev} \circ \big\langle [\![\Gamma \cup \{x : \alpha\} \vdash u : \zeta \Rightarrow \beta]\!], [\![\Gamma \cup \{x : \alpha\} \vdash v : \zeta]\!]\big\rangle \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle \\
&= [\![\Gamma \cup \{x : \alpha\} \vdash u\,v : \beta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle
\end{aligned}$$

(unit) If $s = \langle\rangle$, then $\beta = \mathsf{unit}$, and

$$[\![\Gamma \vdash \langle\rangle[t/x] : \mathsf{unit}]\!] = [\![\Gamma \vdash \langle\rangle : \mathsf{unit}]\!] = \tau_{[\![\Gamma]\!]} : [\![\Gamma]\!] \to \top.$$

Moreover,

$$[\![\Gamma \cup \{x : \alpha\} \vdash \langle\rangle : \mathsf{unit}]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle : [\![\Gamma]\!] \to \top.$$

Since the morphisms from $[\![\Gamma]\!]$ to the terminal object $\top$ are unique, we have

$$[\![\Gamma \vdash \langle\rangle[t/x] : \mathsf{unit}]\!] = [\![\Gamma \cup \{x : \alpha\} \vdash \langle\rangle : \mathsf{unit}]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle.$$

(*pair*) If $s = \langle u, v\rangle$ is a pair, then

$$
\begin{aligned}
&[\![\Gamma \vdash \langle u, v\rangle[t/x] : \zeta \times \eta]\!] \\
&= [\![\Gamma \vdash \langle u[t/x], v[t/x]\rangle : \zeta \times \eta]\!] \\
&= \Big\langle [\![\Gamma \vdash u[t/x] : \zeta]\!], [\![\Gamma \vdash v[t/x] : \eta]\!] \Big\rangle \\
&= \Big\langle [\![\Gamma \cup \{x : \alpha\} \vdash u : \zeta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle, \\
&\quad [\![\Gamma \cup \{x : \alpha\} \vdash v : \eta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle \Big\rangle && \text{(IH)} \\
&= \Big\langle [\![\Gamma \cup \{x : \alpha\} \vdash u : \zeta]\!], [\![\Gamma \cup \{x : \alpha\} \vdash v : \eta]\!] \Big\rangle \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle \\
&= [\![\Gamma \cup \{x : \alpha\} \vdash \langle u, v\rangle : \zeta \times \eta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle
\end{aligned}
$$

(Fst) If $s = \mathsf{Fst}(u)$, then

$$
\begin{aligned}
&[\![\Gamma \vdash \mathsf{Fst}(u)[t/x] : \zeta]\!] \\
&= [\![\Gamma \vdash \mathsf{Fst}\big(u[t/x]\big) : \zeta]\!] \\
&= \pi_1 \circ [\![\Gamma \vdash u[t/x] : \zeta \times \eta]\!] \\
&= \pi_1 \circ \Big([\![\Gamma \cup \{x : \alpha\} \vdash u : \zeta \times \eta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle \Big) && \text{(IH)} \\
&= [\![\Gamma \cup \{x : \alpha\} \vdash \mathsf{Fst}(u) : \zeta \times \eta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!]\big\rangle
\end{aligned}
$$

(Snd) Similar to the case (Fst).

□

Note that the fact that capture-free substitution is interpreted soundly by composition is essential in proving the (β) case for the soundness theorem.

Theorem A.5. (Soundness Theorem) Given a cartesian closed category $\mathcal{C}$, any structure $\mathbb{M}$ in $\mathcal{C}$ is a model of the $\lambda\beta\eta$-theory.

Proof. We show that a structure $\mathbb{M}$ satisfies all rules in $\lambda\beta\eta$ by induction on the rules of $\lambda\beta\eta$. The proofs of the rules ($refl$),(sym) and ($trans$) follows directly from the fact that $=$ is an equivalence relation.

(*app*) By inductive hypothesis, $\mathbb{M}$ satisfies $\Gamma \vdash s = s' : \alpha \Rightarrow \beta$ and $\Gamma \vdash t = t' : \alpha$.

$$
\begin{aligned}
[\![\Gamma \vdash s\,t : \beta]\!] &= \mathsf{ev} \circ \big\langle [\![\Gamma \vdash s : \alpha \Rightarrow \beta]\!], [\![\Gamma \vdash t : \alpha]\!]\big\rangle \\
&= \mathsf{ev} \circ \big\langle [\![\Gamma \vdash s' : \alpha \Rightarrow \beta]\!], [\![\Gamma \vdash t' : \alpha]\!]\big\rangle = [\![\Gamma \vdash s'\,t' : \beta]\!]
\end{aligned}
$$

Thus $\mathbb{M}$ satisfies $\Gamma \vdash s\,t = s'\,t' : \beta$.

(*abs*) By inductive hypothesis, $\mathbb{M}$ satisfies $\Gamma \cup \{x : \alpha\} \vdash s = t : \beta$.

$$
\begin{aligned}
[\![\Gamma \vdash \lambda x.s : \alpha \Rightarrow \beta]\!] &= \lambda\big([\![\Gamma \cup \{x : \alpha\} \vdash s : \beta]\!]\big) \\
&= \lambda\big([\![\Gamma \cup \{x : \alpha\} \vdash t : \beta]\!]\big) \\
&= [\![\Gamma \vdash \lambda x.t : \alpha \Rightarrow \beta]\!]
\end{aligned}
$$

Thus $\mathbb{M}$ satisfies $\Gamma \vdash \lambda x.s = \lambda x.t : \alpha \Rightarrow \beta$.

(β) Recall the (β) rule states that $\Gamma \vdash (\lambda x.s)t = s[t/x] : \beta$, where $\Gamma \cup \{x : \alpha\} \vdash s : \beta$ and $\Gamma \vdash t : \alpha$ are typed terms.

$$\begin{aligned}
[\![\Gamma \vdash (\lambda x.s)t : \beta]\!] &= \mathsf{ev} \circ \big\langle \lambda([\![\Gamma \cup \{x : \alpha\} \vdash s : \beta]\!]), [\![\Gamma \vdash t : \alpha]\!] \big\rangle \\
&= \mathsf{ev} \circ (\lambda([\![\Gamma \cup \{x : \alpha\} \vdash s : \beta]\!]) \times \mathsf{Id}_{[\![\alpha]\!]}) \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!] \big\rangle \\
&= [\![\Gamma \cup \{x : \alpha\} \vdash s : \beta]\!] \circ \big\langle \mathsf{Id}_{[\![\Gamma]\!]}, [\![\Gamma \vdash t : \alpha]\!] \big\rangle \\
&= [\![\Gamma \vdash s\,[t/x] : \beta]\!] && \text{(Sub. Lemma A.4)}
\end{aligned}$$

Thus, $\mathbb{M}$ satisfies $\Gamma \vdash (\lambda x.s)\, t = s\,[t/x] : \beta$.

(η) Assume that $x \notin \mathsf{FV}(s)$, we have

$$\begin{aligned}
[\![\Gamma \vdash \lambda x.s\, x : \alpha \Rightarrow \beta]\!] &= \lambda\Big([\![\Gamma \cup \{x : \alpha\} \vdash s\, x : \beta]\!]\Big) \\
&= \lambda\Big(\mathsf{ev} \circ \big\langle [\![\Gamma \cup \{x : \alpha\} \vdash s : \alpha \Rightarrow \beta]\!], [\![\Gamma \cup \{x : \alpha\} \vdash x : \alpha]\!] \big\rangle\Big) \\
&= \lambda\Big(\mathsf{ev} \circ \big\langle [\![\Gamma \vdash s : \alpha \Rightarrow \beta]\!] \circ \pi_1, \pi_2 \big\rangle\Big) \\
&= \lambda\Big(\mathsf{ev} \circ \big([\![\Gamma \vdash s : \alpha \Rightarrow \beta]\!] \times \mathsf{Id}_{[\![\alpha]\!]}\big)\Big) \\
&= [\![\Gamma \vdash s : \alpha \Rightarrow \beta]\!]
\end{aligned}$$

Thus, $\mathbb{M}$ satisfies $\Gamma \vdash \lambda x.s\, x = s : \alpha \Rightarrow \beta$.

(unit) Consider $[\![\Gamma \vdash s : \mathsf{unit}]\!] = \tau_{[\![\Gamma]\!]} : [\![\Gamma]\!] \to \top$. Since $\top$ is the terminal object, $\tau_{[\![\Gamma]\!]}$ is unique. Therefore, $[\![\Gamma \vdash s : \mathsf{unit}]\!] = [\![\Gamma \vdash \langle\rangle : \mathsf{unit}]\!] = \tau_{[\![\Gamma]\!]}$, and $\mathbb{M}$ satisfies $\Gamma \vdash s = \langle\rangle : \mathsf{unit}$.

($pair$) Let $\Gamma \vdash p : \alpha \times \beta$ be a typed term.

$$\begin{aligned}
[\![\Gamma \vdash \langle \mathsf{Fst}(p), \mathsf{Snd}(p) \rangle : \alpha \times \beta]\!] &= \big\langle [\![\Gamma \vdash \mathsf{Fst}(p) : \alpha]\!], [\![\Gamma \vdash \mathsf{Snd}(p) : \beta]\!] \big\rangle \\
&= \big\langle \pi_1 \circ [\![\Gamma \vdash p : \alpha \times \beta]\!], \pi_2 \circ [\![\Gamma \vdash p : \alpha \times \beta]\!] \big\rangle \\
&= [\![\Gamma \vdash p : \alpha \times \beta]\!]
\end{aligned}$$

(Fst) Let $\Gamma \vdash s : \alpha$ and $\Gamma \vdash t : \beta$.

$$\begin{aligned}
[\![\Gamma \vdash \langle \mathsf{Fst}(\langle s,t \rangle) : \alpha]\!] &= \pi_1 \circ [\![\Gamma \vdash \langle s,t \rangle : \alpha \times \beta]\!] \\
&= \pi_1 \circ \big\langle [\![\Gamma \vdash \langle s \rangle : \alpha]\!], [\![\Gamma \vdash \langle t \rangle : \beta]\!] \big\rangle \\
&= [\![\Gamma \vdash s : \alpha]\!]
\end{aligned}$$

(Snd) Similar to the case (Fst).

□

Knowing that the interpretation actually makes sense, we can define a λ-theory based on any cartesian closed category.

Corollary A.6. Every cartesian closed category $\mathcal{C}$ gives rise to a λ-theory $\mathsf{Th}(\mathcal{C})$.

Proof. Let $\mathbb{M}$ be a structure in $\mathcal{C}$. Define a theory as follows.

$$\mathsf{Th}(\mathcal{C}) := \big\{ \; \Gamma \vdash s = t : \alpha \mid [\![\Gamma \vdash s : \alpha]\!] = [\![\Gamma \vdash t : \alpha]\!] \; \big\}$$

By soundness theorem, $\mathbb{M}$ satisfies all rules in $\lambda\beta\eta$, hence $\lambda\beta\eta$ is included in $\mathsf{Th}(\mathcal{C})$. So, $\mathsf{Th}(\mathcal{C})$ is indeed a λ-theory. □

Soundness tells us that for any cartesian closed category, we can define a λ-theory with respect to it. Now, we consider the converse. Given a λ-theory, we would like to construct a cartesian closed category in which the theory can be modelled soundly in the category. More specifically, we would like to construct the "smallest" such category. We call it the *classifying* category of the theory. The classifying category is the "smallest" in the sense that for any other categories $\mathcal{D}$, in which the theory can be modelled soundly, there is a cartesian closed functor F from the classifying category to $\mathcal{D}$ such that the interpretation of the theory in $\mathcal{D}$ can be expressed as the composition of the interpretation in the classifying category and F.

To establish this, we need to consider the two categories. Given cartesian closed categories $\mathcal{C}$ and $\mathcal{D}$, we define the *category of cartesian closed functors* $\mathsf{CCCat}_{\cong}(\mathcal{C}, \mathcal{D})$ as the category with objects as cartesian closed functors from $\mathcal{C}$ to $\mathcal{D}$, and morphisms as natural isomorphisms. Recall a cartesian closed functor $F : \mathcal{C} \to \mathcal{D}$ is a functor where

- F preserves products via the isomorphism $\Phi_{A,B} := \langle F\pi_1, F\pi_2 \rangle$,
- F preserves exponential via the isomorphism $\Psi_{A,B} := \lambda(F(\mathsf{ev}) \circ \Phi^{-1}_{A \Rightarrow B, A})$.

Let $\mathscr{T}$ be a λ-theory and $\mathcal{C}$ be a cartesian closed category. We define the *category of models* $\mathsf{Mod}_{\cong}(\mathscr{T}, \mathcal{C})$ as the category with objects as models of the λ-theory $\mathscr{T}$ in $\mathcal{C}$, and morphisms as model homomorphisms $h : \mathbb{M} \to \mathbb{N}$ which is given by a collection of isomorphisms $h_\gamma : [\![\gamma]\!]_{\mathbb{M}} \to [\![\gamma]\!]_{\mathbb{N}}$ in $\mathcal{C}$, for each type variable $\gamma \in \mathsf{TV}$, and

$$h_{\alpha \times \beta} := h_\alpha \times h_\beta \qquad \text{and} \qquad h_{\alpha \Rightarrow \beta} := h_\alpha^{-1} \Rightarrow h_\beta := \lambda\big(h_\beta \circ \mathsf{ev} \circ (\mathsf{Id} \times h_\alpha^{-1})\big)$$

Definition A.5 (Classifying Category)**.** Given a λ-theory $\mathscr{T}$, we say a cartesian closed category is *classifying*, denoted as $\mathsf{Cl}(\mathscr{T})$, if there is a "generic" model $\mathbb{G}$ that soundly interpret $\mathscr{T}$ in $\mathsf{Cl}(\mathscr{T})$, and for any cartesian closed category $\mathcal{D}$, there is a natural equivalence

$$\mathsf{CCCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D}) \simeq \mathsf{Mod}_{\cong}(\mathscr{T}, \mathcal{D}).$$

The idea behind this definition is that for any cartesian closed category $\mathcal{D}$, a model of $\mathscr{T}$ in $\mathcal{D}$ can be represented by a functor from $\mathsf{Cl}(\mathscr{T})$ to $\mathcal{D}$. It is not difficult to see that for any given λ-theory, its classifying category and "generic" model are unique up to isomorphism.

We shall first set up the equivalence in the forward direction. Note that this direction is well-defined for any cartesian closed category $\mathcal{C}$.

Definition A.6 (Modelling functors)**.** Let $\mathcal{C}$ and $\mathcal{D}$ be cartesian closed categories, $\mathscr{T}$ be a λ-theory and $\mathbb{M}$ be a model of $\mathscr{T}$ in $\mathcal{C}$. We define a *family of modelling functors* $\mathsf{Ap}_{\mathbb{M}} : \mathsf{CCCat}_{\cong}(\mathcal{C}, \mathcal{D}) \to \mathsf{Mod}_{\cong}(\mathscr{T}, \mathcal{D})$ by defining, for any cartesian closed functor $F : \mathcal{C} \to \mathcal{D}$, a model $\mathsf{Ap}_{\mathbb{M}}F$ of $\mathscr{T}$ in $\mathcal{D}$, where

$$[\![\gamma]\!]_{\mathsf{Ap}_{\mathbb{M}}F} := F([\![\gamma]\!]_{\mathbb{M}}),$$

and for any natural isomorphism $\phi : F \to G$, a model homomorphism $\mathsf{Ap}_{\mathbb{M}}\phi : \mathsf{Ap}_{\mathbb{M}}F \to \mathsf{Ap}_{\mathbb{M}}G$ where

$$\big(\mathsf{Ap}_{\mathbb{M}}\phi\big)_\gamma := \phi_{[\![\gamma]\!]_{\mathbb{M}}}.$$

Remark. It is easy to check that $\mathsf{Ap}_{\mathbb{M}}F$ is indeed a model of $\mathscr{T}$ in $\mathcal{D}$, $\mathsf{Ap}_{\mathbb{M}}\phi$ is indeed a model homomorphism and $\mathsf{Ap}_{\mathbb{M}}$ is a well-defined functor from $\mathsf{CCCat}_{\cong}(\mathcal{C}, \mathcal{D})$ to $\mathsf{Mod}_{\cong}(\mathscr{T}, \mathcal{D})$. We call this $\mathsf{Ap}_{\mathbb{M}}$ a collection of modelling functors since it specifies a family of models of $\mathscr{T}$ in $\mathcal{D}$ such that there is some cartesian closed functor $F : \mathcal{C} \to \mathcal{D}$ that maps to this model via $\mathsf{Ap}_{\mathbb{M}}$.

Now, we prove the completeness theorem by first constructing a cartesian closed category based on the syntax of the given λ-theory $\mathscr{T}$ and then proving that it is classifying by finding the "inverse" for the functor $\mathsf{Ap}_{\mathbb{G}} : \mathsf{CCCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D}) \to \mathsf{Mod}_{\cong}(\mathscr{T}, \mathcal{D})$.

Theorem A.7. (Completeness Theorem) Given a λ-theory $\mathscr{T}$, we can construct a classifying category $\mathsf{Cl}(\mathscr{T})$.

Proof. Define the classifying category $\mathsf{Cl}(\mathscr{T})$ as follows:

- objects are types of $\mathscr{T}$,
- morphisms $\boldsymbol{f} : \alpha \to \beta$ are equivalence classes of typed terms $\big[\{x : \alpha\} \vdash M : \beta\big]$, where two typed terms are equivalent if they are provably equal in $\mathscr{T}$. We write $\{x : \alpha\} \vdash M : \beta$ instead of $\big[\{x : \alpha\} \vdash M : \beta\big]$,
- composition of $\boldsymbol{g} = \{y : \beta\} \vdash N : \gamma$ and $\boldsymbol{f} = \{x : \alpha\} \vdash M : \beta$ is given by
$$\boldsymbol{f} \circ \boldsymbol{g} = \{x : \alpha\} \vdash N[M/y] : \gamma,$$
- the identity morphism of the object α is $\mathbf{Id}_{\boldsymbol{\alpha}} := \{x : \alpha\} \vdash x : \alpha$,
- the product of objects α and β is $\alpha \times \beta$ with projections
$$\boldsymbol{\pi_1} = \{z : \alpha \times \beta\} \vdash \mathsf{Fst}(z) : \alpha \qquad \text{and} \qquad \boldsymbol{\pi_2} = \{z : \alpha \times \beta\} \vdash \mathsf{Snd}(z) : \beta,$$
- the pairing of morphisms $\boldsymbol{f} = \{x : \gamma\} \vdash M : \alpha$ and $\boldsymbol{g} = \{x : \gamma\} \vdash N : \beta$ is
$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \{x : \gamma\} \vdash \langle M, N \rangle : \alpha \times \beta$$
- the exponential of objects β and γ is $\beta \Rightarrow \gamma$ and the evaluating morphism is
$$\mathbf{ev} = \{z : (\beta \Rightarrow \gamma) \times \beta\} \vdash \mathsf{Fst}(z)\ \mathsf{Snd}(z) : \gamma,$$
where for any morphism $\boldsymbol{f} = \{z : \alpha \times \beta\} \vdash M : \gamma$, the exponential mate of $\boldsymbol{f}$ is
$$\boldsymbol{\lambda(f)} = \{x : \alpha\} \vdash \lambda y.(M[\langle x, y \rangle / z]) : \beta \Rightarrow \gamma,$$
where x and y are distinct and fresh variables.

We first need to check that $\mathsf{Cl}(\mathscr{T})$ is indeed a cartesian closed category. Consider morphisms $\alpha \xleftarrow{\boldsymbol{f}} \zeta \xrightarrow{\boldsymbol{g}} \beta$ where $\boldsymbol{f} = \{x : \zeta\} \vdash M : \alpha$ and $\boldsymbol{g} = \{x : \zeta\} \vdash N : \beta$. Then

$$\begin{aligned}
\pi_1 \circ \langle \boldsymbol{f}, \boldsymbol{g} \rangle &= \Big(\{z : \alpha \times \beta\} \vdash \mathsf{Fst}(z) : \alpha\Big) \circ \Big(\{x : \zeta\} \vdash \langle M, N \rangle : \alpha \times \beta\Big) \\
&= \{x : \zeta\} \vdash \mathsf{Fst}(z)[\langle M, N \rangle / z] : \alpha \\
&= \{x : \zeta\} \vdash \mathsf{Fst}(\langle M, N \rangle) : \alpha \\
&= \{x : \zeta\} \vdash M : \alpha = \boldsymbol{f}
\end{aligned}$$

and similarly, $\pi_2 \circ \langle \boldsymbol{f}, \boldsymbol{g} \rangle = \boldsymbol{g}$. To prove that the pairing is indeed unique, it is enough to show that for any $\boldsymbol{h} = \{x : \zeta\} \vdash p : \alpha \times \beta$, we have $\boldsymbol{h} = \langle \pi_1 \circ \boldsymbol{h}, \pi_2 \circ \boldsymbol{h} \rangle$.

$$\begin{aligned}
&\langle \pi_1 \circ \boldsymbol{h}, \pi_2 \circ \boldsymbol{h} \rangle \\
&= \Big\langle \big(\{z : \alpha \times \beta\} \vdash \mathsf{Fst}(z) : \alpha\big) \circ \big(\{x : \zeta\} \vdash p : \alpha \times \beta\big), \\
&\qquad \big(\{z : \alpha \times \beta\} \vdash \mathsf{Snd}(z) : \alpha\big) \circ \big(\{x : \zeta\} \vdash p : \alpha \times \beta\big) \Big\rangle \\
&= \big\langle \{x : \zeta\} \vdash \mathsf{Fst}(p) : \alpha, \{x : \zeta\} \vdash \mathsf{Snd}(p) : \beta \big\rangle \\
&= \{x : \zeta\} \vdash \big\langle \mathsf{Fst}(p), \mathsf{Snd}(p) \big\rangle : \alpha \times \beta \\
&= \{x : \zeta\} \vdash p : \alpha \times \beta = \boldsymbol{h}
\end{aligned}$$

The terminal object of $\mathsf{Cl}(\mathscr{T})$ is the type unit. Indeed, for any type α, there exists a unique morphism

$$\boldsymbol{\tau}_\alpha = \{x : \alpha\} \vdash \langle\rangle : \mathsf{unit}.$$

Now we prove that $\mathsf{Cl}(\mathscr{T})$ is indeed cartesian closed by showing that it has exponentials. For any types β and γ, there is a type $\beta \Rightarrow \gamma$ and morphism $\mathbf{ev} := \{z : (\beta \Rightarrow \gamma) \times \beta\} \vdash \mathsf{Fst}(z)\ \mathsf{Snd}(z) : \gamma$ where for any $\boldsymbol{f} = \{z : \alpha \times \beta\} \vdash M : \gamma$, there is a morphism $\boldsymbol{\lambda}(\boldsymbol{f}) = \{x : \alpha\} \vdash \lambda y.\big(M[\langle x, y\rangle]/z\big) : \beta \Rightarrow \gamma$ and

$$\begin{aligned}
&\mathbf{ev} \circ \big(\boldsymbol{\lambda}(\boldsymbol{f}) \times \mathbf{Id}_{\boldsymbol{\beta}}\big)\\
&= \mathbf{ev} \circ \big\langle \boldsymbol{\lambda}(\boldsymbol{f}) \circ \boldsymbol{\pi_1}, \mathbf{Id}_{\boldsymbol{\beta}} \circ \boldsymbol{\pi_2}\big\rangle\\
&= \mathbf{ev} \circ \Big\langle \{w : \alpha \times \beta\} \vdash \big(\lambda y.(M[\langle x, y\rangle/z])\big)[\mathsf{Fst}(w)/x] : \beta \Rightarrow \gamma, \boldsymbol{\pi_2}\Big\rangle\\
&= \mathbf{ev} \circ \Big\langle \{w : \alpha \times \beta\} \vdash \lambda y.(M[\langle x, y\rangle/z][\mathsf{Fst}(w)/x]) : \beta \Rightarrow \gamma, \{w : \alpha \times \beta\} \vdash \mathsf{Snd}(w) : \beta\Big\rangle\\
&= \mathbf{ev} \circ \Big(\{w : \alpha \times \beta\} \vdash \big\langle \lambda y.(M[\langle x, y\rangle/z][\mathsf{Fst}(w)/x]), \mathsf{Snd}(w)\big\rangle : (\beta \Rightarrow \gamma) \times \beta\Big)\\
&= \{w : \alpha \times \beta\} \vdash \big(\lambda y.(M[\langle x, y\rangle/z][\mathsf{Fst}(w)/x])\big)\ \mathsf{Snd}(w) : \gamma\\
&= \{w : \alpha \times \beta\} \vdash M[\langle x, y\rangle/z][\mathsf{Fst}(w)/x][\mathsf{Snd}(w)/y] : \gamma\\
&= \{w : \alpha \times \beta\} \vdash M[\langle \mathsf{Fst}(w), \mathsf{Snd}(w)\rangle/z] : \gamma\\
&= \{w : \alpha \times \beta\} \vdash M[w/z] : \gamma\\
&= \boldsymbol{f}.
\end{aligned}$$

Finally, to prove that the exponential is unique, we show that for any $\boldsymbol{h} = \{w : \alpha\} \vdash H : \beta \Rightarrow \gamma$,

$$\begin{aligned}
\boldsymbol{\lambda}(\mathbf{ev}) \circ (\boldsymbol{h} \times \mathbf{Id}_{\boldsymbol{\beta}}) &= \boldsymbol{\lambda}(\mathbf{ev}) \circ \langle \boldsymbol{h} \circ \boldsymbol{\pi_1}, \mathbf{Id}_{\boldsymbol{\beta}} \circ \boldsymbol{\pi_2}\rangle\\
&= \boldsymbol{\lambda}\big(\mathbf{ev} \circ \{z : \alpha \times \beta\} \vdash \langle H[\mathsf{Fst}(z)/w], \mathsf{Snd}(z)\rangle : (\beta \Rightarrow \gamma) \times \beta\big)\\
&= \boldsymbol{\lambda}\big(\{z : \alpha \times \beta\} \vdash H[\mathsf{Fst}(z)/w]\ \mathsf{Snd}(z) : \gamma\big)\\
&= \{x : \alpha\} \vdash \lambda y.\Big(\big(H[\mathsf{Fst}(z)/w]\ \mathsf{Snd}(z)\big)[\langle x, y\rangle/z]\Big) : \beta \Rightarrow \gamma\\
&= \{x : \alpha\} \vdash \lambda y.\Big(\big(H[x/w]\ y\big)\Big) : \beta \Rightarrow \gamma\\
&= \{x : \alpha\} \vdash H[x/w] : \beta \Rightarrow \gamma\\
&= \boldsymbol{h}
\end{aligned}$$

Now we show that $\mathsf{Cl}(\mathscr{T})$ is classifying. We define the "generic" model $\mathbb{G}$ of $\mathscr{T}$ in $\mathsf{Cl}(\mathscr{T})$ by setting

$$[\![\gamma]\!]_{\mathbb{G}} := \gamma$$

for all type variables γ, thus $[\![\alpha]\!] = \alpha$ for all types α. Let $\mathcal{D}$ be a cartesian closed category. To establish the equivalence $\mathsf{CCCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D}) \simeq \mathsf{Mod}_{\cong}(\mathscr{T}, \mathcal{D})$ we define a functor $\mathsf{Ap}_{\mathbb{G}}^{-1} : \mathsf{Mod}_{\cong}(\mathscr{T}, \mathcal{D}) \to \mathsf{CCCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D})$, and prove that it is the "inverse" of the modelling functors.

For any model $\mathbb{M}$ of $\mathscr{T}$ in $\mathcal{D}$, we define the functor

$$\begin{aligned}
\mathsf{Ap}_{\mathbb{G}}^{-1}(\mathbb{M}) : \qquad \mathsf{Cl}(\mathscr{T}) &\longrightarrow \mathcal{D}\\
\alpha &\longmapsto [\![\alpha]\!]_{\mathbb{M}}\\
\{x : \alpha\} \vdash M : \beta &\longmapsto [\![\{x : \alpha\} \vdash M : \beta]\!]_{\mathbb{M}} : [\![\alpha]\!]_{\mathbb{M}} \to [\![\beta]\!]_{\mathbb{M}}.
\end{aligned}$$

Soundness tells us that $\mathsf{Ap}_{\mathbb{G}}^{-1}(\mathbb{M})$ is well-defined and it is easy to check that it is a cartesian closed functor. For any model homomorphism $h : \mathbb{M} \to \mathbb{N}$ in $\mathsf{Mod}_{\cong}(\mathscr{T},\mathcal{D})$, we define the natural isomorphism $\mathsf{Ap}_{\mathbb{G}}^{-1}(h) : \mathsf{Ap}_{\mathbb{G}}^{-1}(\mathbb{M}) \to \mathsf{Ap}_{\mathbb{G}}^{-1}(\mathbb{N})$ by setting

$$\big(\mathsf{Ap}_{\mathbb{G}}^{-1}(h)\big)_\alpha := h_\alpha : [\![\alpha]\!]_{\mathbb{M}} \to [\![\alpha]\!]_{\mathbb{N}}.$$

We first show that $\big(\mathsf{Ap}_{\mathbb{G}}^{-1}(h)\big)_\alpha : [\![\alpha]\!]_{\mathbb{M}} \to [\![\alpha]\!]_{\mathbb{N}}$ is a natural transformation. For any morphism $\Gamma \vdash M : \beta$ in $\mathsf{Cl}(\mathscr{T})$, we prove by induction on the derivation of the typed terms $\Gamma \vdash M : \beta$ that the following diagram commutes.

$$\begin{array}{ccc} [\![\Gamma]\!]_{\mathbb{M}} & \xrightarrow{h_\Gamma} & [\![\Gamma]\!]_{\mathbb{N}} \\ {\scriptstyle [\![\Gamma\vdash M:\beta]\!]_{\mathbb{M}}}\downarrow & & \downarrow{\scriptstyle [\![\Gamma\vdash M:\beta]\!]_{\mathbb{N}}} \\ [\![\beta]\!]_{\mathbb{M}} & \xrightarrow{h_\beta} & [\![\beta]\!]_{\mathbb{N}} \end{array}$$

(var) $[\![\Gamma\cup\{x:\alpha\}\vdash x:\alpha]\!]_{\mathbb{N}}\circ h_{\Gamma\times\alpha} = \boldsymbol{\pi_2}\circ(h_\Gamma\times h_\alpha) = h_\alpha\circ\boldsymbol{\pi_2} = h_\alpha\circ[\![\Gamma\cup\{x:\alpha\}\vdash x:\alpha]\!]_{\mathbb{M}}$

(abs) Let $\boldsymbol{n} = [\![\Gamma\cup\{x:\alpha\}\vdash s:\beta]\!]_{\mathbb{N}}$ and $\boldsymbol{m} = [\![\Gamma\cup\{x:\alpha\}\vdash s:\beta]\!]_{\mathbb{M}}$. By induction hypothesis, we know that $\boldsymbol{n}\circ h_{\Gamma\times\alpha} = h_\beta\circ\boldsymbol{m}$, and this is equivalent to $\boldsymbol{n}\circ(h_\Gamma\times\mathbf{Id}) = h_\beta\circ\boldsymbol{m}\circ(\mathbf{Id}\times h_\alpha^{-1})$. We want to prove that

$$\begin{array}{rrcl} & [\![\Gamma\vdash\lambda x.s:\alpha\Rightarrow\beta]\!]_{\mathbb{N}}\circ h_\Gamma & = & h_{\alpha\Rightarrow\beta}\circ[\![\Gamma\vdash\lambda x.s:\alpha\Rightarrow\beta]\!]_{\mathbb{M}} \\ \Longleftrightarrow & \boldsymbol{\lambda}(\boldsymbol{n})\circ h_\Gamma & = & (h_\alpha^{-1}\Rightarrow h_\beta)\circ\boldsymbol{\lambda}(\boldsymbol{m}) \\ & & = & \boldsymbol{\lambda}\big(h_\beta\circ\mathbf{ev}\circ(\mathbf{Id}\times h_\alpha^{-1})\big)\circ\boldsymbol{\lambda}(\boldsymbol{m}). \end{array}$$

We would prove it by showing that $\boldsymbol{\lambda}(\boldsymbol{n})\circ h_\Gamma$ and $\boldsymbol{\lambda}\big(h_\beta\circ\mathbf{ev}\circ(\mathbf{Id}\times h_\alpha^{-1})\big)\circ\boldsymbol{\lambda}(\boldsymbol{m})$ have the same exponential mate.

$$\mathbf{ev}\circ\big(\boldsymbol{\lambda}(\boldsymbol{n})\circ h_\Gamma\times\mathbf{Id}\big) = \mathbf{ev}\circ\big(\boldsymbol{\lambda}(\boldsymbol{n})\times\mathbf{Id}\big)\circ\big(h_\Gamma\times\mathbf{Id}\big) = \boldsymbol{n}\circ\big(h_\Gamma\times\mathbf{Id}\big),$$

and

$$\begin{aligned} &\mathbf{ev}\circ\Big(\Big(\boldsymbol{\lambda}\big(h_\beta\circ\mathbf{ev}\circ(\mathbf{Id}\times h_\alpha^{-1})\big)\circ\boldsymbol{\lambda}(\boldsymbol{m})\Big)\times\mathbf{Id}\Big) \\ &= \mathbf{ev}\circ\Big(\boldsymbol{\lambda}\big(h_\beta\circ\mathbf{ev}\circ(\mathbf{Id}\times h_\alpha^{-1})\big)\times\mathbf{Id}\Big)\circ\big(\boldsymbol{\lambda}(\boldsymbol{m})\times\mathbf{Id}\big) \\ &= h_\beta\circ\mathbf{ev}\circ(\mathbf{Id}\times h_\alpha^{-1})\circ\big(\boldsymbol{\lambda}(\boldsymbol{m})\times\mathbf{Id}\big) \\ &= h_\beta\circ\boldsymbol{m}\circ(\mathbf{Id}\times h_\alpha^{-1}) \end{aligned}$$

By the induction hypothesis, they are equal.

(app) Let $\boldsymbol{n} = [\![\Gamma\vdash s:\alpha\Rightarrow\beta]\!]_{\mathbb{N}}$ and $\boldsymbol{m} = [\![\Gamma\vdash s:\alpha\Rightarrow\beta]\!]_{\mathbb{M}}$, and $\boldsymbol{n'} = [\![\Gamma\vdash t:\alpha]\!]_{\mathbb{N}}$ and $\boldsymbol{m'} = [\![\Gamma\vdash t:\alpha]\!]_{\mathbb{M}}$. So, by induction hypothesis, $\boldsymbol{n}\circ h_\Gamma = h_{\alpha\Rightarrow\beta}\circ\boldsymbol{m}$ and

$\boldsymbol{n'} \circ h_\Gamma = h_\alpha \circ \boldsymbol{m'}$.

$$\begin{aligned}
&[\![\Gamma \vdash st : \beta]\!]_{\mathbb{N}} \\
&= \mathbf{ev} \circ \langle \boldsymbol{n}, \boldsymbol{n'} \rangle \circ h_\Gamma \\
&= \mathbf{ev} \circ \langle \boldsymbol{n} \circ h_\Gamma, \boldsymbol{n'} \circ h_\Gamma \rangle \\
&= \mathbf{ev} \circ \langle h_{\alpha \Rightarrow \beta} \circ \boldsymbol{m}, h_\alpha \circ \boldsymbol{m'} \rangle \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(h_\beta \circ \mathbf{ev} \circ (\mathbf{Id} \times h_\alpha^{-1})\big) \circ \boldsymbol{m}, h_\alpha \circ \boldsymbol{m'} \Big\rangle \\
&= \mathbf{ev} \circ \Big\langle \boldsymbol{\lambda}\big(h_\beta \circ \mathbf{ev} \circ (\mathbf{Id} \times h_\alpha^{-1}) \circ (\boldsymbol{m} \times \mathbf{Id})\big), h_\alpha \circ \boldsymbol{m'} \Big\rangle && \text{(prop A.3)} \\
&= \mathbf{ev} \circ \Big(\boldsymbol{\lambda}\big(h_\beta \circ \mathbf{ev} \circ (\mathbf{Id} \times h_\alpha^{-1}) \circ (\boldsymbol{m} \times \mathbf{Id})\big) \times \mathbf{Id}\Big) \circ \langle \mathbf{Id}, h_\alpha \circ \boldsymbol{m'} \rangle \\
&= h_\beta \circ \mathbf{ev} \circ (\mathbf{Id} \times h_\alpha^{-1}) \circ (\boldsymbol{m} \times \mathbf{Id}) \circ \langle \mathbf{Id}, h_\alpha \circ \boldsymbol{m'} \rangle \\
&= h_\beta \circ \mathbf{ev} \circ (\boldsymbol{m} \times h_\alpha^{-1}) \circ \langle \mathbf{Id}, h_\alpha \circ \boldsymbol{m'} \rangle \\
&= h_\beta \circ \mathbf{ev} \circ \langle \boldsymbol{m} \circ \mathbf{Id}, h_\alpha^{-1} \circ h_\alpha \circ \boldsymbol{m'} \rangle \\
&= h_\beta \circ \mathbf{ev} \circ \langle \boldsymbol{m}, \boldsymbol{m'} \rangle
\end{aligned}$$

(unit) $[\![\Gamma \vdash \langle\rangle : \mathsf{unit}]\!]_{\mathbb{N}} \circ h_\Gamma = \boldsymbol{\tau}_{[\![\Gamma]\!]_{\mathbb{N}}} \circ h_\Gamma : [\![\Gamma]\!]_{\mathbb{M}} \to \mathsf{unit}$ must be equals to $\boldsymbol{\tau}_{[\![\Gamma]\!]_{\mathbb{M}}} : [\![\Gamma]\!]_{\mathbb{M}} \to \mathsf{unit}$ since unit is the terminal object. Note that $h_{\mathsf{unit}} \circ [\![\Gamma \vdash \langle\rangle : \mathsf{unit}]\!]_{\mathbb{M}} = \boldsymbol{\tau}_{[\![\Gamma]\!]_{\mathbb{M}}}$. Thus, $[\![\Gamma \vdash \langle\rangle : \mathsf{unit}]\!]_{\mathbb{N}} \circ h_\Gamma$ and $h_{\mathsf{unit}} \circ [\![\Gamma \vdash \langle\rangle : \mathsf{unit}]\!]_{\mathbb{M}}$ are the same.

(*pair*) Let $\boldsymbol{n} = [\![\Gamma \vdash s : \alpha]\!]_{\mathbb{N}}$ and $\boldsymbol{m} = [\![\Gamma \vdash s : \alpha]\!]_{\mathbb{M}}$, and $\boldsymbol{n'} = [\![\Gamma \vdash t : \beta]\!]_{\mathbb{N}}$ and $\boldsymbol{m'} = [\![\Gamma \vdash t : \beta]\!]_{\mathbb{M}}$. By inductive hypothesis, $\boldsymbol{n} \circ h_\Gamma = h_\alpha \circ \boldsymbol{m}$ and $\boldsymbol{n'} \circ h_\Gamma = h_\beta \circ \boldsymbol{m'}$.

$$\begin{aligned}
[\![\Gamma \vdash \langle s,t \rangle : \alpha \times \beta]\!]_{\mathbb{N}} \circ h_\Gamma &= \langle \boldsymbol{n}, \boldsymbol{n'} \rangle \circ h_\Gamma \\
&= \langle n \circ h_\Gamma, \boldsymbol{n'} \circ h_\Gamma \rangle \\
&= \langle h_\alpha \circ \boldsymbol{m'}, h_\beta \circ \boldsymbol{m'} \rangle \\
&= (h_\alpha \times h_\beta) \circ \langle \boldsymbol{m}, \boldsymbol{m'} \rangle \\
&= h_{\alpha \times \beta} \circ \langle \boldsymbol{m}, \boldsymbol{m'} \rangle
\end{aligned}$$

(Fst) Let $\boldsymbol{n} = [\![\Gamma \vdash p : \alpha \times \beta]\!]_{\mathbb{N}}$ and $\boldsymbol{m} = [\![\Gamma \vdash p : \alpha \times \beta]\!]_{\mathbb{M}}$. Then, by inductive hypothesis, $\boldsymbol{n} \circ h_\Gamma = h_{\alpha \times \beta} \circ \boldsymbol{m}$.

$$\begin{aligned}
[\![\Gamma \vdash \mathsf{Fst}(p) : \alpha]\!]_{\mathbb{N}} \circ h_\Gamma &= \boldsymbol{\pi_1} \circ \boldsymbol{n} \circ h_\Gamma \\
&= \boldsymbol{\pi_1} \circ h_{\alpha \times \beta} \circ \boldsymbol{m} \\
&= \boldsymbol{\pi_1} \circ (h_\alpha \times h_\beta) \circ \boldsymbol{m} \\
&= h_\alpha \circ \boldsymbol{\pi_1} \circ \boldsymbol{m} \\
&= h_\alpha \circ [\![\Gamma \vdash \mathsf{Fst}(p) : \alpha]\!]_{\mathbb{M}}
\end{aligned}$$

(Snd) This is similar to the case (Fst).

Thus, $\mathsf{Ap}_{\mathbb{G}}^{-1}(h)$ is indeed a natural transformation. Since h is a model homomorphism, by definition, $\big(\mathsf{Ap}_{\mathbb{G}}^{-1}(h)\big)_\alpha$ is an isomorphism. So, $\mathsf{Ap}_{\mathbb{G}}^{-1}(h)$ is a natural isomorphism.

The final thing we need to check is that $\mathsf{Ap}_{\mathbb{G}}$ and $\mathsf{Ap}_{\mathbb{G}}^{-1}$ are indeed equivalence. We define natural isomorphisms

$$\mu : \mathsf{Ap}_{\mathbb{G}}\mathsf{Ap}_{\mathbb{G}}^{-1} \cong \mathsf{Id}_{\mathsf{Mod}_{\cong}(\mathscr{T},\mathcal{D})} \qquad \text{and} \qquad \nu : \mathsf{Id}_{\mathsf{CCCat}_{\cong}(\mathsf{Cl}(\mathscr{T}),\mathcal{D})} \cong \mathsf{Ap}_{\mathbb{G}}^{-1}\mathsf{Ap}_{\mathbb{G}}$$

such that for any model $\mathbb{M}$ of $\mathscr{T}$ in $\mathcal{D}$, $\mu_{\mathbb{M}} : \mathsf{Ap}_{\mathbb{G}}\mathsf{Ap}_{\mathbb{G}}^{-1}\mathbb{M} \to \mathbb{M}$ is defined as

$$(\mu_{\mathbb{M}})_\gamma := \mathsf{Id}_{[\![\gamma]\!]_{\mathbb{M}}} : [\![\gamma]\!]_{\mathsf{Ap}_{\mathbb{G}}\mathsf{Ap}_{\mathbb{G}}^{-1}\mathbb{M}} = [\![\gamma]\!]_{\mathbb{M}} \to [\![\gamma]\!]_{\mathbb{M}}$$

and for any cartesian closed functor $F : \mathsf{Cl}(\mathscr{T}) \to \mathcal{D}$, we define

$$(\nu_F)_\alpha := \mathsf{Id}_{F\alpha} : F\alpha \to \big(\mathsf{Ap}_{\mathbb{G}}^{-1}(\mathsf{Ap}_{\mathbb{G}}F)\big)\alpha = F([\![\alpha]\!]_{\mathbb{G}}) = F\alpha.$$

Obviously, μ and ν are natural isomorphisms. Thus,

$$\mathsf{Ap}_{\mathbb{G}} : \mathsf{CCCat}_{\cong}(\mathsf{Cl}(\mathscr{T}), \mathcal{D}) \simeq \mathsf{Mod}_{\cong}(\mathscr{T}, \mathcal{D}) : \mathsf{Ap}_{\mathbb{G}}^{-1}$$

and $\mathsf{Cl}(\mathscr{T})$ is indeed a classifying category with the "generic" model $\mathbb{G}$. □

B Proofs

Lemma 4.1. Let $M, N \in \Lambda^r$ and x be a variable.

(i) $(M[N/x])^d \equiv M^d[N^d/x]$

(ii) $(M\langle N/x\rangle)^d \equiv \dfrac{\partial M^d}{\partial x} \cdot N^d$

Proof. (i) Straightforward induction on the structure of M.

(ii) Prove by induction on the structure of M. The only interesting case is when M is an application.

$$
\begin{aligned}
&\Big(M[\vec{L}, \vec{N}^!]\langle N'/x\rangle\Big)^d \\
\equiv\ &\Big(M\langle N'/x\rangle\ [\vec{L}, \vec{N}^!]\Big)^d + \Big(M\ \big([\vec{L}, \vec{N}^!]\langle N'/x\rangle\big)\Big)^d \\
\equiv\ &\Big(\mathsf{D}^k\big(M\langle N'/x\rangle\big)^d \cdot \vec{L}^d\Big)\ \sum N_i^d + \Big(\sum_j M\ \big[L_j\langle N'/x\rangle, \vec{L}_{-j}, \vec{N}^!\big]\Big)^d + \\
&\Big(\sum_i M\ \big[N_i\langle N'/x\rangle, \vec{L}, \vec{N}^!\big]\Big)^d \\
\equiv\ &\Big(\mathsf{D}^k\big(M\langle N'/x\rangle\big)^d \cdot \vec{L}^d\Big)\ \sum N_i^d + \sum_j \Big(\mathsf{D}^k M^d \cdot \big((L_j\langle N'/x\rangle)^d, \vec{L}^d_{-j}\big)\Big)\ \sum N_i^d + \\
&\sum_i \Big(\mathsf{D}^{k+1} M^d \cdot \big((N_i\langle N'/x\rangle)^d, \vec{L}^d\big)\Big)\ \sum N_i^d \\
\equiv\ &\Big(\mathsf{D}^k\Big(\frac{\partial M^d}{\partial x} \cdot N'^d\Big) \cdot \vec{L}^d\Big)\ \sum N_i^d + \sum_j \Big(\mathsf{D}^k M^d \cdot \Big(\frac{\partial L_j^d}{\partial x} \cdot N'^d, \vec{L}^d_{-j}\Big)\Big)\ \sum N_i^d \\
&\sum_i \Big(\mathsf{D}^{k+1} M^d \cdot \Big(\frac{\partial N_i^d}{\partial x} \cdot N'^d, \vec{L}^d\Big)\Big)\ \sum N_i^d \qquad \text{(IH)} \\
\equiv\ &\Big(\mathsf{D}^k\Big(\frac{\partial M^d}{\partial x} \cdot N'^d\Big) \cdot \vec{L}^d + \sum_j \mathsf{D}^k M^d \cdot \Big(\frac{\partial L_j^d}{\partial x} \cdot N'^d, \vec{L}^d_{-j}\Big)\Big)\ \sum N_i^d + \\
&\sum_i \Big(\mathsf{D}\big(\mathsf{D}^k M^d \cdot \vec{L}^d\big) \cdot \Big(\frac{\partial N_i^d}{\partial x} \cdot N'^d\Big)\Big)\ \sum N_i^d \\
\equiv\ &\Big(\frac{\partial}{\partial x}\big(\mathsf{D}^k M^d \cdot \vec{L}^d\big) \cdot N'^d\Big)\ \sum N_i^d + \Big(\mathsf{D}\big(\mathsf{D}^k M^d \cdot \vec{L}^d\big) \cdot \Big(\frac{\partial \sum N_i^d}{\partial x} \cdot N'^d\Big)\Big)\ \sum N_i^d \\
\equiv\ &\frac{\partial}{\partial x}\Big(\big(\mathsf{D}^k M^d \cdot \vec{L}^d\big)\ \sum N_i^d\Big) \cdot N'^d \\
\equiv\ &\frac{\partial}{\partial x}\big(M[\vec{L}, \vec{N}^!]\big)^d \cdot N'^d
\end{aligned}
$$

□

Proposition 4.2. For any $M, N \in \Lambda^r$, we have

(i) $\Gamma \vdash_{\mathcal{R}} M : \alpha \iff \Gamma \vdash_{\mathcal{D}} M^d : \alpha$,

(ii) If M and N are provably equal in the theory $\lambda\beta\eta_{\mathsf{R}}$, then their translations are also provably equal in $\lambda\beta\eta_{\mathsf{D}}$. i.e.

$$\lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} M = N : \alpha \implies \lambda\beta\eta_{\mathsf{D}} \rhd \Gamma \vdash_{\mathcal{D}} M^d = N^d : \alpha.$$

Proof. (i) Easy induction on the length of the proofs.

(ii) Induction on all the rules in the theory $\lambda\beta\eta_{\mathsf{R}}$. The only interesting rules to consider is the (β_{R}) rule.

Consider $\lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} (\lambda x.M)[\vec{L}, \vec{N}^!] = M\langle L/x\rangle\big[\sum N_i/x\big] : \alpha$. By part (i), we know that $\big((\lambda x.M)[\vec{L}, \vec{N}^!]\big)^d$ and $\big(M\langle L/x\rangle\big[\sum N_i/x\big]\big)^d$ are both differential typed terms.

$$\begin{aligned}
\lambda\beta\eta_{\mathsf{D}} \rhd \big((\lambda x.M)[\vec{L}, \vec{N}^!]\big)^d &\equiv \big(\mathsf{D}^k(\lambda x.M^d) \cdot \vec{L}^d\big) \sum_i N_i^d \\
&= \left(\lambda x.\Big(\frac{\partial^k M^d}{\partial x^k} \cdot \vec{L}^d\Big)\right) \sum_i N_i^d && (\beta_{\mathsf{D}}) \\
&= \Big(\frac{\partial^k M^d}{\partial x^k} \cdot \vec{L}^d\Big) \Big[\sum_i N_i^d/x\Big] && (\beta) \\
&\equiv \Big(M\langle L/x\rangle\big[\sum N_i/x\big]\Big)^d : \alpha && \text{(Lemma 4.1)}
\end{aligned}$$

□

Lemma 4.3. Let $S, T \in \Lambda^d$ and x a variable.

(i) $\lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} \big(S[T/x]\big)^r = S^r[T^r/x] : \alpha$

(ii) $\lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} \Big(\dfrac{\partial S}{\partial x} \cdot T\Big)^r = S^r\langle T^r/x\rangle : \alpha$

Proof. (i) Straightforward induction on the structure of S.

(ii) Induction on the structure of S.

(*app*) Consider the case where $S \equiv sU$ is an application.

$$\begin{aligned}
\Big(\frac{\partial sU}{\partial x} \cdot T\Big)^r &\equiv \left(\Big(\frac{\partial s}{\partial x} \cdot T\Big)U + \Big(\mathsf{D}s \cdot \big(\frac{\partial U}{\partial x} \cdot T\big)\Big)U\right)^r \\
&\equiv \Big(\frac{\partial s}{\partial x} \cdot T\Big)^r \big[(U^r)^!\big] + \Big(\mathsf{D}s \cdot \big(\frac{\partial U}{\partial x} \cdot T\big)\Big)^r \big[(U^r)^!\big] \\
&\equiv \Big(\frac{\partial s}{\partial x} \cdot T\Big)^r \big[(U^r)^!\big] + \Big(\lambda y.\big(s^r\big[\big(\frac{\partial U}{\partial x} \cdot T\big)^r, y^!\big]\big)\Big)\big[(U^r)^!\big] \\
&= \big(s^r\langle T^r/x\rangle\big)\big[(U^r)^!\big] + s^r\big[U^r\langle T^r/x\rangle, (U^r)^!\big] && (\text{IH and } \beta_{\mathsf{R}}) \\
&\equiv \big(s^r\big[(U^r)^!\big]\big)\langle T^r/x\rangle \\
&\equiv (sU)^r\langle T^r/x\rangle
\end{aligned}$$

(D) Consider the case where $S \equiv \mathsf{D}s \cdot u$ is a differential application.

$$
\begin{aligned}
&\Big(\frac{\partial}{\partial x}(\mathsf{D}s \cdot u) \cdot T\Big)^r \\
&\equiv \Big(\mathsf{D}\Big(\frac{\partial s}{\partial x} \cdot T\Big) \cdot u + \mathsf{D}s \cdot \Big(\frac{\partial u}{\partial x} \cdot T\Big)\Big)^r \\
&\equiv \lambda y.\Big(\Big(\frac{\partial s}{\partial x} \cdot T\Big)^r [u^r, y^!]\Big) + \lambda y.\Big(s^r \Big[\Big(\frac{\partial u}{\partial x} \cdot T\Big)^r, y^!\Big]\Big) \\
&= \lambda y.\big(s^r \langle T^r/x\rangle [u^r, y^!]\big) + \lambda y.(s^r [u^r \langle T^r/x\rangle, y^!]) && \text{(IH)} \\
&\equiv \lambda y.\Big(s^r \langle T^r/x\rangle [u^r, y^!] + s^r \big([u^r \langle T^r/x\rangle, y^!] + [u^r, y\langle T^r/x\rangle, y^!]\big)\Big) && (*) \\
&\equiv \lambda y.\Big(s^r \langle T^r/x\rangle [u^r, y^!] + s^r \big([u^r, y^!]\langle T^r/x\rangle\big)\Big) \\
&\equiv \big(\lambda y.(s^r [u^r, y^!])\big)\langle T^r/x\rangle
\end{aligned}
$$

$(*)$ Note that by notation 3.1, $[u^r, y\langle T^r/x\rangle, y^!] \equiv [u^r, 0, y^!] \equiv 0$.

□

Proposition 4.4. For any $s, t \in \Lambda^d$, we have

(i) $\Gamma \vdash_{\mathcal{D}} s : \alpha \iff \Gamma \vdash_{\mathcal{R}} s^r : \alpha$,

(ii) If s and t are provably equal in the theory $\lambda\beta\eta_{\mathsf{D}}$, then their translations are also provably equal in $\lambda\beta\eta_{\mathsf{R}}$. i.e.

$$\lambda\beta\eta_{\mathsf{D}} \rhd \Gamma \vdash_{\mathcal{D}} s = t : \alpha \qquad \Longrightarrow \qquad \lambda\beta\eta_{\mathsf{R}} \rhd \Gamma \vdash_{\mathcal{R}} s^r = t^r : \alpha.$$

Proof. (i) Easy induction on the length of the proof.

(ii) We prove it by induction on all the rules in the theory $\lambda\beta\eta_{\mathsf{D}}$. We would only consider the (β_{D}) rule where $\mathsf{D}(\lambda x.s) \cdot T = \lambda x.\Big(\frac{\partial s}{\partial x} \cdot T\Big)$.

$$
\begin{aligned}
\big(\mathsf{D}(\lambda x.s) \cdot T\big)^r &= \lambda y.\big((\lambda x.s^r)\ [T^r, y^!]\big) \\
&= \lambda y.\big(s^r \langle T^r/x\rangle [y/x]\big) \\
&= \lambda x.\big(s^r \langle T^r/x\rangle\big) \\
&= \lambda x.\Big(\frac{\partial s}{\partial x} \cdot T\Big)^r && \text{(Lemma 4.3)} \\
&= \Big(\lambda x.\Big(\frac{\partial s}{\partial x} \cdot T\Big)\Big)^r
\end{aligned}
$$

□

Proposition 5.1. Let $f : A_1 \times A_2 \to B$ and $g, h : A_1 \to A_2$ be morphisms in a cartesian closed differential category. We have

$$(f \star g) \star h = (f \star h) \star g.$$

Proof. Note that $\langle 0,0\rangle = 0$.

$$
\begin{aligned}
& (f \star g) \star h \\
&= D\Big[D[f] \circ \big\langle \langle 0, g\circ \pi_1\rangle, \mathsf{Id}\big\rangle\Big] \circ \big\langle \langle 0, h\circ\pi_1\rangle, \mathsf{Id}\big\rangle \\
&= D\big[D[f]\big] \circ \Big\langle D\big[\big\langle\langle 0, g\circ\pi_1\rangle, \mathsf{Id}\big\rangle\big], \big\langle\langle 0, g\circ\pi_1\rangle, \mathsf{Id}\big\rangle \circ \pi_2\Big\rangle \circ \big\langle\langle 0, h\circ\pi_1\rangle, \mathsf{Id}\big\rangle && \text{(CD5)} \\
&= D\big[D[f]\big] \circ \Big\langle \big\langle\langle 0, D[g]\circ(\pi_1\times\pi_1)\rangle, \pi_1\big\rangle, \big\langle\langle 0, g\circ\pi_1\rangle, \mathsf{Id}\big\rangle \circ \pi_2\Big\rangle \circ \big\langle\langle 0, h\circ\pi_1\rangle, \mathsf{Id}\big\rangle && \text{(CD3-5)} \\
&= D\big[D[f]\big] \circ \Big\langle \big\langle\langle 0, D[g]\circ\langle 0, \pi_1\rangle\rangle, \langle 0, h\circ\pi_1\rangle\big\rangle, \big\langle\langle 0, g\circ\pi_1\rangle, \mathsf{Id}\big\rangle\Big\rangle \\
&= D\big[D[f]\big] \circ \Big\langle \big\langle\langle 0, 0\rangle, \langle 0, h\circ\pi_1\rangle\big\rangle, \big\langle\langle 0, g\circ\pi_1\rangle, \mathsf{Id}\big\rangle\Big\rangle && \text{(CD2)} \\
&= D\big[D[f]\big] \circ \Big\langle \big\langle 0, \langle 0, h\circ\pi_1\rangle\big\rangle, \big\langle\langle 0, g\circ\pi_1\rangle, \mathsf{Id}\big\rangle\Big\rangle \\
&= D\big[D[f]\big] \circ \Big\langle \big\langle 0, \langle 0, g\circ\pi_1\rangle\big\rangle, \big\langle\langle 0, h\circ\pi_1\rangle, \mathsf{Id}\big\rangle\Big\rangle && \text{(CD7)} \\
&= (f\star h)\star g
\end{aligned}
$$

□

Lemma 5.6. (i) If $x \notin \mathsf{FV}(t)$, then

$$(\mathsf{D}(\lambda x.(t[t'/x'])) \cdot u)\ s = (\mathsf{D}(\lambda x'.t) \cdot ((\mathsf{D}(\lambda x.t') \cdot u)\ s))\ t'[s/x].$$

(ii) if $x \not\equiv y$, $x, y \notin \mathsf{FV}(M)$ and $x, y, z \notin \mathsf{FV}(u) \cup \mathsf{FV}(v)$,

$$\Big(\frac{\partial}{\partial x}\big(M[\langle x, y\rangle/z]\big) \cdot u\Big)\ [v/x] = \Big(\frac{\partial M}{\partial z} \cdot \langle u, 0\rangle\Big)\ [\langle v, y\rangle/z].$$

Proof. (i)

$$
\begin{aligned}
(\mathsf{D}(\lambda x.(t[t'/x'])) \cdot u)\ s &= (\mathsf{D}(\lambda x.((\lambda x'.t)\ t')) \cdot u)\ s \\
&= (\lambda x.(\frac{\partial}{\partial x}((\lambda x'.t)\ t') \cdot u))\ s \\
&= (\lambda x.((\frac{\partial}{\partial x}(\lambda x'.t) \cdot u)\ t' + (\mathsf{D}(\lambda x'.t) \cdot (\frac{\partial t'}{\partial x} \cdot u))\ t'))\ s \\
&= (\lambda x.((\lambda x'.(\frac{\partial t}{\partial x} \cdot u))\ t' + (\lambda x'.(\frac{\partial t}{\partial x'} \cdot (\frac{\partial t'}{\partial x} \cdot u)))\ t'))\ s \\
&= (\lambda x.(0 + (\frac{\partial t}{\partial x'} \cdot (\frac{\partial t'}{\partial x} \cdot u))[t'/x']))\ s \\
&= (\frac{\partial t}{\partial x'} \cdot (\frac{\partial t'}{\partial x} \cdot u))[t'/x'][s/x] \\
(\mathsf{D}(\lambda x'.t) \cdot ((\mathsf{D}(\lambda x.t') \cdot u)\ s))\ t'[s/x] &= (\lambda x'.(\frac{\partial t}{\partial x'} \cdot ((\lambda x.(\frac{\partial t'}{\partial x} \cdot u))\ s)))\ t'[s/x] \\
&= (\frac{\partial t}{\partial x'} \cdot ((\frac{\partial t'}{\partial x} \cdot u)[s/x]))[t'[s/x]/x'] \\
&= (\frac{\partial t}{\partial x'} \cdot (\frac{\partial t'}{\partial x} \cdot u))[s/x][t'[s/x]/x'] \\
&= (\frac{\partial t}{\partial x'} \cdot (\frac{\partial t'}{\partial x} \cdot u))[t'/x'][s/x]
\end{aligned}
$$

(ii) Prove by induction on structure of M.

- $M \equiv w \neq z$
  Note that $w \neq x,y$ since $x,y \notin \mathsf{FV}(M)$

$$\Big(\frac{\partial\ w[\langle x,y\rangle/z]}{\partial x} \cdot u\Big)\ [v/x] = \Big(\frac{\partial w}{\partial x} \cdot u\Big)\ [v/x] = 0 = \Big(\frac{\partial w}{\partial z} \cdot \langle u,0\rangle\Big)\ [\langle v,y\rangle/z]$$

- $M \equiv z$

$$\Big(\frac{\partial\ \langle x,y\rangle}{\partial x} \cdot u\Big)\ [v/x] = \langle u,0\rangle\ [v/x] = \langle u,0\rangle = \langle u,0\rangle\ [\langle v,y\rangle/z] = \Big(\frac{\partial z}{\partial z} \cdot \langle u,0\rangle\Big)\ [\langle v,y\rangle/z]$$

- $M \equiv sT$

$$\begin{aligned}
&\Big(\frac{\partial (sT)}{\partial z} \cdot \langle u,0\rangle\Big)\ [\langle v,y\rangle/z] \\
&= \Big(\Big(\frac{\partial s}{\partial z} \cdot \langle u,0\rangle\Big)\ T\ +\ \Big(\mathsf{D}s \cdot \Big(\frac{\partial T}{\partial z} \cdot \langle u,0\rangle\Big)\Big)\ T\Big)\ [\langle v,y\rangle/z] \\
&= \Big(\frac{\partial s}{\partial z} \cdot \langle u,0\rangle\Big)[\langle v,y\rangle/z]\ T[\langle v,y\rangle/z]\ + \\
&\quad \Big(\mathsf{D}(s[\langle v,y\rangle/z]) \cdot \Big(\Big(\frac{\partial T}{\partial z} \cdot \langle u,0\rangle\Big)[\langle v,y\rangle/z]\Big)\Big)\ T[\langle v,y\rangle/z] \\
&= \Big(\frac{\partial\ s[\langle x,y\rangle/z]}{\partial x} \cdot u\Big)\ [v/x]\ T[\langle v,y\rangle/z]\ + \\
&\quad \Big(\mathsf{D}(s[\langle v,y\rangle/z]) \cdot \Big(\Big(\frac{\partial\ T[\langle x,y\rangle/z]}{\partial x} \cdot u\Big)\ [v/x]\Big)\Big)\ T[\langle v,y\rangle/z] \qquad \text{(IH)} \\
&= \Big(\frac{\partial\ s[\langle x,y\rangle/z]}{\partial x} \cdot u\Big)\ [v/x]\ T[\langle x,y\rangle/z][v/x]\ + \\
&\quad \Big(\mathsf{D}(s[\langle x,y\rangle/z][v/x]) \cdot \Big(\Big(\frac{\partial\ T[\langle x,y\rangle/z]}{\partial x} \cdot u\Big)\ [v/x]\Big)\Big)\ T[\langle x,y\rangle/z][v/x] \\
&= \Big(\Big(\frac{\partial\ s[\langle x,y\rangle/z]}{\partial x} \cdot u\Big)\ T[\langle x,y\rangle/z]\ + \\
&\quad \Big(\mathsf{D}(s[\langle x,y\rangle/z]) \cdot \Big(\frac{\partial\ T[\langle x,y\rangle/z]}{\partial x} \cdot u\Big)\Big)\ T[\langle x,y\rangle/z]\Big)[v/x] \\
&= \Big(\frac{\partial\ (sT)[\langle x,y\rangle/z]}{\partial x} \cdot u\Big)\ [v/x]
\end{aligned}$$

The proof for the cases for $\lambda w.s$, $\mathsf{D}s \cdot t$, 0, $s + T$, $\langle s,t\rangle$, $\mathsf{Fst}(s)$ and $\mathsf{Snd}(s)$ are trivial.

□