# Synthetic Differentiation and its Connection to Iteration in Asynchronous Processes

Isaiah B. Hilsenrath

A THESIS

in

Mathematics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Master of Arts

2025

Supervisor of Thesis
Dr. Henry Towsner
Professor of Mathematics

Co-Supervisor of Thesis
Dr. Andre Scedrov
Professor of Mathematics

Graduate Chair of Mathematics
Dr. Ron Donagi
Thomas A. Scott Professor of Mathematics

# Abstract

In this thesis, we describe three variants of a purely synthetic framework to treat the differentiation of analytic functions: differential $\lambda$-calculus, Taylor $\lambda$-calculus, and differential categories. We present the first explicit description of the relationship between Taylor $\lambda$-calculus and differential $\lambda$-calculus, showing that Taylor $\lambda$-calculus can be thought of a subset of differential $\lambda$-calculus. Separately, we generalize the existing definition of the free cornering on a resource theory, which is a model of asynchronous processes, in the literature to include infinite resources and the ability to borrow resources. We will, then, prove that the category of resource transducers induced by this generalized notion of the free cornering has a canonical differential categorical structure.

# Acknowledgements

I think it is special to have just one person truly guide you; I am very fortunate to have so many amazing people make a difference in my academic journey. I first and foremost want to thank my incredible advisors Henry Towsner and Andre Scedrov. If it wasn't for their unwavering support on this massive effort, this thesis would not be possible. I was captivated by pure logic from the moment I walked into Henry's introductory logic class, and I greatly enjoyed the countless hours I spent learning linear logic in Andre's topics class on logic. I am greatly indebted to Henry and Andre for all the conversations we had over this past year. Thank you Henry and Andre!

I would also like to thank Spencer Breiner and Priyaa Varshinee Srinivasan, who worked with me during the initial stages of this thesis. I am greatly indebted to Spencer for agreeing to spend the last year of my undergrad teaching me everything I now know about category theory and introducing me to applied category theory and categorical logic. And I am greatly indebted to Priyaa for her wonderful mentorship as my supervisor during the 2023 Adjoint School. Priyaa's invaluable insight in linearly distributive categories and differential categories steadily guided the initial stages of this work, and my countless conversations with Priyaa never failed to be fascinating and enjoyable. Thank you Spencer and Priyaa!

In addition, I would like to thank my colleagues in the 2023 Adjoint School, Brandon Baylor, Paige Frederick, Rose Kudzman-Blais, Durgesh Kuma, and Fabian Wiesner for reading with me Chad Nester's work and for all the fun conversations we had.

Last but not least, I would like to thank my parents, Hana and Oliver; my brothers, Saul and Lior; and my sisters, Nama and Ella for their unconditional support and encouragement, especially over these past two years of my PhD.

*To my parents, Hana and Oliver.*

# Table of Contents

# Chapter 1

# Introduction

One of the key concepts that will be studied in this thesis is that of an a *asynchronous processes*. The basic idea of an asynchronous processes is that it is a system that that is composed from simpler processes that take in a certain collection of resources and output a certain (potentially different) collection of resources and/or messages that communicate the state of those resources after the process has concluded. The collection of all possible resources that is used by an asynchronous processes (or a collection of asynchronous processes) is called a *resource theory*, and the behavior of each simpler processes within an asynchronous process can also be influenced by the messages produced by other simpler processes in the system. Those processes that just take in and then output a message, without taking or outputting a resource, is called a *resource transducer*.

In 2023, Chad Nester, an Estonian mathematician at the University of Tartu, defined the *free cornering* in [15], which is a rudimentary graphical model of asynchronous processes. The general idea in this paper is that a symmetric monoidal category can be thought of as a resource theory, where the objects are some collection of resources, morphisms are processes that manipulate resources, and tensoring represents having two resources or performing two processes manipulating resources simultaneously. From a symmetric monoidal category, one can build a double category called the free cornering, where a subset of cells called vertical cells represent processes on resources and a subset of cells called horizontal cells represent messages between distinct parts of the system. The subset of horizontal cells form the category of resource transducers. In [16], Chad Nester showed that the category of resource transducers is an instance of *linear actegories*, a logic of message passing developed by Canadian mathematicians J.R.B. Cockett and Craig Pastro in [6] back in 2007. Importantly, linear actegories are grounded in linear logic, which is the first indication that free cornering and linear logic are intimately tied, and is one of the motivations for the work in this thesis.

Separate from the topic of free cornering, in 2004, Thomas Ehrhard and Laurent Regnier, French mathematicians at the Institut de Mathématique de Luminy, developed a powerful framework to express the differentiation of analytic functions purely syntactically called the differential $\lambda$-calculus in [8]. In attempt to simplify the syntax of differential $\lambda$-calculus and make its connection to linear logic more apparent, Ehrhard and Regnier in 2006 (see [7]) proposed two new frameworks to express the differentiation of analytic functions syntactically: *Taylor $\lambda$-calculus* and *differential interaction nets*, which is also referred to as *differential linear logic*. While these systems seem very different at first glance, as stated toward the end of [7] and as restated in Theorem 4.1.14 in this thesis, Taylor $\lambda$-calculus and differential interaction nets are actually equivalent. Also, in 2006, at the same time that Ehrhard and Regnier were working on these two frameworks of synthetic differentiation, R.F. Blute, J.R.B. Cockett, and R.A.G. Seely, formulated a fourth framework to express the differentiation of analytic functions purely syntactically called the *differential category* (see [1]). It is immediate that the morphisms in the differential category that are formed only by the differential category's structure maps are, in fact, just differential interaction nets. In this way, differential categories and differential interaction nets are equivalent as systems for synthetic differentiation.

Chapter 2 of this thesis will carefully work out the details of differential $\lambda$-calculus, making a few minor improvements to the definitions of differential terms, substitution, and partial differentiation to ensure that the notion of free variables, substitution, and partial differentiation are well-defined. Chapter 3 will, then, carefully work out the details of Taylor $\lambda$-calculus. It is important to note that Taylor $\lambda$-calculus, differential interaction nets, and differential categories are equivalent frameworks for synthetic differentiation. But this raises a natural question? how do these systems relate to the original framework for synthetic differentiation, differential $\lambda$-calculus? It is strongly suggested in [9] that Taylor-$\lambda$ calculus is a subset of differential $\lambda$-calculus. In section 3.4, we provide the first explicit description of this, showing that substitution, partial differentiation, and differential $\beta$-reduction is preserved by the canonical injection from Taylor $\lambda$-calculus into differential $\lambda$-calculus (see Lemma 3.4.5, Lemma 3.4.7, Theorem 3.4.8, and Lemma 3.4.9).

At this point, we have mentioned asynchronous processes and synthetic differentiation, but so far, we have not in any way illustrated the reason these two concepts are related. Let us change that.

A year after Chad Nester published his original work on free cornering, Chad Nester together with Niels Voorneveld, another Estonian mathematician, improved the definition of the free cornering in [16]. In this paper, they define the *free cornering with choice*, which enables the description of control statements (i.e., choice) in an asynchronous process, and the *free cornering with choice and iteration*, which enables the description of loops (i.e., iteration) along with choice in an asynchronous process. Interestingly, according to Lemmas 18-19 in [16] (which are restated in Lemmas 4.3.4 and 4.3.5 in this thesis), the right comultiplication cell, right counit cell, the right comonadic comultiplication cell, and right comonadic counit cell, which are all resource transducers and are defined in Definition 4.3.3, behave surprisingly similarly to the natural transformations contraction, weakening, digging, and dereliction in a differential category. Hence, it is only natural to assume that that category of resource transducers under suitable conditions is, in fact, just a differential category. That is what this thesis will explore in Chapter 4.

In order to show that the category of resource transducers is, in fact, a differential category, it is important to recognize that the category of resource transducers misses two key properties. First, the category of resource transducers is not symmetric, as Chad Nester points out in Observation 6.4 in [15]. Second, the category of resource transducers lacks natural transformations corresponding to co-contraction, co-weakening, and co-dereliction, which appear in differential categories. To show that the category of resource transducers has these properties, we add very natural conditions to our resource theory and correspondingly generalize our notion of a free cornering with choice and iteration. Namely, to add symmetry, it will be necessary to allow resources to be borrowed (in the sense presented in [14]) and allow the existence of infinite resources, and to add co-contraction, co-weakening, and co-dereliction, we will need an additive structure in our resource theory and free cornering. This results in the definition of a resource category, presented in Definition 4.1.16, and the free cornering with choice and iteration of a resource category, presented in Definition 4.2.3. Under these minimal conditions, we find the remarkable fact that the category of resource transducers has a canonical differential categorical structure (see Theorem 4.3.21).

# Chapter 2

# Differential $\lambda$-Calculus

In this chapter, we will introduce the first of four formulations of a purely syntactic framework to treat differentiation of analytic functions, which is primarily based on Thomas Ehrhard and Laurent Regnier's seminal paper [8] on the topic. The first section will give a brief motivation of the syntax of differential $\lambda$-calculus without getting too in the weeds with the precise construction of a differential term. The second section will then provide the precise construction of a term in this calculus. To accomplish this, it will be necessary for us to define what will be called differential pre-terms, whose properties will make up most of the material of the second section. The third section will introduce a recursive definition of the partial derivative of a differential term and will demonstrate that this definition of a partial derivative allows for purely symbolic formulations of the chain rule, the generalized Leibniz rule, and Schwarz's theorem. The fourth and final section of this chapter will introduce a procedure of evaluating a differential term, expanding the usual notion of $\beta$-reduction in ordinary lambda calculus. This section will then conclude with a proof that the differential $\lambda$-calculus satisfies the Church-Rosser property.

## 2.1 Motivating the Syntax of Differential $\lambda$-Calculus

The title of this paper uses quite a peculiar phrase: *synthetic differentiation*. But what precisely does *synthetic* mean, especially in the context of differentiation? The idea, here, is that we want to be able to express ideas or prove properties about differentiation of analytic functions via purely symbolic reasoning, rather than via, say, the usual limit definition of differentiation. So, in this context, *synthetic* means *via purely symbolic reasoning*, and this is precisely what differential $\lambda$-calculus accomplishes. This is a very similar idea to how sequent calculus, originally introduced by German mathematician Gerhard Gentzen in 1935 [10], allows us to draw conclusions about properties of classical logic via syntactic manipulations of sequents without considering the actual semantic notion of truth underpinning classical logic.

Now that we have a rough idea of the goal of synthetic differentiation, which is to capture syntactically the behavior of differentiation, there is a natural follow-up question. How should we construct the syntax to capture the behavior of differentiation in a way that is computationally useful? The first syntactic element we will need is a denumerable set $\mathcal{V}$ of symbols (henceforth called *variables*), which we will use to represent the arguments of functions such as $x, y, z$ and to represent analytic functions themselves such as $f : \mathbb{R}^2 \to \mathbb{R}$. This denumerable set $\mathcal{V}$ of variables will serve as the base case of our differential terms. However, labeling an analytic function by the symbol $f$ does not tell us much about what $f$ actually is, and hence, is not very useful computationally. We should have differential terms that allow us to break down a function into an argument and some simpler differential term that may or may not use this argument. Such a term is known as an *abstraction*, and is written as $\lambda x.s$, where $x$, a variable, is the argument of the function and $s$, a simpler differential term, is some process that may or may not use $x$. For example, using this notation, we would write the function that takes in a number and returns that inputted number (i.e., the identity function) as $\mathbf{I} = \lambda x.x$, and we would write the projection function that takes any number to some value represented by $y$ (i.e., the projection function) as $\mathbf{P} = \lambda x.y$. It seems though, since an abstraction only has one argument, that we can only express functions that take in one parameter, like the identity $\mathbf{I}$ and the projection function $\mathbf{P}$. However, by using multiple abstractions, we can actually express a function that takes any number of arguments. A function $f : \mathbb{R}^n \to \mathbb{R}$ can be written as $f = \lambda x_n.(\ldots(\lambda x_1.s)\ldots)$, where $x_i$ is a variable representing the $i$th argument of $f$ and $s$ is some differential term representing a process that may or may not use the arguments $x_1, \ldots, x_n$. For example, using this notation, we would write the function that takes in two arguments and returns the second input as $\mathbf{K} = \lambda x.(\lambda y.x)$. The last fundamental syntactic element we will need before getting into differentiation is some way to express the composition of two functions. This is accomplished by concatenation, and a term formed this way is called an *application*. For example, the composition of the differential terms representing the analytic functions $s$ and $t$ is written as $st$. To recap, so far, we have that a differential term is one of the following:

(i) a variable;

(ii) an abstraction $\lambda x.s$, where $x$ is a variable and $s$ is a differential term; or

(iii) an application $st$, where $s$ and $t$ are differential terms.

However, at this point, we still have not introduced any syntax representing the derivative of a function. In fact, the only differential terms we have discussed are simply the terms found in ordinary $\lambda$-calculus. Let us fix this. Given a positive integer $n \in \mathbb{N}$, positive numbers $i_1, \ldots, i_n$, a variable $f$, representing an analytic function of type $\mathbb{R}^m \to \mathbb{R}$, and differential terms $t_1, \cdots, t_n$, representing some functions, we introduce the differential term $D_{i_1,\ldots,i_n} f * (t_1, \ldots, t_n)$, which represents the function whose type is the same as $f$ and is given by

$$x_1, \ldots, x_m \mapsto (\partial_{x_{i_1}} \ldots \partial_{x_{i_n}} f)(x_1, \ldots, x_m) \cdot t_1(x_1, \ldots, x_m) \cdot t_n(x_1, \ldots, x_m). \tag{2.1.1}$$

Note that $i_1, \ldots, i_n$ do not necessarily need to be distinct. Similarly, given a positive integer $n \in \mathbb{N}$, a variable $x$, representing an argument of a function, a differential term $s$, representing a process that may or may not use the argument $x$, and differential terms $t_1, \cdots, t_n$, representing some functions, we introduce the differential term $D^n_x \lambda x.s * (t_1, \ldots, t_n)$, which represents the function of type $\mathbb{R} \to \mathbb{R}$ given by

$$x \mapsto (\partial^n_x f)(x) \cdot t_1(x) \cdot t_n(x), \tag{2.1.2}$$

where $f : \mathbb{R} \to \mathbb{R}$ is the one-dimensional analytic function represented by $\lambda x.s$. Observe that (2.1.2) coincides with (2.1.1) when $m = 1$. Hence, a differential term can also be of the following:

(iv) a *differentiated variable* $(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n))$, where $n, i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, $x \in \mathcal{V}$ is a variable, and $t_1, \ldots, t_n$ are differential terms; or

(v) a *differentiated abstraction* $(D^n_1 \lambda x.s * (t_1, \ldots, t_n))$, where $n \in \mathbb{N}$ is a positive integer, $x \in \mathcal{V}$ is a variable, and $s, t_1, \ldots, t_n$ are differential terms.

This gives us a notion of differentiation that very conservatively extends ordinary $\lambda$-calculus. Yet, if we want any shot at expressing even simple behavior of derivatives, such as the product rule, we are still missing one critical ingredient: *sums*. That is, we want differential terms that take the form $\sum_{i=1}^{n}(r_i \circ s_i)$, where $s_i$ are simpler differential terms and $r_i$ are some symbols from the set $\mathcal{R}$ (henceforth called *scalars*), representing real numbers. Hence, a differential term can also be of the following:

(vi) a *scalar multiplication* $r \cdot s$, where $r$ is a scalar and $s$ is a differential term;

(vii) a *sum* $s + t$, where $s, t$ are differential terms; or

(viii) 0, which represent the zero functions.

These eight different kinds of differential terms will make up our calculus, which will be called differential $\lambda$-calculus. With this intuition at hand, let us make this notion of a differential term much more precise.

## 2.2 Differential Pre-Terms and Terms

It is tempting to jump straight into defining differential terms, and it is not entirely unreasonable to do this. But by making this leap, we would be glossing over a critical point: the set of differential terms is not a freely generated system. This should not be surprising. Just note, for instance, that sums of differential terms should clearly be identified by commutativity and associativity. The lack of freeness in the set of differential terms means that we cannot automatically claim that recursively defined function on the set of differential terms, such as partial differentiation, is well-defined. To resolve this issue, we define the set of differential pre-terms, which is freely generated, and then quotient this set by the appropriate relations to obtain the set of differential terms. This approach is motivated by [13], where the authors, Sørensen Morten and Urzyczyn Pawel, quotiented the freely-generated set of pre-terms to obtain the set of ordinary $\lambda$-terms.

**Definition 2.2.1. Set of Scalars**

Let $\mathcal{S}$ be a denumerable set of symbols and let 1 be a distinguished symbol. Then, the *set of scalars* $\mathcal{R}$ on $\mathcal{S}$ is the free set on $\mathcal{S} \cup \{1\}$ generated by two binary connectives $\cdot$ and $+$. ♣

**Definition 2.2.2. Differential Pre-Terms**

Given a denumerable set $\mathcal{V}$ of variables, a set of scalars $\mathcal{R}$, and a distinguished symbol 0, a *simple differential pre-term* and a *complex differential pre-term* are defined via mutual induction as follows:

(i) Given a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, positive integers $i_1, \ldots, i_n \in \mathbb{N}$, and simple differential pre-terms $t_1, \ldots, t_n$, $(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n))$ is a simple differential pre-term.

(ii) Given a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, and simple differential pre-terms $s, t_1, \ldots, t_n$, $(D_1^n \lambda x.s * (t_1, \ldots, t_n))$ is a simple differential pre-term.

(iii) Given a simple differential pre-term $s$ and a complex differential pre-term $\tau$, $(s\tau)$ is a simple differential pre-term.

(iv) 0 is a complex differential pre-term.

(v) Given a simple differential pre-term $s$, $s$ is also a complex differential pre-term.

(vi) Given a scalar $r \in \mathcal{R}$ and a complex differential pre-term $\tau$, $(r \cdot \tau)$ is a complex differential pre-term.

(vii) Given complex differential pre-terms $\sigma$ and $\tau$, $(\sigma + \tau)$ is a complex differential pre-term.

Let $\Delta_{d,\mathcal{V}}$ denote the set of simple differential pre-terms, and let $\mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ denote the set of complex differential pre-terms. ♣

Notice that at virtually every inductive step in the definition of a differential pre-term, outer parentheses are added. This quickly makes differential pre-terms rather unwieldy, so we will adopt some straightforward notational conventions to make differential pre-terms easier to read and write.

**Notation 2.2.3**

We apply the following notational conventions when writing differential pre-terms:

(i) The outermost parentheses are omitted. For example, $(D_{6,2}x * (y, y))$, where $x, y \in \mathcal{V}$ are variables, is abbreviated as $D_{6,2}x * (y, y)$.

(ii) If $n = 1$ in a differential pre-term in the shape of Definition 2.2.2 (i)-(ii), the parenthesis around the list of simple pre-terms at the end of the differential pre-term are dropped. For example, $(D_3\lambda x.x * (y))$, where $x, y \in \mathcal{V}$ are variables, is abbreviated as $(D_3\lambda x.x * y)$.

(iii) Application associates to the right. For example, $((s\tau)\sigma)$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\tau, \sigma \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms, is abbreviated to $(s\tau\sigma)$

(iv) Summation associates to the right. For example, $((\rho + \sigma) + \tau)$, where $\rho, \sigma, \tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms, is abbreviated to $(\rho + \sigma + \tau)$.

(v) Multiplication binds stronger than summation. For example, $((r_1 \cdot \tau) + (r_2 \cdot \sigma))$, where $r_1, r_2 \in \mathcal{R}$ are scalars and $\tau, \sigma \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms, is abbreviated as $(r_1 \cdot \tau + r_2 \cdot \sigma)$. ♠

It is important to note, before getting further, that not all variables were created equal; rather, variables fall into two distinct classes. To build some intuition, think of the function $h := t \mapsto \frac{1}{2}gt^2$ representing the distance a ball falls at time $t$ after it was dropped at time 0 given that the constant of gravitational acceleration is $g$. The symbols $g$ and $t$ are clearly both variables. However, they do not exactly behave in the same way. $g$ is a constant in the function $h$ representing some value, which *can* be specified later. We call $g$ a *free variable* of $h$. On the other hand, $t$ is the argument of the function $h$. We call $t$ the *bound variable* of $h$. That is, constants of functions are *free variable* and arguments of functions are bound variables. Equipped with this intuition, let us translate this into differential λ-calculus. In differential λ-calculus, we would write $h$ as $\frac{1}{2}\lambda t.Pgtt$, where $P$ is some differential pre-term that returns the product of three inputted numbers when evaluated. Hence, in differential λ-calculus, a variable is said to be free when it is not contained in an abstraction whose corresponding variable is the same; otherwise, the variable is said to be bound.

**Definition 2.2.4. The Variables of Differential Pre-Terms**

Define the function $\mathrm{Var}\colon \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle \to \mathcal{V}^*$ that takes a complex differential pre-term to a finite subset of variables by:

$$\mathrm{Var}\left(D_{i_1,\ldots,i_n} x * (t_1,\ldots,t_n)\right) = \{x\} \cup \mathrm{Var}\,(t_1) \cup \cdots \cup \mathrm{Var}\,(t_n), \tag{2.2.1a}$$
$$\mathrm{Var}\left(D_1^n \lambda x.s * (t_1,\ldots,t_n)\right) = \{x\} \cup \mathrm{Var}\,(s)\;\mathrm{Var}\,(t_1) \cup \cdots \cup \mathrm{Var}\,(t_n), \tag{2.2.1b}$$
$$\mathrm{Var}\,(s\mu) = \mathrm{Var}\,(s) \cup \mathrm{Var}\,(\mu), \tag{2.2.1c}$$
$$\mathrm{Var}\,(0) = \varnothing, \tag{2.2.1d}$$
$$\mathrm{Var}\,(r\cdot\mu) = \mathrm{Var}\,(\mu), \tag{2.2.1e}$$
$$\mathrm{Var}\,(\mu+\nu) = \mathrm{Var}\,(\mu) \cup \mathrm{Var}\,(\nu), \tag{2.2.1f}$$

where $x \in \mathcal{V}$ is a variable, $r \in \mathcal{R}$ is a scalar, $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1,\ldots,i_n \in \mathbb{N}$ are positive integers, $s, t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $\mu,\nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. The image of a complex differential pre-term $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ under Var will be called the *variables* of $\tau$. ♣

**Definition 2.2.5. The Free Variables of Differential Pre-Terms**

Define the function $\mathrm{FV}\colon \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle \to \mathcal{V}^*$ that takes a complex differential pre-term to a finite subset of variables by:

$$\mathrm{FV}\left(D_{i_1,\ldots,i_n} x * (t_1,\ldots,t_n)\right) = \{x\} \cup \mathrm{FV}\,(t_1) \cup \cdots \cup \mathrm{FV}\,(t_n), \tag{2.2.2a}$$
$$\mathrm{FV}\left(D_1^n \lambda x.s * (t_1,\ldots,t_n)\right) = (\mathrm{FV}\,(s) - \{x\}) \cup \mathrm{FV}\,(t_1) \cup \cdots \cup \mathrm{FV}\,(t_n), \tag{2.2.2b}$$
$$\mathrm{FV}\,(s\mu) = \mathrm{FV}\,(s) \cup \mathrm{FV}\,(\mu), \tag{2.2.2c}$$
$$\mathrm{FV}\,(0) = \varnothing, \tag{2.2.2d}$$
$$\mathrm{FV}\,(r\cdot\mu) = \mathrm{FV}\,(\mu), \tag{2.2.2e}$$
$$\mathrm{FV}\,(\mu+\nu) = \mathrm{FV}\,(\mu) \cup \mathrm{FV}\,(\nu), \tag{2.2.2f}$$

where $x \in \mathcal{V}$ is a variable, $r \in \mathcal{R}$ is a scalar, $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1,\ldots,i_n \in \mathbb{N}$ are positive integers, $s, t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $\mu,\nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. The image of a complex differential pre-term $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ under FV will be called the *free variables* of $\tau$. ♣

**Definition 2.2.6. The Bound Variables of Differential Pre-Terms**

Define the function $\mathrm{BV}\colon \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle \to \mathcal{V}^*$ that takes a complex differential pre-term to a finite subset of variables by:

$$\mathrm{BV}\left(D_{i_1,\ldots,i_n} x * (t_1,\ldots,t_n)\right) = \mathrm{BV}\,(t_1) \cup \cdots \cup \mathrm{BV}\,(t_n), \tag{2.2.3a}$$
$$\mathrm{BV}\left(D_1^n \lambda x.s * (t_1,\ldots,t_n)\right) = \{x\} \cup \mathrm{BV}\,(s) \cup \mathrm{BV}\,(t_1) \cup \cdots \cup \mathrm{BV}\,(t_n), \tag{2.2.3b}$$
$$\mathrm{BV}\,(s\mu) = \mathrm{BV}\,(s) \cup \mathrm{BV}\,(\mu), \tag{2.2.3c}$$
$$\mathrm{BV}\,(0) = \varnothing, \tag{2.2.3d}$$
$$\mathrm{BV}\,(r\cdot\mu) = \mathrm{BV}\,(\mu), \tag{2.2.3e}$$
$$\mathrm{BV}\,(\mu+\nu) = \mathrm{BV}\,(\mu) \cup \mathrm{BV}\,(\nu), \tag{2.2.3f}$$

where $x \in \mathcal{V}$ is a variable, $r \in \mathcal{R}$ is a scalar, $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1,\ldots,i_n \in \mathbb{N}$ are positive integers, $s, t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $\mu,\nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. The image of a complex differential pre-term $\tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ under BV will be called the *bound variables* of $\tau$. ♣

The motivation for distinguishing free from bound variables spears when we try to define *substation*. In particular, when we want to substitute a certain complex differential pre-term, say $\sigma$ for a given variable, say, $x$ into a complex differential pre-term, say, *tau*, we only want to replace those instances of $x$ in $\tau$ that are free. To understand the reason we only replace free variables when substituting, we go back to our intuition of free variables as constants and bound variables as arguments. Intuitively it is clear constants are distinct from arguments: constants, as opposed to arguments, are fixed values that can be specified later. Specifically, we can substitute a constant with some chosen number. For example, consider again the function $h := t \mapsto \frac{1}{2}gt^2$ (or $h = \frac{1}{2}\lambda t.Pgtt$ in the notation of differential $\lambda$-calculus), which tells us how much a ball traveled as a given time $t$ after being dropped given that the constant of gravitational acceleration is $g$. To express appropriately how much the ball travels at a given time if the ball was dropped on the Moon, we would simply substitute every instance of $g$ in $h$ with $1.625\ m/s^2$, which would give us $h[(1.625\ m/s^2)/g] := t \mapsto \frac{1}{2}(1.625\ m/s^2)t^2$ (or $h = \frac{1}{2}\lambda t.P(1.625\ m/s^2)tt$ in the notation of differential $\lambda$-calculus). However, unlike a constant, it does not make sense to substitute the argument of a function with some number, as it could fundamentally change a potentially very dynamic function into some kind of projection. For example, if we substitute $1\ s$ for every instance of $t$ in $h$ we get the function $t \mapsto \frac{1}{2}g(1s)^2$ (or $\frac{1}{2}\lambda t.Pg(1\ s)(1\ s)$ in the notation of differential $\lambda$-calculus), which is a function that returns $\frac{1}{2}g(1s)^2$ for all inputs. No longer does the function, after this invalid substitution, represent the distance a ball fell given a particular time after the ball is dropped. Instead, this function represents a function that always returns the distance the ball fell after $1\ s$ given any time after the ball is dropped, making the fact that this is a function of time at all rather irrelevant. This is precisely the reason we can only replace the free instances of a variable during substitution.

The choice to only replace free variables and skip entirely bound variables during substitution is absolutely necessary to make sure a differential pre-term maintains its general behavior after substitution, as we demonstrated quite explicitly with the function representing the distance a ball falls after it is dropped. However, it is not quite sufficient. Sometimes, it is simply impossible to sensibly substitute a complex differential pre-term for a particular variable in another complex differential pre-term. For example, suppose that we substitute $t$ for $g$ in $h$. Then, we get the function $t \mapsto \frac{1}{2}t^3$ or ($\lambda t.Pttt$ in the notation of differential $\lambda$-calculus), which returns the inputted time cubed, and has no relation anymore to the distance a ball falls after it is dropped. Clearly, this should be an invalid substitution, as it fundamentally altered the behavior of the function. But what went wrong? The problem with this substitution is that we substituted a constant with the argument of the function. This is called *variable capture*. Hence, to make sure that a substitution is valid, we need to first ensure that no variables are captured. That is, in order for a substitution to be valid, we need to make sure that whenever we replace a free variable, say, $x$ with a complex differential pre-term $\sigma$ in the complex differential pre-term $\tau$, we need to ensure that no free variables of $\sigma$ also appears as the corresponding variable of an abstraction containing that instance of $x$. Specifically, we can only substitute $\sigma$ for $x$ in $\tau$ when $\mathrm{Sub}_{x,\sigma}\,\tau = 1$.

**Definition 2.2.7. Condition for Substitution to be Possible**

Given a variable $x \in \mathcal{V}$ and a simple differential pre-term $t \in \Delta_{d,\mathcal{V}}$, define the function

$\mathrm{Sub}_{x,t}\colon \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle \to \{0,1\}$ by

$$\mathrm{Sub}_{x,t}\,(D_{i_1,\dots,i_n}x * (t_1,\dots,t_n)) = \mathrm{Sub}_{x,t}\,(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,t}\,(t_n), \tag{2.2.4a}$$
$$\mathrm{Sub}_{x,t}\,(D_{i_1,\dots,i_n}y * (t_1,\dots,t_n)) = \mathrm{Sub}_{x,t}\,(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,t}\,(t_n), \tag{2.2.4b}$$
$$\mathrm{Sub}_{x,t}\,(D_1^n\lambda x.s * (t_1,\dots,t_n)) = \mathrm{Sub}_{x,t}\,(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,t}\,(t_n), \tag{2.2.4c}$$
$$\mathrm{Sub}_{x,t}\,(D_1^n\lambda y.s * (t_1,\dots,t_n)) = \begin{cases}\mathrm{Sub}_{x,t}\,(s) \wedge \mathrm{Sub}_{x,t}\,(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,t}\,(t_n) & \text{if } y \notin \mathrm{FV}(t) \vee x \notin \mathrm{FV}\,(s),\\ 0 & \text{otherwise,}\end{cases} \tag{2.2.4d}$$
$$\mathrm{Sub}_{x,t}\,(s\mu) = \mathrm{Sub}_{x,t}\,(s) \wedge \mathrm{Sub}_{x,t}\,(\mu), \tag{2.2.4e}$$
$$\mathrm{Sub}_{x,t}\,(0) = 1, \tag{2.2.4f}$$
$$\mathrm{Sub}_{x,t}\,(r\cdot\mu) = \mathrm{Sub}_{x,t}\,(\mu), \tag{2.2.4g}$$
$$\mathrm{Sub}_{x,t}\,(\mu+\nu) = \mathrm{Sub}_{x,t}\,(\mu) \wedge \mathrm{Sub}_{x,t}\,(\nu), \tag{2.2.4h}$$

where $y \in \mathcal{V}$ is a variable distinct from $x$, $r \in \mathcal{R}$ is a scalar, $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1,\dots,i_n \in \mathbb{N}$ are positive integers, $s, t_1,\dots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $\mu,\nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. ♣

We now define substitution for differential pre-terms. It will, or course, only be defined when no variable is captured as a result of the substitution.

**Definition 2.2.8. Substitution for Differential Pre-Terms**

Given a variable $x \in \mathcal{V}$, a complex differential pre-term $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$, and a simple differential pre-term $t \in \Delta_{d,\mathcal{V}}$ such that $\mathrm{Sub}_{x,t}(\tau) = 1$, the *substitution* $\tau[t/x]$ of $t$ for $x$ in $\tau$ is defined by:

$$(D_{i_1,\dots,i_n}x * (t_1,\dots,t_n))[t/x] = D_{i_1,\dots,i_n}t * (t_1[t/x],\dots,t_n[t/x]), \tag{2.2.5a}$$
$$(D_{i_1,\dots,i_n}y * (t_1,\dots,t_n))[t/x] = D_{i_1,\dots,i_n}y * (t_1[t/x],\dots,t_n[t/x]), \tag{2.2.5b}$$
$$(D_1^n\lambda x.s * (t_1,\dots,t_n))[t/x] = D_1^n\lambda x.s * (t_1[t/x],\dots,t_n[t/x]), \tag{2.2.5c}$$
$$(D_1^n\lambda y.s * (t_1,\dots,t_n))[t/x] = D_1^n\lambda y.s[t/x] * (t_1[t/x],\dots,t_n[t/x]), \tag{2.2.5d}$$
$$(s\mu)[t/x] = s[t/x]\mu[t/x], \tag{2.2.5e}$$
$$0[t/x] = 0, \tag{2.2.5f}$$
$$(r\cdot\mu)[t/x] = r\cdot\mu[t/x], \tag{2.2.5g}$$
$$(\mu+\nu)[t/x] = \mu[t/x] + \nu[t/x], \tag{2.2.5h}$$

where $y \in \mathcal{V}$ is a variable distinct from $x$, $r \in \mathcal{R}$ is a scalar, $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1,\dots,i_n \in \mathbb{N}$ are positive integers, $s, t_1,\dots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $\mu,\nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. ♣

With a solid definition of differential pre-terms, free and bound variables, and substitution, it's about time that we start proving some basic properties about these definitions.

**Lemma 2.2.9**

Let $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ be a complex differential pre-term and $t \in \Delta_{d,\mathcal{V}}$ be a simple differential pre-term. If $x \notin \mathrm{FV}\,(\tau)$, then $t$ can be substituted for $x$ in $\tau$ (i.e., $\mathrm{Sub}_{x,t}\,(\tau) = 1$) and $\tau[t/x] = \tau$. ♡

***Proof***. This follows by induction on $\tau$.

1. Assume that $\tau = D_{i_1,\dots,i_n}y * (t_1,\dots,t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $y \in \mathcal{V}$ is a variable, $i_1,\dots,i_n \in \mathbb{N}$ are positive integers, and $t_1,\dots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential

pre-terms. Since

$$x \notin \mathrm{FV}(\tau) = \mathrm{FV}(D_{i_1,\dots,i_n} y * (t_1,\dots,t_n)) = \{x\} \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n), \tag{2.2.6}$$

we have that $y \neq x$ and $x \notin \mathrm{FV}(t_j)$ for all $1 \leqslant j \leqslant n$, so by the induction hypothesis, $\mathrm{Sub}_{x,t}(t_j) = 1$ and $t_j[t/x] = t_j$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}\mathrm{Sub}_{x,t}(\tau) &= \mathrm{Sub}_{x,t}(D_{i_1,\dots,i_n} y * (t_1,\dots,t_n)) = \mathrm{Sub}_{x,t}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,t}(t_n)\\ &= 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.7}$$

$$\begin{aligned}\tau[t/x] &= (D_{i_1,\dots,i_n} y * (t_1,\dots,t_n))[t/x] = D_{i_1,\dots,i_n} y * (t_1[t/x],\dots,t_n[t/x])\\ &= D_{i_1,\dots,i_n} y * (t_1,\dots,t_n) = \tau.\end{aligned} \tag{2.2.8}$$

2. Assume that $\tau = D_1^n \lambda y.s * (t_1,\dots,t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $y \in \mathcal{V}$ is a variable, and $s, t_1, \dots t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. Since

$$x \notin \mathrm{FV}(\tau) = \mathrm{FV}(D_1^n \lambda y.s * (t_1,\dots,t_n)) = (\mathrm{FV}(s) - \{y\}) \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n), \tag{2.2.9}$$

we have that $x \notin \mathrm{FV}(s) - \{y\}$ and $x \notin \mathrm{FV}(t_j)$ for all $1 \leqslant j \leqslant n$, so by the induction hypothesis, $\mathrm{Sub}_{x,t}(t_j) = 1$ and $t_j[t/x] = t_j$ for all $1 \leqslant j \leqslant n$. If $x \in \mathrm{FV}(s)$, then $y = x$, which implies that

$$\begin{aligned}\mathrm{Sub}_{x,t}(\tau) &= \mathrm{Sub}_{x,t}(D_1^n \lambda x.s * (t_1,\dots,t_n)) = \mathrm{Sub}_{x,t}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,t}(t_n)\\ &= 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.10}$$

$$\begin{aligned}\tau[t/x] &= (D_1^n \lambda x.s * (t_1,\dots,t_n))[t/x] = D_1^n \lambda x.s * (t_1[t/x],\dots,t_n[t/x])\\ &= D_1^n \lambda x.s * (t_1,\dots,t_n) = \tau.\end{aligned} \tag{2.2.11}$$

If $x \notin \mathrm{FV}(s)$, by the induction hypothesis, $\mathrm{Sub}_{x,t}(s) = 1$ and $s[t/x] = s$, so

$$\begin{aligned}\mathrm{Sub}_{x,t}(\tau) &= \mathrm{Sub}_{x,t}(D_1^n \lambda y.s * (t_1,\dots,t_n))\\ &= \mathrm{Sub}_{x,t}(s) \wedge \mathrm{Sub}_{x,t}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,t}(t_n) = 1 \wedge 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.12}$$

$$\begin{aligned}\tau[t/x] &= (D_1^n \lambda y.s * (t_1,\dots,t_n))[t/x]\\ &= D_1^n \lambda y.s[t/x] * (t_1[t/x],\dots,t_n[t/x]) = D_1^n \lambda y.s * (t_1,\dots,t_n) = \tau.\end{aligned} \tag{2.2.13}$$

3. Assume that $\tau = s\mu$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\mu \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term. Since $x \notin \mathrm{FV}(\tau) = \mathrm{FV}(s\mu) = \mathrm{FV}(s) \cup \mathrm{FV}(\mu)$, we have that $x \notin \mathrm{FV}(s)$ and $x \notin \mathrm{FV}(\mu)$, so by the induction hypothesis, $\mathrm{Sub}_{x,t}(s) = \mathrm{Sub}_{x,t}(\mu) = 1$, $s[t/x] = s$, and $\mu[t/x] = \mu$, which implies that $\mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(s\mu) = \mathrm{Sub}_{x,t}(s) \wedge \mathrm{Sub}_{x,t}(\mu) = 1 \wedge 1 = 1$ and $\tau[t/x] = (s\mu)[t/x] = s[t/x]\mu[t/x] = s\mu = \tau$.

4. Assume that $\tau = 0$. Then, $\mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(0) = 1$ and $\tau[t/x] = 0[t/x] = 0$.

5. Assume that $\tau = r \cdot \mu$, where $r \in \mathcal{R}$ is a scalar and $\mu \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term. Since $x \notin \mathrm{FV}(\tau) = \mathrm{FV}(r \cdot \mu) = \mathrm{FV}(\mu)$, by the induction hypothesis, $\mathrm{Sub}_{x,t}(\mu) = 1$ and $\mu[t/x] = \mu$, which implies that $\mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(r \cdot \mu) = \mathrm{Sub}_{x,t}(\mu) = 1$ and $\tau[t/x] = (r \cdot \mu)[t/x] = r \cdot \mu[t/x] = r \cdot \mu = \tau$.

6. Assume that $\tau = \mu + \nu$, where $\mu, \nu \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-terms. Since $x \notin \mathrm{FV}(\tau) = \mathrm{FV}(\mu + \nu) = \mathrm{FV}(\mu) \cup \mathrm{FV}(\nu)$, we have that $x \notin \mathrm{FV}(\mu)$ and $x \notin \mathrm{FV}(\nu)$, so by the induction hypothesis, $\mathrm{Sub}_{x,t}(\mu) = \mathrm{Sub}_{x,t}(\nu) = 1$, $\mu[t/x] = \mu$, and $\nu[t/x] = \nu$, which implies that $\mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(\mu + \nu) = \mathrm{Sub}_{x,t}(\mu) \wedge \mathrm{Sub}_{x,t}(\nu) = 1 \wedge 1 = 1$ and $\tau[t/x] = (\mu + \nu)[t/x] = \mu[t/x] + \nu[t/x] = \mu + \nu = \tau$.

This completes the induction. ❑

**Lemma 2.2.10**

Let $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ be a complex differential pre-term, $t \in \Delta_{d,\mathcal{V}}$ be a simple differential pre-term, and $x \in \mathcal{V}$ be a variable such that $t$ can be substituted for $x$ in $\tau$ (i.e., $\mathrm{Sub}_{x,t}(\tau) = 1$). Then, $y \in \mathrm{FV}(\tau[t/x])$ if and only if either $y \in \mathrm{FV}(\tau)$ and $x \neq y$ or $y \in \mathrm{FV}(t)$ and $x \in \mathrm{FV}(\tau)$. ♡

***Proof.*** This follows by induction on $\tau$. The details are left to the reader. ❏

**Lemma 2.2.11**

Given a complex differential pre-term $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ and a variable $x \in \mathcal{V}$, $x$ can always be substituted for $x$ in $\tau$ (i.e., $\mathrm{Sub}_{x,x}(\tau) = 1$) and $\tau[x/x] = \tau$. ♡

***Proof.*** This follows by induction on $\tau$.

1. Assume that $\tau = D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $y \in \mathcal{V}$ is a variable, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, and $t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. Note that $y[x/x] = y$ if $y \neq x$ and $y[x/x] = x[x/x] = x = y$ if $y = x$. Also, by the induction hypothesis, $\mathrm{Sub}_{x,x}(t_j) = 1$ and $t_j[x/x] = t_j$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}\mathrm{Sub}_{x,x}(\tau) &= \mathrm{Sub}_{x,x}(D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n)) = \mathrm{Sub}_{x,x}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,x}(t_n) \\ &= 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.14}$$

$$\begin{aligned}\tau[x/x] &= (D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n))[x/x] = D_{i_1,\ldots,i_n} y[x/x] * (t_1[x/x], \ldots, t_n[x/x]) \\ &= D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n) = \tau.\end{aligned} \tag{2.2.15}$$

2. Assume that $\tau = D_1^n \lambda y.s * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $y \in \mathcal{V}$ is a variable, and $s, t_1, \ldots t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. By the induction hypothesis, $\mathrm{Sub}_{x,x}(s) = 1$ and $s[x/x] = s$, and for all $1 \leqslant j \leqslant n$, $\mathrm{Sub}_{x,x}(t_j) = 1$ and $t_j[x/x] = t_j$. If $y = x$, then

$$\begin{aligned}\mathrm{Sub}_{x,x}(\tau) &= \mathrm{Sub}_{x,x}(D_1^n \lambda x.s * (t_1, \ldots, t_n)) = \mathrm{Sub}_{x,x}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,x}(t_n) \\ &= 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.16}$$

$$\begin{aligned}\tau[x/x] &= (D_1^n \lambda x.s * (t_1, \ldots, t_n))[x/x] = D_1^n \lambda x.s * (t_1[x/x], \ldots, t_n[x/x]) \\ &= D_1^n \lambda x.s * (t_1, \ldots, t_n) = \tau.\end{aligned} \tag{2.2.17}$$

   If $y \neq x$, then $y \notin \{x\} = \mathrm{FV}(x)$, which implies that

$$\begin{aligned}\mathrm{Sub}_{x,x}(\tau) &= \mathrm{Sub}_{x,x}(D_1^n \lambda y.s * (t_1, \ldots, t_n)) = \mathrm{Sub}_{x,x}(s) \wedge \mathrm{Sub}_{x,x}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,x}(t_n) \\ &= 1 \wedge 1 \wedge \cdots \wedge 1 = 1, \\ \tau[x/x] &= (D_1^n \lambda y.s * (t_1, \ldots, t_n))[x/x] = D_1^n \lambda y.s[x/x] * (t_1[x/x], \ldots, t_n[x/x]) \\ &= D_1^n \lambda y.s * (t_1, \ldots, t_n) = \tau.\end{aligned}$$

3. Assume that $\tau = s\mu$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. By the induction hypothesis, $\mathrm{Sub}_{x,x}(s) = \mathrm{Sub}_{x,x}(\mu) = 1$, $s[x/x] = s$, and $\mu[x/x] = \mu$, which implies that $\mathrm{Sub}_{x,x}(\tau) = \mathrm{Sub}_{x,x}(s\mu) = \mathrm{Sub}_{x,x}(s) \wedge \mathrm{Sub}_{x,x}(\mu) = 1 \wedge 1 = 1$ and $\tau[x/x] = (s\mu)[x/x] = s[x/x]\mu[x/x] = s\mu = \tau$.

4. Assume that $\tau = 0$. Then, $\mathrm{Sub}_{x,x}(\tau) = \mathrm{Sub}_{x,x}(0) = 1$ and $\tau[x/x] = 0[x/x] = 0$.

5. Assume that $\tau = r \cdot \mu$, where $r \in \mathcal{R}$ is a scalar and $\mu \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term. By the induction hypothesis, $\mathrm{Sub}_{x,x}(\mu) = 1$ and $\mu[x/x] = \mu$, which implies that $\mathrm{Sub}_{x,x}(\tau) = \mathrm{Sub}_{x,x}(r \cdot \mu) = \mathrm{Sub}_{x,x}(\mu) = 1$ and $\tau[x/x] = (r \cdot \mu)[x/x] = r \cdot \mu[x/x] = r \cdot \mu = \tau$.

6. Assume that $\tau = \mu + \nu$, where $\mu, \nu \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-terms. By the induction hypothesis, $\mathrm{Sub}_{x,x}(\mu) = \mathrm{Sub}_{x,x}(\nu) = 1$, $\mu[x/x] = \mu$, and $\nu[x/x] = \nu$, which implies that $\mathrm{Sub}_{x,x}(\tau) = \mathrm{Sub}_{x,x}(\mu + \nu) = \mathrm{Sub}_{x,x}(\mu) \wedge \mathrm{Sub}_{x,x}(\nu) = 1 \wedge 1 = 1$ and $\tau[x/x] = (\mu + \nu)[x/x] = \mu[x/x] + \nu[x/x] = \mu + \nu = \tau$.

This ccompletes the induction. ❏

**Lemma 2.2.12**

Let $\tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ be a complex differential pre-term, let $u, v \in \Delta_{d,\mathcal{V}}$ be simple differential pre-terms, and $x, y \in \mathcal{V}$ such that $x \neq y$ and $\mathrm{Sub}_{x,u}(\tau) = \mathrm{Sub}_{y,v}(\tau[u/x]) = \mathrm{Sub}_{y,v}(u) = 1$. If $x \notin \mathrm{FV}(v)$ or $y \notin \mathrm{FV}(\tau)$, then $\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = 1$ and $\tau[u/x][v/y] = \tau[v/y][u[v/y]/x]$. ♡

***Proof.*** This follows by induction on $\tau$.

1. Assume that $\tau = x$. Trivially, $\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(x) = 1$, and since $x \neq y$, $\tau[v/y] = x[v/y] = x$, so $\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = \mathrm{Sub}_{x,u[v/y]}(x) = 1$. Then,

$$\begin{aligned} \tau[u/x][v/y] &= x[u/x][v/y] = u[v/y] = x[u[v/y]/x] \\ &= x[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x]. \end{aligned} \tag{2.2.18}$$

2. Assume that $\tau = y$. Then, $\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(y) = 1$, and since $y \in \{y\} = \mathrm{FV}(y) = \mathrm{FV}(\tau)$, we must have that $x \notin \mathrm{FV}(v) = \mathrm{FV}(y[v/y]) = \mathrm{FV}(\tau[v/y])$, so by Lemma 2.2.9, $\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = 1$ and

$$\tau[v/y][u[v/y]/x] = \tau[v/y] = y[v/y] = y[u/x][v/y] = \tau[u/x][v/y]. \tag{2.2.19}$$

3. Assume that $\tau = z$, where $z \in \mathcal{V}$ is a variable distinct form $x$ and $y$. Then,

$$\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(z) = 1, \tag{2.2.20}$$

$$\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = \mathrm{Sub}_{x,u[v/y]}(z[v/y]) = \mathrm{Sub}_{x,u[v/y]}(z) = 1, \tag{2.2.21}$$

$$\begin{aligned} \tau[u/x][v/y] &= z[u/x][v/y] = z[v/y] = z = z[u[v/y]/x] \\ &= z[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x]. \end{aligned} \tag{2.2.22}$$

4. Assume that $\tau = \lambda x.s$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term. Since $\tau[u/x] = (\lambda x.s)[u/x] = \lambda x.s = \tau$ and $\mathrm{Sub}_{y,v}(\tau[u/x]) = 1$, we have that $\mathrm{Sub}_{y,v}(\tau) = 1$, which implies that $v$ can be substituted for $y$ in $\tau$. Thus, $\tau[v/y] = (\lambda x.s)[v/y] = \lambda x.s[v/y]$, which tells us that $\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = \mathrm{Sub}_{x,u[v/y]}(\lambda x.s[v/y]) = 1$. Putting this all together,

$$\begin{aligned} \tau[u/x][v/y] &= (\lambda x.s)[u/x][v/y] = (\lambda x.s)[v/y] = (\lambda x.s[v/y]) = (\lambda x.s[v/y])[u[v/y]/x] \\ &= (\lambda x.s)[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x]. \end{aligned} \tag{2.2.23}$$

5. Assume that $\tau = \lambda y.s$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term. Then, $\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(\lambda y.s) = 1$. Since $y \neq x$ and $1 = \mathrm{Sub}_{x,u}(\tau) = \mathrm{Sub}_{x,u}(\lambda y.s)$, which implies that $y \notin \mathrm{FV}(u)$ or $x \notin \mathrm{FV}(s)$. If $x \notin \mathrm{FV}(s)$, then $\mathrm{Sub}_{x,u[v,y]}(\tau[v/y]) = \mathrm{Sub}_{x,u[v,y]}((\lambda y.s)[v/y]) = \mathrm{Sub}_{x,u[v,y]}(\lambda y.s) = 1$ and

$$\begin{aligned}\tau[u/x][v/y] &= (\lambda y.s)[u/x][v/y] = (\lambda y.s[u/x])[v/y] = \lambda y.s[u/x] = \lambda y.s \\ &= (\lambda y.s)[u[v/y]/x] = (\lambda y.s)[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x].\end{aligned} \tag{2.2.24}$$

If $y \notin \mathrm{FV}(u)$, then by Lemma 2.2.9, $u[v/x] = u$, which implies that $y \notin \mathrm{FV}(u) = \mathrm{FV}(u[v/x])$, whence $\mathrm{Sub}_{x,u[v,y]}(\tau[v/y]) = \mathrm{Sub}_{x,u[v,y]}(\lambda y.s) = 1$ and

$$\begin{aligned}\tau[u/x][v/y] &= \lambda y.s[u/x] = (\lambda y.s)[u/x] = (\lambda y.s)[u[v/y]/x] \\ &= (\lambda y.s)[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x].\end{aligned} \tag{2.2.25}$$

6. Assume that $\tau = \lambda z.s$, where $z \in \mathcal{V}$ is a variable distinct from $x$ and $y$ and $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term. Let us first show that $\mathrm{Sub}_{y,v}(\tau) = 1$. Since $z \neq y$ and $1 = \mathrm{Sub}_{y,v}(\tau[u/x]) = \mathrm{Sub}_{y,v}((\lambda z.s)[u/x]) = \mathrm{Sub}_{y,v}(\lambda z.s[u/x])$, we have that $z \notin \mathrm{FV}(v)$ or $y \notin \mathrm{FV}(s[u/x])$. If $z \in \mathrm{FV}(v)$, this implies that $y \notin \mathrm{FV}(s[u/x]) \subset (\mathrm{FV}(s) - \{x\}) \cup \mathrm{FV}(u)$, where we used Corollary 2.2.10 for the inclusion, so since $y \neq x$, this tells us that $y \notin \mathrm{FV}(s)$. Thus, $z \notin \mathrm{FV}(v)$ or $y \notin \mathrm{FV}(s)$, which implies that $\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(\lambda z.s) = 1$. Now, let us show that $\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = 1$. Suppose that $z \in \mathrm{FV}(u[v/y])$ and $x \in \mathrm{FV}(\tau[v/y])$. Further, suppose that $x \notin \mathrm{FV}(v)$. By Corollary 2.2.10, $\mathrm{FV}(\tau[v/y]) \subset (\mathrm{FV}(\tau) - y) \cup \mathrm{FV}(v)$ . Since $x \neq y$ and $x \notin \mathrm{FV}(v)$, this implies that $x \in \mathrm{FV}(\tau)$. Instead, suppose that $x \in \mathrm{FV}(v)$. Since $x \notin \mathrm{FV}(v)$ or $y \notin \mathrm{FV}(\tau)$ by hypothesis, this implies that $y \notin \mathrm{FV}(\tau)$, so by Lemma 2.2.9, $\tau[v/y] = \tau$. Hence $x \in \mathrm{FV}(\tau[v/y]) = \mathrm{FV}(\tau)$. Therefore, regardless of whether $x$ is in $\mathrm{FV}(v)$, we have that $x \in \mathrm{FV}(\tau) = \mathrm{FV}(\lambda z.s) = \mathrm{FV}(s) - \{z\}$, which implies that $x \in \mathrm{FV}(s)$, as $z \neq z$. Since $1 = \mathrm{Sub}_{x,u}(\tau) = \mathrm{Sub}_{x,u}(\lambda z\ s)$ by hypothesis and $z \neq x$, we have that $z \notin \mathrm{FV}(u)$ or $x \notin \mathrm{FV}(s)$. Since we know that $x \in \mathrm{FV}(s)$, this implies that $z \notin \mathrm{FV}(u)$. Since $z \in \mathrm{FV}(u[v/y])$, by Lemma 2.2.10, this implies that $y \in \mathrm{FV}(u)$ and $z \in \mathrm{FV}(v)$. Since $x \in \mathrm{FV}(s)$ and $y \in \mathrm{FV}(u)$, by Lemma 2.2.10, $y \in \mathrm{FV}(s[u/x])$. Recall from earlier that $z \notin \mathrm{FV}(v)$ or $y \notin \mathrm{FV}(s[u/x])$, so the prior sentence implies that $z \notin \mathrm{FV}(v)$, contradicting the fact that $z \in \mathrm{FV}(v)$. Hence, $z \notin \mathrm{FV}(u[v/y])$ or $x \notin \mathrm{FV}(\tau[v/y])$, which implies that $\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = 1$. Now, let us show that $\tau[u/x][v/y] = \tau[v/y][u[v/y]/x]$. Since $1 = \mathrm{Sub}_{x,u}(\tau) = \mathrm{Sub}_{x,u}(\lambda z.s)$, we must have that $\mathrm{Sub}_{x,u}(s) = 1$, and since $1 = \mathrm{Sub}_{y,v}(\tau[u/x]) = \mathrm{Sub}_{y,v}((\lambda z.s)[u/x]) = \mathrm{Sub}_{y,v}(\lambda z.s[u/x])$, we must have that $\mathrm{Sub}_{y,v}(s[u/x]) = 1$, so by the induction hypothesis, $s[u/x][v/y] = s[v/y][u[v/y]/x]$, which implies that

$$\begin{aligned}\tau[u/x][v/y] &= (\lambda z.s)[u/x][v/y] = \lambda z.s[u/x][v/y] = \lambda z.s[v/y][u[v/y]/x] \\ &= (\lambda z.s)[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x].\end{aligned} \tag{2.2.26}$$

7. Assume that $\tau = D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $z \in \mathcal{V}$ is a variable, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, and $t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. By hypothesis,

$$\begin{aligned}1 &= \mathrm{Sub}_{x,u}(\tau) = \mathrm{Sub}_{x,u}(D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)) \\ &= \mathrm{Sub}_{x,u}(z) \wedge \mathrm{Sub}_{x,u}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,u}(t_n),\end{aligned} \tag{2.2.27}$$

$$\begin{aligned}1 &= \mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = \mathrm{Sub}_{x,u[v/y]}((D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n))[v/y]) \\ &= \mathrm{Sub}_{x,u[v/y]}((D_{i_1,\ldots,i_n} z[v/y] * (t_1[v/y], \ldots, t_n[v/y]))) \\ &= \mathrm{Sub}_{x,u[v/y]}(z[v/y]) \wedge \mathrm{Sub}_{x,u[v/y]}(t_1[v/y]) \wedge \cdots \wedge \mathrm{Sub}_{x,u[v/y]}(t_n[v/y]),\end{aligned} \tag{2.2.28}$$

which implies that $\mathrm{Sub}_{x,u}(z) = \mathrm{Sub}_{x,u[v/y]}(z[v/y]) = \mathrm{Sub}_{x,u}(t_j) = \mathrm{Sub}_{x,u[v/y]}(t_j[v/y]) = 1$ for all $1 \leqslant j \leqslant n$, so by the induction hypothesis, $\mathrm{Sub}_{y,v}(z) = \mathrm{Sub}_{x,u[v/y]}(z[v/y]) = 1$ and $z[u/x][v/y] = z[v/y][u[v/y]/x]$, and for all $1 \leqslant j \leqslant n$, $\mathrm{Sub}_{y,v}(t_j) = \mathrm{Sub}_{x,u[v/y]}(t_j[v/y]) = 1$ and $t_j[u/x][v/y] = t_j[v/y][u[v/y]/x]$. Hence,

$$\begin{aligned} \mathrm{Sub}_{y,v}(\tau) &= \mathrm{Sub}_{y,v}(D_{i_1,\ldots,i_n} z * (t_1,\ldots,t_n)) \\ &= \mathrm{Sub}_{y,v}(z) \wedge \mathrm{Sub}_{y,v}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{y,v}(t_n) = 1 \wedge 1 \wedge \cdots \wedge 1 = 1, \end{aligned} \tag{2.2.29}$$

$$\begin{aligned} \mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) &= \mathrm{Sub}_{x,u[v/y]}((D_{i_1,\ldots,i_n} z * (t_1,\ldots,t_n))[v/y]) \\ &= \mathrm{Sub}_{x,u[v/y]}(D_{i_1,\ldots,i_n} z[v/y] * (t_1[v/y],\ldots,t_n[v/y])) \\ &= \mathrm{Sub}_{x,u[v/y]}(z[v/y]) \wedge \mathrm{Sub}_{x,u[v/y]}(t_1[v/y]) \wedge \cdots \wedge \mathrm{Sub}_{x,u[v/y]}(t_n[v/y]) \\ &= 1 \wedge 1 \wedge \cdots \wedge 1 = 1, \end{aligned} \tag{2.2.30}$$

$$\begin{aligned} \tau[u/x][v/y] &= (D_{i_1,\ldots,i_n} z * (t_1,\ldots,t_n))[u/x][v/y] \\ &= (D_{i_1,\ldots,i_n} z[u/x][v/y] * (t_1[u/x][v/y],\ldots,t_n[u/x][v/y])) \\ &= (D_{i_1,\ldots,i_n} z[v/y][u[v/y]/x] * (t_1[v/y][u[v/y]/x],\ldots,t_n[v/y][u[v/y]/x])) \\ &= (D_{i_1,\ldots,i_n} z * (t_1,\ldots,t_n))[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x]. \end{aligned} \tag{2.2.31}$$

8. Assume that $\tau = D_1^n \lambda z.s * (t_1,\ldots,t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $z \in \mathcal{V}$ is a variable, and $s, t_1, \ldots t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. By hypothesis,

$$\begin{aligned} 1 = \mathrm{Sub}_{x,u}(\tau) &= \mathrm{Sub}_{x,u}(D_1^n \lambda z.s * (t_1,\ldots,t_n)) \\ &= \mathrm{Sub}_{x,u}(\lambda z.s) \wedge \mathrm{Sub}_{x,u}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,u}(t_n), \end{aligned} \tag{2.2.32}$$

$$\begin{aligned} 1 = \mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) &= \mathrm{Sub}_{x,u[v/y]}((D_1^n \lambda z.s * (t_1,\ldots,t_n))[v/y]) \\ &= \mathrm{Sub}_{x,u[v/y]}((D_1^n (\lambda z.s)[v/y] * (t_1[v/y],\ldots,t_n[v/y]))) \\ &= \mathrm{Sub}_{x,u[v/y]}((\lambda z.s)[v/y]) \wedge \mathrm{Sub}_{x,u[v/y]}(t_1[v/y]) \wedge \cdots \wedge \mathrm{Sub}_{x,u[v/y]}(t_n[v/y]), \end{aligned} \tag{2.2.33}$$

so $\mathrm{Sub}_{x,u}(\lambda z.s) = \mathrm{Sub}_{x,u[v/y]}((\lambda z.s)[v/y]) = \mathrm{Sub}_{x,u}(t_j) = \mathrm{Sub}_{x,u[v/y]}(t_j[v/y]) = 1$ for all $1 \leqslant j \leqslant n$, which implies that by the induction hypothesis, we have that $\mathrm{Sub}_{y,v}(\lambda z.s) = \mathrm{Sub}_{x,u[v/y]}((\lambda z.s)[v/y]) = 1$ and $(\lambda z.s)[u/x][v/y] = (\lambda z.s)[v/y][u[v/y]/x]$, and for all $1 \leqslant j \leqslant n$, $\mathrm{Sub}_{y,v}(t_j) = \mathrm{Sub}_{x,u[v/y]}(t_j[v/y]) = 1$ and $t_j[u/x][v/y] = t_j[v/y][u[v/y]/x]$. Hence,

$$\begin{aligned} \mathrm{Sub}_{y,v}(\tau) &= \mathrm{Sub}_{y,v}(D_1^n \lambda z.s * (t_1,\ldots,t_n)) \\ &= \mathrm{Sub}_{y,v}(\lambda z.s) \wedge \mathrm{Sub}_{y,v}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{y,v}(t_n) = 1 \wedge 1 \wedge \cdots \wedge 1 = 1, \end{aligned} \tag{2.2.34}$$

$$\begin{aligned} \mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) &= \mathrm{Sub}_{x,u[v/y]}((D_1^n \lambda z.s * (t_1,\ldots,t_n))[v/y]) \\ &= \mathrm{Sub}_{x,u[v/y]}(D_1^n (\lambda z.s)[v/y] * (t_1[v/y],\ldots,t_n[v/y])) \\ &= \mathrm{Sub}_{x,u[v/y]}((\lambda z.s)[v/y]) \wedge \mathrm{Sub}_{x,u[v/y]}(t_1[v/y]) \wedge \cdots \wedge \mathrm{Sub}_{x,u[v/y]}(t_n[v/y]) \\ &= 1 \wedge 1 \wedge \cdots \wedge 1 = 1, \end{aligned} \tag{2.2.35}$$

$$\begin{aligned} \tau[u/x][v/y] &= (D_1^n \lambda z.s * (t_1,\ldots,t_n))[u/x][v/y] \\ &= (D_1^n (\lambda z.s)[u/x][v/y] * (t_1[u/x][v/y],\ldots,t_n[u/x][v/y])) \\ &= (D_1^n (\lambda z.s)[v/y][u[v/y]/x] * (t_1[v/y][u[v/y]/x],\ldots,t_n[v/y][u[v/y]/x])) \\ &= (D_1^n \lambda z.s * (t_1,\ldots,t_n))[u[v/y]/x] = \tau[v/y][u[v/y]/x]. \end{aligned} \tag{2.2.36}$$

9. Assume that $\tau = s\mu$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. By hypothesis,

$$1 = \mathrm{Sub}_{x,u}(\tau) = \mathrm{Sub}_{x,u}(s\mu) = \mathrm{Sub}_{x,u}(s) \wedge \mathrm{Sub}_{x,u}(\mu), \tag{2.2.37}$$

$$\begin{aligned} 1 = \mathrm{Sub}_{y,v}(\tau[u/x]) &= \mathrm{Sub}_{y,v}((s\mu)[u/x]) = \mathrm{Sub}_{y,v}(s[u/x]\mu[u/s]) \\ &= \mathrm{Sub}_{y,v}(s[u/x]) \wedge \mathrm{Sub}_{y,v}(\mu[u/s]), \end{aligned} \tag{2.2.38}$$

so $\mathrm{Sub}_{x,u}(s) = \mathrm{Sub}_{y,v}(s[u/x]) = 1$ and $\mathrm{Sub}_{x,u}(\mu) = \mathrm{Sub}_{y,v}(\mu[u/s]) = 1$, which implies by the induction hypothesis, $\mathrm{Sub}_{y,v}(s) = \mathrm{Sub}_{x,u[v/y]}(s[v/y]) = 1$, $s[u/x][v/y] = s[v/y][u[v/y]/x]$, $\mathrm{Sub}_{y,v}(\mu) = \mathrm{Sub}_{x,u[v/y]}(\mu[v/y]) = 1$, and $\mu[u/x][v/y] = \mu[v/y][u[v/y]/x]$. Hence,

$$\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(s\mu) = \mathrm{Sub}_{y,v}(s) \wedge \mathrm{Sub}_{y,v}(\mu) = 1 \wedge 1 = 1, \tag{2.2.39}$$

$$\begin{aligned}\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) &= \mathrm{Sub}_{x,u[v/y]}((s\mu)[v/y]) = \mathrm{Sub}_{x,u[v/y]}(s[v/y])\mu[v/y]) \\ &= \mathrm{Sub}_{x,u[v/y]}(s[v/y])) \wedge \mathrm{Sub}_{x,u[v/y]}(\mu[v/y]) = 1 \wedge 1 = 1,\end{aligned} \tag{2.2.40}$$

$$\begin{aligned}\tau[u/x][v/y] &= (s\mu)[u/x][v/y] = s[u/x][v/y]\mu[u/x][v/y] \\ &= s[v/y][u[v/y]/x]\mu[v/y][u[v/y]/x] = (s\mu)[v/y][u[v/y]/x] \\ &= \tau[v/y][u[v/y]/x].\end{aligned} \tag{2.2.41}$$

10. Assume that $\tau = 0$. Then,

$$\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(0) = 1, \tag{2.2.42}$$

$$\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) = \mathrm{Sub}_{x,u[v/y]}(0[v/y]) = \mathrm{Sub}_{x,u[v/y]}(0) = 1, \tag{2.2.43}$$

$$\tau[u/x][v/y] = 0[u/x][v/y] = 0 = 0[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x]. \tag{2.2.44}$$

11. Assume that $\tau = r \cdot \mu$, where $r \in \mathcal{R}$ is a scalar and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. By hypothesis,

$$1 = \mathrm{Sub}_{x,u}(\tau) = \mathrm{Sub}_{x,u}(r \cdot \mu) = \mathrm{Sub}_{x,u}(\mu), \tag{2.2.45}$$

$$1 = \mathrm{Sub}_{y,v}(\tau[u/x]) = \mathrm{Sub}_{y,v}((r \cdot \mu)[u/x]) = \mathrm{Sub}_{y,v}(r \cdot \mu[u/s]) = \mathrm{Sub}_{y,v}(\mu[u/s]), \tag{2.2.46}$$

which implies that $\mathrm{Sub}_{x,u}(\mu) = \mathrm{Sub}_{y,v}(\mu[u/s]) = 1$, so by the induction hypothesis, we have that $\mathrm{Sub}_{y,v}(\mu) = \mathrm{Sub}_{x,u[v/y]}(\mu[v/y]) = 1$ and $\mu[u/x][v/y] = \mu[v/y][u[v/y]/x]$. Hence,

$$\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(r \cdot \mu) = \mathrm{Sub}_{y,v}(\mu) = 1, \tag{2.2.47}$$

$$\begin{aligned}\mathrm{Sub}_{x,u[v/y]}(\tau[v/y]) &= \mathrm{Sub}_{x,u[v/y]}((r \cdot \mu)[v/y]) = \mathrm{Sub}_{x,u[v/y]}(r \cdot \mu[v/y]) \\ &= \mathrm{Sub}_{x,u[v/y]}(\mu[v/y]) = 1,\end{aligned} \tag{2.2.48}$$

$$\begin{aligned}\tau[u/x][v/y] &= (r \cdot \mu)[u/x][v/y] = r \cdot \mu[u/x][v/y] = r \cdot \mu[v/y][u[v/y]/x] \\ &= (r \cdot \mu)[v/y][u[v/y]/x] = \tau[v/y][u[v/y]/x].\end{aligned} \tag{2.2.49}$$

12. Assume that $\tau = \mu + \nu$, where $\mu, \nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. By hypothesis,

$$1 = \mathrm{Sub}_{x,u}(\tau) = \mathrm{Sub}_{x,u}(\mu + \nu) = \mathrm{Sub}_{x,u}(\mu) \wedge \mathrm{Sub}_{x,u}(\nu), \tag{2.2.50}$$

$$1 = \mathrm{Sub}_{y,v}(\tau[u/x]) = \mathrm{Sub}_{y,v}((\mu + \nu)[u/x]) = \mathrm{Sub}_{y,v}(\mu[u/x] + \nu[u/s]) \tag{2.2.51}$$

$$= \mathrm{Sub}_{y,v}(\mu[u/x]) \wedge \mathrm{Sub}_{y,v}(\nu[u/s]), \tag{2.2.52}$$

which implies that $\mathrm{Sub}_{x,u}(\mu) = \mathrm{Sub}_{y,v}(\mu[u/x]) = 1$ and $\mathrm{Sub}_{x,u}(\nu) = \mathrm{Sub}_{y,v}(\nu[u/s]) = 1$, so by the induction hypothesis, $\mathrm{Sub}_{y,v}(\mu) = \mathrm{Sub}_{x,u[v/y]}(mu[v/y]) = 1$, $\mu[u/x][v/y] = \mu[v/y][u[v/y]/x]$, $\mathrm{Sub}_{y,v}(\nu) = \mathrm{Sub}_{x,u[v/y]}(\nu[v/y]) = 1$, and $\nu[u/x][v/y] = \nu[v/y][u[v/y]/x]$. Hence,

$$\mathrm{Sub}_{y,v}(\tau) = \mathrm{Sub}_{y,v}(\mu + \nu) = \mathrm{Sub}_{y,v}(\mu) \wedge \mathrm{Sub}_{y,v}(\nu) = 1 \wedge 1 = 1, \tag{2.2.53}$$

$$\begin{aligned}\mathrm{Sub}_{x,u[v/y]}\left(\tau[v/y]\right) &= \mathrm{Sub}_{x,u[v/y]}\left((\mu+\nu)[v/y]\right) = \mathrm{Sub}_{x,u[v/y]}\left(\mu[v/y]\right) + \nu[v/y]\right) \\ &= \mathrm{Sub}_{x,u[v/y]}\left(\mu[v/y]\right)\right) \wedge \mathrm{Sub}_{x,u[v/y]}\left(\nu[v/y]\right) = 1 \wedge 1 = 1,\end{aligned} \tag{2.2.54}$$

$$\begin{aligned}\tau[u/x][v/y] &= (\mu+\nu)[u/x][v/y] = \mu[u/x][v/y] + \nu[u/x][v/y] \\ &= \mu[v/y][u[v/y]/x] + \nu[v/y][u[v/y]/x] = (\mu+\nu)[v/y][u[v/y]/x] \\ &= \tau[v/y][u[v/y]/x].\end{aligned} \tag{2.2.55}$$

This completes the induction. ❑

**Lemma 2.2.13**

Let $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ be a complex differential pre-term and let $x, y \in \mathcal{V}$ be variables. If $y$ can be substituted for $x$ in $\tau$ (i.e., $\mathrm{Sub}_{x,y}(\tau) = 1$) and $y \notin \mathrm{FV}(\tau)$, then $x$ can be substituted for $y$ in $\tau[y/x]$ (i.e., $\mathrm{Sub}_{y,x}(\tau[y/x]) = 1$) and $\tau[y/x][x/y] = \tau$. ♡

***Proof.*** We start by proving that $\mathrm{Sub}_{y,x}(\tau[y/x]) = 1$, which we can do via induction on $\tau$.

1. Assume that $\tau = D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $z \in \mathcal{V}$ is a variable, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, and $t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. By hypothesis,

$$1 = \mathrm{Sub}_{x,y}(\tau) = \mathrm{Sub}_{x,y}\left(D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)\right) = \mathrm{Sub}_{x,y}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,y}(t_n), \tag{2.2.56}$$

$$y \notin \mathrm{FV}(\tau) = \mathrm{FV}\left(D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)\right) = \{z\} \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n), \tag{2.2.57}$$

which implies that $\mathrm{Sub}_{x,y}(t_j) = 1$ and $y \notin \mathrm{FV}(t_j)$ for all $1 \leqslant j \leqslant n$, so by the induction hypothesis, $\mathrm{Sub}_{y,x}(t_j[y/x]) = 1$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}\mathrm{Sub}_{y,x}(\tau[y/x]) &= \mathrm{Sub}_{y,x}\left((D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n))[y/x]\right) \\ &= \mathrm{Sub}_{y,x}\left(D_{i_1,\ldots,i_n} z[y/x] * (t_1[y/x], \ldots, t_n[y/x])\right) \\ &= \mathrm{Sub}_{y,x}(t_1[y/x]) \wedge \cdots \wedge \mathrm{Sub}_{y,x}(t_n[y/x]) = 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.58}$$

where we used the fact that $z[y/x]$ is a variable for the third equality.

2. Assume that $\tau = D_1^n \lambda z.s * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $z \in \mathcal{V}$ is a variable, and $s, t_1, \ldots t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. Then,

$$y \notin \mathrm{FV}(\tau) = \mathrm{FV}\left(D_1^n \lambda z.s * (t_1, \ldots, t_n)\right) = (\mathrm{FV}(s) - \{z\}) \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n), \tag{2.2.59}$$

which implies that $y \notin \mathrm{FV}(s) - \{z\}$ and $y \notin \mathrm{FV}(t_j)$ for all $1 \leqslant j \leqslant n$. Suppose that $z = x$. Then,

$$1 = \mathrm{Sub}_{x,y}(\tau) = \mathrm{Sub}_{x,y}\left(D_1^n \lambda x.s * (t_1, \ldots, t_n)\right) = \mathrm{Sub}_{x,y}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,y}(t_n), \tag{2.2.60}$$

which implies that $\mathrm{Sub}_{x,y}(t_j) = 1$ for all $1 \leqslant j \leqslant n$. By the induction hypothesis, this implies that $\mathrm{Sub}_{y,x}(t_j[y/x]) = 1$ for all $1 \leqslant j \leqslant n$. Suppose, also, that $x \neq y$. Since $y \notin \mathrm{FV}(s) - \{z\}$ and $y \neq x = z$, $y \notin \mathrm{FV}(s)$, so by Lemma 2.2.9, $\mathrm{Sub}_{y,x}(s) = 1$. Hence.

$$\begin{aligned}\mathrm{Sub}_{y,x}(\tau[y/x]) &= \mathrm{Sub}_{y,x}\left((D_1^n \lambda x.s * (t_1, \ldots, t_n))[y/x]\right) \\ &= \mathrm{Sub}_{y,x}\left(D_1^n (\lambda x.s)[y/x] * (t_1[y/x], \ldots, t_n[y/x])\right) \\ &= \mathrm{Sub}_{y,x}\left(D_1^n \lambda x.s * (t_1[y/x], \ldots, t_n[y/x])\right) \\ &= \mathrm{Sub}_{y,x}(s) \wedge \mathrm{Sub}_{y,x}(t_1[y/x]) \wedge \cdots \wedge \mathrm{Sub}_{y,x}(t_n[y/x]) \\ &= 1 \wedge 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.61}$$

where we used the fact that $y \notin \mathrm{FV}(s)$ for the fourth equality. Instead, suppose that $x = y$. Then,

$$\begin{aligned}\mathrm{Sub}_{y,x}(\tau[y/x]) &= \mathrm{Sub}_{x,x}((D_1^n \lambda x.s * (t_1, \ldots, t_n))[y/x]) \\ &= \mathrm{Sub}_{x,x}(D_1^n (\lambda x.s)[y/x] * (t_1[y/x], \ldots, t_n[y/x])) \\ &= \mathrm{Sub}_{x,x}(D_1^n \lambda x.s * (t_1[y/x], \ldots, t_n[y/x])) \\ &= \mathrm{Sub}_{y,x}(t_1[y/x]) \wedge \cdots \wedge \mathrm{Sub}_{y,x}(t_n[y/x]) = 1 \wedge \cdots \wedge 1 = 1.\end{aligned} \tag{2.2.62}$$

Now, suppose that $z \neq x$. Then,

$$\begin{aligned}1 = \mathrm{Sub}_{x,y}(\tau) &= \mathrm{Sub}_{x,y}(D_1^n \lambda z.s * (t_1, \ldots, t_n)) \\ &= \mathrm{Sub}_{x,y}(s) \wedge \mathrm{Sub}_{x,y}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,y}(t_n),\end{aligned} \tag{2.2.63}$$

which implies that $\mathrm{Sub}_{x,y}(s) = 1$ and $\mathrm{Sub}_{x,y}(t_j) = 1$ for all $1 \leqslant j \leqslant n$. By the induction hypothesis, this implies that $\mathrm{Sub}_{y,x}(s[y/x]) = 1$ and $\mathrm{Sub}_{y,x}(t_j[y/x]) = 1$ for all $1 \leqslant j \leqslant n$. If $z = y$,

$$\begin{aligned}\mathrm{Sub}_{y,x}(\tau[y/x]) &= \mathrm{Sub}_{z,x}((D_1^n \lambda z.s * (t_1, \ldots, t_n))[y/x]) \\ &= \mathrm{Sub}_{z,x}(D_1^n (\lambda z.s)[y/x] * (t_1[y/x], \ldots, t_n[y/x])) \\ &= \mathrm{Sub}_{z,x}(D_1^n \lambda z.s[y/x] * (t_1[y/x], \ldots, t_n[y/x])) \\ &= \mathrm{Sub}_{y,x}(t_1[y/x]) \wedge \cdots \wedge \mathrm{Sub}_{y,x}(t_n[y/x]) = 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.64}$$

and if $z \neq y$,

$$\begin{aligned}\mathrm{Sub}_{y,x}(\tau[y/x]) &= \mathrm{Sub}_{y,x}((D_1^n \lambda z.s * (t_1, \ldots, t_n))[y/x]) \\ &= \mathrm{Sub}_{y,x}(D_1^n (\lambda z.s)[y/x] * (t_1[y/x], \ldots, t_n[y/x])) \\ &= \mathrm{Sub}_{y,x}(D_1^n \lambda z.s[y/x] * (t_1[y/x], \ldots, t_n[y/x])) \\ &= \mathrm{Sub}_{y,x}(s[y/x]) \wedge \mathrm{Sub}_{y,x}(t_1[y/x]) \wedge \cdots \wedge \mathrm{Sub}_{y,x}(t_n[y/x]) \\ &= 1 \wedge 1 \wedge \cdots \wedge 1 = 1,\end{aligned} \tag{2.2.65}$$

where we used the fact that $z \neq x$ for the fourth equality.

3. Assume that $\tau = s\mu$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. By hypothesis,

$$1 = \mathrm{Sub}_{x,y}(\tau) = \mathrm{Sub}_{x,y}(s\mu) = \mathrm{Sub}_{x,y}(s) \wedge \mathrm{Sub}_{x,y}(\mu), \tag{2.2.66}$$
$$y \notin \mathrm{FV}(\tau) = \mathrm{FV}(s\mu) = \mathrm{FV}(s) \cup \mathrm{FV}(\mu), \tag{2.2.67}$$

which implies that $\mathrm{Sub}_{x,y}(s)$, $y \notin \mathrm{FV}(s)$, $\mathrm{Sub}_{x,y}(\mu) = 1$, and $y \notin \mathrm{FV}(\mu)$, so by the induction hypothesis, $\mathrm{Sub}_{y,x}(s[y/x]) = 1$ and $\mathrm{Sub}_{y,x}(\mu[y/x]) = 1$. Hence,

$$\begin{aligned}\mathrm{Sub}_{y,x}(\tau[y/x]) &= \mathrm{Sub}_{y,x}((s\mu)[y/x]) = \mathrm{Sub}_{y,x}(s[y/x]\mu[y/x]) \\ &= \mathrm{Sub}_{y,x}(s[y/x]) \wedge \mathrm{Sub}_{y,x}(\mu[y/x]) = 1 \wedge 1 = 1.\end{aligned} \tag{2.2.68}$$

4. Assume that $\tau = 0$. Then, $\mathrm{Sub}_{y,x}(\tau[y/x]) = \mathrm{Sub}_{y,x}(0[y/x]) = \mathrm{Sub}_{y,x}(0) = 1$.

5. Assume that $\tau = r \cdot \mu$, where $r \in \mathcal{R}$ is a scalar and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. By hypothesis, $1 = \mathrm{Sub}_{x,y}(\tau) = \mathrm{Sub}_{x,y}(r \cdot \mu) = \mathrm{Sub}_{x,y}(\mu)$ and $y \notin \mathrm{FV}(\tau) = \mathrm{FV}(r \cdot \mu) = \mathrm{FV}(\mu)$, so by the induction hypothesis, $\mathrm{Sub}_{y,x}(\mu[y/x]) = 1$, which implies that

$$\mathrm{Sub}_{y,x}(\tau[y/x]) = \mathrm{Sub}_{y,x}((r \cdot \mu)[y/x]) = \mathrm{Sub}_{y,x}\, r \cdot \mu[y/x]) = \mathrm{Sub}_{y,x}\, \mu[y/x]) = 1. \tag{2.2.69}$$

6. Assume that $\tau = \mu + \nu$, where $\mu, \nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. By hypothesis,

$$1 = \mathrm{Sub}_{x,y}(\tau) = \mathrm{Sub}_{x,y}(\mu + \nu) = \mathrm{Sub}_{x,y}(\mu) \wedge \mathrm{Sub}_{x,y}(\nu), \tag{2.2.70}$$
$$y \notin \mathrm{FV}(\tau) = \mathrm{FV}(\mu + \nu) = \mathrm{FV}(\mu) \cup \mathrm{FV}(\nu), \tag{2.2.71}$$

which implies that $\mathrm{Sub}_{x,y}(\mu)$, $y \notin \mathrm{FV}(\mu)$, $\mathrm{Sub}_{x,y}(\nu) = 1$, and $y \notin \mathrm{FV}(\nu)$, so by the induction hypothesis, $\mathrm{Sub}_{y,x}(\mu[y/x]) = 1$ and $\mathrm{Sub}_{y,x}(\nu[y/x]) = 1$. Hence,

$$\begin{aligned}\mathrm{Sub}_{y,x}(\tau[y/x]) &= \mathrm{Sub}_{y,x}((\mu+\nu)[y/x]) = \mathrm{Sub}_{y,x}(\mu[y/x] + \nu[y/x]) \\ &= \mathrm{Sub}_{y,x}(\mu[y/x]) \wedge \mathrm{Sub}_{y,x}(\nu[y/x]) = 1 \wedge 1 = 1.\end{aligned} \tag{2.2.72}$$

This completes the induction. Hence, $\mathrm{Sub}_{y,x}(\tau[y/x]) = 1$. Since $\mathrm{Sub}_{x,y}(\tau) = \mathrm{Sub}_{y,x}(\tau[y/x]) = \mathrm{Sub}_{y,x}(y) = 1$ and $y \notin \mathrm{FV}(\tau)$, by Lemma 2.2.12, we have that $\mathrm{Sub}_{y,x}(\tau) = \mathrm{Sub}_{x,y[x/y]}(\tau[x/y]) = 1$ and $\tau[y/x][x/y] = \tau[x/y][y[x/y]/x] = \tau[x/y][x/x]$. By Lemma 2.2.11, $\tau[x/y][x/x] = \tau[x/y]$, and since $y \notin \mathrm{FV}\,\tau$, by Lemma 2.2.9, $\tau[x/y] = \tau$. Hence, $\tau[y/x][x/y] = \tau$. ❑

We have proven a lot about differential pre-terms, but it is now time that we start laying the groundwork to define differential terms. Remember the set of differential terms is supposed to be the set of differential pre-terms quotiented by some relation, which we will call $\alpha$-equivalence or $\alpha$-conversion. But how should we define $\alpha$-equivalence? To answer this, we need to understand how differential terms should differ from differential pre-terms. The idea is that a differential term should capture some expression relating to analytic functions that, unlike differential pre-terms, is invariant under minor syntactic changes that do not alter the meaning of the analytic functions the differential term represent. That is, $\alpha$-equivalence should equate differential pre-terms that express the same underlying analytic function. For example, $\lambda x.x$ and $\lambda y.y$ are two differential pre-terms that represent a function that takes in a number and returns the inputted number. That is, both differential terms represent the identity function; the fact that the former differential term has $x$ as a bound variable and the latter has a bound variable has $y$ has a bound variable doesn't change that. Hence, $\lambda x.x$ and $\lambda y.y$ are related by $\alpha$-conversion. More generally, relabeling any bound variable in a differential pre-term that does not result in variable capture should result in an $\alpha$-equivalent differential pre-term. Also, given a variable $f \in \mathcal{V}$ and simple differential pre-terms $t_1, t_2 \in \Delta_{d,\mathcal{V}}$ representing analytic functions of type $\mathbb{R}^2 \to \mathbb{R}$, the differential pre-terms $D_{1,2}f * (t_1, t_2)$ and $D_{2,1}f * (t_2, t_1)$ represent the same underlying analytic function $(\partial_{x_1}\partial_{x_2} f)(x_1, x_2) \cdot t_1(x_1, x_2) \cdot t_2(x_1, x_2) = (\partial_{x_2}\partial_{x_1} f)(x_1, x_2) \cdot t_2(x_1, x_2) \cdot t_1(x_1, x_2)$, where we used the Schwartz Lemma for the equality. Hence, we should have that $D_{1,2}f * (t_1, t_2)$ and $D_{2,1}f * (t_2, t_1)$ are related by $\alpha$-conversion. With this intuition, we now can define $\alpha$-conversion precisely.

**Definition 2.2.14. $\alpha$-Conversion for Differential Pre-Terms**

The relation $=_{d,\alpha}$ is the smallest (with respect to inclusion) transitive and reflexive relation on the set $\mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ of complex differential pre-terms satisfying the following properties:

(i) If $n \in \mathbb{N}_0$ is a nonnegative integer, $x, y \in \mathcal{V}$ are variables, and $s, t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms such that $y \notin \mathrm{FV}(s)$ and $y$ can be substituted for $x$ in $s$ (i.e., $\mathrm{Sub}_{x,y}(s) = 1$), then $D_1^n \lambda x.s * (t_1, \ldots, t_n) =_{d,\alpha} D_1^n \lambda y.s[y/x] * (t_1, \ldots, t_n)$.

(ii) If $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, $x \in \mathcal{V}$ is a variable, and $t_1, \ldots, t_n, v_1, \ldots, v_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms such that $t_j =_{d,\alpha} v_j$ for all $1 \leqslant j \leqslant n$, then $D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n) =_{d,\alpha} D_{i_1,\ldots,i_n} x * (v_1, \ldots, v_n)$.

(iii) If $n \in \mathbb{N}_0$ is a nonnegative integer, $x \in \mathcal{V}$ is a variable, and $s, t_1, \ldots, t_n, u, v_1, \ldots, v_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms such that $s =_{d,\alpha} u$ and $t_j =_{d,\alpha} v_j$ for all $1 \leqslant j \leqslant n$, then $D_1^n \lambda x.s * (t_1, \ldots, t_n) =_{d,\alpha} D_1^n \lambda x.u * (v_1, \ldots, v_n)$.

(iv) If $s, u \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms such that $s =_{d,\alpha} u$ and $\tau, \sigma \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-terms such that $\tau =_{d,\alpha} \sigma$, then $s\tau =_{d,\alpha} u\sigma$.

(v) If $r \in \mathcal{R}$ is a scalar and $\tau, \sigma \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-terms such that $\tau =_{d,\alpha} \sigma$, then $r \cdot \tau =_{d,\alpha} r \cdot \sigma$.

(vi) If $\mu, \nu, \sigma, \tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-terms such that $\tau =_{d,\alpha} \mu$ and $\sigma =_{d,\alpha} \nu$, then $\tau + \sigma =_{d,\alpha} \mu + \nu$.

(vii) If $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, $t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, $x \in \mathcal{V}$ is a variable, and $f \in \mathfrak{S}_n$ is some permutation, then $D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n) =_{d,\alpha} D_{i_1,\ldots,i_n} x * (t_{f(1)}, \ldots, t_{f(n)})$.

(viii) If $n \in \mathbb{N}_0$ is a nonnegative integer, $x \in \mathcal{V}$ is a variable, $s, t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $f \in \mathfrak{S}_n$ is some permutation, then $D_1^n \lambda x.s * (t_1, \ldots, t_n) =_{d,\alpha} D_1^n \lambda x.s * (t_{f(1)}, \ldots, t_{f(n)})$.

(ix) If $\sigma, \tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-terms, then $\sigma + \tau =_{d,\alpha} \tau + \sigma$.

(x) If $\rho, \tau, \sigma \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-terms, then $(\rho + \sigma) + \tau =_{d,\alpha} \rho + (\sigma + \tau)$ and $\rho + (\sigma + \tau) =_{d,\alpha} (\rho + \sigma) + \tau$.

(xi) If $\tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term, then $\tau + 0 =_{d,\alpha} \tau$ and $\tau =_{d,\alpha} \tau + 0$.

(xii) If $r_1, r_2 \in \mathcal{R}$ are scalars and $\tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term, then $r_1 \cdot \tau + r_2 \cdot \tau =_{d,\alpha} (r_1 + r_2) \cdot \tau$ and $(r_1 + r_2) \cdot \tau =_{d,\alpha} r_1 \cdot \tau + r_2 \cdot \tau$.

(xiii) If $r \in \mathcal{R}$ is a scalar and $\tau, \sigma \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-term, then $r \cdot \tau + r \cdot \sigma =_{d,\alpha} r \cdot (\tau + \sigma)$ and $r \cdot (\tau + \sigma) =_{d,\alpha} r \cdot \tau + r \cdot \sigma$.

(xiv) If $r_1, r_2 \in \mathcal{R}$ are scalars and $\tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term, then $r_1 \cdot (r_2 \cdot \tau) =_{d,\alpha} (r_1 \cdot r_2) \cdot \tau$ and $(r_1 \cdot r_2) \cdot \tau =_{d,\alpha} r_1 \cdot (r_2 \cdot \tau)$.

(xv) If $\tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term, then $1 \cdot \tau =_{d,\alpha} \tau$ and $\tau =_{d,\alpha} 1 \cdot \tau$.

(xvi) If $r \in \mathcal{R}$ is a scalar, $r \cdot 0 =_{d,\alpha} 0$ and $0 =_{d,\alpha} r \cdot 0$.

(xvii) If $r_1, r_2 \in \mathcal{R}$ are scalars and $\tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term, then $(r_1 \cdot r_2) \cdot \tau =_{d,\alpha} (r_2 \cdot r_1) \cdot \tau$. ♣

In order to quotient differential-pre terms by $\alpha$-conversion, $\alpha$-conversion cannot be just any relation on differential-pre terms; it needs to be an equivalence relation, and showing that $\alpha$-conversion is an equivalence relation is exactly what we do in the next lemma.

**Lemma 2.2.15**

The relation $=_{d,\alpha}$ is an equivalence relation. ♡

***Proof.*** By construction, $=_{d,\alpha}$ is transitive and reflexive, so all that remains is to show symmetry.

That is, $\sigma =_{d,\alpha} \tau$ given complex differential pre-terms $\sigma, \tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau =_{d,\alpha} \sigma$. This can be shown via induction on $\tau =_{d,\alpha} \sigma$. The only interesting case is when $\tau =_{d,\alpha} \sigma$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, variables $x, y \in \mathcal{V}$, simple differential pre-terms $s, t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ such that $y \notin \mathrm{FV}(s)$, $\mathrm{Sub}_{x,y}(s) = 1$, $\tau = D_1^n \lambda x.s * (t_1, \ldots, t_n)$ and $\sigma = D_1^n \lambda y.s[y/x] * (t_1, \ldots, t_n)$. If $x \in \mathrm{FV}(y) = \{y\}$, then $x = y \neq \mathrm{FV}(s)$, so by Lemma 2.2.10, $x \notin \mathrm{FV}(s[y/x])$. Since $\mathrm{Sub}_{x,y}(s) = 1$ and $x \notin \mathrm{FV}(s[y/x])$, by Lemma 2.2.13, $x$ can be substituted for $y$ in $s[y/x]$ (i.e., $\mathrm{Sub}_{y,x}(s[y/x]) = 1$) and $s[y/x][x/y] = s$. This implies that $\sigma = D_1^n \lambda y.s[y/x] * (t_1, \ldots, t_n) =_{d,\alpha} D_1^n \lambda x.s[y/x][x/y] * (t_1, \ldots, t_n) =_1^n \lambda x.s * (t_1, \ldots, t_n) = \tau$. ❏

Now, we prove some important properties about $\alpha$-equivalence.

**Lemma 2.2.16**

If $\tau, \sigma \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms such that $\tau =_{d,\alpha} \sigma$, then $\mathrm{FV}(\tau) = \mathrm{FV}(\sigma)$. ♡

***Proof.*** This follows by induction on $\tau =_{d,\alpha} \sigma$.

1. Assume that $\tau =_{d,\alpha} \sigma$ because of transitivity. That is, there exist $\rho \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau =_{d,\alpha} \rho$ and $\rho =_{d,\alpha} \sigma$. Then, by the induction hypothesis, $\mathrm{FV}(\tau) = \mathrm{FV}(\rho)$ and $\mathrm{FV}(\rho) = \mathrm{FV}(\sigma)$, so $\mathrm{FV}(\tau) = \mathrm{FV}(\sigma)$.

2. Assume that $\tau =_{d,\alpha} \sigma$ because of reflexivity. That is, $\tau = \sigma$. Then, trivially, $\mathrm{FV}(\tau) = \mathrm{FV}(\sigma)$. Assume now that $\tau =_{d,\alpha} \sigma$ because of symmetry. Then, $\sigma =_{d,\alpha} \tau$, so by the induction hypothesis, $\mathrm{FV}(\sigma) = \mathrm{FV}(\tau)$.

3. Assume that $\tau =_{d,\alpha} \sigma$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, positive integers $i_1, \ldots, i_n \in \mathbb{N}$, simple differential pre-terms $t_1, \ldots, t_n, v_1, \ldots, v_n \in \Delta_{d,\mathcal{V}}$, and a variable $x \in \mathcal{V}$, such that $t_j =_{d,\alpha} v_j$ for all $1 \leqslant j \leqslant n$, $\tau = D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)$, $\sigma = D_{i_1,\ldots,i_n} x * (v_1, \ldots, v_n)$. By the induction hypothesis, $\mathrm{FV}(t_j) = \mathrm{FV}(v_j)$ for all $1 \leqslant j \leqslant n$. Hence,
$$\begin{aligned} \mathrm{FV}(\tau) &= \mathrm{FV}(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) = \{x\} \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n) \\ &= \{x\} \cup \mathrm{FV}(v_1) \cup \cdots \cup \mathrm{FV}(v_n) = \mathrm{FV}(D_{i_1,\ldots,i_n} x * (v_1, \ldots, v_n)) = \mathrm{FV}(\sigma). \end{aligned} \tag{2.2.73}$$

4. Assume that $\tau = \sigma$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, variables $x, y \in \mathcal{V}$, and simple differential pre-terms $s, t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ such that $y \notin \mathrm{FV}(s)$, $\mathrm{Sub}_{x,y}(s) = 1$, $\tau = D_1^n \lambda x.s * (t_1, \ldots, t_n)$, and $\sigma = D_1^n \lambda y.s[y/x] * (t_1, \ldots, t_n)$. Then,
$$\begin{aligned} \mathrm{FV}(\tau) &= \mathrm{FV}(D_1^n \lambda x.s * (t_1, \ldots, t_n)) = (\mathrm{FV}(s) - \{x\}) \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n) \\ &= (\mathrm{FV}(s[y/x]) - \{y\}) \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n) \\ &= \mathrm{FV}(D_1^n \lambda y.s[y/x] * (t_1, \ldots, t_n)) = \mathrm{FV}(\sigma), \end{aligned} \tag{2.2.74}$$
where we used Lemma 2.2.10 for the third equality.

5. Assume that $\tau =_{d,\alpha} \sigma$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, and simple differential pre-terms $s, t_1, \ldots, t_n, u, v_1, \ldots, v_n \in \Delta_{d,\mathcal{V}}$ such that $s =_{d,\alpha} u$, $t_j =_{d,\alpha} v_j$ for all $1 \leqslant j \leqslant n$, $\tau = D_1^n \lambda x.s * (t_1, \ldots, t_n)$, and $\sigma = D_1^n \lambda x.u * (v_1, \ldots, v_n)$. By the induction hypothesis, $\mathrm{FV}(s) = \mathrm{FV}(u)$ and $\mathrm{FV}(t_j) = \mathrm{FV}(v_j)$ for all $1 \leqslant j \leqslant n$, so
$$\begin{aligned} \mathrm{FV}(\tau) &= \mathrm{FV}(D_1^n \lambda x.s * (t_1, \ldots, t_n)) = (\mathrm{FV}(s) - \{x\}) \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n) \\ &= (\mathrm{FV}(u) - \{x\}) \cup \mathrm{FV}(v_1) \cup \cdots \cup \mathrm{FV}(v_n) \\ &= \mathrm{FV}(D_1^n \lambda x.u * (v_1, \ldots, v_n)) = \mathrm{FV}(\sigma). \end{aligned} \tag{2.2.75}$$

6. Assume that $\tau =_{d,\alpha} \sigma$ because there exist simple differential pre-terms $s, u \in \Delta_{d,\mathcal{V}}$ and complex differential pre-terms $\mu, \nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $s =_{d,\alpha} u$, $\mu =_{d,\alpha} \nu$, $\tau = s\mu$, and $\sigma = u\sigma$. By the induction hypothesis, $\mathrm{FV}(s) = \mathrm{FV}(u)$ and $\mathrm{FV}(\mu) = \mathrm{FV}(\nu)$, so

$$\mathrm{FV}(\tau) = \mathrm{FV}(s\mu) = \mathrm{FV}(s) \cup \mathrm{FV}(\mu) = \mathrm{FV}(u) \cup \mathrm{FV}(\nu) = \mathrm{FV}(u\nu) = \mathrm{FV}(\sigma). \tag{2.2.76}$$

7. Assume that $\tau =_{d,\alpha} \sigma$ because there exist a scalar $r \in \mathcal{R}$ and complex differential pre-terms $\mu, \nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\mu =_{d,\alpha} \nu$, $\tau = r \cdot \mu$, and $\sigma = r \cdot \nu$. By the induction hypothesis, $\mathrm{FV}(\mu) = \mathrm{FV}(\nu)$, so

$$\mathrm{FV}(\tau) = \mathrm{FV}(r \cdot \mu) = \mathrm{FV}(\mu) = \mathrm{FV}(\nu) = \mathrm{FV}(r \cdot \nu) = \mathrm{FV}(\sigma). \tag{2.2.77}$$

8. Assume that $\tau =_{d,\alpha} \sigma$ because there exist complex differential pre-terms $\mu, \nu, \alpha, \beta \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\mu =_{d,\alpha} \nu$, $\alpha =_{d,\alpha} \beta$, $\tau = \mu + \alpha$, and $\sigma = \nu + \beta$. By the induction hypothesis, $\mathrm{FV}(\mu) = \mathrm{FV}(\nu)$ and $\mathrm{FV}(\alpha) = \mathrm{FV}(\beta)$, so

$$\begin{aligned}\mathrm{FV}(\tau) &= \mathrm{FV}(\mu + \alpha) = \mathrm{FV}(\mu) \cup \mathrm{FV}(\alpha) \\ &= \mathrm{FV}(\nu) + \mathrm{FV}(\beta) = \mathrm{FV}(\nu + \beta) = \mathrm{FV}(\sigma).\end{aligned} \tag{2.2.78}$$

9. Assume that $\tau =_{d,\alpha} \sigma$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, positive integers $i_1, \ldots, i_n \in \mathbb{N}$, simple differential pre-terms $t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$, a variable $x \in \mathcal{V}$, and a permutation $f \in \mathfrak{S}_n$ such that $\tau = D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)$ and $\sigma = D_{i_1,\ldots,i_n} x * (t_{f(1)}, \ldots, t_{f(n)})$. Then,

$$\begin{aligned}\mathrm{FV}(\tau) &= \mathrm{FV}(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) = \{x\} \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n) \\ &= \{x\} \cup \mathrm{FV}(t_{f(1)}) \cup \cdots \cup \mathrm{FV}(t_{f(n)}) \\ &= \mathrm{FV}(D_{i_1,\ldots,i_n} x * (t_{f(1)}, \ldots, t_{f(n)})) = \mathrm{FV}(\sigma).\end{aligned} \tag{2.2.79}$$

10. Assume that $\tau =_{d,\alpha} \sigma$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, simple differential pre-terms $s, t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$, and a permutation $f \in \mathfrak{S}_n$ such that $\tau = D_1^n \lambda x.s * (t_1, \ldots, t_n)$ and $\sigma = D_1^n \lambda x.s * (t_{f(1)}, \ldots, t_{f(n)})$. Then,

$$\begin{aligned}\mathrm{FV}(\tau) &= \mathrm{FV}(D_1^n \lambda x.s * (t_1, \ldots, t_n)) = (\mathrm{FV}(s) - \{x\}) \cup \mathrm{FV}(t_1) \cup \cdots \cup \mathrm{FV}(t_n) \\ &= (\mathrm{FV}(s) - \{x\}) \cup \mathrm{FV}(t_{f(1)}) \cup \cdots \cup \mathrm{FV}(t_{f(n)}) \\ &= \mathrm{FV}(D_1^n \lambda x.s * (t_{f(1)}, \ldots, t_{f(n)})) = \mathrm{FV}(\sigma).\end{aligned} \tag{2.2.80}$$

11. Assume that $\tau =_{d,\alpha} \sigma$ because there exist complex differential pre-terms $\mu, \nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau = \mu + \nu$ and $\sigma = \nu + \mu$. Then,

$$\begin{aligned}\mathrm{FV}(\tau) &= \mathrm{FV}(\mu + \nu) = \mathrm{FV}(\mu) \cup \mathrm{FV}(\nu) \\ &= \mathrm{FV}(\nu) \cup \mathrm{FV}(\mu) = \mathrm{FV}(\nu + \mu) = \mathrm{FV}(\sigma).\end{aligned} \tag{2.2.81}$$

12. Assume that $\tau =_{d,\alpha} \sigma$ because there exist complex differential pre-terms $\mu, \nu, \xi \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau = (\mu + \nu) + \xi$ and $\sigma = \mu + (\nu + \xi)$. Then,

$$\begin{aligned}\mathrm{FV}(\tau) &= \mathrm{FV}((\mu + \nu) + \xi) = \mathrm{FV}(\mu + \nu) \cup \mathrm{FV}(\xi) = \mathrm{FV}(\mu) \cup \mathrm{FV}(\nu) \cup \mathrm{FV}(\xi) \\ &= \mathrm{FV}(\mu) + \mathrm{FV}(\nu + \xi) = \mathrm{FV}(\mu + (\nu + \xi)) = \mathrm{FV}(\sigma).\end{aligned} \tag{2.2.82}$$

13. Assume that $\tau =_{d,\alpha} \sigma$ because there exists a complex differential pre-term $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau = \mu + 0$ and $\sigma = \mu$. Then,

$$\mathrm{FV}(\tau) = \mathrm{FV}(\mu + 0) = \mathrm{FV}(\mu) + \mathrm{FV}(0) = \mathrm{FV}(\sigma) + \varnothing = \mathrm{FV}(\sigma). \tag{2.2.83}$$

14. Assume that $\tau =_{d,\alpha} \sigma$ because there exist scalars $r_1, r_2 \in \mathcal{R}$ and a complex differential pre-term $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau = r_1 \cdot \mu + r_2 \cdot \mu$ and $\sigma = (r_1 + r_2) \cdot \mu$. Then,

$$\begin{aligned}\mathrm{FV}(\tau) &= \mathrm{FV}(r_1 \cdot \mu + r_2 \cdot \mu) = \mathrm{FV}(r_1 \cdot \mu) \cup \mathrm{FV}(r_2 \cdot \mu) \\ &= \mathrm{FV}(\mu) \cup \mathrm{FV}(\mu) = \mathrm{FV}(\mu) = \mathrm{FV}((r_1 + r_2) \cdot \mu) = \mathrm{FV}(\sigma).\end{aligned} \tag{2.2.84}$$

15. Assume that $\tau =_{d,\alpha} \sigma$ because there exist a scalar $r \in \mathcal{R}$ and complex differential pre-terms $\mu, \nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau = r \cdot \mu + r \cdot \nu$ and $\sigma = r \cdot (\mu + \nu)$. Then,

$$\begin{aligned}\mathrm{FV}(\tau) &= \mathrm{FV}(r \cdot \mu + r \cdot \nu) = \mathrm{FV}(r \cdot \mu) \cup \mathrm{FV}(r \cdot \nu) = \mathrm{FV}(\mu) \cup \mathrm{FV}(\nu) \\ &= \mathrm{FV}(\mu + \nu) = \mathrm{FV}(r \cdot (\mu + \nu)) = \mathrm{FV}(\sigma).\end{aligned} \tag{2.2.85}$$

16. Assume that $\tau =_{d,\alpha} \sigma$ because there exist scalars $r_1, r_2 \in \mathcal{R}$ and a complex differential pre-term $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau = r_1 \cdot (r_2 \cdot \mu)$ and $\sigma = (r_1 \cdot r_2) \cdot \mu$. Then,

$$\mathrm{FV}(\tau) = \mathrm{FV}(r_1 \cdot (r_2 \cdot \mu)) = \mathrm{FV}(r_2 \cdot \mu) = \mathrm{FV}(\mu) = \mathrm{FV}((r_1 \cdot r_2) \cdot \mu) = \mathrm{FV}(\sigma). \tag{2.2.86}$$

17. Assume that $\tau =_{d,\alpha} \sigma$ because there exists a complex differential pre-term $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau = 1 \cdot \mu$ and $\sigma = \mu$. Then,

$$\mathrm{FV}(\tau) = \mathrm{FV}(1 \cdot \mu) = \mathrm{FV}(\mu) = \mathrm{FV}(\sigma). \tag{2.2.87}$$

18. Assume that $\tau =_{d,\alpha} \sigma$ because there exists a scalar $r \in \mathcal{R}$ such that $\tau = r \cdot 0$ and $\sigma = 0$. Then,

$$\mathrm{FV}(\tau) = \mathrm{FV}(r \cdot 0) = \mathrm{FV}(0) = \mathrm{FV}(\sigma). \tag{2.2.88}$$

19. Assume that $\tau =_{d,\alpha} \sigma$ because there exist scalars $r_1, r_2 \in \mathcal{R}$ and a complex differential pre-term $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau = (r_1 \cdot r_2) \cdot \mu$ and $\sigma = (r_2 \cdot r_1) \cdot \mu$. Then,

$$\mathrm{FV}(\tau) = \mathrm{FV}((r_1 \cdot r_2) \cdot \mu) = \mathrm{FV}(\mu) = \mathrm{FV}((r_2 \cdot r_1) \cdot \mu) = \mathrm{FV}(\sigma). \tag{2.2.89}$$

This completes the induction. ❏

**Lemma 2.2.17**

If $\tau, \tau' \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms, $t, t' \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $x \in \mathcal{V}$ is a variable such that $\tau =_{d,\alpha} \tau'$, $t =_{d,\alpha} t'$, and $\mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t'}(\tau') = 1$, then $\tau[t/x] = \tau'[t'/x]$. ♡

***Proof.*** This follows by induction on $\tau =_{d,\alpha} \tau'$. There are only two interesting cases. The first interesting case is when $\tau =_{d,\alpha} \tau'$ by reflexivity. That is, $\tau = \tau'$. We proceed by induction on $\tau$.

1. Assume that $\tau = D_{i_1,\ldots,i_n} y * (t_1,\ldots,t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1,\ldots,i_n \in \mathbb{N}$ are positive integers, $t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $y \in \mathcal{V}$ is a variable. If $y = x$, then $y[t/x] = x[t/x] = t =_{d,\alpha} t' = x[t'/x] = y[t'/x]$, and if $y \neq x$, then $y[t/x] = y = y[t'/x]$, which implies by the reflexivity of $=_{d,\alpha}$ that $y[t/x] =_{d,\alpha} y[t'/x]$. Also, since $1 = \mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(D_{i_1,\ldots,i_n} y * (t_1,\ldots,t_n)) = \mathrm{Sub}_{x,t}(t_1) \cup \cdots \cup \mathrm{Sub}_{x,t}(t_n)$, $\mathrm{Sub}_{x,t}(t_j) = 1$ for all $1 \leqslant j \leqslant n$. Similarly, since $1 = \mathrm{Sub}_{x,t'}(\tau') = \mathrm{Sub}_{x,t'}(\tau)$, $\mathrm{Sub}_{x,t'}(t_j) = 1$ for all $1 \leqslant j \leqslant n$. Hence, by the induction hypothesis, $t_j[t/x] =_{d,\alpha} t_j[t'/x]$. Since $y[t/x] =_{d,\alpha} y[t'/x]$ and $t_j[t/x] =_{d,\alpha} t_j[t'/x]$ for all $1 \leqslant j \leqslant n$, $\tau[t/x] = (D_{i_1,\ldots,i_n} y * (t_1,\ldots,t_n))[t/x] = D_{i_1,\ldots,i_n} y[t/x] * (t_1[t/x],\ldots,t_n[t/x]) =_{d,\alpha} D_{i_1,\ldots,i_n} y[t'/x] * (t_1[t'/x],\ldots,t_n[t'/x]) = (D_{i_1,\ldots,i_n} y * (t_1,\ldots,t_n))[t'/x] = \tau[t'/x] = \tau'[t'/x]$.

2. Assume that $\tau = D_1^n \lambda x.s * (t_1,\ldots,t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $s, t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. Since $1 = \mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(D_1^n \lambda x.s * (t_1,\ldots,t_n)) = \mathrm{Sub}_{x,t}(t_1) \cup \cdots \cup \mathrm{Sub}_{x,t}(t_n)$, $\mathrm{Sub}_{x,t}(t_j) = 1$ for all $1 \leqslant j \leqslant n$. Similarly, $\mathrm{Sub}_{x,t'}(t_j) = 1$ for all $1 \leqslant j \leqslant n$. By the induction hypothesis, this implies that $t_j[t/x] =_{d,\alpha} t_j[t'/x]$. Therefore, since $s =_{d,\alpha} s$ and $t_j[t/x] =_{d,\alpha} t_j[t'/x]$ for all $1 \leqslant j \leqslant n$, $\tau[t/x] = (D_1^n \lambda x.s * (t_1,\ldots,t_n))[t/x] = D_1^n \lambda x.s * (t_1[t/x],\ldots,t_n[t/x]) =_{d,\alpha} D_1^n \lambda x.s * (t_1[t'/x],\ldots,t_n[t'/x]) = (D_1^n \lambda x.s * (t_1,\ldots,t_n))[t'/x] = \tau[t'/x] = \tau'[t'/x]$.

3. Assume now that $\tau = D_1^n \lambda y.s * (t_1,\ldots,t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $y \in \mathcal{V}$ is a variable, and $s, t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. Since $1 = \mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(D_1^n \lambda x.s * (t_1,\ldots,t_n)) = \mathrm{Sub}_{x,t}(s) \cup \mathrm{Sub}_{x,t}(t_1) \cup \cdots \cup \mathrm{Sub}_{x,t}(t_n)$, $\mathrm{Sub}_{x,t}(s) = 1$ and $\mathrm{Sub}_{x,t}(t_j) = 1$ for all $1 \leqslant j \leqslant n$. Similarly, $\mathrm{Sub}_{x,t'}(s) = 1$ and $\mathrm{Sub}_{x,t'}(t_j) = 1$ for all $1 \leqslant j \leqslant n$. By the induction hypothesis, $s[t/x] =_{d,\alpha} [t'/x]$ and $t_j[t/x] =_{d,\alpha} t_j[t'/x]$ for all $1 \leqslant j \leqslant n$, whence $\tau[t/x] = (D_1^n \lambda y.s[t/x] * (t_1,\ldots,t_n))[t/x] = D_1^n \lambda y.s[t/x] * (t_1[t/x],\ldots,t_n[t/x]) =_{d,\alpha} D_1^n \lambda y.s[t'/x] * (t_1[t'/x],\ldots,t_n[t'/x]) = (D_1^n \lambda y.s * (t_1,\ldots,t_n))[t'/x] = \tau[t'/x] = \tau'[t'/x]$.

4. Assume now that $\tau = s\mu$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. Since $\mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(s\mu) = \mathrm{Sub}_{x,t}(s) \wedge \mathrm{Sub}_{x,t}(\mu)$, $\mathrm{Sub}_{x,t}(s) = \mathrm{Sub}_{x,t}(\mu) = 1$. Similarly, $\mathrm{Sub}_{x,t'}(s) = \mathrm{Sub}_{x,t'}(\mu) = 1$. Therefore, by the induction hypothesis, $s[t/x] = s[t'/x]$ and $\mu[t/x] = \mu[t'/x]$, so $\tau[t/x] = (s\mu)[t/x] = s[t/x]\mu[t/x] = s[t'/x]\mu[t'/x] = (s\mu)[t'/x] = \tau[t'/x] = \tau'[t'/x]$.

5. Assume now that $\tau = 0$. Then, $\tau[t/x] = 0[t/x] = 0 = 0[t'/x] = \tau[t'/x] = \tau'[t'/x]$.

6. Assume now that $\tau = r \cdot \mu$, where $r \in \mathcal{R}$ is a scalar and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. Since $1 = \mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(r \cdot \mu) = \mathrm{Sub}_{x,t}(\mu)$ and $\mathrm{Sub}_{x,t'}(\tau') = \mathrm{Sub}_{x,t'}(\tau) = \mathrm{Sub}_{x,t'}(r \cdot \mu) = \mathrm{Sub}_{x,t}(\mu)$, by the induction hypothesis, $\mu[t/x] =_{d,\alpha} \mu[t'/x]$, so $\tau[t/x] = (r \cdot \mu)[t/x] = r \cdot \mu[t/x] =_{d,\alpha} r \cdot \mu[t'/x] = (r \cdot \mu)[t'/x] = \tau[t'/x] = \tau'[t'/x]$.

7. Assume now that $\tau = \mu + \nu$, where $\mu, \nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. Since $\mathrm{Sub}_{x,t}(\tau) = \mathrm{Sub}_{x,t}(\mu + \nu) = \mathrm{Sub}_{x,t}(\mu) \wedge \mathrm{Sub}_{x,t}(\nu)$, $\mathrm{Sub}_{x,t}(\mu) = \mathrm{Sub}_{x,t}(\nu) = 1$. Similarly, $\mathrm{Sub}_{x,t'}(\mu) = \mathrm{Sub}_{x,t'}(\nu) = 1$. Therefore, by the induction hypothesis, $\mu[t/x] = \mu[t'/x]$ and $\nu[t/x] = \nu[t'/x]$, so $\tau[t/x] = (\mu+\nu)[t/x] = \mu[t/x]+\nu[t/x] = \mu[t'/x]+\nu[t'/x] = (\mu+\nu)[t'/x] = \tau[t'/x] = \tau'[t'/x]$.

This completes the induction for the reflexivity case.

The second interesting case is when $\tau =_{d,\alpha} \tau'$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, variables $z, y \in \mathcal{V}$, and simple differential pre-terms $s, t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ such that $y \notin \mathrm{FV}(s)$,

$\mathrm{Sub}_{z,y}(s) = 1$, $\tau = D_1^n \lambda z.s * (t_1, \ldots, t_n)$, and $\tau' = D_1^n \lambda y.s[y/z] * (t_1, \ldots, t_n)$. Since

$$\begin{aligned} 1 = \mathrm{Sub}_{x,t}(\tau) &= \mathrm{Sub}_{x,t}(D_1^n \lambda z.s * (t_1, \ldots, t_n)) \\ &= \mathrm{Sub}_{x,t}(\lambda z.s) \wedge \mathrm{Sub}_{x,t}(t_1) \wedge \cdots \wedge \mathrm{Sub}_{x,t}(t_n), \end{aligned} \tag{2.2.90}$$

so

$$\mathrm{Sub}_{x,t}(\lambda z.s) = \mathrm{Sub}_{x,t}(t_1) = \cdots = \mathrm{Sub}_{x,t}(t_n) = 1, \tag{2.2.91}$$

and since $\mathrm{Sub}_{x,t'}(\tau') = 1$, by an identical argument,

$$\mathrm{Sub}_{x,t'}(\lambda y.s[y/z]) = \mathrm{Sub}_{x,t'}(t_1) = \cdots = \mathrm{Sub}_{x,t'}(t_n) = 1. \tag{2.2.92}$$

Because $t_j =_{d,\alpha} t_j$ for all $1 \leqslant j \leqslant n$, by the induction hypothesis, this implies that $t_j[t/x] =_{d,\alpha} t_j[t'/x]$ for all $1 \leqslant j \leqslant n$. Suppose that $x = z$. Since $y \notin \mathrm{FV}(s)$ and $\mathrm{Sub}_{z,y}(s) = 1$, $\lambda z.s =_{d,\alpha} \lambda y.s[y/z]$, so by Lemma 2.2.16, $\mathrm{FV}(\lambda y.s[y/z]) = \mathrm{FV}(\lambda z.s) = \mathrm{FV}(s) - \{z\} = \mathrm{FV}(s) - \{x\}$, which implies that $z \notin \mathrm{FV}(\lambda y.s[y/z])$, so by Lemma 2.2.9, this implies that $(\lambda y.s[y/z])[t'/x] = \lambda y.s[y/z]$. Hence,

$$\begin{aligned} \tau[t/x] &= (D_1^n \lambda x.s * (t_1, \ldots, t_n))[t/z] = D_1^n \lambda x.s * (t_1[t/x], \ldots, t_n[t/x]) \\ &= D_1^n \lambda z.s * (t_1[t/x], \ldots, t_n[t/x]) =_{d,\alpha} D_1^n \lambda z.s * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &=_{d,\alpha} D_1^n \lambda y.s[y/z] * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &= D_1^n \lambda (y.s[y/z])[t'/x] * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &= (D_1^n \lambda y.s[y/z] * (t_1, \ldots, t_n))[t'/x] = \tau'[t'/x], \end{aligned} \tag{2.2.93}$$

where we used the fact that $y \notin \mathrm{FV}(s)$ and $\mathrm{Sub}_{z,y}(s) = 1$ for the second $\alpha$-conversion. A similar argument can be made when $x = y$. Now, suppose that $x$ is distinct form both $z$ and $y$. Since $\mathrm{Sub}_{x,t}(\lambda z.s) = 1$, $x \notin \mathrm{FV}(s)$ or $z \notin \mathrm{FV}(t)$. Let us first assume that $x \notin \mathrm{FV}(s)$. Since $x \neq y$ and $x \notin \mathrm{FV}(s)$, by Lemma 2.2.10, $x \notin \mathrm{FV}(s[y/z])$. Since $x \notin \mathrm{FV}(s)$ and $x \notin \mathrm{FV}(s[y/z])$, by Lemma 2.2.9, $s[t/x] = s$ and $s[y/z][t'/x] = s[y/z]$. Hence,

$$\begin{aligned} \tau[t/x] &= (D_1^n \lambda x.s * (t_1, \ldots, t_n))[t/x] = D_1^n \lambda x.s[t/x] * (t_1[t/x], \ldots, t_n[t/x]) \\ &= D_1^n \lambda x.s * (t_1[t/x], \ldots, t_n[t/x]) =_{d,\alpha} D_1^n \lambda x.s * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &=_{d,\alpha} D_1^n \lambda y.s[y/z] * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &= D_1^n \lambda y.s[y/z][t'/x] * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &= (D_1^n \lambda y.s[y/z] * (t_1, \ldots, t_n))[t'/x] = \tau'[t'/x], \end{aligned} \tag{2.2.94}$$

where we used the fact that $y \notin \mathrm{FV}(s)$ and $\mathrm{Sub}_{z,y}(s) = 1$ for the second $\alpha$-conversion. Instead, assume that $x \in \mathrm{FV}(s)$. Then, $z \notin \mathrm{FV}(t)$. Since $t =_{d,\alpha} t'$ and $z \notin \mathrm{FV}(t)$, by Lemma 2.2.16, $z \notin \mathrm{FV}(t')$. Also, since $\mathrm{Sub}_{x,t}(\lambda z.s) = \mathrm{Sub}_{x,t'}(\lambda y.s[y/z]) = 1$ and $x \notin z, y$, $\mathrm{Sub}_{x,t}(s) = \mathrm{Sub}_{x,t'}(s[y/z]) = 1$. Because $z \neq x$, $\mathrm{Sub}_{z,y}(s) = \mathrm{Sub}_{x,t'}(s[y/z]) = \mathrm{Sub}_{x,t'}(y) = 1$, and $z \notin \mathrm{FV}(t')$, by Lemma 2.2.12, $1 = \mathrm{Sub}_{z,y[t'/x]}(s[t'/x]) = \mathrm{Sub}_{z,y}(s[t'/x])$ and $s[y/z][t'/x] = s[t'/x][y[\rho/x]/z] = s[t'/x][y/z]$. Since $\mathrm{Sub}_{z,y}(s[t'/x]) = 1$ and $z \notin \mathrm{FV}(s[t'/x][y/z])$, by Lemma 2.2.13, $\mathrm{Sub}_{y,z}(s[t'/x][y/z]) = 1$ and $s[t'/x][y/z][z/y] = s[t'/x]$.

$$\begin{aligned} \tau'[t'/x] &= (D_1^n \lambda y.s[y/z] * (t_1, \ldots, t_n))[t'/x] \\ &= D_1^n \lambda y.s[y/z][t'/x] * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &= D_1^n \lambda y.s[t'/x][y/z] * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &=_{d,\alpha} D_1^n \lambda z.s[t'/x][y/z][z/y] * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &= D_1^n \lambda z.s[t'/x] * (t_1[t'/x], \ldots, t_n[t'/x]) \\ &= (D_1^n \lambda z.s * (t_1, \ldots, t_n))[t'/x] = \tau[t'/x] =_{d,\alpha} \tau[t/x]. \end{aligned} \tag{2.2.95}$$

where we used the fact that $z \notin \mathrm{FV}\,(s[t'/x][y/z])$ and $\mathrm{Sub}_{y,z}\,(s[t'/x][y/z]) = 1$ for the first $\alpha$-conversion, we used the fact that $\mathrm{Sub}_{x,t'}(\tau)$ (as $z \notin \mathrm{FV}\,t'$) for the , and the inductive hypothesis for the second $\alpha$-conversion. Since $=_{d,\alpha}$ is symmetric by Lemma 2.2.15, $\tau[t/x] =_{d,\alpha} \tau'[t'/x]$. ❑

**Lemma 2.2.18**

Let $V \subset \mathcal{V}$ be a subset of variables, and define the function $\mathrm{B}_V \colon \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle \to \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ by:

$$\mathrm{B}_V\,(D_{i_1,\ldots,i_n}x * (t_1,\ldots,t_n)) = D_{i_1,\ldots,i_n}x * (\mathrm{B}_V\,(t_1),\ldots,\mathrm{B}_V\,(t_n)), \tag{2.2.96a}$$
$$\mathrm{B}_V\,(D_1^n\lambda x.s * (t_1,\ldots,t_n)) = D_1^n\lambda y.\,\mathrm{B}_{V\cup\{y\}}\,(s[y/x]) * (\mathrm{B}_V\,(t_1),\ldots,\mathrm{B}_V\,(t_n)), \tag{2.2.96b}$$
$$\mathrm{B}_V\,(s\mu) = \mathrm{B}_V\,(s)\,\mathrm{B}_V\,(\mu), \tag{2.2.96c}$$
$$\mathrm{B}_V\,(0) = 0, \tag{2.2.96d}$$
$$\mathrm{B}_V\,(r\cdot\mu) = r\cdot\mathrm{B}_V\,(\mu), \tag{2.2.96e}$$
$$\mathrm{B}_V\,(\mu+\nu) = \mathrm{B}_V\,(\mu) + \mathrm{B}_V\,(\nu), \tag{2.2.96f}$$

where $x \in \mathcal{V}$ is a variable, $r \in \mathcal{R}$ is a scalar, $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1,\ldots,i_n \in \mathbb{N}$ are positive integers, $s, t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, $\mu,\nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms, and $y \in \mathcal{V}$ is a variable such that $y \notin V \cup \mathrm{Var}\,(s)$. Then, given a complex differential pre-term $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$, $\tau =_{d,\alpha} \mathrm{B}_V\,(\tau)$ and $V \cap \mathrm{BV}\,(\mathrm{B}_V\,(\tau)) = \varnothing$. ♡

***Proof.*** This follows by induction on $\tau$.

1. Assume that $\tau = D_{i_1,\ldots,i_n}x * (t_1,\ldots,t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1,\ldots,i_n \in \mathbb{N}$ are positive integers, $t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $x \in \mathcal{V}$ is a variable. For each $1 \leqslant j \leqslant n$, by the induction hypothesis on $t_j$, $t_j =_{d,\alpha} \mathrm{B}_V\,(t_j)$ and $V \cap \mathrm{BV}\,(\mathrm{B}_V\,(t_j)) = \varnothing$. Hence,

$$\begin{aligned}\tau &= D_{i_1,\ldots,i_n}x * (t_1,\ldots,t_n) =_{d,\alpha} D_{i_1,\ldots,i_n}x * (\mathrm{B}_V\,(t_1),\ldots,\mathrm{B}_V\,(t_n))\\ &= \mathrm{B}_V\,(D_{i_1,\ldots,i_n}x * (t_1,\ldots,t_n)) = \mathrm{B}\,(\tau),\end{aligned} \tag{2.2.97}$$

$$\begin{aligned}V \cap \mathrm{BV}\,(\mathrm{B}_V\,(\tau)) &= V \cap \mathrm{BV}\,(\mathrm{B}_V\,(D_{i_1,\ldots,i_n}x * (t_1,\ldots,t_n)))\\ &= V \cap \mathrm{BV}\,(D_{i_1,\ldots,i_n}x * (\mathrm{B}_V\,(t_1),\ldots,\mathrm{B}_V\,(t_n)))\\ &= V \cap (\mathrm{BV}\,(\mathrm{B}_V\,(t_1)) \cup \cdots \cup \mathrm{BV}\,(\mathrm{B}_V\,(t_n)))\\ &= (V \cap \mathrm{BV}\,(\mathrm{B}_V\,(t_1))) \cup \cdots \cup (V \cap \mathrm{BV}\,(\mathrm{B}_V\,(t_n)))\\ &= \varnothing \cup \cdots \cup \varnothing = \varnothing.\end{aligned} \tag{2.2.98}$$

2. Assume now that $\tau = D_1^n\lambda x.s * (t_1,\ldots,t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $x \in \mathcal{V}$ is a variable, and $s, t_1,\ldots,t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms. Since $y \notin \mathrm{Var}\,(s)$, it is clear that $y$ can be substituted for $x$ in $s$ (i.e., $\mathrm{Sub}_{x,y}\,(s) = 1$). Because $s \notin \mathrm{Var}\,(s) \supset \mathrm{FV}\,(s)$ and $\mathrm{Sub}_{x,y}\,(s) = 1$,

$$\tau = D_1^n\lambda x.s * (t_1,\ldots,t_n) =_{d,\alpha} D_1^n\lambda y.s[y/x] * (t_1,\ldots,t_n). \tag{2.2.99}$$

By the induction hypothesis on $s[y/x]$, we have that $s[y/x] =_{d,\alpha} \mathrm{B}_{V\cup\{y\}}\,(s[y/x])$ and $(V \cup \{y\}) \cap \mathrm{BV}\,(\mathrm{B}_{V\cup\{y\}}\,(s[y/x])) = \varnothing$, and for each $1 \leqslant j \leqslant n$, by the induction hypothesis on $t_j$, we have that $t_j =_{d,\alpha} \mathrm{B}_V\,(t_j)$ and $V \cap \mathrm{BV}\,(\mathrm{B}_V\,(t_j)) = \varnothing$. Hence,

$$\begin{aligned}\tau &=_{d,\alpha} D_1^n\lambda y.s[y/x] * (t_1,\ldots,t_n)\\ &=_{d,\alpha} D_1^n\lambda y.\,\mathrm{B}_{V\cup\{y\}}\,(s[y/x]) * (\mathrm{B}_V\,(t_1),\ldots,\mathrm{B}_V\,(t_n))\\ &= \mathrm{B}_V\,(D_1^n\lambda x.s * (t_1,\ldots,t_n)) = \mathrm{B}\,(\tau),\end{aligned} \tag{2.2.100}$$

$$\begin{aligned}
V \cap \mathrm{BV}\left(\mathrm{B}_V(\tau)\right) &= V \cap \mathrm{BV}\left(\mathrm{B}_V\left(D_1^n \lambda x.s * (t_1, \dots, t_n)\right)\right) \\
&= V \cap \mathrm{BV}\left(D_1^n \lambda y.\, \mathrm{B}_{V\cup\{y\}}\left(s[y/x]\right) * \left(\mathrm{B}_V(t_1), \dots, \mathrm{B}_V(t_n)\right)\right) \\
&= V \cap \left(\{y\} \cup \mathrm{BV}\left(\mathrm{B}_{V\cup\{y\}}\left(s[y/x]\right)\right) \cup \mathrm{BV}\left(\mathrm{B}_V(t_1)\right)\right. \\
&\qquad \left.\cup \dots \cup \mathrm{BV}\left(\mathrm{B}_V(t_n)\right)\right) \\
&= (V \cap \{y\}) \cup \left(V \cap \mathrm{BV}\left(\mathrm{B}_{V\cup\{y\}}\left(s[y/x]\right)\right)\right) \cup \left(V \cap \mathrm{BV}\left(\mathrm{B}_V(t_1)\right)\right) \\
&\qquad \cup \dots \cup \left(V \cap \mathrm{BV}\left(\mathrm{B}_V(t_n)\right)\right) \\
&= \varnothing \cup \varnothing \cup \varnothing \cup \dots \cup \varnothing = \varnothing.
\end{aligned} \tag{2.2.101}$$

3. Assume now that $\tau = s\mu$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. By the induction hypothesis on $s$, $s =_{d,\alpha} \mathrm{B}_V(s)$ and $V \cap \mathrm{BV}(\mathrm{B}_V(s)) = \varnothing$, and by the induction hypothesis on $\mu$, $\mu =_{d,\alpha} \mathrm{B}_V(\mu)$ and $V \cap \mathrm{BV}(\mathrm{B}_V(\mu)) = \varnothing$. Hence,

$$\tau = s\mu =_{d,\alpha} \mathrm{B}_V(s)\,\mathrm{B}_V(\mu) = \mathrm{B}_V(s\mu) = \mathrm{B}(\tau), \tag{2.2.102}$$

$$\begin{aligned}
V \cap \mathrm{BV}\left(\mathrm{B}_V(\tau)\right) &= V \cap \mathrm{BV}\left(\mathrm{B}_V(s\mu)\right) = V \cap \mathrm{BV}\left(\mathrm{B}_V(s)\,\mathrm{B}_V(\mu)\right) \\
&= V \cap \left(\mathrm{BV}\left(\mathrm{B}_V(s)\right) \cup \mathrm{BV}\left(\mathrm{B}_V(\mu)\right)\right) \\
&= \left(V \cap \mathrm{BV}\left(\mathrm{B}_V(s)\right)\right) \cup \left(V \cap \mathrm{BV}\left(\mathrm{B}_V(\mu)\right)\right) \\
&= \varnothing \cup \varnothing = \varnothing.
\end{aligned} \tag{2.2.103}$$

4. Assume now that $\tau = 0$. Then, $\tau = 0 = \mathrm{B}_V(0) = \mathrm{B}_V(\tau)$, so by the reflexivity of $=_{d,\alpha}$, $\tau =_{d,\alpha} \mathrm{B}_V(\tau)$. Also, $V \cap \mathrm{BV}(\mathrm{B}_V(\tau)) = V \cap \mathrm{BV}(\mathrm{B}_V(0)) = V \cap \mathrm{BV}(0) = V \cap \varnothing = \varnothing$.

5. Assume now that $\tau = r \cdot \mu$, where $r \in \mathcal{R}$ is a scalar and $\mu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. By the induction hypothesis on $\mu$, $\mu =_{d,\alpha} \mathrm{B}_V(\mu)$ and $V \cap \mathrm{BV}(\mathrm{B}_V(\mu)) = \varnothing$. Hence, $\tau = r \cdot \mu = r \cdot \mathrm{B}_V(\mu) = \mathrm{B}_V(r \cdot \mu) = \mathrm{B}_V(\tau)$ and $V \cup \mathrm{BV}(\mathrm{B}_V(\tau)) = V \cup \mathrm{BV}(\mathrm{B}_V(r \cdot \mu)) = V \cup \mathrm{BV}(r \cdot \mathrm{B}_V(\mu)) = V \cup \mathrm{BV}(\mathrm{B}_V(\mu)) = \varnothing$.

6. Assume now that $\tau = \mu + \nu$, where $\mu, \nu \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms. By the induction hypothesis on $\mu$, $\mu =_{d,\alpha} \mathrm{B}_V(\mu)$ and $V \cap \mathrm{BV}(\mathrm{B}_V(\mu)) = \varnothing$, and by the induction hypothesis on $\nu$, $\nu =_{d,\alpha} \mathrm{B}_V(\nu)$ and $V \cap \mathrm{BV}(\mathrm{B}_V(\nu)) = \varnothing$. Hence,

$$\tau = \mu + \nu =_{d,\alpha} \mathrm{B}_V(\mu) + \mathrm{B}_V(\nu) = \mathrm{B}_V(\mu + \nu) = \mathrm{B}(\tau), \tag{2.2.104}$$

$$\begin{aligned}
V \cap \mathrm{BV}\left(\mathrm{B}_V(\tau)\right) &= V \cap \mathrm{BV}\left(\mathrm{B}_V(\mu + \nu)\right) = V \cap \mathrm{BV}\left(\mathrm{B}_V(\mu) + \mathrm{B}_V(\nu)\right) \\
&= V \cap \left(\mathrm{BV}\left(\mathrm{B}_V(\mu)\right) \cup \mathrm{BV}\left(\mathrm{B}_V(\nu)\right)\right) \\
&= \left(V \cap \mathrm{BV}\left(\mathrm{B}_V(\mu)\right)\right) \cup \left(V \cap \mathrm{BV}\left(\mathrm{B}_V(\nu)\right)\right) \\
&= \varnothing \cup \varnothing = \varnothing.
\end{aligned} \tag{2.2.105}$$

This completes the induction. ❏

**Lemma 2.2.19**

Let $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ be a complex differential pre-term.

(i) Given a simple differential pre-term $u \in \Delta_{d,\mathcal{V}}$ and a variable $x \in \mathcal{V}$, there exists a complex differential pre-term $\tau' \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ such that $\tau =_{d,\alpha} \tau'$ and $u$ can be substituted for $x$ into $\tau'$ (i.e., $\mathrm{Sub}_{x,u}(\tau') = 1$).

(ii) Given simple differential pre-terms $u, v \in \Delta_{d,\mathcal{V}}$ and variables $x, y \in \mathcal{V}$, there exist a complex differential pre-term $\tau' \in \Delta_{d,\mathcal{V}}$ and a simple differential pre-terms $u' \in \Delta_{d,\mathcal{V}}$ such that $\tau =_{d,\alpha} \tau'$, $u =_{d,\alpha} u'$, $u'$ can be substituted for $x$ in $\tau'$ (i.e., $\mathrm{Sub}_{x,u'}(\tau') = 1$), and $v$ can be substituted for $y$ in $\tau'[u'/x]$ (i.e., $\mathrm{Sub}_{y,v}(\tau'[u'/x]) = 1$). ♡

***Proof of Part (i)***. Let $\tau' = \mathrm{B}_{\mathrm{FV}(u)}(\tau)$. By Lemma 2.2.18, $\tau =_{d,\alpha} \tau'$ and $\mathrm{FV}(u) \cap \mathrm{BV}(\tau') = \varnothing$. Since $\mathrm{FV}(u) \cap \mathrm{BV}(\tau') = \varnothing$, trivially, $u$ can be substituted for $x$ in $\tau'$ (i.e., $\mathrm{Sub}_{x,u}(\tau') = 1$). ❑

***Proof of Part (ii)***. Let $V = \mathrm{FV}(u) \cup \mathrm{FV}(v)$ be a subset of variables, and let $\tau' = \mathrm{B}_V(\tau)$ and $u' = \mathrm{B}_V(u)$. By Lemma 2.2.18, $\tau =_{d,\alpha} \tau'$ and $V \cap \mathrm{BV}(\tau') = \varnothing$. Similarly, by Lemma 2.2.18, $u =_{d,\alpha} u'$ and $V \cap \mathrm{BV}(u') = \varnothing$. Since $u =_{d,\alpha} u'$, by Lemma 2.2.16, $\mathrm{FV}(u) = \mathrm{FV}(u')$. Therefore, $\mathrm{FV}(u') \cap \mathrm{BV}(\tau') = \mathrm{FV}(u) \cap \mathrm{BV}(\tau') \subset (\mathrm{FV}(u) \cup \mathrm{FV}(v)) \cap \mathrm{BV}(\tau') = V \cap \mathrm{BV}(\tau') = \varnothing$, which implies that $u'$ can be substituted for $x$ in $\tau'$ (i.e., $\mathrm{Sub}_{x,u}(\tau') = 1$). Also, $\mathrm{FV}(v) \cap \mathrm{BV}(\tau'[u'/x]) \subset (\mathrm{FV}(u) \cup \mathrm{FV}(v)) \cap (\mathrm{BV}(\tau') \cup \mathrm{BV}(u')) = V \cap (\mathrm{BV}(\tau') \cup \mathrm{BV}(u')) = (V \cap \mathrm{BV}(\tau')) \cup (V \cap \mathrm{BV}(u')) = \varnothing \cup \varnothing = \varnothing$, which implies that $v$ can be substituted for $y$ in $\tau'[u'/x]$ (i.e., $\mathrm{Sub}_{x,u}(\tau') = 1$). ❑

**Lemma 2.2.20**

Let $v \in \Delta_{d,\mathcal{V}}$ be a simple differential pre-term.

(i) If $D_1^n \lambda x.s * (t_1, \ldots, t_n) =_{d,\alpha} v$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $x \in \mathcal{V}$ are variables, and $s, t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, then $v = D_1^n \lambda y.u * (t'_{f(1)}, \ldots, t'_{f(n)})$, where $y \in \mathcal{V}$ is a variable, $u, t'_1, \ldots, t'_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $f \in \mathfrak{S}_n$ such that $t_j =_{d,\alpha} t'_j$ for all $1 \leqslant j \leqslant n$ and there exists a simple differential pre-term $s' \in \Delta_{d,\mathcal{V}}$ such that $s =_{d,\alpha} s'$, $y$ can be substituted for $x$ in $s'$ (i.e., $\mathrm{Sub}_{x,y}(s') = 1$), and $s'[y/x] =_{d,\alpha} u$.

(ii) If $D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n) =_{d,\alpha} v$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, $t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms, and $x \in \mathcal{V}$ is a variable, then $v = D_{i_1,\ldots,i_n} x * (t'_{f(1)}, \ldots, t'_{f(n)})$, where $t_1, \ldots, t_n \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms and $f \in \mathfrak{S}_n$ is a permutation such that $t_j =_{d,\alpha} t'_j$ for all $1 \leqslant j \leqslant n$.

(iii) If $s\mu =_{d,\alpha} v$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\mu$ is a complex differential pre-term, then $v = s'\mu'$, where $s' \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term and $\mu' \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term such that $s =_{d,\alpha} s'$ and $\mu =_{d,\alpha} \mu'$. ♡

***Proof***. Part (i) follows by induction on $D_1^n \lambda x.s * (t_1, \ldots, t_n) =_{d,\alpha} v$. The only interesting case is when this $\alpha$-equivalence holds because of transitivity. That is, there exists a simple differential $w \in \Delta_{d,\mathcal{V}}$ such that $D_1^n \lambda x.s * (t_1, \ldots, t_n) =_{d,\alpha} w$ and $w =_{d,\alpha} v$. By the induction hypothesis on the former $\alpha$-equivalence, there exist a variable $y \in \mathcal{V}$, simple differential pre-terms $u, s', t'_1, \ldots, t'_n \in \Delta_{d,\mathcal{V}}$, and a permutation $f \in \mathfrak{S}_n$ such that $w = D_1^n \lambda y.u * (t'_{f(1)}, \ldots, t'_{f(n)})$, $t_j =_{d,\alpha} t'_j$ for all $1 \leqslant j \leqslant n$, $s =_{d,\alpha} s'$, $\mathrm{Sub}_{x,y}(s') = 1$, $s'[y/x] =_{d,\alpha} u$. Since $w =_{d,\alpha} v$, this implies that $D_1^n \lambda y.u * (t'_{f(1)}, \ldots, t'_{f(n)}) = v$, so by the induction hypothesis, there exist a variable $z \in \mathcal{V}$, simple differential pre-terms $u', s'', t''_1, \ldots, t''_n \in \Delta_{d,\mathcal{V}}$, and a permutation $f' \in \mathfrak{S}_n$ such that $v = D_1^n \lambda z.u' * (t''_{f'(f(1))}, \ldots, t''_{f'(f(n))})$, $t'_j =_{d,\alpha} t''_j$ for all $1 \leqslant j \leqslant n$, $u =_{d,\alpha} s''$, $\mathrm{Sub}_{y,z}(s'') = 1$, $s''[z/y] =_{d,\alpha} u'$. Since $\mathfrak{S}_n$ is closed under composition, $f' \circ f \in \mathfrak{S}_n$. In addition, for each $1 \leqslant j \leqslant n$, $t_j =_{d,\alpha} t'_j$ and $t'_j =_{d,\alpha} t''_j$, so by the transitivity of $=_{d,\alpha}$, $t_j =_{d,\alpha} t''_j$. Finally, by Lemma 2.2.19(ii), there exists a simple differential pre-term $s''' \in \Delta_{d,\mathcal{V}}$ such that $s =_{d,\alpha} s'''$, $\mathrm{Sub}_{x,y}(s''') = 1$, and $\mathrm{Sub}_{y,z}(s'''[y/x]) = 1$. Since $s =_{d,\alpha} s'''$ and $s =_{d,\alpha} s'$, $s''' =_{d,\alpha} s'$, so since $\mathrm{Sub}_{x,y}(s''') = 1$ and $\mathrm{Sub}_{x,y}(s') = 1$, by Lemma 2.2.17, $s'''[y/x] =_{d,\alpha} s'[y/x] =_{d,\alpha} u =_{d,\alpha} s''$. Thus, since $s'''[y/x] =_{d,\alpha} s''$, $\mathrm{Sub}_{y,z}(s'''[y/x]) = 1$, and $\mathrm{Sub}_{y,z}(s'') = 1$, by Lemma 2.2.17 again, $s'''[y/x][z/y] =_{d,\alpha} s'[z/y] =_{d,\alpha} u'$. Now, let us show that $\mathrm{Sub}_{x,z}(s''') = 1$ and $s'''[y/x][z/y] = s'''[z/x]$. Since $D_1^n \lambda x.s * (t_1, \ldots, t_n) =_{d,\alpha} v = D_1^n \lambda z.u' * (t''_{f'(f(1))}, \ldots, t''_{f'(f(n))})$, by the definition of $=_{d,\alpha}$, it must be the case that $y \notin \mathrm{FV}(s)$ or $y \neq x$. Also, since $s =_{d,\alpha} s'''$, by Lemma 2.2.16, $\mathrm{FV}(s) = \mathrm{FV}(s''')$. Hence, $y \notin \mathrm{FV}(s''')$ or $x = y$. Suppose that $x = y$. Then, by Lemma 2.2.11, $s'''[y/x] = s'''[x/x] = s'''$, so $1 = \mathrm{Sub}_{y,z}(s'''[y/z]) = \mathrm{Sub}_{x,z}(s''')$ and $s'''[y/x][z/y] = s'''[z/x]$. Instead, suppose that $x \neq y$. Then, $y \notin \mathrm{FV}(s''')$. Since $\mathrm{Sub}_{x,y}(s''') = \mathrm{Sub}_{y,z}(s'''[y/x]) = \mathrm{Sub}_{y,z}(y) = 1$, $x \neq y$, $y \notin \mathrm{FV}(s''')$, by Lemma 2.2.12, $1 = \mathrm{Sub}_{x,y[z/y]}(s'''[z/y]) = \mathrm{Sub}_{x,z}(s'''[z/y])$ and $s'''[y/x][z/y] = s'''[z/y][y[z/y]/x] = s'''[z/y][z/x]$. Since $y \notin \mathrm{FV}(s''')$, by Lemma 2.2.9, $s'''[z/y] = s'''$, so $1 = \mathrm{Sub}_{x,z}(s'''[z/y]) = \mathrm{Sub}_{x,z}(s''')$ and $s'''[z/y][z/x] = s'''[z/x]$. Hence, in either case, $\mathrm{Sub}_{x,z}(s''') = 1$ and $s'''[y/x][z/y] = s'''[z/x]$. Since $s'''[y/x][z/y] =_{d,\alpha} u'$, this implies that $s'''[z/x] =_{d,\alpha} u'$. In summary, $v = D_1^n \lambda z.u' * (t''_{(f' \circ f)(1))}, \ldots, t''_{(f' \circ f)(n))})$, where $f' \circ f \in \mathfrak{S}_n$, $t_j =_{d,\alpha} t''_j$ for all $1 \leqslant j \leqslant n$, $\mathrm{Sub}_{x,z}(s''') = 1$, and $s'''[z/x] =_{d,\alpha} u'$. This completes the proof in the case of transitivity.

The proof of parts (ii)-(iii) is similar. The details are left to the reader. ❏

**Lemma 2.2.21**

(i) If $\tau \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term, then there exist a nonnegative integer $n \in \mathbb{N}_0$, scalars $a_1, \ldots, a_n \in \mathcal{R}$, and distinct simple differential pre-terms $u_1, \cdots, u_n \in \Delta_{d,\mathcal{V}}$ such that $\tau =_{d,\alpha} a_1 \cdot u_1 + \cdots + a_n \cdot u_n$.

(ii) If $n, m \in \mathbb{N}_0$ are nonnegative integers, $a_1, \ldots, a_n, b_1, \ldots, b_m \in \mathcal{R}$ are scalars, $u_1, \cdots, u_n \in \Delta_{d,\mathcal{V}}$ are distinct simple differential pre-terms, and $v_1, \cdots, v_n \in \Delta_{d,\mathcal{V}}$ are (possible different) distinct simple differential pre-terms such that $a_1 \cdot u_1 + \cdots + a_n \cdot u_n =_{d,\alpha} b_1 \cdot v_1 + \cdots + b_m \cdot v_m$, then $n = m$ and there exists a permutation $f \in \mathfrak{S}_n$ such that $b_j =_{d,\alpha} a_{f(j)}$ and $v_j =_{d,\alpha} u_{f(j)}$ for all $1 \leqslant j \leqslant n = m$. ♡

***Proof of Part (i).*** This follows by induction on $\tau$.

1. Assume that $\tau = s$, where $s \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term. Let $n = 1$, $a_1 = 1$, and $u_1 = s$. Then, $\tau = s =_{d,\alpha} 1 \cdot s = a_1 \cdot u_1$.

2. Assume that $\tau = 0$. Let $n = 0$. Then, the result is trivial.

3. Assume now that $\tau = r \cdot \mu$, where $r \in \mathcal{R}$ is a scalar and $\mu \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ is a complex differential pre-term. By the induction hypothesis on $\mu$, there exists a nonnegative integer $n \in \mathcal{N}_0$, scalars $b_1, \ldots, b_n \in \mathcal{R}$, and distinct simple differential pre-terms $u_1, \cdots, u_n \in \Delta_{d,\mathcal{V}}$ such that $\mu =_{d,\alpha} b_1 \cdot u_1 + \cdots + b_n \cdot u_n$. For each $1 \leqslant j \leqslant n$, let $a_j = r \cdot b_j$. Then, $\tau = r \cdot \mu = r \cdot (a_1 \cdot u_1 + \cdots + a_n \cdot u_n) =_{d,\alpha} r \cdot (b_1 \cdot u_1) + \cdots + r \cdot (b_n \cdot u_n) =_{d,\alpha} (r \cdot b_1) \cdot u_1 + \cdots + (r \cdot b_n) \cdot u_n = a_1 \cdot u_1 + \cdots + a_n \cdot u_n$.

4. Assume now that $\tau = \mu + \nu$, where $\mu, \nu \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential pre-terms. By the induction hypothesis on $\mu$, there exists a nonnegative integer $\ell \in \mathbb{N}_0$, scalars $b_1, \ldots, b_\ell \in \mathcal{R}$, and distinct simple differential pre-terms $v_1, \ldots, v_\ell \in \Delta_{d,\mathcal{V}}$ such that $\mu =_{d,\alpha} b_1 \cdot v_1 + \cdots + b_\ell \cdot v_\ell$. By the induction hypothesis on $\nu$, there exists a nonnegative integer $m \in \mathbb{N}_0$, scalars $c_1, \ldots, c_m \in \mathcal{R}$, and distinct simple differential pre-terms $w_1, \ldots, w_m \in \Delta_{d,\mathcal{V}}$ such that $\nu =_{d,\alpha} c_1 \cdot w_1 + \cdots + c_m \cdot w_m$. Let $J = \{j \in \{1, \ldots, \ell\} : \exists i \in \{1, \ldots, m\} \text{ s.t. } v_j = w_i\}$, let $f\colon J \to \{1, \ldots, m\}$ be the unique injective function such that $v_j = w_{f(j)}$ for each $j \in J$, and let $g : \{1, \ldots, m - |J|\} \to \{1, \ldots, m\} - f(J)$ be the unique order-preserving bijection. Let $n = \ell + m - |J|$. For each $1 \leqslant j \leqslant \ell$, let $u_j = v_j$, and if $j \notin J$, $a_j = b_j$, otherwise $a_j = b_j + c_{f(j)}$.

For each $1 \leqslant j \leqslant m - |J|$, let $a_{\ell+j} = c_{g(j)}$ and $u_{\ell+j} = v_{g(j)}$. Then, $a_1, \dots a_n$ are distinct simple differential terms and

$$\begin{aligned}
\tau = \mu + \nu &= \left(\sum_{j=1}^{\ell} b_j \cdot v_j\right) + \left(\sum_{j=1}^{m} c_j \cdot w_j\right) \\
&=_{d,\alpha} \sum_{j=1, j \notin J}^{\ell} b_j \cdot v_j + \sum_{j \in J} b_j \cdot v_j + \sum_{j \in f(J)} c_j \cdot w_j + \sum_{j=1, j \notin f(J)}^{m} c_j \cdot w_j \\
&= \sum_{j=1, j \notin J}^{\ell} b_j \cdot v_j + \sum_{j \in J} b_j \cdot v_j + \sum_{j \in J} c_{f(j)} \cdot w_{f(j)} + \sum_{j=1}^{m-|J|} c_{g(j)} \cdot w_{g(j)} \\
&= \sum_{j=1, j \notin J}^{\ell} b_j \cdot v_j + \sum_{j \in J} b_j \cdot v_j + \sum_{j \in J} c_{f(j)} \cdot v_j + \sum_{j=1}^{m-|J|} c_{g(j)} \cdot w_{g(j)} \\
&=_{d,\alpha} \sum_{j=1, j \notin J}^{\ell} b_j \cdot v_j + \sum_{j \in J} (b_j \cdot v_j + c_{f(j)} \cdot v_j) + \sum_{j=1}^{m-|J|} c_{g(j)} \cdot w_{g(j)} \\
&=_{d,\alpha} \sum_{j=1, j \notin J}^{\ell} b_j \cdot v_j + \sum_{j \in J} (b_j + c_{f(j)}) \cdot v_j + \sum_{j=1}^{m-|J|} c_{g(j)} \cdot w_{g(j)} \\
&= \sum_{j=1, j \notin J}^{\ell} a_j \cdot u_j + \sum_{j \in J} a_j \cdot u_j + \sum_{j=1}^{m-|J|} a_{\ell+j} \cdot u_{\ell+j} \\
&= a_1 \cdot u_1 + \cdots + a_n \cdot u_n.
\end{aligned} \tag{2.2.106}$$

This completes the induction. ❏

***Proof of Part (ii)***. Trivial. The proof is left to the reader. ❏

With all these properties of $\alpha$-conversion for differential pre-terms defined, we can finally define differential terms, the main element we work with in differential $\lambda$-calculus. Differential terms, as we indicated above, are just differential pre-terms where we identify differential pre-terms by $\alpha$-equivalence.

**Definition 2.2.22. Differential Terms**

The set of *simple differential terms* is the set of simple differential pre-terms $\Delta_{d,\mathcal{V}}$ quotiented by $\alpha$-equivalence $=_{d,\alpha}$. That is,

$$\Lambda_{d,\mathcal{V}} := \{[s]_{d,\alpha} : s \in \Delta_{d,\mathcal{V}}\}, \tag{2.2.107}$$

where $[s]_{d,\alpha} := \{t \in \Delta_{d,\mathcal{V}} : s =_{d,\alpha} t\}$. Similarly, the set of *complex differential terms* is the set of complex differential pre-terms $\mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ quotiented by $\alpha$-equivalence $=_{d,\alpha}$. That is,

$$\mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle := \{[\tau]_{d,\alpha} : \tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle\}, \tag{2.2.108}$$

where $[s]_{d,\alpha} := \{\sigma \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle : \tau =_{d,\alpha} \sigma\}$.

♣

We can now lift the notion of free variables and substitution from complex differential pre-terms to complex differential terms.

**Definition 2.2.23. Free Variables for Differential Terms**

Define the function $\mathrm{FV}\colon \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle \to \mathcal{V}^*$ that takes a complex differential term to a finite subset of variables by

$$\mathrm{FV}\,([\tau']_{d,\alpha}) = \mathrm{FV}\,(\tau'), \tag{2.2.109}$$

where $\tau' \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term. Note that FV is well-defined by Lemma 2.2.16. The image of a complex differential term $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ under FV will be called the *free variables* of $\tau$, and if $\mathrm{FV}\,(\tau) = \varnothing$, then $\tau$ is said to be *closed*. ♣

**Definition 2.2.24. Substitution for Differential Terms**

For a complex differential term $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ and a simple differential term $t \in \Lambda_{d,\mathcal{V}}$, the substitution $\tau[t/x]$ of $t$ for $x$ in $\tau$ is defined by

$$\tau[t/x] = [\tau'[t'/x]]_{d,\alpha}, \tag{2.2.110}$$

where $\tau' \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term and $t' \in \Delta_{d,\mathcal{V}}$ is a simple differential pre-term such that $\tau = [\tau']_{d,\alpha}$, $t = [t']_{d,\alpha}$, and $t'$ can be substituted for $x$ in $\tau'$ (i.e., $\mathrm{Sub}_{x,t'}\,(\tau') = 1$). The existence of such a $\tau' \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ and $t' \in \Delta_{d,\mathcal{V}}$ is guaranteed by Lemma 2.2.19(i), and any such choice of $\tau' \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ and $t' \in \Delta_{d,\mathcal{V}}$ yields the same result by Lemma 2.2.17. ♣

The differential pre-term operations can themselves be lifted to differential terms.

**Notation 2.2.25. Operations on Simple Differential Terms**

Let $n \in \mathbb{N}_0$ be a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ be positive integers, let $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ be simple differential terms, let $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be a complex differential term, and let $x \in \mathcal{V}$ be a variable. Then, $D_{i_1,\ldots,i_n}x * (t_1, \ldots, t_n)$, $D_1^n\lambda x.s * (t_1, \ldots, t_n)$, and $s\tau$ are the unique simple differential terms such that

$$\begin{aligned} D_{i_1,\ldots,i_n}x * (t_1, \ldots, t_n) &:= [D_{i_1,\ldots,i_n}x * (t_1', \ldots, t_n')]_{d,\alpha}, && (2.2.111a)\\ D_1^n\lambda x.s * (t_1, \ldots, t_n) &:= [D_1^n\lambda x.s' * (t_1', \ldots, t_n')]_{d,\alpha}, && (2.2.111b)\\ s\tau &:= [s'\tau']_{d,\alpha}, && (2.2.111c)\end{aligned}$$

where $s, t_1, \ldots, t_n \in \Delta_{d,\mathcal{V}}$ are simple differential pre-terms and $\tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ is a complex differential pre-term such that $s = [s_i]_{d,\alpha}$, $t_j = [t_j']_{d,\alpha}$ for all $1 \leqslant j \leqslant n$, and $\tau = [\tau']_{d,\alpha}$. The uniqueness follows from Definition 2.2.14(ii)-(iv). ♠

**Notation 2.2.26. Operations on Complex Differential Terms**

Let $r \in \mathcal{R}$ be a scalar and $\tau, \sigma \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential terms. Then, $0$, $r \cdot \tau$, and $\tau + \sigma$ are the unique complex differential terms such that

$$\begin{aligned} 0 &:= [0]_{d,\alpha}, && (2.2.112a)\\ r \cdot \tau &= [r \cdot \tau']_{d,\alpha}, && (2.2.112b)\\ \tau + \sigma &= [\tau' + \sigma']_{d,\alpha}, && (2.2.112c)\end{aligned}$$

where $\tau', \sigma' \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ are complex differential pre-terms such that $\tau = [\tau']_{d,\alpha}$ and $\sigma = [\sigma']_{d,\alpha}$. The uniqueness follows from Definition 2.2.14(v)-(vi). ♠

Using this notation, we can show that the equations defining free variables and substitution for differential pre-terms also hold for differential terms.

**Lemma 2.2.27**

Let $n \in \mathbb{N}_0$ be a nonnegative integer, let $i_1, \ldots, i_n \in \mathbb{N}$ be positive integers, let $r_1, \ldots, r_n \in \mathcal{R}$ be scalars, let $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ be simple differential terms, let $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be a complex differential term, and let $x \in \mathcal{V}$ be a variable. Then, the following equations regarding free variables of differential terms are valid:

$$\mathrm{FV}\,(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) = \{x\} \cup \mathrm{FV}\,(t_1) \cup \cdots \cup \mathrm{FV}\,(t_n), \tag{2.2.113a}$$
$$\mathrm{FV}\,(D_1^n \lambda x.s * (t_1, \ldots, t_n)) = (\mathrm{FV}\,(s) - \{x\}) \cup \mathrm{FV}\,(t_1) \cup \cdots \cup \mathrm{FV}\,(t_n), \tag{2.2.113b}$$
$$\mathrm{FV}\,(s\tau) = \mathrm{FV}\,(s) \cup \mathrm{FV}\,(\tau), \tag{2.2.113c}$$
$$\mathrm{FV}\,(r_1 \cdot t_1 + \cdots + r_n \cdot t_n) = \mathrm{FV}\,(t_1) \cup \cdots \cup \mathrm{FV}\,(t_n). \tag{2.2.113d}$$

♡

***Proof.*** This is an immediate result of Notation 2.2.25-2.2.26 and Definitions 2.2.5, 2.2.23. ❑

**Lemma 2.2.28**

Let $n \in \mathbb{N}_0$ be a nonnegative integer, let $i_1, \ldots, i_n \in \mathbb{N}$ be positive integers, let $r_1, \ldots, r_n \in \mathcal{R}$ be scalars, let $s, t, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ be simple differential terms, let $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be a complex differential term, and let $x, y \in \mathcal{V}$ be distinct variables. Then, the following equations regarding substitution of differential terms are valid:

$$(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n))[t/x] = D_{i_1,\ldots,i_n} t * (t_1[t/x], \ldots, t_n[t/x]), \tag{2.2.114a}$$
$$(D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n))[t/x] = D_{i_1,\ldots,i_n} y * (t_1[t/x], \ldots, t_n[t/x]), \tag{2.2.114b}$$
$$(D_1^n \lambda x.s * (t_1, \ldots, t_n))[t/x] = D_1^n \lambda x.s * (t_1[t/x], \ldots, t_n[t/x]), \tag{2.2.114c}$$
$$(D_1^n \lambda y.s * (t_1, \ldots, t_n))[t/x] = D_1^n \lambda y.s[t/x] * (t_1[t/x], \ldots, t_n[t/x]) \quad \text{if } y \notin \mathrm{FV}\,(t), \tag{2.2.114d}$$
$$(s\mu)[t/x] = s[t/x]\mu[t/x], \tag{2.2.114e}$$
$$(r_1 \cdot t_1 + \cdots + r_n \cdot t_n)[t/x] = r_1 \cdot t_1[t/x] + \cdots r_n \cdot t_n[t/x]. \tag{2.2.114f}$$

♡

***Proof.*** This is an immediate result of Notation 2.2.25-2.2.26 and Definitions 2.2.8, 2.2.24. ❑

The next lemma lifts a few properties of differential pre-terms to the level of differential terms.

**Lemma 2.2.29**

Let $n \in \mathbb{N}_0$ be a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ be positive integers, $x, y \in \mathcal{V}$ be variables, $u, v, s, s', t_1, \ldots, t_n, t'_1, \ldots, t'_n \in \Lambda_{d,\mathcal{V}}$ be simple differential terms, and $\tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential terms. Then, the following statements are true:

(i) If $x \notin \mathrm{FV}\,(\tau)$, then $\tau[u/x] = \tau$.

(ii) $y \in \mathrm{FV}\,(\tau[u/x])$ if and only if either $y \in \mathrm{FV}\,(\tau)$ and $x \neq y$ or $y \in \mathrm{FV}\,(u)$ and $x \in \mathrm{FV}\,(\tau)$.

(iii) $\tau[x/x] = \tau$.

(iv) If $y \neq x \notin \mathrm{FV}\,(v)$, then $\tau[u/x][v/y] = \tau[v/y][u[v/y]/x]$.

(v) If $y \notin \mathrm{FV}\,(\tau)$, then $\tau[y/x][u/y] = \tau[u/x]$.

(vi) If $y \notin \mathrm{FV}\,(s)$, then $D_1^n \lambda x.s * (t_1, \ldots, t_n) = D_1^n \lambda y.s[y/x] * (t_1, \ldots, t_n)$.

(vii) $D_1^n \lambda x.s * (t_1, \ldots, t_n) = D_1^n \lambda y.s' * (t'_1, \ldots, t'_n)$ if and only if there exists a permutation $f \in \mathcal{S}_n$ such that $t'_j = f(t_j)$ for all $1 \leqslant j \leqslant n$, $s[y/x] = s'$, and $s'[x/y] = s$.

(viii) $D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n) = D_{i_1,\ldots,i_n} x * (t'_1, \ldots, t'_n)$ if and only if there exists a permutation $f \in \mathcal{S}_n$ such that $t'_j = f(t_j)$ for all $1 \leqslant j \leqslant n$.

(ix) $s\tau = s'\tau'$ if and only if $s = s'$ and $\tau = \tau'$.

(x) Summation makes the set of complex differential terms $\mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ into a commutative monoid with 0 as the unit.

(xi) $\mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a left-$\mathcal{R}$-module on the commutative monoid $(\mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle, +, 0)$.

(xii) There exists a unique nonnegative integer $n \in \mathbb{N}_0$, unique scalars $r_1, \ldots, r_n \in \mathcal{R}$, and unique distinct simple differential terms $s_1, \cdots, s_n \in \Lambda_{d,\mathcal{V}}$ such that $\tau = r_1 \cdot s_1 + \cdots + r_n \cdot s_n$.

**Note** that we define the *support* $\operatorname{supp}(\tau)$ of $\tau$ to be this unique set of distinct simple differential terms. That is, $\operatorname{supp}(\tau) = \{s_1, \ldots, s_n\}$. ♡

***Proof***. Parts (i)-(v) follow from Lemmas 2.2.9-2.2.12 and Lemma 2.2.19. Part (vi) follows from Definition 2.2.14(i) and Lemma 2.2.19(i). The "if" part of Parts (vii)-(ix) follows from Definition 2.2.14(ii)-(iv). The "only if" part of Parts (vii)-(ix) follows from Lemma 2.2.20. Part (x) follows from Definition 2.2.14(ix)-(xi). Part (xi) follows from Definition 2.2.14(xii)-(xv). Part (xi) follows from Lemma 2.2.21. ❏

From now on, expressions involving abstractions, applications, differentiation, scaling, sums, and variables are always understood as differential terms, as defined in Notation 2.2.25-2.2.26. In particular, with the present machinery at hand, we can formulate definitions and write proofs by induction on the structure of differential terms, rather than first introducing the relevant notions for differential pre-terms and then lifting.

## 2.3 Defining the Partial Derivative

In this section, we are going to define two closely related concepts: *linearization* and *partial differentiation*. The reason we need both of these notions when they appear so similar is hard to see at first, so for now we will just focus on talking about linearization. Given differential terms $f$ and $g$ representing analytic functions of type $\mathbb{R}^m \to \mathbb{R}$, intuitively, the linearization $\mathcal{D}_i f * g$ of $f$ with respect to the $i$th argument times $g$ is a function whose type is the same as $f$ and is given by

$$x_1, \ldots, x_n \mapsto (\partial_{x_i} f)(x_1, \ldots, x_m) \cdot g(x_1, \ldots, x_m). \tag{2.3.1}$$

Using this informal definition of linearization and the correspondence between differential terms and analytic function as laid out in section 2.1, one arrives to the following definition of linearization.

**Definition 2.3.1. Linearization for Differential Terms**

Let $\sigma, \tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$ be complex differential terms and $i \in \mathbb{N}$ be a positive integer. The *linearization* $\mathcal{D}_i \tau * \sigma$ of $\tau$ with respect to the $i$th argument times $\sigma$ is defined by:

$$\mathcal{D}_i(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) * t = D_{i,i_1,\ldots,i_n} x * (t, t_1, \ldots, t_n) \tag{2.3.2a}$$

$$\mathcal{D}_i(D_1^n \lambda x.s * (t_1, \ldots, t_n)) * t = \begin{cases} D_1^{n+1} \lambda x.s * (t, t_1, \ldots, t_n) & \text{if } i = 1, \\ D_1^n \lambda x.(\mathcal{D}_{i-1} s * t) * (t_1, \ldots, t_n) & \text{if } i > 1 \wedge x \notin \mathrm{FV}(\sigma), \end{cases} \tag{2.3.2b}$$

$$\mathcal{D}_i(s\mu) * t = (\mathcal{D}_{i+1} s * t)\mu, \tag{2.3.2c}$$

$$\mathcal{D}_i \left( \sum_{j=1}^{m} a_j \cdot s_k \right) * \sum_{k=1}^{n} b_k \cdot t_k = \sum_{j=1}^{m} \sum_{k=1}^{n} (a_j \cdot b_k) \cdot \mathcal{D}_i s_j * t_k. \tag{2.3.2d}$$

where $m, n \in \mathbb{N}_0$ are nonnegative integers, $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathcal{R}$ are scalars, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, $s, t, s_1, \ldots, s_n, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms, and $x \in \mathcal{V}$ is a variable. ♣

Since differential $\lambda$-calculus is a calculus representing analytic functions and partial derivatives commute when applied to analytic functions, the order in which we perform linearization should not matter, and that is exactly what we prove in the next lemma.

**Lemma 2.3.2**

Given positive integers $i, j \in \mathbb{N}$ and simple differential terms $s, u, v \in \Lambda_{d,\mathcal{V}}$,

$$\mathcal{D}_i(\mathcal{D}_j s * u) * v = \mathcal{D}_j(\mathcal{D}_i s * v) * u. \tag{2.3.3}$$ ♡

***Proof.*** This follows by induction on $s$.

1. We prove this via induction on s. Let us assume that $s = D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)$, where $i_1, \ldots, i_n \in \mathbb{N}$ are nonnegative integers, $x \in \mathcal{V}$ is a variable, and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Then,

$$\begin{aligned} \mathcal{D}_i(\mathcal{D}_j s * u) * v &= \mathcal{D}_i(\mathcal{D}_j(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) * u) * v \\ &= \mathcal{D}_i(D_{j,i_1,\ldots,i_n} x * (u, t_1, \ldots, t_n)) * v \\ &= D_{i,j,i_1,\ldots,i_n} x * (v, u, t_1, \ldots, t_n) \\ &= D_{j,i,i_1,\ldots,i_n} x * (u, v, t_1, \ldots, t_n) \\ &= \mathcal{D}_j(D_{i,i_1,\ldots,i_n} x * (v, t_1, \ldots, t_n)) * u \\ &= \mathcal{D}_j(\mathcal{D}_i(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) * v) * u \\ &= \mathcal{D}_j(Dis * v) * u. \end{aligned} \tag{2.3.4}$$

2. Assume that $s = D_1^n \lambda x.t * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $x \in \mathcal{V}$ is a variable, and $t, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Further, suppose that $i = j = 1$. Then,

$$\begin{aligned} \mathcal{D}_i(\mathcal{D}_j s * u) * v &= \mathcal{D}_1(\mathcal{D}_1(D_1^n \lambda x.t * (t_1, \ldots, t_n)) * u) * v \\ &= \mathcal{D}_1(D_1^{n+1} \lambda x.t * (u, t_1, \ldots, t_n)) * v \\ &= D_1^{n+2} \lambda x.t * (v, u, t_1, \ldots, t_n) \\ &= D_1^{n+2} \lambda x.t * (u, v, t_1, \ldots, t_n) \\ &= D1(D_1^{n+1} \lambda x.t * (v, t_1, \ldots, t_n)) * u \\ &= D1(D1(D_1^n \lambda x.t * (t_1, \ldots, t_n)) * v) * u \\ &= \mathcal{D}_i(\mathcal{D}_j s * v) * u. \end{aligned} \tag{2.3.5}$$

Instead, suppose that $i, j > 1$. By the induction hypothesis, $D_{i-1}(\mathcal{D}_{j-1}t*u)*v = \mathcal{D}_{j-1}(D_{i-1}t*v)*u$. Hence,

$$\begin{aligned}\mathcal{D}_i(\mathcal{D}_j s * u) * v &= \mathcal{D}_i(\mathcal{D}_j(D_1^n \lambda x.t * (t_1, \ldots, t_n)) * u) * v \\ &= \mathcal{D}_i(D_1^n \lambda x.(\mathcal{D}_{j-1}t * u) * (t_1, \ldots, t_n)) * v \\ &= D_1^n \lambda x.(D_{i-1}(\mathcal{D}_{j-1}t * u) * v) * (t_1, \ldots, t_n) \\ &= D_1^n \lambda x.(\mathcal{D}_{j-1}(D_{i-1}t * v) * u) * (t_1, \ldots, t_n) \\ &= \mathcal{D}_j(D_1^n \lambda x.(D_{i-1}t * b) * (t_1, \ldots, t_n)) * u \\ &= \mathcal{D}_j(\mathcal{D}_i(D_1^n \lambda x.t * (t_1, \ldots, t_n)) * v) * u \\ &= \mathcal{D}_i(\mathcal{D}_j s * v) * u.\end{aligned} \tag{2.3.6}$$

Now, suppose that $i = 1 < j$. Then,

$$\begin{aligned}\mathcal{D}_i(\mathcal{D}_j s * u) * v &= \mathcal{D}_1(\mathcal{D}_j(D_1^n \lambda x.t * (t_1, \ldots, t_n)) * u) * v \\ &= \mathcal{D}_1(D_1^n \lambda x.(\mathcal{D}_{j-1}t * u) * (t_1, \ldots, t_n)) * v \\ &= D_1^{n+1} \lambda x.(\mathcal{D}_{j-1}t * u) * (v, t_1, \ldots, t_n) \\ &= \mathcal{D}_j(D_1^{n+1} \lambda x.t * (v, t_1, \ldots, t_n)) * u \\ &= \mathcal{D}_j(D_1(D_1^n \lambda x.t * (t_1, \ldots, t_n)) * v) * u \\ &= \mathcal{D}_i(\mathcal{D}_j s * v) * u.\end{aligned} \tag{2.3.7}$$

Finally, suppose $j = 1 < i$. Then,

$$\begin{aligned}\mathcal{D}_i(\mathcal{D}_j s * u) * v &= \mathcal{D}_i(\mathcal{D}_1(D_1^n \lambda x.t * (t_1, \ldots, t_n)) * u) * v \\ &= \mathcal{D}_i(D_1^{n+1} \lambda x.t * (u, t_1, \ldots, t_n)) * v \\ &= D_1^{n+1} \lambda x.(D_{i-1}t * v) * (u, t_1, \ldots, t_n) \\ &= \mathcal{D}_1(D_1^n \lambda x.(D_{i-1}t * v) * (t_1, \ldots, t_n)) * u \\ &= \mathcal{D}_1(\mathcal{D}_i(D_1^n \lambda x.t * (t_1, \ldots, t_n)) * v) * u \\ &= \mathcal{D}_i(\mathcal{D}_j s * v) * u.\end{aligned} \tag{2.3.8}$$

3. Assume that $s = t\mu$, where $t \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ is a complex differential term. By the induction hypothesis, $D_{i+1}(\mathcal{D}_{j+1}a * u) * v = \mathcal{D}_{j+1}(D_{i+1}a * v) * u$. Hence,

$$\begin{aligned}\mathcal{D}_i(\mathcal{D}_j s * u) * v &= \mathcal{D}_i(\mathcal{D}_j t\mu * u) * v \\ &= \mathcal{D}_i(\mathcal{D}_{j+1}t * u)\mu * v \\ &= (D_{i+1}(\mathcal{D}_{j+1}t * u) * v)\mu \\ &= (\mathcal{D}_{j+1}(D_{i+1}t * v) * u)\mu \\ &= \mathcal{D}_j(D_{i+1}t * v)\mu * u \\ &= \mathcal{D}_j(\mathcal{D}_i t\mu * v) * u \\ &= \mathcal{D}_j(\mathcal{D}_i s * v) * u.\end{aligned} \tag{2.3.9}$$

This completes the induction. ❏

**Corollary 2.3.3**

Given nonnegative integers $i, j \in \mathbb{N}$ and complex differential terms $\tau, \sigma, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$,

$$\mathcal{D}_i(\mathcal{D}_j\tau * \sigma) * \rho = \mathcal{D}_j(\mathcal{D}_i\tau * \rho) * \sigma. \tag{2.3.10}$$

♡

***Proof***. By Lemma 2.2.29(xii), there exist unique nonnegative integers $\ell, m, n \in \mathbb{N}_0$, scalars $a_1, \ldots, a_\ell$, $b_1 \ldots, b_m, c_1, \ldots, c_n \in \mathcal{R}$, and simple differential terms $s_1, \ldots, s_\ell, u_1, \ldots, u_m, v_1, \ldots, v_n \in \Lambda_{d,\mathcal{V}}$ such that $\tau = \sum_{x=1}^{\ell} r_x \cdot s_x$, $\sigma = \sum_{y=1}^{m} a_y \cdot u_y$, and $\rho = \sum_{z=1}^{n} b_z \cdot v_z$. Then,

$$\begin{aligned}
\mathcal{D}_i(\mathcal{D}_j\tau * \sigma) * \rho &= \mathcal{D}_i\left(\mathcal{D}_j\left(\sum_{x=1}^{\ell} r_x \cdot s_x\right) * \left(\sum_{y=1}^{m} a_y \cdot u_y\right)\right) * \left(\sum_{z=1}^{n} b_z \cdot v_z\right) \\
&= \sum_{x=1}^{\ell}\sum_{y=1}^{m}\sum_{z=1}^{n}(r_x \cdot a_y \cdot b_z) \cdot \mathcal{D}_i(\mathcal{D}_j s_x * u_y) * v_z \\
&= \sum_{x=1}^{\ell}\sum_{y=1}^{m}\sum_{z=1}^{n}(r_x \cdot a_y \cdot b_z) \cdot \mathcal{D}_j(\mathcal{D}_i s_x * v_z) * u_y \\
&= \sum_{x=1}^{\ell}\sum_{z=1}^{n}\sum_{y=1}^{m}(r_x \cdot a_z \cdot b_y) \cdot \mathcal{D}_j(\mathcal{D}_i s_x * v_z) * u_y \\
&= \mathcal{D}_i\left(\mathcal{D}_j\left(\sum_{x=1}^{\ell} r_x \cdot s_x\right) * \left(\sum_{z=1}^{m} b_z \cdot v_z\right)\right) * \left(\sum_{z=1}^{n} a_y \cdot u_y\right) \\
&= \mathcal{D}_j(\mathcal{D}_i\tau * \rho) * \sigma,
\end{aligned} \tag{2.3.11}$$

where we used Lemma 2.3.2 for the third equality and (2.3.2d) for the second and fifth equality. ❏

At this point, a differentiated variable is built exclusively from a variable and simple differential terms, a differentiated abstraction is built built exclusively from simple differential terms, and an application is built from a simple differential term and a complex differential term. However, we will find it essential, especially in the context of defining $\beta$-reduction for differential $\lambda$-calculus, to consider generalized versions of these two notions built completely from complex differential terms, using the linearity suggested by (2.1.1) and (2.1.2).

**Notation 2.3.4**

Let $n \in \mathbb{N}_0$ be a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ be positive integers, let $\sigma, \tau_1, \ldots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential terms, and let $x \in \mathcal{V}$ be a variable. Then, the *complex differentiated variable* $D_{i_1,\ldots,i_n}\sigma * (\tau_1, \ldots, \tau_n)$, the *complex differentiated abstraction* $D_1^n\lambda x.\sigma * (\tau_1, \ldots, t_n)$, and the *complex differentiated application* $\sigma\tau$ are the complex differential terms given by

$$D_{i_1,\ldots,i_n}\sigma * (\tau_1, \ldots, \tau_n) := D_{i_1}(\cdots(D_{i_n}\sigma * \tau_1)\cdots) * \tau_1, \tag{2.3.12a}$$

$$D_1^n\lambda x.\sigma * (\tau_1, \ldots, \tau_n) := \sum_{j=1}^{m} r_j \cdot D_1(\cdots(D_1\lambda x.s_j * \tau_1)\cdots) * \tau_1, \tag{2.3.12b}$$

$$\sigma\tau := \sum_{j=1}^{m} r_j \cdot s_j\tau, \tag{2.3.12c}$$

where $m \in \mathbb{N}_0$ is the nonnegative integer, $r_1, \cdots, r_n \in \mathcal{R}$ are the scalars, and $s_1, \cdots, s_n \in \Lambda_{d,\mathcal{V}}$

are the simple differential terms such that $\sigma = r_1 \cdot s_1 + \cdots + r_n \cdot s_n$, which we know exist and are unique by Lemma 2.2.29(xii). ♠

The fact that notion of a complex differentiated variable, complex differentiated abstraction, and complex application, as defined in Notation 2.3.4, reduce to the origin definition of a differentiated variable, differentiated abstraction, and complex application, as defined in Notation 2.2.25, is immediate. With this generalization of differentiated variable, differentiated abstraction, and abstraction, we can generalize out usual notion of substitution of a simple differential term for a variable in a complex differential term to a complex differential term for a variable in a complex differential term.

**Definition 2.3.5. Complex Substitution for Differential Terms**

Given a variable $x \in \mathcal{V}$ and complex differential terms $\tau, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, the *substitution* $\tau[\rho/x]$ of $\rho$ for $x$ in $\tau$ is defined by:

$$(D_{i_1,\ldots,i_n} x * (\tau_1, \ldots, \tau_n))[\rho/x] = D_{i_1,\ldots,i_n} \rho * (\tau_1[\rho/x], \ldots, \tau_n[\rho/x]), \tag{2.3.13a}$$
$$(D_{i_1,\ldots,i_n} y * (\tau_1, \ldots, \tau_n))[\rho/x] = D_{i_1,\ldots,i_n} y * (\tau_1[\rho/x], \ldots, \tau_n[\rho/x]), \tag{2.3.13b}$$
$$(D_1^n \lambda x.s * (\tau_1, \ldots, \tau_n))[\rho/x] = D_1^n \lambda x.s * (\tau_1[\rho/x], \ldots, \tau_n[\rho/x]), \tag{2.3.13c}$$
$$(D_1^n \lambda y.s * (\tau_1, \ldots, \tau_n))[\rho/x] = D_1^n \lambda y.\sigma[\rho/x] * (\tau_1[\rho/x], \ldots, \tau_n[\rho/x]) \quad \text{if } y \notin \mathrm{FV}(\rho), \tag{2.3.13d}$$
$$(\sigma\mu)[t/x] = \sigma[\rho/x]\mu[\rho/x], \tag{2.3.13e}$$
$$(r_1 \cdot \tau_1 + \cdots r_n \cdot t_n)[\rho/x] = r_1 \cdot \tau_1[\rho/x] + \cdots r_n \cdot \tau_n[\rho/x]. \tag{2.3.13f}$$

where $y \in \mathcal{V}$ is a variable distinct from $x$, $r_1, \ldots, r_n \in \mathcal{R}$ are scalars, $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, and $\mu, \tau_1, \ldots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms. ♣

Again, the fact that the substitution of a complex differential term for a variable in a complex differential term reduces, as defined in Definition 2.3.5, the substation of a simple differential term for a variable in a complex differential term, as defined in 2.2.24, is immediate. Now, we show that linearization (i.e., $\mathcal{D}$) commutes with substitution.

**Lemma 2.3.6**

If $i \in \mathbb{N}$ is a positive integer and $\tau, \sigma, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms, we have that

$$(\mathcal{D}_i \tau * \sigma)[\rho/x] = \mathcal{D}_i t[\rho/x] * \sigma[\rho/x]. \tag{2.3.14}$$ ♡

***Proof.*** This follows by induction on $\tau$.

1. Assume that $\tau = D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Then,

$$\begin{aligned}
\mathcal{D}_i \tau * \sigma)[\rho/x] &= (\mathcal{D}_i (D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) * \sigma)[\rho/x] \\
&= (D_{i,i_1,\ldots,i_n} x * (u, t_1, \ldots, t_n))[\rho/x] \\
&= D_{i,i_1,\ldots,i_n} v * (u[\rho/x], t_1[\rho/x], \ldots, t_n[\rho/x]) \\
&= \mathcal{D}_i (D_{i_1,\ldots,i_n} v * (t_1[\rho/x], \ldots, t_n[\rho/x])) * \sigma[\rho/x] \\
&= \mathcal{D}_i (D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n))[\rho/x] * \sigma[\rho/x] \\
&= \mathcal{D}_i \tau[\rho/x] * \sigma[\rho/x].
\end{aligned} \tag{2.3.15}$$

2. Assume that $\tau = D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n)$, where $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, $y \in \mathcal{V}$ is a variable such that $y \neq x$, and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Then,

$$\begin{aligned}(\mathcal{D}_i\tau * \sigma)[\rho/x] &= (\mathcal{D}_i(D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n)) * \sigma)[\rho/x] \\ &= (D_{i,i_1,\ldots,i_n} x * (u, t_1, \ldots, t_n))[\rho/x] \\ &= D_{i,i_1,\ldots,i_n} y * (u[\rho/x], t_1[\rho/x], \ldots, t_n[\rho/x]) \\ &= \mathcal{D}_i(D_{i_1,\ldots,i_n} y * (t_1[\rho/x], \ldots, t_n[\rho/x])) * \sigma[\rho/x] \\ &= \mathcal{D}_i(D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n))[\rho/x] * \sigma[\rho/x] \\ &= \mathcal{D}_i\tau[\rho/x] * \sigma[\rho/x].\end{aligned} \tag{2.3.16}$$

3. Assume that $\tau = D_1^n \lambda y.s * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $y \in \mathcal{V}$ is a variable such that $y$ is a fresh variable, and $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Further, suppose that $i = 1$. Then,

$$\begin{aligned}(\mathcal{D}_i\tau * \sigma)[\rho/x] &= (\mathcal{D}_1(D_1^n \lambda y.s * (t_1, \ldots, t_n)) * \sigma)[\rho/x] \\ &= (D_1^{n+1} \lambda y.s * (u, t_1, \ldots, t_n))[\rho/x] \\ &= D_1^{n+1} \lambda y.s[\rho/x] * (u[\rho/x], t_1[\rho/x], \ldots, t_n[\rho/x]) \\ &= \mathcal{D}_1(D_1^n \lambda y.s[\rho/x] * (t_1[\rho/x], \ldots, t_n[\rho/x])) * \sigma[\rho/x] \\ &= \mathcal{D}_1((D_1^n \lambda y.s * (t_1, \ldots, t_n))[\rho/x]) * \sigma[\rho/x] \\ &= \mathcal{D}_i\tau[\rho/x] * \sigma[\rho/x].\end{aligned} \tag{2.3.17}$$

Instead, suppose that $i > 1$. By the induction hypothesis, $(\mathcal{D}_{i-1}s * \sigma)[\rho/x] = \mathcal{D}_{i-1}s[\rho/x] * \sigma[\rho/x]$. Hence,

$$\begin{aligned}(\mathcal{D}_i\tau * \sigma)[\rho/x] &= (\mathcal{D}_i(D_1^n \lambda y.s * (t_1, \ldots, t_n)) * \sigma)[\rho/x] \\ &= (D_1^n \lambda y.(\mathcal{D}_{i-1}s * \sigma) * (t_1, \ldots, t_n))[\rho/x] \\ &= D_1^n \lambda y.(\mathcal{D}_{i-1}s * \sigma)[\rho/x] * (t_1[\rho/x], \ldots, t_n[\rho/x]) \\ &= D_1^n \lambda y.(\mathcal{D}_{i-1}s[\rho/x] * \sigma[\rho/x]) * (t_1[\rho/x], \ldots, t_n[\rho/x]) \\ &= \mathcal{D}_i(D_1^n \lambda y.s[\rho/x] * (t_1[\rho/x], \ldots, t_n[\rho/x])) * \sigma[\rho/x] \\ &= \mathcal{D}_i((D_1^n \lambda y.s * (t_1, \ldots, t_n))[\rho/x]) * \sigma[\rho/x] \\ &= \mathcal{D}_i\tau[\rho/x] * \sigma[\rho/x].\end{aligned} \tag{2.3.18}$$

4. Assume that $\tau = s\mu$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential terms. By the induction hypothesis, $(\mathcal{D}_{i+1}s*\sigma)[\rho/x] = \mathcal{D}_{i+1}s[\rho/x]*\sigma[\rho/x]$. Hence,

$$\begin{aligned}(\mathcal{D}_i\tau * \sigma)[\rho/x] &= (\mathcal{D}_i s\mu * \sigma)[\rho/x] \\ &= ((\mathcal{D}_{i+1}s * \sigma)\mu)[\rho/x] \\ &= (\mathcal{D}_{i+1}s * \sigma)[\rho/x]b[\rho/x] \\ &= (\mathcal{D}_{i+1}s[\rho/x] * \sigma[\rho/x])\mu[\rho/x] \\ &= \mathcal{D}_i s[\rho/x]\mu[\rho/x] * \sigma[\rho/x] \\ &= \mathcal{D}_i(s\mu)[\rho/x] * \sigma[\rho/x] \\ &= \mathcal{D}_i\tau[\rho/x] * \sigma[\rho/x].\end{aligned} \tag{2.3.19}$$

5. Assume that $\tau = \sum_{j=1}^n r_j \cdot s_j$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \ldots, r_n \in \mathcal{R}$ are scalars, and $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis,

$(\mathcal{D}_i s_j * \sigma)[\rho/x] = \mathcal{D}_i s_j[\rho/x] * \sigma[\rho/x]$ for each $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
\mathcal{D}_i\tau * \sigma)[\rho/x] &= \left(\mathcal{D}_i\left(\sum_{j=1}^{n} r_j \cdot s_j\right) * \sigma\right)[\rho/x] \\
&= \left(\sum_{j=1}^{n} r_j \cdot (\mathcal{D}_i s_j * \sigma)\right)[\rho/x] \\
&= \sum_{j=1}^{n} r_j \cdot (\mathcal{D}_i s_j * \sigma)[\rho/x] \\
&= \sum_{j=1}^{n} r_j \cdot (\mathcal{D}_i s_j[\rho/x] * \sigma[\rho/x]) \\
&= \mathcal{D}_i\left(\sum_{j=1}^{n} r_j \cdot s_j[\rho/x]\right) * \sigma[\rho/x] \\
&= \mathcal{D}_i\left(\sum_{j=1}^{n} r_j \cdot s_j\right)[\rho/x] * \sigma[\rho/x] \\
&= \mathcal{D}_i\tau[\rho/x] * \sigma[\rho/x]
\end{aligned} \tag{2.3.20}$$

This completes the induction. ❏

With this lemma, we can lift many of the properties of substitution of a simple differential term for a variable in a complex differential term to substitution of a complex differential term for a variable in a complex differential term.

**Lemma 2.3.7**

Let $x, y \in \mathcal{V}$ be variables and $\tau, \sigma, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential terms. Then, the following statements are true:

(i) If $x \notin \mathrm{FV}(\tau)$, then $\tau[\sigma/x] = \tau$.

(ii) $y \in \mathrm{FV}(\tau[\sigma/x])$ if and only if either $y \in \mathrm{FV}(\tau)$ and $x \neq y$ or $y \in \mathrm{FV}(u)$ and $x \in \mathrm{FV}(\tau)$.

(iii) If $y \neq x \notin \mathrm{FV}(\rho)$, then $\tau[\sigma/x][\rho/y] = \tau[\rho/y][\sigma[\rho/y]/x]$.

(iv) If $y \notin \mathrm{FV}(\tau)$, then $\tau[y/x][\sigma/y] = \tau[\sigma/x]$. ♡

***Proof***. This follows from Lemma 2.2.29(i), (ii), (iv), and (v), Notation 2.3.4, Definition 2.3.5, and Lemma 2.3.6. ❏

We are ready to define partial differentiation for differential terms. Given differential terms $f$ and $g$ representing analytic functions of type $\mathbb{R}^m \to \mathbb{R}$, intuitively, the partial derivative $\partial_x f * g$ of $f$ with respect to the $x$ times $g$ is a function whose type is the same as $f$ and is *equivalent* to the function given by

$$x_1, \ldots, x_n \mapsto (\partial_x f)(x_1, \ldots, x_m) \cdot g(x_1, \ldots, x_m) \tag{2.3.21}$$

after evaluation. This makes sense. But there is something strange about this definition. It seems as though the informal definition for partial differentiation is almost word for word the informal definition of linearization. That is correct. In fact, the only difference in the two informal definitions is the appearance of the phrase "equivalent to [...] after evaluation" in the informal definition for partial differentiation, a phrase that does not appear in the informal definition for linearization. This phrase is key. The difference between linearization and partial differentiation that lineation of a differential term represents preforming partial differentiation on that differential term purely syntactically by taking on or modifying a symbol, while partial differentiation of a differential term returns a differential term whose underlying analytic function is the underlying analytic function of the linearization *after* we apply simple differentiation rules like Leibnitz rule and the chain rule. In this way, linearization is syntactic and partial differentiation is computational. We see a similar dynamic between two notions in ordinary $\lambda$-calculus: the notions of application and substitution. Given a variable $x$ and two ordinary $\lambda$-terms $M$ and $N$, the application $(\lambda x.M)N$ represents the term where we input $N$ into the function $\lambda x.M$, but we do not *actually* replace every instance of $x$ in $M$ with the term $N$, as opposed to the substitution $M[N/x]$ where we are actually modifying $M$ by replacing every instance of $x$ in $M$ with $N$. In this way, application and substation, just like linearization and partial differentiation, represent the same concept from a birds eye view, but the former is symbolic and the latter is computational. Without further ado, we now define partial differentiation.

**Definition 2.3.8. Partial Derivative of Differential Pre-Terms**

Given a variable $x \in \mathcal{V}$ and complex differential term $\sigma, \tau \in \mathcal{R}\langle\Delta_{d,\mathcal{V}}\rangle$, the *partial derivative* $\partial_x \tau * \sigma$ of $\tau$ with respect to $x$ times $\sigma$ is defined by:

$$\begin{aligned}\partial_x(D_{i_1,\dots,i_n} x * (t_1,\dots,t_n)) * \sigma = D_{i_1,\dots,i_n} \sigma * (t_1,\dots,t_n) \\ + \sum_{j=1}^{n} D_{i_1,\dots,i_n} y * (t_1,\dots,\partial_x t_j * \sigma,\dots,t_n),\end{aligned} \tag{2.3.22a}$$

$$\partial_x(D_{i_1,\dots,i_n} y * (t_1,\dots,t_n)) * \sigma = \sum_{j=1}^{n} D_{i_1,\dots,i_n} y * (t_1,\dots,\partial_x t_j * \sigma,\dots,t_n), \tag{2.3.22b}$$

$$\partial_x(D_1^n \lambda x.s * (t_1,\dots,t_n)) * \sigma = \sum_{j=1}^{n} D_1^n \lambda x.s * (t_1,\dots,\partial_x t_j * \sigma,\dots,t_n), \tag{2.3.22c}$$

$$\begin{aligned}\partial_x(D_1^n \lambda y.s * (t_1,\dots,t_n)) * \sigma = D_1^n \lambda y.(\partial_x s * \sigma) * (t_1,\dots,t_n) \quad \text{if } y \notin \mathrm{FV}(\sigma) \\ + \sum_{j=1}^{n} D_1^n \lambda y.s * (t_1,\dots,\partial_x t_j * \sigma,\dots,t_n),\end{aligned} \tag{2.3.22d}$$

$$\partial_x(s\mu) * \sigma = (\partial_x s * \sigma)\mu + (\mathcal{D}_1 s * (\partial_x \mu * \sigma))\mu, \tag{2.3.22e}$$

$$\partial_x\left(\sum_{j=1}^{m} a_j \cdot s_k\right) * \sigma = \sum_{j=1}^{m} a_j \cdot \partial_x s_j * \sigma. \tag{2.3.22f}$$

where $n \in \mathbb{N}_0$ are nonnegative integers, $a_1,\dots,a_n \in \mathcal{R}$ are scalars, $i_1,\dots,i_n \in \mathbb{N}$ are positive integers, $s, s_1,\dots,s_n, t_1,\dots,t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms, and $y \in \mathcal{V}$ is a variable distinct from $x$. ♣

To simplify notation, we are going to introduce a symbol that behaves a lot like the very familiar Kronecker Delta symbol.

**Notation 2.3.9. Kronecker Delta**

Given variables $x, y \in \mathcal{V}$ and a complex differential term $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, we define the complex differential term $\delta_{x,y}\tau$ by

$$\delta_{x,y}\tau \coloneqq \begin{Bmatrix} \tau, & \text{if } x = y \\ 0, & \text{if } x \neq y \end{Bmatrix}. \tag{2.3.23}$$

♠

Now, we prove some important properties of the partial derivative.

**Lemma 2.3.10**

If $i \in \mathbb{N}$ is a positive integer and $\tau, \sigma, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms, we have that

$$\partial_x(\mathcal{D}_i\tau * \sigma) * \rho = \mathcal{D}_i(\partial_x\tau * \rho) * \sigma + \mathcal{D}_i\tau * (\partial_x\sigma * \rho). \tag{2.3.24}$$

♡

***Proof***. We prove this via induction on $\tau$.

1. Assume that $\tau = D_{i_1,\dots,i_n}y * (t_1, \dots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \dots, i_n \in \mathbb{N}$ are positive integers, $y \in \mathcal{V}$ is a variable, and $t_1, \dots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Then,

$$\begin{aligned}
\partial_x(\mathcal{D}_i\tau * \sigma) * \rho &= \partial_x(\mathcal{D}_i(D_{i_1,\dots,i_n}y * (t_1, \dots, t_n)) * \sigma) * \rho \\
&= \partial_x(D_{i,i_1,\dots,i_n}y * (u, t_1, \dots, t_n)) * \rho \\
&= \delta x, yD_{i,i_1,\dots,i_n}\rho * (u, t_1, \dots, t_n) \\
&\quad + D_{i,i_1,\dots,i_n}y * (\partial_x\sigma * \rho, t_1, \dots, t_n) \\
&\quad + \sum_{j=1}^{n} D_{i,i_1,\dots,i_n}y * (u, t_1, \dots, \partial_x uj * \rho, \dots, t_n) \\
&= \mathcal{D}_i(\delta x, yD_{i_1,\dots,i_n}\rho * (t_1, \dots, t_n)) * \sigma \\
&\quad + \mathcal{D}_i(D_{i_1,\dots,i_n}y * (t_1, \dots, t_n)) * (\partial_x\sigma * \rho) \\
&\quad + \sum_{j=1}^{n} \mathcal{D}_i(D_{i_1,\dots,i_n}y * (t_1, \dots, \partial_x uj * \rho, \dots, t_n)) * \sigma \\
&= \mathcal{D}_i\Bigg(\delta x, yD_{i_1,\dots,i_n}\rho * (t_1, \dots, t_n) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\dots,i_n}y * (t_1, \dots, \partial_x uj * \rho, \dots, t_n)\Bigg) * \sigma \\
&\quad + \mathcal{D}_i(D_{i_1,\dots,i_n}y * (t_1, \dots, t_n)) * (\partial_x\sigma * \rho) \\
&= \mathcal{D}_i(\partial_x(D_{i_1,\dots,i_n}y * (t_1, \dots, t_n)) * \rho) * \sigma \\
&\quad + \mathcal{D}_i(D_{i_1,\dots,i_n}y * (t_1, \dots, t_n)) * (\partial_x\sigma * \rho) \\
&= \mathcal{D}_i(\partial\tau * \rho) * \sigma + \mathcal{D}_i\tau * (\partial_x\sigma * \rho).
\end{aligned} \tag{2.3.25}$$

2. Assume that $\tau = D_1^n\lambda y.s * (t_1, \dots, t_n)$, where $n \in \mathbb{N}$ is a nonnegative integer, $y \in \mathcal{V}$ is a variable such that $y \neq x$ and $y \notin FV(u)$, and $t, t_1, \dots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms.

Further, suppose that $i = 1$. Then,

$$\begin{aligned}
\partial_x(\mathcal{D}_i\tau * \sigma) * \rho &= \partial_x(\mathcal{D}_1(D_1^n\lambda y.s * (t_1, \dots, t_n)) * \sigma) * \rho \\
&= \partial_x(\mathcal{D}_1^{n+1}\lambda y.s * (u, t_1, \dots, t_n)) * \rho \\
&= \mathcal{D}_1^{n+1}\lambda y.(\partial_x s * \rho) * (u, t_1, \dots, t_n) \\
&\quad + \mathcal{D}_1^{n+1}\lambda y.s * (\partial_x \sigma * \rho, t_1, \dots, t_n) \\
&\quad + \sum_{i=1}^{n} \mathcal{D}_1^{n+1}\lambda y.s * (u, t_1, \dots, \partial_x ui * \rho, t_n) \\
&= \mathcal{D}_1(D_1^n\lambda y.(\partial_x s * \rho) * (t_1, \dots, t_n)) * \sigma \\
&\quad + \mathcal{D}_1(D_1^n\lambda y.s * (t_1, \dots, t_n)) * (\partial_x \sigma * \rho) \\
&\quad + \sum_{i=1}^{n} \mathcal{D}_1(D_1^n\lambda y.s * (t_1, \dots, \partial_x ui * \rho, t_n)) * \sigma \\
&= \mathcal{D}_1(D_1^n\lambda y.(\partial_x s * \rho) * (t_1, \dots, t_n)) * \sigma \\
&\quad + \sum_{i=1}^{n} \mathcal{D}_1(D_1^n\lambda y.s * (t_1, \dots, \partial_x ui * \rho, t_n)) * \sigma \\
&\quad + \mathcal{D}_1(D_1^n\lambda y.s * (t_1, \dots, t_n)) * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_1\bigg(D_1^n\lambda y.(\partial_x s * \rho) * (t_1, \dots, t_n) \\
&\quad + \sum_{i=1}^{n} D_1^n\lambda y.s * (t_1, \dots, \partial_x ui * \rho, t_n)\bigg) * \sigma \\
&\quad + \mathcal{D}_1(D_1^n\lambda y.s * (t_1, \dots, t_n)) * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_1(\partial_x(D_1^n\lambda y.s * (t_1, \dots, t_n)) * \rho) * \sigma \\
&\quad + \mathcal{D}_1(D_1^n\lambda y.s * (t_1, \dots, t_n)) * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_i(\partial_x \tau * \rho) * \sigma + \mathcal{D}_i\tau * (\partial_x \sigma * \rho).
\end{aligned} \tag{2.3.26}$$

Instead, let us assume that $i > 1$. By the induction hypothesis, $\partial_x(\mathcal{D}_{i-1}s * \sigma) * \rho = \mathcal{D}_{i-1}(\partial_x s * \rho) * \sigma + \mathcal{D}_{i-1}s * (\partial_x \sigma * \rho)$. Hence,

$$\begin{aligned}
\partial_x(\mathcal{D}_i\tau * \sigma) * \rho &= \partial_x(\mathcal{D}_i(D_1^n\lambda y.s * (t_1, \dots, t_n)) * \sigma) * \rho \\
&= \partial_x(D_1^n\lambda y.(\mathcal{D}_{i-1}s * \sigma) * (t_1, \dots, t_n)) * \rho \\
&= D_1^n\lambda y.(\partial_x(\mathcal{D}_{i-1}s * \sigma) * \rho) * (t_1, \dots, t_n) \\
&\quad + \sum_{i=1}^{n} D_1^n\lambda y.(\mathcal{D}_{i-1}s * \sigma) * (t_1, \dots, \partial_x ui * \rho, \dots, t_n) \\
&= D_1^n\lambda y.(\mathcal{D}_{i-1}(\partial_x s * \rho) * \sigma + \mathcal{D}_{i-1}s * (\partial_x \sigma * \rho)) * (t_1, \dots, t_n) \\
&\quad + \sum_{i=1}^{n} D_1^n\lambda y.(\mathcal{D}_{i-1}s * \sigma) * (t_1, \dots, \partial_x ui * \rho, \dots, t_n) \\
&= D_1^n\lambda y.(\mathcal{D}_{i-1}(\partial_x s * \rho) * \sigma) * (t_1, \dots, t_n) \\
&\quad + D_1^n\lambda y.(\mathcal{D}_{i-1}s * (\partial_x \sigma * \rho)) * (t_1, \dots, t_n) \\
&\quad + \sum_{i=1}^{n} D_1^n\lambda y.(\mathcal{D}_{i-1}s * \sigma) * (t_1, \dots, \partial_x ui * \rho, \dots, t_n)
\end{aligned} \tag{2.3.27}$$

$$
\begin{aligned}
&= \mathcal{D}_i(D_1^n \lambda y.(\partial_x s * \rho) * (t_1, \dots, t_n)) * \sigma \\
&\quad + \mathcal{D}_i(D_1^n \lambda y.s * (t_1, \dots, t_n)) * (\partial_x \sigma * \rho) \\
&\quad + \sum_{i=1}^{n} \mathcal{D}_i(D_1^n \lambda y.s * (t_1, \dots, \partial_x ui * \rho, \dots, t_n)) * \sigma \\
&= \mathcal{D}_i(D_1^n \lambda y.(\partial_x s * \rho) * (t_1, \dots, t_n)) * \sigma \\
&\quad + \sum_{i=1}^{n} \mathcal{D}_i(D_1^n \lambda y.s * (t_1, \dots, \partial_x ui * \rho, \dots, t_n)) * \sigma \\
&\quad + \mathcal{D}_i(D_1^n \lambda y.s * (t_1, \dots, t_n)) * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_i\Bigg( D_1^n \lambda y.(\partial_x s * \rho) * (t_1, \dots, t_n) \\
&\quad + \sum_{i=1}^{n} D_1^n \lambda y.s * (t_1, \dots, \partial_x ui * \rho, \dots, t_n) \Bigg) * \sigma \\
&\quad + \mathcal{D}_i(D_1^n \lambda y.s * (t_1, \dots, t_n)) * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_i(\partial_x(D_1^n \lambda y.s * (t_1, \dots, t_n)) * \rho) * \sigma \\
&\quad + \mathcal{D}_i(D_1^n \lambda y.s * (t_1, \dots, t_n)) * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_i(\partial_x \tau * \rho) * \sigma + \mathcal{D}_i \tau * (\partial_x \sigma * \rho).
\end{aligned}
$$

3. Assume that $\tau = s\mu$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ are complex differential terms. By the induction hypothesis, $\partial_x(\mathcal{D}_{i+1} s * \sigma) * \rho = \mathcal{D}_{i+1}(\partial_x s * \rho) * \sigma + \mathcal{D}_{i+1} s * (\partial_x \sigma * \rho)$. Hence,

$$
\begin{aligned}
\partial_x(\mathcal{D}_i \tau * \sigma) * \rho &= \partial_x(\mathcal{D}_i(s\mu) * \sigma) * \rho \\
&= \partial_x((\mathcal{D}_{i+1} s * \sigma)\mu) * \rho \\
&= (\partial_x(\mathcal{D}_{i+1} s * \sigma) * \rho)\mu + (\mathcal{D}_1(\mathcal{D}_{i+1} s * \sigma) * (\partial_x \mu * \rho))\mu \\
&= (\mathcal{D}_{i+1}(\partial_x s * \rho) * \sigma + \mathcal{D}_{i+1} s * (\partial_x \sigma * \rho))\mu \\
&\quad + (\mathcal{D}_1(\mathcal{D}_{i+1} s * \sigma) * (\partial_x \mu * \rho))\mu \\
&= (\mathcal{D}_{i+1}(\partial_x s * \rho) * \sigma)\mu + (\mathcal{D}_{i+1} s * (\partial_x \sigma * \rho))\mu \\
&\quad + (\mathcal{D}_1(\mathcal{D}_{i+1} s * \sigma) * (\partial_x \mu * \rho))\mu \\
&= (\mathcal{D}_{i+1}(\partial_x s * \rho) * \sigma)\mu + (\mathcal{D}_{i+1} s * (\partial_x \sigma * \rho))\mu \\
&\quad + (\mathcal{D}_{i+1}(\mathcal{D}_1 s * (\partial_x \mu * \rho)) * \sigma)\mu \\
&= (\mathcal{D}_{i+1}(\partial_x s * \rho) * \sigma)\mu + (\mathcal{D}_{i+1}(\mathcal{D}_1 s * (\partial_x \mu * \rho)) * \sigma)\mu \\
&\quad + (\mathcal{D}_{i+1} s * (\partial_x \sigma * \rho))\mu \\
&= (\mathcal{D}_{i+1}(\partial_x s * \rho + \mathcal{D}_1 s * (\partial_x \mu * \rho)) * \sigma)\mu + (\mathcal{D}_{i+1} s * (\partial_x \sigma * \rho))\mu \\
&= \mathcal{D}_i(\partial_x s * \rho + \mathcal{D}_1 s * (\partial_x \mu * \rho))\mu) * \sigma + \mathcal{D}_i(a)\mu * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_i((\partial_x s * \rho)\mu + (\mathcal{D}_1 s * (\partial_x \mu * \rho))\mu) * \sigma + \mathcal{D}_i(a)\mu * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_i(\partial_x(a)\mu * \rho) * \sigma + \mathcal{D}_i(a)\mu * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_i(\partial_x \tau * \rho) * \sigma + \mathcal{D}_i \tau * (\partial_x \sigma * \rho)
\end{aligned}
\tag{2.3.28}
$$

4. Assume that $\tau = \sum_{j=1}^{n} r_j \cdot s_j$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \dots, r_n \in \mathcal{R}$ are scalars, and $s_1, \dots, s_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis,

$\partial_x(\mathcal{D}_i s_j * \sigma) * \rho = \mathcal{D}_i(\partial_x s_j * \rho) * \sigma + \mathcal{D}_i s_j * (\partial_x \sigma * \rho)$ for each $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
\partial_x(\mathcal{D}_i \tau * \sigma) * \rho &= \partial_x \left( \mathcal{D}_i \left( \sum_{j=1}^n r_j \cdot s_j \right) * \sigma \right) * \rho \\
&= \partial_x \left( \sum_{j=1}^n r_j \cdot (\mathcal{D}_i s_j * \sigma) \right) * \rho \\
&= \sum_{j=1}^n r_j \cdot (\partial_x(\mathcal{D}_i s_j * \sigma) * \rho) \\
&= \sum_{j=1}^n r_j \cdot (\mathcal{D}_i(\partial_x s_j * \rho) * \sigma + \mathcal{D}_i s_j * (\partial_x \sigma * \rho)) \\
&= \sum_{j=1}^n r_j \cdot (\mathcal{D}_i(\partial_x s_j * \rho) * \sigma) + \sum_{j=1}^n r_j \cdot (\mathcal{D}_i s_j * (\partial_x \sigma * \rho)) \\
&= \mathcal{D}_i \left( \partial_x \left( \sum_{j=1}^n r_j \cdot s_j \right) * \rho \right) * \sigma) + \mathcal{D}_i \left( \sum_{j=1}^n r_j \cdot s_j \right) * (\partial_x \sigma * \rho) \\
&= \mathcal{D}_i(\partial_x \tau * \rho) * \sigma + \mathcal{D}_i \tau * (\partial_x \sigma * \rho)
\end{aligned} \tag{2.3.29}$$

This completes the induction. ❏

**Lemma 2.3.11**

Let $\tau, \sigma \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ be complex differential terms. If $x \notin \mathrm{FV}(\tau)$, then $\partial_x \tau * \sigma = 0$. ♡

***Proof.*** We prove this via induction on $\tau$.

1. Assume that $\tau = D_{i_1,\dots,i_n} y * (t_1, \dots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \dots, i_n \in \mathbb{N}$ are positive integers, $y \in \mathcal{V}$ is a variable such that $x \neq y$, and $t_1, \dots, t_n$ are simple differential terms. By the induction hypothesis, $\partial_x t_j * \sigma = 0$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
\partial_x t * \sigma &= \partial_x(D_{i_1,\dots,i_n} y * (t_1, \dots, t_n)) * \sigma \\
&= \delta_{x,y} D_{i_1,\dots,i_n} \sigma * (t_1, \dots, t_n) + \sum_{i=1}^n D_{i_1,\dots,i_n} y * (t_1, \dots, \partial_x t_j * \sigma, \dots, t_n) \\
&= 0 + \sum_{i=1}^n D_{i_1,\dots,i_n} y * (t_1, \dots, 0, \dots, t_n) = 0 + \sum_{i=1}^n 0 = 0.
\end{aligned} \tag{2.3.30}$$

2. Assume that $\tau = D_1^n(\lambda y.s) * (t_1, \dots, t_n)$, where $n \in N$, $y$ is a variable, and $s, t_1, \dots, t_n$ are simple differential terms. By the induction hypothesis, $\partial_x s * \sigma = 0$ and $\partial_x t_j * \sigma = 0$ for all $1 \leqslant j \leqslant n$. Hence

$$\begin{aligned}
\partial_x t * \sigma &= \partial_x(D_1^n(\lambda y.s) * (t_1, \dots, t_n)) * \sigma \\
&= D_1^n(\lambda y.(\partial_x s * \sigma)) * (t_1, \dots, t_n) + \sum_{i=1}^n D_1^n(\lambda y.s) * (t_1, \dots, \partial_x t_j * \sigma, \dots, t_n) \\
&= D_1^n(\lambda y.0) * (t_1, \dots, t_n) + \sum_{i=1}^n D_1^n(\lambda y.s) * (t_1, \dots, 0, \dots, t_n) = 0 + \sum_{i=1}^n 0 = 0.
\end{aligned} \tag{2.3.31}$$

3. Assume that $\tau = s\mu$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential term. By the induction hypothesis, $\partial_x s * \sigma = 0$ and $\partial_x \mu * \sigma = 0$. Hence,

$$\begin{aligned} \partial_x t * \sigma &= \partial_x s\mu * \sigma = (\partial_x s * \sigma)\mu + (\mathcal{D}_1 s * (\partial_x \mu * \sigma))\mu \\ &= 0\mu + (\mathcal{D}_1 s * 0)\mu = 0 + 0 = 0. \end{aligned} \tag{2.3.32}$$

4. Assume that $\tau = \sum_{j=1}^{n} r_j \cdot s_j$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \ldots, r_n \in \mathcal{R}$ are scalars, and $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $\partial_x s_j = 0$ for each $1 \leqslant j \leqslant n$. Hence,

$$\partial_x \left(\sum_{j=1}^{n} r_j \cdot s_j\right) * \sigma = \sum_{j=1}^{n} r_j \cdot (\partial_x s_j * \sigma) = \sum_{j=1}^{n} r_j \cdot 0 = 0. \tag{2.3.33}$$

This completes the induction. ❏

**Lemma 2.3.12**

Let $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be a complex differential term, let $x \in \mathcal{V}$ be a variable, let $m \in \mathbb{N}_0$ be a nonnegative integer, let $r_1, \ldots, r_m \in \mathcal{R}$ be scalars, and let $s_1, \ldots, s_m \in \Lambda_{d,\mathcal{V}}$ be simple differential terms. Then,

$$\partial_x \tau * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) = \sum_{k=1}^{m} r_k \cdot (\partial_x \tau * s_k). \tag{2.3.34}$$

♡

***Proof.*** We prove this by induction on $\tau$.

1. Assume that $\tau = D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$, $y \in \mathcal{V}$ is a variable, and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $\partial_x t_j * (\sum_{k=1}^{m} r_k \cdot s_k) = \sum_{k=1}^{m} r_k \cdot \partial_x t_j * s_k$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned} \partial_x \tau * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) &= \partial_x (D_{i_1,\ldots,i_n} y * (t_1, \ldots, t_n)) * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) \\ &= \delta_{x,y} D_{i_1,\ldots,i_n} \left(\sum_{k=1}^{m} r_k \cdot s_k\right) * (t_1, \ldots, t_n) \\ &\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} y * \left(t_1, \ldots, \partial_x t_j * \left(\sum_{k=1}^{m} r_k \cdot s_k\right), \ldots, t_n\right) \\ &= \delta_{x,y} D_{i_1,\ldots,i_n} \left(\sum_{k=1}^{m} r_k \cdot s_k\right) * (t_1, \ldots, t_n) \\ &\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} y * \left(t_1, \ldots, \sum_{k=1}^{m} r_k \cdot \partial_x t_j * s_k, \ldots, t_n\right) \\ &= \sum_{k=1}^{m} r_k \cdot \Bigg(\delta_{x,y} D_{i_1,\ldots,i_n} s_k * (t_1, \ldots, t_n) \\ &\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} y * (t_1, \ldots, \partial_x t_j * s_k, \ldots, t_n)\Bigg) \end{aligned} \tag{2.3.35}$$

$$= \sum_{k=1}^{m} r_k \cdot \partial_x(D_{i_1,\ldots,i_n} y * (t_1,\ldots,t_n)) * s_k$$
$$= \sum_{k=1}^{m} r_k \cdot \partial_x \tau * s_k.$$

2. Assume that $\tau = D_1^n \lambda y.s*(t_1,\ldots,t_n)$, where $n \in \mathbb{N}_0$, $y \in \mathcal{V}$ is a variable, and $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $\partial_x s * (\sum_{k=1}^m r_k \cdot s_k) = \sum_{k=1}^m r_k \cdot \partial_x s * s_k$ and $\partial_x t_j * (\sum_{k=1}^m r_k \cdot s_k) = \sum_{k=1}^m r_k \cdot \partial_x t_j * s_k$ for all $1 \leqslant j \leqslant n$. Hence,
$$\begin{aligned}
\partial_x \tau * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) &= \partial_x(D_1^n \lambda y.s * (t_1,\ldots,t_n)) * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) \\
&= D_1^n \lambda y. \left(\partial_x s * \left(\sum_{k=1}^{m} r_k \cdot s_k\right)\right) * (t_1,\ldots,t_n) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda y.s * \left(t_1,\ldots,\partial_x t_j * \left(\sum_{k=1}^{m} r_k \cdot s_k\right),\ldots,t_n\right) \\
&= D_1^n \lambda y. \left(\sum_{k=1}^{m} r_k \cdot \partial_x s * s_k\right) * (t_1,\ldots,t_n) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda y.s * \left(t_1,\ldots,\sum_{k=1}^{m} r_k \cdot \partial_x t_j * s_k,\ldots,t_n\right) \qquad (2.3.36) \\
&= \sum_{k=1}^{m} r_k \cdot (D_1^n \lambda y.(\partial_x s * s_k) * (t_1,\ldots,t_n) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda y.s * (t_1,\ldots,\partial_x t_j * s_k,\ldots,t_n)) \\
&= \sum_{k=1}^{m} r_k \cdot \partial_x(D_1^n \lambda y.s * (t_1,\ldots,t_n)) * s_k \\
&= \sum_{k=1}^{m} r_k \cdot \partial_x \tau * s_k.
\end{aligned}$$

3. Assume that $\tau = s\mu$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential term. By the induction hypothesis, $\partial_x s * (\sum_{k=1}^m r_k \cdot s_k) = \sum_{k=1}^m r_k \cdot \partial_x s * s_k$ and $\partial_x \mu * (\sum_{k=1}^m r_k \cdot s_k) = \sum_{k=1}^m r_k \cdot \partial_x \mu * s_k$. Hence,
$$\begin{aligned}
\partial_x(s\mu) * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) &= \left(\partial_x s * \left(\sum_{k=1}^{m} r_k \cdot s_k\right)\right)\mu \\
&\quad + \left(\mathcal{D}_1 s * \left(\partial_x \mu * \left(\sum_{k=1}^{m} r_k \cdot s_k\right)\right)\right)\mu \qquad (2.3.37) \\
&= \left(\sum_{k=1}^{m} r_k \cdot \partial_x s * s_k\right)\mu \\
&\quad + \left(\mathcal{D}_1 s * \left(\sum_{k=1}^{m} r_k \cdot \partial_x \mu * s_k\right)\right)\mu
\end{aligned}$$

$$= \sum_{k=1}^{m} r_k \cdot ((\partial_x s * s_k)\mu + (\mathcal{D}_1 s * (\partial_x \mu * s_k))\mu)$$
$$= \sum_{k=1}^{m} r_k \cdot \partial_x(s\mu) * s_k$$
$$= \sum_{k=1}^{m} r_k \cdot \partial_x \tau * s_k.$$

4. Assume that $\tau = \sum_{j=1}^{n} a_j \cdot t_j$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $a_1, \ldots, a_n \in \mathcal{R}$ are scalars, and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $\partial_x t_k * (\sum_{k=1}^{m} r_k \cdot s_k) = \sum_{k=1}^{m} r_k \cdot \partial_x t_j * s_k$ for each $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
\partial_x \tau * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) &= \partial_x \left(\sum_{j=1}^{n} a_j \cdot t_j\right) * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) \\
&= \sum_{j=1}^{n} a_j \cdot \partial_x t_j * \left(\sum_{k=1}^{m} r_k \cdot s_k\right) \\
&= \sum_{j=1}^{n} a_j \cdot \sum_{k=1}^{m} r_k \cdot \partial_x t_j * s_k \qquad (2.3.38) \\
&= \sum_{k=1}^{m} r_k \cdot \partial_x \left(\sum_{j=1}^{n} a_j \cdot t_j\right) * s_k \\
&= \sum_{k=1}^{m} r_k \cdot \partial_x \tau * s_k.
\end{aligned}$$

This completes the induction. ❑

**Lemma 2.3.13**

Let $x, y \in \mathcal{V}$ be variables and $\tau, \sigma, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential terms such that $y \notin \mathrm{FV}(\sigma)$. Then,

$$\partial_x(\partial_y \tau * \rho) * \sigma = \partial_y(\partial_x \tau * \sigma) * \rho + \partial_y \tau * (\partial_x \rho * \sigma). \qquad (2.3.39)$$

♡

***Proof***. This is Lemma 4 in [8]. ❑

**Corollary 2.3.14. Synthetic Schwarz lemma**

Let $x, y \in \mathcal{V}$ be variables and $\tau, \sigma, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential terms such that $y \notin \mathrm{FV}(\sigma)$ and $x \notin \mathrm{FV}(\rho)$. Then,

$$\partial_x(\partial_y \tau * \rho) * \sigma = \partial_y(\partial_x \tau * \sigma) * \rho. \qquad (2.3.40)$$

♡

***Proof***. Since $x \notin \mathrm{FV}(\rho)$, by Lemma 2.3.11, $\partial_x \rho * \sigma = 0$. By Lemma 2.3.13, $\partial_x(\partial_y \tau * \rho) * \sigma = \partial_y(\partial_x \tau * \sigma) * \rho + \partial_y \tau * (\partial_x \rho * \sigma) = \partial_y(\partial_x \tau * \sigma) * \rho + \partial_y \tau * 0 = \partial_y(\partial_x \tau * \sigma) * \rho + 0 = \partial_y(\partial_x \tau * \sigma) * \rho$. ❑

**Lemma 2.3.15**

Let $n \in \mathbb{N}_0$ be a nonnegative number, let $i_1, \dots, i_n \in \mathbb{N}$ be positive integers, and let $\tau_1, \sigma_1, \dots, \sigma_n \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$. Then,

$$\begin{aligned} \partial_x(D_{i_1,\dots,i_n}\tau * (\sigma_1, \dots, \sigma_n)) * \sigma = D_{i_1,\dots,i_n}&(\partial_x\tau * \sigma) * (\sigma_1, \dots, \sigma_n) \\ &+ \sum_{j=1}^{n} D_{i_1,\dots,i_n}\tau * (\sigma_1, \dots, \partial_x\sigma_j * \sigma, \dots, \sigma_n). \end{aligned} \tag{2.3.41}$$

♡

***Proof***. This follows immediately by induction on $n$ using Notation 2.3.4 and Lemma 2.3.10. ❏

**Lemma 2.3.16**

Let $x, y \in \mathcal{V}$ be variables, $\tau, \sigma, \rho \in \mathcal{R}\langle \Delta_{d,\mathcal{V}} \rangle$ are complex differential terms such that $x \neq y$ and $y$ is not free in $\sigma$ and $\rho$. Then,

$$\partial_x\tau[\rho/y] * \sigma = (\partial_x\tau * \sigma)[\rho/y] + (\partial_y\tau * (\partial_x\rho * \sigma))[\rho/y]. \tag{2.3.42}$$

♡

***Proof***. We prove this via induction on $\tau$.

1. Assume that $\tau = D_{i_1,\dots,i_n}z * (t_1, \dots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \dots, i_n \in \mathbb{N}$, $z \in \mathcal{V}$ is a variable, and $t_1, \dots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. As we can see,

$$\begin{aligned} (\partial_x\tau * \sigma)[\rho/y] &= (\partial_x(D_{i_1,\dots,i_n}z * (t_1, \dots, t_n)) * \sigma)[\rho/y] \\ &= (\delta_{x,z}D_{i_1,\dots,i_n}\sigma * (t_1, \dots, t_n) \\ &\qquad + \sum_{j=1}^{n} D_{i_1,\dots,i_n}z * (t_1, \dots, \partial_x t_j * \sigma, \dots, t_n))[\rho/y] \\ &= \delta_{x,z}D_{i_1,\dots,i_n}\sigma[\rho/y] * (t_1[\rho/y], \dots, t_n[\rho/y]) \\ &\qquad + \sum_{j=1}^{n} D_{i_1,\dots,i_n}z[\rho/y] * (t_1[\rho/y], \dots, (\partial_x t_j * \sigma)[\rho/y], \dots, t_n[\rho/y]) \end{aligned} \tag{2.3.43}$$

and

$$\begin{aligned} (\partial_y\tau * (\partial_x\rho * \sigma))[\rho/y] &= (\partial_y(D_{i_1,\dots,i_n}z * (t_1, \dots, t_n)) * (\partial_x\rho * \sigma))[\rho/y] \\ &= (\delta_{y,z}D_{i_1,\dots,i_n}(\partial_x\rho * \sigma) * (t_1, \dots, t_n) \\ &\qquad + \sum_{j=1}^{n} D_{i_1,\dots,i_n}z * (t_1, \dots, \partial_x t_j * (\partial_x\rho * \sigma), \dots, t_n)[\rho/y] \\ &= \delta_{y,z}D_{i_1,\dots,i_n}(\partial_x\rho * \sigma)[\rho/y] * (t_1[\rho/y], \dots, t_n[\rho/y]) \\ &\qquad + \sum_{j=1}^{n} D_{i_1,\dots,i_n}z[\rho/y] * (t_1[\rho/y], \\ &\qquad\qquad \dots, (\partial_x t_j * (\partial_x\rho * \sigma))[\rho/y], \dots, t_n[\rho/y]). \end{aligned} \tag{2.3.44}$$

By the induction hypothesis, $\partial_x t_j[\rho/y] * \sigma = (\partial_x t_j * \sigma)[\rho/y] + (\partial_y t_j * (\partial_x\rho * \sigma))[\rho/y]$ for all

$1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
(\partial_x\tau * \sigma)[\rho/y] \\
+(\partial_y\tau * (\partial_x\rho * \sigma))[\rho/y] &= \delta_{x,z} D_{i_1,\ldots,i_n}\sigma[\rho/y] * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y], \\
&\qquad \ldots,(\partial_x t_j * \sigma)[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \delta_{y,z} D_{i_1,\ldots,i_n}(\partial_x\rho * \sigma)[\rho/y] * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y], \\
&\qquad \ldots,(\partial_x t_j * (\partial_x\rho * \sigma))[\rho/y],\ldots,t_n[\rho/y]) \\
&= \delta_{x,z} D_{i_1,\ldots,i_n}\sigma[\rho/y] * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \delta_{y,z} D_{i_1,\ldots,i_n}(\partial_x\rho * \sigma)[\rho/y] * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y], \\
&\qquad \ldots,(\partial_x t_j * \sigma)[\rho/y] + (\partial_x t_j * (\partial_x\rho * \sigma))[\rho/y],\ldots,t_n[\rho/y]) \\
&= \delta_{x,z} D_{i_1,\ldots,i_n}\sigma[\rho/y] * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \delta_{y,z} D_{i_1,\ldots,i_n}(\partial_x\rho * \sigma)[\rho/y] * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y],\ldots,\partial_x t_j[\rho/y] * \sigma,\ldots,t_n[\rho/y]) \\
&= \delta_{x,z} D_{i_1,\ldots,i_n}\sigma * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \delta_{y,z} D_{i_1,\ldots,i_n}(\partial_x\rho * \sigma) * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y],\ldots,\partial_x t_j[\rho/y] * \sigma,\ldots,t_n[\rho/y]),
\end{aligned} \tag{2.3.45}$$

where we used the fact that $y$ is not free in $\sigma$ or $\rho$ for the last equality. If $z \neq y$, then

$$\begin{aligned}
\partial_x\tau[\rho/y] * \sigma &= \partial_x(D_{i_1,\ldots,i_n} z * (t_1,\ldots,t_n))[\rho/y] * \sigma \\
&= \partial_x(D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y],\ldots,t_n[\rho/y])) * \sigma \\
&= \partial_x(D_{i_1,\ldots,i_n} z * (t_1[\rho/y],\ldots,t_n[\rho/y])) * \sigma \\
&= \delta_{x,z} D_{i_1,\ldots,i_n}\sigma * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\qquad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y],\ldots,\partial_x t_j[\rho/y] * \sigma,\ldots,t_n[\rho/y]) \\
&= \delta_{x,z} D_{i_1,\ldots,i_n}\sigma * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\qquad + \delta_{y,z} D_{i_1,\ldots,i_n}(\partial_x\rho * \sigma) * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\qquad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y],\ldots,\partial_x t_j[\rho/y] * \sigma,\ldots,t_n[\rho/y]) \\
&= \delta_{x,z} D_{i_1,\ldots,i_n}\sigma * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\qquad + \delta_{y,z} D_{i_1,\ldots,i_n}(\partial_x\rho * \sigma) * (t_1[\rho/y],\ldots,t_n[\rho/y]) \\
&\qquad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y],\ldots,\partial_x t_j[\rho/y] * \sigma,\ldots,t_n[\rho/y]) \\
&= (\partial_x\tau * \sigma)[\rho/y] + (\partial_y\tau * (\partial_x\rho * \sigma))[\rho/y],
\end{aligned} \tag{2.3.46}$$

and if $z = y$, then

$$\begin{aligned}
\partial_x\tau[\rho/y] * \sigma &= \partial_x(D_{i_1,\dots,i_n} z * (t_1,\dots,t_n))[\rho/y] * \sigma \\
&= \partial_x(D_{i_1,\dots,i_n} z[\rho/y] * (t_1[\rho/y],\dots,t_n[\rho/y])) * \sigma \\
&= \partial_x(D_{i_1,\dots,i_n} y[\rho/y] * (t_1[\rho/y],\dots,t_n[\rho/y])) * \sigma \\
&= \partial_x(D_{i_1,\dots,i_n} \rho * (t_1[\rho/y],\dots,t_n[\rho/y])) * \sigma \\
&= D_{i_1,\dots,i_n}(\partial_x\rho * \sigma) * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\dots,i_n}\rho * (t_1[\rho/y],\dots,\partial_x t_j[\rho/y] * \sigma,\dots,t_n[\rho/y]) \\
&= \delta_{x,z} D_{i_1,\dots,i_n}\sigma * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \delta_{y,z} D_{i_1,\dots,i_n}(\partial_x\rho * \sigma) * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\dots,i_n} y[\rho/y] * (t_1[\rho/y],\dots,\partial_x t_j[\rho/y] * \sigma,\dots,t_n[\rho/y]) \\
&= \delta_{x,z} D_{i_1,\dots,i_n}\sigma * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \delta_{y,z} D_{i_1,\dots,i_n}(\partial_x\rho * \sigma) * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\dots,i_n} z[\rho/y] * (t_1[\rho/y],\dots,\partial_x t_j[\rho/y] * \sigma,\dots,t_n[\rho/y]) \\
&= (\partial_x\tau * \sigma)[\rho/y] + (\partial_y\tau * (\partial_x\rho * \sigma))[\rho/y],
\end{aligned} \tag{2.3.47}$$

where we used Lemma 2.3.15 for the fifth equality.

2. Assume that $\tau = D_1^n \lambda z.s * (t_1,\dots,t_n)$, where $n \in \mathbb{N}_0$, $z$ is a fresh variable, and $s, t_1,\dots,t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $\partial_x s[\rho/y] * \sigma = (\partial_x s * \sigma)[\rho/y] + (\partial_y s * (\partial_x\rho * \sigma))[\rho/y]$ and $\partial_x t_j[\rho/y] * \sigma = (\partial_x t_j * \sigma)[\rho/y] + (\partial_y t_j * (\partial_x\rho * \sigma))[\rho/y]$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
\partial_x\tau[\rho/y] * \sigma &= \partial_x(D_1^n \lambda z.s * (t_1,\dots,t_n))[\rho/y] * \sigma \\
&= \partial_x(D_1^n \lambda z.s[\rho/y] * (t_1[\rho/y],\dots,t_n[\rho/y])) * \sigma \\
&= D_1^n \lambda z.(\partial_x s[\rho/y] * \sigma) * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s[\rho/y] * (t_1[\rho/y],\dots,\partial_x t_j[\rho/y] * \sigma,\dots,t_n[\rho/y]) \\
&= D_1^n \lambda z.((\partial_x s * \sigma)[\rho/y] + (\partial_y s * (\partial_x\rho * \sigma))[\rho/y]) * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s[\rho/y] * (t_1[\rho/y],\dots,(\partial_x t_j * \sigma)[\rho/y] \\
&\quad + (\partial_y t_j * (\partial_x\rho * \sigma))[\rho/y],\dots,t_n[\rho/y]) \\
&= D_1^n \lambda z.(\partial_x s * \sigma)[\rho/y] * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + D_1^n \lambda z.(\partial_y s * (\partial_x\rho * \sigma))[\rho/y] * (t_1[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s[\rho/y] * (t_1[\rho/y],\dots,(\partial_x t_j * \sigma)[\rho/y],\dots,t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s[\rho/y] * (t_1[\rho/y],\dots,(\partial_y t_j * (\partial_x\rho * \sigma))[\rho/y],\dots,t_n[\rho/y])
\end{aligned} \tag{2.3.48}$$

$$
\begin{aligned}
&= D_1^n \lambda z.(\partial_x s * \sigma)[\rho/y] * (t_1[\rho/y], \ldots, t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s[\rho/y] * (t_1[\rho/y], \ldots, (\partial_x t_j * \sigma)[\rho/y], \ldots, t_n[\rho/y]) \\
&\quad + D_1^n \lambda z.(\partial_y s * (\partial_x \rho * \sigma))[\rho/y] * (t_1[\rho/y], \ldots, t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s[\rho/y] * (t_1[\rho/y], \ldots, (\partial_y t_j * (\partial_x \rho * \sigma))[\rho/y], \ldots, t_n[\rho/y]) \\
&= (D_1^n \lambda z.(\partial_x s * \sigma) * (t_1, \ldots, t_n))[\rho/y] \\
&\quad + \sum_{j=1}^{n} (D_1^n \lambda z.s * (t_1, \ldots, \partial_x t_j * \sigma, \ldots, t_n))[\rho/y] \\
&\quad + (D_1^n \lambda z.(\partial_y s * (\partial_x \rho * \sigma)) * (t_1, \ldots, t_n))[\rho/y] \\
&\quad + \sum_{j=1}^{n} (D_1^n \lambda z.s * (t_1, \ldots, (\partial_y t_j * (\partial_x \rho * \sigma)), \ldots, t_n))[\rho/y] \\
&= (D_1^n \lambda z.(\partial_x s * \sigma) * (t_1, \ldots, t_n) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s * (t_1, \ldots, \partial_x t_j * \sigma, \ldots, t_n))[\rho/y] \\
&\quad + (D_1^n \lambda z.(\partial_y s * (\partial_x \rho * \sigma)) * (t_1, \ldots, t_n) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s * (t_1, \ldots, (\partial_y t_j * (\partial_x \rho * \sigma)), \ldots, t_n))[\rho/y] \\
&= (\partial_x (D_1^n \lambda z.s * (t_1, \ldots, t_n)) * \sigma)[\rho/y] \\
&\quad + (\partial_y (D_1^n \lambda z.s * (t_1, \ldots, t_n)) * (\partial_x \rho * \sigma))[\rho/y] \\
&= (\partial_x \tau * \sigma)[\rho/y] + (\partial_y \tau * (\partial_x \rho * \sigma))[\rho/y].
\end{aligned}
$$

3. Assume that $\tau = s\mu$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential term. By the induction hypothesis, $\partial_x s[\rho/y] * \sigma = (\partial_x s * \sigma)[\rho/y] + (\partial_y s * (\partial_x \rho * \sigma))[\rho/y]$ and $\partial_x \mu[\rho/y] * \sigma = (\partial_x \mu * \sigma)[\rho/y] + (\partial_y \mu * (\partial_x \rho * \sigma))[\rho/y]$. Hence,

$$
\begin{aligned}
\partial_x \tau[\rho/y] * \sigma &= \partial_x (s\mu)[\rho/y] * \sigma \\
&= \partial_x (s[\rho/y])\mu[\rho/y] * \sigma \\
&= (\partial_x s[\rho/y] * \sigma)\mu[\rho/y] + (\mathcal{D}_1 s[\rho/y] * (\partial_x \mu[\rho/y] * \sigma))\mu[\rho/y] \\
&= ((\partial_x s * \sigma)[\rho/y] + (\partial_y s * (\partial_x \rho * \sigma))[\rho/y])\mu[\rho/y] \\
&\quad + (\mathcal{D}_1 s[\rho/y] * ((\partial_x \mu * \sigma)[\rho/y] + (\partial_y \mu * (\partial_x \rho * \sigma))[\rho/y]))\mu[\rho/y] \\
&= ((\partial_x s * \sigma)[\rho/y])\mu[\rho/y] + ((\partial_y s * (\partial_x \rho * \sigma))[\rho/y])\mu[\rho/y] \\
&\quad + (\mathcal{D}_1 s[\rho/y] * ((\partial_x \mu * \sigma)[\rho/y]))\mu[\rho/y] \\
&\quad + (\mathcal{D}_1 s[\rho/y] * ((\partial_y \mu * (\partial_x \rho * \sigma))[\rho/y]))\mu[\rho/y] \qquad (2.3.49) \\
&= ((\partial_x s * \sigma)\mu)[\rho/y] + ((\partial_y s * (\partial_x \rho * \sigma))\mu)[\rho/y] \\
&\quad + ((\mathcal{D}_1 s * (\partial_x \mu * \sigma))\mu)[\rho/y] + ((\mathcal{D}_1 s * (\partial_y \mu * (\partial_x \rho * \sigma)))\mu)[\rho/y] \\
&= ((\partial_x s * \sigma)\mu)[\rho/y] + ((\mathcal{D}_1 s * (\partial_x \mu * \sigma))\mu)[\rho/y] \\
&\quad + ((\partial_y s * (\partial_x \rho * \sigma))\mu)[\rho/y] + ((\mathcal{D}_1 s * (\partial_y \mu * (\partial_x \rho * \sigma)))\mu)[\rho/y] \\
&= ((\partial_x s * \sigma)\mu + (\mathcal{D}_1 s * (\partial_x \mu * \sigma))\mu)[\rho/y]
\end{aligned}
$$

$$
\begin{aligned}
&+ ((\partial_y s * (\partial_x \rho * \sigma))\mu + (\mathcal{D}_1 s * (\partial_y \mu * (\partial_x \rho * \sigma)))\mu)[\rho/y] \\
&= (\partial_x(s\mu) * \sigma)[\rho/y] + (\partial_y(s\mu) * (\partial_x \rho * \sigma))[\rho/y] \\
&= (\partial_x \tau * \sigma)[\rho/y] + (\partial_y \tau * (\partial_x \rho * \sigma))[\rho/y].
\end{aligned}
$$

4. Assume that $\tau = \sum_{j=1}^{n} r_j \cdot s_j$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \ldots, r_n \in \mathcal{R}$ are scalars, and $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $\partial_x s_j[\rho/y] * \sigma = (\partial_x s_j * \sigma)[\rho/y] + (\partial_y s_j * (\partial_x \rho * \sigma))[\rho/y]$ for each $1 \leqslant j \leqslant n$. Hence,

$$
\begin{aligned}
\partial_x \tau[\rho/y] * \sigma &= \partial_x \left( \sum_{j=1}^{n} r_j \cdot s_j \right) [\rho/y] * \sigma \\
&= \partial_x \left( \sum_{j=1}^{n} r_j \cdot s_j[\rho/y] \right) * \sigma \\
&= \sum_{j=1}^{n} r_j \cdot \partial_x s_j[\rho/y] * \sigma \\
&= \sum_{j=1}^{n} r_j \cdot ((\partial_x s_j * \sigma)[\rho/y] + (\partial_y s_j * (\partial_x \rho * \sigma))[\rho/y]) \\
&= \sum_{j=1}^{n} r_j \cdot (\partial_x s_j * \sigma)[\rho/y] + \sum_{j=1}^{n} r_j \cdot (\partial_y s_j * (\partial_x \rho * \sigma))[\rho/y] \\
&= \left( \partial_x \left( \sum_{j=1}^{n} r_j \cdot s_j \right) * \sigma \right) [\rho/y] \\
&\quad + \left( \partial_y \left( \sum_{j=1}^{n} r_j \cdot s_j \right) * (\partial_x \rho * \sigma) \right) [\rho/y] \\
&= (\partial_x \tau * \sigma)[\rho/y] + (\partial_y \tau * (\partial_x \rho * \sigma))[\rho/y].
\end{aligned}
\tag{2.3.50}
$$

This completes the induction. ❑

**Corollary 2.3.17**

Let $\tau, \sigma, \rho \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ are complex differential terms and $x, y \in \mathcal{V}$ are variables such that $x \neq y$, $y$ is not free in the terms $\sigma$ and $\rho$, and $x$ is not free in $\rho$. Then,

$$\partial_x \tau[\rho/y] * \sigma = (\partial_x \tau * \sigma)[\rho/y]. \tag{2.3.51}$$

♡

***Proof*.** As we can see, $\partial_x \tau[\rho/y] * \sigma = (\partial_x \tau * \sigma)[\rho/y] + (\partial_y \tau * (\partial_x \rho * \sigma))[\rho/y] = (\partial_x \tau * \sigma)[\rho/y] + (\partial_y \tau * 0)[\rho/y] = (\partial_x \tau * \sigma)[\rho/y] + 0 = (\partial_x \tau * \sigma)[\rho/y]$, where we used Lemma 2.3.16 for the first equality and Lemma 2.3.11 for the second equality. ❑

**Lemma 2.3.18**

Let $\tau, \sigma, \rho \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ be complex differential terms and $x, y \in \mathcal{V}$ be variables such that $x \neq y$, $x \notin \mathrm{FV}(\rho)$, and $y \notin \mathrm{FV}(\tau)$. Then,

$$(\partial_x \tau * \sigma)[\rho/y] = \partial_x \tau * (\sigma[\rho/y]). \tag{2.3.52}$$

♡

***Proof*.** We prove this by induction on $\tau$.

1. Assume that $\tau = D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)$, where $i_1, \ldots, i_n \in \mathbb{N}$, $z \in \mathcal{V}$ is a variable, and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Since $y \notin fv(\tau) = \mathrm{FV}\,(D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)) = \mathrm{FV}\,(z) + \mathrm{FV}\,(t_1) + \ldots + \mathrm{FV}\,(t_n)$, for all $1 \leqslant j \leqslant n$, we have that $y \notin \mathrm{FV}\,(t_j)$, so by the induction hypothesis, $(\partial_x t_j * \sigma)[\rho/y] = \partial_x t_j * \sigma[\rho/y]$. Hence,

$$\begin{aligned}
(\partial_x \tau * \sigma)[\rho/y] &= (\partial_x (D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)) * \sigma)[\rho/y] \\
&= \left( \delta_{x,z} D_{i_1,\ldots,i_n} \sigma * (t_1, \ldots, t_n) + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z * (t_1, \ldots, \partial_x t_j * \sigma, \ldots, t_n) \right) [\rho/y] \\
&= \delta_{x,z} D_{i_1,\ldots,i_n} \sigma[\rho/y] * (t_1[\rho/y], \ldots, t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z[\rho/y] * (t_1[\rho/y], \ldots, (\partial_x t_j * \sigma)[\rho/y], \ldots, t_n[\rho/y]) \\
&= \delta_{x,z} D_{i_1,\ldots,i_n} \sigma[\rho/y] * (t_1, \ldots, t_n) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z * (t_1, \ldots, (\partial_x t_j * \sigma)[\rho/y], \ldots, t_n) \\
&= \delta_{x,z} D_{i_1,\ldots,i_n} \sigma[\rho/y] * (t_1, \ldots, t_n) \\
&\quad + \sum_{j=1}^{n} D_{i_1,\ldots,i_n} z * (t_1, \ldots, \partial_x t_j * \sigma[\rho/y], \ldots, t_n) \\
&= \partial_x (D_{i_1,\ldots,i_n} z * (t_1, \ldots, t_n)) * \sigma[\rho/y] \\
&= \partial_x \tau * \sigma[\rho/y],
\end{aligned} \tag{2.3.53}$$

where we used the fact that $y \notin \mathrm{FV}\,(z)$ and $y \notin \mathrm{FV}\,(t_j)$ for all $1 \leqslant j \leqslant n$ in the fourth equality.

2. Assume that $\tau = D_1^n \lambda z.s * (t_1, \ldots, t_n)$ such that $n \in \mathbb{N}_0$, $z$ is a fresh variable, and $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Since $y \notin fv(\tau) = \mathrm{FV}\,(D_1^n \lambda z.s * (t_1, \ldots, t_n)) = \mathrm{FV}\,(\lambda s.z) + \mathrm{FV}\,(t_1) + \ldots + \mathrm{FV}\,(t_n) = \mathrm{FV}\,(s) - \{z\} + \mathrm{FV}\,(t_1) + \ldots + \mathrm{FV}\,(t_n)$, we have that $y \notin \mathrm{FV}\,(s)$ and $y \notin \mathrm{FV}\,(t_j)$ for all $1 \leqslant j \leqslant n$, so by the induction hypothesis, $(\partial_x s * \sigma)[\rho/y] = \partial_x s * \sigma[\rho/y] and (\partial_x t_j * \sigma)[\rho/y] = \partial_x t_j * \sigma[\rho/y]$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
(\partial_x \tau * \sigma)[\rho/y] &= (\partial_x (D_1^n \lambda z.s * (t_1, \ldots, t_n)) * \sigma)[\rho/y] \\
&= \left( D_1^n \lambda z.(\partial_x s * \sigma) * (t_1, \ldots, t_n) + \sum_{j=1}^{n} D_1^n \lambda z.s * (t_1, \ldots, \partial_x t_j * \sigma, \ldots, t_n) \right) [\rho/y] \\
&= D_1^n \lambda z.(\partial_x s * \sigma)[\rho/y] * (t_1[\rho/y], \ldots, t_n[\rho/y]) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s[\rho/y] * (t_1[\rho/y], \ldots, (\partial_x t_j * \sigma)[\rho/y], \ldots, t_n[\rho/y]) \\
&= D_1^n \lambda z.(\partial_x s * \sigma)[\rho/y] * (t_1, \ldots, t_n) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s * (t_1, \ldots, (\partial_x t_j * \sigma)[\rho/y], \ldots, t_n) \\
&= D_1^n \lambda z.(\partial_x s * \sigma[\rho/y]) * (t_1, \ldots, t_n) \\
&\quad + \sum_{j=1}^{n} D_1^n \lambda z.s * (t_1, \ldots, \partial_x t_j * \sigma[\rho/y], \ldots, t_n) \\
&= \partial_x (D_1^n \lambda z.s * (t_1, \ldots, t_n)) * \sigma[\rho/y] \\
&= \partial_x \tau * \sigma[\rho/y].
\end{aligned} \tag{2.3.54}$$

3. Assume that $\tau = s\mu$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ is a complex differential term. Since $y \notin \mathrm{FV}\,(s\mu) = \mathrm{FV}\,(s) + \mathrm{FV}\,(\mu)$, we know that $y \notin \mathrm{FV}\,(s)$ and $y \notin \mathrm{FV}\,(\mu)$, so by the induction hypothesis, $(\partial_x s * \sigma)[\rho/y] = \partial_x s * (\sigma[\rho/y])$ and $(\partial_x \mu * \sigma)[\rho/y] =$

$\partial_x \mu * (\sigma[\rho/y])$. Hence,

$$\begin{aligned}
(\partial_x s\mu * \sigma)[\rho/y] &= ((\partial_x s * \sigma)\mu + (\mathcal{D}_1 s * (\partial_x \mu * \sigma))\mu)[\rho/y] \\
&= ((\partial_x s * \sigma)[\rho/y])\mu[\rho/y] + (\mathcal{D}_1 s[\rho/y] * (\partial_x \mu * \sigma)[\rho/y])\mu[\rho/y] \\
&= ((\partial_x s * \sigma)[\rho/y])\mu + (\mathcal{D}_1 s * (\partial_x \mu * \sigma)[\rho/y])\mu \\
&= (\partial_x s * \sigma[\rho/y])\mu + (\mathcal{D}_1 s * (\partial_x \mu * \sigma[\rho/y]))\mu \\
&= \partial_x s\mu * \sigma[\rho/y] \\
&= \partial_x \tau * \sigma[\rho/y].
\end{aligned} \tag{2.3.55}$$

4. Assume that $\tau = \sum_{j=1}^{n} r_j \cdot s_j$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \ldots, r_n \in \mathcal{R}$ are scalars, and $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Since $y \notin fv(\tau) = \mathrm{FV}\left(\sum_{j=1}^{n} r_j \cdot s_j\right) = \cup_{j=1}^{n} \mathrm{FV}(s_j)$, by the induction hypothesis, $(\partial_x s_j * \sigma)[\rho/y] = \partial_x s_j * (\sigma[\rho/y])$ for each $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
(\partial_x \tau * \sigma)[\rho/y] &= \left(\partial_x \left(\sum_{j=1}^{n} r_j \cdot s_j\right) * \sigma\right)[\rho/y] \\
&= \left(\sum_{j=1}^{n} r_j \cdot \partial_x s_j * \sigma\right)[\rho/y] \\
&= \sum_{j=1}^{n} r_j \cdot (\partial_x s_j * \sigma)[\rho/y] \\
&= \sum_{j=1}^{n} r_j \cdot \partial_x s_j * \sigma[\rho/y] \\
&= \partial_x \left(\sum_{j=1}^{n} r_j \cdot s_j\right) * \sigma[\rho/y] \\
&= \partial_x \tau * \sigma[\rho/y].
\end{aligned} \tag{2.3.56}$$

This completes the induction. ❏

**Corollary 2.3.19. Synthetic Chain Rule**

Let $\tau, \sigma, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential terms and $x, y \in \mathcal{V}$ be variables such that $x \neq y$ and $x \notin \mathrm{FV}(\rho)$. Then,

$$(\partial_x \tau * \sigma)[\rho/y] = \partial_x \tau[\rho/y] * (\sigma[\rho/y]). \tag{2.3.57}$$

♡

***Proof.*** Take any variable $z \in \mathcal{V}$ such that $z$ is distinct from $x$ and $y$ and is not free in $\sigma$ and $\rho$. Let $\tau' := \tau[z/y]$. Since $z \notin \mathrm{FV}(\tau)$, by Lemma 2.3.7(iv), $\tau'[y/z] = \tau[z/y][y/z] = \tau[y/y]$, and by Lemma 2.2.29(iii), $\tau[y/y] = \tau$. That is, $\tau = \tau'[y/z]$. Hence,

$$\begin{aligned}
\partial_x \tau * \sigma &= \partial_x \tau'[y/z] * \sigma \\
\partial_x \tau * \sigma &= (\partial_x \tau' * \sigma)[y/z] \\
(\partial_x \tau * \sigma)[\rho/y] &= (\partial_x \tau' * \sigma)[y/z][\rho/y] \\
&= (\partial_x \tau' * \sigma)[\rho/y][y[\rho/y]/z] \\
&= (\partial_x \tau' * \sigma)[\rho/y][\rho/z]
\end{aligned} \tag{2.3.58}$$

$$
\begin{aligned}
&= (\partial_x \tau' * \sigma[\rho/y])[\rho/z] \\
&= \partial_x \tau'[\rho/z] * \sigma[\rho/y] \\
&= \partial_x \tau[z/y][\rho/z] * \sigma[\rho/y] \\
&= \partial_x \tau[\rho/y] * \sigma[\rho/y],
\end{aligned}
$$

where the second equality follows from the fact that $x, y, z$ are distinct variables, $z \notin \mathrm{FV}(\sigma)$, and Corollary 2.3.17; the fourth equality follows from the fact that $y \neq z \notin \mathrm{FV}(\rho)$ and Lemma 2.3.7(iv); the sixth equality follows from the fact that $y \notin \mathrm{FV}(\tau') = \mathrm{FV}\,\tau[z/y] \subset (\mathrm{FV}(\tau) - \{y\}) \cup \{z\}$ and Lemma 2.3.18; the seventh equality follows from the fact that $x$ and $z$ are distinct and not free in $\rho$, $z \notin \mathrm{FV}(\sigma[\rho/y]) \subset (\mathrm{FV}(\sigma) - \{y\}) \cup \mathrm{FV}(\rho)$, and Corollary 2.3.17 again; and the ninth equality follows from the fact that $z \notin \mathrm{FV}(\tau)$ and Lemma 2.3.7(iv). ❏

**Lemma 2.3.20**

Let $k \in \mathbb{N}_0$ be a nonnegative integer, $i \in \mathbb{N}$ be a positive integer, $x_1, \dots, x_k \in \mathcal{V}$ be variables, and $\tau, \sigma, \rho_1, \dots, \rho_k \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ be complex differential terms. Then,

$$
\frac{\partial^k (\mathcal{D}_i \tau * \sigma)}{\partial x_1 \dots \partial x_k} * (\rho_1, \dots, \rho_k) = \sum_{r=0}^{k} \sum_{\substack{j_1, \dots, j_k = 1 \\ j_1 < \dots < j_r \\ j_{r+1} < \dots < j_k}}^{k} \mathcal{D}_i \tau_{k, j_1, \dots, j_r} * \sigma_{k, j_{r+1}, \dots, j_k}, \tag{2.3.59}
$$

where

$$
\tau_{k, j_1, \dots, j_r} = \frac{\partial^r \tau}{\partial x_{j_1} \dots \partial x_{j_r}} * (\rho_{j_1}, \dots, \rho_{j_r}), \tag{2.3.60a}
$$

$$
\sigma_{k, j_{r+1}, \dots, j_k} = \frac{\partial^{k-r} \sigma}{\partial x_{j_{r+1}} \dots \partial x_{j_k}} * (\rho_{j_{r+1}}, \dots, \rho_{j_k}). \tag{2.3.60b}
$$

♡

***Proof.*** We prove this via induction on $k$. When $k = 0$, this is trivial. Assume that $k = n + 1$, where $n \in \mathbb{N}$ is a nonnegative integer. Then,

$$
\begin{aligned}
\frac{\partial^k (\mathcal{D}_i \tau * \sigma)}{\partial x_1 \dots \partial x_k} * (\rho_1, \dots, \rho_k) &= \frac{\partial^{n+1} (\mathcal{D}_i \tau * \sigma)}{\partial x_1 \dots \partial x_n \partial x_{n+1}} * (\rho_1, \dots, \rho_n, \rho_{n+1}) \\
&= \partial_{x_{n+1}} \left( \frac{\partial^n (\mathcal{D}_i \tau * \sigma)}{\partial x_1 \dots \partial x_n} * (\rho_1, \dots, \rho_n) \right) * \rho_{n+1} \\
&= \partial_{x_{n+1}} \left( \sum_{r=0}^{n} \sum_{\substack{j_1, \dots, j_n = 1 \\ j_1 < \dots < j_r \\ j_{r+1} < \dots < j_n}}^{n} \mathcal{D}_i \tau_{n, j_1, \dots, j_r} * \sigma_{n, j_{r+1}, \dots, j_n} \right) * \rho_{n+1} \qquad (2.3.61) \\
&= \sum_{r=0}^{n} \sum_{\substack{j_1, \dots, j_n = 1 \\ j_1 < \dots < j_r \\ j_{r+1} < \dots < j_n}}^{n} \partial_{x_{n+1}} (\mathcal{D}_i \tau_{n, j_1, \dots, j_r} * \sigma_{n, j_{r+1}, \dots, j_n}) * \rho_{n+1} \\
&= \sum_{r=0}^{n} \sum_{\substack{j_1, \dots, j_n = 1 \\ j_1 < \dots < j_r \\ j_{r+1} < \dots < j_n}}^{n} \Big( \mathcal{D}_i (\partial_{x_{n+1}} \tau_{n, j_1, \dots, j_r} * \rho_{n+1}) * \sigma_{n, j_{r+1}, \dots, j_n}
\end{aligned}
$$

$$
\begin{aligned}
&\qquad + \mathcal{D}_i \tau_{n,j_1,\ldots,j_r} * (\partial_{x_{n+1}} \sigma_{n,j_{r+1},\ldots,j_n} * \rho_{n+1})\Big) \\
&= \sum_{r=0}^{n} \sum_{\substack{j_1,\ldots,j_n=1 \\ j_1<\ldots<j_r \\ j_{r+1}<\ldots<j_n}}^{n} \Big(\mathcal{D}_i \tau_{n+1,j_1,\ldots,j_r,n+1} * \sigma_{n+1,j_{r+1},\ldots,j_n} \\
&\qquad + \mathcal{D}_i \tau_{n+1,j_1,\ldots,j_r} * \sigma_{n+1,j_{r+1},\ldots,j_n,n+1}\Big) \\
&= \sum_{r=0}^{n+1} \sum_{\substack{j_1,\ldots,j_n,j_{n+1}=1 \\ j_1<\ldots<j_r \\ j_{r+1}<\ldots<j_{n+1}}}^{n+1} \mathcal{D}_i \tau_{n+1,j_1,\ldots,j_r} * \sigma_{n+1,j_{r+1},\ldots,j_{n+1}} \\
&= \sum_{r=0}^{k} \sum_{\substack{j_1,\ldots,j_k=1 \\ j_1<\ldots<j_r \\ j_{r+1}<\ldots<j_k}}^{k} \mathcal{D}_i \tau_{k,j_1,\ldots,j_r} * \sigma_{k,j_{r+1},\ldots,j_k},
\end{aligned}
$$

where we used the induction hypothesis on $n$ for the third equality and Lemma 2.3.10 for the fifth equality. ❑

**Lemma 2.3.21**

Let $k \in \mathbb{N}_0$ be a nonnegative integer, let $x_1, \ldots, x_n \in \mathcal{V}$ be variables, let $s \in \Lambda_{d,\mathcal{V}}$ be a simple differential term, and let $\tau, \sigma_1, \ldots, \sigma_k \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential terms. Then,

$$
\frac{\partial^k(s\tau)}{\partial x_1 \ldots \partial x_k} * (\sigma_1, \ldots, \sigma_k) = \sum_{q=1}^{n} \sum_{0=r_0<\ldots<r_{q+1}=n} \sum_{\substack{j_1,\ldots,j_n=1 \\ \forall 1\leqslant i<q(j_{r_{i-1}+1}<\ldots<r_{j_i})}}^{n} (\mathcal{D}_1^q s' * (\tau_1, \ldots, \tau_n))\tau, \quad (2.3.62)
$$

where

$$
s' = \frac{\partial^{r_1}\tau}{\partial x_{j_1} \ldots \partial x_{j_{r_1}}} * (\sigma_{j_1}, \ldots, \sigma_{j_r}), \quad (2.3.63a)
$$

$$
\tau_i = \frac{\partial^{r_{i+1}-r_i}\sigma}{\partial x_{j_{r_i+1}} \ldots \partial x_{j_{r_{i+1}}}} * (\sigma_{j_{r_i+1}}, \ldots, \sigma_{j_{r_{i+1}}}) \quad \forall 1 \leqslant j \leqslant q. \quad (2.3.63b)
$$

♡

***Proof.*** By induction on $k$. The detailes are left to the reader. ❑

## 2.4 Defining Differential Reduction

In this section, we are doing to talk about a relation on differential terms called $\beta$-reduction, which we represent some notion of evaluating a differential term. The idea is that a differential term should be related via $\beta$-reduction to a differential term that is the result of performing a “single step” of evaluation on the original differential term. But before we can define $\beta$-reduction for differential terms, we need to define two types of relations that behave nicely with the structure of differential $\lambda$-calculus: *linear relations*, which behave nicely with linear combinations, and *contextual relations*, which behave nicely with linear combinations, abstraction, application, and linearization.

**Definition 2.4.1. Linear Relations**

Let $\to_R$ be a relation from complex differential terms to complex differential terms such that the following two properties hold:

(i) $0 \to_R 0$

(ii) Given scalars $r_1, r_2 \in \mathcal{R}$ and $\sigma, \sigma', \tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ complex differential terms such that $\sigma \to_R \sigma'$ and $\tau \to_R \tau'$, $r_1\sigma + r_2\tau \to_R r_1\sigma' + r_2\tau'$.

Then, $\to_R$ is a *linear relation.* ♣

**Definition 2.4.2. Contextual Relations**

Let $\to_R$ be a reflexive, linear relation from complex differential terms to complex differential terms such that

$$\lambda x.\sigma \to_R \lambda x.\sigma' \tag{2.4.1a}$$
$$\sigma\tau \to_R \sigma'\tau' \tag{2.4.1b}$$
$$\mathcal{D}_i\sigma * \tau \to_R \mathcal{D}_i\sigma' * \tau' \tag{2.4.1c}$$

when $x \in \mathcal{V}$ is a variable, $i \in \mathbb{N}$ is a positive integer, and $\sigma, \sigma', \tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms such that $\sigma \to_R \sigma'$ and $\tau \to_R \tau'$. Then, $\to_R$ is a *contextual relation.* ♣

Since contextual relations behave nicely with linear combinations, abstraction, application, and linearization, it is not too hard to see that contextual relations also behave nicely with partial differentiation and substitution.

**Lemma 2.4.3**

Let $\to_R$ be a contextual relation from complex differential terms to complex differential terms, $\sigma, \tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms such that $\tau \to_R \tau'$. Then, $\partial_x\sigma * \tau \to_R \partial_x\sigma * \tau'$. ♡

***Proof.*** We prove this by induction on $\sigma$.

1. Assume that $\sigma = D_{i_1,\ldots,i_n}y * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, $y \in \mathcal{V}$ is a variable, and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are terms. By the induction hypothesis, $\partial_x t_j * \tau \to_R \partial_x t_j * \tau'$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
\partial_x\sigma * \tau &= \partial_x(D_{i_1,\ldots,i_n}y * (t_1, \ldots, t_n)) * \tau \\
&= \delta x, y D_{i_1,\ldots,i_n}\tau * (t_1, \ldots, t_n) + \sum_{i=1}^{n} D_{i_1,\ldots,i_n}y * (t_1, \ldots, \partial_x t_j * \tau, t_n) \\
&= D_{i_1}(\ldots(D i_n \tau * t_n)\ldots) * t_1 \\
&\qquad + \sum_{i=1}^{n} D_{i_1}(\ldots(\ldots(D i_n y * t_n)\ldots) * (\partial_x t_j * \tau)\ldots) * t_1 \\
&\to_R D_{i_1}(\ldots(D i_n \tau' * t_n)\ldots) * t_1 \\
&\qquad + \sum_{i=1}^{n} D_{i_1}(\ldots(\ldots(D i_n y * t_n)\ldots) * (\partial_x t_j * \tau')\ldots) * t_1 \\
&= \cdots = \partial_x\sigma * u'.
\end{aligned} \tag{2.4.2}$$

2. Assume that $\sigma = D_1^n \lambda z.s * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $z \in \mathcal{V}$ is a fresh variable and $s, t_1, \ldots, t_n$ are simple differential terms. By the induction hypothesis, $\partial_x s * \tau \to_R \partial_x s * \tau'$ and $\partial_x t_j * \tau \to_R \partial_x t_j * \tau'$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
\partial_x \sigma * \tau &= \partial_x (D_1^n \lambda z.s * (t_1, \ldots, t_n)) * \tau \\
&= D_1^n \lambda z.(\partial_x s * \tau) * (t_1, \ldots, t_n) + \sum_{i=1}^{n} D_1^n \lambda z.s * (t_1, \ldots, \partial_x t_j * \tau, \ldots, t_n) \\
&= \mathcal{D}_1(\ldots(\mathcal{D}_1 \lambda z.(\partial_x s * \tau) * t_1)\ldots) * t_n \\
&\quad + \sum_{i=1}^{n} \mathcal{D}_1(\ldots(\mathcal{D}_1(\ldots(\mathcal{D}_1 \lambda z.s * t_1)\ldots) * (\partial_x t_j * \tau))\ldots) * t_n \\
&\to_R \mathcal{D}_1(\ldots(\mathcal{D}_1 \lambda z.(\partial_x s * \tau') * t_1)\ldots) * t_n \\
&\quad + \sum_{i=1}^{n} \mathcal{D}_1(\ldots(\mathcal{D}_1(\ldots(\mathcal{D}_1 \lambda z.s * t_1)\ldots) * (\partial_x t_j * \tau'))\ldots) * t_n \\
&= \cdots = \partial_x \sigma * \tau'.
\end{aligned} \tag{2.4.3}$$

3. Assume that $\sigma = s\mu$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential terms and $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential terms. By the induction hypothesis, $\partial_x s * \tau \to_R \partial_x s * \tau'$ and $\partial_x \mu * \tau \to_R \partial_x \mu * \tau'$. Hence,

$$\begin{aligned}
\partial_x \sigma * \tau = \partial_x s\mu * \tau &= (\partial_x s * \tau)\mu + (\mathcal{D}_1 s * (\partial_x \mu * \tau))\mu \\
&\to_R (\partial_x s * \tau')\mu + (\mathcal{D}_1 s * (\partial_x \mu * \tau'))\mu = \cdots = \partial_x \sigma * \tau'.
\end{aligned} \tag{2.4.4}$$

4. Assume that $\tau = \sum_{j=1}^{n} r_j \cdot s_j$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \ldots, r_n \in \mathcal{R}$ are scalars, and $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $\partial_x s_j * \tau \to_R \partial_x s_j * \tau'$ for each $1 \leqslant j \leqslant n$. Hence,

$$\partial_x \sigma = \partial_x \left( \sum_{j=1}^{n} r_j \cdot s_j \right) * \tau = \sum_{j=1}^{n} r_j \cdot \partial_x s_j * \tau \to_R \sum_{j=1}^{n} r_j \cdot \partial_x s_j * \tau' = \cdots = \partial_x \sigma * \tau'. \tag{2.4.5}$$

This completes the induction. ❏

**Lemma 2.4.4**

Let $\to_R$ be a contextual relation from complex differential terms to complex differential terms, $\sigma, \tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential termssuch that $\tau \to_R \tau'$. Then, $\sigma[\tau/x] \to_R \sigma[\tau'/x]$. ♡

***Proof***. We prove this via induction on $\sigma$.

1. Assume that $\sigma = D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $t_j[\tau/x] = t_j[\tau'/x]$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned}
\sigma[\tau/x] &= (D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n))[\tau/x] = D_{i_1,\ldots,i_n} \tau * (t_1[\tau/x], \ldots, t_n[\tau/x]) \\
&= \mathcal{D}_1(\ldots(\mathcal{D}_1 \tau * t_1[\tau/x])\ldots) * t_n[\tau/x] \\
&\to_R \mathcal{D}_1(\ldots(\mathcal{D}_1 \tau' * t_1[\tau/x])\ldots) * t_n[\tau'/x] = \cdots = \sigma[\tau'/x]).
\end{aligned} \tag{2.4.6}$$

2. Assume that $\sigma = D_{i_1,\dots,i_n} y * (t_1, \dots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \dots, i_n \in \mathbb{N}$ are positive integers, $y \in \mathcal{V}$ is a variable such that $y \neq x$, and $t_1, \dots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $t_j[\tau/x] = t_j[\tau'/x]$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned} \sigma[\tau/x] &= (D_{i_1,\dots,i_n} y * (t_1, \dots, t_n))[\tau/x] = D_{i_1,\dots,i_n} y * (t_1[\tau/x], \dots, t_n[\tau/x]) \\ &= \mathcal{D}_1(\dots(\mathcal{D}_1 y * t_1[\tau/x])\dots) * t_n[\tau/x] \\ &\to_R \mathcal{D}_1(\dots(\mathcal{D}_1 y * t_1[\tau/x])\dots) * t_n[\tau'/x] = \dots = \sigma[\tau'/x]. \end{aligned} \tag{2.4.7}$$

3. Assume that $\sigma = D_1^n \lambda z.s * (t_1, \dots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $z \in \mathcal{V}$ is a fresh variable and $s, t_1, \dots, t_n$ are simple differential terms. By the induction hypothesis, $s[\tau/x] \to_R s[\tau'/x]$ and $t_j[\tau/x] \to_R t_j[\tau'/x]$ for all $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned} \sigma[\tau/x] &= (D_1^n \lambda z.s * (t_1, \dots, t_n))[\tau/x] = D_1^n \lambda z.s[\tau/x] * (t_1[\tau/x], \dots, t_n[\tau/x]) \\ &= \mathcal{D}_1(\dots(\mathcal{D}_1 \lambda z.s[\tau/x] * t_1[\tau/x])\dots) * t_n[\tau/x] \\ &\to_R \mathcal{D}_1(\dots(\mathcal{D}_1 \lambda z.s[\tau'/x] * t_1[\tau'/x])\dots) * t_n[\tau'/x] = \dots = \sigma[\tau'/x]. \end{aligned} \tag{2.4.8}$$

4. Assume that $\sigma = s\mu$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential terms and $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential terms. By the induction hypothesis, $s[\tau/x] \to_R s[\tau'/x]$ and $\mu[\tau'/x] \to_R \mu[\tau'/x]$. Hence,

$$\sigma[\tau/x] = (s\mu)[\tau/x] = s[\tau/x]\mu[\tau/x] \to_R s[\tau'/x]\mu[\tau'/x] = (s\mu)[\tau'/x] = \sigma[\tau'/x]. \tag{2.4.9}$$

5. Assume that $\tau = \sum_{j=1}^n r_j \cdot s_j$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \dots, r_n \in \mathcal{R}$ are scalars, and $s_1, \dots, s_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. By the induction hypothesis, $s_j[\tau/x] \to_R s_j[\tau'/x]$ for each $1 \leqslant j \leqslant n$. Hence,

$$\begin{aligned} \sigma[\tau/x] &= \left(\sum_{j=1}^n r_j \cdot s_j\right)[\tau/x] = \sum_{j=1}^n r_j \cdot s_j[\tau/x] \\ &\to_R \sum_{j=1}^n r_j \cdot s_j[\tau'/x] = \left(\sum_{j=1}^n r_j \cdot s_j\right)[\tau'/x] = \sigma[\tau'/x]. \end{aligned} \tag{2.4.10}$$

This completes the induction. ❑

We are interested in extending a relation $\to_R$ from simple to complex differential terms to a relation $\to_S$ from complex to complex differential terms. The first and obvious way to do this is by linearity. That is, given complex differential terms $\tau, \sigma \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, if $\tau$ is a linear combination of simple differential terms $s_1, \dots, s_n \in \Lambda_{d,\mathcal{V}}$, $\tau \to_S \sigma$ when there exists $\sigma_1, \dots, \sigma_n$ such that $\sigma$ is obtained by replacing $s_i$ in $\tau$ with $\sigma_i$ and $s_i \to_R \sigma_i$ for each $1 \leqslant i \leqslant n$. This gives us the linear extension as laid out in Definition 2.4.5. However, let us say that $\to_R$ represents a single computational step, and we want $\to_S$ to also represent a single computation. Then, given complex differential terms $\tau, \sigma \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, if $\tau$ is a linear combination of simple differential terms $s_1, \dots, s_n \in \Lambda_{d,\mathcal{V}}$, we want $\tau \to_S \sigma$ if there exists a single $i_0$ such that there exists $\sigma_{i_0}$ such that $\sigma$ is obtained by replacing $s_{i_0}$ in $\tau$ with $\sigma_{i_0}$ and $s_{i_0} \to_R \sigma_{i_0}$, This gives us the pseudo-linear extension as laid out in Definition 2.4.6.

**Definition 2.4.5. The Linear Extension of a Differential Relation**

Given a relation $\to_R$ from simple differential terms to complex differential terms, the *linear extension* $\to_{\bar{R}}$ of $\to_R$ is the smallest relation from complex differential terms to complex differential terms containing $\to_R$ such that

$$\sum_{j=1}^{n} r_j \cdot s_j \to_{\bar{R}} \sum_{j=1}^{n} r_j \cdot \sigma_j \tag{2.4.11}$$

whenever $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \ldots, r_n \in \mathcal{R}$ are scalars, $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms, and $\tau_1, \ldots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms such that $s_j \to_R \sigma_j$ for each $1 \leqslant j \leqslant n$.

**Note:** $\to_{\bar{R}}$ is the smallest linear relation containing $\to_R$.

♣

**Definition 2.4.6. The Pseudo-Linear Extension of a Differential Relation**

Given a relation $\to_R$ from simple differential terms to complex differential terms, the *pseudo-linear extension* $\to_{\tilde{R}}$ of $\to_R$ is the smallest relation from complex differential terms to complex differential terms containing $\to_R$ such that

$$r \cdot s + \tau \to_{\tilde{R}} r \cdot \sigma + \tau \tag{2.4.12}$$

whenever $r \in \mathcal{R}$ is a scalar, $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential term, and $\sigma, \tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms such that $s \to_{\mathcal{R}} \sigma$.

**Note:** $\to_{\tilde{R}}$ is not a linear relation as $0 \not\to_{\tilde{R}} 0$.

♣

We, now, define differential pre-$\beta$-reduction which relates simple differential terms to a complex differential terms that results by performing “a single step” of computation on the simple differential term.

**Definition 2.4.7. Differential Pre-$\beta$-Reduction**

Define differential pre-$\beta$-reduction $\to_{\beta^1}$ to be the smallest (with respect to inclusion) relation from simple differential terms to complex differential terms satisfying the following properties:

(i) Given a simple differential term $s \in \Lambda_{d,\mathcal{V}}$ and complex differential terms $\sigma, \tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s \to_{\beta^1} \sigma$,
$$s\tau \to_{\beta^1} \sigma\tau. \tag{2.4.13}$$

(ii) Given a simple differential term $s \in \Lambda_{d,\mathcal{V}}$ and complex differential terms $\tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\tau \to_{\tilde{\beta}^1} \tau'$,
$$s\tau \to_{\beta^1} s\tau'. \tag{2.4.14}$$

(iii) Given a variable $x \in \mathcal{V}$, a simple differential term $s \in \Lambda_{d,\mathcal{V}}$, and a complex differential term $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$,
$$(\lambda x.s)\tau \to_{\beta^1} s[\tau/x]. \tag{2.4.15}$$

(iv) Given a positive integer $n \in \mathbb{N}$, a variable $x \in \mathcal{V}$, positive integers $i_1, \ldots, i_n \in \mathbb{N}$, simple differential terms $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, an integer $1 \leqslant j \leqslant n$, and a complex differential term $\tau_j \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $t_j \to_{\beta^1} \tau_j$,
$$D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_j, \ldots, t_n) \to_{\beta^1} D_{i_1,\ldots,i_n} x * (t_1, \ldots, \tau_j, \ldots, t_n). \tag{2.4.16}$$

(v) Given a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, simple differential terms $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, and a complex differential term $\sigma \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s \to_{\beta^1} \sigma$,

$$D_1^n \lambda x.s * (t_1, \ldots, t_n) \to_{\beta^1} D_1^n \lambda x.\sigma * (t_1, \ldots, t_n). \tag{2.4.17}$$

(vi) Given a positive integer $n \in \mathbb{N}$, a variable $x \in \mathcal{V}$, simple differential terms $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, an integer $1 \leqslant j \leqslant n$, and a complex differential term $\tau_j \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $t_j \to_{\beta^1} \tau_j$,

$$D_1^n \lambda x.s * (t_1, \ldots, t_j, \ldots, t_n) \to_{\beta^1} D_1^n \lambda x.s * (t_1, \ldots, \tau_j, \ldots, t_n). \tag{2.4.18}$$

(vii) Given a positive integer $n \in \mathbb{N}$, a variable $x \in \mathcal{V}$, simple differential terms $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, and $1 \leqslant j \leqslant n$,

$$D_1^n \lambda x.s * (t_1, \ldots, t_n) \to_{\beta^1} D_1^{n-1} \lambda x.(\partial_x s * t_j) * (t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n). \tag{2.4.19}$$

♣

While differential pre-$\beta$-reduction cannot be contextual (as it is not a relation from complex differential terms to complex differential terms), differential pre-$\beta$-reduction does, in fact, respect application, application, and linearization just like a contextual relation would.

**Lemma 2.4.8**

Let $x \in \mathcal{V}$ be a variable, $i \in \mathbb{N}$ be a positive integer, $s, t \in \Lambda_{d,\mathcal{V}}$ be simple differential terms, and $\sigma, \tau, \tau' \in \Lambda_{d,\mathcal{V}}$ be complex differential terms. Then, the following statements are true:

(i) If $s \to_{\beta^1} \sigma$, $s\tau \to_{\beta^1} \sigma\tau$.

(ii) If $\tau \to_{\tilde{\beta}^1} \tau'$, $s\tau \to_{\beta^1} s\tau'$.

(iii) If $s \to_{\beta^1} \sigma$, $\lambda x.s \to_{\beta^1} \lambda x.\sigma$.

(iv) If $s \to_{\beta^1} \sigma$, $\mathcal{D}_i s * t \to_{\beta^1} \mathcal{D}_i \sigma * t$.

(v) If $t \to_{\beta^1} \tau$, $\mathcal{D}_i s * t \to_{\beta^1} \mathcal{D}_i s * \tau$.

We say that any relation $\to_R$ from simple differential terms to complex differential terms that satisfy these five properties is *weakly contextual*. Hence, $\to_{\beta^1}$ is weakly contextual. ♡

***Proof of Part (i).*** This just follows from Definition 2.4.7(i). ❑

***Proof of Part (ii).*** This just follows from Definition 2.4.7(ii). ❑

***Proof of Part (iii).*** This just follows from Definition 2.4.7(v). ❑

***Proof of Part (iv).*** This follows by induction on $s \to_{\beta^1} \sigma$.

1. Assume that $s \to_{\beta^1} \sigma$ because there exist a simple differential term $u \in \Lambda_{d,\mathcal{V}}$ and complex differential terms $\mu, \nu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s = u\nu$, $\sigma = \mu\nu$, and $u \to_{\beta^1} \mu$. Since $u \to_{\beta^1} \mu$, by the induction hypothesis, $\mathcal{D}_{i+1} u * t \to_{\beta^1} \mathcal{D}_{i+1} \mu * t$. Hence, $\mathcal{D}_i s * t = \mathcal{D}_i(u\nu) * t = (\mathcal{D}_{i+1} u * t)\nu \to_{\beta^1} (\mathcal{D}_{i+1} \mu * t)\nu = \mathcal{D}_i(\mu\nu) * t = \mathcal{D}_i \sigma * t$.

2. Assume that $s \to_{\beta^1} \sigma$ because there exist a simple differential term $u \in \Lambda_{d,\mathcal{V}}$ and complex differential terms $\nu, \nu' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s = u\nu$, $\sigma = u\nu'$, and $\nu \to_{\tilde{\beta}^1} \nu'$. Then, $\mathcal{D}_i s * t = \mathcal{D}_i(u\nu) * t = (\mathcal{D}_{i+1} u * t)\nu \to_{\beta^1} (\mathcal{D}_{i+1} u * t)\nu' = \mathcal{D}_i(u\nu') * t = \mathcal{D}_i \sigma * t$.

3. Assume that $s \to_{\beta^1} \sigma$ because there exist a variable $x \in \mathcal{V}$, which we can choose to not be free in $t$ via $\alpha$-conversion, a simple differential terms $u \in \Lambda_{d,\mathcal{V}}$, and a complex differential terms $\nu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s = (\lambda x.u)\nu$ and $\sigma = u[\nu/x]$. Then, $\mathcal{D}_i s * t = \mathcal{D}_i((\lambda x.u)\nu) * t = (\mathcal{D}_{i+1}(\lambda x.u) * t)\nu = (\lambda x.\mathcal{D}_{i+1-1} u * t)\nu = (\lambda x.\mathcal{D}_i u * t)\nu \to_{\beta^1} = (\mathcal{D}_i u * t)[\nu/x] = \mathcal{D}_i u[\nu/x] * t[\nu/x] = \mathcal{D}_i \sigma * t$, where we used Lemma 2.3.6 for the second to last equality and the fact that $x \notin \mathrm{FV}(t)$ and Lemma 2.3.7(i) for the last equality.

4. Assume that $s \to_{\beta^1} \sigma$ because there exist a positive integers $n \in \mathbb{N}$, a variable $x \in \mathcal{V}$, positive integers $i_1, \ldots, i_n \in \mathbb{N}$, simple differential terms $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, and complex differential terms $\tau_1 \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s = D_{i_1,\ldots,i_n} x * (t_1, t_2, \ldots, t_n)$, $\sigma = D_{i_1,\ldots,i_n} * (\tau_1, t_2, \ldots, t_n)$, and $t_1 \to_{\beta^1} \tau_1$. Then, $\mathcal{D}_i s * t = \mathcal{D}_i(D_{i_1,\ldots,i_n} x * (t_1, t_2, \ldots, t_n) * t = D_{i,i_1,\ldots,i_n} x * (t, t_1, t_2, \ldots, t_n) \to_{\beta^1} D_{i,i_1,\ldots,i_n} x * (t, \tau_1, t_2, \ldots, t_n) = \mathcal{D}_i(D_{i_1,\ldots,i_n} x * (\tau_1, t_2, \ldots, t_n) * t = \mathcal{D}_i \sigma * t$.

5. Assume that $s \to_{\beta^1} \sigma$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, simple differential terms $u, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, and a complex differential term $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s = D_1^n \lambda x.u * (t_1, \ldots, t_n)$, $\sigma = D_1^n \lambda x.\mu * (t_1, \ldots, t_n)$., and $u \to_{\beta^1} \mu$. Further, suppose that $i = 1$. Then, $\mathcal{D}_i s * t = \mathcal{D}_1(D_1^n \lambda x.u * (t_1, \ldots, t_n)) * t = D_1^{n+1} \lambda x.u * (t, t_1, \ldots, t_n) \to_{\beta^1} D_1^{n+1} \lambda x.\mu * (t, t_1, \ldots, t_n) = \mathcal{D}_1(D_1^n \lambda x.\mu * (t_1, \ldots, t_n)) * t = \mathcal{D}_i \sigma * t$. Instead, suppose that $i > 1$. Since $u \to_{\beta^1} \mu$, by the induction hypothesis, $\mathcal{D}_{i-1} u * t \to_{\beta^1} \mathcal{D}_{i-1}\mu * t$. Hence, $\mathcal{D}_i s * t = \mathcal{D}_i(D_1^n \lambda x.u * (t_1, \ldots, t_n)) * t = D_1^n \lambda x.(\mathcal{D}_{i-1} u * t) * (t_1, \ldots, t_n) \to_{\beta^1} D_1^n \lambda x.(\mathcal{D}_{i-1}\mu * t) * (t_1, \ldots, t_n) = \mathcal{D}_i(D_1^n \lambda x.\mu * (t_1, \ldots, t_n)) * t = \mathcal{D}_i \sigma * t$.

6. Assume that $s \to_{\beta^1} \sigma$ because there exist a positive integer $n \in \mathbb{N}$, a variable $x \in \mathcal{V}$, simple differential terms $u, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, and a complex differential term $\tau_1 \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s = D_1^n \lambda x.u * (t_1, t_2, \ldots, t_n)$, $\sigma = D_1^n \lambda x.u * (\tau_1, t_2, \ldots, t_n)$., and $t_1 \to_{\beta^1} \tau_1$. Further, suppose that $i = 1$. Then, $\mathcal{D}_i s * t = \mathcal{D}_1(D_1^n \lambda x.u * (t_1, t_2, \ldots, t_n)) * t = D_1^{n+1} \lambda x.u * (t, t_1, t_2, \ldots, t_n) \to_{\beta^1} D_1^{n+1} \lambda x.u * (t, \tau_1, t_2, \ldots, t_n) = \mathcal{D}_1(D_1^n \lambda x.u * (\tau_1, t_2, \ldots, t_n)) * t = \mathcal{D}_i \sigma * t$. Instead, suppose that $i > 1$. Then, $\mathcal{D}_i s * t = \mathcal{D}_i(D_1^n \lambda x.u * (t_1, t_2, \ldots, t_n)) * t = D_1^n \lambda x.(\mathcal{D}_{i-1} u * t) * (t_1, t_2, \ldots, t_n) \to_{\beta^1} D_1^n \lambda x.(\mathcal{D}_{i-1} u * t) * (\tau_1, t_2, \ldots, t_n) = \mathcal{D}_i(D_1^n \lambda x.u * (\tau_1, t_2, \ldots, t_n)) * t = \mathcal{D}_i \sigma * t$.

7. Assume that $s \to_{\beta^1} \sigma$ because there exist a positive integer $n \in \mathbb{N}$, a variable $x \notin \mathrm{FV}(t)$, and simple differential terms $u, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ such that $s = D_1^n \lambda x.u * (t_1, t_2, \ldots, t_n)$ and $D_1^{n-1} \lambda x.(\partial_x u * t_1) * (t_2, \ldots, t_n)$. Further, suppose that $i = 1$. Then, $\mathcal{D}_i s * t = \mathcal{D}_1(D_1^n \lambda x.u * (t_1, t_2, \ldots, t_n)) * t = D_1^{n+1} \lambda x.u * (t, t_1, t_2, \ldots, t_n) \to_{\beta^1} D_1^n \lambda x.(\partial_x u * t_1) * (t, t_2, \ldots, t_n) = \mathcal{D}_1(D_1^{n-1} \lambda x.(\partial_x u * t_1) * (t_2, \ldots, t_n)) * t = \mathcal{D}_i \sigma * t$. Instead, suppose that $i > 1$. Then,

$$
\begin{aligned}
\mathcal{D}_i s * t &= \mathcal{D}_i(D_1^n \lambda x.u * (t_1, t_2, \ldots, t_n)) * t \\
&= D_1^n \lambda x.(\mathcal{D}_i u * t) * (t_1, t_2, \ldots, t_n) \\
&\to_{\beta^1} D_1^{n-1} \lambda x.(\partial_x(\mathcal{D}_{i-1} u * t) * t_1) * (t_2, \ldots, t_n) \\
&= D_1^{n-1} \lambda x.(\partial_x(\mathcal{D}_{i-1} u * t) * t_1) * (t_2, \ldots, t_n) \\
&= D_1^{n-1} \lambda x.(\mathcal{D}_{i-1}(\partial_x u * t_1) * t + \mathcal{D}_{i-1} u * (\partial_x t * t_1)) * (t_2, \ldots, t_n) \\
&= D_1^{n-1} \lambda x.(\mathcal{D}_{i-1}(\partial_x u * t_1) * t + \mathcal{D}_{i-1} u * 0) * (t_2, \ldots, t_n) \\
&= \mathcal{D}_i(D_1^{n-1} \lambda x.(\partial_x u * t_1) * (t_2, \ldots, t_n)) * t \\
&= \mathcal{D}_i \sigma * t,
\end{aligned}
\tag{2.4.20}
$$

where we used Lemma 2.3.10 for the fourth from last equality and the fact that $x \notin \mathrm{FV}(t)$ and Lemma 2.3.11 for the third form last equality.

This completes the induction. ❏

***Proof of Part (v).*** This follows by induction on $s$.

1. Assume that $s = D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $i_1, \ldots, i_n \in \mathbb{N}$ are positive integers, $x \in \mathcal{V}$ is a variable, and $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Then, $\mathcal{D}_i s * t = \mathcal{D}_i(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) * t = D_{i,i_1,\ldots,i_n} x * (t, t_1, \ldots, t_n) \to_{\beta^1} Di, i_1, \ldots, i_n x * (\tau, t_1, \ldots, t_n) = \mathcal{D}_i(D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n)) * \tau = \mathcal{D}_i s * \tau$.

2. Assume that $s = D_1^n \lambda x.u * (t_1, \ldots, t_n)$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $x \in \mathcal{V}$ is a fresh variable and $u, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Further, suppose that $i = 1$. Then, $\mathcal{D}_i s * t = \mathcal{D}_1(D_1^n \lambda x.u * (t_1, \ldots, t_n)) * t = D_1^{n+1} \lambda x.u * (t, t_1, \ldots, t_n) \to_{\beta^1} D_1^{n+1} \lambda x.u * (\tau, t_1, \ldots, t_n) = \mathcal{D}_1(D_1^n \lambda x.u * (t_1, \ldots, t_n)) * t = \mathcal{D}_i s * \tau$. Instead, suppose that $i > 1$. By the induction hypothesis, $D_{i-1} s * t \to_{\beta^1} D_{i-1} s * \tau$. Hence, $\mathcal{D}_i s * t = \mathcal{D}_i(D_1^n \lambda x.u * (t_1, \ldots, t_n)) * t = D_1^n \lambda x.(D_{i-1} s * t) * (t_1, \ldots, t_n) \to_{\beta^1} D_1^n \lambda x.(D_{i-1} s * \tau) * (t_1, \ldots, t_n) = \mathcal{D}_i(D_1^n \lambda x.u * (t_1, \ldots, t_n)) * \tau = \mathcal{D}_i s * \tau$. In either case, $\mathcal{D}_i s * t \to_{\beta^1} \mathcal{D}_i s * \tau$.

3. Assume that $s = u\nu$, where $u \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\nu \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ is a complex differential term. By the induction hypothesis, $\mathcal{D}_{i+1} u * t \to_{\beta^1} \mathcal{D}_{i+1} u * \tau$. Hence, $\mathcal{D}_i s * t = \mathcal{D}_i(u\nu) * t = (\mathcal{D}_{i+1} u * t)\nu \to_{\beta^1} (\mathcal{D}_{i+1} u * \tau)\nu = \mathcal{D}_i(u\nu) * \tau = \mathcal{D}_i s * \tau$.

This completes the induction. ❏

**Notation 2.4.9. The Reflexive and Transitive Closure of a Relation on $\mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$**

Let $\to_R$ be a relation on complex differential terms. Then,

1. $\twoheadrightarrow_R$ is the reflexive and transitive closure of $\to_R$, and

2. we say that $\sigma \twoheadrightarrow_{R^+} \tau$ when $\sigma$ reduces to $\tau$ in at least one step (i.e., there exists a complex differential term $\rho \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\sigma \to_R \rho \twoheadrightarrow_R \tau$).

♠

We can now finally define differential $\beta$-reduction, which captures a "a single step" of evaluation. Differential $\beta$-reduction is just the pseudo-linear extension of differential pre-$\beta$-reduction.

**Definition 2.4.10. Differential $\beta$-Reduction**

Let differential $\beta$-Reduction $\to_\beta$ be the relation on complex differential terms $\to_{\tilde{\beta}_1}$.

♣

Since differential pre-$\beta$-reduction is weekly contextual, we would expect that differential $\beta$-reduction is a contextual relation, and it is! The next few lemmas prove this.

**Lemma 2.4.11**

Let $\sigma, \sigma', \tau, \tau' \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ are complex differential terms and $r_1, r_2 \in \mathcal{R}$ be scalars such that $\sigma \twoheadrightarrow_\beta \sigma'$ and $\tau \twoheadrightarrow_\beta \tau'$, then $r_1 \cdot \sigma + r_2 \cdot \tau \twoheadrightarrow_\beta r_1 \cdot \sigma' + r_2 \cdot \tau'$.

♡

***Proof.*** This follows by induction.

1. Assume that $\sigma \twoheadrightarrow_\beta \sigma'$ and $\tau \twoheadrightarrow_\beta \tau'$ because $\sigma \to_\beta \sigma'$ and $\tau \to_\beta \tau'$. That is, $\sigma \to_{\tilde{\beta}^1} \sigma'$ and $\tau \to_{\tilde{\beta}^1} \tau'$. Then, there exist simple differential terms $u, v \in \Lambda_{d,\mathcal{V}}$, complex differential terms $\mu, \nu, \xi, \rho \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, and scalars $r_3, r_4 \in \mathcal{R}$ such that $\sigma = r_3 \cdot u + \xi$, $\sigma' = r_3 \cdot \mu + \xi$, $\tau = r_4 \cdot v + \rho$, $\tau' = r_4 \cdot \nu + \rho$, $u \to_{\beta^1} \mu$, and $v \to_{\beta^1} \nu$. Hence, $r_1 \cdot \sigma + r_2 \cdot \tau = r_1 \cdot (r_3 \cdot u + \xi) + r_2 \cdot (r_4 \cdot v + \rho) = (r_1 \cdot r_3) \cdot u + (r_2 \cdot r_4) \cdot v + r_1 \cdot \xi + r_2 \cdot \rho \to_{\tilde{\beta}^1} (r_1 \cdot r_3) \cdot \mu + (r_2 \cdot r_4) \cdot v + r_1 \cdot \xi + r_2 \cdot \rho \to_{\tilde{\beta}^1} (r_1 \cdot r_3) \cdot \mu + (r_2 \cdot r_4) \cdot \nu + r_1 \cdot \xi + r_2 \cdot \rho = r_1 \cdot (r_3 \cdot \mu + \xi) + r_2 \cdot (r_4 \cdot \nu + \rho) = r_1 \cdot \sigma' + r_2 \cdot \tau'$.

2. Assume that $\sigma \twoheadrightarrow_\beta \sigma'$ and $\tau \twoheadrightarrow_\beta \tau'$ because $\sigma \to_\beta \sigma'$ and $\tau = \tau'$. Since $\sigma \to_\beta \sigma'$, $\sigma \to_{\tilde{\beta}^1} \sigma'$, which means that there exist simple differential term $u \in \Lambda_{d,\mathcal{V}}$, complex differential terms $\mu, \xi \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, and scalars $r_3 \in \mathcal{R}$ such that $\sigma = r_3 \cdot u + \xi$, $\sigma' = r_3 \cdot \mu + \xi$, and $u \to_{\beta^1} \mu$. Hence, $r_1 \cdot \sigma + r_2 \cdot \tau = r_1 \cdot (r_3 \cdot u + \xi) + r_2 \cdot \tau = (r_1 \cdot r_3) \cdot u + r_2 \cdot \tau + r_1 \cdot \xi \to_{\tilde{\beta}^1} (r_1 \cdot r_3) \cdot \mu + r_2 \cdot \tau + r_1 \cdot \xi \to_{\tilde{\beta}^1} (r_1 \cdot r_3) \cdot \mu + r_2 \cdot \tau' + r_1 \cdot \xi = r_1 \cdot (r_3 \cdot \mu + \xi) + r_2 \cdot \tau' = r_1 \cdot \sigma' + r_2 \cdot \tau'$.

3. Assume that $\sigma \twoheadrightarrow_\beta \sigma'$ and $\tau \twoheadrightarrow_\beta \tau'$ because $\sigma = \sigma'$ and $\tau = \tau'$. Then $r_1 \cdot \sigma + r_2 \cdot \tau = r_1 \cdot \sigma' + r_2 \cdot \tau'$, so by the reflexivity of $\twoheadrightarrow_\beta$, we know that $r_1 \cdot \sigma + r_2 \cdot \tau \twoheadrightarrow_\beta r_1 \cdot \sigma' + r_2 \cdot \tau'$.

4. Assume that $\sigma \twoheadrightarrow_\beta \sigma'$ and $\tau \twoheadrightarrow_\beta \tau'$ because there exists term $\rho$ such that $\sigma \twoheadrightarrow_\beta \rho \twoheadrightarrow_\beta \sigma'$. Since $\sigma \twoheadrightarrow_\beta \rho$ and $\tau \twoheadrightarrow_\beta \tau'$, by the induction hypothesis, $r_1 \cdot \sigma + r_2 \cdot \tau \twoheadrightarrow_\beta r_1 \cdot \rho + r_2 \cdot \tau'$, and since $\rho \twoheadrightarrow_\beta \sigma'$ and $\tau' \twoheadrightarrow_\beta \tau'$ (as $\twoheadrightarrow_\beta$ is reflexive), by the induction hypothesis, $r_1 \cdot \rho + r_2 \cdot \tau' \twoheadrightarrow_\beta r_1 \cdot \sigma' + r_2 \cdot \tau'$. By the transitivity of $\twoheadrightarrow_\beta$, this implies that $r_1 \cdot \sigma + r_2 \cdot \tau \twoheadrightarrow_\beta r_1 \cdot \sigma' + r_2 \cdot \tau'$.

The remaining cases are similar. This completes the induction. ❑

**Lemma 2.4.12**

Let $\tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be complex differential term such that $\tau \twoheadrightarrow_\beta \tau'$. Then, there exists a nonnegative integer $n \in \mathbb{N}_0$, scalars $r_1, \ldots, r_n \in \mathcal{R}$, simple differential terms $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$, and complex differential terms $\sigma_1, \ldots, \sigma_n$ such that $\tau = \sum_{j=1}^{n} r_j \cdot s_j$, $\tau' = \sum_{j=1}^{n} r_j \cdot \sigma_j$, and $s_j \to_\beta \sigma_j$ for each $1 \leqslant j \leqslant n$. ♡

***Proof.*** This follows by a simple induction on $\tau \twoheadrightarrow_\beta \tau'$ and Lemma 2.2.29. ❑

**Lemma 2.4.13**

the relation $\twoheadrightarrow_\beta$ on complex differential terms is contextual. ♡

***Proof.*** Since the relation $\to_\beta$ on complex differential terms is reflexive, $0 \to_\beta 0$. Because of this fact and Lemma 2.4.11, $\to_\beta$ is linear. By Lemma 2.4.12, the fact that $\to_{\beta^1}$ is weakly contextual (which was shown in Lemma 2.4.8), and the linearity of $\to_\beta$, it follows that $\to_\beta$ satisfies (2.4.1). Hence, $\to_\beta$ is contextual. ❑

Differential $\beta$-reduction also satisfies the Church-Rosser property, but to show this is unfortunately not very straightforward. To show that $\beta$-reduction satisfies the Church-Rosser property, we first need to define a relation $\to_\rho$ from simple to complex differential terms called parallel differential pre-$\beta$ reduction. Then, we extend this relation by linearity given us the relation $\to_{\tilde{\rho}}$ on complex differential terms called parallel differential $\beta$-reduction. Next, we show that the reflexive and transitive closure of parallel differential $\beta$-reduction is the same as that for differential $\beta$-reduction. Finally, we show that parallel differential pre-$\beta$ reduction is confluent, which implies that differential $\beta$-reduction also satisfies the Church-Rosser property.

**Definition 2.4.14. Parallel Differential $\beta$-Reduction**

Define parallel differential pre-$\beta$-reduction $\to_\rho$ to be the smallest (with respect to inclusion) reflexive relation from simple differential terms to complex differential terms satisfying the following properties:

(i) Given a simple differential term $s \in \Lambda_{d,\mathcal{V}}$ and complex differential terms $\sigma, \tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s \to_\rho \sigma$ and $\tau \to_{\bar\rho} \tau'$,

$$s\tau \to_\rho \sigma\tau'. \tag{2.4.21}$$

(ii) Given a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, simple differential terms $s, t_1, \ldots, t_n, t'_j \in \Lambda_{d,\mathcal{V}}$, and complex differential terms $\sigma, \tau, \tau', \tau_1, \ldots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s \to_\rho \sigma$, $\tau \to_{\bar\rho} \tau'$, and $t_j \to_\rho \tau_j$ for all $1 \leqslant j \leqslant n$,

$$(D_1^n \lambda x.s * (t_1, \ldots, t_n))\tau \to_\rho ((\partial_x^n \sigma * (\tau_1, \ldots, \tau_n))[\tau'/x]. \tag{2.4.22}$$

(iii) Given a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, positive integers $i_1, \ldots, i_n \in \mathbb{N}$, simple differential terms $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, and a complex differential term $\tau_1, \ldots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $t_j \to_\rho \tau_j$ for all $1 \leqslant j \leqslant n$,

$$D_{i_1,\ldots,i_n} x * (t_1, \ldots, t_n) \to_\rho D_{i_1,\ldots,i_n} x * (\tau_1, \ldots, \tau_n). \tag{2.4.23}$$

(iv) Given a nonnegative integer $n \in \mathbb{N}_0$, a nonnegative integer $m \leqslant n$, a variable $x \in \mathcal{V}$, simple differential terms $s, t_1, \ldots, t_n, \in \Lambda_{d,\mathcal{V}}$, complex differential terms $\sigma, \tau_1, \ldots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $s \to_\rho \sigma$ and $t_j \to_\rho \tau_j$ for all $1 \leqslant j \leqslant n$, and distinct positive integers $j_1, \ldots, j_n \in \{1, \ldots, n\}$ such that $j_1 < \ldots < j_m$ and $j_{m+1} < \cdots < j_n$,

$$D_1^n \lambda x.s * (t_1, \ldots, t_n) \to_\rho D^{n-m} \lambda x.((\partial_x^m \sigma * (\tau_{j_1}, \ldots, \tau_{j_m})) * (\tau_{j_{m+1}}, \ldots, \tau_{j_n}). \tag{2.4.24}$$

♣

**Lemma 2.4.15**

The relation $\to_{\bar\rho}$ on complex differential terms is contextual. ♡

***Proof.*** Replacing $\to_{\beta^1}$ in the proof of Lemma 2.4.8 with $\to_\rho$ and making the appropriate corresponding modifications, we can conclude that $\to_\rho$ is weakly contextual. Since $\to_\rho$ is weakly contextual and $\to_{\bar\rho}$ is the linear closure of $\to_\rho$, it follows that $\to_{\bar\rho}$ is contextual. ❏

**Lemma 2.4.16**

Let $u \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential term. If $u \to_{\beta^1} \mu$, then $u \to_\rho \mu$. ♡

***Proof.*** We prove this by induction on $k$.

1. Assume that $u \to_{\beta^1} \mu$ because there exist a simple differential term $s$ and complex differential terms $\tau, \sigma \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = s\tau$, $\mu = \sigma\tau$, and $s \to_{\beta^1} \sigma$. By the induction hypothesis, $s \to_\rho \sigma$. Since $s \to_\rho \sigma$ and $\tau \to_\rho \tau$ (as $\to_\rho$ is reflexive), by Definition 2.4.14(i), $u = (s)\tau \to_\rho \sigma\tau$.

2. Assume that $u \to_{\beta^1} \mu$ because there exist a simple differential term $s \in \Lambda_{d,\mathcal{V}}$ and complex differential terms $\tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = s\tau$, $\mu = s\tau'$, and $\tau \to_{\tilde\beta^1} \tau'$. Then, there

exist a simple differential term $v$, complex differential terms $\nu, \rho$, and scalars $r_1, r_2 \in \mathcal{R}$ such that $\tau = r_1 \cdot v + \rho$, $\tau' = r_2 \cdot \nu + \rho$, and $v \to_{\beta^1} \nu$. By the induction hypothesis, $v \to_\rho \nu$. Since $s \to_\rho s$ (as $\to_\rho$ is reflexive) and $\tau = r_1 \cdot v + \rho \to_{\tilde{\rho}} \cdot \nu + \rho = \tau'$, by Definition 2.4.14(i), $u = s\tau \to_\rho s\tau' = \mu$.

3. Assume that $u \to_{\beta^1} \mu$ because there exist a variable $x \in \mathcal{V}$, a simple differential term $s \in \Lambda_{d,\mathcal{V}}$, and a complex differential term $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = (\lambda x.s)\tau$ and $\mu = s[\tau/x]$. Since $s \to_\rho s$ and $\tau \to_\rho \tau$, by Definition 2.4.14(ii), $u = (\lambda x.s)\tau \to_\rho s[\tau/x] = \mu$.

4. Assume that $u \to_{\beta^1} \mu$ because there exist a positive integer $n \in \mathbb{N}$, a variable $x \in \mathcal{V}$, positive integers $i_1, \ldots, i_n \in \mathbb{N}$, simple differential terms $t_1, \ldots, t_n$, and a complex differential term $\tau_1' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = D_{i_1,\ldots,i_n} x * (t_1, t_2, \ldots, t_n)$ and $\mu = D_{i_1,\ldots,i_n} x * (\tau_1, t_2, \ldots, t_n)$ such that $t_1 \to_{\beta^1} \tau_1$. By the induction hypothesis, $t_1 \to_\rho \tau_1$. Since $t_1 \to_\rho \tau_1$ and $t_j \to_\rho t_j$ for all $2 \leqslant j \leqslant n$, by Definition 2.4.14(iii), $u = D_{i_1,\ldots,i_n} x * (t_1, t_2, \ldots, t_n) \to_\rho D_{i_1,\ldots,i_n} x * (\tau_1, t_2, \ldots, t_n) = \mu$.

5. Assume that $u \to_{\beta^1} \mu$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, a simple differential terms $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, and a complex differential term $\sigma \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = D_1^n \lambda x.s * (t_1, \ldots, t_n)$, $\mu = D_1^n \lambda x.\sigma * (t_1, \ldots, t_n)$, and $s \to_{\beta^1} \sigma$. By the induction hypothesis, $s \to_\rho \sigma$. Since $s \to_\rho \sigma$ and $t_i \to_\rho t_i$ for all $1 \leqslant i \leqslant n$, by Definition 2.4.14(iv), $u = D_1^n \lambda x.s * (t_1, \ldots, t_n) \to_\rho D_1^n \lambda x.\sigma * (t_1, \ldots, t_n) = \mu$.

6. Assume that $u \to_{\beta^1} \mu$ because there exist a positive integer $n \in \mathbb{N}$, a variable $x \in \mathcal{V}$ simple differential terms $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$, and complex differential term $\tau_1 \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = D_1^n \lambda x.s * (t_1, t_2, \ldots, t_n)$, $\mu = D_1^n \lambda x.s * (\tau_1, t_2, \ldots, t_n)$, and $t_1 \to_{\beta^1} \tau_1$. By the induction hypothesis, $t_1 \to_\rho \tau_1$. Since $s \to_\rho s$, $t_1 \to_\rho \tau_1$, and $t_j \to_\rho t_j$ for all $2 \leqslant j \leqslant n$, by Definition 2.4.14(iv), $u = D_1^n \lambda x.s * (t_1, t_2, \ldots, t_n) \to_\rho D_1^n \lambda x.s * (\tau_1, t_2, \ldots, t_n) = \mu$.

7. Assume that $u \to_{\beta^1} \mu$ because there exist a positive integer $n \in \mathbb{N}$, a variable $x \in \mathcal{V}$, and simple differential terms $t_1, \ldots, t_n$, and $n > N$ such that $u = D_1^n \lambda x.s * (t_1, \ldots, t_n)$ and $\mu = D_1^n \lambda x.(\partial_x s * t_j) * (t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n)$. Since $s \to_\rho s$ and $t_i \to_\rho t_i$ for all $1 \leqslant i \leqslant n$, we have that $u = D_1^n \lambda x.s * (t_1, \ldots, t_n) = D_1^n \lambda x.(\partial_x s * t_j) * (t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n) = \mu$.

This completes the induction. ❏

**Corollary 2.4.17**

Let $\mu, \nu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential term. If $\mu \to_{\tilde{\beta}^1} \nu$, then $\mu \to_{\tilde{\rho}} \nu$. ♡

***Proof.*** Since $\mu \to_{\tilde{\beta}^1} \nu$, there exist scalar $r \in \mathcal{R}$, simple differential term $s \in \Lambda_{d,\mathcal{V}}$, and complex differential terms $\sigma, \tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\mu = r \cdot s + \tau$, $\nu = r \cdot \sigma + \tau$, and $s \to_{\beta^1} \sigma$. By Lemma 2.4.16, $s \to_\rho \sigma$. Since $\to_{\tilde{\rho}}$ contains $\to_\rho$, this implies that $s \to_{\tilde{\rho}} \sigma$. Since $\to_{\tilde{\rho}}$ is reflexive, $\tau \to_{\tilde{\rho}} \tau$. Since $\to_{\tilde{\rho}}$ is linear, this implies that $\mu = r \cdot s + \tau = r \cdot s + 1 \cdot \tau \to_{\tilde{\rho}} r \cdot \sigma + 1 \cdot \tau = r \cdot \sigma + \tau = \nu$. ❏

**Lemma 2.4.18**

Let $u \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\mu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential term. If $u \to_\rho \mu$, then $u \twoheadrightarrow_\beta \mu$. ♡

***Proof.*** We prove this by induction on $u \to_\rho \mu$.

1. Assume that $u \to_\rho \mu$ because $u = \mu$. Since $\twoheadrightarrow_\beta$ is reflexive, this implies that $u \twoheadrightarrow_\beta \mu$

2. Assume that $u \to_\rho \mu$ because there exist a simple differential term $s \in \Lambda_{d,\mathcal{V}}$, and complex differential terms $\sigma, \tau, \tau' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = s\tau$, $\mu = (\sigma)\tau'$, $s \to_\rho \sigma$, and $\tau \to_\rho \tau'$. By the induction hypothesis, $s \twoheadrightarrow_\beta \sigma$ and $\tau \twoheadrightarrow_\beta \tau'$. Since $\twoheadrightarrow_\beta$ is contextual,

$$u = s\tau \twoheadrightarrow_\beta \sigma\tau \twoheadrightarrow_\beta \sigma\tau' = \mu. \tag{2.4.25}$$

3. Assume that $u \to_\rho \mu$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, simple differential terms $s, t_1, \dots, t_n \in \Lambda_{d,\mathcal{V}}$, and complex differential terms $\tau, \tau', \sigma, \tau_1, \dots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = (D_1^n \lambda x.s * (t_1, \dots, t_n))\tau$, $\mu = ((\partial_x)^n \sigma * (\tau_1, \dots, \tau_n))[\tau'/x]$, $\tau \to_{\bar\rho} \tau'$, $s \to_\rho \sigma$, and $t_j \to_\rho \tau_j$ for all $1 \leqslant j \leqslant n$. By the induction hypothesis, $\tau \twoheadrightarrow_\beta \tau'$, $s \twoheadrightarrow_\beta \sigma$, and $t_j \twoheadrightarrow_\beta \tau_j$ for all $1 \leqslant j \leqslant n$. Since $\twoheadrightarrow_\beta$ is contextual,

$$\begin{aligned} u &= (D_1^n \lambda x.s * (t_1, \dots, t_n))\tau = (\mathcal{D}_1(\cdots(\mathcal{D}_1 \lambda x.s * t_n)\cdots) * t_n)\tau \\ &\twoheadrightarrow_\beta (\mathcal{D}_1(\cdots(\mathcal{D}_1 \lambda x.s * t_1)\cdots) * t_n)\tau' \twoheadrightarrow_\beta (\mathcal{D}_1(\cdots(\mathcal{D}_1 \lambda x.\sigma * t_n)\cdots) * t_n n)\tau' \\ &\twoheadrightarrow_\beta (\mathcal{D}_1(\cdots(\mathcal{D}_1 \lambda x.\sigma * \tau_1)\cdots) * t_n)\tau' \\ &\twoheadrightarrow_\beta \dots \twoheadrightarrow_\beta (\mathcal{D}_1(\cdots(\mathcal{D}_1 \lambda x.\sigma * \tau_1)\cdots) * \tau_n)\tau' \\ &\to_{\beta^1} \dots \to_{\beta^1} \lambda x.(\partial_x(\cdots(\partial_x \sigma * \tau_1)\cdots) * \tau_n)\tau' = (\lambda x.(\partial_x)^n \sigma * (\tau_1, \dots, \tau_n))\tau' \\ &\to_{\beta^1} ((\partial_x)^n \sigma * (\tau_1, \dots, \tau_n))[\tau'/x] = \mu. \end{aligned} \tag{2.4.26}$$

4. Assume that $u \to_\rho \mu$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, positive integers $i_1, \dots, i_n \in \mathbb{N}$, and simple differential terms $t_1, \dots, t_n \in \Lambda_{d,\mathcal{V}}$, and complex differential terms $\tau_1, \dots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $u = D_{i_1,\dots,i_n} x * (t_1, \dots, t_n)$, $\mu = D_{i_1,\dots,i_n} x * (\tau_1, \dots, \tau_n)$, and $t_j \to_\rho \tau_j$ for all $1 \leqslant j \leqslant n$. By the induction hypothesis, $t_j =\twoheadrightarrow_\beta \tau_j$ for all $1 \leqslant j \leqslant m$. Since $\twoheadrightarrow_\beta$ is contextual,

$$\begin{aligned} u &= D_{i_1,\dots,i_n} x * (t_1, \dots, t_n) = \mathcal{D}_{i_n}(\cdots(\mathcal{D}_{i_1} x * t_1)\cdots) * t_n \\ &\twoheadrightarrow_\beta \dots \twoheadrightarrow_\beta \mathcal{D}_{i_n}(\cdots(\mathcal{D}_{i_1} x * \tau_1)\cdots) * \tau_n = D_{i_1,\dots,i_n} x * (\tau_1, \dots, \tau_n) = \mu. \end{aligned} \tag{2.4.27}$$

5. Assume that $u \to_\rho \mu$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, a nonnegative integer $m \leqslant n$, a variable $x \in \mathcal{V}$, simple differential terms $s, t_1, \dots, t_n, \in \Lambda_{d,\mathcal{V}}$, complex differential terms $\sigma, \tau_1, \dots, \tau_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, and distinct positive integers $j_1, \dots, j_n \in \{1, \dots, n\}$, $u = D_1^n \lambda x.s * (t_1, \dots, t_n)$, $\mu = D^{n-m} \lambda x.((\partial_x^m \sigma * (\tau_{j_1}, \dots, \tau_{j_m})) * (\tau_{j_{m+1}}, \dots, \tau_{j_n})$, $s \to_\rho \sigma$, $t_j \to_\rho \tau_j$ for all $1 \leqslant j \leqslant n$, and $j_1 \leqslant \dots \leqslant j_m$ and $j_{m+1} \leqslant \cdots \leqslant j_n$. By the induction hypothesis, $s \twoheadrightarrow_\beta \sigma$, and $t_i \twoheadrightarrow_\beta \tau_i'$ for all $1 \leqslant i \leqslant m$. Since $\to_\rho$ is contextual,

$$\begin{aligned} u &= D_1^n \lambda x.s * (t_1, \dots, t_n) = \mathcal{D}_1(\cdots(\mathcal{D}_1 \lambda x.s * t_1)\cdots) * t_n \\ &\twoheadrightarrow_\beta \mathcal{D}_1(\cdots(\mathcal{D}_1 \lambda x.\sigma * t_1)\cdots) * t_n \twoheadrightarrow_\beta \dots \twoheadrightarrow_\beta \mathcal{D}_1(\cdots(\mathcal{D}_1 \lambda x.\sigma * \tau_1)\cdots) * \tau_n \\ &= D_1^n \lambda x.\sigma * (\tau_1, \dots, \tau_n) \\ &\to_{\beta^1} \dots \to_{\beta^1} D^{n-m} \lambda x.((\partial_x^m \sigma * (\tau_{j_1}, \dots, \tau_{j_m})) * (\tau_{j_{m+1}}, \dots, \tau_{j_n}) = \mu. \end{aligned} \tag{2.4.28}$$

This completes the induction. ❏

**Corollary 2.4.19**

Let $\mu, \nu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ is a complex differential term. If $\mu \to_{\bar\rho} \nu$, then $\mu \to_\beta \nu$. ♡

***Proof.*** Since $\mu \to_{\bar{\rho}} \nu$, there exists a nonnegative integer $n \in \mathbb{N}_0$, scalars $r_1, \ldots, r_n \in \mathcal{R}$, simple differential terms $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$, and complex differential terms $\sigma_1, \ldots, \sigma_n \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\mu = \sum_{j=1}^{n} r_j \cdot s_j$, $\nu = \sum_{j=1}^{n} r_j \cdot \sigma_j$, and $s_j \to_\rho \sigma_j$ for all $1 \leqslant j \leqslant n$. Then, by Lemma 2.4.18, $s_j \to_\beta \sigma_j$ for all $1 \leqslant j \leqslant n$, so since $\to_\beta$ is linear (by Lemma 2.4.13), this implies that $\mu = \sum_{j=1}^{n} r_j \cdot s_j \to_\beta \sum_{j=1}^{n} r_j \cdot \sigma_j \to \nu$. ❏

**Lemma 2.4.20**

The reflexive and transitive closure of $\to_{\bar{\rho}}$ (i.e., $\twoheadrightarrow_\rho$) is $\twoheadrightarrow_\beta$. ♡

***Proof.*** Take any complex differential term $\sigma, \tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\sigma \twoheadrightarrow_\beta \tau$. Then, there exists complex differential terms $\mu_0, \ldots, \mu_n$ such that $\sigma = \mu_0 \to_{\tilde{\beta}^1} \cdots \to_{\tilde{\beta}^1} \mu_n = \tau$. By Corollary 2.4.17, this implies that $\sigma = \mu_0 \to_{\bar{\rho}} \cdots \to_{\bar{\rho}} \mu_n = \tau$. That is, $\sigma \twoheadrightarrow_{\bar{\rho}} \tau$. Conversely, take any complex differential term $\sigma, \tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\sigma \twoheadrightarrow_{\bar{\rho}} \tau$. Then, there exists complex differential terms $\mu_0, \ldots, \mu_n$ such that $\sigma = \mu_0 \to_{\bar{\rho}} \cdots \to_{\bar{\rho}} \mu_n = \tau$. By Corollary 2.4.19, this implies that $\sigma = \mu_0 \to_\beta \cdots \to_\beta \mu_n = \tau$. That is, $\sigma \twoheadrightarrow_\beta \tau$. ❏

**Lemma 2.4.21**

Let $x \in \mathcal{V}$ be a variable and let $\tau, \tau', \sigma, \sigma' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms such that $\tau \to_{\bar{\rho}} \tau'$ and $\tau \to_{\bar{\rho}} \tau'$. Then,

(i) $\partial_x \tau * \sigma \to_{\bar{\rho}} \partial_x \tau' * \sigma'$ and

(ii) $\tau[\sigma/x] \to_{\bar{\rho}} \tau'[\sigma'/x]$. ♡

***Proof of Part (i).*** This is Lemma 15 in [8]. ❏

***Proof of Part (ii).*** This is Lemma 14 in [8]. ❏

**Lemma 2.4.22**

Let $x \in \mathcal{V}$ be a variable and let $\tau, \tau', \sigma, \sigma' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ are complex differential terms such that $\tau \to_\beta \tau'$ and $\tau \to_\beta \tau'$. Then,

(i) $\partial_x \tau * \sigma \to_\beta \partial_x \tau' * \sigma'$ and

(ii) $\tau[\sigma/x] \to_\beta \tau'[\sigma'/x]$. ♡

***Proof of Part (i).*** The first step is to show that this is true when $\tau = \tau'$, which follows by induction on $\sigma \to_\beta \sigma'$. Note that The reflexive and transitive closure of $\to_{\bar{\rho}}$ is $\to_\beta$ by Lemma 2.4.20.

1. Assume that $\sigma \to_\beta \sigma'$ because of reflexivity. Then, $\sigma = \sigma'$, which implies that $\partial_x \tau * \sigma = \partial_x \tau * \sigma'$, so by the reflexivity of $\to_\beta$, $\partial_x \tau * \sigma \to_\beta \partial_x \tau * \sigma'$.

2. Assume that $\sigma \to_\beta \sigma'$ because of transitivity. Then, there exists a complex differential term $\sigma'' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\sigma \to_\beta \sigma'' \to_\beta \sigma'$. By the induction hypothesis, $\partial_x \tau * \sigma \to_\beta \partial_x \tau * \sigma''$ and $\partial_x \tau * \sigma'' \to_\beta \partial_x \tau * \sigma'$. Hence, by the transitivity of $\to_\beta$, this implies that $\partial_x \tau * \sigma \to_\beta \partial_x \tau * \sigma'$.

3. Assume that $\sigma \to_\beta \sigma'$ because $\sigma \to_{\bar{\rho}} \sigma'$. Since $\to_{\bar{\rho}}$ is reflexive, $\tau \to_{\bar{\rho}} \tau'$ also. Hence, by Lemma 2.4.21(i), $\partial_x \tau * \sigma \to_{\bar{\rho}} \partial_x \tau * \sigma'$. Since $\to_{\bar{\rho}}$ is contained in $\to_\beta$, this implies that $\partial_x \tau * \sigma \to_\beta \partial_x \tau * \sigma'$.

This completes the induction. Now, we can prove the full statement of Lemma 2.4.22(i) via induction on $\tau \to_\beta \tau'$.

1. Assume that $\tau \to_\beta \tau'$ because of reflexivity. Then, $\tau = \tau'$, which implies that $\partial_x \tau * \sigma \to_\beta \partial_x \tau * \sigma' = \partial_x \tau' * \sigma'$.

2. Assume that $\tau \to_\beta \tau'$ because of transitivity. Then, there exists a complex differential term $\tau'' \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\tau \to_\beta \tau'' \to_\beta \tau'$. By the induction hypothesis, $\partial_x \tau * \sigma \to_\beta \partial_x \tau'' * \sigma'$ and $\partial_x \tau'' * \sigma' \to_\beta \partial_x \tau' * \sigma'$. Hence, by the transitivity of $\to_\beta$, this implies that $\partial_x \tau * \sigma \to_\beta \partial_x \tau' * \sigma'$.

3. Assume that $\tau \to_\beta \tau'$ because $\tau \to_{\bar{\rho}} \tau'$. Since $\to_{\bar{\rho}}$ is reflexive, $\sigma \to_{\bar{\rho}} \sigma$. Hence, by Lemma 2.4.21(i), $\partial_x \tau * \sigma \to_{\bar{\rho}} \partial_x \tau' * \sigma$. Since $\to_{\bar{\rho}}$ is contained in $\to_\beta$, this implies that $\partial_x \tau * \sigma \to_\beta \partial_x \tau' * \sigma$. Also, since $\sigma \to_\beta \sigma'$, $\partial_x \tau' * \sigma' \to_\beta \partial_x \tau' * \sigma'$. By the transitivity of $\to_\beta$, this implies that $\partial_x \tau * \sigma \to_\beta \partial_x \tau' * \sigma'$.

This completes the induction. ❏

***Proof of Part (ii)***. The proof is very similar to the proof of Lemma 2.4.22(i). ❏

**Definition 2.4.23**

Let $\to_R$ be a relation from simple (resp. complex) differential terms to complex differential terms and $\to_S$ be a relation from complex differential terms to complex differential terms. Given any simple (resp. complex) differential term $s \in \Lambda_{d,\mathcal{V}}$ (resp. $\in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$), positive integer $n \in \mathbb{N}$, and complex differential terms $\tau_1, \ldots, \tau_n \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $s \to_R \tau_j$ for all $1 \leqslant j \leqslant n$, suppose that there exists a corresponding complex differential term $\tau \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\tau_j \to_S \tau$ for all $1 \leqslant j \leqslant n$. Then, the pair $(\to_R, \to_S)$ is called *multi-confluent*. ♣

**Lemma 2.4.24**

Let $\to_R$ be a relation from simple differential terms to complex differential terms and $\to_S$ be a linear relation from complex differential terms to complex differential terms such that $(\to_R, \to_S)$ is multi-confluent. Then, $(\to_{\bar{R}}, \to_S)$ is also multi-confluent. ♡

***Proof***. Take any positive integer $n \in \mathbb{N}$ and complex differential terms $\sigma, \tau_1, \ldots, \tau_n \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\sigma \to_{\bar{R}} \tau_j$ for all $1 \leqslant j \leqslant n$. By Lemma 2.2.29(xii), there exists a unique nonnegative integer $n \in \mathbb{N}_0$, unique scalars $r_1, \ldots, r_m \in \mathcal{R}$, and unique simple differential terms $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$ such that $\sigma = \sum_{i=1}^m r_i \cdot s_i$. Since $\sigma \to_{\bar{R}} \tau_j$ for all $1 \leqslant j \leqslant n$, this implies that for each $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$, there exist a nonnegative integer $\ell_{i,j} \in \mathbb{N}_0$, scalars $r_{i,j,1}, \ldots, r_{i,j,\ell_{i,j}} \in \mathcal{R}$ such that $r_i = \sum_{k=1}^{\ell_{i,j}} r_{i,j,k}$ and complex differential terms $\sigma_{i,j,1}, \ldots, \sigma i, j, n \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $s_i \to_R \sigma_{i,j,k}$ and $\tau_j = \sum_{i=1}^m \sum_{k=1}^{\ell_{i,j}} r_{i,j,k} \cdot \sigma_{i,j,k}$. Take any $1 \leqslant i \leqslant m$. Since $(\to_R, \to_S)$ is *multi-confluent* and $s_i \to_R \sigma_{i,j,k}$ for all $1 \leqslant j \leqslant n$ and $1 \leqslant k \leqslant \ell_{i,j}$, there exists a complex differential term $\sigma_i \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\sigma_{i,j,k} \to_S \sigma_i$ for all $1 \leqslant j \leqslant n$ and $1 \leqslant k \leqslant \ell_{i,j}$. Let $\tau = \sum_{i=1}^m r_i \cdot \sigma_i$. Because $\to_S$ is linear, this implies that

$$\tau_j = \sum_{i=1}^m \sum_{k=1}^{\ell_{i,j}} (r_{i,j,k} \cdot \sigma_{i,j,k}) \to_S \sum_{i=1}^m \sum_{k=1}^{\ell_{i,j}} (r_{i,j,k} \cdot \sigma_i) = \sum_{i=1}^m \left( \sum_{k=1}^{\ell_{i,j}} r_{i,j,k} \right) \cdot \sigma_i = \sum_{i=1}^m r_i \cdot \sigma_i = \tau \qquad (2.4.29)$$

for each $1 \leqslant j \leqslant n$. Hence, $(\to_{\bar{R}}, \to_S)$ is multi-confluent. ❏

**Lemma 2.4.25**

$(\to_\rho, \to_{\bar\rho})$ is multi-confluent. That is, given a simple differential term $s \in \Lambda_{d,\mathcal{V}}$, a positive integer $n \in \mathbb{N}$, and complex differential terms $\mu_1, \ldots, \mu_n \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $u \to_\rho \tau_j$ for all $1 \leqslant j \leqslant n$, there exists a corresponding complex differential term $\mu \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\tau_j \to_{\bar\rho} \tau$ for all $1 \leqslant j \leqslant n$. ♡

***Proof.*** This follows by induction on the height of complex differential term $u \in \Lambda_{d,\mathcal{V}}$.

1. Assume that $u = D_{i_1,\ldots,i_m} x * (t_1, \ldots, t_m)$, where $m \in \mathbb{N}_0$ is a nonnegative integer, $x \in \mathcal{V}$ is a variable, and $t_1, \ldots, t_m \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Then, for each $1 \leqslant j \leqslant m$, since $u \to_\rho \mu_j$, by the definition of $\to_\rho$, there exists complex differential terms $\tau_{j,1}, \ldots, \tau_{j,n} \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\mu_j = D_{i_1,\ldots,i_m} x * (\tau_{j,1}, \ldots, \tau_{j,m})$ and $t_i \to_\rho \tau_{j,i}$ for all $1 \leqslant i \leqslant m$. Now, take any $1 \leqslant i \leqslant m$. Then, $t_i \to_\rho \tau_{j,i}$ for all $1 \leqslant j \leqslant n$, so by the induction hypothesis, there exists a complex differential term $\tau_i \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\tau_{j,i} \to_{\bar\rho} \tau_i$. Define the complex differential term $\mu = D_{i_1,\ldots,i_m} x * (\tau_1, \ldots, \tau_m) \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$. Then, for each $1 \leqslant j \leqslant n$, this implies that $\mu_j = D_{i_1,\ldots,i_m} x * (\tau_{j,1}, \ldots, \tau_{j,m}) \to_{\bar\rho} D_{i_1,\ldots,i_m} x * (\tau_1, \ldots, \tau_m) = \mu$.

2. Assume that $u = D_1^m \lambda x.s * (t_1, \ldots, t_m)$, where $m \in \mathbb{N}_0$ is a nonnegative integer, $x \in \mathcal{V}$ is a variable, and $s, t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms. Then, for each $1 \leqslant j \leqslant n$, since $u \to_\rho \mu_j$, there exist complex differential terms $\sigma_j, \tau_{j,1}, \ldots, \tau_{j,m} \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $s \to_\rho \sigma_j$ and $t_i \to_\rho \tau_{j,i}$ for all $1 \leqslant i \leqslant m$, a nonnegative integer $0 \leqslant \ell_j \leqslant m$, and distinct positive integers $k_{j,1}, \ldots, k_{j,\ell_j} \in 1, \ldots, m$ such that $k_{j,1} < \ldots < k_{j,\ell_j}$, $k_{j,\ell_j+1} < \ldots < k_{j,m}$, $u = D_1^m \lambda x.s * (t_1, \ldots, t_m)$, and $\mu_j = D^{m-\ell_j} \lambda x.((\partial_x^{\ell_j} \sigma_j * (\tau_{k_{j,1}}, \ldots, \tau_{k_{j,\ell_j}})) * (\tau_{k_{j,\ell_j+1}}, \ldots, \tau_{k_{j,m}})$. Since $s \to_\rho \sigma_j$ for all $1 \leqslant j \leqslant n$, by the induction hypothesis, there exists a complex differential term $\sigma \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\sigma_j \to_{\bar\rho} \sigma$ for all $1 \leqslant j \leqslant n$. For each $1 \leqslant i \leqslant m$, since $t_i \to_\rho \tau_{j,i}$ for all $1 \leqslant j \leqslant n$, by the induction hypothesis, there exists a complex differential term $\tau_i \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\tau_{j,i} \to_{\bar\rho} \tau_i$ for all $1 \leqslant j \leqslant n$. Define the complex differential term $\mu = \lambda x.\partial_x^m \sigma * (\tau_1, \ldots, \tau_n)$. Then, for each $1 \leqslant j \leqslant n$, this implies that $\mu_j = D^{m-\ell_j} \lambda x.((\partial_x^{\ell_j} \sigma_j * (\tau_{k_{j,1}}, \ldots, \tau_{k_{j,\ell_j}})) * (\tau_{k_{j,\ell_j+1}}, \ldots, \tau_{k_{j,m}}) \to_{\bar\rho} \lambda x.\partial_x^m \sigma * (\tau_1, \ldots, \tau_n) = \mu$.

3. Assume that $u = s\tau$, where $s \in \Lambda_{d,\mathcal{V}}$ is a simple differential term and $\tau \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ is a complex differential term.

   (a) Suppose that for each $1 \leqslant j \leqslant n$, there exist complex differential terms $\sigma_j, \tau_j \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\mu_j = \sigma_j \tau_j'$. Then, for each $1 \leqslant j \leqslant n$, since $s\tau = u \to_\rho \mu_n = \sigma_j \tau_j'$, it must be the case that $s \to_\rho \sigma_j$ and $\tau \to_{\bar\rho} \tau_j'$. Since $s \to_\rho \sigma_j$ for all $1 \leqslant j \leqslant n$, by the induction hypothesis, there exists a complex differential term $\sigma \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\sigma_j \to_{\bar\rho} \sigma$. Since $\tau \to_{\bar\rho} \tau_j$ for all $1 \leqslant j \leqslant n$, by the induction hypothesis and by Lemma 2.4.24, there exists a complex differential term $\tau' \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$ such that $\tau_j' \to_{\bar\rho} \tau'$. Define the complex differential term $\mu = \sigma\tau \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$. Then, for each $1 \leqslant j \leqslant n$, $\mu_j = \sigma_j \tau_j' \to_{\bar\rho} \sigma\tau' = \mu$.

   (b) Suppose otherwise. Then, there exists a nonnegative integer $m \in \mathbb{N}_0$, a variable $x \in \mathcal{V}$, complex differential terms $v_1, \ldots, v_n, \omega_{1,1}, \ldots, \omega_{n,1}, \ldots, \omega_{1,m}, \ldots, \omega_{n,m}, \tau_1', \ldots, \tau_n' \in \mathcal{R}\langle \Lambda_{d,\mathcal{V}} \rangle$, simple differential terms $v, w_1, \ldots, w_n \in \Lambda_{d,\mathcal{V}}$, positive integer $2 \leqslant q \leqslant n$, integers $0 \leqslant \ell_q, \ldots, \ell_m \leqslant m$,

      i. $s = D_1^m \lambda x.v * (w_1, \ldots, w_m)$,

      ii. $\mu_j = (\partial_x^m \nu_j * (\omega_{j,1}, \ldots, \omega_{j,n}))[\tau_j'/x]$ for $1 \leqslant j < q$,

      iii. $\mu_j = (D^{m-\ell_j} \lambda x.((\partial_x^{\ell_j} \nu_j * (\omega_{k_{j,1}}, \ldots, \omega_{k_{j,\ell_j}})) * (\omega_{k_{j,\ell_j+1}}, \ldots, \omega_{k_{j,m}}))\tau_j'$ for $q \leqslant j \leqslant n$,

iv. $v \to_\rho \nu_j$ for $1 \leqslant j \leqslant n$,

v. $w_i \to_\rho \omega_{j,i}$ for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant n$, and

vi. $\tau \to_{\bar{\rho}} \tau'_j$ for $1 \leqslant j \leqslant n$,

where for each $q \leqslant j \leqslant n$, $k_{j,1}, \ldots, k_{j,m}$ are distinct integers ranging from 1 to $m$ such that $k_{j,1} < \ldots < k_{j,\ell_j}$ and $k_{j,\ell_j+1} < \ldots < k_{j,m}$. Since $v \to_\rho \nu_j$ for $1 \leqslant j \leqslant n$, by the induction hypothesis, there exists a complex differential term $\nu \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\nu_j \to_\rho \nu$. Since $w_i \to_\rho \omega_{j,i}$ for all $1 \leqslant j \leqslant n$ and $1 \leqslant i \leqslant n$, by the induction hypothesis, there exists a complex differential term $\omega_i \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\omega_{j,i} \to_\rho \omega_i$ for each $1 \leqslant i \leqslant n$. Since $\tau \to_\rho \tau'_j$ for $1 \leqslant j \leqslant n$, by the induction hypothesis and Lemma 2.4.24, there exists a complex differential term $\tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $\tau'_j \to_\rho \tau'$. Let $\mu = (\partial_x^m \nu * (\omega_1, \ldots, \omega_n))[\tau'/x]$. Then, using Lemma 2.4.21, we can see that for each $1 \leqslant j < q$, $\mu_j = \partial_x^m \nu_j * (\omega_{j,1}, \ldots, \omega_{j,n}))[\tau'_j/x] \to_{\bar{\rho}} (\partial_x^m \nu * (\omega_1, \ldots, \omega_n))[\tau'/x] = \mu$, and for each each $q \leqslant j \leqslant n$, $\mu_j = (D^{m-\ell_j}\lambda x.((\partial_x^{\ell_j} \nu_j * (\omega_{k_{j,1}}, \ldots, \omega_{k_{j,\ell_j}})) * (\omega_{k_{j,\ell_j+1}}, \ldots, \omega_{k_{j,m}}))\tau'_j \to_{\bar{\rho}} (\partial_x^m \nu * (\omega_1, \ldots, \omega_n))[\tau'/x] = \mu$.

This completes the induction. ❏

**Lemma 2.4.26**

The relation $\to_{\bar{\rho}}$ on complex differential terms is confluent. ♡

***Proof*.** By Lemma 2.4.25, $(\to_\rho, \to_{\bar{\rho}})$ is multi-confluent, so by Lemma 2.4.24, $(\to_{\bar{\rho}}, \to_{\bar{\rho}})$, which means that $\to_{\bar{\rho}}$ is confluent. ❏

**Lemma 2.4.27**

Let $\to_R$ be a confluent relation on a set $S$, then $\twoheadrightarrow_R$ satisfies the Church-Rosser property. That is, given $A, B, C \in S$ such that $C \twoheadrightarrow_R A$ and $C \twoheadrightarrow_R B$, there exists $D \in S$ such that $A \twoheadrightarrow_R D$ and $B \twoheadrightarrow_R D$. ♡

***Proof*.** Since $C \twoheadrightarrow_R A$ and $C \twoheadrightarrow_R B$, there exists $m, n \in \mathbb{N}_0$ and $E_{0,0}, E_{1,0}, \ldots, E_{m,0}, E_{0,1}, \ldots, E_{0,n} \in S$ such that $C = E_{0,0} \to_R \cdots \to_R E_{m,0} = A$ and $C = E_{0,0} \to_R \cdots \to_R E_{0,n} = B$. For each $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$, we want to define inductively an element $E_{i,j}$ in $S$ such that $E_{i-1,j} \to_R E_{i,j}$ and $E_{i,j-1} \to_R E_{i,j}$, so without further ado, take any $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. Since $E_{i-1,j-1} \to_R E_{i,j-1}$, $E_{i-1,j-1} \to_R E_{j-1,i}$, and $\to_R$, we know that there exists $F \in S$ such that $E_{j-1,i} \to F$ and $E_{j,i-1} \to F$, so naturally, we choose to set $E_{i,j} = F$. Let $D = E_{m,n}$. Then, $A = E_{m,0} \to_R \cdots \to_R E_{m,n-1} \to_R E_{m,n} = D$ and $B = E_{0,n} \to_R \cdots \to_R E_{m-1,n} \to_R E_{m,n} = D$. ❏

**Theorem 2.4.28. The Church-Rosser Property of Differential $\beta$-Reduction**

The relation $\to_\beta$ satisfies the Church-Rosser property. ♡

***Proof*.** By Lemma 2.4.26, $\to_{\bar{\rho}}$ is confluent, so by Lemma 2.4.27, $\twoheadrightarrow_{\bar{\rho}}$ satisfies the Church-Rosser property. By Lemma 2.4.20, the relation $\twoheadrightarrow_{\bar{\rho}}$ is equivalent to $\twoheadrightarrow_\beta$. Hence, $\twoheadrightarrow_\beta$ also satisfies the Church-Rosser property. ❏

# Chapter 3

# Taylor $\lambda$-Calculus

In the prior chapter, we carefully constructed differential $\lambda$-calculus, motivated its syntax as a synthetic framework to represent the differentiation of analytic function, proved a number of lemmas, and showed that computations in differential $\lambda$-calculus is well-behaved by showing that differential $\beta$-reduction satisfies the Church-Rosser property. In this chapter, we are going to define another formulation of a purely syntactic framework to treat differentiation of analytic functions. This formulation is called Taylor $\lambda$-calculus, and just like its cousin differential $\lambda$-calculus, it was proposed by Thomas Ehrhard and Laurent Regnier's (see [9]). The first section will motivate the syntax of Taylor $\lambda$-calculus and relate it to differential $\lambda$-calculus. The second section will introduce multisets and operations on multisets and prove some elementary properties of these operations. The third section will rigorously define simple and complex Taylor terms and poly-terms. Then, in the third section we will define partial differentiation of Taylor terms and show that partial differentiation behaves like a nondeterministic substitution. Finally, in the fourth section, we will show that Taylor $\lambda$-calculus is a subset of differential $\lambda$-calculus.

## 3.1 Motivating the Syntax of Taylor $\lambda$-Calculus

Before we give a general description of the syntax of Taylor $\lambda$-calculus, we need to ask a key question: why do we need a new variant of $\lambda$-calculus, and more specifically, why do need a second variant of a $\lambda$-calculus for expressing analytic functions, the first being differential $\lambda$-calculus. Differential $\lambda$-calculus, which was discussed at length in chapter 2, is already a decent framework for dealing with the differentiation of analytic functions syntactically. However, there are two incredibly important properties that are missing from differential $\lambda$-calculus: first, the ability to directly express and work with polynomials, and second, linearity of every structure in the calculus. Taylor $\lambda$-calculus aims to be fix this, and this is the reason we are interested in investigating it.

So, to get our discussion on the syntax of Taylor $\lambda$-calculus started, we note that, just as with ordinary differential $\lambda$-calculus, a term in Taylor $\lambda$-term represents an analytic function and we have two obvious types of Taylor terms:

(i) a variable and

(ii) an abstraction $\lambda x.s$, where $x$ is a variable and $s$ is a Taylor term.

We address, now, the first property we want to see in Taylor $\lambda$-calculus: the ability to express polynomials. To do this we really should be able to express finite products of analytic functions. Given a nonnegative integer $n \in \mathbb{N}_0$ and Taylor terms $f_1, \ldots, f_n$, representing analytic functions of type $\mathbb{R}^m \to \mathbb{R}$, we introduce the Taylor term $f_1 \dots f_n$, which is the multiset whose elements are $f_1, \dots, f_n$. The Taylor term $f_1 \dots f_n$ represents a function of type $\mathbb{R}^m \to \mathbb{R}$ and is given by:

$$x_1, \dots, x_n \mapsto \prod_{i=1}^{n} f_i(x_1, \dots, x_n). \tag{3.1.1}$$

With the ability to write polynomials in Taylor $\lambda$-calculus, we now turn to the second property we want to see in Taylor $\lambda$-calculus: the linearity of every structure in the calculus. Recall that in differential $\lambda$-calculus, we have that

$$\begin{aligned} &D_{i_1,\dots,i_n} x * \left( \sum_{j_1=1}^{m_1} r_{1,j_1} \cdot t_{1,j_1}, \dots, \sum_{j_n=1}^{m_n} r_{n,j_n} \cdot t_{n,j_n} \right) \\ &\qquad = \sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} (r_{1,j_1} \cdots r_{n,j_n}) \cdot D_{i_1,\dots,i_n} x * (t_{1,j_1}, \dots, t_{n,j_n}), \end{aligned} \tag{3.1.2a}$$

$$\begin{aligned} &D_1^n \lambda x. \left( \sum_{j_0=1}^{m_0} r_{0,j_0} \cdot s_{j_0} \right) * \left( \sum_{j_1=1}^{m_1} r_{1,j_1} \cdot t_{1,j_1}, \dots, \sum_{j_n=1}^{m_n} r_{n,j_n} \cdot t_{n,j_n} \right) \\ &\qquad = \sum_{j_0=1}^{m_0} \sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} (r_{0,j_0} r_{1,j_1} \cdots r_{n,j_n}) \cdot D_1^n \lambda x. s_{j_0} * (t_{1,j_1}, \dots, t_{n,j_n}), \end{aligned} \tag{3.1.2b}$$

where $n, m_0, \dots, m_n \in \mathbb{N}_0$ are nonnegative integers, $r_{0,1}, \dots, r_{0,m_0}, \dots, r_{n,1}, \dots, r_{n,m_n} \in \mathcal{R}$ are scalars, $s_{0,1}, \dots, s_{0,m_0}, t_{0,1}, \dots, t_{0,m_0}, \dots, t_{n,1}, \dots, t_{n,m_n} \in \Lambda_{d,\mathcal{V}}$ are simple differential terms, and $x \in \mathcal{V}$ is a variable. Hence, a complex differentiated variable and a complex differentiated abstraction is linear in all their arguments. However, notice that

$$f(x + y) \neq fx + fy \tag{3.1.3}$$

in differential $\lambda$-calculus, where $x, y, f \in \mathcal{V}$ are variables. Hence, application in differential $\lambda$-calculus is not linear in all its arguments. The underlying reason for this is clear: analytic functions are not necessarily linear. Consequently, if we want every structure in Taylor $\lambda$-calculus to be linear in all its arguments, we cannot have application. Instead, given a nonnegative number $\mathbb{N}_0$ and Taylor term $f, t_1, \dots, t_n$, representing analytic functions of type $\mathbb{R}^m \to \mathbb{R}$, we introduce the Taylor term $\langle f \rangle t_1 \dots t_n$, which is a constant function of with the same type as $f$ and is given by:

$$x_1, \dots, x_m \mapsto \cdots (\partial_{x_{i_n}} \cdot (\partial_{x_{i_2}} ((\partial_{x_{i_1}} f) \cdot t_1) \cdot t_2) \cdots) \cdot t_n)(0, \dots, 0), \tag{3.1.4}$$

where $i_1, i_2, \dots, i_n \in \{1, \dots, m\}$ are some not necessarily distinct integers. Intuitively, $\langle f \rangle t_1 \dots t_n$ is linear in its $n + 1$ arguments. This is our notion of application in Taylor $\lambda$-calculus.

Also, just like differential $\lambda$-calculus, Taylor $\lambda$-calculus should have simpler Taylor terms that are linear combinations of Taylor terms. In summary, a Taylor term is one of the following:

(i) a variable;

(ii) an abstraction $\lambda x.s$, where $x$ is a variable and $s$ is a Taylor term;

(iii) a *product* (or *poly-term*) $t_1 \dots t_n$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $t_1, \dots, t_n$ are Taylor terms.

(iv) a *differentiated application* $\langle s \rangle T$, where $s$ is a Taylor Term and $T$ is a poly-term.

(v) a *scalar multiplication* $r \cdot s$, where $r$ is a scalar and $s$ is a Taylor term;

(vi) a *sum* $s + t$, where $s, t$ are Taylor terms; or

(vii) 0, which represent the zero functions.

These seven different kinds of differential terms will make up our calculus, which will be called Taylor $\lambda$-calculus. With this intuition at hand, let us make this notion of a differential term much more precise.

## 3.2 Working with Multisets

One of the key difference between Taylor $\lambda$-calculus and differential $\lambda$-calculus, is the presence of poly-terms in Taylor $\lambda$-calculus, which are just mutlisets of Taylor terms, so to treat prcicely Taylor $\lambda$-calculus, we need to work a bit on multisets.

**Definition 3.2.1. Multiset**

Given a set $I$, the set $\mathcal{M}_{\text{fin}}(I)$ of *multisets* over $I$ is the free commutative monoid over $I$. ♣

There are a handful of important operators on multisets.

**Notation 3.2.2. Operations on Multisets**

Let $I$ be some set, $i_0$ be a distinguished element in $I$, $S, T \in \mathcal{M}_{\text{fin}}(I)$ be multisets over $I$, $R$ is a commutative monoid, and $x \in R^I$ be an $I$-indexed family of elements in $R$.

1. The *exponential* of $x$ be $S$ is $x^S \coloneqq \prod_{i \in I}[x(i)]^{S(i)} \in R$.

2. The *cardinality* of $S$ is $|S| \coloneqq \sum_{i \in I} S(i) \in \mathbb{N}_0$.

3. The *factorial* of $S$ is $S! \coloneqq \prod_{i \in I} S(i) \in \mathbb{N}_0$

4. The *multinational coefficient* of $S$ is $[S] \coloneqq \frac{|S|!}{S!} \in \mathbb{N}_0$.

5. The *product* of $S$ and $T$ is the multiset $ST \in \mathcal{M}_{\text{fin}}(I)$ such that $(ST)(i) = S(i) + T(i)$ also.

6. $T$ is *less than (or equal to)* $S$ if and only if $T(i)$ is less than (or equal to) $S(i)$ for all $i \in I$.

7. If $T \leqslant S$, the *binomial coefficient* of $S$ and $T$ is $\binom{S}{T} = \frac{S!}{}\left(_{T!(S-T)!}\right)$.

8. If $S(i_0) \geqslant 1$, then $S - i_0 \in \mathcal{M}_{\text{fin}}(I)$ is the multiset such that $(S - i_0)(i_0) = S(i_0) - 1$ and $(S - i_0)(i) = S(i)$ for all $i \in I - \{i\}$ ♠

With a notion of the cardinality of a multiset, we can and should talk about the set of multisets with a specific cardinality.

**Definition 3.2.3. The Set of Multiset of Specific Height**

Given a set $I$ and a nonnegative integer $n \in \mathbb{N}_0$,

$$\mathcal{M}_n(I) \coloneqq \{S \in \mathcal{M}_{\text{fin}}(I) : |S| = n\} \tag{3.2.1a}$$

$$\mathcal{M}_{\leqslant n}(I) \coloneqq \{S \in \mathcal{M}_{\text{fin}}(I) : |S| \leqslant n\} \tag{3.2.1b}$$
$$\mathcal{M}_{\geqslant n} I \coloneqq \{S \in \mathcal{M}_{\text{fin}}(I) : |S| \geqslant n\}. \tag{3.2.1c}$$

♣

In the next few lemmas, we will see that the operators on multisets lift many of the properties of natural numbers, including the Binomial Theorem and Pascal's Theorem.

**Lemma 3.2.4**

Let $I$ be some set and let $S, T \in \mathcal{M}_{\text{fin}}(I)$ be multisets over $I$. Then,

$$\binom{S}{T} = \prod_{i \in I} \binom{S(i)}{T(i)}. \tag{3.2.2}$$

♡

***Proof.*** As we can see,

$$\begin{aligned}\binom{S}{T} &= \frac{S!}{T!(S-T)!} = \frac{\prod_{i\in I} S(u)!}{\prod_{i\in I} T(i)! \cdot \prod_{i\in I}(S(i)-T(i))!} \\ &= \prod_{i\in I} \frac{S(i)!}{[T(i)!(S(i)-T(i))!} = \prod_{i\in I}\binom{S(i)}{T(i)}.\end{aligned} \tag{3.2.3}$$

This completes the proof. ❏

**Lemma 3.2.5**

Let $I$ be some set and let $S, T \in \mathcal{M}_{\text{fin}}(I)$ be multisets over $I$. Suppose that $T \leqslant S$ Then,

$$\binom{S}{T} = \binom{S}{S-T}. \tag{3.2.4}$$

♡

***Proof.*** Take any $i \in I$. Since $T \leqslant S$, $T(i) \leqslant S(i)$, so

$$\binom{S(i)}{T(i)} = \binom{S(i)}{S(i)-T(i)} \tag{3.2.5}$$

Since our choice of $i \in I$ was arbitrary, this implies that

$$\binom{S}{T} = \prod_{i\in I}\binom{S(i)}{T(i)} = \prod_{i\in I}\binom{S(i)}{S(i)-T(i)} = \prod_{i\in I}\binom{S(i)}{(S-T)(i)} = \binom{S}{S-T}, \tag{3.2.6}$$

where we used Lemma 3.2.4 for the first and last equalities. ❏

**Lemma 3.2.6. Binomial Theorem for Multisets**

Let $I$ be some set, $R$ be a commutative semiring, $x, y \in R^I$ be $I$-indexed families of elements in $R$, and $S \in \mathcal{M}_{\text{fin}}(I)$ be a multiset over $I$. Then,

$$(x+y)^S = \sum_{T \leqslant S} \binom{S}{T} x^T y^{S-T}. \tag{3.2.7}$$

♡

***Proof.*** As we can see,

$$
\begin{aligned}
(x+y)^S &= \prod_{u\in I}[(x+y)(i)]^S(i) = \prod_{i\in I}[x(i)+y(i)]^S(i) \\
&= \prod_{i\in I}\sum_{\substack{k_1+k_2=S(i)\\ k_1,k_2>=0}}\binom{S(i)}{k_1,k_2}[x(u)]^{k_1}[y(u)]^{k_2} = \prod_{i\in I}\sum_{k=0}^{S(i)}\binom{S(i)}{k}[x(i)]^k[y(i)]^{S(i)-k} \\
&= \sum_{T\leqslant S}\prod_{i\in I}\binom{S(i)}{T(i)}[x(i)]^{T(i)}[y(u)]^{S(i)-T(i)} \\
&= \sum_{T\leqslant S}\prod_{i\in I}\binom{S(i)}{T(i)}\cdot\prod_{i\in I}[x(i)]^{T(i)}\cdot\prod_{i\in I}[y(i)]^{(S-T)(i)} = \sum_{T\leqslant S}\binom{S}{T}x^Ty^{S-T},
\end{aligned}
\tag{3.2.8}
$$

where we used Lemma 3.2.4 for the last equality. ❏

**Lemma 3.2.7**

Let $I$ be some set, $i_0$ be a distinguished element in $I$, and $S \in \mathcal{M}_{\text{fin}}(I)$ be a multiset over $I$. Suppose that $S(i_0) \geqslant 1$. Then,

$$\binom{S}{S-i_0} = S(i_0). \tag{3.2.9}$$

♡

***Proof.*** Since $S(i_0) \geqslant 1$, we know that

$$\binom{S(i_0)}{S(i_0)-1} = S(i_0). \tag{3.2.10}$$

Hence,

$$
\begin{aligned}
\binom{S}{S-i_0} &= \prod_{i\in I}\binom{S(i)}{(S-i_0)(i)} = \binom{S(i_0)}{(S-i_0)(i_0)}\cdot\prod_{i\in I-\{i_0\}}\binom{S(i)}{(S-i_0)(i)} \\
&= \binom{S(i)}{S(i)-1}\cdot\prod_{i\in I-\{i\}}\binom{S(i)}{S(i)} = S(i_0)\cdot\prod_{i\in I-\{i_0\}}1 = S(i_0)\cdot 1 = S(i_0).
\end{aligned}
\tag{3.2.11}
$$

where we used Lemma 3.2.4 for the first equality. ❏

**Lemma 3.2.8. Pascal's formula for Multisets**

Let $I$ be some set, $i_0$ be a distinguished element in $I$, and $S, T \in \mathcal{M}_{\text{fin}}(I)$ be multisets over $I$. Suppose that $T \leqslant S$ and $S(i_0) > T(i_0) \geqslant 1$. Then,

$$\binom{S}{T} = \binom{S-i_0}{T} + \binom{S-i_0}{T-i_0} \tag{3.2.12}$$

♡

***Proof.*** Since $S(i_0) > T(i_0) \geqslant 1$, by the ordinary Pascal's formula,

$$\binom{S(i_0)}{T(i_0)} = \binom{S(i_0)-1}{T} + \binom{S(i_0)-1}{T(i_0)-1}. \tag{3.2.13}$$

Hence,

$$\begin{aligned}
\binom{S}{T} &= \prod_{i\in I}\binom{S(i)}{T(i)} = \binom{S(i_0)}{T(i_0)}\cdot\prod_{i\in I-\{i_0\}}\binom{S(i)}{T(i)} \\
&= \left[\binom{S(i_0)-1}{T(i_0)} + \binom{S(i_0)-1}{T(i_0)-1}\right]\cdot\prod_{i\in I-\{i_0\}}\binom{S(i)}{T(i)} \\
&= \left[\binom{(S-i_0)(i_0)}{T(i_0)} + \binom{(S-i_0)(i_0)}{(T-i_0)(i_0)}\right]\cdot\prod_{i\in I-\{i_0\}}\binom{S(i)}{T(i)} \\
&= \binom{(S-i_0)(i_0)}{T(i_0)}\cdot\prod_{i\in I-\{i_0\}}\binom{S(i)}{T(i)} + \binom{(S-i_0)(i_0)}{(T-i_0)(i_0)}\cdot\prod_{i\in I-\{i_0\}}\binom{S(i)}{T(i)} \\
&= \binom{(S-i_0)(i_0)}{T(i_0)}\cdot\prod_{i\in I-\{i_0\}}\binom{(S-i_0)(i)}{T(i)} + \binom{(S-i_0)(i_0)}{(T-i_0)(i_0)}\cdot\prod_{i\in I-\{i_0\}}\binom{(S-i_0)(i)}{(T-i_0)(i)} \\
&= \prod_{i\in I}\binom{(S-i_0)(i)}{T(i)} + \prod_{i\in I}\binom{(S-i_0)(i)}{(T-i_0)(i)} = \binom{S-i_0}{T} + \binom{S-i_0}{U-i_0},
\end{aligned} \tag{3.2.14}$$

where we used Lemma 3.2.4 for the first and last equalities. ❏

**Lemma 3.2.9**

Let $I$ be some set, $i_0$ be a distinguished element in $I$, and $T, U, V \in \mathcal{M}_{\text{fin}}$ be multisets over $I$. Suppose that $UV = T$, $i_0 \in U$, and $i_0 \notin V$. Then,

$$\binom{T-i_0}{U-i_0} = \binom{T}{U}. \tag{3.2.15}$$

♡

***Proof.*** Since $i_0 \notin V$, $V(i_0) = 0$, and since $UV = T$, $T(i_0) = U(i_0) + V(i_0) = U(i_0) + 0 = U(i_0)$, which implies that $(T - i_0)(i_0) = T(i_0) - 1 = U(i_0) - 1 = (U - i_0)(i_0)$. Since $T(i_0) = U(i_0)$ and $(T - i_0)(i_0) = (U - i_0)(i_0)$, we know that

$$\binom{T(i_0)}{U(i_0)} = \binom{(T-i_0)(i_0)}{(U-i_0)(i_0)} = 1, \tag{3.2.16}$$

so

$$\begin{aligned}
\binom{T-i_0}{U-i_0} &= \prod_{i\in I}\binom{(T-i_0)(i)}{(U-i_0)(i)} = \binom{(T-i_0)(i_0)}{(U-i_0)(i_0)}\cdot\prod_{i\in I-\{i_0\}}\binom{(T-i_0)(i)}{(U-i_0)(i)} \\
&= \binom{T(i_0)}{U(i_0)}\cdot\prod_{i\in I-\{i_0\}}\binom{T(i)}{U(i)} = \prod_{i\in I}\binom{T(i)}{U(i)} = \binom{T}{U},
\end{aligned} \tag{3.2.17}$$

where we used Lemma 3.2.4 for the first and last equalities. ❏

**Lemma 3.2.10**

Let $I$ be some set, $i_0$ be a distinguished element in $I$, and $T, U, V \in \mathcal{M}_{\text{fin}}$ be multisets over $I$. Suppose that $UV = T$ and $i_0 \notin U$. Then,

$$\binom{T-i_0}{U} = \binom{T}{U}. \tag{3.2.18}$$

♡

***Proof.*** Note that

$$\binom{(T-i_0)(i_0)}{0} = \binom{T(i_0)}{0} = 1. \tag{3.2.19}$$

Since $i_0 \notin U$, $U(i_0) = 0$, so

$$\begin{aligned}\binom{T-i_0}{U} &= \prod_{i\in I}\binom{(T-i_0)(i)}{U(i)} = \binom{(T-i_0)(i_0)}{U(i_0)} \cdot \prod_{i\in I-\{i_0\}}\binom{(T-i_0)(i)}{U(i)} \\ &= \binom{(T-i_0)(i_0)}{0} \cdot \prod_{i\in I-\{i_0\}}\binom{T(i)}{U(i)} = \binom{T(i_0)}{0} \cdot \prod_{i\in I-\{i_0\}}\binom{T(i)}{U(i)} \\ &= \binom{T(i_0)}{U(i_0)} \cdot \prod_{i\in I-\{i_0\}}\binom{T(i)}{U(i)} = \prod_{i\in I}\binom{T(i)}{U(i)} = \binom{T}{U},\end{aligned} \tag{3.2.20}$$

where we used Lemma 3.2.4 for the first and last equalities. ❏

Of all the lemmas about multisets presented in this section, the last three are important to keep in mind, as we will see that they are essential to the proof of the Generalized Synthetic Leibnitz Rule for Taylor λ-calculus (i.e., Corollary 3.3.34) presented at the end of the next section.

## 3.3 Defining Taylor λ-Calculus and the Partial Derivative

It is finally time to define Taylor λ-terms. Following the structure of section 2.2, we define Taylor pre-terms, α-equivalence for Taylor pre-terms that equates Taylor pre-terms that represent the same underlying analytic function, and Taylor terms, which is the quotient of Taylor pre-terms by α-equivalence. It should be noted that there are three kinds of Taylor pre-terms: simple Taylor pre-terms, which represent basic analytic functions; simple Taylor pre-poly-terms, which represents finite products of analytic functions; and complex Taylor pre-poy-terms, which represent linear combinations of products of analytic functions.

**Definition 3.3.1. Taylor Pre-Terms**

Given a denumerable set $\mathcal{V}$ of variables, a positive semiring of scalars $\mathcal{R}$, and a distinguished symbol 0, a *simple Taylor pre-term*, a *simple Taylor pre-poly-term*, and a *complex Taylor pre-poly-term* are defined via mutual induction as follows:

(i) Given a variable $x \in \mathcal{V}$, $x$ is a simple Taylor pre-term.

(ii) Given a simple Taylor pre-term $s$ and simple Taylor pre-poly-term $T$, $(\langle s\rangle T)$ is a simple Taylor pre-term.

(iii) Given a variable $x \in \mathcal{V}$ and a simple Taylor pre-term $s$, $(\lambda x.s)$ is a simple Taylor pre-poly-term

(iv) The empty multiset of simple Taylor pre-term 1 is a simple Taylor pre-poly-term.

(v) Given a simple Taylor pre-poly-term $T$ and a simple Taylor pre-term $s$, the concatenation $Ts$ is a simple Taylor pre-poly-term.

(vi) 0 is a complex Taylor pre-poly-term.

(vii) Given a simple Taylor pre-poly-term $S$, $S$ is a complex Taylor pre-poly-term.

(viii) Given a scalar $r \in \mathcal{R}$ and a complex Taylor pre-poly-term $\sigma$, $(r \cdot \sigma)$ is a complex Taylor pre-poly-term.

(ix) Given complex Taylor pre-poly-terms $\sigma$ and $\tau$, $(\sigma+\tau)$ is a complex Taylor pre-poly-term. Note that $+$ here is syntactic and does not refer to concatenation of multisets.

Let $\Delta_{t,\mathcal{V}}$ denote the set of simple Taylor pre-terms, let $\Delta^!_{t,\mathcal{V}}$ denote the set of simple Taylor pre-poly-terms, and let $\mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ denote the set of complex Taylor pre-poly-terms. Note that $\Delta^!_{t,\mathcal{V}} = \mathcal{M}_{\text{fin}}(\Delta_{t,\mathcal{V}})$.

♣

It should be noted that we identify the set of simple Taylor terms $\Delta_{t,\mathcal{V}}$ with the subset $\mathcal{M}_1(\Delta_{t,\mathcal{V}})$ of simple Taylor poly-terms $\Delta^!_{t,\mathcal{V}} = \mathcal{M}_{\text{fin}}(\Delta_{t,\mathcal{V}})$, and we identity the set of simple Taylor poly-terms $\Delta^!_{t,\mathcal{V}}$ with the obvious subset of complex Taylor poly-terms $\mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$.

**Definition 3.3.2. $\alpha$-Conversion for Taylor Pre-Terms**

The relation $=_{t,\alpha}$ is the smallest (with respect to inclusion) transitive, reflexive, and symmetric relation on the set $\mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ of complex Taylor pre-poly-terms satisfying the following properties:

(i) If $x, y \in \mathcal{V}$ are variables and $s \in \Delta_{t,\mathcal{V}}$ is a simple Taylor pre-terms such that $y \notin \mathrm{FV}(s)$ and $y$ can be substituted for $x$ in $s$, then $\lambda x.s =_{t,\alpha} \lambda y.s[y/x]$.

(ii) If $s, t \in \Delta_{t,\mathcal{V}}$ are simple Taylor pre-terms such that $s =_{t,\alpha} t$ and $S, T \in \Delta^!_{t,\mathcal{V}}$ are simple pre-poly-terms such that $S =_{t,\alpha} T$, then $\langle s\rangle S =_{t,\alpha} \langle t\rangle T$.

(iii) If $x \in \mathcal{V}$ is a variable and $s, t \in \Delta_{t,\mathcal{V}}$ are simple Taylor pre-terms such that $s =_{t,\alpha} t$, then $\lambda x.s =_{t,\alpha} \lambda x.t$.

(iv) If $s, t \in \Delta_{t,\mathcal{V}}$ are simple Taylor pre-terms such that $s =_{t,\alpha} t$ and $S, T \in \Delta^!_{t,\mathcal{V}}$ are simple pre-poly-terms such that $S =_{t,\alpha} T$, then $Ss =_{t,\alpha} Tt$.

(v) If $r \in \mathcal{R}$ is a scalar and $\tau, \sigma \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ are complex Taylor pre-poly-terms such that $\tau =_{t,\alpha} \sigma$, then $r \cdot \tau =_{t,\alpha} r \cdot \sigma$.

(vi) If $\mu, \nu, \sigma, \tau \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ are complex Taylor pre-poly-terms such that $\tau =_{t,\alpha} \mu$ and $\sigma =_{t,\alpha} \nu$, then $\tau + \sigma =_{t,\alpha} \mu + \nu$.

(vii) If $\sigma, \tau \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ are complex Taylor pre-poly-terms, then $\sigma + \tau =_{t,\alpha} \tau + \sigma$.

(viii) If $\rho, \tau, \sigma \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ are complex Taylor pre-poly-terms, then $(\rho+\sigma)+\tau =_{t,\alpha} \rho+(\sigma+\tau)$.

(ix) If $\tau \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ is a complex Taylor pre-poly-term, then $\tau + 0 =_{t,\alpha} \tau$ and $\tau =_{t,\alpha} \tau + 0$.

(x) If $r_1, r_2 \in \mathcal{R}$ are scalars and $\tau \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ is a complex Taylor pre-poly-term, then $r_1 \cdot \tau + r_2 \cdot \tau =_{t,\alpha} (r_1 + r_2) \cdot \tau$.

(xi) If $r \in \mathcal{R}$ is a scalar and $\tau, \sigma \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ are complex Taylor pre-poly-term, then $r \cdot \tau + r \cdot \sigma =_{t,\alpha} r \cdot (\tau + \sigma)$.

(xii) If $r_1, r_2 \in \mathcal{R}$ are scalars and $\tau \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ is a complex Taylor pre-poly-term, then $r_1 \cdot (r_2 \cdot \tau) =_{t,\alpha} (r_1 \cdot r_2) \cdot \tau$.

(xiii) If $\tau \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ is a complex Taylor pre-poly-term, then $1 \cdot \tau =_{t,\alpha} \tau$.

(xiv) If $r \in \mathcal{R}$ is a scalar, $r \cdot 0 =_{t,\alpha} 0$.

(xv) If $r_1, r_2 \in \mathcal{R}$ are scalars and $\tau \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ is a complex Taylor pre-poly-term, then $(r_1 \cdot r_2) \cdot \tau =_{t,\alpha} (r_2 \cdot r_1) \cdot \tau$. ♣

With the definition of Taylor pre-terms and $\alpha$-equivalence laid out, we can finally define Taylor terms, the main element we work with in Taylor $\lambda$-calculus. Taylor terms, as we stated at the begging of the section, are just Taylor pre-terms where we identify Taylor pre-terms by $\alpha$-equivalence.

**Definition 3.3.3. Taylor Terms**

Given $\sigma \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$, let $[\sigma]_{t,\alpha}$ be the collection of complex pre-poly-terms that are $\alpha$-equivalent to $\sigma$. That is, given $\sigma \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$, let

$$[\sigma]_{t,\alpha} \coloneqq \{\tau \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle : \sigma =_{t,\alpha} \tau\}. \tag{3.3.1}$$

The sets of *simple Taylor terms* $\Lambda_{t,\mathcal{V}}$, *simple Taylor poly-terms* $\Lambda^!_{t,\mathcal{V}}$, and *complex Taylor poly-terms* $\mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ are the sets of simple Taylor pre-terms $\Delta_{t,\mathcal{V}}$, simple Taylor pre-poly-terms $\Delta^!_{t,\mathcal{V}}$, and complex Taylor pre-poly-terms $\mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ quotiented by $\alpha$-equivalence $=_{t,\alpha}$, respectively. That is,

$$\Lambda_{t,\mathcal{V}} \coloneqq \{[s]_{t,\alpha} : s \in \Delta_{t,\mathcal{V}}\}, \tag{3.3.2a}$$
$$\Lambda^!_{t,\mathcal{V}} \coloneqq \{[T]_{t,\alpha} : T \in \Delta^!_{t,\mathcal{V}}\}, \tag{3.3.2b}$$
$$\mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle \coloneqq \{[\sigma]_{t,\alpha} : \sigma \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle\}. \tag{3.3.2c}$$
♣

The basic operations on Taylor pre-terms can be lifted to Taylor terms in the canonical way.

**Notation 3.3.4. Operations on Taylor Terms**

Let $s \in \Lambda_{t,\mathcal{V}}$ be a simple Taylor term, $T \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, $r \in \mathcal{R}$ be a scalar, and let $\sigma, \tau \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ be complex Taylor poly-terms. Then, $\lambda x.s$ and $\langle s\rangle T$ are the unique simple Taylor terms, $1$ and $sT$ are the unique simple Taylor poly-terms, and $0$, $r \cdot \sigma$, and $\sigma + \tau$ are the unique complex Taylor poly-terms such that

$$\lambda x.s = [\lambda x.s']_{t,\alpha} \in \Lambda_{t,\mathcal{V}}, \tag{3.3.3a}$$
$$\langle s\rangle T = [\langle s'\rangle T']_{t,\alpha} \in \Lambda_{t,\mathcal{V}}, \tag{3.3.3b}$$
$$1 = [1]_{t,\alpha} \in \Lambda^!_{t,\mathcal{V}}, \tag{3.3.3c}$$
$$sT = [s'T']_{t,\alpha} \in \Lambda^!_{t,\mathcal{V}}, \tag{3.3.3d}$$
$$0 = [0]_{t,\alpha} \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle, \tag{3.3.3e}$$
$$r \cdot \sigma = [r \cdot \sigma']_{t,\alpha} \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle, \tag{3.3.3f}$$
$$\sigma + \tau = [\sigma' + \tau']_{t,\alpha}, \tag{3.3.3g}$$

where $s' \in \Delta_{t,\mathcal{V}}$ is a simple Taylor pre-term , $T' \in \Delta^!_{t,\mathcal{V}}$ is a simple Taylor pre-poly-term, and $\sigma', \tau' \in \mathcal{R}\langle\Delta^!_{t,\mathcal{V}}\rangle$ are complex Taylor pre-terms such that $s = [s']_{t,\alpha}$, $T = [T']_{t,\alpha}$, $\sigma = [\sigma']_{t,\alpha}$, and $\tau = {\tau'}_{t,\alpha}$. The uniqueness follows from Definition 3.3.2(ii)-(vi). ♠

At this point, there is an odd asymmetry in Taylor $\lambda$-calculus. In Taylor $\lambda$-calculus, there are simple and complex Taylor poly-terms, where a complex Taylor poly-terms is just a linear combination of simple Taylor poly-terms, but in Taylor $\lambda$-calculus, at this point, there is only simple Taylor terms. There isn't, at this point complex Taylor terms. Let's fix this. Remember a simple Taylor term is just a complex Taylor term of cardinality 1 and a complex Taylor term should be a linear combination of simple Taylor terms. Hence, a complex Taylor term can viewed as a complex Taylor poly-term whose support has cardinality 1.

**Definition 3.3.5. Complex Taylor terms**

The set of complex Taylor terms $\mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ are the linear combinations of simple Taylor terms. That is,

$$\mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle \coloneqq \{\tau \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle : |S| = 1\ \forall S \in \operatorname{supp}(\tau)\}. \tag{3.3.4}$$

♣

Recall that every structure in Taylor $\lambda$-calculus is multi-linear, so we extend the operations of abstraction, differentiated application, and concatenation from simple Taylor terms to complex Taylor terms by linearity.

**Notation 3.3.6**

Let $x \in \mathcal{V}$ be a variable, $s \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ be a complex Taylor term, and $T \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ be a complex Taylor poly-term. Then, the *complex abstraction* $\lambda x.s$ and the *complex differentiated application* $\langle s\rangle\tau$ are the complex Taylor terms and the *complex concatenation* $\tau s$ is the complex Taylor poly-term such that

$$\lambda x.s = \sum_{i=1}^{m} a_i \cdot \lambda x.s_i, \tag{3.3.5a}$$

$$\langle s\rangle\tau = \sum_{i=1}^{m}\sum_{j=1}^{n} (a_i \cdot b_j) \cdot \langle s_i\rangle T_j, \tag{3.3.5b}$$

$$\tau s = \sum_{i=1}^{n} s_i \cdot \tau s_i, \tag{3.3.5c}$$

where $m, n \in \mathbb{N}_0$ are nonnegative integers, $a_1, \ldots, a_m, b_1, \ldots, b_n \in \mathcal{R}$ are scalars, $s_1, \ldots, s_m \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms, and $T_1, \ldots, T_n \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms such that $s = \sum_{i=1}^{m} a_i \cdot s_i$ and $\tau = \sum_{i=1}^{n} b_i \cdot T_i$, which we know exist and are unique by Lemma 3.3.10(xii).

♠

With these operations on Taylor terms, we will no longer need to make explicit reference to Taylor pre-terms, and from now on we will only work with Taylor terms directly. We now define the size, the set of free variables, and the substitution of. simple Taylor term for a variable for Taylor terms. The proofs that the latter two operations on Taylor terms are well-defined are virtually identical to the proofs that those operation are well-defined in differential $\lambda$-calculus, and are left to the reader.

**Definition 3.3.7. Size of Simple Taylor poly-Terms**

Define the function size: $\Lambda^!_{t,\mathcal{V}} \to \mathbb{N}_0$ that takes a simple Taylor poly-term to a nonnegative number by:

$$\operatorname{size}(x) = 1, \tag{3.3.6a}$$

$$\operatorname{size}(\lambda x.s) = 1 + \operatorname{size}(s), \tag{3.3.6b}$$
$$\operatorname{size}(\langle s \rangle T) = 1 + \operatorname{size}(s) + \operatorname{size}(T), \tag{3.3.6c}$$
$$\operatorname{size}(t_1 \dots t_n) = n + \sum_{i=1}^{n} \operatorname{size}(t_i), \tag{3.3.6d}$$

where $x \in \mathcal{V}$ is a variable, $s, t_1, \dots, t_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms, and $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-terms. The image of a simple Taylor poly-term $S \in \Lambda^!_{t,\mathcal{V}}$ under size is called the *size* of $S$.

♣

**Definition 3.3.8. Free Variables of Taylor Terms**

Define the function size: $\mathcal{R}\langle \Lambda^!_{t,\mathcal{V}} \rangle \to \mathcal{V}^*$ that takes a complex Taylor poly-term to a finite subset of variables by:

$$\operatorname{FV}(x) = 1, \tag{3.3.7a}$$
$$\operatorname{FV}(\lambda x.s) = \operatorname{FV}(s) - \{x\}, \tag{3.3.7b}$$
$$\operatorname{FV}(\langle s \rangle T) = \operatorname{FV}(s) \cup \operatorname{FV}(T), \tag{3.3.7c}$$
$$\operatorname{FV}(t_1 \dots t_n) = \bigcup_{i=1}^{n} \operatorname{FV}(t_i), \tag{3.3.7d}$$
$$\operatorname{FV}\left(\sum_{i=1}^{n} r_i \cdot S_i\right) = \bigcup_{i=1}^{n} \operatorname{FV}(S_i), \tag{3.3.7e}$$

where $x \in \mathcal{V}$ is a variable, $n \in \mathbb{N}_0$ is a nonnegative integer, $s, t_1, \dots, t_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms, $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-terms, $r_1, \dots, r_n \in \mathcal{R}$ are scalars, and $S_1, \dots, S_n \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms. The image of a complex Taylor poly-term $\tau \in \Lambda^!_{t,\mathcal{V}}$ under FV is called the *free variables* of $\tau$.

♣

**Definition 3.3.9. Substitution for Taylor Terms**

Given a variable $x \in \mathcal{V}$, a complex Taylor poly-term $\sigma \in \mathcal{R}\langle \Lambda^!_{t,\mathcal{V}} \rangle$, and a complex Taylor term $t \in \mathcal{R}\langle \Lambda_{t,\mathcal{V}} \rangle$, the *substitution* $\sigma[t/x]$ of $t$ for $x$ in $\sigma$ is defined by:

$$x[t/x] = t, \tag{3.3.8a}$$
$$y[t/x] = y, \tag{3.3.8b}$$
$$(\lambda x.s)[t/x] = \lambda x.s, \tag{3.3.8c}$$
$$(\lambda y.s)[t/x] = \lambda y.s[t/x] \qquad \text{if } y \notin \operatorname{FV}(t), \tag{3.3.8d}$$
$$(\langle s \rangle T)[t/x] = \langle s[t/x] \rangle T[t/x], \tag{3.3.8e}$$
$$(t_1 \dots t_n)[t/x] = t_1[t/x] \dots t_n[t/x], \tag{3.3.8f}$$
$$\left(\sum_{i=1}^{n} r_i \cdot S_i\right)[t/x] = \sum_{i=1}^{n} r_i \cdot S_i[t/x], \tag{3.3.8g}$$

where $y \in \mathcal{V}$ is a variable distinct from $x$, $n \in \mathbb{N}_0$ is a nonnegative integer, $s, t_1, \dots, t_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms, $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-terms, $r_1, \dots, r_n \in \mathcal{R}$ are scalars, and $S_1, \dots, S_n \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms.

♣

There are a number of elementary properties of free variables, substitution, and uniqueness of Taylor terms that we will use throughout the remainder of this chapter.

**Lemma 3.3.10**

Let $m, n \in \mathbb{N}_0$ be nonnegative integers, $x, y \in \mathcal{V}$ be variables, $s, s', t_1, \ldots, t_m, t'_1, \ldots, t'_n \in \Lambda_{t,\mathcal{V}}$ be simple Taylor terms, $u, v \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ be complex Taylor terms, $T, T' \in \Lambda^!_{t,\mathcal{V}}$ be simple poly-terms, and $\tau \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ be a complex Taylor poly-terms. Then, the following statements are true:

(i) If $x \notin \mathrm{FV}(\tau)$, then $\tau[u/x] = \tau$.

(ii) $y \in \mathrm{FV}(\tau[u/x])$ if and only if either $y \in \mathrm{FV}(\tau)$ and $x \neq y$ or $y \in \mathrm{FV}(u)$ and $x \in \mathrm{FV}(\tau)$.

(iii) $\tau[x/x] = \tau$.

(iv) If $y \neq x \notin \mathrm{FV}(v)$, then $\tau[u/x][v/y] = \tau[v/y][u[v/y]/x]$.

(v) If $y \notin \mathrm{FV}(\tau)$, then $\tau[y/x][u/y] = \tau[u/x]$.

(vi) If $y \notin \mathrm{FV}(s)$, then $D_1^n \lambda x.s * (t_1, \ldots, t_n) = D_1^n \lambda y.s[y/x] * (t_1, \ldots, t_n)$.

(vii) $\lambda x.s = \lambda y.s'$ if and only if $s'[x/y] = s$.

(viii) $\langle s\rangle T = \langle s'\rangle T'$ if and only if $s = s'$ and $T = T'$.

(ix) $t_1 \ldots t_m = t'_1 \ldots t'_n$ if and only if $n = m$ and $t_i = t'_i$ for all $1 \leqslant i \leqslant m = n$.

(x) Summation makes the set of complex Taylor poly-terms $\mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ into a commutative monoid with 0 as the unit.

(xi) $\mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ is a left-$\mathcal{R}$-module on the commutative monoid $(\mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle, +, 0)$.

(xii) There exists a unique nonnegative integer $\ell \in \mathbb{N}_0$, unique scalars $r_1, \ldots, r_\ell \in \mathcal{R}$, and unique distinct simple Taylor terms $s_1, \cdots, s_\ell \in \Lambda_{d,\mathcal{V}}$ such that $\tau = r_1 \cdot s_1 + \cdots + r_\ell \cdot s_\ell$.

**Note** that we define the *support* $\mathrm{supp}(\tau)$ of $\tau$ to be this unique set of distinct simple differential terms. That is, $\mathrm{supp}(\tau) = \{s_1, \ldots, s_\ell\}$. ♡

***Proof.*** The proof for this is analogous to the proof of Lemma 2.2.29. ❏

Now, we define the partial derivative on Taylor terms. Again, the fact that this operation is well-defined is obvious and is left to the reader.

**Definition 3.3.11. Partial Derivative for Taylor Terms**

Given a variable $x \in \mathcal{V}$, a complex Taylor poly-term $\sigma \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$, and a complex Taylor term $t \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$, the *partial derivative* $\partial_x \sigma * t$ of $\sigma$ with respect to $x$ times $t$ is defined by:

$$\partial_x(x) * t = t, \tag{3.3.9a}$$
$$\partial_x(y) * t = 0, \tag{3.3.9b}$$
$$\partial_x(\lambda y.s) * t = \lambda y.\partial_x(s) * t \qquad \text{if } y \notin \mathrm{FV}(t), \tag{3.3.9c}$$
$$\partial_x(\langle s\rangle T) * t = \langle \partial_x(s) * t\rangle T + \langle s\rangle(\partial_x(T) * t), \tag{3.3.9d}$$

$$\partial_x(t_1 \dots t_n) * t = \sum_{i=1}^{n} t_1 \dots t_{i-1}(\partial_x(t_i) * t)t_{i+1} \dots t_n, \tag{3.3.9e}$$

$$\partial_x\left(\sum_{i=1}^{n} r_i \cdot S_i\right) * t = \sum_{i=1}^{n} r_i \cdot \partial_x(S_i) * t, \tag{3.3.9f}$$

where $y \in \mathcal{V}$ is a variable distinct from $x$, $n \in \mathbb{N}_0$ is a nonnegative integer, $s, t_1, \dots, t_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms, $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-terms, $r_1, \dots, r_n \in \mathcal{R}$ are scalars, and $S_1, \dots, S_n \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms. ♣

With the definition of the partial derivative of Taylor terms at hand, we define a number of properties of the partial derivative, including the Synthetic Schwartz Lemma for Taylor λ-calculus (i.e., Corollary 3.3.16).

**Lemma 3.3.12**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, $x \in \mathcal{V}$ be a variable, and $t \in \Lambda_{t,\mathcal{V}}$ be a simple Taylor term. Then, for $U \in \operatorname{supp}(\partial_x S * t)$, one has $\operatorname{size}(U) = \operatorname{size}(S) + \operatorname{size}(t) - 1$. ♡

***Proof***. We prove this via induction on $S$.

1. Assume that $S = x$. Then, $\partial_x S * t = \partial_x x * t = t$. Take any $U \in \operatorname{supp}(\partial_x S * t) = \operatorname{supp}(t) = \{t\}$, which implies that $U = t$. Hence,

$$\begin{aligned} \operatorname{size}(S) + \operatorname{size}(t) - 1 &= \operatorname{size}(x) + \operatorname{size}(t) - 1 = 1 + \operatorname{size}(t) - 1 \\ &= \operatorname{size}(t) = \operatorname{size}(U). \end{aligned} \tag{3.3.10}$$

2. Assume that $S = y$, where $y \in \mathcal{V}$ is a variable such that $y \neq x$. Then, $\operatorname{supp}(\partial_x S * t) = \operatorname{supp}(\partial_x y * t) = \operatorname{supp}(0) = \varnothing$, so our result holds trivially.

3. Assume that $S = \lambda y.s$, where $y \in \mathcal{V}$ is a fresh variable and $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. Then, $\partial_x S * t = \partial_x \lambda y.s * t = \lambda y.\partial_x s * t$. Take any $U \in \operatorname{supp}(\partial_x S * t) = \operatorname{supp}(\lambda y.\partial_x s * t)$. Then, there exists a simple Taylor term $v \in \operatorname{supp}(\partial_x s * t)$ such that $U = \lambda y.v$. By the induction hypothesis, $\operatorname{size}(v) = \operatorname{size}(s) + \operatorname{size}(t) - 1$. Hence,

$$\begin{aligned} \operatorname{size}(S) + \operatorname{size}(t) - 1 &= \operatorname{size}(\lambda y.s) + \operatorname{size}(t) - 1 = 1 + \operatorname{size}(s) + \operatorname{size}(t) - 1 \\ &= 1 + \operatorname{size}(v) = \operatorname{size}(\lambda y.v) = \operatorname{size}(U). \end{aligned} \tag{3.3.11}$$

4. Assume that $S = \langle s \rangle T$, where $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-term. Then, $\partial_x S * t = \partial_x \langle s \rangle T * t = \langle \partial_x s * t \rangle T + \langle s \rangle (\partial_x T * t)$. Take any $U \in \operatorname{supp}(\partial_x S * t) = \operatorname{supp}(\langle \partial_x s * t \rangle T + \langle s \rangle (\partial_x T * t)) = \operatorname{supp}(\langle \partial_x s * t \rangle T) \cup \operatorname{supp}(\langle s \rangle (\partial_x T * t))$. Suppose that there exists a simple Taylor term $v \in \operatorname{supp}(\partial_x s * t)$ such that $U = \langle v \rangle T$. By the induction hypothesis, $\operatorname{size}(v) = \operatorname{size}(s) + \operatorname{size}(t) - 1$, which implies that

$$\begin{aligned} \operatorname{size}(S) + \operatorname{size}(t) - 1 &= \operatorname{size}(\langle s \rangle T) + \operatorname{size}(t) - 1 \\ &= 1 + \operatorname{size}(s) + \operatorname{size}(T) + \operatorname{size}(t) - 1 \\ &= 1 + \operatorname{size}(v) + \operatorname{size}(T) = \operatorname{size}(\langle v \rangle T) = \operatorname{size}(U). \end{aligned} \tag{3.3.12}$$

Instead, suppose that there exists a simple Taylor poly-term $V \in \operatorname{supp}(\partial_x T * t)$ such that $U = \langle s \rangle V$. By the induction hypothesis, $\operatorname{size}(v) = \operatorname{size}(s) + \operatorname{size}(t) - 1$, which implies that

$$\begin{aligned}\operatorname{size}(S) + \operatorname{size}(t) - 1 &= \operatorname{size}(\langle s \rangle T) + \operatorname{size}(t) - 1 \\ &= 1 + \operatorname{size}(s) + \operatorname{size}(T) + \operatorname{size}(t) - 1 \\ &= 1 + \operatorname{size}(s) + \operatorname{size}(V) = \operatorname{size}(\langle s \rangle V) = \operatorname{size}(U).\end{aligned} \tag{3.3.13}$$

5. Assume that $S = s_1 \ldots s_n$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $s_1, \ldots, s_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. Then, $\partial_x S * t = \partial_x s_1 \ldots s_n * t = \sum_{i=1}^n s_1 \ldots (\partial_x s_i * t) \ldots s_n$. Take any

$$\begin{aligned}U \in \operatorname{supp}(\partial_x S * t) &= \operatorname{supp}\left(\sum_{i=1}^n s_1 \ldots (\partial_x s_i * t) \ldots s_n\right) \\ &= \bigcup_{i \in I} \operatorname{supp}(s_1 \ldots (\partial_x s_i * t) \ldots s_n).\end{aligned} \tag{3.3.14}$$

Then, there exists $1 \leqslant i_0 \leqslant n$ and simple Taylor term $v \in \partial_x s_i * t$ such that $U = s_1 \ldots v \ldots s_n$. By the induction hypothesis, $\operatorname{size}(v) = \operatorname{size}(s_i) + \operatorname{size}(t) - 1$, which implies that

$$\begin{aligned}\operatorname{size}(S) + \operatorname{size}(t) - 1 &= \operatorname{size}(s_1 \ldots s_n) + \operatorname{size}(t) - 1 \\ &= n + \sum_{i=0}^n \operatorname{size}(s_i) + \operatorname{size}(t) - 1 \\ &= n + \sum_{i=0, i \neq i_0}^n \operatorname{size}(s_i) + \operatorname{size}(s_{i_0}) + \operatorname{size}(t) - 1 \\ &= n + \sum_{i=0, i \neq i_0}^n \operatorname{size}(s_i) + \operatorname{size}(v) \\ &= \operatorname{size}(s_1 \ldots \eta \ldots s_n) = \operatorname{size}(U).\end{aligned} \tag{3.3.15}$$

This completes the induction. ❑

**Lemma 3.3.13**

Let $\sigma \in \mathcal{R}\langle \Lambda^!_{t,\mathcal{V}} \rangle$ be a complex Taylor poly-term, $x \in \mathcal{V}$ be a variable, $n \in \mathbb{N}_0$ be a nonnegative integer, $r_1, \ldots, r_n \in \mathcal{R}$ be scalars, and $t_1, \ldots, t_n \in \Lambda_{t,\mathcal{V}}$ be simple Taylor terms. Then, $\partial_x \sigma * (\sum_{i=1}^n r_i \cdot t_i) = \sum_{i=1}^n r_i \cdot (\partial_x \sigma * t_i)$. ♡

***Proof.*** We prove this via induction on $\sigma$.

1. Assume that $\sigma = x$. Then,

$$\begin{aligned}\partial_x \sigma * \left(\sum_{i=1}^n r_i \cdot t_i\right) &= \partial_x x * \left(\sum_{i=1}^n r_i \cdot t_i\right) = \sum_{i=1}^n r_i \cdot t_i \\ &= \sum_{i=1}^n r_i \cdot (\partial_x x * t_i) = \sum_{i=1}^n r_i \cdot (\partial_x \sigma * t_i).\end{aligned} \tag{3.3.16}$$

2. Assume that $\sigma = y$, where $y \in \mathcal{V}$ is a variable distinct from $x$. Then,

$$\begin{aligned}\partial_x\sigma * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) &= \partial_x y * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) = 0 = \sum_{i=1}^{n} r_i \cdot 0 \\ &= \sum_{i=1}^{n} r_i \cdot (\partial_x y * t_i) = \sum_{i=1}^{n} r_i \cdot (\partial_x\sigma * t_i).\end{aligned} \tag{3.3.17}$$

3. Assume that $\sigma = \lambda y.s$, where $y \in \mathcal{V}$ is a fresh variable and $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. By the induction hypothesis, $\partial_x s * (\sum_{i=1}^{n} r_i \cdot t_i) = \sum_{i=1}^{n} r_i \cdot (\partial_x s * t_i)$, which implies that

$$\begin{aligned}\partial_x\sigma * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) &= \partial_x\lambda y.s * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) = \lambda y.\partial_x s * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) \\ &= \lambda y. \sum_{i=1}^{n} r_i \cdot (\partial_x s * t_i) = \sum_{i=1}^{n} r_i \cdot \lambda y.(\partial_x s * t_i) \\ &= \sum_{i=1}^{n} r_i \cdot (\partial_x\lambda y.s * t_i) = \sum_{i=1}^{n} r_i \cdot (\partial_x\sigma * t_i).\end{aligned} \tag{3.3.18}$$

4. Assume that $\sigma = \langle s\rangle T$, where $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $T \in \Lambda^{!}_{t,\mathcal{V}}$ is a simple Taylor poly-term. By the induction hypothesis, $\partial_x s * (\sum_{i=1}^{n} r_i \cdot t_i) = \sum_{i=1}^{n} r_i \cdot (\partial_x s * t_i)$ and $\partial_x T * (\sum_{i=1}^{n} r_i \cdot t_i) = \sum_{i=1}^{n} r_i \cdot (\partial_x T * t_i)$, which implies that

$$\begin{aligned}\partial_x\sigma * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) &= \partial_x\langle s\rangle T * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) \\ &= \langle \partial_x s * \left(\sum_{i=1}^{n} r_i \cdot t_i\right)\rangle T + \langle s\rangle \left(\partial_x T * \left(\sum_{i=1}^{n} r_i \cdot t_i\right)\right) \\ &= \langle \sum_{i=1}^{n} r_i \cdot (\partial_x s * t_i)\rangle T + \langle s\rangle \left(\sum_{i=1}^{n} r_i \cdot (\partial_x T * t_i)\right) \\ &= \sum_{i=1}^{n} (r_i \cdot \langle \partial_x s * t_i\rangle T + \langle s\rangle(\partial_x T * t_i)) \\ &= \sum_{i=1}^{n} r_i \cdot \partial_x\langle s\rangle T * t_i = \sum_{i=1}^{n} r_i \cdot \partial_x\sigma * t_i.\end{aligned} \tag{3.3.19}$$

5. Assume that $\sigma = s_1 \ldots s_m$, where $m \in \mathbb{N}_0$ is a nonnegative integer and $s_1, \ldots, s_m \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. By the induction hypothesis, $\partial_x s_j * (\sum_{i=1}^{n} r_i \cdot t_i) = \sum_{i=1}^{n} r_i \cdot (\partial_x s_j * t_i)$ for all $1 \leqslant j \leqslant m$, which implies that

$$\begin{aligned}\partial_x\sigma * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) &= \partial_x s_1 \ldots s_m * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) \\ &= \sum_{i=1}^{m} s_1 \ldots \left(\partial_x s_j * \left(\sum_{i=1}^{n} r_i \cdot t_i\right)\right) \ldots s_m \\ &= \sum_{i=1}^{m} s_1 \ldots \left(\sum_{i=1}^{n} r_i \cdot (\partial_x s_j * t_i)\right) \ldots s_m\end{aligned} \tag{3.3.20}$$

$$= \sum_{i=1}^{n} r_i \cdot \sum_{i=1}^{m} s_1 \ldots (\partial_x s_j * t_i) \ldots s_m$$
$$= \sum_{i=1}^{n} r_i \cdot \partial_x s_1 \ldots s_m * t_i = \sum_{i=1}^{n} r_i \cdot \partial_x \sigma * t_i.$$

6. Assume that $\sigma = \sum_{j=1}^{m} a_j \cdot T_j$, where $m \in \mathbb{N}_0$ is a nonnegative integer, $a_1, \ldots, a_m$ are scalars, and $T_1, \ldots, T_m \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms. By the induction hypothesis, $\partial_x T_j * (\sum_{i=1}^{n} r_i \cdot t_i) = \sum_{i=1}^{n} r_i \cdot (\partial_x T_j * t_i)$ for all $1 \leqslant j \leqslant m$, which implies that

$$\begin{aligned}
\partial_x \sigma * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) &= \partial_x \left(\sum_{j=1}^{m} a_j \cdot T_j\right) * \left(\sum_{i=1}^{n} r_i \cdot t_i\right) \\
&= \sum_{j=1}^{m} a_j \cdot \left(\partial_x T_j * \left(\sum_{i=1}^{n} r_i \cdot t_i\right)\right) = \sum_{j=1}^{m} a_j \cdot \sum_{i=1}^{n} r_i \cdot (\partial_x T_j * t_i) \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} a_j \cdot (r_i \cdot (\partial_x T_j * t_i)) = \sum_{j=1}^{m} \sum_{i=1}^{n} (a_j \cdot r_i) \cdot (\partial_x T_j * t_i) \qquad (3.3.21) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} (r_i \cdot a_j) \cdot (\partial_x T_j * t_i) = \sum_{i=1}^{n} r_i \cdot \sum_{j=1}^{m} a_j \cdot (\partial_x T_j * t_i) \\
&= \sum_{i=1}^{n} r_i \cdot \left(\partial_x \left(\sum_{j=1}^{m} a_j \cdot T_j\right) * t_i\right) = \sum_{i=1}^{n} r_i \cdot (\partial_x \sigma * t_i).
\end{aligned}$$

This completes the induction. ❏

**Lemma 3.3.14**

Let $\sigma \in \mathcal{R}\langle \Lambda^!_{t,\mathcal{V}} \rangle$ be a complex Taylor poly-term, $x \in \mathcal{V}$ be a variable, and $t \in \mathcal{R}\langle \Lambda_{t,\mathcal{V}} \rangle$ be simple Taylor term. Suppose that $x \notin \mathrm{FV}(\sigma)$. Then, $\partial_x \sigma * t = 0$. ♡

***Proof.*** We prove this via induction on $\sigma$.

1. Assume that $\sigma = y$, where $y \in \mathcal{V}$ is a variable. Since $x \notin \mathrm{FV}(\sigma) = \mathrm{FV}(y) = \{y\}$, we have that $y \neq x$, which implies that $\partial_x \sigma * t = \partial_x y * t = 0$.

2. Assume that $\sigma = \lambda y.s$, where $y \in \mathcal{V}$ is a fresh variable. Since $x \notin \mathrm{FV}(\sigma) = \mathrm{FV}(\lambda y.s) = \mathrm{FV}(s) - \{y\}$ and $y \in \mathcal{V}$ is a fresh variable, $x \notin \mathrm{FV}(s)$, so by the induction hypothesis, $\partial_x s * t = 0$, which implies that $\partial_x \sigma * t = \partial_x \lambda y.s * t = \lambda y.\partial_x s * t = \lambda y.0 = 0$.

3. Assume that $\sigma = \langle s \rangle T$, where $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-term. Since $x \notin \mathrm{FV}(\sigma) = \mathrm{FV}(\langle s \rangle T) = \mathrm{FV}(s) \cup \mathrm{FV}(T)$, by the induction hypothesis, $\partial_x s * t = 0$ and $\partial_x T * t = 0$, which implies that $\partial_x \sigma * t = \partial_x \langle s \rangle T * t = \langle \partial_x s * t \rangle T + \langle s \rangle (\partial_x T * t) = \langle 0 \rangle T + \langle s \rangle 0 = 0$.

4. Assume that $\sigma = s_1 \ldots s_n$, where $s_1, \ldots, s_n$ are simple Taylor terms. For each $1 \leqslant j \leqslant n$, $x \notin \mathrm{FV}(\sigma) = \mathrm{FV}(s_1 \ldots s_n) = \bigcup_{i=1}^{n} \mathrm{FV}(s_i) \supset \mathrm{FV}(s_j)$, so by the induction hypothesis, $\partial_x s_j * t = 0$. Hence, $\partial_x \sigma * t = \partial_x s_1 \ldots s_n * t = \sum_{i=1}^{n} s_1 \ldots (\partial_x s_i * t) \ldots s_n = \sum_{i=1}^{n} s_1 \ldots 0 \ldots s_n = \sum_{i=1}^{n} 0 = 0$.

5. Assume that $\sigma = \sum_{i=1}^{n} r_i \cdot T_i$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \ldots, r_n$ are scalars, and $T_1, \ldots, T_n \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms. For each $1 \leqslant j \leqslant n$, $x \notin \mathrm{FV}(\sigma) = \mathrm{FV}(\sum_{i=1}^{n} r_i \cdot T_i) = \bigcup_{i=1}^{n} \mathrm{FV}(s_i) \supset \mathrm{FV}(s_j)$, so by the induction hypothesis, $\partial_x s_j * t = 0$. Hence, $\partial_x \sigma * t = \partial_x (\sum_{i=1}^{n} r_i \cdot T_i) * t = \sum_{i=1}^{n} r_i \cdot \partial_x T_i * t = \sum_{i=1}^{n} r_i \cdot 0 = 0$.

This completes the induction. ❑

**Lemma 3.3.15**

Let $\sigma \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ be a complex Taylor poly-term, $s, t \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ be complex Taylor terms, $x, y \in \mathcal{V}$ be variables such that $x \notin \mathrm{FV}(t)$. Then,

$$\partial_y(\partial_x\sigma * s) * t = \partial_x(\partial_y\sigma * t) * s + \partial_x\sigma * (\partial_y s * t). \tag{3.3.22}$$

♡

***Proof***. We prove this via induction on $\sigma$.

1. Assume that $\sigma = x$. Then,

$$\begin{aligned}\partial_x(\partial_y\sigma * t) * s + \partial_x\sigma * (\partial_y s * t) &= \partial_x(\partial_y x * t) * s + \partial_x x * (\partial_y s * t)\\ &= \partial_x(0) * s + \partial_y s * t = 0 + \partial_y s * t = \partial_y s * t\\ &= \partial_y(\partial_x x * s) * t = \partial_y(\partial_x\sigma * s) * t.\end{aligned} \tag{3.3.23}$$

2. Assume that that $\sigma = y$. Then,

$$\begin{aligned}\partial_x(\partial_y\sigma * t) * s + \partial_x\sigma * (\partial_y s * t) &= \partial_x(\partial_y y * t) * s + \partial_x y * (\partial_y s * t) = \partial_x t * s + 0\\ &= 0 + 0 = 0 = \partial_y 0 * t = \partial_y(\partial_x y * s) * t\\ &= \partial_y(\partial_x\sigma * s) * t,\end{aligned} \tag{3.3.24}$$

where we used that $x \notin \mathrm{FV}(t)$ for the third equality.

3. Assume that $\sigma = z$, where $z \in \mathcal{V}$ is a variable distinct from $x, y$. Then,

$$\begin{aligned}\partial_x(\partial_y\sigma * t) * s + \partial_x\sigma * (\partial_y s * t) &= \partial_x(\partial_y z * t) * s + \partial_x z * (\partial_y s * t) = \partial_x 0 * s + 0\\ &= 0 + 0 = 0 = \partial_y 0 * t = \partial_y(\partial_x z * s) * t\\ &= \partial_y(\partial_x\sigma * s) * t.\end{aligned} \tag{3.3.25}$$

4. Assume that $\sigma = \lambda z.u$, where $z \in \mathcal{V}$ is a fresh variable and $u \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. By the induction hypothesis, $\partial_y(\partial_x u * s) * t = \partial_x(\partial_y u * t) * s + \partial_x u * (\partial_y s * t)$, which implies that

$$\begin{aligned}\partial_y(\partial_x\sigma * s) * t &= \partial_y(\partial_x\lambda z.u * s) * t = \lambda z.\partial_y(\partial_x u * s) * t\\ &= \lambda z.[\partial_x(\partial_y u * t) * s + \partial_x u * (\partial_y s * t)]\\ &= \lambda z.\partial_x(\partial_y u * t) * s + \lambda z.\partial_x u * (\partial_y s * t)\\ &= \partial_x(\partial_y\lambda z.u * t) * s + \partial_x\lambda z.u * (\partial_y s * t)\\ &= \partial_x(\partial_y\sigma * t) * s + \partial_x\sigma * (\partial_y s * t).\end{aligned} \tag{3.3.26}$$

5. The case where $\sigma = \langle u\rangle U$, where $u \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $U \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-term was shown in Lemma 2 of [9].

6. Assume that $\sigma = u_1 \dots u_n$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $u_1, \dots, u_n \in stt$ are simple Taylor terms. By the induction hypothesis, $\partial_y(\partial_x u_i * s) * t = \partial_x(\partial_y u_i * t) * s + \partial_x u_i * (\partial_y s * t)$ for all $1 \leqslant i \leqslant n$, which implies that

$$\begin{aligned}
\partial_y(\partial_x \sigma * s) * t &= \partial_y(\partial_x u_1 \dots u_n * s) * t \\
&= \partial_y \left( \sum_{i=1}^{n} u_1 \dots (\partial_x u_i * s) \dots u_n \right) * t \\
&= \sum_{i=1}^{n} \partial_y(u_1 \dots (\partial_x u_i * s) \dots u_n) * t \\
&= \sum_{i=1}^{n} \Bigg[ \sum_{j=1}^{i-1} (u_1 \dots (\partial_y u_j * t) \dots (\partial_x u_i * s) \dots u_n) \\
&\quad + (u_1 \dots (\partial_y(\partial_x u_i * s) * t) \dots u_n) \\
&\quad + \sum_{j=i+1}^{n} (u_1 \dots (\partial_x u_i * s) \dots (\partial_y u_j * t) \dots u_n) \Bigg] \\
&= \sum_{i=1}^{n} \Bigg[ \sum_{j=1}^{i-1} (u_1 \dots (\partial_y u_j * t) \dots (\partial_x u_i * s) \dots u_n) \\
&\quad + u_1 \dots (\partial_x(\partial_y u_i * t) * s + \partial_x u_i * (\partial_y s * t)) \dots u_n \\
&\quad + \sum_{j=i+1}^{n} (u_1 \dots (\partial_x u_i * s) \dots (\partial_y u_j * t) \dots u_n) \Bigg] \\
&= \sum_{i=1}^{n} \Bigg[ \sum_{j=1}^{i-1} (u_1 \dots (\partial_y u_j * t) \dots (\partial_x u_i * s) \dots u_n) \qquad (3.3.27) \\
&\quad + u_1 \dots (\partial_x(\partial_y u_i * t) * s) \dots u_n \\
&\quad + u_1 \dots (\partial_x u_i * (\partial_y s * t)) \dots u_n \\
&\quad + \sum_{j=i+1}^{n} (u_1 \dots (\partial_x u_i * s) \dots (\partial_y u_j * t) \dots u_n) \Bigg] \\
&= \sum_{i=1}^{n} \Bigg[ \sum_{j=1}^{i-1} (u_1 \dots (\partial_y u_j * t) \dots (\partial_x u_i * s) \dots u_n) \\
&\quad + u_1 \dots (\partial_x(\partial_y u_i * t) * s) \dots u_n \\
&\quad + \sum_{j=i+1}^{n} (u_1 \dots (\partial_x u_i * s) \dots (\partial_y u_j * t) \dots u_n) \Bigg] \\
&\quad + \sum_{i=1}^{n} u_1 \dots (\partial_x u_i * (\partial_y s * t)) \dots u_n \\
&= \sum_{i=1}^{n} \Bigg[ \sum_{j=1}^{i-1} (u_1 \dots (\partial_y u_j * t) \dots (\partial_x u_i * s) \dots u_n) \\
&\quad + u_1 \dots (\partial_x(\partial_y u_i * t) * s) \dots u_n \\
&\quad + \sum_{j=i+1}^{n} (u_1 \dots (\partial_x u_i * s) \dots (\partial_y u_j * t) \dots u_n) \Bigg]
\end{aligned}$$

$$
\begin{aligned}
&+ \partial_x u_1 \dots u_n * (\partial_y s * t) \\
&= \sum_{i=1}^{n} \Bigg[ \sum_{j=1}^{i-1} (u_1 \dots (\partial_y u_j * t) \dots (\partial_x u_i * s) \dots u_n) \\
&\quad + u_1 \dots (\partial_x(\partial_y u_i * t) * s) \dots u_n \\
&\quad + \sum_{j=i+1}^{n} (u_1 \dots (\partial_x u_i * s) \dots (\partial_y u_j * t) \dots u_n) \Bigg] \\
&\quad + \partial_x \sigma * (\partial_y s * t) \\
&= \sum_{j=1}^{n} \Bigg[ \sum_{i=j+1}^{n} (u_1 \dots (\partial_y u_j * t) \dots (\partial_x u_i * s) \dots u_n) \\
&\quad + u_1 \dots (\partial_x(\partial_y u_i * t) * s) \dots u_n \\
&\quad + \sum_{i=1}^{j-1} (u_1 \dots (\partial_x u_i * s) \dots (\partial_y u_j * t) \dots u_n) \Bigg] \\
&\quad + \partial_x \sigma * (\partial_y s * t) \\
&= \sum_{j=1}^{n} [\partial_x u_1 \dots (\partial_y u_j * t) \dots u_n * s] + \partial_x \sigma * (\partial_y s * t) \\
&= \partial_x \Bigg[ \sum_{j=1}^{n} [u_1 \dots (\partial_y u_j * t) \dots u_n \Bigg] * s + \partial_x \sigma * (\partial_y s * t) \\
&= \partial_x (\partial_y u_1 \dots u_n * t) * s + \partial_x \sigma * (\partial_y s * t) \\
&= \partial_x (\partial_y \sigma * t) * s + \partial_x \sigma * (\partial_y s * t).
\end{aligned}
$$

7. Assume that $\sigma = \sum_{i=1}^{n} r_i \cdot T_i$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $T_1, \dots, T_n \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms. By the induction hypothesis, $\partial_y(\partial_x T_i * s) * t = \partial_x(\partial_y T_i * t) * s + \partial_x T_i * (\partial_y s * t)$ for all $1 \leqslant i \leqslant n$, which implies that

$$
\begin{aligned}
\partial_y(\partial_x \sigma * s) * t &= \partial_y \left( \partial_x \left( \sum_{i=1}^{n} r_i \cdot T_i \right) * s \right) * t = \sum_{i=1}^{n} r_i \cdot [\partial_y(\partial_x T_i * s) * t] \\
&= \sum_{i=1}^{n} r_i \cdot [\partial_x(\partial_y T_i * t) * s + \partial_x T_i * (\partial_y s * t)] \\
&= \sum_{i=1}^{n} r_i \cdot [\partial_x(\partial_y T_i * t) * s] + \sum_{i=1}^{n} r_i \cdot [\partial_x T_i * (\partial_y s * t)] \\
&= \partial_x (\partial_y \left( \sum_{i=1}^{n} r_i \cdot T_i \right) * t) * s + \partial_x \left( \sum_{i=1}^{n} r_i \cdot T_i \right) * (\partial_y s * t) \\
&= \partial_x (\partial_y \sigma * t) * s + \partial_x \sigma * (\partial_y s * t).
\end{aligned}
\tag{3.3.28}
$$

This completes the induction. ❏

**Corollary 3.3.16. Synthetic Schwarz Lemma**

Let $\sigma \in \mathcal{R}\langle \Lambda^!_{t,\mathcal{V}} \rangle$ be a complex Taylor poly-term, $s, t \in \mathcal{R}\langle \Lambda_{t,\mathcal{V}} \rangle$ be complex Taylor terms,

$x, y \in \mathcal{V}$ be variables. Suppose that $x \notin \mathrm{FV}(t)$ and $y \notin \mathrm{FV}(s)$. Then,

$$\partial_y(\partial_x \sigma * s) * t = \partial_x(\partial_y \sigma * t) * s. \tag{3.3.29}$$

♡

***Proof.*** As we can see, $\partial_y(\partial_x\sigma * s) * t = \partial_x(\partial_y\sigma * t) * s + \partial_x\sigma * (\partial_y s * t) = \partial_x(\partial_y\sigma * t) * s + \partial_x\sigma * 0 = \partial_x(\partial_y\sigma * t) * s + 0 = \partial_x(\partial_y\sigma * t) * s$, where we used Lemma 3.3.15 for the first equality and Lemma 3.3.14 for the second equality. ❏

Oftentimes, especially, when talking about substitution or partial differentiation, we are less interested in the existence of a variable, say $x \in \mathcal{V}$, in a complex Taylor poly-term, say $\sigma \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$, and more interested in the number of times $x$ appears in $\sigma$.

**Definition 3.3.17. Degree of a Taylor Term**

Given a variable $x \in \mathcal{V}$, define the function $\deg_x \colon \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle \to \mathbb{N}_0$ that takes a complex Taylor poly-term to a nonnegative number by:

$$\deg_x(x) = 1, \tag{3.3.30a}$$
$$\deg_x(y) = 0, \tag{3.3.30b}$$
$$\deg_x(\lambda x.s) = 0, \tag{3.3.30c}$$
$$\deg_x(\lambda y.s) = \deg_x(s), \tag{3.3.30d}$$
$$\deg_x(\langle s\rangle T) = \mathrm{FV}(s) + \mathrm{FV}(T), \tag{3.3.30e}$$
$$\deg_x(t_1 \dots t_n) = \sum_{i=1}^{n} \deg_x(t_i), \tag{3.3.30f}$$
$$\deg_x\left(\sum_{i=1}^{n} r_i \cdot S_i\right) = \sum_{i=1}^{n} \deg_x(S_i), \tag{3.3.30g}$$

where $y \in \mathcal{V}$ is a variable distinct from $x$, $n \in \mathbb{N}_0$ is a nonnegative integer, $s, t_1, \dots, t_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms, $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-terms, $r_1, \dots, r_n \in \mathcal{R}$ are scalars, and $S_1, \dots, S_n \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms. The image of a complex Taylor poly-term $\tau \in \Lambda^!_{t,\mathcal{V}}$ under FV is called the *free variables* of $\tau$.

♣

We, now, prove some elementary properties of the degree of Taylor terms.

**Lemma 3.3.18**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term and $x \in \mathcal{V}$. $\deg_x(S) = 0$ if and only if $x \notin \mathrm{FV}(S)$. ♡

***Proof.*** This a a simple induction on $S$. The details are left to the reader. ❏

**Lemma 3.3.19**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor (poly-)term and let $t \in \Lambda_{t,\mathcal{V}}$ be a simple Taylor term. Let $x \in \mathcal{V}$ be a variable and let $n_S \coloneqq \deg_x(S)$. Then, there exists simple Taylor (poly-)terms $S_1, \dots, S_{n_S} \in \Lambda^!_{t,\mathcal{V}}$ such that $\partial_x S * t = \sum_{i=1}^{n_S} S_i$ and $\deg_x(S_i) = \deg_x(S) + \deg_x(t) - 1$ for each $1 \leqslant i \leqslant n_S$.

♡

***Proof.*** We prove this via induction on $S$.

1. Assume that $S = x$. Then, $\deg_x(S) = \deg_x(x) = 1$ and $\partial_x S * t = \partial_x x * t = t$. Let $S_1 = t$. Then, $\deg_x(S) + \deg_x(t) - 1 = \deg_x(x) + \deg_x(t) - 1 = 1 + \deg_x(t) - 1 = \deg_x(t) = \deg_x(S_1)$.

2. Assume that $S = y$, where $y \in \mathcal{V}$ is a variable distinct from $x \in \mathcal{V}$. Then, $n_S = \deg_x(S) = 0$ and $\partial_x S * t = \partial_x y * t = 0$.

3. Assume that $S = \lambda y.s$, where $y \in \mathcal{V}$ is a fresh variable and $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. Let $n_s \coloneqq \deg_x(s)$. By the induction hypothesis, there exists simple Taylor terms $s_1, \dots, s_n \in \Lambda_{t,\mathcal{V}}$ such that $\partial_x s * t = \sum_{i=1}^{n_s} s_i$ and $\deg_x(s_i) = \deg_x(s) + \deg_x(t) - 1$ for each $1 \leqslant i \leqslant n$. Notice that $n_S = \deg_x(S) = \deg_x(\lambda y.s) = \deg_x(s) = n_s$ and $\partial_x S * t = \partial_x \lambda y.s * t = \lambda y.\partial_x s * t = \lambda y.\left(\sum_{i=1}^{n_s} s_i\right) = \sum_{i=1}^{n_S} \lambda y.s_i$. For each $1 \leqslant i \leqslant n_S = n_s$, define the simple Taylor term $S_i \coloneqq \lambda y.s_i \in \Lambda_{t,\mathcal{V}}$. Then, for each $1 \leqslant i \leqslant n$, $\deg_x(S_i) = \deg_x(\lambda y.s_i) = \deg_x(s_i) = \deg_x(s) + \deg_x(t) - 1 = \deg(\lambda y.s) + \deg_x(t) - 1 = \deg_x(S) + \deg_x(t) - 1$.

4. The case where $S = \langle s \rangle T$, where $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-term was shown in Lemma 2 of [9].

5. Assume that $S = s_1 \dots s_m$, where $m \in \mathbb{N}_0$ is a nonnegative integer and $s_1, \dots, s_m \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. For each $1 \leqslant i \leqslant m$, let $n_{s_i} \coloneqq \deg_x(s_i)$. By the induction hypothesis, for each $1 \leqslant i \leqslant m$, there exists simple Taylor terms $s_{i,1}, \dots, s_{i,n_{s_i}} \in \Lambda_{t,\mathcal{V}}$ such that $\partial_x s_i * t = \sum_{j=1}^{n_{s_i}} s_{i,j}$ and $\deg_x(s_{i,j}) = \deg_x(s_i) + \deg_x(t) - 1$ for $1 \leqslant j \leqslant n_{s_i}$. Notice that $n_S = \deg_x(S) = \deg_x(s_1 \dots s_m) = \deg_x(s_1) + \dots + \deg_x(s_m) = n_{s_1} + \dots + n_{s_m}$. As we can see, $\partial_x(S) * t = \partial_x(s_1 \dots s_m) * t = \sum_{i=1}^{m} s_1 \dots (\partial_x s_i * t) \dots s_m = \sum_{i=1}^{m} s_1 \dots \left(\sum_{j=1}^{n_{s_i}} s_{i,j}\right) \dots s_m = \sum_{i=1}^{m} \sum_{j=1}^{n_{s_i}} s_1 \dots s_{ij} \dots s_m$. For each $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n_{s_i}$, define the simple Taylor poly-term $S_{i,j} \coloneqq s_1 \dots s_{i-1} s_{i,j} s_{i+1} \dots s_m \in \Lambda^!_{t,\mathcal{V}}$. Then, for each $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n_{s_i}$,

$$\begin{aligned}
\deg_x(S_{i,j}) &= \deg_x(s_1 \dots s_{i-1} s_{i,j} s_{i+1} \dots s_m) = \deg_x(s_{ij}) + \sum_{i=1, i \neq j}^{m} \deg_x(s_i) \\
&= \deg_x(s_i) + \deg_x(t) - 1 + \sum_{i=1, i \neq j}^{m} \deg_x(s_i) \\
&= \sum_{i=1}^{m} \deg_x(s_i) + \deg_x(t) - 1 = \deg_x(s_1 \dots s_m) + \deg_x(t) - 1 \\
&= \deg_x(S) + \deg_x(t) - 1.
\end{aligned} \tag{3.3.31}$$

This completes the induction. ❏

**Lemma 3.3.20**

If $S \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-term such that $\deg_x(S) = 0$, then $S[0/x] = S$. ♡

***Proof.*** Since $\deg_x(S) = 0$, by Lemma 3.3.18, $x \notin \mathrm{FV}(S)$, which implies that $S[0/x] = \sigma$. ❏

**Lemma 3.3.21**

If $S$ is a (poly-)term such that $\deg_x(S) \geqslant 1$, then, $S[0/x] = 0$. ♡

***Proof.*** We prove this via induction on $S$.

1. Assume that $S = x$. Then,
$$S[0/x] = x[0/x] = 0. \tag{3.3.32}$$

2. Assume that $S = y$, where $y \in \mathcal{V}$ is a variable distinct from $x$. Then,
$$\deg_x(S) = \deg_x(y) = 0 < 1. \tag{3.3.33}$$

3. Assume that $S = \lambda y.s$, where $y \in \mathcal{V}$ is a fresh variable and $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. Since $y$ is fresh, $\deg_x(s) = \deg_x(\lambda y.s) = \deg_x(S)$, so by the induction hypothesis, $s[0/x] = 0$, which implies that
$$S[0/x] = (\lambda y.s)[0/x] = \lambda y.s[0/x] = \lambda y.0 = 0. \tag{3.3.34}$$

4. Assume that $S = \langle s \rangle T$, where $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-term. Since $1 \leqslant \deg_x(S) = \deg_x \langle s \rangle T \;= \deg_x(s) + \deg_x(T)$, $\deg_x(s) \geqslant 1$ or $\deg_x(T) \geqslant 1$. Suppose that $\deg_x(s) \geqslant 1$. Then, by the induction hypothesis, $s[0/x] = 0$, so
$$(S[0/x] = (\langle s \rangle T)[0/x] = \langle s[0/x] \rangle T[0/x] = \langle 0 \rangle T[0/x] = 0. \tag{3.3.35}$$
Instead, suppose that $\deg_x(T) \geqslant 1$. Then, by the induction hypothesis, $T[0/x] = 0$, so
$$S[0/x] = \langle s \rangle T)[0/x] = \langle s[0/x] \rangle T[0/x] = \langle s[0/x] \rangle 0 = 0. \tag{3.3.36}$$
In either case, $S[0/x] = 0$.

5. Assume that $S = s_1 \dots s_n$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $s_1, \dots, s_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. Since $1 \leqslant \deg_x(S) = \deg_x(s_1 \dots s_n) = \deg_x(s_1) + \dots + \deg_x(s_n)$, there must exist $1 \leqslant i_0 \leqslant n$ such that $\deg_x(s_{i_0}) \geqslant 1$. By the induction hypothesis, this implies that $s_{i_0}[0/x] = 0$, so
$$\begin{aligned} S[0/x] &= (s_1 \dots s_{i_0-1} s_{i_0} s_{i_0+1} \dots s_n)[0/x] \\ &= s_1[0/x] \dots s_{i_0-1}[0/x] s_{i_0}[0/x] s_{i_0+1}[0/x] \dots s_n[0/x] \\ &= s_1[0/x] \dots s_{i_0-1}[0/x] 0 s_{i_0+1}[0/x] \dots s_n[0/x] = 0. \end{aligned} \tag{3.3.37}$$

This completes the induction. ❏

**Lemma 3.3.22**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term and let $x, x_1, \dots, x_m \in \mathcal{V}$ be distinct variables. Then,
$$\deg_x(S[x/x_1, \dots, x_m]) = \deg_x(S) + \sum_{i=1}^{m} \deg_{x_i}(S). \tag{3.3.38}$$
♡

***Proof***. We prove this via induction on $S$.

1. Assume that $S = x$. Then,
$$\begin{aligned} \deg_x(S) + \sum_{i=1}^{m} \deg_{x_i}(S) = \deg_x(x) + \sum_{i=1}^{m} \deg_{x_i}(x) = 1 + \sum_{i=1}^{m} 0 = 1 = \deg_x(x) \\ = \deg_x(x[x/x_1, \dots, x_m]) = \deg_x(S[x/x_1, \dots, x_m]). \end{aligned} \tag{3.3.39}$$

2. Assume that $S = x_{i_0}$, where $1 \leqslant i_0 \leqslant m$. Then,

$$\begin{aligned}\deg_x(S) + \sum_{i=1}^{m} \deg_{x_i}(S) &= \deg_x(x_{i_0}) + \sum_{i=1}^{m} \deg_{x_i}(x_{i_0}) \\ &= \deg_x(x_{i_0}) + \sum_{i=1, i\neq i_0}^{m} \deg_{x_i}(x_{i_0}) + \deg_{x_{i_0}}(x_{i_0}) \\ &= 0 + \sum_{i=1,i\neq i_0}^{m} 0 + 1 = 1 = \deg_x(x) \\ &= \deg_x(x_{i_0}[x/x_1,\ldots,x_m]) = \deg_x(S[x/x_1,\ldots,x_m]).\end{aligned} \tag{3.3.40}$$

3. Assume that $S = y$, where $y \in \mathcal{V}$ is a variable distinct from $x, x_1, \ldots, x_m$. Then,

$$\begin{aligned}\deg_x(S) + \sum_{i=1}^{m} \deg_{x_i}(S) &= \deg_x(y) + \sum_{i=1}^{m} \deg_{x_i}(y) = 0 + \sum_{i=1}^{m} 0 = 0 = \deg_x(y) \\ &= \deg_x(y[x/x_1,\ldots,x_m]) = \deg_x(S[x/x_1,\ldots,x_m]).\end{aligned} \tag{3.3.41}$$

4. Assume that $S = \lambda y.s$, where $y \in \mathcal{V}$ is a fresh variable and $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. By the induction hypothesis, $\deg_x(s[x/x_1,\ldots,x_m]) = \deg_x(s) + \sum_{i=1}^{m} \deg_{x_i}(s)$, which implies that

$$\begin{aligned}\deg_x(S[x/x_1,\ldots,x_m]) &= \deg_x((\lambda y.s)[x/x_1,\ldots,x_m]) = \deg_x(\lambda y.s[x/x_1,\ldots,x_m]) \\ &= \deg_x(s[x/x_1,\ldots,x_m]) = \deg_x(s) + \sum_{i=1}^{m} \deg_{x_i}(s) \\ &= \deg_x(\lambda y.s) + \sum_{i=1}^{m} \deg_{x_i}(\lambda y.s) = \deg_x(S) + \sum_{i=1}^{m} \deg_{x_i}(S).\end{aligned} \tag{3.3.42}$$

5. Assume that $S = \langle s \rangle T$, where $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $T \in \Lambda^{!}_{t,\mathcal{V}}$ is a simple Taylor poly-term. By the induction hypothesis, $\deg_x(s[x/x_1,\ldots,x_m]) = \deg_x(s) + \sum_{i=1}^{m} \deg_{x_i}(s)$ and $\deg_x(T[x/x_1,\ldots,x_m]) = \deg_x(T) + \sum_{i=1}^{m} \deg_{x_i}(T)$, which implies that

$$\begin{aligned}\deg_x(S[x/x_1,\ldots,x_m]) &= \deg_x((\langle s \rangle T)[x/x_1,\ldots,x_m]) \\ &= \deg_x(\langle s[x/x_1,\ldots,x_m] \rangle T[x/x_1,\ldots,x_m]) \\ &= \deg_x(s[x/x_1,\ldots,x_m]) + \deg_x(T[x/x_1,\ldots,x_m]) \\ &= \deg_x(s) + \sum_{i=1}^{m} \deg_{x_i}(s) + \deg_x(T) + \sum_{i=1}^{m} \deg_{x_i}(T) \\ &= \deg_x(s) + \deg_x(T) + \sum_{i=1}^{m} \deg_{x_i}(s) + \deg_{x_i}(T) \\ &= \deg_x(\langle s \rangle T) + \sum_{i=1}^{m} \deg_{x_i}(\langle s \rangle T) \\ &= \deg_x(S) + \sum_{i=1}^{m} \deg_{x_i}(S).\end{aligned} \tag{3.3.43}$$

6. Assume that $S = s_1 \dots s_n$, where $s_1, \dots, s_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. By the induction hypothesis, $\deg_x(s_i[x/x_1, \dots, x_m]) = \deg_x(s_i) + \sum_{j=1}^m \deg_{x_j}(s_i)$ for all $1 \leqslant i \leqslant m$, which implies that

$$\begin{aligned}
\deg_x(S[x/x_1,\dots,x_m]) &= \deg_x((s_1 \dots s_m)[x/x_1,\dots,x_m]) \\
&= \deg_x(s_1[x/x_1,\dots,x_m] \dots s_m[x/x_1,\dots,x_m]) \\
&= \sum_{i=1}^m \deg_x(s_i[x/x_1,\dots,x_m]) \\
&= \sum_{i=1}^m [\deg_x(s_i) + \sum_{j=1}^m \deg_{x_j}(s_i)] \\
&= \sum_{i=1}^m \deg_x(s_i) + \sum_{i=1}^m \sum_{j=1}^m \deg_{x_j}(s_i) \\
&= \sum_{i=1}^m \deg_x(s_i) + \sum_{j=1}^m \sum_{i=1}^m \deg_{x_j}(s_i) \\
&= \deg_x(s_1 \dots s_m) + \sum_{j=1}^m \deg_{x_j}(s_1 \dots s_m) \\
&= \deg_x(S) + \sum_{j=1}^m \deg_{x_i}(S).
\end{aligned} \tag{3.3.44}$$

This completes the induction. ❏

Recall that differentiated abstraction syntactically represents the underlying analytic function given in (3.1.4). We now define the computation equivalent using the partial differentiation operation of differentiated abstraction, called *big step differentiation*.

**Definition 3.3.23. Big Step Differential**

Let $\sigma \in \mathcal{R}\langle \Lambda^!_{t,\mathcal{V}} \rangle$ be a complex Taylor poly-term, let $x \in \mathcal{V}$ be a variable, and let $T \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term such that $x \notin \mathrm{FV}(T)$. Then, there exists simple Taylor terms $t_1, \dots, t_n \in \Lambda_{t,\mathcal{V}}$, where $n = |T|$, such that $T = t_1 \dots t_n$ and $x \notin \mathrm{FV}(t_i)$ for all $1 \leqslant i \leqslant n$. Then, the *big step differential* of $\sigma$ with respect to $x$ times $T$ is

$$\partial_x(\sigma, T) \coloneqq (\partial_x(\cdots(\partial_x \sigma * t_1)\cdots) * t_n)[0/x]. \tag{3.3.45}$$

Note that the order of the partial derivatives does not matter by the Synthetic Schwartz Lemma (i.e., Corollary 3.3.16). ♣

Now, we prove some lemmas about big step differentiation, ultimately showing that degree of the complex Taylor poly-term is directly related to when its big step differentiation is zero.

**Lemma 3.3.24**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor (poly-)term and let $x \in \mathcal{V}$ be a variable. Define $n_S \coloneqq \deg_x(S)$.

let $k$ be an integer $\leqslant n_S$ and let $t_1, \dots, t_k$ be simple terms. Define

$$m_{S,k} = \prod_{i=0}^{k-1} \left( n_S + \sum_{j=1}^{i} \deg_x(t_j) - i \right) \tag{3.3.46}$$

Then, there exist simple Taylor (poly-)terms $S_1, \dots, S_{m_{S,k}} \in \Lambda^!_{t,\mathcal{V}}$ such that $\partial_x^k S * (t_1, \dots, t_k) = \sum_{i=1}^{m_{S,k}} S_i$ and $\deg_x(S_i) = \deg_x(S) + \sum_{j=1}^{k} \deg_x(t_j) - k$ for all $1 \leqslant i \leqslant m_{S,k}$. ♡

***Proof.*** This follows by induction on $k$.

1. The case where $k = 0$ is trivial.

2. The case where $k = 1$ was shown in Lemma 3.3.19.

3. Assume that $k = \ell + 1$, where $\ell \in \mathbb{N}$ is a positive integer. Define

$$m_{S,\ell} \coloneqq \prod_{i=0}^{\ell-1} \left( n_S + \sum_{j=1}^{i} \deg_x(t_j) - i \right). \tag{3.3.47}$$

By the induction hypothesis, there exist simple Taylor (poly-)terms $S_1, \dots, S_{m_{S,\ell}} \in \Lambda^!_{t,\mathcal{V}}$ such that $\partial_x^\ell S * (t_1, \dots, t_k) = \sum_{i=1}^{m_{S,\ell}} S_i$ and $\deg_x(S_i) = \deg_x(S) + \sum_{j=1}^{\ell} \deg_x(t_j) - \ell$ for all $1 \leqslant i \leqslant m_{S,\ell}$. Define $n' = \deg_x(S) + \sum_{j=1}^{\ell} \deg_x(t_j) - \ell$. For each $1 \leqslant i \leqslant m_{S,\ell}$, notice that $\deg_x(S_i) = \deg_x(S) + \sum_{j=1}^{\ell} \deg_x(t_j) - \ell = n'$, so by 3.3.19, there exist a simple Taylor (poly-)terms $S_{i,1}, \dots, S_{i,n'} \in \Lambda^!_{t,\mathcal{V}}$ such that $\partial_x S_i * t_{\ell+1} = \sum_{j=1}^{n'} S_{i,j}$ and $\deg_x(S_{i,j}) = \deg_x(S_i) + \deg_x(t_{\ell+1}) - 1$ for $1 \leqslant j \leqslant n'$. Hence,

$$\begin{aligned} \partial_x^{\ell+1} S * (t_1, \dots, t_\ell, t_{\ell+1}) &= \partial_x(\partial_x^\ell S * (t_1, \dots, t_\ell)) * t_{\ell+1} \\ &= \partial_x \left( \sum_{i=1}^{m_{S,\ell}} S_i \right) * t_{\ell+1} \\ &= \sum_{i=1}^{m_{S,\ell}} \partial_x S_i * t_{\ell+1} \\ &= \sum_{i=1}^{m_{S,\ell}} \sum_{j=1}^{n'} S_{i,j} \end{aligned} \tag{3.3.48}$$

and

$$\begin{aligned} \deg(S_{i,j}) &= \deg_x(S_i) + \deg_x(t_{\ell+1}) - 1 \\ &= \deg_x(S) + \sum_{j=1}^{\ell} \deg_x(t_j) - \ell + \deg_x(t_{\ell+1}) - 1 \\ &= \deg_x(S) + \sum_{j=1}^{\ell+1} \deg_x(t_j) - (\ell + 1) \\ &= \deg_x(S) + \sum_{j=1}^{k} \deg_x(t_j) - k \end{aligned} \tag{3.3.49}$$

for all $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n'$. Note that

$$\begin{aligned} m_{S,\ell} \cdot n' &= \prod_{i=0}^{\ell-1} \left( n_S + \sum_{j=1}^{i} \deg_x (t_j) - i \right) \cdot n' \\ &= \prod_{i=0}^{\ell-1} \left( n_S + \sum_{j=1}^{i} \deg_x (t_j) - i \right) \cdot \left( \deg_x (S) + \sum_{j=1}^{\ell} \deg_x (t_j) - \ell \right) \\ &= \prod_{i=0}^{\ell} \left( n_S + \sum_{j=1}^{i} \deg_x (t_j) - i \right) \\ &= m_{S,k}. \end{aligned} \tag{3.3.50}$$

This completes the induction. ❏

**Corollary 3.3.25**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, define $n_S \coloneqq \deg_x (S)$, and let $t_1, \ldots, t_{n_S} \in \Lambda_{t,\mathcal{V}}$ be simple Taylor terms, and let $x \in \mathcal{V}$ be a variable. Suppose that $x \notin \mathrm{FV}(t_i)$ for all $1 \leqslant i \leqslant \ell$. Then, $x \notin \mathrm{FV}(\partial_x^{n_S} S * (t_1, \ldots, t_{n_S}))$. ♡

***Proof***. Take any $1 \leqslant i \leqslant \ell$. Since $x \notin \mathrm{FV}(t_i)$, by Lemma 3.3.18, $\deg_x t_i = 0$. Since our choice of $1 \leqslant i \leqslant \ell$ was arbitrary, this implies that $\deg_x t_i = 0$ for all $1 \leqslant i \leqslant \ell$. By Lemma 3.3.24, there exists a nonnegative integer $m_{S,n_S} \in \mathbb{N}_0$ and simple Taylor poly-terms $S_1, \ldots, S_{m_{S,n_S}} \in \Lambda^!_{t,\mathcal{V}}$ such that $\partial_x^{n_S} S*(t_1, \ldots, t_{n_S}) = \sum_{i=1}^{m_{S,n_S}} S_i$ and $\deg_x (S_i) = \deg_x (S) + \sum_{j=1}^{k} \deg_x (t_j) - n_S = n_S + \sum_{j=1}^{k} 0 - n_S = 0$ for all $1 \leqslant i \leqslant m_{S,n_S}$. Take any $1 \leqslant i \leqslant m_{s,n_S}$. Since $\deg_x (S_i) = 0$, by Lemma 3.3.18, $x \notin \mathrm{FV}(S_i)$. Hence, $\mathrm{FV}(\partial_x^{n_S} S * (t_1, \ldots, t_{n_S})) = \mathrm{FV}(\sum_{i=1}^{m_{S,n_S}} S_i) = \cup_{i=1}^{m_{S,n_S}} \mathrm{FV}(S_i) \not\ni x$. ❏

**Corollary 3.3.26**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, $\ell \in \mathbb{N}$ be a positive integer, $t_1, \ldots, t_\ell \in \Lambda_{t,\mathcal{V}}$ be simple Taylor terms, and $x \in \mathcal{V}$ be a variable. Define $n_S \coloneqq \deg_x (S)$. Suppose that $n_S < \ell$ and $x \notin \mathrm{FV}(t_i)$ for all $1 \leqslant i \leqslant \ell$. Then, $\partial_x^{\ell} S * (t_1, \ldots, t_\ell) = 0$. ♡

***Proof***. By Corollary 3.3.25, $x \notin \mathrm{FV}(\partial_x^{n_S} S * (t_1, \ldots, t_{n_S}))$, so by Lemma 3.3.14,

$$\partial_x^{n_S+1} S * (t_1, \ldots, t_{n_S}, t_{n_S+1}) = 0. \tag{3.3.51}$$

Hence,

$$\begin{aligned} \partial_x^{\ell} S * (t_1, \ldots, t_\ell) &= \partial_x^{\ell - n_s - 2} (\partial_x^{n_s+1} S * (t_1, \ldots, t_{n_s}, t_{n_s+1})) * (t_{n_s+2}, \ldots, \ell) \\ &= \partial_x^{\ell - n_s - 2} 0 * (t_{n_s+2}, \ldots, \ell) \\ &= 0. \end{aligned} \tag{3.3.52}$$

This completes the proof. ❏

**Lemma 3.3.27**

Let $S, T \in \Lambda^!_{t,\mathcal{V}}$ be simple Taylor (poly-)term and $x \in \mathcal{V}$ be a variable. Suppose that $x \notin$

FV $(T)$. Define $n_S \coloneqq \deg_x(S)$ and $\ell \coloneqq |T|$. Then,

$$\partial_x(S,T) = \begin{Bmatrix} \partial_x^{n_S} S * (t_1, \ldots, t_{n_S}), & \text{if } n_S = \ell \\ 0, & \text{if } n_s \neq \ell \end{Bmatrix}. \tag{3.3.53}$$

***Proof.*** Since $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor (poly-)term, $\ell = |T|$, and $x \notin \mathrm{FV}(T)$, there exists simple Taylor terms $t_1, \ldots, t_\ell$ such that $T = t_1 \ldots t_\ell$ and $x \notin \mathrm{FV}(t_i)$ for all $1 \leqslant i \leqslant \ell$.

1. Assume that $n_s < \ell$. By Lemma 3.3.24, there exists a nonnegative integer $m_{S,\ell} \in \mathbb{N}_0$ and simple Taylor poly-terms $S_1, \ldots, S_{m_{S,\ell}} \in \Lambda^!_{t,\mathcal{V}}$ such that $\partial_x^\ell S * (t_1, \ldots, t_\ell) = \sum_{i=1}^{m_{S,\ell}} S_i$ and $\deg_x(S_i) = \deg_x(S) + \sum_{j=1}^k \deg_x(t_j) - \ell = n_s + \sum_{j=1}^k 0 - \ell \geqslant 1 > 0$ for all $1 \leqslant i \leqslant m_{S,\ell}$. Since $\deg_x(S_i) > 0$ for all $1 \leqslant i \leqslant m_{S,\ell}$, by Lemma 3.3.21, $S_i[0/x] = 0$ for all $1 \leqslant i \leqslant m_{S,\ell}$. Hence,

$$\partial_x(S,T) = (\partial_x^\ell S * (t_1, \ldots, t_\ell))[0/x] = \left(\sum_{i=1}^{m_{S,\ell}} S_i\right)[0/x] = \sum_{i=1}^{m_{S,\ell}} S_i[0/x] = \sum_{i=1}^{m_{S,\ell}} 0 = 0. \tag{3.3.54}$$

2. Assume that $n_s = \ell$. By Corollory 3.3.25, $x \notin \mathrm{FV}(\partial_x^{n_S} S * (t_1, \ldots, t_{n_S}))$, so by Lemmas 3.3.18 and 3.3.20,

$$\partial_x(S,T) = (\partial_x^{n_S} S * (t_1, \ldots, t_{n_S}))[0/x] = \partial_x^{n_S} S * (t_1, \ldots, t_{n_S}). \tag{3.3.55}$$

3. Assume that $n_S > \ell$. By Corollory 3.3.26, $\partial_x^\ell S * (t_1, \ldots, t_\ell) = 0$, which implies that

$$\partial_x(S,T) = (\partial_x^\ell S * (t_1, \ldots, t_\ell))[0/x] = 0[0/x] = 0. \tag{3.3.56}$$

This completes the proof. ❏

We, now, work toward one of the most remarkable results of Taylor λ-calculus, Theorem 3.3.30, the fact that the partial derivatives of a simple Taylor poly-term $S$ with respect to a variable $x$ times a simple Taylor term $t$ is a sum of $\deg_x S$ many elements, where each element in the sum is just $S$ where we replace one of the instances of $x$ with the simple Taylor term $t$. This is the reason we can think of partial differentiation as some kind-of non-deterministic substitution that conserves resources, and is the reason partial differentiation in Taylor λ-calculus is sometimes called *linear substitution* (see section 1.2.10 in [9]).

**Lemma 3.3.28**

let $S \in \Lambda^!_{t,\mathcal{V}}$ be a (poly-)term, $T \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, and $x \in \mathcal{V}$ be a variable. Define $n \coloneqq |T|$. Since $T \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, there exists $t_1, \ldots, t_n \in \Lambda_{t,\mathcal{V}}$ be a simple Taylor term such that $T = t_1 \ldots t_n$. Given $U \in \mathrm{supp}(\partial_x^n \sigma * (t_1, \ldots, t_n))$, $\mathrm{size}(U) = \mathrm{size}(S) + \mathrm{size}(T) - n$. ♡

***Proof.*** This follows by induction on $n$.

1. The case where $n = 0$ is trivial.

2. The case where $n = 1$ is just Lemma 3.3.12.

3. Assume that $n = k+1$, where $k \in \mathbb{N}$ is a positive integer. Since $U \in \operatorname{supp}(\partial_x^n S * (t_1, \ldots, t_n)) = \operatorname{supp}(\partial_x^{k+1} S * (t_1, \ldots, t_n)) = \operatorname{supp}(\partial_x(\partial_x^k S * (t_1, \ldots, t_k)) * t_{k+1})$, there exists simple Taylor (poly-)term $V \in \operatorname{supp}(\partial_x^k S * (t_1, \ldots, t_k))$ such that $U \in \operatorname{supp}(\partial_x V * t_{k+1})$. Since $V \in \operatorname{supp}(\partial_x^k S * (t_1, \ldots, t_k))$, by the induction hypothesis, $\operatorname{size}(V) = \operatorname{size}(S) + \operatorname{size}(t_1 \ldots t_k) - k$. Since $U \in \operatorname{supp}(\partial_x V * t_{k+1})$, by Lemma 3.3.12, $\operatorname{size}(U) = \operatorname{size}(V) + \operatorname{size}(t_{k+1}) - 1$. Hence,

$$\begin{aligned}\operatorname{size}(U) &= \operatorname{size}(V) + \operatorname{size}(t_{k+1}) - 1 = \operatorname{size}(S) + \operatorname{size}(t_1 \ldots t_k) - k + \operatorname{size}(t_{k+1}) - 1 \\ &= \operatorname{size}(S) + \operatorname{size}(t_1 \ldots t_k t_{k+1}) - (k+1) = \operatorname{size}(S) + \operatorname{size}(t_1 \ldots t_n) - n \\ &= \operatorname{size}(S) + \operatorname{size}(T) - n.\end{aligned} \tag{3.3.57}$$

This completes the induction. ❏

**Corollary 3.3.29**

Let $S, T \in \Lambda^!_{t,\mathcal{V}}$ be poly-terms and $x \in \mathcal{V}$ be a variable. Suppose that $x \notin \mathrm{FV}(T)$ and $\deg_x(S) = |T|$. Given $U \in \operatorname{supp}(\partial_x(\sigma, T))$, we have that $\operatorname{size}(U) = \operatorname{size}(S) + \operatorname{size}(T) - |T|$. ♡

***Proof***. Since $\deg_x(S) = |T|$, by Lemma 3.3.27, $\partial_x(S, T) = \partial_x^{|T|} S * T$. Hence, $U \in \operatorname{supp}(\partial_x(\sigma, T)) = \operatorname{supp}(\partial_x^{|T|} S * T)$, so by Lemma 3.3.28, $\operatorname{size}(U) = \operatorname{size}(S) + \operatorname{size}(T) - |T|$. ❏

**Theorem 3.3.30. Partial Differentiation is Linear Substitution**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, let $x \in \mathcal{V}$ be a variable, define $n_S := \deg_x(S)$, let $t \in \Lambda_{t,\mathcal{V}}$ be a simple Taylor term, let $x_1, \ldots, x_{n_S}$ be distinct variables distinct from $x$ and not free in $t$, and let $S'$ be the $x$-linearization of $S$ in $x_1, \ldots, x_{n_S}$. Then,

$$\partial_x S * t = \sum_{i=1}^{n_S} S'[t/x_i][x/x_1, \ldots, x_{n_S}]. \tag{3.3.58}$$

♡

***Proof***. We prove this by induction on $S$.

1. Assume that $S = x$. Then, $n_S = \deg_x(S) = \deg_x(x) = 1$, $S' = x_1$, and $\partial_x S * t = \partial_x x * t = t = t[x/x_1] = x_1[t/x_1][x/x_1] = S'[t/x_1][x/x_1] = \sum_{i=1}^{1} S'[t/x_i][x/x_1] = \sum_{i=1}^{n_S} S'[t/x_i][x/x_1]$.

2. Assume that $S = y$, where $y \in \mathcal{V}$ is a variable distinct from $x$. Then, $n_S = \deg_x(S) = \deg_x(y) = 0$, $S' = y$ and $\partial_x S * t = \partial_x y * t = 0$.

3. Assume that $S = \lambda y.s$, where $y \in \mathcal{V}$ is a fresh variable and $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. Define $n_s = \deg_x(s)$. Notice that $n_S = \deg_x(S) = \deg_x(\lambda y.s) = \deg_x(s) = n_s$. Let $s'$ be the $x$-linearization of $s$ in $x_1, \ldots, x_{n_s}$. By the induction hypothesis, $\partial_x s * t = \sum_{i=1}^{n_S} s'[t/x_i][x/x_1, \ldots, x_{n_S}]$. Since $S' = \lambda y.s'$, this implies that

$$\begin{aligned}\partial_x S * t = \partial_x(\lambda y.s) * t = \lambda y.(\partial_x s * t) &= \lambda y.\sum_{i=1}^{n} s'[t/x_i][x/x_1, \ldots, x_{n_S}] \\ &= \sum_{i=1}^{n} \lambda y.s'[t/x_i][x/x_1, \ldots, x_{n_S}] = \sum_{i=1}^{n} (\lambda y.s')[t/x_i][x/x_1, \ldots, x_{n_S}] \\ &= \sum_{i=1}^{n} S'[t/x_i][x/x_1, \ldots, x_{n_S}].\end{aligned} \tag{3.3.59}$$

4. Assume that $S = \langle s \rangle T$, where $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $T \in \Lambda^{!}_{t,\mathcal{V}}$ is a simple Taylor poly-term. Define $n_s = \deg_x(s)$ and $n_T = \deg_x(T)$. Notice that $n_S = \deg_x(S) = \deg_x(\langle s \rangle T) = \deg_x(s) + \deg_{(T)} = n_s + n_T$. Let $s'$ be the $x$-linearization of $s$ in $x_1, \ldots, x_{n_s}$ and $t'$ be the $x$-lineraization of $t$ in $x_{n_s+1}, \ldots, x_{n_S}$. By the induction hypothesis, $\partial_x s * t = \sum_{i=1}^{n_s} s'[t/x_i][x/x_1, \ldots, x_{n_s}]$ and $\partial_x T * t = \sum_{i=n_s+1}^{n_S} T'[t/x_i][x/x_{n_s+1}, \ldots, x_{n_S}]$. Hence,

$$\begin{aligned}
\partial_x S * t &= \partial_x \langle s \rangle T * t = \langle \partial_x s * t \rangle T + \langle s \rangle (\partial_x T * t) \\
&= \langle \sum_{i=n_s+1}^{n_S} s'[t/x_i][x/x_1, \ldots, x_{n_s}] \rangle T \\
&+ \langle s \rangle \left( \sum_{i=n_s+1}^{n_S} T'[t/x_i][x/x_{n_s+1}, \ldots, x_{n_S}] \right) \\
&= \sum_{i=n_s+1}^{n_S} \langle s'[t/x_i][x/x_1, \ldots, x_{n_s}] \rangle T \\
&+ \sum_{i=n_s+1}^{n_S} \langle s \rangle (T'[t/x_i][x/x_{n_s+1}, \ldots, x_{n_S}]) \\
&= \sum_{i=1}^{n_s} \langle s'[t/x_i][x/x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}] \rangle T \\
&\quad + \sum_{i=n_s+1}^{n_S} \langle s \rangle (T'[t/x_i][x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}]) \\
&= \sum_{i=1}^{n_s} \langle s'[t/x_i][x/x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}] \rangle \\
&\quad (T'[t/x_i][x/x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}]) \\
&\quad + \sum_{i=n_s+1}^{n_S} \langle s'[t/x_i][x/x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}] \rangle \\
&\quad (T'[t/x_i][x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}]) \\
&= \sum_{i=1}^{n_s} (\langle s' \rangle T')[t/x_i][x/x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}] \\
&\quad + \sum_{i=n_s+1}^{n_S} (\langle s' \rangle T')[t/x_i][x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}]) \\
&= \sum_{i=1}^{n_s} S'[t/x_i][x/x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}] \\
&\quad + \sum_{i=n_s+1}^{n_S} S'[t/x_i][x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}]) \\
&= \sum_{i=1}^{n_S} S'[t/x_i][x/x_1, \ldots, x_{n_s}, x_{n_s+1}, \ldots, x_{n_S}].
\end{aligned} \tag{3.3.60}$$

5. Assume that $S = s_1 \ldots s_m$, where $n \in \mathbb{N}_0$ is a nonegative integer and $s_1, \ldots, s_m \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. For each $1 \leqslant i \leqslant m$ Define $n_{s_i} = \deg_x(s_i)$. For each $0 \leqslant i \leqslant m$, let $k_i = \sum_{j=1}^{i} n_{s_i}$. Notice that $n_S = \deg_x(S) = \deg_x(s_i \ldots s_m) = \sum_{j=1}^{m} \deg_x(s_i) = \sum_{j=1}^{m} n_{s_i} =$

$k_m$. For each $1 \leqslant i \leqslant m$, let $s'_i$ be the $x$-lineraization of $s_i$ in $x_{k_{i-1}}, \ldots, x_{k_i}$. By the induction hypothesis, for each $1 \leqslant i \leqslant m$, $\partial_x s_i * t = \sum_{j=k_{i-1}}^{k_i} s'_i[t/x_j][x/x_{k_{i-1}}, \ldots, x_{i,k_i}]$. Hence,

$$\begin{aligned}
\partial_x S * t = \partial_x s_1 \ldots s_{n_S} * t &= \sum_{i=1}^{m} s_1 \ldots (\partial_x s_i * t) \ldots s_{n_S} \\
&= \sum_{i=1}^{m} s_1 \ldots \left( ( \sum_{j=k_{i-1}}^{k_i} s'_i[t/x_i][x/x_{k_{i-1}}, \ldots, x_{k_i}] \right) \ldots s_{n_S} \\
&= \sum_{i=1}^{m} \sum_{j=k_{i-1}}^{k_i} s_1 \ldots (s'_i[t/x_i][x/x_{k_{i-1}}, \ldots, x_{k_i}]) \ldots s_{n_S} \\
&= \sum_{i=1}^{m} \sum_{j=k_{i-1}}^{k_i} s_1 \ldots (s'_i[t/x_i][x/x_1, \ldots, x_{k_m}]) \ldots s_{n_S} \\
&= \sum_{i=1}^{m} \sum_{j=k_{i-1}}^{k_i} s'_1[t/x_i][x/x_1, \ldots, x_{k_m}] \ldots (s'_i[t/x_i][x/x_1, \ldots, x_{k_m}]) \\
&\qquad \ldots s'_{n_S}[t/x_i][x/x_1, \ldots, x_{k_m}] \\
&= \sum_{i=1}^{m} \sum_{j=k_{i-1}}^{k_i} (s'_1 \ldots s'_i \ldots s'_{n_S})[t/x_i][x/x_1, \ldots, x_{k_m}] \\
&= \sum_{i=1}^{m} \sum_{j=k_{i-1}}^{k_i} S'[t/x_i][x/x_1, \ldots, x_{n_S}] \\
&= \sum_{i=1}^{n_S} S'[t/x_i][x/x_1, \ldots, x_{n_S}].
\end{aligned} \tag{3.3.61}$$

This completes the induction. ❏

**Corollary 3.3.31**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, let $x \in \mathcal{V}$ be a variable, define $n_S \coloneqq \deg_x(S)$, let $m$ be a positive integer such that $m \leqslant n$ let $t_1, \ldots, t_m \in \Lambda_{t,\mathcal{V}}$ be a simple Taylor term, let $x_1, \ldots, x_{n_S}$ be distinct variables distinct from $x$ and not free in $t_1, \ldots, t_m$, let $S'$ be the $x$-linearization of $S$ in $x_1, \ldots, x_{n_S}$, and let $\mathcal{A}_m$ be the set of injections from $1, \ldots, m$ to $1, \ldots, n$. Then,

$$\partial_x^m S * (t_1, \ldots, t_m) = \sum_{f \in \mathcal{A}_m} S'[t_1/x_{f(1)}, \ldots, t_m/x_{f(m)}][x/x_1, \ldots, x_{n_S}]. \tag{3.3.62}$$

♡

***Proof***. We prove this via induction on $m$.

1. The case where $m = 1$ is done in Theorem 3.3.30.

2. Assume that $m = k + 1$, where $k$ is some positive integer. Then,

$$\begin{aligned}
\partial_x^m S * (t_1, \ldots, t_m) &= \partial_x^{k+1} S * (t_1, \ldots, t_{k+1}) \\
&= \partial_x \left( \partial_x^k S * (t_1, \ldots, t_k) \right) * t_{k+1}
\end{aligned} \tag{3.3.63}$$

$$= \partial_x \left( \sum_{f \in \mathcal{A}_k} S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][x/x_1, \ldots, x_{n_S}] \right) * t_{k+1}$$
$$= \sum_{f \in \mathcal{A}_k} \partial_x (S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][x/x_1, \ldots, x_{n_S}]) * t_{k+1},$$

where we used the induction hypothesis in the third equality. Given $f \in \mathcal{A}_m$, notice that

$$\deg_x (S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][x/x_1, \ldots, x_{n_S}]) = n_S - k. \tag{3.3.64}$$

Now, let us define the simple Taylor poly-term $(S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][x/x_1, \ldots, x_{n_S}])'$ to be the $x$-linearization of $S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][x/x_1, \ldots, x_{n_S}]$ in $\{x_j \coloneqq [n_S] - f([k])\}$. Trivially,

$$(S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][x/x_1, \ldots, x_{n_S}])' = S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}] \tag{3.3.65}$$

Thus, by Theorem 3.3.30,

$$\begin{aligned}
\partial_x^m S * (t_1, \ldots, t_m) &= \sum_{f \in \mathcal{A}_k} \partial_x (S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][x/x_1, \ldots, x_{n_S}]) * t_{k+1} \\
&= \sum_{f \in \mathcal{A}_k} \sum_{j \in [n_S] - f([k])} (S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][x/x_1, \ldots, x_{n_S}])' \\
&\qquad [t_{k+1}/x_j][x/x_1, \ldots, x_{n_S}] \\
&= \sum_{f \in \mathcal{A}_k} \sum_{j \in [n_S] - f([k])} S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}][t_{k+1}/x_j] \\
&\qquad [x/x_1, \ldots, x_{n_S}] \\
&= \sum_{f \in \mathcal{A}_k} \sum_{j \in [n_S] - f([k])} S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}, t_{k+1}/x_j] \\
&\qquad [x/x_1, \ldots, x_{n_S}] \\
&= \sum_{f \in \mathcal{A}_{k+1}} S'[t_1/x_{f(1)}, \ldots, t_k/x_{f(k)}, t_{k+1}/x_{f(k+1)}][x/x_1, \ldots, x_{n_S}] \\
&= \sum_{f \in \mathcal{A}_m} S'[t_1/x_{f(1)}, \ldots, t_m/x_{f(m)}][x/x_1, \ldots, x_{n_S}].
\end{aligned} \tag{3.3.66}$$

This completes the induction. ❏

**Corollary 3.3.32**

Let $S, T \in \Lambda^!_{t,\mathcal{V}}$ be simple Taylor poly-terms, let $x \in \mathcal{V}$ be a variable, define $n_S \coloneqq \deg_x (S)$, let $x_1, \ldots, x_{n_S}$ be distinct variables distinct from $x$, and let $S'$ be the $x$-linearization of $S$ in $x_1, \ldots, x_{n_S}$. Suppose that $x, x_1, \ldots, x_n \notin \mathrm{FV}(T)$ and $\deg_x (S) = |T|$. Then,

$$\partial_x (S, T) = \sum_{f \in \mathfrak{S}_{n_S}} S'[t_1/x_{f(1)}, \ldots, t_m/x_{f(n_S)}]. \tag{3.3.67}$$

♡

***Proof.*** Since $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor (poly-)term and $|T| = n_s$, there exists simple Taylor terms $t_1, \ldots, t_\ell$ such that $T = t_1 \ldots t_\ell$. As we can see,

$$\begin{aligned}\partial_x(S,T) &= \partial_x^{n_S} S * (t_1, \ldots, t_{n_S}) \\ &= \sum_{f \in \mathcal{A}_{n_S}} S'[t_1/x_{f(1)}, \ldots, t_{n_S}/x_{f(n_S)}][x/x_1, \ldots, x_{n_S}] \\ &= \sum_{f \in \mathfrak{S}_{n_S}} S'[t_1/x_{f(1)}, \ldots, t_{n_S}/x_{f(n_S)}],\end{aligned} \tag{3.3.68}$$

where we used the fact that $\deg_x(S) = |T|$ and Lemma 3.3.27 for the first equality; Corollory 3.3.31 for the second equality; and the fact that $x_1, \ldots, x_n \notin \mathrm{FV}(T)$ for the third equality. ❏

To conclude this section on Taylor $\lambda$-calculus and the partial derivative operator for it, we prove a synthetic version of the Generalized Leibniz Rule.

**Lemma 3.3.33. Synthetic Leibniz Rule**

Let $\sigma \in \mathcal{R}\langle \Lambda^!_{t,\mathcal{V}} \rangle$ be a complex poly-term, $t$ be a simple Taylor term, $x, x_1, x_2 \in \mathcal{V}$ be a variable. Suppose that $x_1 \neq x_2$, $x \notin \mathrm{FV}(\sigma)$, and $x_1, x_2 \notin \mathrm{FV}(t)$. Then,

$$\partial_x \sigma[x/x_1, x_2] * t = (\partial_{x_1}\sigma * t)[x/x_1, x_2] + (\partial_{x_2}\sigma * t)[x/x_1, x_2]. \tag{3.3.69}$$

♡

***Proof***. We prove this via induction on $\sigma$.

1. Assume that $\sigma = y$, where $y \in \mathcal{V}$ is a variable distinct from $x, x_1, x_2$. Then,

$$\begin{aligned}\partial_x \sigma[x/x_1, x_2] * t &= \partial_x y[x/x_1, x_2] * t = \partial_x y * t = 0 = 0 + 0 \\ &= 0[x/x_1, x_2] + 0[x/x_1, x_2] \\ &= (\partial_{x_1} y * t)[x/x_1, x_2] + (\partial_{x_2} y * t)[x/x_1, x_2] \\ &= (\partial_{x_1}\sigma * t)[x/x_1, x_2] + (\partial_{x_2}\sigma * t)[x/x_1, x_2].\end{aligned} \tag{3.3.70}$$

2. Assume that $\sigma = x_1$ and $x_1$ is distinct from $x$. Then,

$$\begin{aligned}\partial_x \sigma[x/x_1, x_2] * t &= \partial_x y[x/x_1, x_2] * t = \partial_x x_1[x/x_1, x_2] * t = \partial_x x * t \\ &= t = t + 0 = t[x/x_1, x_2] + 0[x/x_1, x_2] \\ &= (\partial_{x_1} x_1 * t)[x/x_1, x_2] + (\partial_{x_2} x_1 * t)[x/x_1, x_2] \\ &= (\partial_{x_1} y * t)[x/x_1, x_2] + (\partial_{x_2} y * t)[x/x_1, x_2] \\ &= (\partial_{x_1}\sigma * t)[x/x_1, x_2] + (\partial_{x_2}\sigma * t)[x/x_1, x_2].\end{aligned} \tag{3.3.71}$$

3. The case where $y = x_2$ and $x_2$ is distinct form $x$ is similar to the prior case.

4. Assume that $\sigma = \lambda y.s$, where $y$ is a fresh variable and $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. By the induction hypothesis, $\partial_x s[x/x_1, x_2] * t = (\partial_{x_1} s * t)[x/x_1, x_2] + (\partial_{x_2} s * t)[x/x_1, x_2]$, which implies that

$$\begin{aligned}\partial_x \sigma[x/x_1, x_2] * t &= \partial_x(\lambda y.s)[x/x_1, x_2] * t = \lambda y.(\partial_x s[x/x_1, x_2] * t) \\ &= \lambda y.((\partial_{x_1} s * t)[x/x_1, x_2] + (\partial_{x_2} s * t)[x/x_1, x_2]) \\ &= \lambda y.(\partial_{x_1} s * t)[x/x_1, x_2] + \lambda y.(\partial_{x_2} s * t)[x/x_1, x_2] \\ &= (\partial_{x_1} \lambda y.s * t)[x/x_1, x_2] + (\partial_{x_2} \lambda y.s * t)[x/x_1, x_2] \\ &= (\partial_{x_1}\sigma * t)[x/x_1, x_2] + (\partial_{x_2}\sigma * t)[x/x_1, x_2].\end{aligned} \tag{3.3.72}$$

5. Assume that $\sigma = \langle s \rangle T$, where $s \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term and $T \in \Lambda^!_{t,\mathcal{V}}$ is a simple Taylor poly-term. By the induction hypothesis, $\partial_x s[x/x_1,x_2] * t = (\partial_{x_1} s * t)[x/x_1,x_2] + (\partial_{x_2} s * t)[x/x_1,x_2]$ and $\partial_x T[x/x_1,x_2] * s = (\partial_{x_1} T * t)[x/x_1,x_2] + (\partial_{x_2} T * t)[x/x_1,x_2]$, which implies that

$$\begin{aligned}
\partial_x \sigma[x/x_1,x_2] * t &= \partial_x(\langle s \rangle T)[x/x_1,x_2] * t \\
&= \partial_x \langle s[x/x_1,x_2] \rangle T[x/x_1,x_2] * t \\
&= \langle \partial_x s[x/x_1,x_2] * t \rangle T[x/x_1,x_2] + \langle s[x/x_1,x_2] \rangle (\partial_x T[x/x_1,x_2] * t) \\
&= \langle (\partial_{x_1} s * t)[x/x_1,x_2] + (\partial_{x_2} s * t)[x/x_1,x_2] \rangle T[x/x_1,x_2] \\
&\qquad + \langle s[x/x_1,x_2] \rangle ((\partial_{x_1} T * t)[x/x_1,x_2] + (\partial_{x_2} T * t)[x/x_1,x_2]) \\
&= \langle (\partial_{x_1} s * t)[x/x_1,x_2] \rangle T[x/x_1,x_2] \\
&\qquad + \langle (\partial_{x_2} s * t)[x/x_1,x_2] \rangle T[x/x_1,x_2] \\
&\qquad + \langle s[x/x_1,x_2] \rangle (\partial_{x_1} T * t)[x/x_1,x_2] \\
&\qquad + \langle s[x/x_1,x_2] \rangle (\partial_{x_2} T * t)[x/x_1,x_2] \\
&= (\langle (\partial_{x_1} s * t) \rangle T)[x/x_1,x_2] + (\langle (\partial_{x_2} s * t) \rangle T)[x/x_1,x_2] \\
&\qquad + (\langle s \rangle (\partial_{x_1} T * t))[x/x_1,x_2] + (\langle s \rangle (\partial_{x_2} T * t))[x/x_1,x_2] \\
&= (\langle (\partial_{x_1} s * t) \rangle T)[x/x_1,x_2] + (\langle s \rangle (\partial_{x_1} T * t))[x/x_1,x_2] \\
&\qquad + (\langle (\partial_{x_2} s * t) \rangle T)[x/x_1,x_2] + (\langle s \rangle (\partial_{x_2} T * t))[x/x_1,x_2] \\
&= (\partial_{x_1} \langle s \rangle T * t)[x/x_1,x_2] + (\partial_{x_2} \langle s \rangle T * t)[x/x_1,x_2] \\
&= (\partial_{x_1} \sigma * t)[x/x_1,x_2] + (\partial_{x_2} \sigma * t)[x/x_1,x_2].
\end{aligned} \tag{3.3.73}$$

6. Assume that $\sigma = s_1 \dots s_n$. For each $1 \leqslant i \leqslant n$, by the induction hypothesis, $\partial_x si[x/x_1,x_2] * t = (\partial_{x_1} si * t)[x/x_1,x_2] + (\partial_{x_2} si * t)[x/x_1,x_2]$. Hence,

$$\begin{aligned}
\partial_x \sigma[x/x_1,x_2] * t &= \partial_x(s_1 \dots s_n)[x/x_1,x_2] * t = \partial_x s_1[x/x_1,x_2] \dots s_n[x/x_1,x_2] * t \\
&= \sum_{i=1}^{n} s_1[x/x_1,x_2] \dots (\partial_x s_i[x/x_1,x_2] * t) \dots s_n[x/x_1,x_2] \\
&= \sum_{i=1}^{n} (s_1[x/x_1,x_2] \dots ((\partial_{x_1} s_i * t)[x/x_1,x_2] \\
&\qquad + (\partial_{x_2} s_i * t)[x/x_1,x_2]) \dots s_n[x/x_1,x_2]) \\
&= \sum_{i=1}^{n} s_1[x/x_1,x_2] \dots (\partial_{x_1} s_i * t)[x/x_1,x_2] \dots s_n[x/x_1,x_2] \\
&\qquad + \sum_{i=1}^{n} s_1[x/x_1,x_2] \dots (\partial_{x_2} s_i * t)[x/x_1,x_2] \dots s_n[x/x_1,x_2] \\
&= \left( \sum_{i=1}^{n} s_1 \dots (\partial_{x_1} s_i * t) \dots s_n \right) [x/x_1,x_2] \\
&\qquad + \left( \sum_{i=1}^{n} s_1 \dots (\partial_{x_2} s_i * t) \dots s_n \right) [x/x_1,x_2] \\
&= (\partial_{x_1} s_1 \dots s_n * t)[x/x_1,x_2] + (\partial_{x_2} s_1 \dots s_n * t)[x/x_1,x_2] \\
&= (\partial_{x_1} \sigma * t)[x/x_1,x_2] + (\partial_{x_2} \sigma * t)[x/x_1,x_2].
\end{aligned} \tag{3.3.74}$$

7. Assume that $\sigma = \sum_{i=1}^{n} r_i \cdot T_i$, where $n \in \mathbb{N}_0$ is a nonnegative integer, $r_1, \ldots, r_n$ are scalars, and $T_1, \ldots, T_n \in \Lambda^!_{t,\mathcal{V}}$ are simple Taylor poly-terms. For each $1 \leqslant i \leqslant n$, by the induction hypothesis, $\partial_x T_i[x/x_1, x_2] * t = (\partial_{x_1}\sigma * t)[x/x_1, x_2] + (\partial_{x_2} T_i * t)[x/x_1, x_2]$. Hence,

$$\begin{aligned}
\partial_x \sigma[x/x_1, x_2] * t &= \partial_x \left(\sum_{i=1}^{n} r_i \cdot T_i\right)[x/x_1, x_2] * t = \sum_{i=1}^{n} r_i \cdot (\partial_x T_i[x/x_1, x_2] * t) \\
&= \sum_{i=1}^{n} r_i \cdot ((\partial_{x_1}\sigma * t)[x/x_1, x_2] + (\partial_{x_2} T_i * t)[x/x_1, x_2]) \\
&= \sum_{i=1}^{n} r_i \cdot (\partial_{x_1}\sigma * t)[x/x_1, x_2] + \sum_{i=1}^{n} r_i \cdot (\partial_{x_2} T_i * t)[x/x_1, x_2] \\
&= \left(\partial_{x_1}\left(\sum_{i=1}^{n} r_i \cdot T_i\right) * t\right)[x/x_1, x_2] \\
&\quad + \left(\partial_{x_2}\left(\sum_{i=1}^{n} r_i \cdot T_i\right) * t\right)[x/x_1, x_2] \\
&= (\partial_{x_1}\sigma * t)[x/x_1, x_2] + (\partial_{x_2}\sigma * t)[x/x_1, x_2].
\end{aligned} \tag{3.3.75}$$

This completes the induction. ❑

**Corollary 3.3.34. Generalized Synthetic Leibniz Rule**

Let $\sigma \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ be a complex Taylor poly-term, $T \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term, and $x, x_1, x_2 \in \mathcal{V}$ be variables. Suppose that $x_1$ and $x_2$ are distinct, $x \notin \mathrm{FV}(\sigma)$, and $x, x_1, x_2 \notin \mathrm{FV}(T)$. Then,

$$\partial_x^{|T|}\sigma[x/x_1, x_2] * T = \left(\sum_{UV=T}\binom{T}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|V|}\sigma * V) * U\right)[x/x_1, x_2]. \tag{3.3.76}$$

♡

***Proof***. We prove this via induction on $|T|$.

1. The case where $|T| = 1$ was done in Lemma 3.3.33.

2. Assume that $|T| > 1$. Then, there exist a simple Taylor poly-term $S \in \Lambda^!_{t,\mathcal{V}}$ and a simple Taylor term $t \in \Lambda_{t,\mathcal{V}}$ such that $T = tS$. As we can see,

$$\begin{aligned}
\partial_x^{|T|}\sigma[x/x_1, s_2] * T &= \partial_x^{|tS|}\sigma[x/x_1, s_2] * tS \\
&= \partial_x^{|S|}(\partial_x\sigma[x/x_1, x_2] * t) * S \\
&= \partial_x^{|S|}((\partial_{x_1}\sigma * t)[x/x_1, x_2] + (\partial_{x_2}\sigma * t)[x/x_1, x_2]) * S \\
&= \partial_x^{|S|}((\partial_{x_1}\sigma * t)[x/x_1, x_2]) * S + \partial_x^{|S|}((\partial_{x_2}\sigma * t)[x/x_1, x_2]) * S \\
&= \left(\sum_{UV=S}\binom{S}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|V|}(\partial_{x_1}\sigma * t) * V) * U\right)[x/x_1, x_2] \\
&\quad + \left(\sum_{UV=S}\binom{S}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|V|}(\partial_{x_2}\sigma * t) * V) * U\right)[x/x_1, x_2] \\
&= \left(\sum_{UV=S}\binom{S}{U}\partial_{x_1}^{|tU|}(\partial_{x_2}^{|V|}\sigma * V) * tU\right)[x/x_1, x_2]
\end{aligned} \tag{3.3.77}$$

$$
\begin{aligned}
&+\left(\sum_{UV=S}\binom{S}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|tV|}\sigma * tV) * U\right)[x/x_1,x_2] \\
&=\left(\sum_{tUV=tS}\binom{tS-t}{tU-t}\partial_{x_1}^{|tU|}(\partial_{x_2}^{|V|}\sigma * V) * tU\right)[x/x_1,x_2] \\
&\quad+\left(\sum_{UtV=tS}\binom{tS-t}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|tV|}\sigma * tV) * U\right)[x/x_1,x_2] \\
&=\left(\sum_{\substack{\tilde{U}V=T\\ t\in\tilde{U}}}\binom{T-t}{\tilde{U}-t}\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma * V) * \tilde{U}\right)[x/x_1,x_2] \\
&\quad+\left(\sum_{\substack{U\tilde{V}=T\\ t\in\tilde{V}}}\binom{T-t}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|\tilde{V}|}\sigma * \tilde{V}) * U\right)[x/x_1,x_2] \\
&=\left(\sum_{\substack{\tilde{U}V=T\\ t\in\tilde{U},t\in V}}\binom{T-t}{\tilde{U}-t}\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma * V) * \tilde{U}\right)[x/x_1,x_2] \\
&\quad+\left(\sum_{\substack{\tilde{U}V=T\\ t\in\tilde{U},t\notin V}}\binom{T-t}{\tilde{U}-t}\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma * V) * \tilde{U}\right)[x/x_1,x_2] \\
&\quad+\left(\sum_{\substack{U\tilde{V}=T\\ t\in U,t\in\tilde{V}}}\binom{T-t}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|\tilde{V}|}\sigma * \tilde{V}) * U\right)[x/x_1,x_2]+ \\
&\quad\left(\sum_{\substack{U\tilde{V}=T\\ t\notin U,t\in\tilde{V}}}\binom{T-t}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|\tilde{V}|}\sigma * \tilde{V}) * U\right)[x/x_1,x_2] \\
&=\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\in\tilde{U},t\in\tilde{V}}}\binom{T-t}{\tilde{U}-t}\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma * \tilde{V}) * \tilde{U}\right)[x/x_1,x_2] \\
&\quad+\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\in\tilde{U},t\notin\tilde{V}}}\binom{T-t}{\tilde{U}-t}\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma * \tilde{V}) * \tilde{U}\right)[x/x_1,x_2] \\
&\quad+\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\in\tilde{U},t\in\tilde{V}}}\binom{T-t}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|\tilde{V}|}\sigma * \tilde{V}) * \tilde{U}\right)[x/x_1,x_2]
\end{aligned}
$$

$$
\begin{aligned}
&+\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\notin\tilde{U},t\in\tilde{V}}}\binom{T-t}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|\tilde{V}|}\sigma*\tilde{V})*\tilde{U}\right)[x/x_1,x_2]\\
=&\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\in\tilde{U},t\in\tilde{V}}}\left[\binom{T-t}{\tilde{U}-t}+\binom{T-t}{U}\right]\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma*\tilde{V})*\tilde{U}\right)[x/x_1,x_2]\\
&+\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\in\tilde{U},t\notin\tilde{V}}}\binom{T-t}{\tilde{U}-t}\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma*\tilde{V})*\tilde{U}\right)[x/x_1,x_2]\\
&+\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\notin\tilde{U},t\in\tilde{V}}}\binom{T-t}{U}\partial_{x_1}^{|U|}(\partial_{x_2}^{|\tilde{V}|}\sigma*\tilde{V})*\tilde{U}\right)[x/x_1,x_2]\\
=&\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\in\tilde{U},t\in\tilde{V}}}\binom{T}{\tilde{U}}\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma*\tilde{V})*\tilde{U}\right)[x/x_1,x_2]\\
&+\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\in\tilde{U},t\notin\tilde{V}}}\binom{T}{\tilde{U}}\partial_{x_1}^{|\tilde{U}|}(\partial_{x_2}^{|V|}\sigma*\tilde{V})*\tilde{U}\right)[x/x_1,x_2]\\
&+\left(\sum_{\substack{\tilde{U}\tilde{V}=T\\ t\notin\tilde{U},t\in\tilde{V}}}\binom{T}{\tilde{U}}\partial_{x_1}^{|U|}(\partial_{x_2}^{|\tilde{V}|}\sigma*\tilde{V})*\tilde{U}\right)[x/x_1,x_2]\\
=&\left(\sum_{\tilde{U}\tilde{V}=T}\binom{T}{\tilde{U}}\partial_{x_1}^{|U|}(\partial_{x_2}^{|\tilde{V}|}\sigma*\tilde{V})*\tilde{U}\right)[x/x_1,x_2],
\end{aligned}
$$

where we used Lemma 3.3.33 for the third equality, the induction hypothesis for the fifth equality, Corollary 3.3.16 for the sixth equality, and Lemmas 3.2.8-3.2.10 for the second to last equality.

This completes the induction. ❏

In summary, in this section we have defined Taylor pre-term, $\alpha$-equivalence; Taylor terms (which is just Taylor pre-term quotiented by $\alpha$-equivalence), substitution, partial differentiation, and big step differentiation (which is the computational version of differentiated application) and proved a synthetic versions of Schwartz Lemma and Generalized Leibnitz Lemma in Taylor $\lambda$-calculus and showed the partial differentiation can be understated as a kind-of resource based substation. In the next section, we will define a $\beta$-reduction for Taylor $\lambda$-calculus and use this to relate Taylor $\lambda$-calculus with differential $\lambda$-calculus.

## 3.4 Reductions in Taylor $\lambda$-Calculus

Recall that $\beta$-reduction represents a single step of evaluation in Taylor $\lambda$-calculus. Seeing differentiated abstraction as syntactically representing a construction relating to partial differentiation and the partial differentiation as computational, we should expect $\beta$-reduction to related differentiated abstraction to one application of a partial derivative. Note that (3.4.5)-(3.4.6) is motivated by (3.1.4) in section 3.1.

**Definition 3.4.1. Taylor $\beta$-Reduction**

Define Taylor $\beta$-reduction $\to_\beta$ to be the smallest (with respect to inclusion) relation from complex Taylor (poly-)terms to complex Taylor (poly-)terms satisfying the following properties:

(i) Given a simple Taylor term $s \in \Lambda_{t,\mathcal{V}}$, a simple Taylor poly-term $T \in \Lambda^!_{t,\mathcal{V}}$, and a complex Taylor term $s' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $s \to_\beta s'$,

$$\langle s\rangle T \to_\beta \langle s'\rangle T. \tag{3.4.1}$$

(ii) Given a simple Taylor term $s \in \Lambda_{t,\mathcal{V}}$, a simple Taylor poly-term $T \in \Lambda^!_{t,\mathcal{V}}$, and a complex Taylor poly-term $\tau \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ such that $T \to_\beta \tau$,

$$\langle s\rangle T \to_\beta \langle s\rangle\tau. \tag{3.4.2}$$

(iii) Given a variable $x \in \mathcal{V}$, a simple differential term $s \in \Lambda_{t,\mathcal{V}}$, and a complex Taylor term $s' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $s \to_\beta s'$,

$$\lambda x.s \to_\beta \lambda x.s'. \tag{3.4.3}$$

(iv) Given a simple Taylor term $s \in \Lambda_{t,\mathcal{V}}$, a simple Taylor poly-term $T \in \Lambda^!_{t,\mathcal{V}}$, and a complex Taylor term $s' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $s \to_\beta s'$,

$$Ts \to_\beta Ts'. \tag{3.4.4}$$

(v) Given a variable $x \in \mathcal{V}$ and a simple differential term $s \in \Lambda_{t,\mathcal{V}}$,

$$\langle\lambda x.s\rangle 1 \to_\beta s[0/x]. \tag{3.4.5}$$

(vi) Given a variable $x \in \mathcal{V}$, simple differential term $s, t \in \Lambda_{t,\mathcal{V}}$, and a simple Taylor poly-term $T \in \Lambda^!_{t,\mathcal{V}}$,

$$\langle\lambda x.s\rangle tT \to_\beta \langle\lambda x.\partial_x s * t\rangle T. \tag{3.4.6}$$

(vii) Given a simple Taylor (poly-)term $S \in \Lambda^!_{t,\mathcal{V}}$, complex Taylor (poly-)terms $\sigma, \tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, and a scalar $r \in \mathcal{R}$ such that $s \to_{\mathcal{R}} \sigma$,

$$r \cdot S + \tau \to_\beta r \cdot \sigma + \tau. \tag{3.4.7}$$

♣

One of the great properties of Taylor $\beta$-reduction is that it strictly decreases the size of a Taylor term, directly giving us strong normalization for Taylor $\beta$-reduction, a property is much more difficult to show for differential $\beta$-reduction (see Sec. 5 of [8]).

**Lemma 3.4.2**

Let $S \in \Lambda^!_{t,\mathcal{V}}$ be a simple Taylor poly-term and let $\sigma \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ be a complex Taylor poly-term such that $S \to_\beta \sigma$. Then, for $U \in \mathrm{supp}\,(\sigma)$, one has $\mathrm{size}\,(U) < \mathrm{size}\,(S)$. ♡

***Proof***. We prove this by induction on $S \to_\beta \sigma$.

1. Assume that $S \to_\beta \sigma$ because there exist a simple Taylor term $s \in \Lambda_{t,\mathcal{V}}$, a simple Taylor poly-term $T \in \Lambda^!_{t,\mathcal{V}}$, and a complex Taylor term $s' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $s \to_\beta s'$, $S = \langle s\rangle T$, and $\sigma = \langle s'\rangle T$. Since $U \in \mathrm{supp}\,(\sigma) = \mathrm{supp}\,(\langle s'\rangle T)$, there exists a simple Taylor term $u \in \mathrm{supp}\,(s')$ such that $U = \langle u\rangle T$. Since $s \to_\beta s'$ and $u \in \mathrm{supp}\,(s')$, by the induction hypothesis, $\mathrm{size}\,(u) < \mathrm{size}\,(s)$, so $\mathrm{size}\,(U) = \mathrm{size}\,(\langle u\rangle T) = 1 + \mathrm{size}\,(u) + \mathrm{size}\,(T) < 1 + \mathrm{size}\,(s) + \mathrm{size}\,(T) = \mathrm{size}\,(\langle s\rangle T) = \mathrm{size}\,(S)$.

2. Assume that $S \to_\beta \sigma$ because there exist a simple Taylor term $s \in \Lambda_{t,\mathcal{V}}$, a simple Taylor poly-term $T \in \Lambda^!_{t,\mathcal{V}}$, and a complex Taylor poly-term $\tau \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$ such that $T \to_\beta \tau$, $S = \langle s\rangle T$, and $\sigma = \langle s\rangle\tau$. Since $U \in \mathrm{supp}\,(\sigma) = \mathrm{supp}\,(\langle s'\rangle\tau)$, there exists a simple Taylor poly-term $V \in \mathrm{supp}\,(\tau)$ such that $U = \langle s\rangle V$. Since $T \to_\beta \tau$ and $V \in \mathrm{supp}\,(\tau)$, by the induction hypothesis, $\mathrm{size}\,(V) < \mathrm{size}\,(T)$, so $\mathrm{size}\,(U) = \mathrm{size}\,(\langle s\rangle V) = 1 + \mathrm{size}\,(s) + \mathrm{size}\,(V) < 1 + \mathrm{size}\,(s) + \mathrm{size}\,(T) = \mathrm{size}\,(\langle s\rangle T) = \mathrm{size}\,(S)$.

3. Assume that $S \to_\beta \sigma$ because there exist a variable $x \in \mathcal{V}$, a simple differential term $s \in \Lambda_{t,\mathcal{V}}$, and a complex Taylor term $s' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $s \to_\beta s'$, $S = \lambda x.s$, and $\sigma = \lambda x.s'$. Since $U \in \mathrm{supp}\,(\sigma) = \mathrm{supp}\,(\lambda x.s')$, there exists a simple Taylor term $u \in \mathrm{supp}\,(s')$ such that $U = \lambda x.u$. Since $s \to_\beta s'$ and $u \in \mathrm{supp}\,(s')$, by the induction hypothesis, $\mathrm{size}\,(u) < \mathrm{size}\,(s)$, so $\mathrm{size}\,(U) = \mathrm{size}\,(\lambda x.u) = 1 + \mathrm{size}\,(u) < 1 + \mathrm{size}\,(s) = \mathrm{size}\,(\lambda x.s) = \mathrm{size}\,(S)$.

4. Assume that $S \to_\beta \sigma$ because there exist a simple Taylor term $s \in \Lambda_{t,\mathcal{V}}$, a simple Taylor poly-term $T \in \Lambda^!_{t,\mathcal{V}}$, and a complex Taylor term $s' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $s \to_\beta s'$, $S = Ts$, and $\sigma = Ts'$. Since $U \in \mathrm{supp}\,(\sigma) = \mathrm{supp}\,(Ts')$, there exists a simple Taylor term $u \in \mathrm{supp}\,(s')$ such that $U = Tu$. Since $s \to_\beta s'$ and $u \in \mathrm{supp}\,(s')$, by the induction hypothesis, $\mathrm{size}\,(u) < \mathrm{size}\,(s)$, so $\mathrm{size}\,(U) = \mathrm{size}\,(Tu) = 1 + \mathrm{size}\,(T) + \mathrm{size}\,(u) < 1 + \mathrm{size}\,(T) + \mathrm{size}\,(u) = \mathrm{size}\,(Ts) = \mathrm{size}\,(S)$.

5. Assume that $S \to_\beta \sigma$ because there exist a variable $x \in \mathcal{V}$ and a simple differential term $s \in \Lambda_{t,\mathcal{V}}$ such that $S = \langle\lambda x.s\rangle 1$ and $\sigma = s[0/x]$. If $\deg_x(s) > 0$, by Lemma 3.3.20, $\sigma = s[0/x] = 0$, which implies that $U \in \mathrm{supp}\,\sigma = \mathrm{supp}\,0 = \varnothing$, giving us a contradiction. Hence, $\deg_x(s) = 0$, so by Lemma 3.3.20, $x \notin \mathrm{FV}\,(s)$, so by Lemma 3.3.10(i), $\sigma = s[0/x] = s$. Since $U \in \mathrm{supp}\,\sigma = \mathrm{supp}\,s = \{s\}$, $U = s$. Hence, $\mathrm{size}\,(S) = \mathrm{size}\,(\langle\lambda x.s\rangle 1) = 1 + \mathrm{size}\,(\lambda x.s) + \mathrm{size}\,(1) = 1 + 1 + \mathrm{size}\,(s) + 0 = 2 + \mathrm{size}\,(s) = 2 + \mathrm{size}\,(u)$, which implies that $\mathrm{size}\,(u) < \mathrm{size}\,(S)$.

6. Assume that $S \to_\beta \sigma$ because there exist a variable $x \in \mathcal{V}$, simple differential term $s, t \in \Lambda_{t,\mathcal{V}}$, and a simple Taylor poly-term $T \in \Lambda^!_{t,\mathcal{V}}$ such that $S = \langle\lambda x.s\rangle tT$ and $\sigma = \langle\lambda x.\partial_x s * t\rangle T$. Since $U \in \mathrm{supp}\,(\sigma) = \mathrm{supp}\,(\langle\lambda x.\partial_x s * t\rangle T)$, there exists a simple Taylor term $u \in \mathrm{supp}\,(\partial_x s * t)$ such that $U = \langle\lambda x.u\rangle T$. Since $u \in \mathrm{supp}\,(\partial_x s * t)$, by Lemma 3.3.12, $\mathrm{size}\,(u) = \mathrm{size}\,(s) + \mathrm{size}\,(t) - 1$, so $\mathrm{size}\,(S) = \mathrm{size}\,(\langle\lambda x.s\rangle tT) = 1 + \mathrm{size}\,(\lambda x.s) + \mathrm{size}\,(tT) = 1 + 1 + \mathrm{size}\,(s) + 1 + \mathrm{size}\,(t) + \mathrm{size}\,(T) = 4 + \mathrm{size}\,(s) + \mathrm{size}\,(t) - 1 + \mathrm{size}\,(T) = 4 + \mathrm{size}\,(u) + \mathrm{size}\,(T) = 3 + \mathrm{size}\,(\lambda x.u) + \mathrm{size}\,(T) = 2 + \mathrm{size}\,(\langle\lambda x.u\rangle T) = 2 + \mathrm{size}\,(U)$, which implies that $\mathrm{size}\,(U) < \mathrm{size}\,(S)$.

This completes the induction. ❑

**Theorem 3.4.3. Strong Normalization of Taylor $\beta$-Reduction**

Taylor $\beta$-reduction is strongly normalizing. ♡

***Proof.*** Suppose that Taylor $\beta$-reduction is not strongly normalizing. Then, there exists complex Taylor terms $\sigma_1, \sigma_2, \sigma_3, \ldots \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $\sigma_i \to_\beta \sigma_{i+1}$ for all $i \in \mathbb{N}$. This implies that there exist simple Taylor terms $S_1, S_2, S_3, \ldots \in \Lambda_{t,\mathcal{V}}$ and complex Taylor terms $\tau_2, \tau_3, \tau_4, \ldots \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $S_{i-1} \to_\beta \tau_i$ and $S_i \in \operatorname{supp}(\tau_i)$ for all $i \geqslant 2$. Let $n \coloneqq \operatorname{size}(S_1) + 2$. By applying Lemma 3.4.2, $\operatorname{size} S_n \leqslant \operatorname{size} S_1 - n + 1 = \operatorname{size} S_1 - (\operatorname{size}(S_1) + 2) + 1 = -1$, contradicting the fact that size of a Taylor term is never negative. Hence, Taylor $\beta$-reduction is strongly normalizing. ❏

With the introduction of $\beta$-reduction, we have finally defined every key structure of Taylor $\lambda$-calculus. We defined Taylor $\lambda$-calculus, defined partial differentiation, showed that it satisfies the synthetic Schwartz Rule, Chain Rule, and Leibnitz Rule, define a notion of evaluating Taylor terms, and shows that $\beta$-reduction is strongly normalizing. It, like it's cousin differential $\lambda$-calculus, is a good framework at synthetically dealing with analytic functions. But there is a natural question now: is Taylor $\lambda$-calculus unrelated to differential $\lambda$-calculus? The answer is no. In fact, Taylor $\lambda$-calculus can be seen as a subset of differential $\lambda$-calculus.

**Definition 3.4.4. The Injection From Complex Taylor to Differential Terms**

Define the injection $(-)^\star \colon \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle \to \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ that takes a complex Taylor term to complex differential terms by:

$$(x)^\star = x, \tag{3.4.8a}$$
$$(\lambda x.s)^\star = \lambda x.s^\star, \tag{3.4.8b}$$
$$(\langle s\rangle t_1 \ldots t_n)^\star = (\mathcal{D}_1(\cdots(\mathcal{D}_1 s^\star * t_1^\star)\cdots) * (t_n)^\star)0, \tag{3.4.8c}$$
$$\left(\sum_{i=1}^{n} r_i \cdot s_i\right)^\star = \sum_{i=1}^{n} r_i \cdot s_i^\star, \tag{3.4.8d}$$

where $x \in \mathcal{V}$ is a variable, $n \in \mathbb{N}_0$ is a nonnegative integer, $s, s_1, \ldots, s_n, t_1, \ldots, t_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms, and $r_1, \ldots, r_n \in \mathcal{R}$ are scalars. ♣

This map between complex Taylor terms and complex differential terms preserves substitution, partial differentiation, and importantly $\beta$-reduction, which is exactly what we would expect from a canonical injection from Taylor $\lambda$-calculus to differential $\lambda$-calculus.

**Lemma 3.4.5**

Given a variable $x \in \mathcal{V}$ and complex Taylor terms $s, t \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$, $(s[t/x])^\star = s^\star[t^\star/x]$. ♡

***Proof.*** This follows by induction on $s$.

1. Assume that $s = x$. Then,

$$(s[t/x])^\star = (x[t/x])^\star = t^\star = x[t^\star/x] = s^\star[t^\star/x]. \tag{3.4.9}$$

2. Assume that $s = y$, where $y \in \mathcal{V}$ is a variable distinct from $x$. Then,

$$(s[t/x])^\star = (y[t/x])^\star = (y)^\star = y = y[t^\star/x] = (y)^\star[t^\star/x] = s^\star[(t^\star/x)]. \tag{3.4.10}$$

3. Assume that $s = \lambda y.u$, where $y \in \mathcal{V}$ is a fresh variable and $u \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. By the induction hypothesis on $u$, $(u[t/x])^\star = u^\star[t^\star/x]$, so

$$\begin{aligned}(s[t/x])^\star &= ((\lambda y.u)[t/x])^\star = (\lambda y.u[t/x])^\star = \lambda y.(u[y/x])^\star = \lambda y.u^\star[t^\star/x] \\ &= (\lambda y.u^\star)[t^\star/x] = (\lambda y.u)^\star[t^\star/x] = s^\star[t^\star/x].\end{aligned} \tag{3.4.11}$$

4. Assume that $s = \langle u\rangle v_1, \ldots, v_n$, where $n \in \mathbb{N}_0$ be a nonnegative integer and $u, v_1, \ldots, v_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. Then, by the induction hypothesis on $u$, $(u[t/x])^\star = u^\star[t^\star/x]$, and for each $1 \leqslant i \leqslant n$, by the induction hypothesis on $v_i$, $(v_i[t/x])^\star = (v_i)^\star[t^\star/x]$. Hence,

$$\begin{aligned}(s[t/x])^\star &= ((\langle u\rangle v_1 \ldots v_n)[t/x])^\star = (\langle u[t/x]\rangle v_1[t/x] \ldots v_n[t/x])^\star \\ &= (\mathcal{D}_1(\cdots(\mathcal{D}_1(u[t/x])^\star * (v_1[t/x])^\star)\cdots) * (v_n[t/x])^\star)0 \\ &= (\mathcal{D}_1(\cdots(\mathcal{D}_1 u^\star[t^\star/x]) * (v_1^\star[t^\star/x])\cdots) * v_n^\star[t^\star/x])0[t^\star/x] \\ &= (\cdots(\mathcal{D}_1 u^\star) * (v_1^\star)\cdots) * v_n^\star)0)[t^\star/x] \\ &= (\langle u\rangle v_1 \ldots v_n)^\star[t^\star/x] = s^\star[t^\star/x],\end{aligned} \tag{3.4.12}$$

where we used Lemma 2.3.6 for the fifth equality.

5. Assume that $s = \sum_{i=1}^n r_i \cdot u_i$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $u_1, \ldots, u_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms and $r_1, \ldots, r_n \in \mathcal{R}$ are scalars. For each $1 \leqslant i \leqslant n$, by the induction hypothesis on $u_i$, $(u_i[t/x])^\star = (u_i)^\star[t^\star/x]$. Hence,

$$\begin{aligned}(s[t/x])^\star &= \left(\left(\sum_{i=1}^n r_i \cdot u_i\right)[t/x]\right)^\star = \left(\sum_{i=1}^n r_i \cdot u_i[t/x]\right)^\star = \sum_{i=1}^n r_i \cdot (u_i[t/x])^\star \\ &= \sum_{i=1}^n r_i \cdot (u_i)^\star[t^\star/x] = \left(\sum_{i=1}^n r_i \cdot (u_i)^\star\right)[t^\star/x] \\ &= \left(\sum_{i=1}^n r_i \cdot u_i\right)^\star [t^\star/x] = s^\star[t^\star/x].\end{aligned} \tag{3.4.13}$$

This completes the induction. ❏

**Lemma 3.4.6**

Given complex differential term $\sigma, \tau \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$, $\partial_x(\sigma 0) * \tau = (\partial_x \sigma * \tau)0$ ♡

***Proof***. By Lemma 2.2.29(xii), there exist unique nonnegative integers $n \in \mathbb{N}_0$, scalars $r_1, \ldots, r_n \in \mathcal{R}$, and simple differential terms $s_1, \ldots, s_n \in \Lambda_{d,\mathcal{V}}$ such that $\sigma = \sum_{i=1}^n r_i \cdot s_i$. As a result,

$$\begin{aligned}\partial_x(\sigma 0) * \tau &= \partial_x\left(\left(\sum_{i=1}^n r_i \cdot s_i\right)0\right) * \tau = \partial_x\left(\sum_{i=1}^n r_i \cdot s_i 0\right) * \tau = \sum_{i=1}^n r_i \cdot \partial_x(s_i 0) * \tau \\ &= \sum_{i=1}^n r_i \cdot ((\partial_x s_i * \tau)0 + (\mathcal{D}_1 s_i * (\partial_x 0 * \tau))0) \\ &= \sum_{i=1}^n r_i \cdot ((\partial_x s_i * \tau)0 + (\mathcal{D}_1 s_i * 0)0) = \sum_{i=1}^n r_i \cdot ((\partial_x s_i * \tau)0 + 00) \\ &= \sum_{i=1}^n r_i \cdot ((\partial_x s_i * \tau)0 + 0) = \sum_{i=1}^n r_i \cdot ((\partial_x s_i * \tau)0)\end{aligned} \tag{3.4.14}$$

$$= \left(\sum_{i=1}^{n} r_i \cdot \partial_x s_i * \tau\right) 0 = \left(\partial_x \left(\sum_{i=1}^{n} r_i \cdot s_i\right) * \tau\right) 0 = (\partial_x \sigma * \tau)0.$$

This completes the proof. ❏

**Lemma 3.4.7**

Given a variable $x \in \mathcal{V}$ and complex Taylor terms $s, t \in \mathcal{R}\langle\Lambda^!_{t,\mathcal{V}}\rangle$, $(\partial_x s * t)^\star = \partial_x s^\star * t^\star$. ♡

***Proof***. This follows by induction on $s$.

1. Assume that $s = x$. Then,

$$(\partial_x s * t)^\star = (\partial_x x * t)^\star = t^\star = \partial_x x * t^\star = \partial_x (x)^\star * t^\star = \partial_x s^\star * t^\star. \tag{3.4.15}$$

2. Assume that $s = y$, where $y \in \mathcal{V}$ is a variable distinct from $x$. Then,

$$(\partial_x s * t)^\star = (\partial_x y * t)^\star = (y)^\star = y = \partial_x y * t^\star = \partial_x (y)^\star * t^\star = \partial_x s^\star * t^\star. \tag{3.4.16}$$

3. Assume that $s = \lambda y.u$, where $y \in \mathcal{V}$ is a fresh variable and $u \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. By the induction hypothesis on $u$, $(\partial_x u * t)^\star = \partial_x u^\star * t^\star$, so

$$\begin{aligned}(\partial_x s * t)^\star &= (\partial_x (\lambda y.u) * t)^\star = (\lambda y.\partial_x u * t)^\star = \lambda y.(\partial_x u * t)^\star = \lambda y.\partial_x u^\star * t^\star \\ &= \partial_x (\lambda y.u^\star) * t^\star = \partial_x (\lambda y.u)^\star * t^\star = \partial_x s^\star * t^\star.\end{aligned} \tag{3.4.17}$$

4. Assume that $s = \langle u\rangle v_1, \ldots, v_n$, where $n \in \mathbb{N}_0$ be a nonnegative integer and $u, v_1, \ldots, v_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. Then, by the induction hypothesis on $u$, $(\partial_x u * t)^\star = \partial_x u^\star * t^\star$, and for each $1 \leqslant i \leqslant n$, by the induction hypothesis on $v_i$, $(\partial_x v_i * t)^\star = \partial_x (v_i)^\star * t^\star$. Hence,

$$\begin{aligned}(\partial_x s * t)^\star &= (\partial_x (\langle u\rangle v_1, \ldots, v_n) * t)^\star \\ &= (\langle \partial_x u * t\rangle v_1, \ldots, v_n + \langle u\rangle(\partial_x v_1, \ldots, v_n))^\star \\ &= \left(\langle \partial_x u * t\rangle v_1, \ldots, v_n + \langle u\rangle \left(\sum_{i=1}^{n} v_1 \ldots v_{i-1}(\partial_x v_i * t) v_{i+1} \ldots, v_n\right)\right)^\star \\ &= \left(\langle \partial_x u * t\rangle v_1, \ldots, v_n + \sum_{i=1}^{n} \langle u\rangle v_1 \ldots v_{i-1}(\partial_x v_i * t) v_{i+1} \ldots, v_n\right)^\star \\ &= (\langle \partial_x u * t\rangle v_1, \ldots, v_n)^\star + \sum_{i=1}^{n} (\langle u\rangle v_1 \ldots v_{i-1}(\partial_x v_i * t) v_{i+1} \ldots, v_n)^\star \\ &= (\mathcal{D}_{1,\ldots,1}(\partial_x u * t)^\star * (v_1^\star, \ldots, v_n^\star))0 \\ &\qquad + \sum_{i=1}^{n} (\mathcal{D}_{1,\ldots,1} u^\star * (v_1^\star, \ldots v_{i-1}^\star, (\partial_x v_i * t)^\star, v_{i+1}^\star, \ldots, v_n^\star))0 \\ &= \Bigg(\mathcal{D}_{1,\ldots,1}(\partial_x u^\star * t^\star) * (v_1^\star, \ldots, v_n^\star) \\ &\qquad + \sum_{i=1}^{n} \mathcal{D}_{1,\ldots,1} u^\star * (v_1^\star, \ldots v_{i-1}^\star, \partial_x v_i^\star * t^\star, v_{i+1}^\star, \ldots, v_n^\star)\Bigg)0 \\ &= (\partial_x (\mathcal{D}_{1,\ldots,1} u^\star * (v_1^\star, \ldots, v_n^\star)) * t^\star)0\end{aligned} \tag{3.4.18}$$

$$
\begin{aligned}
&= \partial_x((\mathcal{D}_{1,\ldots,1}u^\star * (v_1^\star,\ldots,v_n^\star))0) * t^\star \\
&= \partial_x(\langle s\rangle v_1 \ldots v_n)^\star * t^\star \\
&= \partial_x s^\star * t^\star,
\end{aligned}
$$

where we used Lemma 3.4.6 for the third to last equality and Lemma 2.3.15 for the second to last equality.

5. Assume that $s = \sum_{i=1}^n r_i \cdot u_i$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $u_1, \ldots, u_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms and $r_1, \ldots, r_n \in \mathcal{R}$ are scalars. For each $1 \leqslant i \leqslant n$, by the induction hypothesis on $u_i$, $(\partial_x u_i * t)^\star = \partial_x(u_i)^\star * t^\star$. Hence,

$$
\begin{aligned}
(\partial_x s * t)^\star = \left(\partial_x\left(\sum_{i=1}^n r_i \cdot u_i\right) * t\right)^\star &= \left(\sum_{i=1}^n r_i \cdot (\partial_x u_i * t)\right)^\star = \sum_{i=1}^n r_i \cdot (\partial_x u_i * t)^\star \\
&= \sum_{i=1}^n r_i \cdot (\partial_x (u_i)^\star * t^\star) = \partial_x\left(\sum_{i=1}^n r_i \cdot (u_i)^\star\right) * t^\star \\
&= \partial_x\left(\sum_{i=1}^n r_i \cdot u_i\right)^\star * t^\star = \partial_x s^\star * t^\star.
\end{aligned}
\tag{3.4.19}
$$

This completes the induction. ❏

**Theorem 3.4.8**

Given complex Taylor terms $s, s'$ such that $s \to_\beta s'$, one has $s^\star \to_\beta (s')^\star$. ♡

***Proof.*** We prove this by induction on $s \to_\beta \sigma$.

1. Assume that $s \to_\beta \sigma$ because there exist a nonnegative integer $n \in \mathbb{N}_0$, simple Taylor terms $u, v_1, \ldots, v_n \in \Lambda_{t,\mathcal{V}}$, and a complex Taylor term $u' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $u \to_\beta u'$, $s = \langle u\rangle v_1 \ldots v_n$, and $\sigma = \langle u'\rangle v_1 \ldots v_n$. Since $u \to_\beta u'$, by the induction hypothesis, $u^\star \to_\beta u'^\star$, so

$$
\begin{aligned}
s^\star = (\langle u\rangle v_1 \ldots v_n)^\star &= (\mathcal{D}_1(\cdots(\mathcal{D}_1 u^\star * v_1^\star)\cdots) * v_n^\star)0 \\
&\to_\beta (\mathcal{D}_1(\cdots(\mathcal{D}_1 u'^\star * v_1^\star)\cdots) * v_n^\star)0 \\
&= (\langle u'\rangle v_1 \ldots v_n)^\star = \sigma^\star.
\end{aligned}
\tag{3.4.20}
$$

where we used the fact that differential $\beta$-reduction is contextual (see Lemma 2.4.13) for $\to_\beta$.

2. Assume that $s \to_\beta \sigma$ because there exist a positive integer $n \in \mathbb{N}_0$, simple Taylor terms $u, v_1, \ldots, v_n \in \Lambda_{t,\mathcal{V}}$, and a complex Taylor term $v_1' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $v_1 \to_\beta v_1'$, $s = \langle u\rangle v_1 v_2 \ldots v_n$, and $\sigma = \langle u\rangle v_1' v_2 \ldots v_n$. Since $u \to_\beta u'$, by the induction hypothesis, $u^\star \to_\beta u'^\star$, so

$$
\begin{aligned}
s^\star = (\langle u\rangle v_1 v_2 \ldots v_n)^\star &= (\mathcal{D}_1(\cdots(\mathcal{D}_1(\mathcal{D}_1 u^\star * v_1^\star) * v_2^\star)\cdots) * v_n^\star)0 \\
&\to_\beta (\mathcal{D}_1(\cdots(\mathcal{D}_1(\mathcal{D}_1 u^\star * v_1^\star) * v_2^\star)\cdots) * v_n^\star)0 \\
&= (\langle u\rangle v_1' v_2 \ldots v_n)^\star = \sigma^\star.
\end{aligned}
\tag{3.4.21}
$$

where we used the fact that differential $\beta$-reduction is contextual (see Lemma 2.4.13) for $\to_\beta$.

3. Assume that $s \to_\beta \sigma$ because there exist a variable $x \in \mathcal{V}$, a simple differential term $u \in \Lambda_{t,\mathcal{V}}$, and a complex Taylor term $u' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $u \to_\beta u'$, $s = \lambda x.u$, and $\sigma = \lambda x.u'$. Since $u \to_\beta u'$, by the induction hypothesis, $u^\star \to_\beta u'^\star$, so by (2.4.17),

$$s^\star = (\lambda x.u)^\star = \lambda x.u^\star \to_\beta \lambda x.u'^\star = (\lambda x.u')^\star = \sigma^\star. \tag{3.4.22}$$

4. Assume that $s \to_\beta \sigma$ because there exist a variable $x \in \mathcal{V}$ and a simple differential term $u \in \Lambda_{t,\mathcal{V}}$ such that $s = \langle\lambda x.u\rangle 1$ and $\sigma = u[0/x]$. Then,

$$\begin{aligned} u^\star &= (\langle\lambda x.u\rangle 1)^\star = (\lambda x.u)^\star 0 = (\lambda x.u^\star)0 \\ &\to_\beta u^\star[0/x] = u^\star[0^\star/x] = (u[0/x])^\star = \sigma^\star, \end{aligned} \tag{3.4.23}$$

where we used (2.4.15) for the $\to_\beta$ and Lemma 3.4.5 for the second to last equality.

5. Assume that $s \to_\beta \sigma$ because there exist a variable $x \in \mathcal{V}$, a positive integer $n \in \mathbb{N}_0$, and simple differential term $u, v_1, \ldots, v_n \in \Lambda_{t,\mathcal{V}}$ such that $s = \langle\lambda x.u\rangle v_1 \ldots v_n$ and $\sigma = \langle\lambda x.\partial_x u * v_1\rangle v_2 \ldots v_n$. Then,

$$\begin{aligned} s^\star &= (\langle\lambda x.u\rangle v_1 \ldots v_n)^\star = (\mathcal{D}_1(\cdots(\mathcal{D}_1 u^\star * v_1^\star)\cdots) * v_n^\star)0 \\ &= (D_1^n \lambda x.u^\star * (v_1^\star, \ldots, v_n^\star))0 \\ &\to_\beta (D_1^{n-1}\lambda x.(\partial_x u^\star * v_1^\star) * (v_2^\star, \ldots, v_n^\star))0 \\ &= (\mathcal{D}_1(\cdots(\mathcal{D}_1(\partial_x u^\star * v_1^\star) * v_2^\star)\cdots) * v_n^\star)0 \\ &= \cdots(\mathcal{D}_1(\partial_x u * v_1)^\star * v_2^\star)\cdots) * v_n^\star)0 \\ &= (\langle\lambda x.\partial_x u * v_1\rangle v_2 \ldots v_n)^\star = \sigma^\star, \end{aligned} \tag{3.4.24}$$

where we used (2.4.19) for the $\to_\beta$ and we used Lemma 3.4.7 for the second to last equality.

This completes the induction. ❏

**Lemma 3.4.9**

Let $\sigma, \sigma' \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ be simple (resp. complex) differential terms and $s \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ be a simple (resp. complex) Taylor terms such that $\sigma = s^\star$ and $\sigma \twoheadrightarrow_\beta \sigma'$. Then, there exists a simple (resp. complex) Taylor term $s' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $\sigma' = s'^\star$ and $s \twoheadrightarrow_\beta s'$. ♡

***Proof.*** We prove this via induction on $t$.

1. Assume that $s = x$. Since $x = (x)^\star = s^\star = \sigma \twoheadrightarrow_\beta \sigma'$, we know that $\sigma' = x$. Let $s' = x$. Then, $\sigma' = x = (x)^\star = s'^\star$ and $s = x \twoheadrightarrow_\beta x = s'$.

2. Assume that $s = \lambda y.u$, where $y \in \mathcal{V}$ is a fresh variable and $u \in \Lambda_{t,\mathcal{V}}$ is a simple Taylor term. Since $\lambda y.u^\star = (\lambda y.u)^\star = s^\star = \sigma \twoheadrightarrow_\beta = \sigma'$, we know that $\sigma' = \lambda y.t$, where $t \in \Lambda_{d,\mathcal{V}}$ is a simple differential term such that $u^\star \twoheadrightarrow_\beta t$. By the induction hypothesis, this implies that there exists a simple Taylor term $v \in \Lambda_{t,\mathcal{V}}$ such that $t = v^\star$ and $u \twoheadrightarrow_\beta v$. Let $s' = \lambda x.v$. Then, $\sigma' = \lambda y.t = \lambda y.v^\star = (\lambda y.v)^\star = s'^\star$ and $s = \lambda x.u \twoheadrightarrow_\beta \lambda x.v = s'$.

3. Assume that $s = \langle u\rangle v_1 \ldots v_n$, where $n \in \mathbb{N}_0$ be a nonnegative integer and $u, v_1, \ldots, v_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms. There are two cases.

(a) Suppose that $u$ is an abstraction. Then, $u = \lambda x.w$, where $x \in \mathcal{V}$ is a fresh variable and $w \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ is a simple Taylor term. Since $(\mathcal{D}_1^n \lambda x.w^\star * (v_1^\star, \ldots, v_n^\star))0 = (\langle\lambda x.w\rangle v_1 \ldots v_n)^\star = s^\star = \sigma \twoheadrightarrow_\beta \sigma'$, we know that $\sigma' = (\mathcal{D}_1^{n-m} \lambda x.(\partial_x^m t_0 * (t_1, \ldots, t_m)) * (t_{m+1}, \ldots, t_n))0$, where $m \in \mathbb{N}_0$ is a nonnegative integer and $t_0, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms such that $w^\star \twoheadrightarrow_\beta t_0$ and $v_i^\star \twoheadrightarrow_\beta t_i$ for all $1 \leqslant i \leqslant n$. By the induction hypothesis, there exists simple Taylor terms $w', v_1', \ldots, v_n' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $t_0 = w'^\star$, $t_i = v_i'^\star$ for all $1 \leqslant i \leqslant n$, $w \twoheadrightarrow_\beta w'$, and $v_i \twoheadrightarrow_\beta v_i'$ for all $1 \leqslant i \leqslant n$. Let $s' = \langle\lambda x.\partial_x^m w' * (v_1', \ldots, v_m')\rangle v_{m+1}' \ldots v_n'$. Then, $\sigma' = (\mathcal{D}_1^{n-m} \lambda x.(\partial_x^m t_0 * (t_1, \ldots, t_m)) * (t_{m+1}, \ldots, t_n))0 = (\mathcal{D}_1^{n-m} \lambda x.(\partial_x^m w'^\star * (v_1'^\star, \ldots, v_m'^\star)) * (t_{m+1}'^\star, \ldots, t_n'^\star))0 = \langle\lambda x.\partial_x^m w'^\star * (v_1'^\star, \ldots, v_m'^\star)\rangle v_{m+1}'^\star \ldots v_n'^\star = s'^\star$ and $s = \langle\lambda x.w\rangle v_1 \ldots v_n \twoheadrightarrow_\beta \langle\lambda x.\partial_x^m w' * (v_1', \ldots, v_m')\rangle v_{m+1}' \ldots v_n' = s'$.

(b) Suppose otherwise. Since $(\mathcal{D}_1^n u^\star * (v_1^\star, \ldots, v_n^\star))0 = (\langle u\rangle v_1 \ldots v_n)^\star = s^\star = \sigma \twoheadrightarrow_\beta \sigma'$, this implies that $\sigma' = (\mathcal{D}_1^n t_0 * (t_1, \ldots, t_n))0$, where $t_0, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms such that $u^\star \twoheadrightarrow_\beta t_0$ and $v_i^\star \twoheadrightarrow_\beta t_i$ for all $1 \leqslant i \leqslant n$. By the induction hypothesis, there exists simple Taylor terms $u', v_1', \ldots, v_n' \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $t_0 = u'^\star$, $t_i = v_i'^\star$ for all $1 \leqslant i \leqslant n$, $u \twoheadrightarrow_\beta u'$, and $v_i \twoheadrightarrow_\beta v_i'$ for all $1 \leqslant i \leqslant n$. Let $s' = \langle u'\rangle v_1' \ldots v_n'$. Then, $\sigma' = (\mathcal{D}_1^n t_0 * (t_1, \ldots, t_n))0 = (\mathcal{D}_1^n u'^\star * (v_1'^\star, \ldots, t_n'^\star))0 = (\langle u\rangle v_1' \ldots v_n')^\star = s'^\star$ and $s = \langle u\rangle v_1 \ldots v_n \twoheadrightarrow_\beta \langle u'\rangle v_1' \ldots v_n' = s'$.

4. Assume that $s = \sum_{i=1}^n r_i \cdot u_i$, where $n \in \mathbb{N}_0$ is a nonnegative integer and $u_1, \ldots, u_n \in \Lambda_{t,\mathcal{V}}$ are simple Taylor terms and $r_1, \ldots, r_n \in \mathcal{R}$ are scalars. Since $\sum_{i=1}^n r_i \cdot u_i^\star = (\sum_{i=1}^n r_i \cdot u_i)^\star = s^\star = \sigma \twoheadrightarrow_\beta \sigma'$, we know that $\sigma' = \sum_{i=1}^n r_i \cdot t_i$, where $t_1, \ldots, t_n \in \Lambda_{d,\mathcal{V}}$ are simple differential terms such that $u_i^\star \twoheadrightarrow_\beta t_i$ for all $1 \leqslant i \leqslant n$. For each $1 \leqslant i \leqslant n$, since $u_i^\star \twoheadrightarrow_\beta t_i$, by the induction hypothesis, there exists a simple Taylor term $v_i \in \Lambda_{t,\mathcal{V}}$ such that $t_i = v_i^\star$ and $u_i \twoheadrightarrow_\beta v_i$. Let $s' = \sum_{i=1}^n r_i \cdot v_i$. Then, $\sigma' = \sum_{i=1}^n r_i \cdot t_i = \sum_{i=1}^n r_i \cdot v_i^\star = (\sum_{i=1}^n r_i \cdot v_i)^\star = s'^\star$ and $s = \sum_{i=1}^n r_i \cdot u_i \twoheadrightarrow_\beta \sum_{i=1}^n r_i \cdot v_i = s'$.

This completes the induction. ❏

Since the canonical injection $(-)^\star$ from complex Taylor terms to complex differential terms behaves well with $\beta$-reduction and differential $\beta$-reduction has the Church-Rosser property, we immediately get that Taylor $\beta$-reduction has the Church-Rosser property.

**Theorem 3.4.10. The Church-Rosser Property of Taylor $\beta$-Reduction**

Taylor $\beta$-reduction satisfies the Church-Rosser property. ♡

***Proof***. Take any complex Taylor terms $a, b, c \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $a \twoheadrightarrow_\beta b$ and $b \twoheadrightarrow_\beta c$. By Theorem 3.4.8, $a^\star \twoheadrightarrow_\beta b^\star$ and $a^\star \twoheadrightarrow c^\star$. Since differential $\beta$-reduction satisfies the Church-Rosser property (i.e. Theorem 2.4.28), there exists a complex differential term $\sigma \in \mathcal{R}\langle\Lambda_{d,\mathcal{V}}\rangle$ such that $b^\star \twoheadrightarrow_\beta \sigma$ and $c^\star \twoheadrightarrow_\beta \sigma$. By Lemma 3.4.9, there exists complex Taylor terms $d, e \in \mathcal{R}\langle\Lambda_{t,\mathcal{V}}\rangle$ such that $\sigma = d^\star = e^\star$ and $b \twoheadrightarrow_\beta d$ and $c \twoheadrightarrow_\beta e$. Since $d^\star = e^\star$ and $(-)^\star$ is injective, $d = e$. Hence, $b \twoheadrightarrow_\beta d$ and $c \twoheadrightarrow_\beta d$. ❏

# Chapter 4

# Differential Categories and Free Cornering

In the previous chapter, we introduced Taylor $\lambda$-calculus; we motivated its syntax as a variant of differential $\lambda$-calculus where we can easily express polynomials and every operator in the calculus is multi-linear; we generalized many properties of the binomial coefficient for natural numbers to multisets; we proved a number of properties of Taylor $\lambda$-calculus, like the Generalized Leibnitz Rule and strong normalization; we showed formally that Taylor $\lambda$-calculus is a subset of differential $\lambda$-calculus; and we used that fact to prove the Taylor $\lambda$-calculus satisfies the Church-Rosser property. In this chapter, we will introduce yet another formulation of a purely syntactic way to treat differentiation of analytic functions called differential categories, and we will show that a double-categorical model of asynchronous processes called the free cornering has a canonical differential categorical structure in it. In the first section, we will define various important categorical strictures, like linear categories, differential categories, and most important resource categories, which will be the main category of interest in this chapter. In the second section, we will define the free cornering with choice and iteration of a resource category, generalizing the definition of a free cornering with choice and iteration of a distributive category proposed by Chad Nester and Niels Voorneveld (see [16]). In the final section, we will construct the canonical differential structure in the free cornering.

## 4.1 Introducing Linear Categories, Differential Categories, and Resource Categories

In this chapter, we are interested in exploring a model of asynchronous processes. But before we can talk about asynchronous processes, we need to define the resources that we manipulate in some asynchronous processes. In Chad Nester's seminal work on modeling asynchronous processes via his framework of free cornering (see [15]), he proposed that a resource theory should just be a symmetric monoidal category, where objects are resources, morphisms are processes transforming one resource into another, and tensoring is possessing two objects or running two processes simultaneously. However, when Chad Nester and Niels Voorneveld wanted to expand the definition of free cornering to express control statements (i.e., choice) and loops (i.e., iteration), they found it necessary to define a resource theory to be a distributive category, where the co-product of two resources expresses having either of the resources but not necessarily both. This variation in the definition of the resource theory is small, but without it the ability to express choice and iteration would not be possible. Unfortunately, this definition of a resource theory lacks the ability to express an infinite amount of resources, which is very important when it comes to proving properties about iteration. To remedy this, we will define a resource theory to be a resource category, which is a minor variant of a differential category, which itself is a variant of a linear category, which is a categorical model of linear logic. Naturally, instead of defining a resource category with its over a dozen natural transformation and numerous coherence conditions directly, we are going to introduce it in pieces, defining multiple important categories along the way.

The motivation of these categorical concepts come from linear logic. Linear logic is similar, in a very informal sense, to classical logic. Importantly just like classical logic, linear logic has an idempotent negation ($(-)^{\perp}$). But, unlike classical logic, in linear logic we drop the structure rules of weakening, contraction, and, sometimes, exchange. The issue is that once we do this there becomes two distinct ways to express the notion of the binary connective "and" ($\wedge$), which are "times" ($\otimes$) and "with" ($\&$); two distinct ways to express the notion of the binary connective "or" ($\vee$), which are "parr" ($\parr$) and "plus" ($\oplus$); two distinct ways to express the notion of the constant "true" ($\top$), which are "one" ($\mathbb{1}$) and "top" ($\top$); and two distinct ways to express the notion of the constant "false" ($\perp$), which are "bottom" ($\perp$) and "zero" ($\mathbf{0}$). Since we don't have contraction and weakening, propositions in linear logic act like resources, which makes linear logic the perfect setting for describing a resource theory. The resource theoretic interpretation of linear logic is described in [11]. However, sometimes it is useful to talk about resources where we can perform contraction and weakening, which can be interpreted as an infinite spout or an infinite pool of a particular resource. For this reason, in linear logic, we also have the unary connectives "bang" (!) and "whimper" (?). Propositions under "bang" and "whimper" have some notion of weakening and contraction. In the resource theoretic interpretation of linear logic $!A$ acts like an infinite spout of the resource $A$ and $?A$ acts like an infinite pool of the resource $A$. With such a large amount of connectives and constants in linear logic, it is useful to split the connectives and constants into three groups:

(i) $\otimes$, $\parr$, $\mathbb{1}$, and $\perp$ form the multiplicative fragment of linear logic;

(ii) $\&$, $\oplus$, $\top$, and $\mathbf{0}$ form the additive fragment of linear logic; and

(iii) ! and ? form the exponential fragment of linear logic.

Each category we are about to introduce provides a semantics of the logic we get when reducing linear logic to some subset of its connectives and constants.

We start by introducing a linearly distributive category, which is the categorical semantics of the multiplicative fragment of linear logic without negation or exchange. A full treatment of non-symmetric linearly distributive categories can be found in [17].

**Definition 4.1.1. Linearly Distributive Category**

A *linearly distributive category* $\mathfrak{C}$ is a category that has

(i) functors $\otimes\colon \mathfrak{C}^2 \to \mathfrak{C}$ and $\parr\colon \mathfrak{C}^2 \to \mathfrak{C}$ called *times* and *par*, respectively;

(ii) distinguish objects $\mathbb{1}$ and $\perp$ called *one* and *bottom*, respectively;

(iii) natural isomorphisms $\lambda^{\otimes}_A\colon \mathbb{1} \otimes A \to A$ and $\lambda^{\parr}_A\colon \perp \parr A \to A$ called the *left times unitor* and *par left unitor*, respectively;

(iv) natural isomorphisms $\rho^{\otimes}_A\colon A \otimes \mathbb{1} \to A$ and $\rho^{\parr}_A\colon A \parr \perp \to A$ called the *right times unitor* and *par right unitor*, respectively;

(v) natural isomorphisms $\alpha^{\otimes}_{A,B,C} \colon A \otimes (B \otimes C) \to (A \otimes B) \otimes C$ and $\alpha^{\parr}_{A,B,C} \colon A \parr (B \parr C) \to (A \otimes B) \parr C$ called the *times right associator* and *par right associator*, respectively; and

(vi) natural transformations $\partial^{L}_{A,B,C} \colon A \otimes (B \parr C) \to (A \otimes B) \parr C$ and $\partial^{R}_{A,B,C} \colon (B \parr C) \otimes A \to B \parr (C \otimes A)$ called the *left distributor* and *right distributor*, respectively,

such that $(\mathfrak{C}, \otimes, \lambda^{\otimes}, \rho^{\otimes}, \alpha^{\otimes}, \mathbb{1})$ and $(\mathfrak{C}, \parr, \lambda^{\parr}, \rho^{\parr}, \alpha^{\parr}, \perp)$ are monoidal categories subject to a handful of coherence conditions, which are detailed in [17].

We can categorically express the exchange rule by introducing symmetric braiding natural transformations to the "times" and "par" monoidal structures. This results in a symmetric linearly distributive category. A full treatment of a symmetric linearly distributive categories can be found in [3].

**Definition 4.1.2. Symmetric Linearly Distributive Category**

A *symmetric linearly distributive category* $\mathfrak{C}$ is a linearly distributive category that has

(i) natural isomorphisms $\sigma^{\otimes}_{A,B} \colon A \otimes B \to B \otimes Z$ and $\sigma^{\parr}_{A,B} \colon A \parr B \to B \parr A$ called *times symmetric braiding* and *par symmetric braiding*, respectively,

such that $(\mathfrak{C}, \otimes, \lambda^{\otimes}, \rho^{\otimes}, \alpha^{\otimes}, \sigma^{\otimes}, \mathbb{1})$ and $(\mathfrak{C}, \parr, \lambda^{\parr}, \rho^{\parr}, \alpha^{\parr}, \sigma^{\parr}, \perp)$ are symmetric monoidal categories and

$$\partial^{R}_{A,B,C} = \sigma^{\parr}_{C \otimes A, B} \circ (\sigma^{\otimes}_{A,C} \parr 1_B) \circ \partial^{L}_{A,C,B} \circ (1_A \otimes \sigma^{\parr}_{B,C}) \circ \sigma^{\otimes}_{B \parr C, A}. \tag{4.1.1}$$

To add negation, we need to define the linear dual of an object in a linearly distributive category.

**Definition 4.1.3. Linear Dual**

Let $\mathfrak{C}$ be a linearly distributive category and let $X, Y$ be objects in $\mathfrak{C}$. Then, $Y$ is a *left linear dual* to $X$ (or $X$ is a *right linear dual* $Y$), which is written as $(\eta_{X,Y}, \epsilon_{X,Y}) \colon X \vdash Y$, if there exists morphisms $\eta_{X,Y} \colon \mathbb{1} \to Y \parr X$ and $\epsilon_{X,Y} \colon X \otimes Y \to \perp$ in $\mathfrak{C}$ such that

$$\begin{tikzcd} Y \arrow[r, "(\lambda^{\otimes}_Y)^{-1}"] \arrow[d, equal] & \mathbb{1} \otimes Y \arrow[r, "\eta_{X,Y} \otimes 1_Y"] & (Y \parr X) \otimes Y \arrow[d, "\partial^{R}_{Y,X,Y}"] \\ Y & Y \parr \perp \arrow[l, "\rho^{\parr}_Y"] & Y \parr (X \otimes Y) \arrow[l, "1_Y \parr \epsilon_{X,Y}"] \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r, "(\rho^{\otimes}_X)^{-1}"] \arrow[d, equal] & X \otimes \mathbb{1} \arrow[r, "1_X \otimes \eta_{X,Y}"] & X \otimes (Y \parr X) \arrow[d, "\partial^{L}_{X,Y,X}"] \\ X & \perp \parr X \arrow[l, "\lambda^{\parr}_X"] & (A \otimes B) \parr X. \arrow[l, "\epsilon_{X,Y} \parr 1_X"] \end{tikzcd} \tag{4.1.2}$$

**Lemma 4.1.4. Symmetry and Duality**

Let $\mathfrak{C}$ be a linearly distributive category.

(i) If $(\eta, \epsilon) \colon X \vdash Y$ and $(\eta', \epsilon') \colon X \vdash Z$, then $Y \cong Z$.

(ii) When $\mathfrak{C}$ is also symmetric, $(\eta, \epsilon) \colon X \vdash Y$ if and only if $(\sigma^{\parr}_{Y,X} \circ \eta, \epsilon \circ \sigma^{\otimes}_{Y,X}) \colon Y \vdash X$.

***Proof.*** See Lemma 2.10 in [5]. ❑

When every object in a (symmetric) linearly distributive category has a left linear dual, then the (symmetric) linearly distributive category is a (symmetric) $*$-autonomous category, which is a categorical semantics of the multiplicative fragment of linear logic with negation (and exchange). A good description of $*$-autonomous categories can be found in [5].

**Definition 4.1.5. $*$-Autonomous Category**

A (symmetric) *$*$-autonomous category* $\mathfrak{C}$ is a (symmetric) linearly distributive category such that for every object $X$ in $\mathfrak{C}$, there exists an object $X^\perp$ in $\mathfrak{C}$ and morphisms $\eta_X \colon \mathbb{1} \to X^\perp ⅋ X$ and $\epsilon_X \colon X \otimes X^\perp \to \perp$ in $\mathfrak{C}$ such that $(\eta_X, \epsilon_X) \colon X \vdash X^\perp$. When $\mathfrak{C}$ is symmetric, define $\eta'_X \coloneqq \sigma^{⅋}_{X^\perp, X} \circ \eta_X$ and $\epsilon'_X \coloneqq \epsilon_X \circ \sigma^{⅋}_{X,X}$ for each object $X$ in $\mathfrak{C}$. By Lemma 4.1.4, $(\eta'_X, \epsilon'_X) \colon X^\perp \vdash X$. ♣

There are many interesting example of $*$-autonomous categories, like the category of finiteness spaces, the category of coherence spaces, and even an abelian group realized as a symmetric monoidal category. Nevertheless, the most important $*$-autonomous category for this thesis is the compact closed category, which is just a $*$-autonomous category whose "times" and "parr" monoidal structures coincide.

**Theorem 4.1.6. A Compact Closed Category is a *-Autonomous Category**

If $(\mathfrak{C}, \otimes, \lambda, \rho, \alpha, \sigma, I, \eta, \epsilon)$ is a compact closed category, then

$$(\mathfrak{C}, \otimes, \lambda, \rho, \alpha, \sigma, I, \otimes, \lambda, \rho, \alpha, \sigma, I, \alpha, \alpha^{-1}, \eta, \epsilon) \tag{4.1.3}$$

is a $*$-autonomous category. ♡

***Proof.*** Trivial. ❏

We, now, generalize the notion of a transpose functor from the setting of a compact closed category (see Definition 3.9 in [12]) to the setting of a $*$-autonomous category in the obvious way. Note that when our $*$-autonomous category is the compact closed category of natural numbers and matrices, the transpose functor on a matrix is the transpose of the matrix.

**Definition 4.1.7. The Transpose Functor**

Let $\mathfrak{C}$ be a $*$-autonomous category. Define $(-)^\perp \colon \mathfrak{C} \to \mathfrak{C}$ as follows:

(i) given an object $X$ in $\mathfrak{C}$, $(X)^\perp = X^\perp$; and

(ii) given a morphism $f \colon X \to Y$ in $\mathfrak{C}$,

$$\begin{aligned} Y^\perp \xrightarrow{(f)^\perp} X^\perp = Y^\perp &\xrightarrow{(\lambda^{\otimes}_{Y^\perp})^{-1}} \mathbb{1} \otimes Y^\perp \xrightarrow{\eta_X \otimes 1_{Y^\perp}} (X^\perp ⅋ X) \otimes Y^\perp \\ &\xrightarrow{(1_{X^\perp} ⅋ f) \otimes 1_{Y^\perp}} (X^\perp ⅋ Y) \otimes Y^\perp \xrightarrow{\partial^R_{X^\perp, Y, X}} X^\perp ⅋ (Y \otimes Y^\perp) \\ &\xrightarrow{1_{X^\perp} ⅋ \epsilon_Y} X^\perp ⅋ \perp \xrightarrow{\rho_{X^\perp}} X^\perp. \end{aligned} \tag{4.1.4}$$

Then, $(-)^\perp$ is a contravariant functor called the *transpose functor*. ♣

Now, we introduce the linear category (see definition 1.4 in [18]), which is a categorical semantics for the multiplicative and additive fragment of linear logic with negation and exchange. A description of linear categories can be found in [18].

**Definition 4.1.8. Linear Category**

A *linear category* $\mathfrak{C}$ is a symmetric $*$-autonomous category that has

(i) functors $\&\colon \mathfrak{C}^2 \to \mathfrak{C}$ and $\oplus\colon \mathfrak{C}^2 \to \mathfrak{C}$, called *with* and *plus*, respectively;

(ii) a terminal object $\top$, called *top*;

(iii) an initial object $\mathbf{0}$, called *zero*;

(iv) natural iso. $\partial^{\&}_{A,B,C}\colon (A \mathbin{\&} B) \parr C \to (A \parr C) \mathbin{\&} (B \parr C)$ and $\partial^{\oplus}_{A,B,C}\colon (A \oplus B) \otimes C \to (A{\otimes}C){\oplus}(B{\otimes}C)$, called the *with right distributor* and *plus right distributor*, respectively;

(v) natural trans. $\sigma_{0,A,B}\colon A\&B \to A$ and $\sigma_{1,A,B}\colon A\&B \to A$, called the *bottom projections*;

(vi) natural trans. $\rotatebox[origin=c]{180}{$\sigma$}_{0,A,B}\colon A \to A \oplus B$ and $\rotatebox[origin=c]{180}{$\sigma$}_{1,A,B}\colon B \to A \oplus B$, called the *top projections*

s.t. $(\mathfrak{C}, \otimes, \lambda^{\otimes}, \rho^{\otimes}, \alpha^{\otimes}, \sigma^{\otimes}, \mathbb{1}, \oplus, \rotatebox[origin=c]{180}{$\pi$}^0, \rotatebox[origin=c]{180}{$\pi$}^0, \partial^{\oplus}, \mathbf{0})$ and $(\mathfrak{C}, \parr, \lambda^{\parr}, \rho^{\parr}, \alpha^{\parr}, \sigma^{\parr}, \mathbb{1}, \&, \pi^0, \pi^0, \partial^{\&}, \top)$ are symmetric distributive categories and

$$((-) \mathbin{\&} (-)) \cong ((-)^{\perp} \oplus (-)^{\perp})^{\perp}, \qquad \top \cong \mathbf{0}^{\perp}, \qquad \partial^{\&}_{(-),(-),(-)} \cong (\partial^{\oplus}_{(-)^{\perp},(-)^{\perp},(-)^{\perp}})^{\perp},$$

$$\pi^0_{(-),(-)} \cong (\rotatebox[origin=c]{180}{$\pi$}^0_{(-)^{\perp},(-)^{\perp}})^{\perp}, \qquad \pi^1_{(-),(-)} \cong (\rotatebox[origin=c]{180}{$\pi$}^1_{(-)^{\perp},(-)^{\perp}})^{\perp}.$$

♣

To express "bang," we define a coalgebra modality (see Definition 2.1 in [2]).

**Definition 4.1.9. Coalgebra Modality**

Let $\mathfrak{C}$ be a symmetric monoidal category. Then, the quintuple $(!, \delta, \mu, \Delta, e)$, where $!\colon \mathcal{C} \to \mathcal{C}$ is an endofunctor and $\delta\colon A \to !!A$, $\mu\colon !A \to A$, $\Delta_A\colon !A \to !A \otimes !A$, $e_A\colon !A \to I$ are natural transformations, is called a *coalgebra modality* on $\mathfrak{C}$ when $(!, \delta, \nu)$ is a comonad, $(!A, \Delta_A, e_A)$ is a cocommutative comonoid for all objects $A$ in $\mathfrak{C}$ and $\delta$ is a comonoid homomorphism. We refer to ! as *bang*, $\delta$ as *digging*, $\Delta$ as *contraction*, and $e$ as *weakening*.

♣

A linear category with a coalgebra modality has all the connectives an constants of linear logic (with whimper induced by duality). At this point, none of the categorical constructs we have defined are unique to differential categories. Let us change that. We, now, introduce a bialgebra modality (see Definition 3.2 in [2]), a co-dereliction (see Definition 5.1 in [2]), and a differential category (see Theorem 5.5 in [2]).

**Definition 4.1.10. Bialgebra Modality**

Let $\mathfrak{C}$ be an additive symmetric monoidal category. Then, the septuple $(!, \delta, \mu, \Delta, e, \nabla, u)$, where $(!, \delta, \mu, \Delta, e)$ is a coalgebra modality and $\nabla_A\colon !A \otimes !A \to !A$ and $u_A\colon I \to !X$ are natural transformations, is called a *bialgebra modality* on $\mathfrak{C}$ when $(!A, \nabla_A, u_A, \Delta_A, e_A)$ is a bialgebra and the following diagram commutes

$$\begin{tikzcd}
!A \otimes !X \arrow[r, "\nabla_A"] \arrow[dr, "\rho_A \circ (\mu_A \otimes e_A) + \lambda_A \circ (e_A \otimes \mu_A)"'] & !A \arrow[d, "\mu_A"] \\
& A
\end{tikzcd} \tag{4.1.5}$$

for all objects $A$ in $\mathfrak{C}$. We refer to $\nabla$ as *co-contraction* and $u$ as *co-weakening*.

♣

**Definition 4.1.11. Co-Dereliction**

Let $\mathfrak{C}$ be an additive symmetric monoidal category and let $(!, \delta, \mu, \Delta, e, \nabla, u)$ be a bialgebra modality on $\mathfrak{C}$. Then, the natural transformation $\nu_A \colon A \to !A$ is a *co-dereliction* when the four diagrams commute:

(i) *Constant rule*:

$$\begin{tikzcd} A \arrow[r, "\nu_A"] \arrow[dr, "0"'] & !A \arrow[d, "e_A"] \\ & I \end{tikzcd} \tag{4.1.6}$$

(ii) *Leibniz Rule* (or *Product Rule*):

$$\begin{tikzcd} A \arrow[r, "\nu_A"] \arrow[dr, "\rho_A \circ (\nu_A \otimes u_A) + \lambda_A \circ (u_A \otimes \nu_A)"'] & !A \arrow[d, "\Delta_A"] \\ & !A \otimes !A \end{tikzcd} \tag{4.1.7}$$

(iii) *Linear Rule*:

$$\begin{tikzcd} A \arrow[r, "\nu_A"] \arrow[dr, equal] & !A \arrow[d, "\mu_A"] \\ & A \end{tikzcd} \tag{4.1.8}$$

(iv) *Chain Rule*:

$$\begin{tikzcd} !A \otimes A \arrow[r, "\Delta_A \otimes \nu_A"] \arrow[d, "1_A \otimes \nu_A"'] & (!A \otimes !A) \otimes !A \arrow[r, "\alpha^{-1}_{!A,!A,!A}"] & !A \otimes (!A \otimes !A) \arrow[r, "1_A \otimes \nabla_A"] & !A \otimes !A \arrow[d, "\delta_A \otimes \nu_{!A}"] \\ !A \otimes !A \arrow[d, "\nabla_A"'] & & & !!A \otimes !!A \arrow[d, "\nabla_{!A}"] \\ !A \arrow[rrr, "\delta_A"'] & & & !!A \end{tikzcd} \tag{4.1.9}$$

♣

**Definition 4.1.12. Differential Category**

A differential category $\mathfrak{C}$ is an additive $*$-autonomous category with a bialgebra modality $(!, \delta, \mu, \Delta, e, \nabla, u)$ and a co-dereliction $\nu$ on the additive symmetric monoidal category $(\mathfrak{C}, \otimes, \lambda^{\otimes}, \rho^{\otimes}, \alpha^{\otimes}, \sigma^{\otimes}, \mathbb{1})$. We define

(i) an endofunctor $? \colon \mathfrak{C} \to \mathfrak{C}$, called *whimper*, by $?(-) = (!(-)^{\perp})^{\perp}$,

(ii) a natural trans. $\rotatebox[origin=c]{180}{$\delta$}_A \colon ??A \to A$, called *rotated digging*, by $\rotatebox[origin=c]{180}{$\delta$}_{(-)} \cong (\rotatebox[origin=c]{180}{$\delta$}_{(-)^{\perp}})^{\perp}$,

(iii) a natural trans. $\rotatebox[origin=c]{180}{$\mu$}_A \colon A \to ?A$, called *rotated dereliction*, by $\rotatebox[origin=c]{180}{$\mu$}_{(-)} \cong (\mu_{(-)^{\perp}})^{\perp}$,

(iv) a natural trans. $\nabla'_A \colon ?A \parr ?A \to ?A$, called *rotated contraction*, by $\nabla'_{(-)} \cong (\Delta_{(-)^{\perp}})^{\perp}$,

(v) a natural trans. $\rotatebox[origin=c]{180}{$e$}_A \colon \perp \to ?A$, called *rotated weakening*, by $\rotatebox[origin=c]{180}{$e$}_{(-)} \cong (e_{(-)^{\perp}})^{\perp}$,

(vi) a natural trans. $\Delta'_A : ?A \to ?A \parr ?A$, called *rotated co-contraction*, by $\Delta'_{(-)} \cong (\nabla^!_{(-)^\perp})^\perp$,

(vii) a natural trans. $n_A : ?A \to \bot$, called *rotated co-weakening*, by $n_{(-)} \cong (u_{(-)^\perp})^\perp$, and

(viii) a natural trans. $\rotatebox[origin=c]{180}{$\nu$}_A : ?A \to A$, called *rotated co-dereliction*, by $\rotatebox[origin=c]{180}{$\nu$} \cong (\nu_{(-)^\perp})^\perp$. ♣

The most interesting morphisms in a differential category are those formed only by the structure morphisms. Such morphisms are called differential interaction nets.

**Definition 4.1.13. Differential Interaction Net**

Let $\mathfrak{C}$ be a differential category. A *differential interaction net* is any morphism in $\mathfrak{C}$ built only from $1, \otimes, \lambda^\otimes, \rho^\otimes, \alpha^\otimes, \sigma^\otimes, \mathbb{1}, \parr, \lambda^\parr, \rho^\parr, \alpha^\parr, \sigma^\parr, \bot, \partial^L, \partial^R, \eta, \epsilon, \delta, \mu, \Delta, e, \nabla, u, \rotatebox[origin=c]{180}{$\delta$}, \rotatebox[origin=c]{180}{$\eta$}, \nabla', \rotatebox[origin=c]{180}{$\partial$}, \Delta', n$. ♣

This thesis will not elaborate on the way in which differential interaction nets provide a synthetic formulation of differentiation of analytic model. The interested reader can read [7] to see this. However, we will mention two important facts about differential interaction nets. First, co-dereliction acts as a kind-of generalized derivative. Second, there exists a map $(-)^\bullet_{\mathfrak{C}}$, as defined in Section 6.3 in [7], from Taylor $\lambda$-terms to differential interaction nets in a differential category $\mathfrak{C}$ such that the following property holds.

**Theorem 4.1.14. Equivalence of Differential Categories and Taylor $\lambda$-Calculus**

Given complex Taylor poly-terms $\tau, \tau' \in \mathcal{R}\langle \Lambda^!_{t,\mathcal{V}} \rangle$ such that $\tau \to_\beta \tau'$, $(\tau)^\bullet_{\mathfrak{C}} = (\tau')^\bullet_{\mathfrak{C}}$ in any differential category $\mathfrak{C}$. ♡

***Proof.*** This is mentioned in the last sentence of Section 6.3 in [7]. ❑

For this thesis, we will spend a lot of time dealing with a differential interaction net called the co-deriving transform, which is the morphism formed by first performing a contraction and then performing a dereliction.

**Definition 4.1.15. Co-deriving Transform**

Given a differential category $\mathfrak{C}$, define the natural transformations $d^\circ_A : !A \to !A \otimes A$ and $d'^\circ_A : !A \to A \otimes !A$ by $d^\circ_A = (1_{!A} \otimes \mu_A) \circ \Delta_A$ and $d'^\circ_A = (\mu_A \otimes 1_{!A}) \circ \Delta_A$. We refer to $d^\circ$ as the *left co-deriving transform* and $d'^\circ$ as the *right co-deriving transform*. Similarly, define the natural transformations $\rotatebox[origin=c]{180}{$d$}'^\circ_A : A \parr ?A \to ?A$ and $\rotatebox[origin=c]{180}{$d$}^\circ_A : ?A \parr A \to ?A$ by $\rotatebox[origin=c]{180}{$d$}'^\circ_A = (d^\circ_{A^\perp})^\perp$ and $\rotatebox[origin=c]{180}{$d$}^\circ_A = (d'^\circ_{A^\perp})^\perp$. We refer to $\rotatebox[origin=c]{180}{$d$}'^\circ$ as the *left rotated co-deriving transform* and $\rotatebox[origin=c]{180}{$d$}$ as the *right rotated co-deriving transform*. ♣

We are now ready to define a resource category. A resource category is just a differential category subject to essentially one new idea: If $!A$ is an infinite spout of a resource $A$ then, in essence, $!A$ should be equivalent to either doing nothing or pulling one instance of $A$ out of the spout $!A$. That is,

$$!A = I \mathbin{\&} (!A \otimes A) = I \mathbin{\&} (A \otimes !A) \tag{4.1.10}$$

for all objects (i.e., *resources* $A$) in the resource category. This idea is motivated by the description of the message corresponding to iteration discussed after Definition 12 in [16]. Note that for simplicity we will assume that the "times" and "par" monoidal structures in the resource theory are the same.

**Definition 4.1.16. Resource Category**

A resource category $\mathfrak{C}$ is a differential and linear category whose $\star$-autonomous structure is given by a strict compact closed category (as described in Theorem 4.1.6) such that the following holds for all objects $A$ in $\mathfrak{C}$:

$$!A = I \,\&\, (!A \otimes A), \qquad e_A = \sigma_{0,I,!A\otimes A}, \qquad d^{\circ}_A = \sigma_{1,I,!A\otimes A}, \tag{4.1.11a}$$

$$!A = I \,\&\, (A \otimes !A), \qquad e_A = \sigma_{0,I,A\otimes !A}, \qquad d'^{\circ}_A = \sigma_{1,I,A\otimes !A}. \tag{4.1.11b}$$

♣

The coherence conditions for "bang" specified by 4.1.11a and 4.1.11b induce similar coherence conditions for "whimper."

**Lemma 4.1.17**

Let $\mathfrak{C}$ be a resource category. Given an object $A$ in $\mathfrak{C}$,

$$?A = I \oplus (A \otimes ?A), \qquad \reflectbox{$e$}_A = \reflectbox{$\sigma$}_{0,I,A\otimes ?A}, \qquad p'^{\circ}_A = \reflectbox{$\sigma$}_{1,I,A\otimes ?A}, \tag{4.1.12a}$$

$$?A = I \oplus (?A \otimes A), \qquad \reflectbox{$e$}_A = \sigma_{0,I,?A\otimes A}, \qquad p^{\circ}_A = \sigma_{1,I,?A\otimes A}. \tag{4.1.12b}$$

♡

***Proof.*** By (4.1.11a) on $A^{\perp}$,

$$\begin{aligned}
!(A^{\perp}) &= I \,\&\, (!(A^{\perp}) \otimes A^{\perp}) & e_{A^{\perp}} &= \sigma_{0,I,!(A^{\perp})\otimes A^{\perp}} & d^{\circ}_{A^{\perp}} &= \sigma_{1,I,!(A^{\perp})\otimes A^{\perp}} \\
(!(A^{\perp}))^{\perp} &= (I \,\&\, (!(A^{\perp}) \otimes A^{\perp}))^{\perp} & (e_{A^{\perp}})^{\perp} &= (\sigma_{0,I,!(A^{\perp})\otimes A^{\perp}})^{\perp} & (d^{\circ}_{A^{\perp}})^{\perp} &= (\sigma_{1,I,!(A^{\perp})\otimes A^{\perp}})^{\perp} \\
?A &= I^{\perp} \oplus (!(A^{\perp}) \otimes A^{\perp}))^{\perp} & \reflectbox{$e$}_A &= \reflectbox{$\sigma$}_{0,I^{\perp},(!(A^{\perp})\otimes A^{\perp})^{\perp}} & p'^{\circ}_A &= \reflectbox{$\sigma$}_{1,I^{\perp},(!(A^{\perp})\otimes A^{\perp})^{\perp}} \\
?A &= I \oplus A^{\perp\perp} \otimes (!(A^{\perp}))^{\perp} & \reflectbox{$e$}_A &= \reflectbox{$\sigma$}_{0,I^{\perp},A^{\perp\perp}\otimes(!(A^{\perp}))^{\perp}} & p'^{\circ}_A &= \reflectbox{$\sigma$}_{1,I^{\perp},A^{\perp\perp}\otimes(!(A^{\perp}))^{\perp}} \\
?A &= I \oplus (A \otimes ?A) & \reflectbox{$e$}_A &= \reflectbox{$\sigma$}_{0,I,A\otimes ?A} & p'^{\circ}_A &= \reflectbox{$\sigma$}_{1,I,A\otimes ?A},
\end{aligned}$$

and By (4.1.11b) on $A^{\perp}$,

$$\begin{aligned}
!(A^{\perp}) &= I \,\&\, (A^{\perp} \otimes !(A^{\perp})) & e_{A^{\perp}} &= \sigma_{0,I,A^{\perp}\otimes !(A^{\perp})} & d'^{\circ}_{A^{\perp}} &= \sigma_{1,I,A^{\perp}\otimes !(A^{\perp})} \\
(!(A^{\perp}))^{\perp} &= (I \,\&\, (A^{\perp} \otimes !(A^{\perp})))^{\perp} & (e_{A^{\perp}})^{\perp} &= (\sigma_{0,I,A^{\perp}\otimes !(A^{\perp})})^{\perp} & (d'^{\circ}_{A^{\perp}})^{\perp} &= (\sigma_{1,I,A^{\perp}\otimes !(A^{\perp})})^{\perp} \\
?A &= I^{\perp} \oplus (A^{\perp} \otimes !(A^{\perp})))^{\perp} & \reflectbox{$e$}_A &= \reflectbox{$\sigma$}_{0,I^{\perp},(A^{\perp}\otimes !(A^{\perp}))^{\perp}} & p^{\circ}_A &= \reflectbox{$\sigma$}_{1,I^{\perp},(A^{\perp}\otimes !(A^{\perp}))^{\perp}} \\
?A &= I \oplus (!(A^{\perp}))^{\perp} \otimes A^{\perp\perp} & \reflectbox{$e$}_A &= \reflectbox{$\sigma$}_{0,I^{\perp},(!(A^{\perp}))^{\perp}\otimes A^{\perp\perp}} & p^{\circ}_A &= \reflectbox{$\sigma$}_{1,I^{\perp},(!(A^{\perp}))^{\perp}\otimes A^{\perp\perp}} \\
?A &= I \oplus (?A \otimes A) & \reflectbox{$e$}_A &= \reflectbox{$\sigma$}_{0,I,?A\otimes A} & p^{\circ}_A &= \reflectbox{$\sigma$}_{1,I,?A\otimes A},
\end{aligned}$$

This completes the proof. ❏

We define a resource theory to be a resource category. With a good definition of a resource theory, we are ready to turn to defining a model for asynchronous processes which manipulate the resources in a resource theory, which is what we will do in the next section.

## 4.2 Adding a Compact Closed Structure on the Category of Resource Transducers

Before we can define a model for asynchronous processes, we first need to recognize that in order define an asynchronous processes, we need to have some notion of messages that communicate the state of some resources. In [16], the messages corresponding to a resource theory $\mathfrak{C}$ is defined in

Definition 12 and is called the $\mathfrak{C}$-valued exchanges $\mathfrak{C}_*^{\triangleright\blacktriangleleft}$. The intuition behind the $\mathfrak{C}$-valued exchanges $\mathfrak{C}_*^{\triangleright\blacktriangleleft}$ is that the message $A^\triangleright$ tells the system to send the resource $A$ to the left, the message $A^\blacktriangleleft$ tells the system to send the resource $A$ to the right, the message $U \otimes V$ is the concatenation of the messages $U$ and $V$, $U + V$ is the message that could either send message $U$ or message $V$ from the left, $U \times V$ is the message that could either send message $U$ or message $V$ from the right, $U^+$ is the iteration of the message $U$ from the left, and $U^\times$ is the iteration of the message $U$ from the right.

**Definition 4.2.1. $\mathfrak{C}$-Valued Exchanges**

Let $\mathfrak{C}$ be a resource category. The $\mathfrak{C}$*-valued exchange* $\mathfrak{C}_*^{\triangleright\blacktriangleleft}$ *with choice and iteration* is the set with elements generated by:

$$\frac{A \in \mathrm{Ob}(\mathfrak{C})}{A^\triangleright \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}} \qquad \frac{A \in \mathrm{Ob}(\mathfrak{C})}{A^\blacktriangleleft \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}} \qquad \frac{}{I \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}} \qquad \frac{U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft} \qquad V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}}{U \otimes V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}} \qquad \frac{U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft} \qquad V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}}{U + V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}}$$

$$\frac{U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft} \qquad V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}}{U \times V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}} \qquad \frac{U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}}{U^+ \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}} \qquad \frac{U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}}{U^\times \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}}$$

subject to the following equations:

$$I \otimes U = U \qquad U \otimes I = U \qquad (U \otimes W) \otimes V = U \otimes (W \otimes V) \qquad U^+ = I + (U \otimes U^+)$$

$$U^\times = I \times (U \otimes U^\times)$$

♣

As we mentioned above, every message should communicates the state of some resources. In Definition 4.2.2, we make this precise by defining the forward-time underlying resource corresponding to a message. Note that, following Chad Nester's interpretation in [14], we are going to interpret sending information to the left as moving forward in time, sending information to the right as moving back in time, and if $A$ represents having a resource, $A^\perp$ represents borrowing a resource $A$ we do not yet possess.

**Definition 4.2.2. The Forward-Time Underlying Resource**

Let $\mathfrak{C}$ be a resource theory. The function $\widetilde{(-)} \colon \mathfrak{C}_*^{\triangleright\blacktriangleleft} \to \mathrm{Ob}(\mathfrak{C})$ from the $\mathfrak{C}$-valued exchanges $\mathfrak{C}_*^{\triangleright\blacktriangleleft}$ to the objects of $\mathfrak{C}$ is defined inductively as follows:

$$\widetilde{I} = I, \tag{4.2.1a}$$

$$\widetilde{A^\triangleright} = A, \tag{4.2.1b}$$

$$\widetilde{A^\blacktriangleleft} = A^\perp, \tag{4.2.1c}$$

$$\widetilde{U \otimes V} = \widetilde{V} \otimes \widetilde{U}, \tag{4.2.1d}$$

$$\widetilde{U + V} = \widetilde{U} \oplus \widetilde{V}, \tag{4.2.1e}$$

$$\widetilde{U \times V} = \widetilde{U} \,\&\, \widetilde{V}, \tag{4.2.1f}$$

$$\widetilde{U^+} = ?\widetilde{U}, \tag{4.2.1g}$$

$$\widetilde{U^\times} = !\widetilde{U}, \tag{4.2.1h}$$

where $A \in \mathrm{Ob}(\mathfrak{C})$ is an object in $\mathfrak{C}$ and $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ are $\mathfrak{C}$-valued exchanges. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$, we call $\widetilde{U}$ the *forward-time underlying resource* of $U$.

♣

Now, we define the free cornering with choice and iteration of a resource theory $\mathfrak{C}$. The free cornering will be a single-object double category whose horizontal monoid will be resources, its vertical monoid will be messages and a cell represents a processes that transforms a resource on the cell's top edge using the information provided by the messages on the cell's side edges to output the resource on the cell's bottom edge.

**Definition 4.2.3. The Free Cornering With Choice and Iteration**

The free cornering $\ulcorner\mathfrak{C}\urcorner$ with choice and iteration of a resource category $\mathfrak{C}$ is an *additive single-object double category*, where the horizontal morphisms are $\mathrm{Ob}(\mathfrak{C})$, the vertical morphisms are $\mathfrak{C}_*^{\triangleright\blacktriangleleft}$, and the 2-morphisms (or *cells*) are defined in the following.

(i) For each morphism $f \colon A \to B$ in $\mathfrak{C}$, there exists the inclusion cell $\ulcorner f\urcorner \in \ulcorner\mathfrak{C}\urcorner(I\,{}^{A}_{B}\,I)$. These inclusion cells are subject to the equations:

$$\ulcorner g \circ f\urcorner = \frac{\ulcorner f\urcorner}{\ulcorner g\urcorner}, \quad \ulcorner 1_A\urcorner = 1_A, \quad \ulcorner f \otimes g\urcorner = \ulcorner f\urcorner | \ulcorner g\urcorner, \quad \ulcorner f + g\urcorner = \ulcorner f\urcorner + \ulcorner g\urcorner, \quad \ulcorner 0\urcorner = 0, \tag{4.2.2}$$

where $f, g$ are morphisms in $\mathfrak{C}$ and $A$ is an object in $\mathfrak{C}$.

(ii) For each object $A$ in $\mathfrak{C}$, there exist the four cells $\mathrm{get}_L^A \colon \ulcorner\mathfrak{C}\urcorner(A\,{}^{I}_{A^{\triangleright}}\,I)$, $\mathrm{put}_R^A \colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{A}_{I}\,A^{\triangleright})$, $\mathrm{get}_R^A \colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{I}_{A}\,A^{\blacktriangleleft})$, and $\mathrm{put}_L^A \colon \ulcorner\mathfrak{C}\urcorner(A^{\blacktriangleleft}\,{}^{A}_{I}\,I)$, which are called *corner cells* and are subject to the equations:

$$\frac{\mathrm{get}_L^A}{\mathrm{put}_R^A} = \mathrm{id}_{A^{\triangleright}}, \quad \mathrm{put}_R^A \,|\, \mathrm{get}_L^A = 1_A = \mathrm{put}_L^A \,|\, \mathrm{get}_R^A, \quad \frac{\mathrm{get}_R^A}{\mathrm{put}_L^A} = \mathrm{id}_{A^{\blacktriangleleft}}, \tag{4.2.3}$$

which are called the *yanking equations*.

(iii) For each pair of $\mathfrak{C}$-valued exchanges $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$, there exist the cells $\amalg_{0,U,V} \colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{I}_{I}\,U{+}V)$ and $\amalg_{1,U,V} \colon \ulcorner\mathfrak{C}\urcorner(V\,{}^{I}_{I}\,U{+}V)$, which are called *injection cells*. Moreover, for each pair of cells $\alpha \colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{A}_{B}\,W)$ and $\beta \colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{A}_{B}\,W)$, there exists a unique cell $\alpha \boxplus \beta \colon \ulcorner\mathfrak{C}\urcorner(U{+}V\,{}^{A}_{V}\,W)$, called the *left choice cell*, such that

$$\amalg_{0,V,U} \,|\alpha \boxplus \beta = \alpha \quad \text{and} \quad \amalg_{1,V,U} \,|\alpha \boxplus \beta = \beta. \tag{4.2.4}$$

(iv) For each pair of $\mathfrak{C}$-valued exchanges $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$, there exist the cells $\pi_{0,U,V} \colon \ulcorner\mathfrak{C}\urcorner(U{\times}V\,{}^{I}_{I}\,U)$ and $\pi_{1,U,V} \colon \ulcorner\mathfrak{C}\urcorner(U{\times}V\,{}^{I}_{I}\,V)$, which are called *projection cells*. Moreover, for each pair of cells $\alpha \colon \ulcorner\mathfrak{C}\urcorner(W\,{}^{A}_{B}\,U)$ and $\beta \colon \ulcorner\mathfrak{C}\urcorner(W\,{}^{A}_{B}\,V)$, there exists a unique cell $\alpha \boxtimes \beta \colon \ulcorner\mathfrak{C}\urcorner(W\,{}^{A}_{V}\,U{\times}V)$, called the *right choice cell*, such that

$$\alpha \boxtimes \beta | \pi_{0,V,U} = \alpha \quad \text{and} \quad \alpha \boxtimes \beta | \pi_{1,V,U} = \beta. \tag{4.2.5}$$

(v) For each pair of cells $\alpha \colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{A}_{C}\,V)$ and $\beta \colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{B}_{C}\,V)$, there exists a unique cell $[\alpha, \beta] \colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{A \oplus B}_{C}\,V)$, called the *top choice cell*, such that

$$\frac{\ulcorner \varrho_{0,A,B}\urcorner}{[\alpha, \beta]} = \alpha \quad \text{and} \quad \frac{\ulcorner \varrho_{1,A,B}\urcorner}{[\alpha, \beta]} = \beta. \tag{4.2.6}$$

(vi) For each pair of cells $\alpha\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{C}_{A}\,V)$ and $\beta\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{C}_{B}\,V)$, there exists a unique cell $\langle\alpha,\beta\rangle\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{C}_{A\&B}\,V)$, called the *bottom choice cell*, such that

$$\frac{\langle\alpha,\beta\rangle}{\ulcorner\sigma_{0,A,B}\urcorner} = \alpha \quad\text{and}\quad \frac{\langle\alpha,\beta\rangle}{\ulcorner\sigma_{1,A,B}\urcorner} = \beta. \tag{4.2.7}$$

(vii) For each quadruple of cells $\alpha\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{A}_{A}\,V)$, $\theta\colon \ulcorner\mathfrak{C}\urcorner(W\,{}^{A}_{B}\,X)$, $\iota\colon \ulcorner\mathfrak{C}\urcorner(V{\otimes}X\,{}^{I}_{C}\,X)$, and $\kappa\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{B\otimes C}_{B}\,I)$, there exists a unique cell $\alpha^{+}_{\theta,\iota,\kappa}\colon \ulcorner\mathfrak{C}\urcorner(U^{+}{\otimes}W\,{}^{A}_{B}\,X)$, called the *left iteration cell*, such that

$$\frac{\rotatebox{180}{$\pi$}_{0,U}}{\mathrm{id}_W}\Big|\alpha^{+}_{\theta,\iota,\kappa} = \theta \quad\text{and}\quad \frac{\rotatebox{180}{$\pi$}_{1,U}}{\mathrm{id}_W}\Big|\alpha^{+}_{\theta,\iota,\kappa} = \frac{\frac{\alpha}{\alpha^{+}_{\theta,\iota,\kappa}}\Big|\iota}{\kappa}, \tag{4.2.8}$$

where $\rotatebox{180}{$\pi$}_{0,U}\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{I}_{I}\,U^{+})$ is defined by $\rotatebox{180}{$\pi$}_{0,U} \coloneqq \rotatebox{180}{$\pi$}_{0,I,U\otimes U^{+}}$ and is called the *left stop cell* and $\rotatebox{180}{$\pi$}_{1,U}\colon \ulcorner\mathfrak{C}\urcorner(U{\otimes}U^{+}\,{}^{I}_{I}\,U^{+})$ is defined by $\rotatebox{180}{$\pi$}_{1,U} \coloneqq \rotatebox{180}{$\pi$}_{1,I,U\otimes U^{+}}$ and is called the *left step cell.*

(viii) For each quadruple of cells $\alpha\colon \ulcorner\mathfrak{C}\urcorner(V\,{}^{A}_{A}\,U)$, $\theta\colon \ulcorner\mathfrak{C}\urcorner(X\,{}^{A}_{B}\,W)$, $\iota\colon \ulcorner\mathfrak{C}\urcorner(X\,{}^{I}_{C}\,V{\otimes}X)$, and $\kappa\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{C\otimes B}_{B}\,I)$, there exists a unique cell $\alpha^{\times}_{\theta,\iota,\kappa}\colon \ulcorner\mathfrak{C}\urcorner(X\,{}^{A}_{B}\,U^{\times}{\otimes}W)$, called the *right iteration cell*, such that

$$\alpha^{\times}_{\theta,\iota,\kappa}\Big|\frac{\pi_{0,U}}{\mathrm{id}_W} = \theta \quad\text{and}\quad \alpha^{\times}_{\theta,\iota,\kappa}\Big|\frac{\pi_{1,U}}{\mathrm{id}_W} = \frac{\iota\Big|\frac{\alpha}{\alpha^{\times}_{\theta,\iota,\kappa}}}{\kappa}, \tag{4.2.9}$$

where $\pi_{0,U}\colon \ulcorner\mathfrak{C}\urcorner(U^{\times}\,{}^{I}_{I}\,I)$ is defined by $\pi_{0,U} \coloneqq \pi_{0,I,U\otimes U^{\times}}$ and is called the *right stop cell* and $\pi_{1,U}\colon \ulcorner\mathfrak{C}\urcorner(U^{\times}\,{}^{I}_{I}\,U{\otimes}U^{\times})$ is defined by $\pi_{1,U} \coloneqq \pi_{1,I,U\otimes U^{\times}}$ and is called the *right step cell.*

(ix) For each quadruple of cells $\alpha\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{A}_{B}\,U)$, $\theta\colon \ulcorner\mathfrak{C}\urcorner(V\,{}^{B}_{C}\,U)$, $\iota\colon \ulcorner\mathfrak{C}\urcorner(W\,{}^{C\otimes B}_{C}\,I)$, and $\kappa\colon \ulcorner\mathfrak{C}\urcorner(V\,{}^{I}_{I}\,V{\otimes}W)$, there exists a unique cell $\alpha^{?}_{\theta,\iota,\kappa}\colon \ulcorner\mathfrak{C}\urcorner(V\,{}^{B\otimes ?A}_{C}\,U)$, called the *top iteration cell*, such that

$$\frac{1_B\big|\ulcorner \reflectbox{$\partial$}_{\tilde{A}}\urcorner}{\alpha^{?}_{\theta,\iota,\kappa}} = \theta \quad\text{and}\quad \frac{1_B\big|\ulcorner p_{\tilde{A}}\urcorner}{\alpha^{?}_{\theta,\iota,\kappa}} = \kappa\Big|\frac{\alpha^{?}_{\theta,\iota,\kappa}|\alpha}{\iota}. \tag{4.2.10}$$

(x) For each quadruple of cells $\alpha\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{A}_{B}\,U)$, $\theta\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{B}_{C}\,V)$, $\iota\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{B\otimes C}_{C}\,W)$, and $\kappa\colon \ulcorner\mathfrak{C}\urcorner(V{\otimes}W\,{}^{I}_{I}\,V)$, there exists a unique cell $\alpha'^{?}_{\theta,\iota,\kappa}\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{?A\otimes B}_{C}\,V)$, called the *reversed top iteration cell*, such that

$$\frac{\ulcorner \reflectbox{$\partial$}_{\tilde{A}}\urcorner\big|1_B}{\alpha'^{?}_{\theta,\iota,\kappa}} = \theta \quad\text{and}\quad \frac{\ulcorner p'_{\tilde{A}}\urcorner\big|1_B}{\alpha'^{?}_{\theta,\iota,\kappa}} = \frac{\alpha|\alpha'^{?}_{\theta,\iota,\kappa}}{\iota}\Big|\kappa. \tag{4.2.11}$$

(xi) For each quadruple of cells $\alpha\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{B}_{A}\,U)$, $\theta\colon \ulcorner\mathfrak{C}\urcorner(V\,{}^{C}_{B}\,U)$, $\iota\colon \ulcorner\mathfrak{C}\urcorner(W\,{}^{C}_{C\otimes B}\,I)$, and $\kappa\colon \ulcorner\mathfrak{C}\urcorner(V\,{}^{I}_{I}\,W{\otimes}V)$, there exists a unique cell $\alpha^{!}_{\theta,\iota,\kappa}\colon \ulcorner\mathfrak{C}\urcorner(V\,{}^{C}_{B\otimes !A}\,U)$, called the *bottom iteration cell*, such that

$$\frac{\alpha^{!}_{\theta,\iota,\kappa}}{1_B\big|\ulcorner e_{\tilde{A}}\urcorner} = \theta \quad\text{and}\quad \frac{\alpha^{!}_{\theta,\iota,\kappa}}{1_B\big|\ulcorner d^{\circ}_{\tilde{A}}\urcorner} = \kappa\Big|\frac{\iota}{\alpha^{!}_{\theta,\iota,\kappa}|\alpha}. \tag{4.2.12}$$

(xii) For each quadruple of cells $\alpha\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{B}_{A}\,U)$, $\theta\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{C}_{B}\,V)$, $\iota\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{C}_{B\otimes C}\,W)$, and $\kappa\colon \ulcorner\mathfrak{C}\urcorner(W\otimes V\,{}^{I}_{I}\,V)$, there exists a unique cell $\alpha'^{!}_{\theta,\iota,\kappa}\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{C}_{!A\otimes B}\,V)$, called the *reversed bottom iteration cell*, such that

$$\frac{\alpha'^{!}_{\theta,\iota,\kappa}}{\ulcorner e_{\tilde{A}}\urcorner|1_B} = \theta \quad \text{and} \quad \frac{\alpha'^{!}_{\theta,\iota,\kappa}}{\ulcorner d'^{\circ}_{\tilde{A}}\urcorner|1_B} = \frac{\iota}{\alpha|\alpha'^{!}_{\theta,\iota,\kappa}}|\kappa. \tag{4.2.13}$$

♣

## Definition 4.2.4. The Category of Resource Transducers

Given a resource category $\mathfrak{C}$, the monoidal category $\mathbf{H}\ulcorner\mathfrak{C}\urcorner$ formed by the horizontal cells of $\ulcorner\mathfrak{C}\urcorner$ is called the *category of resource transducers*.

♣

Recall, again, that there is a correspondence between messages and underlying resources, so ideally, we should have cells in our free cornering that allow us to convert readily between a message and it underlying. These cells are called the *generalized corners*.

## Definition 4.2.5. The Generalized Corners

Let $\mathfrak{C}$ be a resource theory. We define $\mathrm{get}^{U}_{L}\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{I}_{\tilde{U}}\,I)$, $\mathrm{put}^{U}_{R}\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{\tilde{U}}_{I}\,U)$, $\mathrm{get}^{U}_{R}\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{I}_{\tilde{U}^{\perp}}\,U)$, and $\mathrm{put}^{U}_{L}\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{\tilde{U}^{\perp}}_{I}\,I)$ inductively. Let $A \in \mathrm{Ob}\,\mathfrak{C}$ be an object in $\mathfrak{C}$ and let $U, V \in \mathfrak{C}^{\triangleright\blacktriangleleft}_{*}$ be $\mathfrak{C}$-valued exchanges.

(i) We have the following base cases:

$$\mathrm{get}^{I}_{L} = 1_I, \qquad \mathrm{get}^{A^{\triangleright}}_{L} = \mathrm{get}^{A}_{L}, \qquad \mathrm{get}^{A^{\blacktriangleleft}}_{L} = \frac{\ulcorner\eta'_A\urcorner}{\mathrm{put}^{A}_{L}|1_{A^{\perp}}}, \tag{4.2.14a}$$

$$\mathrm{put}^{I}_{R} = 1_I, \qquad \mathrm{put}^{A^{\triangleright}}_{R} = \mathrm{put}^{A}_{R}, \qquad \mathrm{put}^{A^{\blacktriangleleft}}_{R} = \frac{1_{A^{\perp}}|\mathrm{get}^{A}_{R}}{\ulcorner\epsilon'_A\urcorner}, \tag{4.2.14b}$$

$$\mathrm{get}^{I}_{R} = 1_I, \qquad \mathrm{get}^{A^{\triangleright}}_{R} = \frac{\ulcorner\eta_A\urcorner}{1_{A^{\perp}}|\mathrm{put}^{A}_{R}}, \qquad \mathrm{get}^{A^{\blacktriangleleft}}_{R} = \mathrm{get}^{A}_{R}, \tag{4.2.14c}$$

$$\mathrm{put}^{I}_{L} = 1_I, \qquad \mathrm{put}^{A^{\triangleright}}_{L} = \frac{\mathrm{get}^{A}_{L}|1_{A^{\perp}}}{\ulcorner\epsilon_A\urcorner}, \qquad \mathrm{put}^{A^{\blacktriangleleft}}_{L} = \mathrm{put}^{A}_{L}. \tag{4.2.14d}$$

(ii) $\mathrm{get}^{U\otimes V}_{L}\colon \ulcorner\mathfrak{C}\urcorner(U\otimes V\,{}^{I}_{\tilde{V}\otimes\tilde{U}}\,I)$ is defined by

$$\mathrm{get}^{U\otimes V}_{L} = \frac{\mathrm{get}^{U}_{L}}{\mathrm{get}^{V}_{L}|1_{\tilde{U}}}. \tag{4.2.15}$$

(iii) $\mathrm{put}^{U\otimes V}_{R}\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{\tilde{V}\otimes\tilde{U}}_{I}\,U\otimes V)$ is defined by

$$\mathrm{put}^{U\otimes V}_{R} = \frac{1_{\tilde{V}}|\mathrm{put}^{U}_{R}}{\mathrm{put}^{V}_{R}}. \tag{4.2.16}$$

(iv) $\mathrm{get}^{U\otimes V}_{R}\colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{I}_{\tilde{U}^{\perp}\otimes\tilde{V}^{\perp}}\,U\otimes V)$ is defined by

$$\mathrm{get}^{U\otimes V}_{R} = \frac{\mathrm{get}^{U}_{R}}{1_{\tilde{U}^{\perp}}|\mathrm{get}^{V}_{R}}. \tag{4.2.17}$$

(v) $\mathrm{put}_L^{U\otimes V} : \ulcorner\mathfrak{C}\lrcorner\left(U{\otimes}V \,\substack{\tilde U^\perp\otimes\tilde V^\perp\\ I}\, I\right)$ is defined by

$$\mathrm{put}_L^{U\otimes V} = \frac{\mathrm{put}_L^U \,|1_{\tilde V^\perp}}{\mathrm{put}_L^V}. \tag{4.2.18}$$

(vi) $\mathrm{get}_L^{U+V} : \ulcorner\mathfrak{C}\lrcorner\left(U{+}V \,\substack{I\\ \tilde U\oplus\tilde V}\, I\right)$ is the unique left choice cell such that

$$\iota_{0,U,V}\,|\,\mathrm{get}_L^{U+V} = \frac{\mathrm{get}_L^U}{\ulcorner\rho_{0,\tilde U,\tilde V}\lrcorner} \quad\text{and}\quad \iota_{1,U,V}\,|\,\mathrm{get}_L^{U+V} = \frac{\mathrm{get}_L^V}{\ulcorner\rho_{1,\tilde U,\tilde V}\lrcorner}. \tag{4.2.19}$$

(vii) $\mathrm{put}_R^{U+V} : \ulcorner\mathfrak{C}\lrcorner\left(I \,\substack{\tilde U\oplus\tilde V\\ I}\, U{+}V\right)$ is the unique top choice cell such that

$$\frac{\ulcorner\rho_{0,\tilde U,\tilde V}\lrcorner}{\mathrm{put}_R^{U+V}} = \mathrm{put}_R^U\,|\,\iota_{0,U,V} \quad\text{and}\quad \frac{\ulcorner\rho_{1,\tilde U,\tilde V}\lrcorner}{\mathrm{put}_R^{U+V}} = \mathrm{put}_R^V\,|\,\iota_{1,U,V}\,. \tag{4.2.20}$$

(viii) $\mathrm{get}_R^{U+V} : \ulcorner\mathfrak{C}\lrcorner\left(I \,\substack{I\\ \tilde U^\perp\&\tilde V^\perp}\, U{+}V\right)$ is the unique bottom choice cell such that

$$\frac{\mathrm{get}_R^{U+V}}{\ulcorner\sigma_{0,\tilde U^\perp,\tilde V^\perp}\lrcorner} = \mathrm{get}_R^U\,|\,\iota_{0,U,V} \quad\text{and}\quad \frac{\mathrm{get}_R^{U+V}}{\ulcorner\sigma_{1,\tilde U^\perp,\tilde V^\perp}\lrcorner} = \mathrm{get}_R^V\,|\,\iota_{1,U,V}\,. \tag{4.2.21}$$

(ix) $\mathrm{put}_L^{U+V} : \ulcorner\mathfrak{C}\lrcorner\left(U{+}V \,\substack{\tilde U^\perp\&\tilde V^\perp\\ I}\, I\right)$ is the unique left choice cell such that

$$\iota_{0,U,V}\,|\,\mathrm{put}_L^{U+V} = \frac{\ulcorner\sigma_{0,\tilde U^\perp,\tilde V^\perp}\lrcorner}{\mathrm{put}_L^U} \quad\text{and}\quad \iota_{1,U,V}\,|\,\mathrm{put}_L^{U+V} = \frac{\ulcorner\sigma_{1,\tilde U^\perp,\tilde V^\perp}\lrcorner}{\mathrm{put}_L^V}. \tag{4.2.22}$$

(x) $\mathrm{get}_L^{U\times V} : \ulcorner\mathfrak{C}\lrcorner\left(U{\times}V \,\substack{I\\ \tilde U\&\tilde V}\, I\right)$ is the unique bottom choice cell such that

$$\frac{\mathrm{get}_L^{U\times V}}{\ulcorner\sigma_{0,\tilde U,\tilde V}\lrcorner} = \pi_{0,U,V}|\,\mathrm{get}_L^U \quad\text{and}\quad \frac{\mathrm{get}_L^{U\times V}}{\ulcorner\sigma_{1,\tilde U,\tilde V}\lrcorner} = \pi_{1,U,V}|\,\mathrm{get}_L^V\,. \tag{4.2.23}$$

(xi) $\mathrm{put}_R^{U\times V} : \ulcorner\mathfrak{C}\lrcorner\left(I \,\substack{\tilde U\&\tilde V\\ I}\, U{\times}V\right)$ is the unique right choice cell such that

$$\mathrm{put}_R^{U\times V}\,|\pi_{0,U,V} = \frac{\ulcorner\sigma_{0,\tilde U,\tilde V}\lrcorner}{\mathrm{put}_R^U} \quad\text{and}\quad \mathrm{put}_R^{U\times V}\,|\pi_{1,U,V} = \frac{\ulcorner\sigma_{1,\tilde U,\tilde V}\lrcorner}{\mathrm{put}_R^V}. \tag{4.2.24}$$

(xii) $\mathrm{get}_R^{U\times V} : \ulcorner\mathfrak{C}\lrcorner\left(I \,\substack{I\\ \tilde U^\perp\oplus\tilde V^\perp}\, U{\times}V\right)$ is the unique right choice cell such that

$$\mathrm{get}_R^{U\times V}\,|\pi_{0,U,V} = \frac{\mathrm{get}_R^U}{\ulcorner\rho_{0,\tilde U^\perp,\tilde V^\perp}\lrcorner} \quad\text{and}\quad \mathrm{get}_R^{U\times V}\,|\pi_{1,U,V} = \frac{\mathrm{get}_R^V}{\ulcorner\rho_{1,\tilde U^\perp,\tilde V^\perp}\lrcorner}. \tag{4.2.25}$$

(xiii) $\mathrm{put}_L^{U\times V} : \ulcorner\mathfrak{C}\lrcorner(U\times V \;{}^{\tilde U^\perp \oplus \tilde V^\perp}_{I}\; I)$ is the unique top choice cell such that

$$\frac{\ulcorner \rotatebox{180}{$\rho$}_{0,\tilde U^\perp,\tilde V^\perp} \lrcorner}{\mathrm{put}_L^{U\times V}} = \pi_{0,U,V} \,|\, \mathrm{put}_L^U \quad\text{and}\quad \frac{\ulcorner \rotatebox{180}{$\rho$}_{1,\tilde U^\perp,\tilde V^\perp} \lrcorner}{\mathrm{put}_L^{U\times V}} = \pi_{1,U,V} \,|\, \mathrm{put}_L^V . \tag{4.2.26}$$

(xiv) $\mathrm{get}_L^{U^+} : \ulcorner\mathfrak{C}\lrcorner(U^+ \;{}^{I}_{?\tilde U}\; I)$ is the unique left iteration cell such that

$$\rotatebox{180}{$\pi$}_{0,U} \,|\, \mathrm{get}_L^{U^+} = \ulcorner \reflectbox{$e$}_{\tilde U} \lrcorner \quad\text{and}\quad \rotatebox{180}{$\pi$}_{1,U} \,|\, \mathrm{get}_L^{U^+} = \frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^+}} \,|\, \mathrm{get}_L^U}{\ulcorner p_{\tilde U} \lrcorner} . \tag{4.2.27}$$

(xv) $\mathrm{put}_R^{U^+} : \ulcorner\mathfrak{C}\lrcorner(I \;{}^{?\tilde U}_{I}\; U^+)$ is the unique top iteration cell such that

$$\frac{\ulcorner \reflectbox{$e$}_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^+}} = \rotatebox{180}{$\pi$}_{0,U} \quad\text{and}\quad \frac{\ulcorner p_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^+}} = \mathrm{put}_R^{U^+} \,|\, \frac{\mathrm{put}_R^U}{\mathrm{id}_{U^+}} \,|\, \rotatebox{180}{$\pi$}_{1,U} . \tag{4.2.28}$$

(xvi) $\mathrm{get}_R^{U^+} : \ulcorner\mathfrak{C}\lrcorner(I \;{}^{I}_{!(\tilde U^\perp)}\; U^+)$ is the unique reversed bottom iteration cell such that

$$\frac{\mathrm{get}_R^{U^+}}{\ulcorner e_{\tilde U^\perp} \lrcorner} = \rotatebox{180}{$\pi$}_{0,U} \quad\text{and}\quad \frac{\mathrm{get}_R^{U^+}}{\ulcorner d'^{\circ}_{\tilde U^\perp} \lrcorner} = \frac{\mathrm{get}_R^U}{1_{\tilde U^\perp}} \,|\, \frac{\mathrm{id}_U}{\mathrm{get}_R^{U^+}} \,|\, \rotatebox{180}{$\pi$}_{1,U} . \tag{4.2.29}$$

(xvii) $\mathrm{put}_L^{U^+} : \ulcorner\mathfrak{C}\lrcorner(U^+ \;{}^{!(\tilde U^\perp)}_{I}\; I)$ is the unique left iteration cell such that

$$\rotatebox{180}{$\pi$}_{0,U} \,|\, \mathrm{put}_L^{U^+} = \ulcorner e_{\tilde U^\perp} \lrcorner \quad\text{and}\quad \rotatebox{180}{$\pi$}_{1,U} \,|\, \mathrm{get}_L^{U^+} = \frac{\ulcorner d'^{\circ}_{\tilde U^\perp} \lrcorner}{\frac{\mathrm{put}_L^U \,|\, 1_{!(\tilde U^\perp)}}{\mathrm{put}_L^{U^+}}} . \tag{4.2.30}$$

(xviii) $\mathrm{get}_L^{U^\times} : \ulcorner\mathfrak{C}\lrcorner(U^\times \;{}^{I}_{!\tilde U}\; I)$ is the unique bottom iteration cell such that

$$\frac{\mathrm{get}_L^{U^\times}}{\ulcorner e_{\tilde U} \lrcorner} = \pi_{0,U} \quad\text{and}\quad \frac{\mathrm{get}_L^{U^\times}}{\ulcorner d^{\circ}_{\tilde U} \lrcorner} = \pi_{1,U} \,|\, \frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}} \,|\, \frac{\mathrm{get}_L^U}{1_{\tilde U}} . \tag{4.2.31}$$

(xix) $\mathrm{put}_R^{U^\times} : \ulcorner\mathfrak{C}\lrcorner(I \;{}^{!\tilde U}_{I}\; U^\times)$ is the unique right iteration cell such that

$$\mathrm{put}_R^{U^\times} \,|\, \pi_{0,U} = \ulcorner e_{\tilde U} \lrcorner \quad\text{and}\quad \mathrm{put}_R^{U^\times} \,|\, \pi_{1,U} = \frac{\ulcorner d^{\circ}_{\tilde U} \lrcorner}{\frac{1_{!\tilde U} \,|\, \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}} . \tag{4.2.32}$$

(xx) $\mathrm{get}_R^{U^\times} : \ulcorner\mathfrak{C}\lrcorner(I \;{}^{I}_{?(\tilde U^\perp)}\; U^\times)$ is the unique right iteration cell such that

$$\mathrm{get}_R^{U^\times} \,|\, \pi_{0,U} = \ulcorner \reflectbox{$e$}_{\tilde U^\perp} \lrcorner \quad\text{and}\quad \mathrm{get}_R^{U^\times} \,|\, \pi_{1,U} = \frac{\mathrm{get}_R^U \,|\, \frac{\mathrm{id}_U}{\mathrm{get}_R^{U^\times}}}{\ulcorner p'_{\tilde U^\perp} \lrcorner} . \tag{4.2.33}$$

(xxi) $\mathrm{put}_L^{U^\times} : \ulcorner\mathfrak{C}\lrcorner(U^\times \, {}^{?(\tilde{U}^\perp)}_{I} \, I)$ is the unique reversed top iteration cell such that

$$\frac{\ulcorner \partial_{\tilde{U}^\perp} \lrcorner}{\mathrm{put}_L^{U^\times}} = \pi_{0,U} \quad \text{and} \quad \frac{\ulcorner p'_{\tilde{U}^\perp} \lrcorner}{\mathrm{put}_L^{U^\times}} = \pi_{1,U} \Big| \frac{\mathrm{put}_L^U}{\mathrm{id}_{U^\times}} \Big| \mathrm{put}_L^{U^\times} . \tag{4.2.34}$$

♣

Note that there is an issue with our inductive definition of the generalized corners: the $\mathfrak{C}_*^{\triangleright\blacktriangleleft}$-valued exchanges are not freely generated. In particular, we have the fact $U^\times = I \times (U \otimes U^\times)$ and $U^+ = I + (U \otimes U^+)$ for all $\mathfrak{C}$-valued exchanges $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$. Hence, to show that out inductive definition of the generalized corners are well-defined, we need to show that the following eight equalities hold. We will see that the fact that these equalities do hold is an almost immediate consequence of the coherence conditions of a resource category.

**Lemma 4.2.6. The Generalized Corner is Well-Defined.**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U$,

$$\begin{aligned}
&\mathrm{get}_L^{U^\times} = \mathrm{get}_L^{I\times(U\otimes U^\times)}, &\quad &\mathrm{put}_R^{U^\times} = \mathrm{put}_R^{I\times(U\otimes U^\times)}, &\quad &\mathrm{get}_R^{U^\times} = \mathrm{get}_R^{I\times(U\otimes U^\times)},\\
&\mathrm{put}_L^{U^\times} = \mathrm{put}_L^{I\times(U\otimes U^\times)}, & &\mathrm{get}_L^{U^+} = \mathrm{get}_L^{I+(U\otimes U^+)}, & &\mathrm{put}_R^{U^+} = \mathrm{put}_R^{I+(U\otimes U^+)},\\
&\mathrm{get}_R^{U^+} = \mathrm{get}_R^{I+(U\otimes U^+)}, & &\mathrm{put}_L^{U^+} = \mathrm{put}_L^{I+(U\otimes U^+)}, & &
\end{aligned}$$

♡

***Proof***. We prove each once individually by uniqueness of the right or left choice cell.

1. By (4.1.11a), we have that

$$\frac{\mathrm{get}_L^{U^\times}}{\ulcorner \sigma_{0,I,!\tilde{U}\otimes\tilde{U}} \lrcorner} = \frac{\mathrm{get}_L^{U^\times}}{\ulcorner e_{\tilde{U}} \lrcorner} = \pi_{0,U} = \pi_{0,U}|1_I = \pi_{0,I,U\otimes U^\times}|\mathrm{get}_L^I, \tag{4.2.36}$$

$$\frac{\mathrm{get}_L^{U^\times}}{\ulcorner \sigma_{1,I,!\tilde{U}\otimes\tilde{U}} \lrcorner} = \frac{\mathrm{get}_L^{U^\times}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner} = \pi_{1,U} \Big| \frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}} \Big| \frac{\mathrm{get}_L^U}{1_{\tilde{U}}} = \pi_{1,U} \Big| \frac{\mathrm{get}_L^U}{\mathrm{get}_L^{U^\times} \,|1_{\tilde{U}}} = \pi_{1,I,U\otimes U^\times}|\mathrm{get}_L^{U\otimes U^\times}, \tag{4.2.37}$$

so by the uniqueness of the bottom choice cell and (4.2.23), $\mathrm{get}_L^{U^\times} = \mathrm{get}_L^{I\times(U\otimes U^\times)}$.

2. By (4.1.11a), we have that

$$\mathrm{put}_R^{U^\times}|\pi_{0,I,U\otimes U^\times} = \mathrm{put}_R^{U^\times}|\pi_{0,U} = \ulcorner e_{\tilde{U}} \lrcorner = \frac{\ulcorner e_{\tilde{U}} \lrcorner}{1_I} = \frac{\ulcorner \sigma_{0,I,!\tilde{U}\otimes\tilde{U}} \lrcorner}{\mathrm{put}_R^I}, \tag{4.2.38}$$

$$\mathrm{put}_R^{U^\times}|\pi_{1,I,U\otimes U^\times} = \mathrm{put}_R^{U^\times}|\pi_{1,U} = \frac{\ulcorner d^\circ_{\tilde{U}} \lrcorner}{\dfrac{1_{!\tilde{U}}|\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}} = \frac{\ulcorner \sigma_{1,I,!\tilde{U}\otimes\tilde{U}} \lrcorner}{\mathrm{put}_R^{U\otimes U^+}}, \tag{4.2.39}$$

so by the uniqueness of the right choice cell and (4.2.24), $\mathrm{put}_R^{U^\times} = \mathrm{put}_R^{I\times(U\otimes U^\times)}$.

3. By (4.1.12a), we have that

$$\mathrm{get}_R^{U^\times}|\pi_{0,I,U\otimes U^\times} = \mathrm{get}_R^{U^\times}|\pi_{0,U} = \ulcorner \partial_{\tilde{U}^\perp} \lrcorner = \frac{1_I}{\ulcorner \partial_{\tilde{U}^\perp} \lrcorner} = \frac{\mathrm{get}_R^I}{\ulcorner \rho_{0,I,\tilde{U}^\perp\otimes?(\tilde{U}^\perp)} \lrcorner}, \tag{4.2.40}$$

$$\mathrm{get}_R^{U^\times} | \pi_{1,I,U\otimes U^\times} = \mathrm{get}_R^{U^\times} | \pi_{1,U} = \frac{\mathrm{get}_R^U \,|\, \frac{\mathrm{id}_U}{\mathrm{get}_R^{U^\times}}}{\ulcorner p'_{\tilde{U}^\perp} \lrcorner} = \frac{\mathrm{get}_R^{U\otimes U^\times}}{\ulcorner \rho_{0,I,\tilde{U}^\perp \otimes ?(\tilde{U}^\perp)} \lrcorner}, \tag{4.2.41}$$

so by the uniqueness of the right choice cell and (4.2.25), $\mathrm{get}_R^{U^\times} = \mathrm{get}_R^{I\times(U\otimes U^\times)}$.

4. By (4.1.12a), we have that

$$\frac{\ulcorner \rho_{0,I,\tilde{U}^\perp \otimes ?(\tilde{U}^\perp)} \lrcorner}{\mathrm{put}_L^{U^\times}} = \frac{\ulcorner \partial_{\tilde{U}^\perp} \lrcorner}{\mathrm{put}_L^{U^\times}} = \pi_{0,U} = \pi_{0,U} | 1_I = \pi_{0,I,U\otimes U^\times} | \mathrm{put}_L^I, \tag{4.2.42}$$

$$\begin{aligned} \frac{\ulcorner \rho_{1,I,\tilde{U}^\perp \otimes ?(\tilde{U}^\perp)} \lrcorner}{\mathrm{put}_L^{U^\times}} &= \frac{\ulcorner p'_{\tilde{U}^\perp} \lrcorner}{\mathrm{put}_L^{U^\times}} = \pi_{1,U} | \frac{\mathrm{put}_L^U}{\mathrm{id}_{U^\times}} | \mathrm{put}_L^{U^\times} = \pi_{1,U} | \frac{\mathrm{put}_L^U}{\mathrm{id}_{U^\times}} | \frac{1_{?\tilde{U}^\perp}}{\mathrm{put}_L^{U^\times}} \\ &= \pi_{1,U} | \frac{\mathrm{put}_L^U \,| 1_{?\tilde{U}^\perp}}{\mathrm{put}_L^V} = \pi_{1,I,U\otimes U^\times} | \mathrm{put}_L^{U\otimes U^\times}, \end{aligned} \tag{4.2.43}$$

so by the uniqueness of the top choice cell and (4.2.26), $\mathrm{put}_L^{U^\times} = \mathrm{put}_L^{I\times(U\otimes U^\times)}$.

5. By (4.1.12b), we have that

$$\mu_{0,I,U\otimes U^+} | \mathrm{get}_L^{U^+} = \mu_{0,U} | \mathrm{get}_L^{U^+} = \ulcorner \partial_{\tilde{U}} \lrcorner = \frac{1_I}{\ulcorner \partial_{\tilde{U}} \lrcorner} = \frac{\mathrm{get}_L^I}{\ulcorner \rho_{0,\tilde{U},?\tilde{U}\otimes\tilde{U}} \lrcorner}, \tag{4.2.44}$$

$$\begin{aligned} \mu_{1,I,U\otimes U^+} | \mathrm{get}_L^{U^+} &= \mu_{1,U} | \mathrm{get}_L^{U^+} = \frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^+}} | \mathrm{get}_L^U}{\ulcorner p_{\tilde{U}} \lrcorner} = \frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^+}} | \frac{\mathrm{get}_L^U}{1_{\tilde{U}}}}{\ulcorner p_{\tilde{U}} \lrcorner} \\ &= \frac{\frac{\mathrm{get}_L^U}{\mathrm{get}_L^{U^+} | 1_{\tilde{U}}}}{\ulcorner p_{\tilde{U}} \lrcorner} = \frac{\mathrm{get}_L^{U\otimes U^+}}{\ulcorner \rho_{1,\tilde{U},?\tilde{U}\otimes\tilde{U}} \lrcorner}, \end{aligned} \tag{4.2.45}$$

so by the uniqueness of the left choice cell and (4.2.19), $\mathrm{get}_L^{U^+} = \mathrm{get}_L^{I+(U\otimes U^+)}$.

6. By (4.1.12b), we have that

$$\frac{\ulcorner \rho_{0,I,?\tilde{U}\otimes\tilde{U}} \lrcorner}{\mathrm{put}_R^{U^+}} = \frac{\ulcorner \partial_{\tilde{U}} \lrcorner}{\mathrm{put}_R^{U^+}} = \mu_{0,U} = 1_I | \mu_{0,U} = \mathrm{put}_R^I | \mu_{0,I,U\otimes U^+}, \tag{4.2.46}$$

$$\begin{aligned} \frac{\ulcorner \rho_{1,I,?\tilde{U}\otimes\tilde{U}} \lrcorner}{\mathrm{put}_R^{U^+}} &= \frac{\ulcorner p_{\tilde{U}} \lrcorner}{\mathrm{put}_R^{U^+}} = \mathrm{put}_R^{U^+} | \frac{\mathrm{put}_R^U}{\mathrm{id}_{U^+}} | \mu_{1,U} = \frac{1_{?\tilde{U}}}{\mathrm{put}_R^{U^+}} | \frac{\mathrm{put}_R^U}{\mathrm{id}_{U^+}} | \mu_{1,U} \\ &= \frac{1_{?\tilde{U}} | \mathrm{put}_R^U}{\mathrm{put}_R^{U^+}} | \mu_{1,U} = \mathrm{put}_R^{U\otimes U^+} | \mu_{1,I,U\otimes U^+}, \end{aligned} \tag{4.2.47}$$

so by the uniqueness of the top choice cell and (4.2.20), $\mathrm{put}_R^{U^+} = \mathrm{put}_R^{I+(U\otimes U^+)}$.

7. By (4.1.11b), we have that

$$\frac{\mathrm{get}_R^{U^+}}{\ulcorner \sigma_{0,I,\tilde{U}^\perp \otimes !(\tilde{U}^\perp)} \lrcorner} = \frac{\mathrm{get}_R^{U^+}}{\ulcorner e_{\tilde{U}^\perp} \lrcorner} = \mu_{0,U} = 1_I | \mu_{0,U} = \mathrm{get}_R^I | \mu_{0,I,U\otimes U^+}, \tag{4.2.48}$$

$$\frac{\mathrm{get}_R^{U^+}}{\ulcorner \sigma_{1,I,\tilde{U}^\perp\otimes!(\tilde{U}^\perp)} \urcorner} = \frac{\mathrm{get}_R^{U^+}}{\ulcorner d'^{\circ}_{\tilde{U}^\perp} \urcorner} = \frac{\mathrm{get}_R^{U}}{1_{\tilde{U}^\perp}} \Big| \frac{\mathrm{id}_U}{\mathrm{get}_R^{U^+}} \Big| \nu_{1,U} = \frac{\mathrm{get}_R^{U}}{1_{\tilde{U}^\perp} \,|\, \mathrm{get}_R^{U^+}} \Big| \nu_{1,U}$$
$$= \mathrm{get}_R^{U\otimes U^+} \,|\, \nu_{1,I,U\otimes U^+}, \tag{4.2.49}$$

so by the uniqueness of the bottom choice cell and (4.2.21), $\mathrm{get}_R^{U^+} = \mathrm{get}_R^{I+(U\otimes U^+)}$.

8. By (4.1.11b), we have that

$$\nu_{0,I,U\otimes U^+} \,|\, \mathrm{put}_L^{U^+} = \nu_{0,U} \,|\, \mathrm{put}_L^{U^+} = \ulcorner e_{\tilde{U}^\perp} \urcorner = \frac{\ulcorner e_{\tilde{U}^\perp} \urcorner}{1_I} = \frac{\ulcorner \sigma_{0,I,\tilde{U}^\perp\otimes!(\tilde{U}^\perp)} \urcorner}{\mathrm{put}_L^{I}} \tag{4.2.50}$$

$$\nu_{1,I,U\otimes U^+} \,|\, \mathrm{put}_L^{U^+} = \nu_{1,U} \,|\, \mathrm{get}_L^{U^+} = \frac{\ulcorner d'^{\circ}_{\tilde{U}^\perp} \urcorner}{\frac{\mathrm{put}_L^{U} \,|\, 1_{!(\tilde{U}^\perp)}}{\mathrm{put}_L^{U^+}}} = \frac{\ulcorner \sigma_{1,I,\tilde{U}^\perp\otimes!(\tilde{U}^\perp)} \urcorner}{\mathrm{put}_L^{U\otimes U^+}} \tag{4.2.51}$$

so by the uniqueness of the left choice cell and (4.2.22), $\mathrm{put}_L^{U^+} = \mathrm{put}_L^{I+(U\otimes U^+)}$.

This completes the proof. ❏

With the definition of the generalized corners at hand, the first property we prove is that the yanking equations stated in Definition 4.2.3 generalize. The intuition, here, is that getting a message from the left and then putting the message to the right should be the same as leaving the message intact, and putting a resource to the right and then getting the resource from the left should be the same as leaving the resource intact.

**Theorem 4.2.7. Generalized Yanking Equations**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $W \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} = \mathrm{id}_U, \tag{4.2.52a}$$

$$\mathrm{put}_R^W \,|\, \mathrm{get}_L^W = 1_{\tilde{U}}, \tag{4.2.52b}$$

$$\mathrm{put}_L^W \,|\, \mathrm{get}_R^W = 1_{\tilde{U}^\perp}, \tag{4.2.52c}$$

$$\frac{\mathrm{get}_R^W}{\mathrm{put}_L^W} = \mathrm{id}_U. \tag{4.2.52d}$$

♡

***Proof***. We prove (4.2.52a) by induction on $U$.

1. Assume that $W = I$. Then,

$$\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} = \frac{\mathrm{get}_L^I}{\mathrm{put}_R^I} = \frac{1_I}{1_I} = 1_I = \mathrm{id}_I = \mathrm{id}_W. \tag{4.2.53}$$

2. Assume that $W = A^{\triangleright}$, where $A \in \mathrm{Ob}(\mathfrak{C})$ is an object in $\mathfrak{C}$. Then,

$$\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} = \frac{\mathrm{get}_L^{A^\triangleright}}{\mathrm{put}_R^{A^\triangleright}} = \frac{\mathrm{get}_L^A}{\mathrm{put}_R^A} = \mathrm{id}_{A^\triangleright} = \mathrm{id}_W, \tag{4.2.54}$$

where we used (4.2.3) for the last equality.

3. Assume that $W = A^{\blacktriangleleft}$, where $A \in \mathrm{Ob}(\mathfrak{C})$ is an object in $\mathfrak{C}$. Then,

$$\begin{aligned}
\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} &= \frac{\mathrm{get}_L^{A^{\blacktriangleleft}}}{\mathrm{put}_R^{A^{\blacktriangleleft}}} = \frac{\frac{\eta'_A}{\mathrm{put}_L^A|1_{A^\perp}}}{\frac{1_{A^\perp}|\mathrm{get}_R^A}{\epsilon'_A}} = \frac{\frac{1_I|1_I}{\frac{\eta'_A|1_I}{\mathrm{put}_L^A|1_{A^\perp}|1_I}}}{\frac{1_I|1_{A^\perp}|\mathrm{get}_R^A}{1_I|\epsilon'_A}} = \frac{\frac{\frac{1_I}{\eta'_A}}{\frac{\mathrm{put}_L^A}{1_I}\Big|\frac{1_{A^\perp}}{1_{A^\perp}}}\Bigg|\frac{\frac{1_I}{1_I}}{\frac{1_I}{\mathrm{get}_R^A}}}{1_I|\epsilon'_A} = \frac{\frac{\frac{1_I}{\eta'_A}}{\mathrm{put}_L^A|1_{A^\perp}}\Bigg|\frac{\frac{\mathrm{get}_R^A}{1_A}}{1_A}}{1_I|\epsilon'_A} \\
&= \frac{\frac{1_I|\mathrm{get}_R^A}{\eta'_A|1_A}}{\frac{\mathrm{put}_L^A|1_{A^\perp}|1_A}{1_I|\epsilon'_A}} = \frac{\frac{\mathrm{get}_R^A}{\eta'_A|1_A}}{\frac{\frac{\mathrm{put}_L^A|1_{A^\perp}|1_A}{1_I|\epsilon'_A}}{1_I|1_I}} = \frac{\frac{\mathrm{get}_R^A}{\eta'_A|1_A}}{\frac{\frac{\mathrm{put}_L^A}{1_I}}{1_I}\Bigg|\frac{\frac{1_{A^\perp}|1_A}{\epsilon'_A}}{1_I}} = \frac{\frac{\mathrm{get}_R^A}{\eta'_A|1_A}}{\frac{1_A}{\mathrm{put}_L^A}\Big|\frac{\epsilon'_A}{1_I}} = \frac{\frac{\mathrm{get}_R^A}{\eta'_A|1_A}}{\frac{1_A|\epsilon'_A}{\mathrm{put}_L^A|1_I}} \\
&= \frac{\frac{\mathrm{get}_R^A}{\eta'_A|1_A}}{\frac{1_A|\epsilon'_A}{\mathrm{put}_L^A}} = \frac{\frac{\mathrm{get}_R^A}{\ulcorner(1_A\otimes\epsilon'_A)\circ(\eta'_A\otimes 1_A)\lrcorner}}{\mathrm{put}_L^A} = \frac{\frac{\mathrm{get}_R^A}{\ulcorner 1_A \lrcorner}}{\mathrm{put}_L^A} = \frac{\frac{\mathrm{get}_R^A}{1_A}}{\mathrm{put}_L^A} = \frac{\mathrm{get}_R^A}{\mathrm{put}_L^A} = \mathrm{id}_{A^{\blacktriangleleft}} = \mathrm{id}_W,
\end{aligned} \tag{4.2.55}$$

where we used (4.2.3) for the last equality.

4. Assume that $W = U \otimes V$, where $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ are $\mathfrak{C}$-valued exchanges. Then,

$$\begin{aligned}
\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} &= \frac{\mathrm{get}_L^{U\otimes V}}{\mathrm{put}_R^{U\otimes V}} = \frac{\frac{\mathrm{get}_L^U}{\mathrm{get}_L^V|1_{\tilde U}}}{\frac{1_{\tilde V}|\mathrm{put}_R^U}{\mathrm{put}_R^V}} = \frac{\frac{\mathrm{id}_U|\mathrm{get}_L^U}{\mathrm{get}_L^V|1_{\tilde U}}}{\frac{1_{\tilde V}|\mathrm{put}_R^U}{\mathrm{put}_R^V|\mathrm{id}_V}} = \frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^V}}{\frac{1_{\tilde V}}{\mathrm{put}_R^V}}\Bigg|\frac{\frac{\mathrm{get}_L^U}{1_{\tilde U}}}{\frac{\mathrm{put}_R^U}{\mathrm{id}_V}} = \frac{\frac{\mathrm{id}_U}{1_I}}{\frac{\mathrm{get}_L^V}{\mathrm{put}_R^V}}\Bigg|\frac{\frac{\mathrm{get}_L^U}{\mathrm{put}_R^U}}{\frac{1_I}{\mathrm{id}_V}} \\
&= \frac{\mathrm{id}_U}{\mathrm{id}_V}\Big|\frac{\mathrm{id}_U}{\mathrm{id}_V} = \frac{\mathrm{id}_U|\mathrm{id}_U}{\mathrm{id}_V|\mathrm{id}_V} = \frac{\mathrm{id}_U}{\mathrm{id}_V} = \mathrm{id}_{U\otimes V},
\end{aligned} \tag{4.2.56}$$

where we used the induction hypothesis on $U$ and $V$ for the sixth equality.

5. Assume that $W = U + V$, where $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ are $\mathfrak{C}$-valued exchanges. Then,

$$\begin{aligned}
\text{ɥ}_{0,U,V}\Big|\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} &= \frac{\text{ɥ}_{0,U,V}}{1_I}\Big|\frac{\mathrm{get}_L^{U+V}}{\mathrm{put}_R^{U+V}} = \frac{\text{ɥ}_{0,U,V}|\mathrm{get}_L^{U+V}}{1_I|\mathrm{put}_R^{U+V}} = \frac{\frac{\mathrm{get}_L^U}{\ulcorner\varrho_{0,\tilde U,\tilde V}\lrcorner}}{\mathrm{put}_R^{U+V}} = \frac{\mathrm{get}_L^U}{\mathrm{put}_R^U|\text{ɥ}_{0,U,V}} \\
&= \frac{\mathrm{get}_L^U|1_I}{\mathrm{put}_R^U|\text{ɥ}_{0,U,V}} = \frac{\mathrm{get}_L^U}{\mathrm{put}_R^U}\Big|\frac{1_I}{\text{ɥ}_{0,U,V}} = \mathrm{id}_U|\text{ɥ}_{0,U,V} = \text{ɥ}_{0,U,V},
\end{aligned} \tag{4.2.57}$$

where we used (4.2.19) for the third equality, (4.2.20) for the fourth equality, and the induction hypothesis on $U$ for the sixth equality, and

$$\begin{aligned}
\text{ɥ}_{1,U,V}\Big|\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} &= \frac{\text{ɥ}_{1,U,V}}{1_I}\Big|\frac{\mathrm{get}_L^{U+V}}{\mathrm{put}_R^{U+V}} = \frac{\text{ɥ}_{1,U,V}|\mathrm{get}_L^{U+V}}{1_I|\mathrm{put}_R^{U+V}} = \frac{\frac{\mathrm{get}_L^V}{\ulcorner\varrho_{1,\tilde U,\tilde V}\lrcorner}}{\mathrm{put}_R^{U+V}} = \frac{\mathrm{get}_L^V}{\mathrm{put}_R^V|\text{ɥ}_{1,U,V}} \\
&= \frac{\mathrm{get}_L^V|1_I}{\mathrm{put}_R^V|\text{ɥ}_{1,U,V}} = \frac{\mathrm{get}_L^V}{\mathrm{put}_R^V}\Big|\frac{1_I}{\text{ɥ}_{1,U,V}} = \mathrm{id}_V|\text{ɥ}_{1,U,V} = \text{ɥ}_{1,U,V},
\end{aligned} \tag{4.2.58}$$

where we used (4.2.19) for the third equality, (4.2.20) for the fourth equality, and the induction hypothesis on $V$ for the sixth equality. Also,

$$\text{ɥ}_{0,U,V}|\mathrm{id}_W = \text{ɥ}_{0,U,V}|\mathrm{id}_{U+V} = \text{ɥ}_{0,U,V} \ \text{ and } \ \text{ɥ}_{1,U,V}|\mathrm{id}_W = \text{ɥ}_{1,U,V}|\mathrm{id}_{U+V} = \text{ɥ}_{1,U,V}\,. \tag{4.2.59}$$

Hence, by the uniqueness of the left choice cell,

$$\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} = \text{ɥ}_{0,U,V} \boxplus \text{ɥ}_{1,U,V} = \mathrm{id}_W. \tag{4.2.60}$$

6. Assume that $W = U \times V$, where $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ are $\mathfrak{C}$-valued exchanges. Then,

$$\begin{aligned}
\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W}\Big|\pi_{0,U,V} &= \frac{\mathrm{get}_L^{U\times V}}{\mathrm{put}_R^{U\times V}}\Big|\frac{1_I}{\pi_{0,U,V}} = \frac{\mathrm{get}_L^{U\times V}\,|1_I}{\mathrm{put}_R^{U\times V}\,|\pi_{0,U,V}} = \frac{\mathrm{get}_L^{U\times V}}{\frac{\ulcorner\sigma_{0,\tilde{U},\tilde{V}}\lrcorner}{\mathrm{put}_R^U}} = \frac{\pi_{0,U,V}|\,\mathrm{get}_L^U}{\mathrm{put}_R^U} \\
&= \frac{\pi_{0,U,V}|\,\mathrm{get}_L^U}{1_I|\,\mathrm{put}_R^U} = \frac{\pi_{0,U,V}}{1_I}\Big|\frac{\mathrm{get}_L^U}{\mathrm{put}_R^U} = \pi_{0,U,V}|\mathrm{id}_U = \pi_{0,U,V},
\end{aligned} \tag{4.2.61}$$

where we used (4.2.24) for the third equality, (4.2.23) for the fourth equality, and the induction hypothesis on $U$ for the sixth equality, and

$$\begin{aligned}
\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W}\Big|\pi_{1,U,V} &= \frac{\mathrm{get}_L^{U\times V}}{\mathrm{put}_R^{U\times V}}\Big|\frac{1_I}{\pi_{1,U,V}} = \frac{\mathrm{get}_L^{U\times V}\,|1_I}{\mathrm{put}_R^{U\times V}\,|\pi_{1,U,V}} = \frac{\mathrm{get}_L^{U\times V}}{\frac{\ulcorner\sigma_{1,\tilde{U},\tilde{V}}\lrcorner}{\mathrm{put}_R^V}} = \frac{\pi_{1,U,V}|\,\mathrm{get}_L^V}{\mathrm{put}_R^V} \\
&= \frac{\pi_{1,U,V}|\,\mathrm{get}_L^V}{1_I|\,\mathrm{put}_R^V} = \frac{\pi_{1,U,V}}{1_I}\Big|\frac{\mathrm{get}_L^V}{\mathrm{put}_R^V} = \pi_{1,U,V}|\mathrm{id}_V = \pi_{1,U,V},
\end{aligned} \tag{4.2.62}$$

where we used (4.2.24) for the third equality, (4.2.23) for the fourth equality, and the induction hypothesis on $V$ for the sixth equality. Also,

$$\mathrm{id}_W|\pi_{0,U,V} = \mathrm{id}_{U\times V}|\pi_{0,U,V} = \pi_{0,U,V} \text{ and } \mathrm{id}_W|\pi_{1,U,V} = \mathrm{id}_{U\times V}|\pi_{1,U,V} = \pi_{1,U,V}. \tag{4.2.63}$$

Hence, by the uniqueness of the left choice cell,

$$\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} = \pi_{0,U,V} \boxtimes \pi_{1,U,V} = \mathrm{id}_W. \tag{4.2.64}$$

7. Assume that $W = U^+$, where $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ is a $\mathfrak{C}$-valued exchange. Then,

$$\iota_{0,U}\,\Big|\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} = \frac{\iota_{0,U}}{1_I}\Big|\frac{\mathrm{get}_L^{U^+}}{\mathrm{put}_R^{U^+}} = \frac{\iota_{0,U}\,|\,\mathrm{get}_L^{U^+}}{1_I|\,\mathrm{put}_R^{U^+}} = \frac{\ulcorner\partial_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^+}} = \iota_{0,U}, \tag{4.2.65}$$

where we used (4.2.27) for the third equality and (4.2.28) for the fourth equality, and

$$\begin{aligned}
\iota_{1,U}\,\Big|\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} &= \frac{\iota_{1,U}}{1_I}\Big|\frac{\mathrm{get}_L^{U^+}}{\mathrm{put}_R^{U^+}} = \frac{\iota_{1,U}\,|\,\mathrm{get}_L^{U^+}}{1_I|\,\mathrm{put}_R^{U^+}} = \frac{\frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^+}}|\,\mathrm{get}_L^U}{\ulcorner p_{\tilde{U}}\lrcorner}}{\mathrm{put}_R^{U^+}} = \frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^+}}|\,\mathrm{get}_L^U}{\mathrm{put}_R^{U^+}\,\Big|\frac{\mathrm{put}_R^U}{\mathrm{id}_{U^+}}\Big|\,\iota_{1,U}} \\
&= \frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^+}}|\,\mathrm{get}_L^U\,|1_I}{\mathrm{put}_R^{U^+}\,\Big|\frac{\mathrm{put}_R^U}{\mathrm{id}_{U^+}}\Big|\,\iota_{1,U}} = \frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^+}}{\mathrm{put}_R^{U^+}}}\Bigg|\frac{\frac{\mathrm{get}_L^U}{\mathrm{put}_R^U}}{1_{U^+}}\Bigg|\frac{1_I}{\pi_{1,0}} = \frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^+}}{\mathrm{put}_R^{U^+}}}\Bigg|\frac{\mathrm{id}_U}{1_{U^+}}\Bigg|\frac{1_I}{\pi_{1,0}} \\
&= \frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^+}}{\mathrm{put}_R^{U^+}}}\Bigg|\pi_{1,0},
\end{aligned} \tag{4.2.66}$$

where we used (4.2.27) for the third equality, (4.2.28) for the fourth equality, and the induction hypothesis on $V$ for the sixth equality. Also,

$$\iota_{0,U}\,|\mathrm{id}_W = \iota_{0,U}\,|\mathrm{id}_{U^+} = \iota_{0,U} \text{ and } \iota_{1,U}\,|\mathrm{id}_W = \iota_{1,U}\,|\mathrm{id}_{U^+} = \frac{\mathrm{id}_U}{\mathrm{id}_{U^+}}\Big|\,\iota_{1,U}\,. \tag{4.2.67}$$

Hence, by the uniqueness of the left iteration cell,

$$\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} = (\mathrm{id}_U)^+_{\amalg_{0,U},\amalg_{1,U},1_I} = \mathrm{id}_{U^+}. \tag{4.2.68}$$

8. Assume that $W = U^\times$, where $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ is a $\mathfrak{C}$-valued exchange. Then,

$$\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W}|\pi_{0,U} = \frac{\mathrm{get}_L^{U^\times}}{\mathrm{put}_R^{U^\times}}|\frac{1_I}{\pi_{0,U}} = \frac{\mathrm{get}_L^{U^\times}\ |1_I}{\mathrm{put}_R^{U^\times}\ |\pi_{0,U}} = \frac{\mathrm{get}_L^{U^\times}}{\ulcorner e_{\tilde U} \lrcorner} = \pi_{0,U}, \tag{4.2.69}$$

where we used (4.2.32) for the third equality and (4.2.31) for the fourth equality, and

$$\begin{aligned}
\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W}|\pi_{1,U} &= \frac{\mathrm{get}_L^{U^\times}}{\mathrm{put}_R^{U^\times}}|\frac{1_I}{\pi_{1,U}} = \frac{\mathrm{get}_L^{U^\times}\ |1_I}{\mathrm{put}_R^{U^\times}\ |\pi_{1,U}} = \frac{\mathrm{get}_L^{U^\times}}{\frac{\ulcorner d^\circ_{\tilde U} \lrcorner}{\frac{1_{!\tilde U}|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}} = \frac{\pi_{1,U}|\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}|\frac{\mathrm{get}_L^U}{1_{\tilde U}}}{\frac{1_{!\tilde U}|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}, \\
&= \frac{\pi_{1,U}|\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}|\frac{\mathrm{get}_L^U}{1_{\tilde U}}}{\frac{1_I|1_{!\tilde U}|\,\mathrm{put}_R^U}{1_I|\,\mathrm{put}_R^{U^\times}\ |\mathrm{id}_{U^\times}}} = \frac{\pi_{1,U}}{\frac{1_I}{1_I}}|\frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\frac{1_{!\tilde U}}{\mathrm{put}_R^{U^\times}}}|\frac{\frac{\mathrm{get}_L^U}{1_{\tilde U}}}{\frac{\mathrm{put}_R^U}{\mathrm{id}_{U^\times}}} = \pi_{1,U}|\frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{put}_R^{U^\times}}}|\frac{\frac{\mathrm{get}_L^U}{\mathrm{put}_R^U}}{\mathrm{id}_{U^\times}} \\
&= \pi_{1,U}|\frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{put}_R^{U^\times}}}|\frac{\mathrm{id}_U}{\mathrm{id}_{U^\times}} = \pi_{1,U}|\frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{put}_R^{U^\times}}},
\end{aligned} \tag{4.2.70}$$

where we used (4.2.32) for the third equality, (4.2.31) for the fourth equality, and the induction hypothesis on $V$ for the eighth equality. Also,

$$\mathrm{id}_W|\pi_{0,U} = \mathrm{id}_{U^\times}|\pi_{0,U} = \pi_{0,U} \text{ and } \mathrm{id}_W|\pi_{1,U} = \mathrm{id}_{U^\times}|\pi_{1,U} = \pi_{1,U}|\frac{\mathrm{id}_U}{\mathrm{id}_{U^\times}}. \tag{4.2.71}$$

Hence, by the uniqueness of the left iteration cell,

$$\frac{\mathrm{get}_L^W}{\mathrm{put}_R^W} = (\mathrm{id}_U)^\times_{\pi_{0,U},\pi_{1,U},1_I} = \mathrm{id}_{U^\times}. \tag{4.2.72}$$

This completes the induction. The other three follow similarly. ❏

An immediate consequence of the yanking equation is that a message is equivalent to its forward-time underlying resource sent left.

**Lemma 4.2.8. A Message is Equiv. to Sending it's Underlying Resource Left**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$U \cong (\tilde U)^\triangleright \tag{4.2.73}$$

in the category of resource transducers $\mathbf{H}\ulcorner\mathfrak{C}\lrcorner$. ♡

***Proof.*** Define the cells $\alpha\colon \ulcorner\mathfrak{C}\lrcorner(U\,{}^{I}_{I}\,(\tilde{U})^{\triangleright})$ and $\beta\colon \ulcorner\mathfrak{C}\lrcorner((\tilde{U})^{\triangleright}\,{}^{I}_{I}\,U)$

$$\alpha = \frac{\mathrm{get}^{U}_{L}}{\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}} \quad \text{and} \quad \beta = \frac{\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}}{\mathrm{put}^{U}_{R}}. \tag{4.2.74}$$

Then,

$$\alpha|\beta = \frac{\mathrm{get}^{U}_{L}}{\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}}\Big|\frac{\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}}{\mathrm{put}^{U}_{R}} = \frac{\mathrm{get}^{U}_{L}}{\frac{\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}}{1_I}}\Bigg|\frac{1_I}{\frac{\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}}{\mathrm{put}^{U}_{R}}} = \frac{\mathrm{get}^{U}_{L}\,|1_I}{\frac{\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}\,|\,\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}}{1_I|\,\mathrm{put}^{U}_{R}}} = \frac{\mathrm{get}^{U}_{L}}{\frac{1_{\tilde{U}}}{\mathrm{put}^{U}_{R}}} = \frac{\mathrm{get}^{U}_{L}}{\mathrm{put}^{U}_{R}} = \mathrm{id}_U. \tag{4.2.75a}$$

$$\beta|\alpha = \frac{\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}}{\mathrm{put}^{U}_{R}}\Big|\frac{\mathrm{get}^{U}_{L}}{\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}} = \frac{\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}}{\frac{\mathrm{put}^{U}_{R}}{1_I}}\Bigg|\frac{1_I}{\frac{\mathrm{get}^{U}_{L}}{\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}}} = \frac{\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}\,|1_I}{\frac{\mathrm{put}^{U}_{R}\,|\,\mathrm{get}^{U}_{L}}{1_I|\,\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}}} = \frac{\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}}{\frac{1_{\tilde{U}}}{\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}}} = \frac{\mathrm{get}^{(\tilde{U})^{\triangleright}}_{L}}{\mathrm{put}^{(\tilde{U})^{\triangleright}}_{R}} = \mathrm{id}_{(\tilde{U})^{\triangleright}}. \tag{4.2.75b}$$

Hence, $\alpha$ and $\beta$ are inverses of each other in $\mathbf{H}\ulcorner\mathfrak{C}\lrcorner$, implying that $U \cong (\tilde{U})^{\triangleright}$. ❏

Using the generalized corner, we can define a canonical functor between a resource category and it's category of resource transducers called the cornering functor. Given a process $f\colon A \to B$ in the resource category, its image under the cornering functor is the resource transducer that gets a message $A$ from the left performs the process $f$ on it and sends the output $B$ to the right.

**Definition 4.2.9. The Cornering Functor**

Let $\mathfrak{C}$ be a resource theory. The *cornering functor* is the additive strict anti-symmetric monoidal functor $(-)^{\triangleright}\colon \mathfrak{C} \to \mathbf{H}\ulcorner\mathfrak{C}\lrcorner$ such that

(i) Given an object $A$ in $\mathfrak{C}$, $(A)^{\triangleright} = A^{\triangleright}$, and

(ii) given a morphism $f\colon A \to B$ in $\mathfrak{C}$,

$$(f)^{\triangleright} = \frac{\mathrm{get}^{A^{\triangleright}}_{L}}{\frac{\ulcorner f\lrcorner}{\mathrm{put}^{B^{\triangleright}}_{R}}}. \tag{4.2.76}$$

♣

Now, we turn to constructing a canonical compact closed categorical structure on the category of resource transducers. To do this, we will first need to show that every message $U$ has a left dual. We will find that the left dual of a message is itself but sent in the opposite direction, and by that, we mean the reversed-time message, which is defined below.

**Definition 4.2.10. Reversed-Time Message**

Let $\mathfrak{C}$ be a resource theory. The function $\widetilde{(-)}^{\#}\colon \mathfrak{C}^{\triangleright\blacktriangleleft}_{*} \to \mathfrak{C}^{\triangleright\blacktriangleleft}_{*}$ from $\mathfrak{C}$-valued exchanges $\mathfrak{C}^{\triangleright\blacktriangleleft}_{*}$ to $\mathfrak{C}$-valued exchanges $\mathfrak{C}^{\triangleright\blacktriangleleft}_{*}$ is defined inductively as follows:

$$(I)^{\#} = I, \tag{4.2.77a}$$

$$(A^{\triangleright})^{\#} = A^{\blacktriangleleft}, \tag{4.2.77b}$$

$$(A^{\blacktriangleleft})^{\#} = A^{\triangleright}, \tag{4.2.77c}$$

$$(U \otimes V)^{\#} = V^{\#} \otimes U^{\#}, \tag{4.2.77d}$$

$$(U + V)^{\#} = U^{\#} \times V^{\#}, \tag{4.2.77e}$$

$$(U \times V)^{\#} = U^{\#} + V^{\#}, \tag{4.2.77f}$$
$$(U^{+})^{\#} = (U^{\#})^{\times}, \tag{4.2.77g}$$
$$(U^{\times})^{\#} = (U^{\#})^{+}, \tag{4.2.77h}$$

where $A \in \mathrm{Ob}\,(\mathfrak{C})$ is an object in $\mathfrak{C}$ and $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ are $\mathfrak{C}$-valued exchanges. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$, we refer to $U^{\#}$ as $U$ *in reverse.* ♣

As we would expect, the forward-time underlying resource of the reversed-time message is the dual of the forward-time underlying resource of the base message.

**Lemma 4.2.11**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $W \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\widetilde{W\#} = (\tilde{W})^{\perp}. \tag{4.2.78}$$ ♡

***Proof.*** We prove this via induction on $W$.

1. Assume that $W = I$. Then, $\widetilde{I\#} = \tilde{I} = I = I^{\perp} = (\tilde{I})^{\perp} = (\tilde{W})^{\perp}$.

2. Assume that $W = A^{\triangleright}$, where $A \in \mathrm{Ob}\,(\mathfrak{C})$ is an object in $\mathfrak{C}$. Then, $\widetilde{W\#} = \widetilde{(A^{\triangleright})\#} = \widetilde{A^{\blacktriangleleft}} = A^{\perp} = (\widetilde{A^{\triangleright}})^{\perp} = (\tilde{W})^{\perp}$.

3. Assume that $W = A^{\blacktriangleleft}$, where $A \in \mathrm{Ob}\,(\mathfrak{C})$ is an object in $\mathfrak{C}$. Then, $\widetilde{W\#} = \widetilde{(A^{\blacktriangleleft})\#} = \widetilde{A^{\triangleright}} = A = (A^{\perp})^{\perp} = (\widetilde{A^{\blacktriangleleft}})^{\perp} = (\tilde{W})^{\perp}$.

4. Assume that $W = U \otimes V$, where $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ are $\mathfrak{C}$-valued exchanges. By the induction hypothesis, $\widetilde{U\#} = (\tilde{U})^{\perp}$ and $\widetilde{V\#} = (\tilde{V})^{\perp}$, so $\widetilde{W\#} = \widetilde{(U \otimes V)\#} = \widetilde{V\# \otimes U\#} = \widetilde{U\#} \otimes \widetilde{V\#} = (\tilde{U})^{\perp} \otimes (\tilde{V})^{\perp} = (\tilde{V} \otimes \tilde{U})^{\perp} = (\widetilde{U \otimes V})^{\perp} = (\tilde{W})^{\perp}$.

5. Assume that $W = U + V$, where $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ are $\mathfrak{C}$-valued exchanges. By the induction hypothesis, $\widetilde{U\#} = (\tilde{U})^{\perp}$ and $\widetilde{V\#} = (\tilde{V})^{\perp}$, so $\widetilde{W\#} = \widetilde{(U + V)\#} = \widetilde{U\# \times V\#} = \widetilde{U\#} \times \widetilde{V\#} = (\tilde{U})^{\perp} \,\&\, (\tilde{V})^{\perp} = (\tilde{U} \oplus \tilde{V})^{\perp} = (\widetilde{U + V})^{\perp} = (\tilde{W})^{\perp}$.

6. Assume that $W = U \times V$, where $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ are $\mathfrak{C}$-valued exchanges. By the induction hypothesis, $\widetilde{U\#} = (\tilde{U})^{\perp}$ and $\widetilde{V\#} = (\tilde{V})^{\perp}$, so $\widetilde{W\#} = \widetilde{(U \times V)\#} = \widetilde{U\# + V\#} = \widetilde{U\#} + \widetilde{V\#} = (\tilde{U})^{\perp} \oplus (\tilde{V})^{\perp} = (\tilde{U} \,\&\, \tilde{V})^{\perp} = (\widetilde{U \times V})^{\perp} = (\tilde{W})^{\perp}$.

7. Assume that $W = U^{+}$, where $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ is a $\mathfrak{C}$-valued exchange. By the induction hypothesis, $\widetilde{U\#} = (\tilde{U})^{\perp}$, so $\widetilde{W\#} = \widetilde{(U^{+})\#} = \widetilde{(U\#)^{\times}} = !\widetilde{U\#} = !(\tilde{U}^{\perp}) = (?\tilde{U})^{\perp} = (\widetilde{U^{+}})^{\perp} = (\tilde{W})^{\perp}$.

8. Assume that $W = U^{\times}$, where $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$ is a $\mathfrak{C}$-valued exchange. By the induction hypothesis, $\widetilde{U\#} = (\tilde{U})^{\perp}$, so $\widetilde{W\#} = \widetilde{(U^{\times})\#} = \widetilde{(U\#)^{+}} = ?\widetilde{U\#} = ?(\tilde{U}^{\perp}) = (!\tilde{U})^{\perp} = (\widetilde{U^{\times}})^{\perp} = (\tilde{W})^{\perp}$.

This completes the induction. ❑

To show that the left dual of a message is its reversed-time message, we need to define cup and cap cells of the correct type. We define the cap cell to be the cell that gets the reversed-time message from the right and puts the forward-time message to the right, and we define the cup cell to be the cell that gets the forward-time message from the left and puts the reversed-time message to the left.

**Definition 4.2.12. Cap and Cup Cell**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$, we define the *cap cell* $\eta_U \colon \ulcorner\mathfrak{C}\urcorner(I\,{}^{I}_{I}\,U^{\#}{\otimes}U)$ by

$$\eta_U = \frac{\mathrm{get}_R^{U^\#}}{\mathrm{put}_R^{U}}, \tag{4.2.79}$$

and we define the *cup cell* $\epsilon_U \colon \ulcorner\mathfrak{C}\urcorner(U{\otimes}U^{\#}\,{}^{I}_{I}\,I)$ by

$$\epsilon_U = \frac{\mathrm{get}_L^{U}}{\mathrm{put}_L^{U^\#}}. \tag{4.2.80}$$

The fact that the get and put cells actually compose in the definition of the cap and cup cells is an immediate consequence of Lemma 4.2.11. ♣

By the generalized yanking equations, we get the desired duality.

**Lemma 4.2.13. Every Message is Dual to it's Reverse Time Message**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$(\eta_U, \epsilon_U) \colon U \vdash U^\# \tag{4.2.81}$$

in $\mathbf{H}\ulcorner\mathfrak{C}\urcorner$. ♡

***Proof.*** As we can see,

$$\frac{\eta_U}{\mathrm{id}_{U^\#}} \Big| \frac{\mathrm{id}_{U^\#}}{\epsilon_U} = \frac{\frac{\mathrm{get}_R^{U^\#}}{\mathrm{put}_R^{U}}}{\mathrm{id}_{U^\#}} \Big| \frac{\frac{\mathrm{id}_{U^\#}}{\mathrm{get}_L^{U}}}{\mathrm{put}_L^{U^\#}} = \frac{\frac{\mathrm{get}_R^{U^\#} \mid \mathrm{id}_{U^\#}}{\mathrm{put}_R^{U} \mid \mathrm{get}_L^{U}}}{\mathrm{id}_{U^\#} \mid \mathrm{put}_L^{U^\#}} = \frac{\frac{\mathrm{get}_R^{U^\#}}{1_{\tilde U}}}{\mathrm{put}_L^{U^\#}} = \frac{\mathrm{get}_R^{U^\#}}{\mathrm{put}_L^{U^\#}} = \mathrm{id}_{U^\#}, \tag{4.2.82}$$

and

$$\frac{\mathrm{id}_U}{\eta_U} \Big| \frac{\epsilon_U}{\mathrm{id}_U} = \frac{\frac{\mathrm{id}_U}{\mathrm{get}_R^{U^\#}}}{\mathrm{put}_R^{U}} \Big| \frac{\frac{\mathrm{get}_L^{U}}{\mathrm{put}_L^{U^\#}}}{\mathrm{id}_U} = \frac{\frac{\mathrm{id}_U \mid \mathrm{get}_L^{U}}{\mathrm{get}_R^{U^\#} \mid \mathrm{put}_L^{U^\#}}}{\mathrm{put}_R^{U} \mid \mathrm{id}_U} = \frac{\frac{\mathrm{get}_L^{U}}{1_{\tilde U^\perp}}}{\mathrm{put}_R^{U}} = \frac{\mathrm{get}_L^{U}}{\mathrm{put}_R^{U}} = \mathrm{id}_U. \tag{4.2.83}$$

Since the snake-equations are satisfied, we are done. ❏

This tells us that every message has a dual in the category of resource transducers, but we cannot yet say that the category of resource transducers is compact closed. We still need to show that the category of resource transducers is symmetric. But before we do that, we have all the necessary data to talk about the transpose of some resource transducers.

**Lemma 4.2.14. The Transpose of the Projection Cell is the Injection Cell**

Let $\mathfrak{C}$ be a resource theory. Given $\mathfrak{C}$-valued exchanges $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\pi^{\perp}_{0,U,V} = \iota_{0,U^\#,V^\#}, \tag{4.2.84a}$$

$$\pi^{\perp}_{1,U,V} = \iota_{1,U^\#,V^\#}. \tag{4.2.84b}$$

♡

***Proof.*** We first prove (4.2.84a). As we can see,

$$
\begin{aligned}
\pi^{\perp}_{0,U,V} &= \frac{\eta_{U\times V}}{\mathrm{id}_{U^\#}} \Big| \frac{\frac{\mathrm{id}_{U^\#+V^\#}}{\pi_{0,U,V}}}{\mathrm{id}_{U^\#}} \Big| \frac{\mathrm{id}_{U^\#+V^\#}}{\epsilon_U} = \frac{\frac{\mathrm{get}_R^{U^\#+V^\#}}{\mathrm{put}_R^{U\times V}}}{\mathrm{id}_{U^\#}} \Big| \frac{\frac{\mathrm{id}_{U^\#+V^\#}}{\pi_{0,U,V}}}{\mathrm{id}_{U^\#}} \Big| \frac{\mathrm{id}_{U^\#+V^\#}}{\epsilon_U} \\
&= \frac{\frac{\mathrm{get}_R^{U^\#+V^\#} \,|\mathrm{id}_{U^\#+V^\#}}{\mathrm{put}_R^{U\times V} \,|\pi_{0,U,V}}}{\mathrm{id}_{U^\#}|\mathrm{id}_{U^\#}} \Big| \frac{\mathrm{id}_{U^\#+V^\#}}{\epsilon_U} = \frac{\frac{\mathrm{get}_R^{U^\#+V^\#}}{\frac{\ulcorner \sigma_{0,\tilde{U},\tilde{V}} \lrcorner}{\mathrm{put}_R^{U}}}}{\mathrm{id}_{U^\#}} \Big| \frac{\mathrm{id}_{U^\#+V^\#}}{\epsilon_U} \\
&= \frac{\frac{\mathrm{get}_R^{U^\#} \,|\amalg_{0,U^\#,V^\#}}{\mathrm{put}_R^{U}}}{\mathrm{id}_{U^\#}} \Big| \frac{\mathrm{id}_{U^\#+V^\#}}{\epsilon_U} = \frac{\mathrm{get}_R^{U^\#} \,|\amalg_{0,U^\#,V^\#} \,|\mathrm{id}_{U^\#+V^\#}}{\frac{\mathrm{put}_R^{U}}{\mathrm{id}_{U^\#}}\Big|\epsilon_U} \qquad (4.2.85) \\
&= \frac{\mathrm{get}_R^{U^\#} \,|\mathrm{id}_{U^\#}|\amalg_{0,U^\#,V^\#}}{\frac{\mathrm{put}_R^{U}}{\mathrm{id}_{U^\#}}\Big|\epsilon_U|1_I} = \frac{\frac{\mathrm{get}_R^{U^\#}}{\mathrm{put}_R^{U}}}{\mathrm{id}_{U^\#}} \Big| \frac{\mathrm{id}_{U^\#}}{\epsilon_U} \Big| \frac{\amalg_{0,U^\#,V^\#}}{1_I} \\
&= \frac{\eta_U}{\mathrm{id}_{U^\#}} \Big| \frac{\mathrm{id}_{U^\#}}{\epsilon_U} \Big| \amalg_{0,U^\#,V^\#} = \mathrm{id}_{U^\#}|\amalg_{0,U^\#,V^\#} = \amalg_{0,U^\#,V^\#},
\end{aligned}
$$

where we used the snake equation in $\mathbf{H}\ulcorner\mathfrak{C}\lrcorner$ for the second to last equality. (4.2.84b) follows by a similar argument. ❏

**Lemma 4.2.15. The Transpose of the Right Choice Cell is the Left Choice Cell**

Let $\mathfrak{C}$ be a resource theory. Given a pair of cells $\alpha\colon \ulcorner\mathfrak{C}\lrcorner(W\,{}^{A}_{B}\,U)$ and $\beta\colon \ulcorner\mathfrak{C}\lrcorner(W\,{}^{A}_{B}\,V)$, we have that

$$(\alpha \boxtimes \beta)^{\perp} = \alpha^{\perp} \boxplus \beta^{\perp}. \qquad (4.2.86)$$

♡

***Proof.*** By Lemma 4.2.14 and the fact that transpose is a contravention functor in $\mathbf{H}\ulcorner\mathfrak{C}\lrcorner$,

$$\amalg_{0,U^\#,V^\#}\,|(\alpha \boxtimes \beta)^{\perp} = \pi^{\perp}_{0,U,V}|(\alpha \boxtimes \beta)^{\perp} = [(\alpha \boxtimes \beta)|\pi_{0,U,V}]^{\perp} = \alpha^{\perp}, \qquad (4.2.87a)$$

$$\amalg_{1,U^\#,V^\#}\,|(\alpha \boxtimes \beta)^{\perp} = \pi^{\perp}_{1,U,V}|(\alpha \boxtimes \beta)^{\perp} = [(\alpha \boxtimes \beta)|\pi_{1,U,V}]^{\perp} = \beta^{\perp}, \qquad (4.2.87b)$$

so by the uniqueness of the left choice cell, $(\alpha \boxtimes \beta)^{\perp} = \alpha^{\perp} \boxplus \beta^{\perp}$. ❏

Now, we define the braiding of the category of resource transducers, which is essentially the cornering of the symmetric braiding of the underlying resource category, and prove that this braiding is natural and symmetric in the category of resource transducers.

**Definition 4.2.16. The Braiding Cell**

Let $\mathfrak{C}$ be a resource theory. Given $\mathfrak{C}$-valued exchanges $U, V \in \mathfrak{C}^{\triangleright\blacktriangleleft}_{*}$, we have the *braiding* cell $\sigma_{U,V}\colon \ulcorner\mathfrak{C}\lrcorner(U\otimes V\,{}^{I}_{I}\,V\otimes U)$, which is defined by

$$\sigma_{U,V} = \frac{\mathrm{get}_L^{U\otimes V}}{\frac{\ulcorner \sigma_{\tilde{V},\tilde{U}} \lrcorner}{\mathrm{put}_R^{V\otimes U}}}. \qquad (4.2.88)$$

♣

**Lemma 4.2.17**

Let $\mathfrak{C}$ be a resource theory. Given cells $h\colon \ulcorner\mathfrak{C}\urcorner(X\,{}^{I}_{I}\,Z)$ and $g\colon \ulcorner\mathfrak{C}\urcorner(Y\,{}^{I}_{I}\,W)$,

$$\frac{h}{\mathrm{id}_W}|\sigma_{Z,W} = \sigma_{X,W}|\frac{\mathrm{id}_W}{h}, \tag{4.2.89a}$$

$$\frac{\mathrm{id}_X}{g}|\sigma_{X,W} = \sigma_{X,Y}|\frac{g}{\mathrm{id}_X}. \tag{4.2.89b}$$

♡

***Proof.*** (4.2.89a) and (4.2.89b) follow by induction on the cells $h$ and $g$, respectively. ❑

**Corollary 4.2.18. Naturality of the Braiding Cell**

Let $\mathfrak{C}$ be a resource theory. Given cells $h\colon \ulcorner\mathfrak{C}\urcorner(X\,{}^{I}_{I}\,Z)$ and $g\colon \ulcorner\mathfrak{C}\urcorner(Y\,{}^{I}_{I}\,W)$,

$$\frac{h}{g}|\sigma_{Z,W} = \sigma_{X,Y}|\frac{g}{h} \tag{4.2.90}$$

♡

***Proof.*** As we can see,

$$\begin{aligned}\frac{h}{g}|\sigma_{Z,W} &= \frac{\mathrm{id}_X|h}{g|\mathrm{id}_W}|\sigma_{Z,W} = \frac{\mathrm{id}_X}{g}|\frac{h}{\mathrm{id}_W}|\sigma_{Z,W} = \frac{\mathrm{id}_X}{g}|\sigma_{X,W}|\frac{\mathrm{id}_W}{h} \\ &= \sigma_{X,Y}|\frac{g}{\mathrm{id}_X}|\frac{\mathrm{id}_W}{h} = \sigma_{X,Y}|\frac{g|\mathrm{id}_W}{\mathrm{id}_X|h} = \sigma_{X,Y}|\frac{g}{h},\end{aligned} \tag{4.2.91}$$

where we used (4.2.89a) for the third equality and (4.2.89b) for the fourth equality. ❑

**Lemma 4.2.19. The Braiding Cell is it's Own Inverse**

Let $\mathfrak{C}$ be a resource theory. $\mathfrak{C}$-valued exchanges $U, V \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\sigma_{U,V}|\sigma_{V,U} = \mathrm{id}_{U\otimes V} \tag{4.2.92}$$

♡

***Proof.*** As we can see,

$$\begin{aligned}\sigma_{U,V}|\sigma_{V,U} &= \frac{\mathrm{get}_L^{U\otimes V}}{\dfrac{\ulcorner\sigma_{\tilde V,\tilde U}\urcorner}{\mathrm{put}_R^{V\otimes U}}}\Bigg|\frac{\mathrm{get}_L^{V\otimes U}}{\dfrac{\ulcorner\sigma_{\tilde U,\tilde V}\urcorner}{\mathrm{put}_R^{U\otimes V}}} = \frac{\dfrac{\mathrm{get}_L^{U\otimes V}}{\dfrac{\ulcorner\sigma_{\tilde V,\tilde U}\urcorner}{\mathrm{put}_R^{V\otimes U}}}}{\dfrac{1_I}{1_I}}\Bigg|\frac{\dfrac{1_I}{\dfrac{1_I}{\mathrm{get}_L^{V\otimes U}}}}{\dfrac{\ulcorner\sigma_{\tilde U,\tilde V}\urcorner}{\mathrm{put}_R^{U\otimes V}}} = \frac{\dfrac{\mathrm{get}_L^{U\otimes V}|1_I}{\dfrac{\ulcorner\sigma_{\tilde V,\tilde U}\urcorner|1_I}{\mathrm{put}_R^{V\otimes U}|\mathrm{get}_L^{V\otimes U}}}}{\dfrac{1_I|\ulcorner\sigma_{\tilde U,\tilde V}\urcorner}{1_I|\mathrm{put}_R^{U\otimes V}}} = \frac{\dfrac{\mathrm{get}_L^{U\otimes V}}{\dfrac{\ulcorner\sigma_{\tilde V,\tilde U}\urcorner}{1_{\tilde V\otimes\tilde U}}}}{\dfrac{\ulcorner\sigma_{\tilde U,\tilde V}\urcorner}{\mathrm{put}_R^{U\otimes V}}} \\ &= \frac{\mathrm{get}_L^{U\otimes V}}{\dfrac{\ulcorner\sigma_{\tilde U,\tilde V}\circ 1_{\tilde V\otimes\tilde U}\circ\sigma_{\tilde V,\tilde U}\urcorner}{\mathrm{put}_R^{U\otimes V}}} = \frac{\mathrm{get}_L^{U\otimes V}}{\dfrac{\ulcorner\sigma_{\tilde U,\tilde V}\circ\sigma_{\tilde V,\tilde U}\urcorner}{\mathrm{put}_R^{U\otimes V}}} = \frac{\mathrm{get}_L^{U\otimes V}}{\dfrac{\ulcorner 1_{\tilde U\otimes\tilde V}\urcorner}{\mathrm{put}_R^{U\otimes V}}} = \frac{\mathrm{get}_L^{U\otimes V}}{\dfrac{1_{\tilde U\otimes\tilde V}}{\mathrm{put}_R^{U\otimes V}}} = \frac{\mathrm{get}_L^{U\otimes V}}{\mathrm{put}_R^{U\otimes V}} = \mathrm{id}_{U\otimes V},\end{aligned} \tag{4.2.93}$$

where we used the yanking equation in the fourth and last equality. ❑

Putting this all together, we get the following theorem.

**Theorem 4.2.20. The Category of Resource Transducers is a KCC**

Let $\mathfrak{C}$ be a resource theory. Then $\mathbf{H}\ulcorner\mathfrak{C}\urcorner$ is a compact closed category (i.e., *KCC*).

♡

***Proof.*** This is an immediate consequence of Lemmas 4.2.13 and 4.2.19 and Corollary 4.2.18. ❑

## 4.3 Showing that the Category of Resource Transducers is a Differential Category

In the prior section, we showed that the category of resource transducers is compact closed, and hence, by Theorem 4.1.6, it is $*$-autonomous, so to show that the category of resource transducers is a differential category, all we need to show is that there exists a bialgebra and a corresponding co-dereliction in it. The first step is to define a functor that act like "bang" or "whimper." By our definition of the forward-time underlying resource, we know that iteration acts like "bang" and "whimper," so iteration should play a key role in our definition of the differential category. To make an iteration functor, we need some way to take an ordinary cell and turn it into a loop. Fortunately, there is a canonical way to do this, as pointed out in Remark 3 of [16].

**Definition 4.3.1. The Canonical Iteration Cell**

Let $\mathfrak{C}$ be a resource theory. Take a cell $\alpha\colon \ulcorner\mathfrak{C}\lrcorner\left(U\,{}^{W}_{W}\,V\right)$.

(i) The *right canonical $\alpha$-iteration cell* $\alpha^{\times}\colon \ulcorner\mathfrak{C}\lrcorner\left(U^{\times}\,{}^{W}_{W}\,V^{\times}\right)$ is the unique right iteration cell such that

$$\alpha^{\times}\,|\,\pi_{0,V} = \pi_{0,U} \quad\text{and}\quad \alpha^{\times}\,|\,\pi_{1,V} = \pi_{1,U}\,|\,\frac{\alpha}{\alpha^{\times}}. \tag{4.3.1}$$

That is $\alpha^{\times} = (\alpha)^{\times}_{\pi_{0,U},\pi_{1,U},1_I}$.

(ii) The *left canonical $\alpha$-iteration cell* $\alpha^{\times}\colon \ulcorner\mathfrak{C}\lrcorner\left(U^{+}\,{}^{W}_{W}\,V^{+}\right)$ is the unique left iteration cell such that

$$\rotatebox[origin=c]{180}{$\pi$}_{0,U}\,|\,\alpha^{+} = \rotatebox[origin=c]{180}{$\pi$}_{0,V} \quad\text{and}\quad \rotatebox[origin=c]{180}{$\pi$}_{1,U}\,|\,\alpha^{+} = \frac{\alpha}{\alpha^{+}}\,|\,\rotatebox[origin=c]{180}{$\pi$}_{1,V}\,. \tag{4.3.2}$$

That is $\alpha^{+} = (\alpha)^{+}_{\rotatebox[origin=c]{180}{$\pi$}_{0,V},\rotatebox[origin=c]{180}{$\pi$}_{1,V},1_I}$.

♣

This canonical iteration cell induces a canonical iteration endofunctor in the category of resource transducers.

**Definition 4.3.2. The Iteration Functors**

The *right iteration functor* $(-)^{\times}\colon \mathbf{H}\ulcorner\mathfrak{C}\lrcorner \to \mathbf{H}\ulcorner\mathfrak{C}\lrcorner$ is the functor such that

(i) given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}^{\triangleright\blacktriangleleft}_{*}$, $(U)^{\times} = U^{\times}$, and

(ii) given a horizontal cell $\alpha\colon \ulcorner\mathfrak{C}\lrcorner\left(U\,{}^{I}_{I}\,V\right)$, $(\alpha)^{\times} = \alpha^{\times}$,

and the *left iteration functor* $(-)^{+}\colon \mathbf{H}\ulcorner\mathfrak{C}\lrcorner \to \mathbf{H}\ulcorner\mathfrak{C}\lrcorner$ is the functor such that

(iii) given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}^{\triangleright\blacktriangleleft}_{*}$, $(U)^{+} = U^{+}$, and

(iv) given a horizontal cell $\alpha\colon \ulcorner\mathfrak{C}\lrcorner\left(U\,{}^{I}_{I}\,V\right)$, $(\alpha)^{+} = \alpha^{+}$.

♣

In our category of resource transducers, we have 14 structure cells involving the iteration type. It should be noted that structure cells (u)-(iv) and (viii)-(xi) already appear in [16] and that the right iteration functor together with the first seven structure cells gives us all data we need to define the category of resource transducers is a differential category.

**Definition 4.3.3. The Structure Cells**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$, we have the following *structure cells* in the free cornering $\ulcorner\mathfrak{C}\lrcorner$ with choice and iteration.

(i) The *right comultiplication cell* $\Delta_U^\times \colon \ulcorner\mathfrak{C}\lrcorner(U^\times\,{}^{I}_{I}\,U^\times \otimes U^\times)$, which is defined by

$$\Delta_U^\times = (\mathrm{id}_U)^\times_{\mathrm{id}_{U^\times}, \pi_{1,U}, 1_I}. \tag{4.3.3}$$

(ii) The *right counit cell* $e_U^\times \colon \ulcorner\mathfrak{C}\lrcorner(U^\times\,{}^{I}_{I}\,I)$, which is defined by

$$e_U^\times = \pi_{0,U}. \tag{4.3.4}$$

(iii) The *right comonadic counit cell* $\mu_U^\times \colon \ulcorner\mathfrak{C}\lrcorner(U^\times\,{}^{I}_{I}\,U)$, which is defined by

$$\mu_U^\times = \pi_{1,U} \Big| \frac{\mathrm{id}_U}{\pi_{0,U}}. \tag{4.3.5}$$

(iv) The *right comonadic comultiplication cell* $\delta_U^\times \colon \ulcorner\mathfrak{C}\lrcorner(U^\times\,{}^{I}_{I}\,U^{\times\times})$, which is defined by

$$\delta_U^\times = (\mathrm{id}_{U^\times})^\times_{\pi_{0,U}, \Delta_U^\times, 1_I}. \tag{4.3.6}$$

(v) The *right multiplication cell* $\nabla_U^\times \colon \ulcorner\mathfrak{C}\lrcorner(U^\times \otimes U^\times\,{}^{I}_{I}\,U^\times)$, which is defined by

$$\nabla_U^\times = (\mathrm{id}_U)^\times_{\frac{\pi_{0,U}}{\pi_{0,U}}, \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}} + \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}} \big| \frac{\sigma_{U^\times, U}}{\mathrm{id}_{U^\times}}, 1_I}. \tag{4.3.7}$$

(vi) The *right unit cell* $u_U^\times \colon \ulcorner\mathfrak{C}\lrcorner(I\,{}^{I}_{I}\,U^\times)$, which is defined by

$$u_U^\times = (0)^\times_{1_I, 1_I, 1_I}. \tag{4.3.8}$$

(vii) The *right plastic cell* $\nu_U^\times \colon \ulcorner\mathfrak{C}\lrcorner(U\,{}^{I}_{I}\,U^\times)$, which is defined by

$$\nu_U^\times = 0 \boxtimes \frac{\mathrm{id}_U}{u_U^\times}. \tag{4.3.9}$$

(viii) The *left multiplication cell* $\nabla_U^+ \colon \ulcorner\mathfrak{C}\lrcorner(U^+ \otimes U^+\,{}^{I}_{I}\,U^+)$, which is defined by

$$\nabla_U^+ = (\mathrm{id}_U)^+_{\mathrm{id}_{U^+}, \rotatebox[origin=c]{180}{$\pi$}_{1,U}, 1_I}. \tag{4.3.10}$$

(ix) The *left unit cell* $\rotatebox[origin=c]{180}{$e$}_U^+ \colon \ulcorner\mathfrak{C}\lrcorner(I\,{}^{I}_{I}\,U^+)$, which is defined by

$$\rotatebox[origin=c]{180}{$e$}_U^+ = \rotatebox[origin=c]{180}{$\pi$}_{0,U}. \tag{4.3.11}$$

(x) The *left monadic unit cell* $\reflectbox{$\eta$}_U^+ \colon \mathfrak{C}(U \,{}^{I}_{I}\, U^+)$, which is defined by

$$\reflectbox{$\eta$}_U^+ = \frac{\mathrm{id}_U}{\rotatebox[origin=c]{180}{$\pi$}_{0,U}} \Big| \rotatebox[origin=c]{180}{$\pi$}_{1,U} \,. \tag{4.3.12}$$

(xi) The *left monadic multiplication cell* $\reflectbox{$\varrho$}_U^+ \colon \mathfrak{C}(U^{++} \,{}^{I}_{I}\, U^+)$, which is defined by

$$\reflectbox{$\varrho$}_U^+ = (\mathrm{id}_{U^+})^+_{\rotatebox[origin=c]{180}{$\pi$}_{0,U}, \nabla_U^+, 1_I}. \tag{4.3.13}$$

(xii) The *left comultiplication cell* $\Delta_U^+ \colon \mathfrak{C}(U^+ \,{}^{I}_{I}\, U^+ \otimes U^+)$, which is defined by

$$\Delta_U^+ = (\mathrm{id}_U)^+_{\frac{\rotatebox[origin=c]{180}{$\pi$}_{0,U}}{\rotatebox[origin=c]{180}{$\pi$}_{0,U}}, \frac{\rotatebox[origin=c]{180}{$\pi$}_{1,U}}{\mathrm{id}_{U^+}} + \frac{\sigma_{U,U^+}}{\mathrm{id}_U^+} | \frac{\mathrm{id}_{U^+}}{\rotatebox[origin=c]{180}{$\pi$}_{1,U}}, 1_I}. \tag{4.3.14}$$

(xiii) The *left counit cell* $n_U^+ =\colon \mathfrak{C}(U^+ \,{}^{I}_{I}\, I)$, which is defined by

$$n_U^+ = u_U^\times = (0)^+_{1_I, 1_I, 1_I}. \tag{4.3.15}$$

(xiv) The *left plastic cell* $\rotatebox[origin=c]{180}{$\nu$}_U^+ \colon \mathfrak{C}(U^+ \,{}^{I}_{I}\, U)$, which is defined by

$$\rotatebox[origin=c]{180}{$\nu$}_U^+ = 0 \boxplus \frac{\mathrm{id}_U}{n_U^+}. \tag{4.3.16}$$

♣

Even though we have all the maps to define a differential categorical structure on the category of resource transducers, we are still a bit away from claiming that the category of resource transducers is, in fact, a differential category. We still need to show that $((-)^\times, \delta^\times, \mu^\times, \Delta^\times, e^\times, \nabla^\times, u^\times)$ satisfies all the properties of a bialgebra modality and $\nu^\times$ satisfies all the properties of a co-dereliction on that bialgebra. To get the ball rolling, we restate two lemmas from [16] .

**Lemma 4.3.4**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

(i) $(U^\times, \Delta_U^\times, e_U^\times)$ is a comonoid,

(ii) $(U^+, \nabla_U^+, \rotatebox[origin=c]{180}{$e$}_U^+)$ is a monoid.

Also, Given a cell $h \in \mathfrak{C}(U \,{}^{I}_{I}\, V)$

(iii) $h^\times | \Delta_V^\times = \Delta_U^\times | \frac{h^\times}{h^\times}$, and

(iv) $\nabla_U^+ | h^+ = \frac{h^+}{h^+} | \nabla_V$.

♡

***Proof.*** This is shown in Lemma 18 of [16]. ❏

**Lemma 4.3.5**

Let $\mathfrak{C}$ be a resource theory. Then,

(i) $((-)^\times, \delta^\times, \mu^\times)$ is a comonad in $\mathbf{H}\ulcorner\mathfrak{C}\lrcorner$, and

(ii) $((-)^+, \varrho^+, \eta^+)$ is a monad in $\mathbf{H}\ulcorner\mathfrak{C}\lrcorner$,

and given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\rhd\blacktriangleleft}$,

(iii) $\delta_U^\times | \Delta_{U^\times}^\times = \Delta_U^\times | \frac{\delta_U^\times}{\delta_U^\times}$ (i.e., $\delta^\times$ is a comonoid homomorphism), and

(iv) $\nabla_{U^+}^+ | \varrho_U^+ = \frac{\varrho_U^+}{\varrho_U^+} | \nabla_U^+$ (i.e., $\varrho^+$ is a monoid homomorphism). ♡

***Proof.*** This is shown in Lemma 19 of [16]. ❑

Importantly, these lemmas prove the naturality of the right comultiplication cell, the right comonadic counit cell, the right comonadic comultiplication cell, We, now, prove the naturality of the right counit cell, the right unit cell, the right multiplication cell, and the right plastic cell.

**Lemma 4.3.6. Naturality of The Right Counit Cell**

Let $\mathfrak{C}$ be a resource theory. Given a cell $h\colon \ulcorner\mathfrak{C}\lrcorner\left(U\,{}^{I}_{I}\,V\right)$,

$$h^\times | e_V^\times = e_U^\times \tag{4.3.17}$$ ♡

***Proof.*** $h^\times | e_V^\times = h^\times | \pi_{0,V} = \pi_{0,U} = e_U^\times$. ❑

**Lemma 4.3.7. Naturality of the Right Unit Cell**

Let $\mathfrak{C}$ be a resource theory. Given a cell $h\colon \ulcorner\mathfrak{C}\lrcorner\left(U\,{}^{I}_{I}\,V\right)$,

$$u_U^\times | h^\times = u_V^\times \tag{4.3.18}$$ ♡

***Proof.*** As we can see,

$$u_U^\times | h^\times | \pi_{0,V} = u_U^\times | \pi_{0,U} = (0)^\times_{1_I, 1_I, 1_I} | \pi_{0,U} = 1_I, \tag{4.3.19a}$$

$$u_U^\times | h^\times | \pi_{1,V} = u_U^\times | \pi_{1,U} | \frac{h}{h^\times} = (0)^\times_{1_I, 1_I, 1_I} | \pi_{1,U} | \frac{h}{h^\times} = \frac{1_I | \frac{0}{1_I}}{1_I} | \frac{h}{h^\times} = 0 = \frac{1_I | \frac{0}{u_U^\times | h^\times}}{1_I}, \tag{4.3.19b}$$

so by the uniqueness the right iteration cell, $u_U^\times | h^\times = (0)^\times_{1_I, 1_I, 1_I} = u_V^\times$. ❑

**Lemma 4.3.8. Naturality of the Right Multiplication Cell**

Let $\mathfrak{C}$ be a resource theory. Given a cell $h\colon \ulcorner\mathfrak{C}\lrcorner\left(U\,{}^{I}_{I}\,V\right)$,

$$\frac{h^\times}{h^\times} | \nabla_V^\times = \nabla_U^\times | h^\times \tag{4.3.20}$$ ♡

***Proof.*** As we can see,

$$\frac{h^\times}{h^\times}|\nabla_V^\times|\pi_{0,V} = \frac{h^\times}{h^\times}|(\mathrm{id}_V)^\times_{\frac{\pi_{0,V}}{\pi_{0,V}},\frac{\pi_{1,V}}{\mathrm{id}_{V^\times}}+\frac{\mathrm{id}_{V^\times}}{\pi_{1,V}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}},1_I}|\pi_{0,V} = \frac{h^\times}{h^\times}|\frac{\pi_{0,V}}{\pi_{0,V}} = \frac{h^\times|\pi_{0,V}}{h^\times|\pi_{0,V}} = \frac{\pi_{0,U}}{\pi_{0,U}}, \tag{4.3.21}$$

and

$$\begin{aligned}
\frac{h^\times}{h^\times}|\nabla_V^\times|\pi_{1,V} &= \frac{h^\times}{h^\times}|(\mathrm{id}_V)^\times_{\frac{\pi_{0,V}}{\pi_{0,V}},\frac{\pi_{1,V}}{\mathrm{id}_{V^\times}}+\frac{\mathrm{id}_{V^\times}}{\pi_{1,V}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}},1_I}|\pi_{1,V} \\
&= \frac{h^\times}{h^\times}|\left(\frac{\pi_{1,V}}{\mathrm{id}_{V^\times}}+\frac{\mathrm{id}_{V^\times}}{\pi_{1,V}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}}\right)|\frac{\mathrm{id}_V}{(\mathrm{id}_V)^\times_{\frac{\pi_{0,V}}{\pi_{0,V}},\frac{\pi_{1,V}}{\mathrm{id}_{V^\times}}+\frac{\mathrm{id}_{V^\times}}{\pi_{1,V}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}},1_I}} \\
&= \frac{h^\times}{h^\times}|\left(\frac{\pi_{1,V}}{\mathrm{id}_{V^\times}}+\frac{\mathrm{id}_{V^\times}}{\pi_{1,V}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}}\right)|\frac{\mathrm{id}_V}{\nabla_V^\times} \\
&= \frac{h^\times}{h^\times}|\frac{\pi_{1,V}}{\mathrm{id}_{V^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} + \frac{h^\times}{h^\times}|\frac{\mathrm{id}_{V^\times}}{\pi_{1,V}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} \\
&= \frac{h^\times|\pi_{1,V}}{h^\times|\mathrm{id}_{V^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} + \frac{h^\times|\mathrm{id}_{V^\times}}{h^\times|\pi_{1,V}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} \\
&= \frac{\pi_{1,U}|\frac{h}{h^\times}}{\mathrm{id}_{U^\times}|h^\times}|\frac{\mathrm{id}_V}{\nabla_V^\times} + \frac{\mathrm{id}_{U^\times}|h^\times}{\pi_{1,U}|\frac{h}{h^\times}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} \\
&= \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}|\frac{h}{\frac{h^\times}{h^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} + \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\frac{h^\times}{h}}{h^\times}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} \\
&= \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}|\frac{h}{\frac{h^\times}{h^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} + \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\frac{h^\times}{h}|\sigma_{V^\times,V}}{h^\times|\mathrm{id}_{V^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} \\
&= \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}|\frac{h}{\frac{h^\times}{h^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} + \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\sigma_{U^\times,U}|\frac{h}{h^\times}}{\mathrm{id}_{U^\times}|h^\times}|\frac{\mathrm{id}_V}{\nabla_V^\times} \\
&= \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}|\frac{h}{\frac{h^\times}{h^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} + \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}}|\frac{h}{\frac{h^\times}{h^\times}}|\frac{\mathrm{id}_V}{\nabla_V^\times} \\
&= \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}|\frac{h|\mathrm{id}_V}{\frac{h^\times}{h^\times}|\nabla_V^\times} + \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}}|\frac{h|\mathrm{id}_V}{\frac{h^\times}{h^\times}|\nabla_V^\times} \\
&= \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}|\frac{h}{\frac{h^\times}{h^\times}|\nabla_V^\times} + \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}}|\frac{h}{\frac{h^\times}{h^\times}|\nabla_V^\times} \\
&= \left(\frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}+\frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}}\right)|\frac{h}{\frac{h^\times}{h^\times}|\nabla_V^\times},
\end{aligned} \tag{4.3.22}$$

where we used Corollary 4.2.18 for the eighth equality. Hence, by the uniqueness of the right iteration cell,

$$\frac{h^\times}{h^\times}|\nabla_V^\times = (h)^\times_{\frac{\pi_{0,V}}{\pi_{0,V}},\frac{\pi_{1,V}}{\mathrm{id}_{V^\times}}+\frac{\mathrm{id}_{V^\times}}{\pi_{1,V}}|\frac{\sigma_{V^\times,V}}{\mathrm{id}_{V^\times}},1_I}. \tag{4.3.23}$$

Also,

$$\nabla_U^\times|h^\times|\pi_{0,V} = \nabla_U^\times|\pi_{0,U} = (\mathrm{id}_U)^\times_{\frac{\pi_{0,U}}{\pi_{0,U}},\frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}+\frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}},1_I}|\pi_{0,U} = \frac{\pi_{0,U}}{\pi_{0,U}} \tag{4.3.24}$$

and

$$\begin{aligned}\nabla^{\times}_{U}|h^{\times}|\pi_{1,V} &= \nabla^{\times}_{U}|\pi_{1,V}|\frac{h}{h^{\times}} = (\mathrm{id}_U)^{\times}_{\frac{\pi_{0,U}}{\pi_{0,U}},\frac{\pi_{1,U}}{\mathrm{id}_{U^{\times}}}+\frac{\mathrm{id}_{U^{\times}}}{\pi_{1,U}}|\frac{\sigma_{U^{\times},U}}{\mathrm{id}_{U^{\times}}},1_I}|\pi_{1,V}|\frac{h}{h^{\times}} \\ &= \left(\frac{\pi_{1,U}}{\mathrm{id}_{U^{\times}}}+\frac{\mathrm{id}_{U^{\times}}}{\pi_{1,U}}|\frac{\sigma_{U^{\times},U}}{\mathrm{id}_{U^{\times}}}\right)|\frac{\mathrm{id}_U}{(\mathrm{id}_U)^{\times}_{\frac{\pi_{0,U}}{\pi_{0,U}},\frac{\pi_{1,U}}{\mathrm{id}_{U^{\times}}}+\frac{\mathrm{id}_{U^{\times}}}{\pi_{1,U}}|\frac{\sigma_{U^{\times},U}}{\mathrm{id}_{U^{\times}}},1_I}}|\frac{h}{h^{\times}} \\ &= \left(\frac{\pi_{1,U}}{\mathrm{id}_{U^{\times}}}+\frac{\mathrm{id}_{U^{\times}}}{\pi_{1,U}}|\frac{\sigma_{U^{\times},U}}{\mathrm{id}_{U^{\times}}}\right)|\frac{\mathrm{id}_U}{\nabla^{\times}_{U}}|\frac{h}{h^{\times}} = \left(\frac{\pi_{1,U}}{\mathrm{id}_{U^{\times}}}+\frac{\mathrm{id}_{U^{\times}}}{\pi_{1,U}}|\frac{\sigma_{U^{\times},U}}{\mathrm{id}_{U^{\times}}}\right)|\frac{\mathrm{id}_U|h}{\nabla^{\times}_{U}|h^{\times}} \\ &= \left(\frac{\pi_{1,U}}{\mathrm{id}_{U^{\times}}}+\frac{\mathrm{id}_{U^{\times}}}{\pi_{1,U}}|\frac{\sigma_{U^{\times},U}}{\mathrm{id}_{U^{\times}}}\right)|\frac{h}{\nabla^{\times}_{U}|h^{\times}},\end{aligned} \tag{4.3.25}$$

so by the uniqueness of the right iteration cell,

$$\nabla^{\times}_{U}|h^{\times} = (h)^{\times}_{\frac{\pi_{0,V}}{\pi_{0,V}},\frac{\pi_{1,V}}{\mathrm{id}_{V^{\times}}}+\frac{\mathrm{id}_{V^{\times}}}{\pi_{1,V}}|\frac{\sigma_{V^{\times},V}}{\mathrm{id}_{V^{\times}}},1_I}. \tag{4.3.26}$$

Therefore, $\frac{h^{\times}}{h^{\times}}|\nabla^{\times}_{V} = \nabla^{\times}_{U}|h^{\times}$.

❏

**Lemma 4.3.9. Naturality of the Right Plastic Cell**

Let $\mathfrak{C}$ be a resource theory. Given a cell $h\colon \ulcorner\mathfrak{C}\urcorner(U\,{}^{I}_{I}\,V)$,

$$h|\nu^{\times}_{V} = \nu^{\times}_{U}|h^{\times} \tag{4.3.27}$$

♡

***Proof*.** As we can see,

$$h|\nu^{\times}_{V}|\pi_{0,V} = h|\left(0\boxtimes\frac{\mathrm{id}_V}{u^{\times}_{V}}\right)|\pi_{0,V} = h|0 = 0 \tag{4.3.28a}$$

$$h|\nu^{\times}_{V}|\pi_{1,V} = h|\left(0\boxtimes\frac{\mathrm{id}_V}{u^{\times}_{V}}\right)|\pi_{1,V} = h|\frac{\mathrm{id}_V}{u^{\times}_{V}} = \frac{h}{1_I}|\frac{\mathrm{id}_V}{u^{\times}_{V}} = \frac{h|\mathrm{id}_V}{1_I|u^{\times}_{V}} = \frac{h}{u^{\times}_{V}}, \tag{4.3.28b}$$

so by the uniqueness of this right choice cell,

$$h|\nu^{\times}_{V} = 0\boxtimes\frac{h}{u^{\times}_{V}}. \tag{4.3.29}$$

Similarly,

$$\nu^{\times}_{U}|h^{\times}|\pi_{0,V} = \nu^{\times}_{U}|\pi_{0,U} = \left(0\boxtimes\frac{\mathrm{id}_U}{u^{\times}_{U}}\right)|\pi_{0,U} = 0, \tag{4.3.30a}$$

$$\nu^{\times}_{U}|h^{\times}|\pi_{1,V} = \nu^{\times}_{U}|\pi_{1,U}|\frac{h}{h^{\times}} = \left(0\boxtimes\frac{\mathrm{id}_U}{u^{\times}_{U}}\right)|\pi_{1,U}|\frac{h}{h^{\times}} = \frac{\mathrm{id}_U}{u^{\times}_{U}}|\frac{h}{h^{\times}} = \frac{\mathrm{id}_U|h}{u^{\times}_{U}|h^{\times}} = \frac{h}{u^{\times}_{V}}, \tag{4.3.30b}$$

where we used Lemma 4.3.7 for the last equality of (4.3.30b), so by the uniqueness of this right choice cell,

$$\nu^{\times}_{U}|h^{\times} = 0\boxtimes\frac{h}{u^{\times}_{V}}. \tag{4.3.31}$$

Hence, $h|\nu^{\times}_{V} = \nu^{\times}_{U}|h^{\times}$. ❏

Finally, we show the first seven structure cells are the cornering of contraction, weakening, dereliction, digging, co-contraction, co-wakening, and co-dereliction, respectively.

**Lemma 4.3.10. The Right Comultiplication Cell is Contraction**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\Delta_U^\times = \frac{\mathrm{get}_L^{U^\times}}{\frac{\ulcorner\Delta_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times\otimes U^\times}}} \cong (\Delta_{\tilde{U}})^{\triangleright}. \tag{4.3.32}$$

♡

***Proof.*** As we can see,

$$\begin{aligned}
\frac{\mathrm{get}_L^{U^\times}}{\frac{\ulcorner\Delta_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times\otimes U^\times}}} \Big| \frac{\pi_{0,U}}{\mathrm{id}_{U^\times}} &= \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^\times}}} \Big| \frac{\pi_{0,U}}{\mathrm{id}_{U^\times}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^\times}} \Big| \frac{\pi_{0,U}}{\mathrm{id}_{U^\times}}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times} \mid \pi_{0,U}}{\mathrm{put}_R^{U^\times} \mid \mathrm{id}_{U^\times}}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \mid \ulcorner e_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}} \\
&= \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner(1_{!\tilde{U}}\otimes e_{\tilde{U}})\circ\Delta_{\tilde{U}}\lrcorner}}{\mathrm{put}_R^{U^\times}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner 1_{!\tilde{U}}\lrcorner}}{\mathrm{put}_R^{U^\times}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{1_{!\tilde{U}}}}{\mathrm{put}_R^{U^\times}} = \frac{\mathrm{get}_L^{U^\times}}{\mathrm{put}_R^{U^\times}} = \mathrm{id}_{U^\times}.
\end{aligned} \tag{4.3.33}$$

and

$$\begin{aligned}
\frac{\mathrm{get}_L^{U^\times}}{\frac{\ulcorner\Delta_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times\otimes U^\times}}} \Big| \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}} &= \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^\times}}} \Big| \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^\times}} \Big| \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times} \mid \pi_{1,U}}{\mathrm{put}_R^{U^\times} \mid \mathrm{id}_{U^\times}}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \Big| \frac{\ulcorner d^\circ_{\tilde{U}}\lrcorner}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}{\mathrm{put}_R^{U^\times}}} \\
&= \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{\frac{1_{!\tilde{U}}}{1_{!\tilde{U}}} \Big| \frac{\ulcorner d^\circ_{\tilde{U}}\lrcorner}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}{\mathrm{put}_R^{U^\times}}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{\frac{1_{!\tilde{U}} \mid \ulcorner d^\circ_{\tilde{U}}\lrcorner}{1_{!\tilde{U}} \Big| \frac{1_{!\tilde{U}} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}{\mathrm{put}_R^{U^\times}}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner(1_{!\tilde{U}}\otimes d^\circ_{\tilde{U}})\circ\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \Big| \frac{1_{!\tilde{U}} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}{\mathrm{put}_R^{U^\times}}} = \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner(\Delta_{\tilde{U}}\otimes 1_{!\tilde{U}})\circ d^\circ_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \Big| \frac{1_{!\tilde{U}} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}{\mathrm{put}_R^{U^\times}}} \\
&= \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner d^\circ_{\tilde{U}}\lrcorner}}{\frac{\frac{\ulcorner\Delta_{\tilde{U}}\lrcorner \mid 1_{!\tilde{U}}}{1_{!\tilde{U}} \Big| \frac{1_{!\tilde{U}} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}{\mathrm{put}_R^{U^\times}}} = \frac{\pi_{1,U} \Big| \frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}} \Big| \frac{\mathrm{get}_L^U}{1_{\tilde{U}}}}{\frac{\frac{\ulcorner\Delta_{\tilde{U}}\lrcorner \mid 1_{!\tilde{U}}}{1_{!\tilde{U}} \Big| \frac{1_{!\tilde{U}} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}{\mathrm{put}_R^{U^\times}}} = \pi_{1,U} \Big| \frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}} \Big| \frac{\mathrm{get}_L^U}{1_{\tilde{U}}}}{\frac{\frac{\ulcorner\Delta_{\tilde{U}}\lrcorner \mid 1_{!\tilde{U}}}{\frac{1_{!\tilde{U}}}{1_{!\tilde{U}}} \Big| \frac{1_{!\tilde{U}} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times} \mid \mathrm{id}_{U^\times}}}}{\mathrm{put}_R^{U^\times}}} \\
&= \pi_{1,U} \Big| \frac{\frac{\frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\Delta_{\tilde{U}}}}{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times}} \Big| \frac{\frac{\mathrm{get}_L^U}{\mathrm{put}_R^U}}{\mathrm{id}_{U^\times}}}{\mathrm{put}_R^{U^\times}} = \pi_{1,U} \Big| \frac{\frac{\frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\Delta_{\tilde{U}}}}{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times}} \Big| \frac{\mathrm{id}_U}{\mathrm{id}_{U^\times}}}{\mathrm{put}_R^{U^\times}} \\
&= \pi_{1,U} \Big| \frac{\mathrm{id}_U}{\frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner\Delta_{\tilde{U}}\lrcorner}}{\frac{1_{!\tilde{U}} \mid \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^\times}}}} = \pi_{1,U} \Big| \frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^\times}}{\frac{\ulcorner\Delta_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times\otimes U^\times}}}}
\end{aligned} \tag{4.3.34}$$

where we used **[cd.3]** in [4] for the seventh equality. Hence, by the uniqueness property of the right iteration cell, this implies that

$$\frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner \Delta_{\tilde{U}} \lrcorner}}{\mathrm{put}_R^{U^\times \otimes U^\times}} = (\mathrm{id}_U)^\times_{\mathrm{id}_{U^\times}, \pi_{U,1}, 1_I} = \Delta_U^\times. \tag{4.3.35}$$

This completes the proof. ❏

**Lemma 4.3.11. The Right Counit Cell is Weakening**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$e_U^\times = \frac{\mathrm{get}_L^{U^\times}}{\ulcorner e_{\tilde{U}} \lrcorner} \cong (e_{\tilde{U}})^\triangleright. \tag{4.3.36}$$

♡

***Proof***. By (4.2.31) and (4.3.4),

$$\frac{\mathrm{get}_L^{U^\times}}{\ulcorner e_{\tilde{U}} \lrcorner} = \pi_{0,U} = e_U^\times \tag{4.3.37}$$

This completes the proof. ❏

**Lemma 4.3.12. The Right Comonadic Counit Cell is Dereliction**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\mu_U^\times = \frac{\mathrm{get}_L^{U^\times}}{\frac{\ulcorner \mu_{\tilde{U}} \lrcorner}{\mathrm{put}_R^U}} \cong (\mu_{\tilde{U}})^\triangleright. \tag{4.3.38}$$

♡

***Proof***. As we can see,

$$\begin{aligned}
\eta_{U^\times} \Big| \frac{\mathrm{id}_{(U^\#)^+}}{\mu_U^\times} &= \frac{\mathrm{get}_R^{(U^\#)^+}}{\mathrm{put}_R^{U^\times}} \Big| \frac{\mathrm{id}_{(U^\perp)^+}}{\pi_{1,U} \big| \frac{\mathrm{id}_U}{\pi_{0,U}}} = \frac{\mathrm{get}_R^{(U^\#)^+} \,|\mathrm{id}_{(U^\perp)^+}}{\mathrm{put}_R^{U^\times}\, |\pi_{1,U} \big| \frac{\mathrm{id}_U}{\pi_{0,U}}} = \frac{\mathrm{get}_R^{(U^\#)^+}}{\mathrm{put}_R^{U^\times}\, |\pi_{1,U} \big| \frac{\mathrm{id}_U}{\pi_{0,U}}} \\
&= \frac{\mathrm{get}_R^{(U^\#)^+}}{\frac{\frac{\ulcorner d^\circ_{\tilde{U}} \lrcorner}{1_{!\tilde{U}} \,|\, \mathrm{put}_R^U}}{\mathrm{put}_R^{U^\times}} \Big| \frac{\mathrm{id}_U}{\pi_{0,U}}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{1_{!\tilde{U}} \,|\, \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}} \Big| \frac{\mathrm{id}_U}{\pi_{0,U}}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{1_{!\tilde{U}} \,|\, \mathrm{put}_R^U \,|\mathrm{id}_U}{\mathrm{put}_R^{U^\times}\, |\pi_{0,U}}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{1_{!\tilde{U}} \,|\, \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}\, |\pi_{0,U}}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{1_{!\tilde{U}} \,|\, \mathrm{put}_R^U}{\ulcorner e_{\tilde{U}} \lrcorner}} \\
&= \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{1_{!\tilde{U}} \,|\, \mathrm{put}_R^U}{\ulcorner e_{\tilde{U}} \lrcorner \,|1_I}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{1_{!\tilde{U}}}{\ulcorner e_{\tilde{U}} \lrcorner} \Big| \frac{\mathrm{put}_R^U}{1_I}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{\ulcorner e_{\tilde{U}} \lrcorner}{1_I} \Big| \frac{1_U}{\mathrm{put}_R^U}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{\ulcorner e_{\tilde{U}} \lrcorner \,|1_U}{1_I \,|\, \mathrm{put}_R^U}} = \frac{\mathrm{get}_R^{(U^\#)^+}}{\frac{\ulcorner (e_U \otimes 1_U) \circ d^\circ_U \lrcorner}{\mathrm{put}_R^U}} = \frac{\mathrm{get}_R^{(U^\#)^+}}{\frac{\ulcorner \mu_{\tilde{U}} \lrcorner}{\mathrm{put}_R^U}} \\
&= \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{1_{\tilde{U}^\perp}}}{\frac{\ulcorner \mu_{\tilde{U}} \lrcorner}{\mathrm{put}_R^U}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\mathrm{put}_R^{U^\times} \,|\, \mathrm{get}_L^{U^\times}}}{\frac{\ulcorner \mu_{\tilde{U}} \lrcorner}{\mathrm{put}_R^U}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+} \,|\mathrm{id}_{(U^\#)^+}}{\mathrm{put}_R^{U^\times} \,|\, \mathrm{get}_L^{U^\times}}}{\frac{\ulcorner \mu_{\tilde{U}} \lrcorner}{\mathrm{put}_R^U}} = \frac{\frac{\mathrm{get}_R^{(U^\#)^+}}{\mathrm{put}_R^{U^\times}} \Big| \frac{\mathrm{id}_{(U^\#)^+}}{\mathrm{get}_L^{U^\times}}}{\frac{\ulcorner \mu_{\tilde{U}} \lrcorner}{\mathrm{put}_R^U}}
\end{aligned} \tag{4.3.39}$$

$$= \frac{\eta_{U^\times} \,\Big|\, \dfrac{\mathrm{id}_{(U\#)^+}}{\mathrm{get}_L^{U^\times}}}{\dfrac{\mu_{\tilde{U}}}{\mathrm{put}_R^{U}}} = \eta_{U^\times} \,\Big|\, \frac{\mathrm{id}_{(U\#)^+}}{\dfrac{\mathrm{get}_L^{U^\times}}{\dfrac{\mu_{\tilde{U}}}{\mathrm{put}_R^{U}}}},$$

where we used **[cd.1]** in [4] for the seventh equality. Since the names of the horizontal cells in (4.3.38) are equal, (4.3.38) holds. ❏

**Lemma 4.3.13**

Let $\mathfrak{C}$ be a differential category. Given an object $A \in \mathrm{Ob}(\mathfrak{C})$,

$$d^\circ_{!A} \circ \delta_A = (\delta_A \otimes 1_{!A}) \circ \Delta_A. \tag{4.3.40}$$

♡

***Proof.*** As we can see,

$$\begin{aligned} d^\circ_{!A} \circ \delta_A &= (1_{!!A} \otimes \mu_{!A}) \circ \Delta_{!A} \circ \delta_A = (1_{!!A} \otimes \mu_{!A}) \circ (\delta_A \otimes \delta_A) \circ \Delta_A \\ &= \Delta_A \circ ((1_{!!A} \circ \delta_A) \otimes (\mu_{!A} \circ \delta_A)) = (\delta_A \otimes 1_{!A}) \circ \Delta_A, \end{aligned} \tag{4.3.41}$$

where we used the comonadic unit law of the comonad $(!, \delta_A, \mu_A)$ for the last equality. ❏

**Lemma 4.3.14. The Right Comonadic Comultiplication Cell is Digging**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\delta_U^\times = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}}}{\mathrm{put}_R^{U^{\times\times}}} \cong (\delta_{\tilde{U}})^\triangleright. \tag{4.3.42}$$

♡

***Proof.*** As we can see,

$$\begin{aligned} \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}}}{\mathrm{put}_R^{U^{\times\times}}} \Big| \pi_{0,U^\times} &= \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}}}{\mathrm{put}_R^{U^{\times\times}}} \Big| \frac{1_I}{\pi_{0,U^\times}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}} \Big| 1_I}{\mathrm{put}_R^{U^{\times\times}} \,\Big|\, \pi_{0,U^\times}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}}}{e_{\widetilde{U^\times}}} \\ &= \frac{\mathrm{get}_L^{U^\times}}{e_{\widetilde{U^\times}} \circ \delta_{\tilde{U}}} = \frac{\mathrm{get}_L^{U^\times}}{e_{!\tilde{U}} \circ \delta_{\tilde{U}}} = \frac{\mathrm{get}_L^{U^\times}}{e_{\tilde{U}}} = \pi_{0,U}, \end{aligned} \tag{4.3.43}$$

where we used Lemma 2.2 in [2] for the second to last equality, and

$$\begin{aligned} \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}}}{\mathrm{put}_R^{U^{\times\times}}} \Big| \pi_{1,U^\times} &= \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}}}{\mathrm{put}_R^{U^{\times\times}}} \Big| \frac{1_I}{\pi_{1,U^\times}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}} \Big| 1_I}{\mathrm{put}_R^{U^{\times\times}} \,\Big|\, \pi_{1,U^\times}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}}}{\dfrac{d^\circ_{\widetilde{U^\times}}}{\dfrac{1_{!\widetilde{U^\times}} \,|\, \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^{\times\times}}}}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\delta_{\tilde{U}}}}{\dfrac{d^\circ_{!\tilde{U}}}{\dfrac{1_{!!\tilde{U}} \,|\, \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^{\times\times}}}}} \\ &= \frac{\dfrac{\mathrm{get}_L^{U^\times}}{d^\circ_{!\tilde{U}} \circ \delta_{\tilde{U}}}}{\dfrac{1_{!!\tilde{U}} \,|\, \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^{\times\times}}}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{(\delta_{\tilde{U}} \otimes 1_{!\tilde{U}}) \circ \Delta_A}}{\dfrac{1_{!!\tilde{U}} \,|\, \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^{\times\times}}}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\Delta_{\tilde{U}}}}{\dfrac{\delta_{\tilde{U}} \,|\, 1_{!\tilde{U}}}{\dfrac{1_{!!\tilde{U}} \,|\, \mathrm{put}_R^{U^\times}}{\mathrm{put}_R^{U^{\times\times}}}}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\Delta_{\tilde{U}}}}{\dfrac{\dfrac{\delta_{\tilde{U}}}{1_{!!\tilde{U}}} \,\Big|\, \dfrac{1_{!\tilde{U}}}{\mathrm{put}_R^{U^\times}}}{\mathrm{put}_R^{U^{\times\times}}}} = \frac{\dfrac{\mathrm{get}_L^{U^\times}}{\Delta_{\tilde{U}}}}{\dfrac{\dfrac{1_{!\tilde{U}}}{\delta_{\tilde{U}}} \,\Big|\, \dfrac{\mathrm{put}_R^{U^\times}}{1_I}}{\mathrm{put}_R^{U^{\times\times}}}} \end{aligned} \tag{4.3.44}$$

$$= \frac{\frac{\frac{\mathrm{get}^{U^\times}_{L}}{\Delta_{\tilde U}}}{1_{!\tilde U} \mid \mathrm{put}^{U^\times}_R}}{\frac{\delta_{\tilde U} \mid 1_I}{\mathrm{put}^{U^{\times\times}}_R}} = \frac{\frac{\frac{\mathrm{get}^{U^\times}_{L}}{\Delta_{\tilde U}}}{1_{!\tilde U} \mid \mathrm{put}^{U^\times}_R}}{\frac{\delta_{\tilde U}}{\mathrm{put}^{U^{\times\times}}_R}} = \frac{\frac{\frac{\mathrm{get}^{U^\times}_{L}}{\Delta_{\tilde U}}}{1_{!\tilde U} \mid \mathrm{put}^{U^\times}_R}}{\frac{\frac{1_{!\tilde U}}{\delta_{\tilde U}}}{\mathrm{put}^{U^{\times\times}}_R}} = \frac{\frac{\frac{\mathrm{get}^{U^\times}_{L}}{\Delta_{\tilde U}}}{1_{!\tilde U} \mid \mathrm{put}^{U^\times}_R} \mid \mathrm{id}_{U^\times}}{\frac{\frac{\mathrm{put}^{U^\times}_L \mid \mathrm{get}^{U^\times}_R}{1_I \mid \delta_{\tilde U}}}{1_I \mid \mathrm{put}^{U^{\times\times}}_R}}$$

$$= \frac{\frac{\frac{\mathrm{get}^{U^\times}_{L}}{\Delta_{\tilde U}}}{1_{!\tilde U} \mid \mathrm{put}^{U^\times}_R}}{\frac{\frac{\mathrm{put}^{U^\times}_L}{1_I}}{1_I}} \Bigg| \frac{\mathrm{id}_{U^\times}}{\frac{\frac{\mathrm{get}^{U^\times}_R}{\delta_{\tilde U}}}{\mathrm{put}^{U^{\times\times}}_R}} = \frac{\Delta^\times_U}{1_I} \Bigg| \frac{\mathrm{id}_{U^\times}}{\frac{\frac{\mathrm{get}^{U^\times}_R}{\delta_{\tilde U}}}{\mathrm{put}^{U^{\times\times}}_R}} = \Delta^\times_U \Bigg| \frac{\mathrm{id}_{U^\times}}{\frac{\frac{\mathrm{get}^{U^\times}_R}{\delta_{\tilde U}}}{\mathrm{put}^{U^{\times\times}}_R}},$$

where we used Lemma 4.3.13 for the sixth equality and Lemma 4.3.10 for the second to last equality. Hence, by the uniqueness of the right iteration cell, this implies that

$$\frac{\frac{\mathrm{get}^{U^\times}_{L}}{\delta_{\tilde U}}}{\mathrm{put}^{U^{\times\times}}_R} = (\mathrm{id}_{U^\times})^\times_{\pi_{0,U},\Delta^\times_U,1_I} = \delta^\times_U. \tag{4.3.45}$$

This completes the proof. ❑

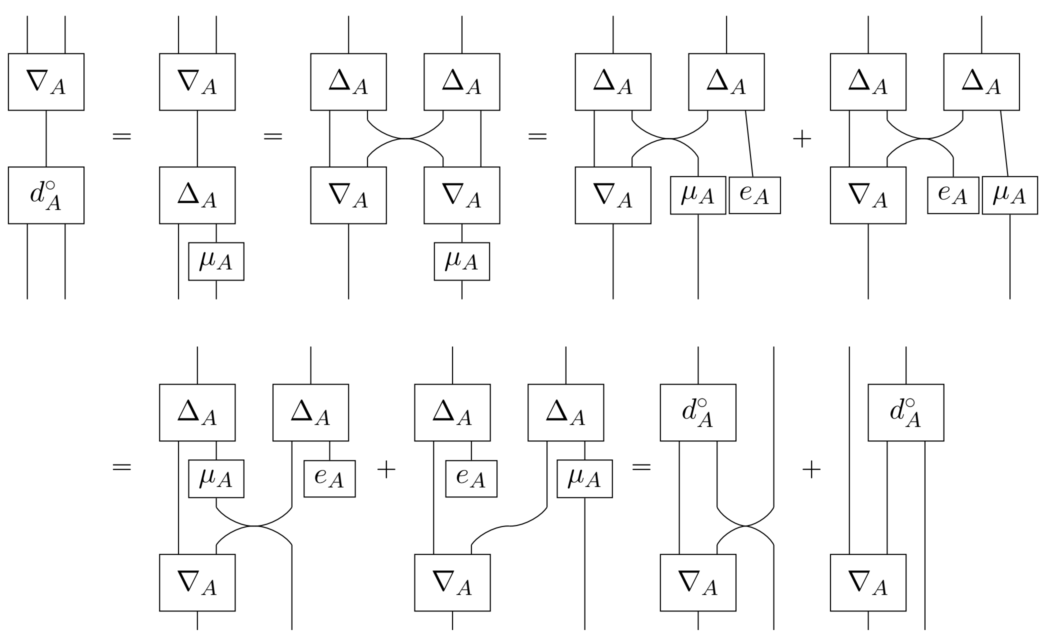
**Lemma 4.3.15**

Let $\mathfrak{C}$ be a differential category. Given an object $A \in \operatorname{Ob}\mathfrak{C}$,

$$\begin{aligned} d^\circ_A \circ \nabla_A = {} & (\nabla_A \otimes 1_A) \circ (1_{!A} \otimes d^\circ_A) \\ & + (\nabla_A \otimes 1_A) \circ \alpha_{!A,!A,A} \circ (1_{!A} \otimes \sigma_{A,!A}) \circ \alpha^{-1}_{!A,A,!A} \circ (d^\circ_A \otimes 1_{!A}). \end{aligned} \tag{4.3.46}$$

♡

***Proof***. Using monoidal graphical calculus, we have that:




(4.3.47)

This completes the proof. ❑

**Lemma 4.3.16. The Right Multiplication Cell is Co-Contraction**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\nabla_U^\times = \frac{\mathrm{get}_L^{U^\times\otimes U^\times}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times}}} \cong (\nabla_{\tilde U})^{\triangleright}. \tag{4.3.48}$$

♡

***Proof***. First, note that

$$\begin{aligned}
\frac{\mathrm{get}_L^{U^\times\otimes U^\times}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times}}} \Big| \pi_{0,U}
&= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times}}} \Big| \pi_{0,U}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times} \mid \pi_{0,U}}}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\ulcorner e_{\tilde U} \lrcorner}}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\ulcorner e_{\tilde U} \circ \nabla_{\tilde U} \lrcorner}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\ulcorner e_{\tilde U} \otimes e_{\tilde U} \lrcorner} \\
&= \frac{\frac{\mathrm{id}_{U^\times} \mid \mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\ulcorner e_{\tilde U} \lrcorner \mid \ulcorner e_{\tilde U} \lrcorner}
= \frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner e_{\tilde U} \lrcorner}} \Big| \frac{\mathrm{get}_L^{U^\times}}{\frac{1_{!\tilde U}}{\ulcorner e_{\tilde U} \lrcorner}}
= \frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner e_{\tilde U} \lrcorner}} \Big| \frac{\mathrm{get}_L^{U^\times}}{\ulcorner e_{\tilde U} \lrcorner}
= \frac{\mathrm{id}_{U^\times}}{\pi_{0,U}} \Big| \pi_{0,U} = \frac{\pi_{0,U}}{\pi_{0,U}},
\end{aligned} \tag{4.3.49}$$

Now, note that

$$\begin{aligned}
\frac{\mathrm{get}_L^{U^\times\otimes U^\times}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times}}} \Big| \pi_{1,U}
&= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times}}} \Big| \pi_{1,U}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times} \mid \pi_{1,U}}}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\ulcorner d^\circ_{\tilde U} \lrcorner}}{\frac{1_{!\tilde U} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\ulcorner d^\circ_{\tilde U} \circ \nabla_{\tilde U} \lrcorner}{\frac{1_{!\tilde U} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}} \\
&= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\ulcorner (\nabla_{\tilde U}\otimes 1_{\tilde U})\circ(1_{!\tilde U}\otimes d^\circ_{\tilde U}) + (\nabla_{\tilde U}\otimes 1_{\tilde U})\circ(1_{!\tilde U}\otimes \sigma_{\tilde U,!\tilde U})\circ(d^\circ_{\tilde U}\otimes 1_{!\tilde U}) \lrcorner}{\frac{1_{!\tilde U} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}} \\
&= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\frac{1_{!\tilde U} \mid \ulcorner d^\circ_{\tilde U} \lrcorner}{\ulcorner \nabla_{\tilde U} \lrcorner \mid 1_{\tilde U}} + \frac{\ulcorner d^\circ_{\tilde U} \lrcorner \mid 1_{!\tilde U}}{\frac{1_{!\tilde U} \mid \ulcorner \sigma_{\tilde U,!\tilde U} \lrcorner}{\ulcorner \nabla_{\tilde U} \lrcorner \mid 1_{\tilde U}}}}{\frac{1_{!\tilde U} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\frac{1_{!\tilde U} \mid \ulcorner d^\circ_{\tilde U} \lrcorner}{\ulcorner \nabla_{\tilde U} \lrcorner \mid 1_{\tilde U}}}{\frac{1_{!\tilde U} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}
+ \frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\frac{\ulcorner d^\circ_{\tilde U} \lrcorner \mid 1_{!\tilde U}}{\frac{1_{!\tilde U} \mid \ulcorner \sigma_{\tilde U,!\tilde U} \lrcorner}{\ulcorner \nabla_{\tilde U} \lrcorner \mid 1_{\tilde U}}}}{\frac{1_{!\tilde U} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}
\end{aligned} \tag{4.3.50}$$

To simplify further, let us simplify each term separately. As we can see,

$$\frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U}}}{\frac{\frac{1_{!\tilde U} \mid \ulcorner d^\circ_{\tilde U} \lrcorner}{\ulcorner \nabla_{\tilde U} \lrcorner \mid 1_{\tilde U}}}{\frac{1_{!\tilde U} \mid \mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}
= \frac{\frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{get}_L^{U^\times}}{1_{!\tilde U}}} \Big| \frac{\mathrm{get}_L^{U^\times}}{\frac{1_{!\tilde U}}{\ulcorner d^\circ_{\tilde U} \lrcorner}}}{\frac{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{1_{!\tilde U}}}{\mathrm{put}_R^{U^\times}} \Big| \frac{\frac{1_{\tilde U}}{\mathrm{put}_R^U}}{\mathrm{id}_{U^\times}}}
= \frac{\frac{\mathrm{id}_{U^\times}}{\mathrm{get}_L^{U^\times}} \Big| \frac{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner d^\circ_{\tilde U} \lrcorner}}{1_{!\tilde U} \mid 1_{\tilde U}}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times}} \Big| \frac{\mathrm{put}_R^U}{\mathrm{id}_{U^\times}}}
= \frac{\frac{\mathrm{id}_{U^\times}}{\mathrm{get}_L^{U^\times}} \Big| \frac{\pi_{1,U} \big| \frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}} \big| \frac{\mathrm{get}_L^U}{1_{\tilde U}}}{1_{!\tilde U} \mid 1_{\tilde U}}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times}} \Big| \frac{\mathrm{put}_R^U}{\mathrm{id}_{U^\times}}}
= \frac{\frac{\mathrm{id}_{U^\times} \mid \pi_{1,U} \big| \frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}} \big| \frac{\mathrm{get}_L^U}{1_{\tilde U}}}{\mathrm{get}_L^{U^\times} \mid 1_{!\tilde U} \mid 1_{\tilde U}}}{\frac{\ulcorner \nabla_{\tilde U} \lrcorner}{\mathrm{put}_R^{U^\times}} \Big| \frac{\mathrm{put}_R^U}{\mathrm{id}_{U^\times}}} \tag{4.3.51}$$

$$
\begin{aligned}
&= \frac{\frac{\pi_{1,U}\,|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}\,|\,\frac{\mathrm{get}_L^U}{1_{\tilde U}}}{\mathrm{get}_L^{U^\times}\,|1_{!\tilde U}|1_{\tilde U}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}\,|\,\frac{\mathrm{put}_R^U}{\mathrm{id}_{U^\times}}}
= \frac{\frac{\pi_{1,U}\,|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\mathrm{get}_L^{U^\times}\,|1_{!\tilde U}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}\,\Big|\,\frac{\frac{\mathrm{get}_L^U}{1_{\tilde U}}}{\frac{1_{\tilde U}}{\frac{\mathrm{put}_R^U}{\mathrm{id}_{U^\times}}}}
= \frac{\frac{\pi_{1,U}\,|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\mathrm{get}_L^{U^\times}\,|1_{!\tilde U}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}\,\Big|\,\frac{\frac{\mathrm{get}_L^U}{\mathrm{put}_R^U}}{\mathrm{id}_{U^\times}}
= \frac{\frac{\pi_{1,U}\,|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\mathrm{get}_L^{U^\times}\,|1_{!\tilde U}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}\,\Big|\,\frac{\mathrm{id}_U}{\mathrm{id}_{U^\times}} \\
&= \frac{\frac{\pi_{1,U}\,|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\mathrm{get}_L^{U^\times}\,|1_{!\tilde U}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}
= \frac{\pi_{1,U}\,\Big|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\mathrm{id}_{U^\times}\,\Big|\,\frac{\mathrm{get}_L^{U^\times}\,|1_{!\tilde U}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}}
= \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}\,\Big|\,\frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^\times}}{\frac{\mathrm{get}_L^{U^\times}\,|1_{!\tilde U}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}}}
= \frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}\,\Big|\,\frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^\times\otimes U^\times}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}},
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{\frac{\mathrm{get}_L^{U^\times}}{\mathrm{get}_L^{U^\times}\,|1_{!\tilde U}}}{\frac{\ulcorner d^\circ_{\tilde U}\urcorner\,|1_{!\tilde U}}{\frac{1_{!\tilde U}\,|\,\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}{\frac{\ulcorner\nabla_{\tilde U}\urcorner\,|1_{\tilde U}}{\frac{1_{!\tilde U}\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}}}
= \frac{\frac{\mathrm{get}_L^{U^\times}}{\frac{\mathrm{get}_L^{U^\times}}{\ulcorner d^\circ_{\tilde U}\urcorner}\,|\,\frac{1_{!\tilde U}}{1_{!\tilde U}}}}{\frac{1_{!\tilde U}\,|\,\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}{\frac{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{1_{!\tilde U}}\,|\,\frac{1_{\tilde U}}{\mathrm{put}_R^U}}{\mathrm{put}_R^{U^\times}}}}
= \frac{\frac{\mathrm{id}_{U^\times}\,|\,\mathrm{get}_L^{U^\times}}{\pi_{1,U}\,|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}\,|\,\frac{\mathrm{get}_L^U}{1_{\tilde U}}\,|\,\frac{1_{!\tilde U}}{1_{!\tilde U}}}}{\frac{1_{!\tilde U}\,|\,\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}{\frac{\ulcorner\nabla_{\tilde U}\urcorner\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}
= \frac{\frac{\mathrm{id}_{U^\times}\,|\,\mathrm{get}_L^{U^\times}}{\pi_{1,U}\,|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}\,|\,\frac{\mathrm{get}_L^U}{1_{\tilde U}}\,|\,\frac{1_{!\tilde U}}{1_{!\tilde U}}}}{\frac{1_{!\tilde U}\,|\,\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}{\frac{\ulcorner\nabla_{\tilde U}\urcorner\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}
= \frac{\frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\mathrm{get}_L^{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}\,|\,\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{1_{\tilde U}|1_{!\tilde U}}}}{\frac{1_{!\tilde U}\,|\,\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}{\frac{\ulcorner\nabla_{\tilde U}\urcorner\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}} \\
&= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{get}_L^{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}\,|\,\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{1_{\tilde U}|1_{!\tilde U}}}}{\frac{1_{!\tilde U}\,|\,\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}{\frac{\ulcorner\nabla_{\tilde U}\urcorner\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}}
= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{get}_L^{U^\times}}{\frac{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{1_{!\tilde U}}\,|\,\frac{\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{1_{\tilde U}|1_{!\tilde U}}}{\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}
= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{get}_L^{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}\,|\,\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}} \\
&= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{id}_{U^\times}\,|\,\mathrm{get}_L^{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}\,|\,\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U^\times}}}
= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}\,\Big|\,\frac{\mathrm{get}_L^{U^\times}}{\frac{\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}}{1_{!\tilde U}\,|\,\mathrm{put}_R^U}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}
= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}\,\Big|\,\frac{\mathrm{get}_L^{U^\times}}{\frac{\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}}{1_{!\tilde U}\,|\,\mathrm{put}_R^U}}}{\frac{1_{!\tilde U}|1_{!\tilde U}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}} \qquad (4.3.52) \\
&= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}\,\Big|\,\frac{\mathrm{get}_L^{U^\times}}{\frac{\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}}{1_{!\tilde U}\,|\,\mathrm{put}_R^U}}\,\Big|\mathrm{id}_U}{\frac{1_{!\tilde U}\,|\,\mathrm{put}_R^{U^\times}\,|\,\mathrm{get}_L^{U^\times}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}}
= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{id}_U}{\frac{\mathrm{get}_L^{U^\times}}{1_{!\tilde U}}}}\,\Big|\,\frac{\frac{\frac{\mathrm{get}_L^{U^\times}}{\frac{\mathrm{get}_L^U\,|1_{!\tilde U}}{\ulcorner\sigma_{\tilde U,!\tilde U}\urcorner}}}{1_{!\tilde U}\,|\,\mathrm{put}_R^U}}{\mathrm{put}_R^{U^\times}}\,\Big|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}} \\
&= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}\,\Big|\sigma_{U^\times,U}\Big|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}
= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\,\Big|\,\frac{\frac{\mathrm{id}_{U^\times}}{\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}\,\Big|\,\frac{\sigma_{U^\times,U}\,|\,\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{1_{!\tilde U}}}{\frac{\ulcorner\nabla_{\tilde U}\urcorner}{\mathrm{put}_R^{U^\times}}}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}} \Big| \cfrac{\cfrac{\frac{\mathrm{id}_{U^\times}}{\mathrm{id}_U}|\sigma_{U^\times,U}|\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\mathrm{get}_L^{U^\times}|1_{!\tilde{U}}}}{\cfrac{\ulcorner\nabla_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}} = \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}} \Big| \cfrac{\cfrac{\sigma_{U^\times,U}|\frac{\mathrm{id}_U}{\mathrm{get}_L^{U^\times}}}{\mathrm{id}_{U^\times}|\mathrm{get}_L^{U^\times}|1_{!\tilde{U}}}}{\cfrac{\ulcorner\nabla_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}} \\
&= \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}} \Big| \frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}} \Big| \cfrac{\mathrm{id}_U}{\cfrac{\mathrm{get}_L^{U^\times}}{\cfrac{\mathrm{get}_L^{U^\times}|1_{!\tilde{U}}}{\cfrac{\ulcorner\nabla_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}}}} = \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}} \Big| \frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}} \Big| \cfrac{\mathrm{id}_U}{\cfrac{\mathrm{get}_L^{U^\times\otimes U^\times}}{\cfrac{\ulcorner\nabla_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}}}.
\end{aligned}
$$

Putting the last three expressions together, we get that

$$
\cfrac{\mathrm{get}_L^{U^\times\otimes U^\times}}{\cfrac{\ulcorner\nabla_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}}\Big|\pi_{1,U} = \left(\frac{\pi_{1,U}}{\mathrm{id}_{U^\times}} + \frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}\Big|\frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}}\right)\Big| \cfrac{\mathrm{id}_U}{\cfrac{\mathrm{get}_L^{U^\times\otimes U^\times}}{\cfrac{\ulcorner\nabla_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}}}. \tag{4.3.53}
$$

Hence, by the uniqueness of the right iteration cell,

$$
\cfrac{\mathrm{get}_L^{U^\times\otimes U^\times}}{\cfrac{\ulcorner\nabla_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}} = (\mathrm{id}_U)^\times_{\frac{\pi_{0,U}}{\pi_{0,U}},\frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}+\frac{\mathrm{id}_{U^\times}}{\pi_{1,U}}|\frac{\sigma_{U^\times,U}}{\mathrm{id}_{U^\times}},1_I} = \nabla_U^\times. \tag{4.3.54}
$$

This completes the proof. ❏

**Lemma 4.3.17**

Let $\mathfrak{C}$ be a differential category. Given an object $A \in \mathrm{Ob}(\mathfrak{C})$,

$$
d_A^\circ \circ u_A = 0. \tag{4.3.55}
$$

♡

***Proof.*** As we can see,

$$
\begin{aligned}
d_A^\circ \circ u_A &= (1_{!A}\otimes\mu_A)\circ\Delta_A\circ u_A = (1_{!A}\otimes\mu_A)\circ(u_A\otimes u_A) = (1_{!A}\circ u_A)\otimes(\mu_A\circ u_A) \\
&= (1_{!A}\circ u_A)\otimes(\mu_A\circ u_A) = (1_{!A}\circ u_A)\otimes 0 = 0,
\end{aligned} \tag{4.3.56}
$$

where we used Lemma 3.3 in [2] for the second to last equality. ❏

**Lemma 4.3.18. The Right Unit Cell is Co-Weakening**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$
u_U^\times = \frac{\ulcorner u_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}} \cong (u_{\tilde{U}})^\triangleright. \tag{4.3.57}
$$

♡

***Proof.*** As we can see,

$$
\frac{\ulcorner u_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}}\Big|\pi_{0,U} = \frac{\ulcorner u_{\tilde{U}}\lrcorner}{\mathrm{put}_R^{U^\times}|\pi_{0,U}} = \frac{\ulcorner u_{\tilde{U}}\lrcorner}{\ulcorner e_U\lrcorner} = \ulcorner e_U\circ u_{\tilde{U}}\lrcorner = \ulcorner 1_I\lrcorner = 1_I, \tag{4.3.58}
$$

and

$$\frac{\ulcorner u_{\tilde{U}} \lrcorner}{\mathrm{put}_R^{U\times}}|\pi_{1,U} = \frac{\ulcorner u_{\tilde{U}} \lrcorner}{\mathrm{put}_R^{U\times}\,|\pi_{1,U}} = \frac{\frac{\ulcorner u_{\tilde{U}} \lrcorner}{\ulcorner d^\circ_{\tilde{U}} \lrcorner}}{\frac{1_{!\tilde{U}}\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U\times}}} = \frac{\ulcorner d^\circ_{\tilde{U}} \circ u_{\tilde{U}} \lrcorner}{\frac{1_{!\tilde{U}}\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U\times}}} = \frac{\ulcorner 0 \lrcorner}{\frac{1_{!\tilde{U}}\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U\times}}}$$
$$= \frac{0}{\frac{1_{!\tilde{U}}\,|\,\mathrm{put}_R^U}{\mathrm{put}_R^{U\times}}} = 0 = \frac{\frac{\ulcorner 0 \lrcorner}{\frac{\ulcorner u_{\tilde{U}} \lrcorner}{\mathrm{put}_R^{U\times}}}|1_I}{1_I}, \tag{4.3.59}$$

so by the uniqueness of the right iteration cell, this implies that

$$u_U^\times = (0)^\times_{1_I,1_I,1_I} = \frac{\ulcorner u_{\tilde{U}} \lrcorner}{\mathrm{put}_R^{U\times}}. \tag{4.3.60}$$

This completes the proof. ❏

**Lemma 4.3.19**

Let $\mathfrak{C}$ be a differential category. Given an object $A \in \mathrm{Ob}(\mathfrak{C})$,

$$d^\circ_A \circ \nu_A = (u_A \otimes 1_A) \circ \lambda_A. \tag{4.3.61}$$

♡

***Proof***. As we can see,

$$\begin{aligned}
d^\circ_A \circ \nu_A &= (1_{!A} \otimes \mu_A) \circ \Delta_A \circ \nu_A \\
&= (1_{!A} \otimes \mu_A) \circ (\nu_A \circ u_A) \circ \rho_A + (1_{!A} \otimes \mu_A) \circ (u_A \circ \nu_A) \circ \lambda_A \\
&= ((1_{!A} \circ \nu_A) \otimes (\mu_A \circ u_A)) \circ \rho_A + ((1_{!A} \circ u_A) \otimes (\mu_A \circ \nu_A)) \circ \lambda_A \\
&= (\nu_A \otimes 0) \circ \rho_A + (u_A \otimes 1_A) \circ \lambda_A \\
&= 0 + (u_A \otimes 1_A) \circ \lambda_A \\
&= (u_A \otimes 1_A) \circ \lambda_A,
\end{aligned} \tag{4.3.62}$$

where we used Lemma 3.3 in [2] for the fourth equality. ❏

**Lemma 4.3.20. The Right Plastic Cell is Co-Dereliction**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in \mathfrak{C}_*^{\triangleright\blacktriangleleft}$,

$$\nu_U^\times = \frac{\frac{\mathrm{get}_L^U}{\ulcorner \nu_{\tilde{U}} \lrcorner}}{\mathrm{put}_R^{U\times}} \cong (\nu_{\tilde{U}})^{\triangleright}. \tag{4.3.63}$$

♡

***Proof***. As we can see,

$$\frac{\frac{\mathrm{get}_L^U}{\ulcorner \nu_{\tilde{U}} \lrcorner}}{\mathrm{put}_R^{U\times}}|\pi_{0,U} = \frac{\frac{\mathrm{get}_L^U}{\ulcorner \nu_{\tilde{U}} \lrcorner}}{\mathrm{put}_R^{U\times}}\Big|\frac{1_I}{\pi_{0,U}} = \frac{\frac{\mathrm{get}_L^U}{\ulcorner \nu_{\tilde{U}} \lrcorner}|1_I}{\mathrm{put}_R^{U\times}\,|\pi_{0,U}} = \frac{\frac{\mathrm{get}_L^U}{\ulcorner \nu_{\tilde{U}} \lrcorner}}{\ulcorner e_{\tilde{U}} \lrcorner} = \frac{\mathrm{get}_L^U}{\ulcorner e_{\tilde{U}} \circ \nu_{\tilde{U}} \lrcorner} = \frac{\mathrm{get}_L^U}{\ulcorner 0 \lrcorner} = \frac{\mathrm{get}_L^U}{0} = 0, \tag{4.3.64}$$

where we used the constant rule of $\nu$ for the fifth equality, and

$$\begin{aligned}\frac{\frac{\mathrm{get}_L^U}{\nu_{\tilde U}}}{\mathrm{put}_R^{U\times}}\Big|\pi_{1,U} &= \frac{\frac{\mathrm{get}_L^U}{\nu_{\tilde U}}}{\mathrm{put}_R^{U\times}}\Big|\frac{1_I}{\pi_{1,U}} = \frac{\frac{\mathrm{get}_L^U}{\nu_{\tilde U}}\big|1_I}{\mathrm{put}_R^{U\times}\,|\pi_{1,U}} = \frac{\frac{\mathrm{get}_L^U}{\nu_{\tilde U}}}{\frac{d^\circ_{\tilde U}}{\frac{1_{!\tilde U}|\mathrm{put}_R^U}{\mathrm{put}_R^{U\times}}}} = \frac{\frac{\mathrm{get}_L^U}{d^\circ_{\tilde U}\circ\nu_{\tilde U}}}{\frac{1_{!\tilde U}|\mathrm{put}_R^U}{\mathrm{put}_R^{U\times}}} = \frac{\frac{\mathrm{get}_L^U}{u_{\tilde U}\otimes 1_{\tilde U}}}{\frac{1_{!\tilde U}|\mathrm{put}_R^U}{\mathrm{put}_R^{U\times}}}\\ &= \frac{\frac{\mathrm{id}_U|\mathrm{get}_L^U}{u_{\tilde U}|1_{\tilde U}}}{\frac{1_{!\tilde U}|\mathrm{put}_R^U}{\mathrm{put}_R^{U\times}}} = \frac{\frac{\frac{\mathrm{id}_U}{u_{\tilde U}}}{1_{!\tilde U}}\Big|\frac{\frac{\mathrm{get}_L^U}{1_{\tilde U}}}{\mathrm{put}_R^U}}{\mathrm{put}_R^{U\times}} = \frac{\frac{\frac{\mathrm{id}_U}{1_I}}{u_{\tilde U}}\Big|\frac{\frac{\mathrm{get}_L^U}{\mathrm{put}_R^U}}{1_I}}{\mathrm{put}_R^{U\times}} = \frac{\frac{\mathrm{id}_U}{u_{\tilde U}}\Big|\frac{\mathrm{id}_U}{1_I}}{\mathrm{put}_R^{U\times}} = \frac{\frac{\mathrm{id}_U|\mathrm{id}_U}{u_{\tilde U}|1_I}}{\mathrm{put}_R^{U\times}} = \frac{\frac{\mathrm{id}_U}{u_{\tilde U}}}{\mathrm{put}_R^{U\times}} = \frac{\mathrm{id}_U}{u_U^\times},\end{aligned} \tag{4.3.65}$$

where we used Lemma 4.3.19 for the fifth equality and 4.3.18 for the last equality. Hence, by the uniqueness of the right choice cell,

$$\frac{\frac{\mathrm{get}_L^U}{\nu_{\tilde U}}}{\mathrm{put}_R^{U\times}} = 0 \boxtimes \frac{\mathrm{id}_U}{u_U^\times} = \nu_U^\times. \tag{4.3.66}$$

This completes the proof. ❏

We, now, have everything we need to prove the main result of this thesis.

**Theorem 4.3.21**

Let $\mathfrak{C}$ be a resource category. Then, $\mathbf{H}\overset{\ulcorner\urcorner}{\underset{\llcorner\lrcorner}{\mathfrak{C}}}$ is a differential category. ♡

***Proof*.** By Theorem 4.2.20, $\mathbf{H}\overset{\ulcorner\urcorner}{\underset{\llcorner\lrcorner}{\mathfrak{C}}}$ is a compact closed category, so by Theorem 4.1.6, $\mathbf{H}\overset{\ulcorner\urcorner}{\underset{\llcorner\lrcorner}{\mathfrak{C}}}$ is a $*$-autonomous category. By Lemmas 4.3.4-4.3.9, $\Delta^\times$, $e^\times$, $\mu^\times$, $\delta^\times$, $\nabla^\times$, $u^\times$, $\nu^\times$ are natural transformations. Since $(!, \delta, \mu, \Delta, e, \nabla, u)$ is a bialgebra and $\nu$ is a co-dereliction corresponding to this bialgebra, by Lemmas 4.3.10-4.3.20, $((-)^\times, \delta^\times, \mu^\times, \Delta^\times, e^\times, \nabla^\times, u^\times)$ is a bialgebra and $\nu^\times$ is a co-dereliction corresponding to this bialgebra. Hence, $\mathbf{H}\overset{\ulcorner\urcorner}{\underset{\llcorner\lrcorner}{\mathfrak{C}}}$ is a differential category. ❏

**Corollary 4.3.22. The Right Step Cell is a Co-Deriving Transform**

Let $\mathfrak{C}$ be a resource theory. Given a $\mathfrak{C}$-valued exchange $U \in (\mathfrak{C})_*^{\triangleright\blacktriangleleft}$,

$$\pi_{1,U} = \Delta_U^\times\Big|\frac{\mu_U^\times}{\mathrm{id}_{U^\times}}. \tag{4.3.67}$$

♡

***Proof*.** As we can see,

$$\begin{aligned}\Delta_U^\times\Big|\frac{\mu_U^\times}{\mathrm{id}_{U^\times}} &= \Delta_U^\times\Big|\frac{\pi_{1,U}\big|\frac{\mathrm{id}_U}{\pi_{0,U}}}{\mathrm{id}_{U^\times}} = \Delta_U^\times\Big|\frac{\pi_{1,U}\big|\frac{\mathrm{id}_U}{\pi_{0,U}}}{\mathrm{id}_{U^\times}|\mathrm{id}_{U^\times}} = \Delta_U^\times\Big|\frac{\pi_{1,U}}{\mathrm{id}_{U^\times}}\Big|\frac{\frac{\mathrm{id}_U}{\pi_{0,U}}}{\mathrm{id}_{U^\times}} = \pi_{1,U}\Big|\frac{\mathrm{id}_U}{\Delta_U^\times}\Big|\frac{\mathrm{id}_U}{\frac{\pi_{0,U}}{\mathrm{id}_{U^\times}}}\\ &= \pi_{1,U}\Big|\frac{\mathrm{id}_U|\mathrm{id}_U}{\Delta_U^\times\big|\frac{\pi_{0,U}}{\mathrm{id}_{U^\times}}} = \pi_{1,U}\Big|\frac{\mathrm{id}_U}{\mathrm{id}_{U^\times}} = \pi_{1,U}.\end{aligned} \tag{4.3.68}$$

This completes the proof. ❏

# Bibliography


[1] R. F. Blute, J.R.B. Cockett, and R.A.G. Seely. Differential categories. *Mathematical Structures in Computer Science*, 16(6):1049–1083, 2006.

[2] R.F. Blute, J.R.B. Cockett, J.-S.P. Lemay, and R.A.G. Seely. Differential categories revisited. *Applied Categorical Structures*, 28(2):171–235, Apr 2020.

[3] R.F. Blute, J.R.B. Cockett, R.A.G. Seely, and T.H. Trimble. Natural deduction and coherence for weakly distributive categories. *Journal of Pure and Applied Algebra*, 113(3):229–296, 1996.

[4] J. R. B. Cockett and J.-S. Lemay. Integral categories and calculus categories. *Mathematical Structures in Computer Science*, 29(2):243–308, 2019.

[5] J.R.B. Cockett, Cole Comfort, and Priyaa Srinivasan. Dagger linear logic for categorical quantum mechanics. *Logical Methods in Computer Science*, Volume 17, Issue 4, Nov 2021.

[6] J.R.B. Cockett and Craig Pastro. The logic of message-passing. *Science of Computer Programming*, 74(8):498–533, 2009. Special Issue on Mathematics of Program Construction (MPC 2006).

[7] T. Ehrhard and L. Regnier. Differential interaction nets. *Theoretical Computer Science*, 364(2):166–195, 2006. Logic, Language, Information and Computation.

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

[9] Thomas Ehrhard and Laurent Regnier. Uniformity and the taylor expansion of ordinary lambda-terms. *Theoretical Computer Science*, 403(2):347–372, 2008.

[10] Gerhard Gentzen. Untersuchungen über das logische schließen. i. *Mathematische Zeitschrift*, 39(1):176–210, Dec 1935.

[11] J.-Y. Girard. Linear logic: its syntax and semantics. In Jean-Yves Girard, Yves Lafont, and Laurent Regnier, editors, *Advances in Linear Logic*, London Mathematical Society Lecture Note Series, page 1–42. Cambridge University Press, 1995.

[12] Chris Heunen and Jamie Vicary. *Categories for Quantum Theory: An Introduction.* Oxford University Press, 11 2019.

[13] Sørensen Morten Heine and Urzyczyn Pawel. *Lectures on the Curry-Howard Isomorphism.* Number v. 149 in Studies in Logic and the Foundations of Mathematics. Elsevier Science, 2006.

[14] Chad Nester. Situated transition systems. In Kohei Kishida, editor, Proceedings of the Fourth International Conference on *Applied Category Theory,* Cambridge, United Kingdom, 12-16th July 2021, volume 372 of *Electronic Proceedings in Theoretical Computer Science*, pages 103–115. Open Publishing Association, 2022.

[15] Chad Nester. Concurrent process histories and resource transducers. *Logical Methods in Computer Science*, Volume 19, Issue 1, Jan 2023.

[16] Chad Nester and Niels Voorneveld. Protocol choice and iteration for the free cornering. *Journal of Logical and Algebraic Methods in Programming*, 137:100942, 2024.

[17] Robert R. Schneck. Natural deduction and coherence for non-symmetric linearly distributive categories. *Theory and Applications of Categories [electronic only]*, 6:105–146, 1999.

[18] R.A.G. Seely. Linear logic, *-autonomous categories and cofree coalgebras. In John W. Gray and Andre Scedrov, editors, *Categories in Computer Science and Logic*, pages 371–383. American Mathematical Society, 1989.