Université Sorbonne Paris Nord
École doctorale sciences, technologies, santé - galilée

Indexing Differential Linear Logic by Differential Operators

Indexer la logique linéaire par des opérateurs différentiels

Thèse de doctorat
présentée par

Simon Mirwasser
Laboratoire d'Informatique de Paris Nord

Pour obtenir le grade de
Docteur en Informatique

soutenue à Villetaneuse le 19 mars 2026
devant le jury composé de :

Giulio Guerrieri	Rapporteur	Associate Professor, University of Sussex
Laurent Regnier	Rapporteur	Professeur, Université d'Aix-Marseille
Claudia Faggian	Examinatrice	Chargée de rcherche, CNRS
Michele Pagani	Examinateur	Professeur, ENS de Lyon
Stefano Guerrini	Directeur de thèse	Professeur, Université Sorbonne Paris Nord
Marie Kerjean	Co-encadrante de thèse	Chargée de rcherche, CNRS

Contents

Résumé en français

Titre : Indexer la logique linéaire par des opérateurs différentiels

Mots clés : Logique Linéaire - Théorie de la Démonstration - Sémantique Dénotationnelle - Équations Différentielles

Résumé : La correspondance de Curry-Howard-Lambek met en évidence un lien structurel profond entre preuves et programmes. Dans ce cadre, la logique linéaire, puis la logique linéaire différentielle mettent en relation la notion de ressource en informatique avec la notion de linéarité entre espaces vectoriels en mathématiques. Cette relation a été étendue, afin de proposer des modèles mathématiques toujours plus riches, afin de représenter des notions toujours plus complexes.

Cette thèse étend ces notions afin de proposer un système logique dans lequel on peut appliquer et résoudre des équations différentielles aux dérivées partielles. Pour ce faire, elle se base sur un modèle mathématique de la logique linéaire différentielle qui utilise des fonctions lisses et des distributions pour représenter les programmes non linéaires. Ce modèle permet une étude approfondie des opérateurs différentiels linéaires aux dérivées partielles à coefficients constants, et de leurs solutions.

Afin d'étudier un système de preuve avec de tels opérateurs, nous nous basons sur la logique linéaire graduée, qui étend la logique linéaire en un système paramétré, ce qui permet une étude plus précise des programmes qui sont représentés. Nous proposons une extension de ce système, qui permet d'interpréter l'opération de différentiation en utilisant des opérateurs différentiels en tant que paramètres. Nous définissons un modèle de cette extension qui utilise les fonctions lisses et les distributions, dans lequel les règles logiques sont interprétées par la résolution d'équations différentielles ou l'application d'opérateurs différentiels.

Le système précédent est cependant limité au premier ordre, à cause de la nature des dérivées partielles, qui ont besoins d'espaces ayant une base canonique pour être interprétées. Nous proposons donc un deuxième système dans lequel nous pouvons manipuler des objets d'ordre supérieur. Pour ce faire, nous nous basons sur des considérations sémantiques, qui proviennent de l'étude de la transformée de Laplace en logique linéaire différentielle. Guidés par ces intuitions, nous raffinons la syntaxe de notre premier système et définissons une logique que nous interprétons dans les espaces de Köthe, qui sont un modèle classique de la logique linéaire différentielle auquel nous pouvons appliquer des opérateurs différentiels même à l'ordre supérieur.

Nous définissons enfin un dernier système, qui permet lui aussi d'étudier des objets d'ordre supérieur mais sans transformée de Laplace. Ce système utilise un ensemble général de paramètres, et nous étudions les propriétés que cet ensemble doit respecter afin d'obtenir un comportement dynamique intéressant.

Abstract

Title: Indexed differential linear logic by differential operators

Keywords: Linear Logic - Proof Theory - Denotational Semantics - Differential equations

Abstract: The Curry-Howard-Lambek correspondence highlights a deep structural link between proofs and programs. In this context, linear logic, followed by differential linear logic, links the concept of resources in computer science with the concept of linearity between vector spaces in mathematics. This relationship has been extended to offer increasingly rich mathematical models that can represent increasingly complex concepts.

This thesis extends these concepts to propose a logical system in which partial differential equations can be applied and solved. To do so, it is based on a mathematical model of linear differential logic that uses smooth functions and distributions to represent non-linear programs. This model allows for an in-depth study of linear partial differential operators with constant coefficients and their solutions.

In order to study a proof system with such operators, we rely on graded linear logic, which extends linear logic to a parameterized system, allowing for a more precise study of the programs that are represented. We propose an extension of this system, which allows the differentiation operation to be interpreted using differential operators as parameters. We define a model of this extension that uses smooth functions and distributions, in which logical rules are interpreted by solving differential equations or applying differential operators.

However, the previous system is limited to first order, due to the nature of partial derivatives, which require spaces with a canonical basis in order to be interpreted. We therefore propose a second system in which we can manipulate higher-order objects. To do this, we rely on semantic considerations, which come from the study of the Laplace transform in differential linear logic. Guided by these intuitions, we refine the syntax of our first system and define a logic that we interpret in Köthe spaces, which are a classical model of differential linear logic to which we can apply differential operators even at higher orders.

Finally, we define a last system, which also allows us to study higher-order objects but without Laplace transforms. This system uses a general set of parameters, and we study the properties that this set must satisfy in order to obtain interesting dynamic behavior.

Introduction

The work carried out in this thesis builds upon several decades of significant advances in programming language theory. In this introduction, we review some of these advances in order to justify the directions of research taken during this thesis. We then describe the contributions presented in this manuscript.

Semantics of programming languages

The motivations behind the semantics of programming languages stem from the fact that a program must be understood by a machine in order to function. It is then written in a language whose formalism is designed for computers and not necessarily for humans. Such a language may be far away from the *natural language*. But while a program needs to be understood by a machine, one would like to ensure that the program is doing what it has been designed for. This task may be hard to perform because of the nature of a programming language. Semantics of programming languages appear to tackle this issue. There are several notions of semantics that are used to study programs and their computations. The earliest manifestation is operational semantics [Flo67]. It consists of expressing logical propositions of a program and proving in a logical framework that these propositions are preserved by the operations made by the program during its execution. Another notion of semantics, which is at the core of this thesis, is the denotational semantics [Sco77]. When one defines a denotational semantics for a program, it associates a mathematical value to each instruction of this program, called *denotations*. Usually, a program is then interpreted by a mathematical function in the context of the denotational semantics. It is useful, since it allows forgetting about the syntax of the program when reasoning about it. In addition, one may use well known mathematical spaces for the denotations of the program. It makes the program easier to study, and a lot of properties of such spaces can easily be used to deduce nice results on the program. To make all of this more concrete, let us take the example of the λ-calculus.

The λ-calculus is a theoretical programming language that was introduced by Church in the 1930s [Chu32] while studying the foundations of mathematics. The strength is this language is that, despite being extremely simple to describe, it is *Turing complete*, meaning that it is as powerful as any other programming language: any computable program can be encoded in the λ-calculus. This makes it one of the tools favored by logicians for studying general properties of programming languages. The λ-calculus consists of a grammar that defines the *terms* (or λ-terms) that are the programs of the calculus, and a rewriting rule over the terms that allow to execute the terms in order to perform a computation. The grammar states that a term is either a variable, a term applied to another term, or the *abstraction* of a term t depending on a parameter variable x. This notion of abstraction corresponds to the notion of function: the abstracted term should be seen as the function that maps the parameter variable x to the term t. On this notion of term, one defines a rule called the β-reduction, which allows reducing a term into another. In the 1970s, Scott defined a denotational model for the λ-calculus [Sco82]. The difficulty of providing such a model comes from the fact that in λ-calculus each term should be understood as a function. If $D \to D$ is the space where the terms are interpreted, a term can also be an argument of such functions, and then be interpreted in D. However, it is not possible to provide a non-trivial set D isomorphic to $D \to D$. Scott's solution was to consider domains and continuous functions over them, which is, in fact, a model of λ-calculus. This model has been crucial in the history of denotational semantics.

The Curry-Howard-Lambek correspondence

The idea we started to develop above while introducing the denotational semantics is the strong relation between the world of computer science and the one of mathematics or logic. This relation is highlighted by the Curry-Howard-Lambek correspondence [Cur34, Ros14,How69]. Originally called the Curry-Howard correspondence, or the proof/program correspondence, this notion builds bridges between logical proofs and computer programs. This correspondence states that the notion of proof and program are equivalent. It also extends to types and formulas: a proof of a formula is equivalent to a program, and the type of this program is equivalent to the proven formula. One possible framework to observe this correspondence is the simply typed λ-calculus. We have described before the untyped λ-calculus, but it is possible to add a notion of typing on the λ-terms. This typing comes with a grammar and a set of rules, which allow typing the terms. In this framework, one can inductively define an isomorphism between a typed term and a proof in natural deduction of a formula in minimal logic.

This idea of bridges between logical concepts and computer science ones has been extended to many well-known notions. One that will be important in our work is the link between the β-reduction of a term and the cut elimination of a proof. In logic, the cut elimination is a rewriting procedure on proofs where the goal is to eliminate from the proof the occurrences of the cut rule, which is the rule that allows to *apply* a proof to another. It is possible to extend the isomorphism from simply typed λ-calculus to intuitionistic minimal logic in order to take into account the β-reduction. The isomorphism will then turn the β-reduction between two terms into a step of the cut elimination procedure between the two corresponding proofs. This gives to the logical proofs a *computational content*. The cut elimination procedure corresponds to a computation over a proof, and the result of this procedure to the result of the computation. In a sense, proofs can compute. Another classic example is the one of the `call/cc` operator which allows to save the memory of the environment. Griffin has proved that having such an operator in a programming language is equivalent at the logical side of being able to prove the Pierce law, which is known for being deductible if and only if the logic is classical [Gri90].

The Curry-Howard correspondence that relates logic and computer science can be extended to category theory. This was noticed by Lambek in the 1970s. A type or a formula can be seen as an object of a category, and then a program of input A and output B can be seen as a morphism between the interpretation of A and the one of B in this category. The β-reduction (or the cut elimination) corresponds then to the equality of morphisms. Following this point of view, the isomorphism between simply typed λ-calculus and intuitionistic minimal logic has a categorical extension, which is the notion of Cartesian closed category.

Quantitative semantics and mathematical models

Going back to the semantics defined by Scott for λ-calculus, some other approaches have been tried to define a denotational model of λ-calculus. One was proposed by Girard in the 1980s, using the notion of normal functors [Gir88a]. Most denotational models allow to deduce *qualitative* properties of the programs they represent, like which output will be produced depending on the input. But with the model of normal functors, some *quantitative* properties can be proved on programs, such as properties on the computation time, the amount of time a resource is used by the program or even probability results on the output. This approach with normal functors is based on the idea of using power series to interpret the programs, whose coefficients are sets. This idea has been generalized in many directions, for many purposes. In particular, focusing on the type of elements used as coefficients of the power series may lead to highly different models.

The study of normal functors by Girard led him to define the so-called linear logic [Gir87]. This framework formalizes the study of the use of resources by a program. It can be model thanks to linear algebra. The programs consuming exactly once their input are represented by linear maps between vector spaces and the others may use several times their input, and they may be model using analytic functions. They are functions equal to their Taylor series, and can then be represented thanks to power series. These ideas have been formalized by Ehrhard, using two different frameworks: Köthe sequence spaces [Ehr02] and finiteness spaces [Ehr05]. Other quantitative models of linear logic use elements of a semiring as coefficients (see Lamarche [Lam92] and Laird *et al.* [Lai16, LMMP13]).

Some other possible aspects of program analysis are well suited to be modeled in the context of quantitative semantics. Probabilities in the context of quantum computing can be taken into account through completely positive maps, as shown by Pagani *et al.* [PSV14]. For probabilistic programming, Danos *et al.* have extended the notion of coherent spaces to probabilities [DE11]. One final idea of quantitativity that should be highlighted in this thesis is about differentiation, which has been very fruitful in the field of semantics. We describe some of these results in the next section.

From discreteness to continuity via the differentiation

The notion of discreteness is, by nature, at the core of the study of programming languages. The machine computes discrete operations, and, moreover, each data is finitely represented in a computer. This intuitively clashes with the models used in quantitative semantics. An analytic function can be infinitely differentiated, but the notion of differentiation is strongly related to the continuous facet of functions. However, being able to manipulate continuous objects with a computer has a lot of interesting applications. Following the work of Ehrhard on vectorial models of linear logic, Ehrhard and Regnier have defined and studied the *differential λ-calculus* [ER03]. In this setting, the semantics have inspired the syntax into defining an operator of differentiation as a primitive operation on the programs. This operation allows computing the Taylor expansion of a term. In a sense, one can see that these kinds of constructions push the boundaries of the Curry-Howard-Lambek correspondence into more and more interesting mathematical frameworks. The work on differential λ-caculus allows to consider a program as a sum of polynomials. These polynomials are quite simple to study, and this idea simplifies a lot program analysis: considering a complex term t, one can approach it thanks to its Taylor expansion and deduced many properties from its approximants, which are easier to study.

Another line of research also initiated by Ehrhard and Regnier has been the definition of *differential linear logic* [ER06]. In this context, the notion of differentiation does not occur at the level of the program, but at the level of the type. In linear logic, the formulas using exactly once their hypotheses are linear, but it is also possible to consider non-linear formulas, which may use several times their hypothesis. Differential linear logic extends this setting by adding the possibility to *linearize* a formula. To perform this linearization, the proof system has to be endowed with the operation of differentiation: the differential of a proof is its linear approximation. Studying this system, Ehrhard and Regnier have noticed that this differentiation had to respect several properties. Elegantly, those correspond to the syntactical counterpart of well-known axioms of differential calculus. In particular, the syntax of differential linear logic has to fulfill the Leibniz rule (or product rule), and the chain rule about differentiation of the composition of functions. Semantically, the model of linear logic made of Köthe sequences is already a model of differential linear logic. It is the case, since one can define an operation of differentiation, which respects usual rules of differential calculus. Other models of differential linear logic, such as finiteness spaces [Ehr05], have been studied, closing more and more the gap between logic and functional analysis, and, by extension, between discreteness and continuity in theoretical computer science.

Even if the framework of Köthe sequence spaces is well suited to perform differentiation, this operation is strongly based on the countable basis of these spaces. The differential in this context is possible by using linear decomposition over these bases. But in many mathematical spaces this is not possible to do such a decomposition. Blute *et al.* have used convenient spaces to provide a new model of differential linear logic [BET12]. This setting is based on the notion of bornology, which is a notion orthogonal to topology. Bornologies, abstracting the notion of bounded sets, allows for a Cartesian closed category of smooth functions While the notion of differentiation is in a sense more natural here, this model is intuitionistic and then does not fully reflect differential linear logic which is a classical logic. A solution to this point has been provided by Kerjean using distribution theory and nuclear Fréchet spaces [Ker18a]. Here, the control comes from topological properties of distributions. This theory enjoys a natural definition of differentiation as well, and the spaces considered are reflexive. Reflexive spaces in the semantics are the interpretation of the classical setting of the logic. In addition, this model led Kerjean to study an extension of differential linear logic. Distribution theory is the favored mathematical setting to study linear partial differential equations [Hor63]. Adapting her model with distributions to take into account these differential equations, Kerjean extended the syntax of differential linear logic by indexing some connectives by a differential operator [Ker18a]. This gives an elegant theory where the notion of linearity from linear logic is stretched to provide a theory where solutions and parameters of differential equations can interact. The starting point of this thesis was to study this theory in detail in order to generalize it both at a syntactical and a semantical level.

Contribution of the thesis and outline

In this thesis, we extend the work by Kerjean of the logic D-DiLL [Ker18a] which adds the notion of partial differential operators in differential linear logic. Technically, this addition was made in the syntax by indexing the exponential connectives of linear logic by a differential operator. These connectives allow the interpretation of the usual non-linear implication in linear logic. In Kerjean's model with distributions, these were the formulas interpreted by either distributions or by their dual which is the space of smooth functions. Using the fact that the space of solutions of a linear partial differential operator is a space of distributions, the indexation gives the possibility to use this space of solutions as the interpretation of some formula. However, some pieces were missing in this theory. In particular, one important property in the world of both denotational semantics and functional programming is to be able to work with objects at *higher order*. Higher order is a paradigm where objects and programs can be mixed, where functions and their inputs and outputs have the same nature. Defining a logical system which fulfills this property comes with a lot of technical points that should be verified. Moreover, it was limited to only one operator, leading to some asymmetry of the rules.

This thesis has been guided by the hope that generalizing the logic D-DiLL may lead to filling this gap. Such a hope arose following the remark that the idea of indexing the exponential connectives has already been explored from the beginning of linear logic. At first, several refinements on the formulas of linear logic such as elementary linear logic [Gir98], light linear logic [Gir98] or soft linear logic [Laf04] have been studied. Girard *et al.* defined bounded linear logic [GSS91] where they index the exponentials to represent computation time. In particular, indexing by elements of a semiring led to the definition of graded linear logic [GS14]. This is a framework that has been studied at several levels [GKO+16, BP15] and used as a typing system [BGMZ14] which works in particular at higher order. We then tried to unify what has been done in graded linear logic and the logic D-DiLL to use the nice structural properties of the former and apply them to the latter.

Question: how to unify differential linear logic and graded linear logic, using differential operators?

This, of course, comes with many difficulties, at many levels.

Higher-order and the promotion rule The question of higher order has always brought difficulties within differential linear logic. Being able to consider semantical higher-order objects comes with the syntactical promotion rule of (differential) linear logic. Semantically, this rule corresponds to the composition of functions, and its interaction with the codereliction corresponds to the chain rule. The chain rule allows one to compute the differential of a composition of functions by the following equality

$$D_a(g \circ f) = D_{f(a)}g \circ D_a f$$

where the differentiation does not simply distribute on the composition, but the point of differentiation is changed. This is what makes this rule harder to express in the syntax of differential linear logic.

Partial differential equations at higher order Having higher-order objects, and considering linear partial differential equations, comes with difficulties as well. In order to unify the graded framework with the differential one, we replace the operator of differentiation D_a by a linear partial differential equation. A linear partial differential equation is a functional equation of the form

$$a_m \frac{\partial^{|\alpha_m|}}{\partial x_1^{\alpha_m(1)} \dots \partial x_n^{\alpha_m(n)}}(f) + \dots + a_1 \frac{\partial^{|\alpha_1|}}{\partial x_1^{\alpha_1(1)} \dots \partial x_n^{\alpha_1(n)}}(f) = g$$

where the a_i are functions and where we seek for a function f satisfying this equation. Usually, the functions a_i, f and g are smooth maps from $\mathbb{R}^n$ to $\mathbb{R}$ and the elements x_i are the elements of the canonical basis of $\mathbb{R}^n$, along with the function f is differentiated. But at higher order, this is not that simple. The functions a_i, f and g may be functions from, for example, the space $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ to the space $\mathcal{C}^\infty(\mathbb{R}^l, \mathbb{R})$ which are infinite dimensional vector spaces. For those spaces, there is, of course, no canonical basis and knowing how to interpret the elements x_i is not an easy thing. Such differential equations at higher order have not been studied by mathematicians, and even the spaces of smooth functions or distributions with an infinite dimensional domain are not well known.

Cut elimination and semiring structures In a graded setting, the study of the cut elimination procedure comes with algebraic considerations on the algebraic structure, which contain the indices. In graded linear logic, each cut elimination case is built similarly to its equivalent in linear logic, but uses an equality on the indices that correspond to an algebraic axiom of a semiring. Incorporating the structural rules of differential logic, along with indices, expands the number of cut elimination cases that should correspond to algebraic axioms. But the axioms of a semiring are not sufficient to represent these algebraic equations. Moreover, the interactions between the promotion and the costructural rules cannot be simply described have to change when we use grades and the dynamics of the logic is then not the same. In addition, the linear partial differential operators that we consider do not satisfy every axiom of a semiring, so some work has to be made to use them as indices.

We have been able to solve these difficulties by providing partial solutions to them in several directions. We first study a unified framework where the promotion rule is not considered, thus without higher-order objects. We then include this promotion rule in a graded and differential setting, where the indices of the exponential connectives are polarized. This polarity allows us to control the interactions between the rules that we have, and is semantically based on the fact that we can relate to the formulas of different polarities using the Laplace transform. Finally, we consider a system without polarities, and we study the restrictions that need to be made at the level of the indices to have a system with a cut elimination. Let us give the outline of the thesis, explaining the directions taken in each chapter.

Chapter 1 gives the background needed to read this manuscript. We present the several logics at stake in this thesis. We introduce linear logic, graded linear logic and differential linear logic. Furthermore, we present their syntax and explain their cut elimination, which represents the computation contents of the proofs of these logics. We also give the categorical semantics for all of them. Finally, we present some well-known models of these logics, which will be useful for the rest of the manuscript. For linear logic, we define the category of relations, which is the most famous model. We show how it extends to graded linear logic for a general semiring. Then we present two models of differential linear logic: the one from Ehrhard with Köthe spaces, and the one with smooth functions and distributions from Kerjean.

Chapter 2 consists of building a unified framework for graded and differential linear logic using differential operators, but removing the promotion rule. We first introduce the logic D-DiLL from Kerjean [Ker18a] by giving its syntax and a model based on distributions and smooth functions. We then present results co-published with Breuvart and Kerjean [BKM23] which have been invited to a special issue at LMCS. Here, the higher-order question is left out of the picture. The goal is to unify the ideas from graded linear logic and from D-differential linear logic into the same syntax but without the promotion rule. We extend D-DiLL to a graded logic. Removing the rule responsible for higher order has allowed us to simplify the algebraic structure of the indices into a simple commutative monoid. We then generalize the rules of D-DiLL to be indexed by a monoid of differential operators, which gives the logic IDiLL. Then we focus on the semantical side. We adapt the model from Kerjean of distributions and smooth functions with a differential operator. We generalize it in order to consider the actions of several operators together with their interactions. We then take a second approach and define a system where we add differential indexed rules to the linear ones of graded linear logic. We provide such a syntax based on the remark that the rule of subtyping of graded linear logic can be used to represent the action of a linear partial differential operator. We define a cut elimination procedure for this logic based on the interactions of differential linear logic. Therefore, we prove that this logic enjoys a cut elimination procedure under the condition that the monoid of indices respects some algebraic axioms. Finally, we show that those two approaches are, in fact, equivalent. This provides a cut elimination procedure for IDiLL.

Chapter 3 builds a possible framework of a differential linear logic graded by differential operators with a higher-order flavor, introducing a duality on the indices of the exponential formulas. It is based on results from a preprint co-written with Kerjean. Naively adding a promotion rule to IDiLL does not seem to be possible. Syntactically we found ourselves struggling at the level of the cut elimination because of new interactions between the promotion and the new differential counterpart to the subtyping rule and because of algebraic equations at the level of the indices. We have fortunately been saved thanks to semantical hints from a model studied by Kerjean and Lemay [KL23]. They had studied a model of differential linear logic that came from work from Ouerdiane *et al.* [GHOR00]. This model is close to our considerations, as it is both differential and graded. It uses some class of functions to bound smooth maps, and these bounds form the set of indices of a graded linear logic. This model makes use of the so-called *Laplace transform*, in its higher-order form. It has led Kerjean and Lemay to remark a nice correspondence between the structural and the costructural rules that are isomorphic thanks to the Laplace transform. Using this idea, we have built a framework where exponential formulas are related by two transformations: the negation and the Laplace transform.

It is materialized in the syntax by an extension of the indices: we use two copies of the same semiring, related by the Laplace transform. From these formulas, we provide a logical system which is graded, differential, and goes to higher order thanks to a promotion rule. We have been able to add a *copromotion* rule in this setting, which is the syntactical counterpart to the notion of codigging defined by Kerjean and Lemay [KL23]. We then obtain a double indexed differential linear logic named DIDiLL. Thanks to this point of view, with two copies of the semiring, the interactions of the logical rules are highly simplified. We have then been able to easily study these interactions for the cut elimination procedure of this logic. In this chapter we define two models of this logic. The first one consists of the combination of Köthe spaces with differential operators. We endow linear partial differential operators with a relaxed structure of semiring. This relaxed structure corresponds exactly to the algebraic axioms required by the logic DIDiLL. We define how to apply those differential operators to Köthe sequences, and prove that it actually gives a model of DIDiLL. To do so, the copromotion rule has to be removed, as we do not know how to interpret the codigging in the context of Köthe spaces. Finally, we prove that the setting from Ouerdiane *et al* [GHOR00] with the Laplace transform forms a model of our logic. We use a semiring of control functions for the indexation. Here we had to turn DIDiLL into a polarized flavor.

Chapter 4 takes another point of view where we define a system graded by a single copy of the semiring and no polarization, and we study the algebraic axioms needed to perform the cut elimination cases. Due to the removal of the polarities, there are many more cases compared to DIDiLL, which lead to an algebraic structure with many new axioms. In addition, in some cases, the dynamics is quite different from its non-graded version in differential linear logic. We give the dynamics for each case, together with the algebraic equalities that need to be satisfied to perform the cut elimination cases, leading us to define the notion of differential semiring. We give some examples of usual semirings that are differential semirings and prove that the semiring of linear partial differential operators with constant coefficients is not a differential semiring.

Chapter 1

Background: Linear Logic, Gradation and Differentiation

Contents

In this chapter, we present the material needed for the definitions and results given in this thesis. The first notion we introduce in linear logic. We give the syntax of this logic, defining the grammar of its formulas and the inference rules of linear logic. We then describe its cut elimination procedure, which represents the calculus of linear logic. Several versions of the syntax of linear logic exist, depending on the situation one wants to represent. We present another version, which is polarized linear logic that uses polarization on the formulas, which is a technique that will be useful in the rest of this manuscript. After these syntactical definitions, we describe the semantics of linear logic by introducing its categorical semantics and presenting a concrete model, which is the relational model.

A second system that we present is graded linear logic. It is an extension of linear logic that uses a semiring to represent more precise situations than linear logic. Depending on the semiring used, graded linear logic can correspond to an analysis of effects and coeffects, represent probabilistic programs and so on. We give its syntax and its cut elimination, which are based on the one of linear logic. We then show how to extend the relational model to get a model of graded linear logic.

The last system we introduce in differential linear logic, which is at the core of this thesis. Differential linear logic is another extension of linear logic, which adds some inference rules that correspond at the level of the model to have a notion of differentiation. Every system that we will define and study in this manuscript is based on differential linear logic. We present this system, giving the new rules and defining its cut elimination. Then we show how the categorical semantics of linear logic can be extended into a semantics for differential linear logic. We then present two concrete models of this logic that will be used in the next chapters. The first one is a model of Köthe spaces, which is based on constructions over vector spaces with their basis. The second one uses smooth functions and distributions.

1.1 Linear logic

Linear logic (LL) was defined by Girard in 1987 [Gir87]. It came from a denotational study of λ-calculus in terms of normal functors [Gir88b]. One way of seeing linear logic is as a refinement of classical and intuitionistic logic. Its calculus contains both the nice dualities of classical logic together with some constructiveness from intuitionistic logic. In classical logic, the "and" and the "or" can be given in an additive or in a multiplicative flavor. The additive corresponds to having the same context, the multiplicative to have a different one. Thanks to the structural rules, the weakening and the contraction, these two flavor are equivalent. But these structural rules are the reason for the non-constructivism of classical sequent calculus. In order to remove this non-constructivism, linear logic refines the use of the structural rules. As a consequence, it has two "and" connectives: a multiplicative and an additive one. And it is the same for the "or" connective. Removing structural rules leads to a logic where the available formulas are consumed, and then hints at a logic of resources. But to be able to embed both classical and intuitionistic logic, the contraction and weakening need to be possible, at least on some formulas. Linear logic is then made up of some formulas that are consumed, and that cannot be weakened or contracted, and some other which are not consumed and can be used as much as one wants. This distinction comes with a refinement of the implication. While in intuitionistic logic, a formula $A \Rightarrow B$ is interpreted as "B is true whenever A is", it says nothing about how many times one needs A to produce B. Linear logic refines this by introducing a linear implication $\multimap$. Now $A \multimap B$ can be seen as "if you give me A I will give you B". There is also a new operator in linear logic, the ! (read bang). A formula $!A$ should be understood as having as many A as wanted. Then, it makes sense in terms of resources to be able to apply a contraction or a weakening on such a formula. In linear logic, the formulas of the form $!A$, named the exponential formulas, will be the only ones on which these structural rules will be possible. From this new operator !, one recovers the classical implication through

$$A \Rightarrow B \equiv !A \multimap B$$

in the sense that having $!A$ allows not to be careful about the amount of time A is needed to produce B. Let us now present how these ideas are formalized into linear logic.

1.1.1 Syntax of classical linear logic

To define the syntax of classical linear logic, we first define the grammar of its formulas. They are based on a countable set of atoms $\mathcal{X}$, and several connectives.

$$A, B \triangleq X \mid X^{\perp} \mid \qquad \text{(atoms)}$$

$$1 \mid \bot \mid A \parr B \mid A \otimes B \mid \qquad \text{(multiplicative)}$$
$$0 \mid \top \mid A \,\&\, B \mid A \oplus B \mid \qquad \text{(additive)}$$
$$!A \mid ?A \qquad \text{(exponential)}$$

with $X \in \mathcal{X}$. The connectives are split into several groups: multiplicative, additive or exponential. One defines inductively the negation of a linear logic formula by

$$\begin{array}{rcl} (X)^\perp & \triangleq & X^\perp \\ (1)^\perp & \triangleq & \bot \\ (0)^\perp & \triangleq & \top \\ (A \parr B)^\perp & \triangleq & A^\perp \otimes B^\perp \\ (A \,\&\, B)^\perp & \triangleq & A^\perp \oplus B^\perp \\ (!A)^\perp & \triangleq & ?A^\perp \end{array} \qquad \begin{array}{rcl} (X^\perp)^\perp & \triangleq & X \\ (\bot)^\perp & \triangleq & 1 \\ (\top)^\perp & \triangleq & 0 \\ (A \otimes B)^\perp & \triangleq & A^\perp \parr B^\perp \\ (A \oplus B)^\perp & \triangleq & A^\perp \,\&\, B^\perp \\ (?A)^\perp & \triangleq & !A^\perp \end{array}$$

One easily proves that this negation operation is involutive: for each formula A,

$$(A^\perp)^\perp = A.$$

While in classical logic, for example, the implication is defined as a connective of the logic, it is not the case in linear logic. But through the connectives of linear logic and the negation, one defines

$$A \multimap B \triangleq A^\perp \parr B$$

as the *linear implication* of linear logic.

1.1.1.1 Sequent calculus

The grammar of linear logic is endowed by inference rules, which describe the sequent calculus of linear logic. These rules are given in Figure 1.1. We use capital Greek letters $\Gamma, \Delta, \Xi, \ldots$ to denote finite sequences of formulas. When we put an operator in front of a capital Greek letter like $?\Gamma$ it represents the fact that the formulas in the sequence are preceded by this operator, $?\Gamma$ have the form $?A_1, \ldots, ?A_n$. The sequent rules are divided into several fragments. The simplest one is the multiplicative fragment: the identity rules of Figure 1.1a with the multiplicative rules of Figure 1.1b forms multiplicative linear logic (MLL). Adding the additive rules of Figure 1.1c gives the multiplicative additive fragment (MALL). The full linear logic (LL) is given by adding the exponential rules of Figure 1.1d.

Remark 1.1.1. After this syntactical presentation of linear logic, the reader might regret a few gaps. We do not try to fully present linear logic here. Linear logic can be, for example, extended to second order through quantifiers at the level of formulas, which have their own syntactic rules. It can also be restricted to an intuitionistic logic, and its rule system is different as well. Here, we do not want to be exhaustive, but we rather chose to present a syntax as close as possible to the context of this thesis.

1.1.1.2 Specific isomorphisms

Defining the equivalence of formulas in linear logic as

$$A \equiv B \triangleq (A \multimap B) \otimes (B \multimap A)$$

$$\frac{}{\vdash A, A^\perp}\ ax \qquad \frac{\vdash \Gamma, A \quad \vdash \Delta, A^\perp}{\vdash \Gamma, \Delta}\ cut$$

(a) Identity rules

$$\frac{}{\vdash 1}\ 1 \qquad \frac{\vdash \Gamma}{\vdash \Gamma, \perp}\ \perp \qquad \frac{\vdash \Gamma, A, B}{\vdash \Gamma, A \parr B}\ \parr \qquad \frac{\vdash \Gamma, A \quad \vdash \Delta, B}{\vdash \Gamma, \Delta, A \otimes B}\ \otimes$$

(b) Multiplicative rules

$$\frac{}{\vdash \Gamma, \top}\ \top \qquad \frac{\vdash \Gamma, A}{\vdash \Gamma, A \oplus B}\ \oplus_L \qquad \frac{\vdash \Gamma, B}{\vdash \Gamma, A \oplus B}\ \oplus_R \qquad \frac{\vdash \Gamma, A \quad \vdash \Gamma, B}{\vdash \Gamma, A \,\&\, B}\ \&$$

(c) Additive rules

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?A}\ \mathsf{w} \qquad \frac{\vdash \Gamma, ?A, ?A}{\vdash \Gamma, ?A}\ \mathsf{c} \qquad \frac{\vdash \Gamma, A}{\vdash \Gamma, ?A}\ \mathsf{d} \qquad \frac{\vdash ?\Gamma, A}{\vdash ?\Gamma, !A}\ \mathsf{p}$$

(d) Exponential rules

Figure 1.1: Sequent rules of linear logic

one can prove in linear logic that we have a De Morgan duality between the connectives.

$$\begin{array}{rclcrcl}
1^\perp & \equiv & \perp & \qquad & \perp^\perp & \equiv & 1 \\
0^\perp & \equiv & \top & & \top^\perp & \equiv & 0 \\
(A \otimes B)^\perp & \equiv & A^\perp \parr B^\perp & & (A \parr B)^\perp & \equiv & A^\perp \otimes B^\perp \\
(A \oplus B)^\perp & \equiv & A^\perp \,\&\, B^\perp & & (A \,\&\, B)^\perp & \equiv & A^\perp \oplus B^\perp \\
(!A)^\perp & \equiv & ?A^\perp & & (?A)^\perp & \equiv & !A^\perp
\end{array}$$

which are obviously divided between the additive dualities and the multiplicative dualities. Other usual isomorphisms in logic can be proved: the distributivity between the connectives. This distributivity is here for multiplicatives over the additives.

$$\begin{array}{rclcrcl}
A \otimes (B \oplus C) & \equiv & (A \otimes B) \oplus (A \oplus C) & \qquad & A \otimes 0 & \equiv & 0 \\
A \parr (B \,\&\, C) & \equiv & (A \parr B) \,\&\, (A \parr C) & & A \parr \top & \equiv & \top
\end{array}$$

Finally, two crucial isomorphisms that we chose to highlight are the Seely isomorphisms.

$$!(A \,\&\, B) \equiv !A \otimes !B \qquad\qquad !\top \equiv 1$$

This syntactical isomorphism will come back later when we define the semantics of linear logic.

1.1.2 Cut elimination of linear logic

From the sequent rules of linear logic, one can define a cut elimination procedure that represents the dynamics of the logic. We give the cut elimination cases for the MALL subpart in Figure 1.2 and for the exponential rules in Figure 1.3.

These cut elimination cases represent the elimination of the cuts on *principal formulas*, the formulas on which the rule before the cut acts. But this does not represent every

$$\dfrac{\dfrac{}{\vdash A, A^{\perp}}\,ax \qquad \genfrac{}{}{0pt}{}{\pi}{\vdash \Gamma, A}}{\vdash \Gamma, A}\,cut \quad\rightsquigarrow\quad \genfrac{}{}{0pt}{}{\pi}{\vdash \Gamma, A}$$

$$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\pi}{\vdash \Gamma}}{\vdash \Gamma, \perp}\,\perp \qquad \dfrac{}{\vdash 1}\,1}{\vdash \Gamma}\,cut \quad\rightsquigarrow\quad \genfrac{}{}{0pt}{}{\pi}{\vdash \Gamma}$$

$$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\pi_1}{\vdash \Gamma, A^{\perp}, B^{\perp}}}{\vdash \Gamma, A^{\perp} \parr B^{\perp}}\,\parr \qquad \dfrac{\genfrac{}{}{0pt}{}{\pi_2}{\vdash \Delta, A} \qquad \genfrac{}{}{0pt}{}{\pi_3}{\vdash \Xi, B}}{\vdash \Delta, \Xi, A \otimes B}\,\otimes}{\vdash \Gamma, \Delta, \Xi}\,cut \quad\rightsquigarrow\quad \dfrac{\dfrac{\genfrac{}{}{0pt}{}{\pi_1}{\vdash \Gamma, A^{\perp}, B^{\perp}} \qquad \genfrac{}{}{0pt}{}{\pi_2}{\vdash \Delta, A}}{\vdash \Gamma, \Delta, B^{\perp}}\,cut \qquad \genfrac{}{}{0pt}{}{\pi_3}{\vdash \Xi, B}}{\vdash \Gamma, \Delta, \Xi}\,cut$$

$$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\pi_1}{\vdash \Gamma, A^{\perp}} \qquad \genfrac{}{}{0pt}{}{\pi_2}{\vdash \Gamma, B^{\perp}}}{\vdash \Gamma, A^{\perp} \,\&\, B^{\perp}}\,\& \qquad \dfrac{\genfrac{}{}{0pt}{}{\pi_3}{\vdash \Delta, A}}{\vdash \Delta, A \oplus B}\,\oplus_L}{\vdash \Gamma, \Delta}\,cut \quad\rightsquigarrow\quad \dfrac{\genfrac{}{}{0pt}{}{\pi_1}{\vdash \Gamma, A^{\perp}} \qquad \genfrac{}{}{0pt}{}{\pi_3}{\vdash \Delta, A}}{\vdash \Gamma, \Delta}\,cut$$

$$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\pi_1}{\vdash \Gamma, A^{\perp}} \qquad \genfrac{}{}{0pt}{}{\pi_2}{\vdash \Gamma, B^{\perp}}}{\vdash \Gamma, A^{\perp} \,\&\, B^{\perp}}\,\& \qquad \dfrac{\genfrac{}{}{0pt}{}{\pi_3}{\vdash \Delta, B}}{\vdash \Delta, A \oplus B}\,\oplus_R}{\vdash \Gamma, \Delta}\,cut \quad\rightsquigarrow\quad \dfrac{\genfrac{}{}{0pt}{}{\pi_2}{\vdash \Gamma, B^{\perp}} \qquad \genfrac{}{}{0pt}{}{\pi_3}{\vdash \Delta, B}}{\vdash \Gamma, \Delta}\,cut$$

Figure 1.2: Cut elimination of MALL

possible cut: a formula in the context can be cut. For example in the following derivation

$$\dfrac{\dfrac{\dfrac{}{\vdash X_1, X_1^{\perp}}\,ax}{\vdash X_1, X_1^{\perp}, ?X_2}\,\mathsf{w} \qquad \dfrac{\dfrac{}{\vdash X_1, X_1^{\perp}}\,ax}{\vdash X_1, X_1^{\perp}, ?X_2}\,\mathsf{w}}{\vdash X_1, X_1^{\perp}, ?X_2, ?X_2}\,cut$$

the cut occurs on X_1 but the formulas introduced by the principal formulas are the two $?X_2$. To deal with this possibility, some commutations have to be made on the rules before applying the cut elimination cases that we have defined. In this specific case, it corresponds to a commutation on the cut with one weakening

$$\dfrac{\dfrac{\dfrac{}{\vdash X_1, X_1^{\perp}}\,ax}{\vdash X_1, X_1^{\perp}, ?X_2}\,\mathsf{w} \qquad \dfrac{\dfrac{}{\vdash X_1, X_1^{\perp}}\,ax}{\vdash X_1, X_1^{\perp}, ?X_2}\,\mathsf{w}}{\vdash X_1, X_1^{\perp}, ?X_2, ?X_2}\,cut \quad\rightsquigarrow\quad \dfrac{\dfrac{\dfrac{\dfrac{}{\vdash X_1, X_1^{\perp}}\,ax}{\vdash X_1, X_1^{\perp}, ?X_2}\,\mathsf{w} \qquad \dfrac{}{\vdash X_1, X_1^{\perp}}\,ax}{\vdash X_1, X_1^{\perp}, ?X_2}\,cut}{\vdash X_1, X_1^{\perp}, ?X_2, ?X_2}\,\mathsf{w}$$

and we can now apply the cut elimination case of the axiom rule to finish the reduction. Many proof rewritings of this kind have to be defined if one wants to study a complete cut

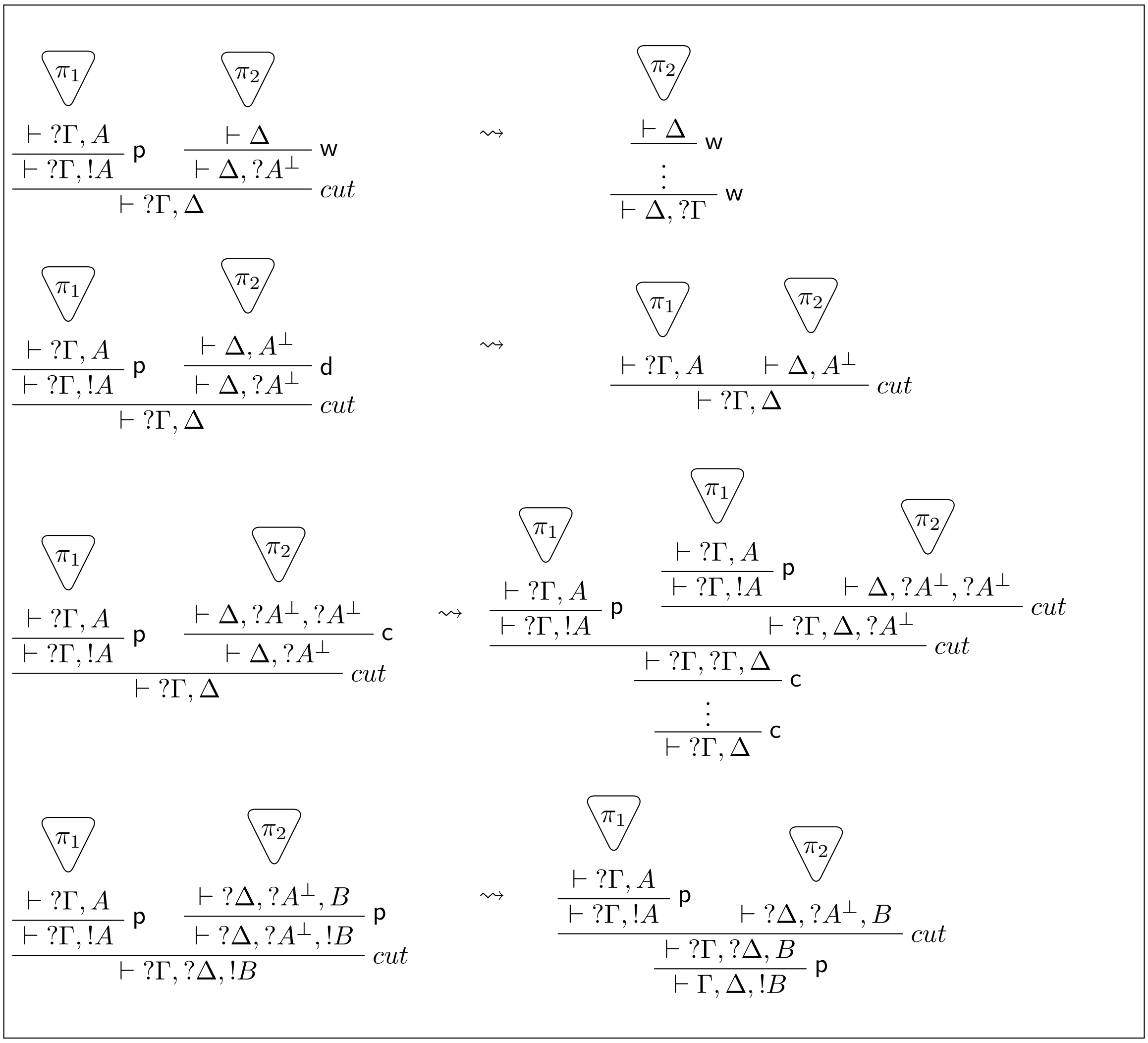


Figure 1.3: Cut elimination for the exponential rules of linear logic

elimination procedure for LL sequent calculus. A full description of these commutation cases can be found in Di Guardia's thesis [DG24]. From these rewriting, one proves a theorem of cut elimination.

Theorem 1.1.2. *For each proof π of a sequent $\vdash \Gamma$ of LL, there is a cut-free proof π' of $\vdash \Gamma$ of LL such that $\pi \rightsquigarrow^* \pi'$.*

Note that to be completely formal, the commutation cases have to be added to the relation $\rightsquigarrow$ to prove such a theorem. In this case, the relation $\rightsquigarrow$ is called *weakly normalizing*.

1.1.3 Polarized linear logic

The notion of polarized linear logic appears in the context of proof search, with the notion of focalization [And90]. The idea is that, when one is looking for a proof for a particular sequent, reducing the number of possible rules that can be used to prove this sequent simplifies the research for such a proof. This is one of the benefits of the cut elimination: since each provable sequent admits a cut-free proof, one does not need to

$$\begin{array}{rcl} P, Q & \triangleq & X \mid 1 \mid 0 \mid P \otimes Q \mid P \oplus Q \mid \, !N \\ N, M & \triangleq & X^\perp \mid \perp \mid \top \mid N \parr M \mid N \,\&\, M \mid \, ?P \end{array}$$

Figure 1.4: Grammar of the formulas of $\mathsf{LL_{pol}}$ and LLP

consider the cut rule while looking for a proof of a sequent. Another way to restrict the rules to consider is by polarizing the logic. In this context, it works by splitting the set of formulas into two: the positive and the negative formulas. It is done by considering reversible connectives and focalized proofs. A connective $\circledast$ is said to be *reversible* if for each provable sequent $\vdash \Gamma, A$ such that $\circledast$ is principal in A, there is a proof of $\vdash \Gamma, A$ where the last rule is the introduction of $\circledast$. As expressed by Laurent [Lau02], the reversible connectives of linear logic are $\parr, \&, \perp, \top$. In the polarized setting, they will correspond to the *negative* connectives. To prove such a result, we consider a general proof π of, for instance, $\vdash \Gamma, A \parr B$. We then have the following proof in LL

$$\dfrac{\dfrac{\dfrac{\pi}{\vdash \Gamma, A \parr B} \qquad \dfrac{\dfrac{}{\vdash A, A^\perp}\,ax \qquad \dfrac{}{\vdash B, B^\perp}\,ax}{\vdash \Gamma, A^\perp \otimes B^\perp, A, B}\,\otimes}{\vdash \Gamma, A, B}\,cut}{\vdash \Gamma, A \parr B}\,\parr$$

which shows that $\parr$ is reversible, and similar constructions can be given for $\&, \perp$ and $\top$. With the idea of proof search, this property means that one can decompose the negative connectives in a sequent without losing provability. Dually, the connectives $\otimes, \oplus, 0, 1$ are the *positive* connectives of linear logic. They are the connectives which are not reversible in LL[1]. Moreover, they are essential to define the notion of focused proof, which is crucial in proof search. But to introduce focused proofs, one needs to define polarized formulas and proofs. In his PhD thesis, Laurent introduced two polarized systems for linear logic [Lau02]. Both are based on the same grammar of formulas, given in Figure 1.4. The formulas P, Q are built from positive connectives and the formulas N, M from negative ones. For the exponential connectives, their polarity comes from Girard's translation from intuitionistic logic to linear logic. This translation uses ideas from polarized linear logic. Fortunately, these polarities work well with our semantic use of polarized linear logic, as it is shown in Section 1.3.5. The logic $\mathsf{LL_{pol}}$ has the same rules as linear logic, but the formulas are restricted to their polarity. For instance, the rule $\otimes$ has the form

$$\dfrac{\vdash \Gamma, P \qquad \vdash \Delta, Q}{\vdash \Gamma, \Delta, P \otimes Q}\,\otimes$$

where P and Q are positive. It works similarly for every other rule of LL. From this system, Laurent defined LLP [Lau04] by giving the following polarity to the exponential rules form

$$\dfrac{\vdash \Gamma}{\vdash \Gamma, N}\,\mathsf{w} \qquad \dfrac{\vdash \Gamma, N, N}{\vdash \Gamma, N}\,\mathsf{c} \qquad \dfrac{\vdash \Gamma, P}{\vdash \Gamma, ?P}\,\mathsf{d} \qquad \dfrac{\vdash \mathcal{N}, N}{\vdash \mathcal{N}, !N}\,\mathsf{p}$$

where $\mathcal{N}$ is a multiset of negative formulas. In LLP, one shows that each provable sequent contains at most one positive formula. This allows to use the focalization technique for proof search.

[1]Note that we do not consider the case of exponential connectives yet.

One possible application of polarized linear logic is about linear logic translations. We stated earlier that linear logic is a refinement of classical and intuitionistic logic. One way to see that is through translations between those systems. First, forgetting the additivity or the multiplicativity of the connectives of LL, one easily defines translations from classical linear logic to classical sequent calculus, and from intuitionistic linear logic to intuitionistic sequent calculus. With these translations, each provable sequent in classical (resp. intuitionistic) linear logic has an equivalent provable sequent in classical (resp. intuitionistic) sequent calculus. The difficulty is to prove such a property the other way around. Since the contraction and the weakening are restricted to exponential formulas in LL, proofs in classical and intuitionistic logic cannot be directly translated. Girard has defined several translations that we do not detail here, but we want to highlight the following. Girard has introduced the calculus LC which refines classical sequent calculus by having a deterministic cut elimination procedure [Gir91] This calculus is strongly based on the notion of polarized formulas. One possible way to translate classical linear logic into LL is to use LC and its polarities to translate it into LLP.

1.1.4 Categorical semantics

The purpose of the categorical semantics of a logic is to give a new understanding of the syntax of this logic, using categorical language. Several definitions can be given for one logical system. In the context of linear logic, many definitions of categorical axiomatizations have been given: Lafont categories, Seely categories, linear categories . . . We will focus here on Seely categories, as this is the axiomatization that fits well with differential linear logic. A more complete presentation has been made by Melliès [Mel09].

1.1.4.1 Categorical definitions

In this section, we recall some basic notions of category theory, and used them to define the object that we need to interpret linear logic in a categorical framework.

Definition 1.1.3. *Let* $\mathbf{C}$ *and* $\mathbf{D}$ *be two categories. The* product category $\mathbf{C} \times \mathbf{D}$ *is the category whose objects are the pairs* (A, B) *with* A *an object of* $\mathbf{C}$ *and* B *an object of* $\mathbf{D}$*. Its morphisms are the pairs* $(f, g) : (A_1, B_1) \to (A_2, B_2)$ *where* $f : A_1 \to A_2$ *is a morphism of* $\mathbf{C}$ *and* $g : B_1 \to B_2$ *is a morphism of* $\mathbf{D}$*. The composition is defined by*

$$(f_2, g_2) \circ (f_1, g_1) = (f_2 \circ f_1, g_2 \circ g_1)$$

and the identity by

$$id_{(A,B)} = (id_A, id_B).$$

A bifunctor *if a functor whose domain is a product category.*

Definition 1.1.4. *A* symmetric monoidal category *is a category* $\mathbf{C}$ *equipped with a bifunctor* $\_ \otimes \_ : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ *and an object* 1 *together with the isomorphisms*

- $\alpha_{A,B,C} : (A \otimes B) \otimes C \simeq A \otimes (B \otimes C)$
- $\lambda_A : 1 \otimes A \simeq A$
- $\rho_A : A \otimes 1 \simeq A$
- $s_{A,B} : A \otimes B \simeq B \otimes A$

such that the following diagrams commute.
The pentagon identity

$$\begin{array}{ccccc} & & (A \otimes B) \otimes (C \otimes D) & & \\ & \nearrow^{\alpha_{A\otimes B,C,D}} & & \searrow^{\alpha_{A,B,C\otimes D}} & \\ ((A \otimes B) \otimes C) \otimes D & & & & A \otimes (B \otimes (C \otimes D)) \\ \downarrow{\scriptstyle \alpha_{A,B,C}\otimes id_D} & & & & \uparrow{\scriptstyle id_A \otimes \alpha_{B,C,D}} \\ (A \otimes (B \otimes C)) \otimes D & & \xrightarrow{\alpha_{A,B\otimes C,D}} & & A \otimes ((B \otimes C) \otimes D) \end{array}$$

the triangle identity

$$\begin{array}{ccc} (A \otimes 1) \otimes B & \xrightarrow{\alpha_{A,1,B}} & A \otimes (1 \otimes B) \\ & {\scriptstyle \rho_A \otimes id_B} \searrow \quad \swarrow {\scriptstyle id_A \otimes \lambda_B} & \\ & A \otimes B & \end{array}$$

the hexagon identity

$$\begin{array}{ccccc} (A \otimes B) \otimes C & \xrightarrow{\alpha_{A,B,C}} & A \otimes (B \otimes C) & \xrightarrow{s_{A,B\otimes C}} & (B \otimes C) \otimes A \\ \downarrow{\scriptstyle s_{A,B}\otimes id_C} & & & & \downarrow{\scriptstyle \alpha_{B,C,A}} \\ (B \otimes A) \otimes C & \xrightarrow{\alpha_{B,A,C}} & B \otimes (A \otimes C) & \xrightarrow{id_B \otimes s_{A,C}} & B \otimes (C \otimes A) \end{array}$$

the unit coherence

$$\begin{array}{ccc} A \otimes 1 & \xrightarrow{s_{A,1}} & 1 \otimes A \\ & {\scriptstyle \rho_A} \searrow \quad \swarrow {\scriptstyle \lambda_A} & \\ & A & \end{array}$$

and the inverse law

$$\begin{array}{ccc} & B \otimes A & \\ & {\scriptstyle s_{A,B}} \nearrow \quad \searrow {\scriptstyle s_{B,A}} & \\ A \otimes B & \overset{id_{A\otimes B}}{=\!=\!=\!=\!=} & A \otimes B \end{array}$$

Remark 1.1.5. Note that restricting Definition 1.1.6 of a symmetric monoidal category by removing the isomorphisms $s_{A,B}$, and keeping only the pentagon and the triangle identity, one gets the definition of a *monoidal category.*

Definition 1.1.6. *Let* $(\mathbf{C}, \&, \top, \alpha_{\mathbf{C}}, \lambda_{\mathbf{C}}, \rho_{\mathbf{C}}, s_{\mathbf{C}})$ *and* $(\mathbf{D}, \otimes, 1, \alpha_{\mathbf{D}}, \lambda_{\mathbf{D}}, \rho_{\mathbf{D}}, s_{\mathbf{D}})$ *be two symmetric monoidal categories,* $!$ *a functor from* $\mathbf{C}$ *to* $\mathbf{D}$*,* m^2 *a natural transformation from* $!(\_) \otimes !(\_)$ *to* $!(\_ \,\&\, \_)$ *and* m^0 *a morphism from* 1 *to* $!\top$*. We say that* $!$ *is a* symmetric monoidal functor *when the diagrams*

$$\begin{array}{ccc} (!A \otimes !B) \otimes !C & \xrightarrow{\alpha_{\mathbf{D}}} & !A \otimes (!B \otimes !C) \\ \downarrow{\scriptstyle m^2_{A,B} \otimes id_{!C}} & & \downarrow{\scriptstyle id_{!A} \otimes m^2_{B,C}} \\ !(A \,\&\, B) \otimes !C & & !A \otimes !(B \,\&\, C) \\ \downarrow{\scriptstyle m^2_{A\&B,C}} & & \downarrow{\scriptstyle m^2_{A,B\&C}} \\ !((A \,\&\, B) \,\&\, C) & \xrightarrow[!\alpha_{\mathbf{C}}]{} & !(A \,\&\, (B \,\&\, C)) \end{array} \qquad \begin{array}{ccc} !A \otimes !B & \xrightarrow{s_{\mathbf{D}}} & !B \otimes !A \\ \downarrow{\scriptstyle m^2_{A,B}} & & \downarrow{\scriptstyle m^2_{A,B}} \\ !(A \,\&\, B) & \xrightarrow[!s_{\mathbf{C}}]{} & !(B \,\&\, A) \end{array}$$

$$\begin{array}{ccc} !A\otimes 1 & \xrightarrow{\rho_{\mathbf{D}}} & !A \\ {\scriptstyle id_{!A}\otimes m^0}\downarrow & & \uparrow{\scriptstyle !\rho_{\mathbf{C}}} \\ !A\otimes !\top & \xrightarrow[m^2_{A,\top}]{} & !(A \,\&\, \top)\end{array} \qquad \begin{array}{ccc} 1\otimes !B & \xrightarrow{\lambda_{\mathbf{D}}} & !B \\ {\scriptstyle m^0\otimes id_{!B}}\downarrow & & \uparrow{\scriptstyle !\lambda_{\mathbf{C}}} \\ !\top\otimes !B & \xrightarrow[m^2_{\top,B}]{} & !(\top \,\&\, B)\end{array}$$

commute. If m^2 and m^0 are isomorphisms, we say that $!\_$ is strong.

Definition 1.1.7. *Let $\mathbf{C}$ and $\mathbf{D}$ be two categories, $F : \mathbf{C} \to \mathbf{C}$ and $G : \mathbf{D} \to \mathbf{C}$ two functors. If there are two natural transformations*

$$\varepsilon : FG \to 1_{\mathbf{C}} \qquad \eta : 1_{\mathbf{C}} \to GF$$

such that

$$\varepsilon F \circ F\eta = 1_F \qquad G\varepsilon \circ \eta G = 1_G$$

then there is an adjunction *between $\mathbf{C}$ and $\mathbf{D}$. In this case, ε is called the counit and ε the unit of the adjunction and we say that F is left adjoint to G and G is right adjoint to F, denoted by $F \dashv G$.*

Definition 1.1.8. *Let $\mathbf{C}$ be a symmetric monoidal category. Then $\mathbf{C}$ is a* symmetric monoidal closed category *if for each object A of $\mathbf{C}$, the functor $\_ \otimes A : \mathbf{C} \to \mathbf{C}$ has a right adjoint. This right adjoint will be denoted $A \multimap \_$.*

Having such an adjunction corresponds to having the following isomorphism

$$\mathbf{C}(A \otimes B, C) \cong \mathbf{C}(B, A \multimap C) \tag{1.1}$$

for each objects A, B, C of $\mathbf{C}$. Using this isomorphism on $id_{A \multimap B} : (A \multimap B) \to (A \multimap B)$, on gets a map

$$ev_{A,B} : A \otimes (A \multimap B) \to B$$

called the *evaluation map*. From this map, on defines a morphism

$$(A \multimap B) \otimes A \xrightarrow{s_{A,B}} A \otimes (A \multimap B) \xrightarrow{ev_{A,B}} B$$

and using, again, the equation 1.1 we get a morphism

$$\partial_{A,B} : A \to (A \multimap B) \multimap B$$

Definition 1.1.9. *A symmetric monoidal closed category is* $\star$-autonomous *if there is an object $\bot$ such that for each object A the morphism $\partial_{A,\bot}$ is an isomorphism. The object $\bot$ is called a* dualizing object*.*

Definition 1.1.10. *A Seely category is a*

- *symmetric monoidal closed category $(\mathbf{C}, \otimes, 1, \alpha, \lambda, \rho, s)$*
- *which is $\star$-autonomous with a dualizing object $\bot$*
- *which has finite products, denoted by $\&$ with $\top$ as terminal object and π_1 and π_2 as projections*
- *with a comonad $(!, \mathsf{der}, \mathsf{dig})$*
- *with two natural isomorphisms m^2 and m^0 making $!$ a strong symmetric monoidal functor from $(\mathbf{C}, \&, \top)$ to $(\mathbf{C}, \otimes, 1)$*

and such that the following diagram commutes

$$
\begin{array}{ccc}
!A \otimes !B & \xrightarrow{m^2_{A,B}} & !(A \mathbin{\&} B) \\
 & & \big\downarrow \scriptstyle \mathsf{dig}_{A \& B} \\
\scriptstyle \mathsf{dig}_A \otimes \mathsf{dig}_B \big\downarrow & & !!(A \mathbin{\&} B) \\
 & & \big\downarrow \scriptstyle !\langle !\pi_1, !\pi_2 \rangle \\
!!A \otimes !!B & \xrightarrow[m^2_{!A,!B}]{} & !(!A \mathbin{\&} !B)
\end{array}
$$

Proposition 1.1.11. *Let* $\mathbf{C}$ *be a Seely categeory, with comonad* $(!, \mathsf{der}, \mathsf{dig})$. *One defined the category* $\mathbf{C}_!$ *whose objets are those of* $\mathbf{C}$ *and whose morphisms between* X *and* Y *are the morphisms of* $\mathbf{C}$ *between* $!X$ *and* Y *(*$\mathbf{C}_!(X, Y) = \mathbf{C}(!X, Y)$*). Its identites are given by* der *and the composition of two morphisms*

$$f : !X \to Y \qquad\qquad g : !Y \to Z$$

is given by

$$!X \xrightarrow{\mathsf{dig}} !!X \xrightarrow{!f} !Y \xrightarrow{g} Z.$$

This category is the Kleisli category of the comonad $(!, \mathsf{der}, \mathsf{dig})$*, and it is Cartesian closed when* $\mathbf{C}$ *is a Seely categeory.*

Remark 1.1.12. The categorical axiomatization of linear logic as Seely categories is not the only possible axiomatization. As said at the beginning of this section, Lafont categories or linear categories are other possible frameworks. One other way to study the categorical models of LL is to start from linear-nonlinear adjunctions. This corresponds to start by considering an ajonction between a category $\mathbf{L}$ where MALL will be interpreted and a category $\mathbf{C}$ for the exponential part. In the context of Seely categories, this adjunction is the one between the Seely category $\mathbf{C}$ and the Kleisli category of the comonad $(!, \mathsf{der}, \mathsf{dig})$.

1.1.4.2 Interpretation of linear logic

Thanks to the previous definitions, we have all the ingredients necessary to interpret the formulas and the proofs of linear logic in a categorical framework. We recall here what is precisely presented by Laurent [Lau08]. Let us fix a Seely category $\mathbf{C}$. First, we want to interpret linear logic formulas as objects of $\mathbf{C}$. We will denote by $[\![A]\!]$ the interpretation of the formula A in the category $\mathbf{C}$. To interpret MLL we need the monoidal closed structure of $\mathbf{C}$.

- The variables are interpreted through a function $\nu : \mathcal{X} \to \mathbf{C}$, which gives $[\![X]\!] \triangleq \nu(X)$ and $[\![X^\perp]\!] \triangleq [\![X]\!] \multimap \perp$.
- The units 1 and $\perp$ are respectively interpreted by $[\![1]\!] \triangleq 1$ and $[\![\perp]\!] \triangleq 1 \multimap \perp$ where 1 is the neutral of $\otimes$ and $\perp$ the dualizing object.
- The tensor of linear logic is naturally interpreted by the monoidal product of $\mathbf{C}$: $[\![A \otimes B]\!] \triangleq [\![A]\!] \otimes [\![B]\!]$.
- The parr is interpreted thanks to the linear implication $[\![A \parr B]\!] = [\![A^\perp]\!] \multimap [\![B]\!]$.

The proofs of MLL are interpreted by morphisms in $\mathbf{C}$. The interpretation of a proof π of a sequent $\vdash A_1, \dots A_n$ will be written $[\![\pi]\!]$, and will be a morphism of type $1 \to [\![A_1]\!] \parr \cdots \parr [\![A_n]\!]$. We explain how to construct this morphism inductively by giving a general construction for each inductive rule. We strongly use the $\star$-autonomous structure of $\mathbf{C}$ which allows to turn a morphism of type $1 \to A \parr B$ into one of type $A^\perp \to B$ or $B^\perp \to A$ or $A^\perp \otimes B^\perp \to \perp$.

- The rule ax is interpreted by the identity $id_A : [\![A]\!] \to [\![A]\!]$.
- The cut takes two morphisms $f : [\![\Gamma]\!] \to [\![A]\!]$ and $g : [\![A]\!] \to [\![\Delta]\!]$ and composes them giving $f; g : [\![\Gamma]\!] \to [\![\Delta]\!]$.
- The rule 1 is interpreted by the identity $id_1 : 1 \to 1$.
- The rule $\perp$ takes $f : [\![\Gamma]\!] \to [\![A]\!]$ and composes it with $\partial_{A,\perp}$ to have $f; \partial_{A,\perp} : [\![\Gamma]\!] \to [\![A]\!] \parr \perp$.
- The tensor rule $\otimes$ is interpreted by taking $f : [\![\Gamma]\!] \to [\![A]\!]$ and $g : [\![\Delta]\!] \to [\![B]\!]$ and applying the bifunctor $\_ \otimes \_$ to get $f \otimes g : [\![\Gamma]\!] \otimes [\![\Delta]\!] \to f \otimes g$.
- The parr rule does not change the interpretation of the proof; its premise is already interpreted by a map $f : 1 \to [\![\Gamma]\!] \parr [\![A]\!] \parr [\![B]\!]$.

We are then able to interpret MLL thanks to the symmetric monoidal closed structure and the $\star$-autonomy of $\mathbf{C}$. Using the additive structure of $\mathbf{C}$, we extend our interpretation to MALL.

- The additive units $\top$ and 0 are interpreted by $[\![\top]\!] \triangleq \top$ and $[\![0]\!] \triangleq \top \multimap \perp$.
- The & is interpreted by the Cartesian product $[\![A \,\&\, B]\!] \triangleq [\![A]\!] \,\&\, [\![B]\!]$ and the $\oplus$ by its negation $[\![A \oplus B]\!] \triangleq (([\![A]\!] \multimap \perp) \,\&\, ([\![B]\!] \multimap \perp) \multimap \perp)$.

For the proofs, we follow the same method as for MALL.

- For the rule $\top$, we use the terminality of $\top$ which implies a morphism from $[\![\Gamma]\!]$ to $\top$.
- For the &, we combine two morphisms $f : [\![\Gamma]\!] \to [\![A]\!]$ and $g : [\![\Gamma]\!] \to [\![B]\!]$ with the Cartesian product of f and g which gives $\langle f, g\rangle : [\![\Gamma]\!] \to [\![A]\!] \,\&\, [\![B]\!]$.
- For the rule $\oplus_1$, we take $f : [\![A]\!]^\perp \to [\![\Gamma]\!]$ and precompose it with the left projection of the Cartesian product to get $\pi_1; f : [\![A]\!]^\perp \,\&\, [\![B]\!]^\perp \to [\![\Gamma]\!]$.
- For $\oplus_2$, we make the same construction that we made for $\oplus_1$ but using the right projection π_2.

We can interpret then interpret the additive fragment of linear logic using the additive structure of Seely categories. Finally, let us define how to interpret exponential formulas and proofs.

- The exponential connectives are naturally interpreted thanks to the functor !, we define $[\![!A]\!] \triangleq ![\![A]\!]$ and $[\![?A]\!] \triangleq [\![A]\!] \multimap \perp$.

For the proofs, we use the comonad $(!, \mathsf{der}, \mathsf{dig})$ together with the strong monoidality of the functor !.

- For the weakening, taking a morphism $f : 1 \to [\![\Gamma]\!]$ and using the terminality of $\top$ with the monoidality of $!\_$ we define

$$![\![A]\!] \xrightarrow{!t_{[\![A]\!]}} !\top \xrightarrow{m^0} 1 \xrightarrow{f} [\![\Gamma]\!]$$

  where t is the morphism coming from the terminality.

- For the contraction, we take a morphism $f : ![\![A]\!] \otimes ![\![A]\!] \to [\![\Gamma]\!]$ and use the diagonal morphism Δ coming from the cartesian product & together with the strength of $!\_$ to define

$$![\![A]\!] \xrightarrow{!\Delta_{[\![A]\!]}} !([\![A]\!] \,\&\, [\![A]\!]) \xrightarrow{(m^2_{[\![A]\!],[\![A]\!]})^{-1}} ![\![A]\!] \otimes ![\![A]\!] \xrightarrow{f} [\![\Gamma]\!]$$

  which gives us the interpretation of the contraction.

- For the dereliction, we take a morphism $f : [\![A]\!] \to [\![\Gamma]\!]$ and we simply use the unit of the comonad which gives $\mathsf{der}; f : ![\![A]\!] \to [\![\Gamma]\!]$.

- Finally, for the promotion rule, the definition is more tedious because we have to work on the context, and not only one the conclusion. We decompose this context by taking $f : ![\![A_1]\!] \otimes \cdots \otimes ![\![A_n]\!] \to [\![A]\!]$ and we use the multiplication of the comonad dig. We give the precise construction in Figure 1.5, but the main idea is to precompose $!f$ by dig.

This concludes the description of the interpretation of linear logic by Seely categories. Note that each part of the definition of a Seely category corresponds to a subsystem of linear logic. One can then remove some axioms of Seely categories to get a categorical axiomatization of this subsystems. The symmetric monoidal closed part corresponds to the multiplicative part of LL, the Cartesian one to the additive part of LL and the strong symmetric monoidal comonad to the exponential part. One can also remove the $\star$-autonomy to get a model of intuitionistic linear logic.

For this interpretation of linear logic from Seely categories, one proves a correction theorem, which states a compatibility between this interpretation and the cut elimination procedure.

Theorem 1.1.13. *Let π and π' be two proofs of linear logic such that $\pi \rightsquigarrow \pi'$. Then $[\![\pi]\!] = [\![\pi']\!]$.*

This theorem is proved by treating every cut elimination case in an induction, and using categorical diagrams for each case of the induction.

1.1.5 A concrete model: the category of relations

One of the most studied models of linear logic is the category of relations. It is simple to define and to study, and has to nice property of being quantitative. We present it here as an example of a Seely category. It may, however, frustrate the reader to have a non-vectorial model to illustrate linear logic. This feeling should be relieved in Section 1.3 when we introduce Köthe spaces and convenient spaces as models of differential linear logic, which are vectorial models are, in fact, models of linear logic since they are a model of differential linear logic. The presentation that we give here follows what is done in the linear logic handbook [Pro23].

$$
\begin{array}{c}
!\llbracket A_1\rrbracket \otimes !\llbracket A_2\rrbracket \otimes !\llbracket A_3\rrbracket \cdots \otimes !\llbracket A_n\rrbracket \\
\downarrow m^2_{\llbracket A_1\rrbracket,\llbracket A_2\rrbracket} \otimes id_{!\llbracket A_3\rrbracket} \cdots \otimes id_{!\llbracket A_n\rrbracket} \\
!(\llbracket A_1\rrbracket \,\&\, \llbracket A_2\rrbracket) \otimes !\llbracket A_3\rrbracket \cdots \otimes !\llbracket A_n\rrbracket \\
\downarrow m^2_{\llbracket A_1\rrbracket \& \llbracket A_2\rrbracket,\llbracket A_3\rrbracket} \otimes \cdots \otimes id_{!\llbracket A_n\rrbracket} \\
\vdots \\
\downarrow m^2_{\llbracket A_1\rrbracket \& \cdots \& \llbracket A_{n-1}\rrbracket,\llbracket A_n\rrbracket} \\
!(\llbracket A_1\rrbracket \,\&\, \llbracket A_2\rrbracket \,\&\, \llbracket A_3\rrbracket \,\&\, \cdots \,\&\, \llbracket A_n\rrbracket) \\
\downarrow \mathsf{dig}_{\llbracket A_1\rrbracket \& \cdots \& \llbracket A_n\rrbracket} \\
!!(\llbracket A_1\rrbracket \,\&\, \llbracket A_2\rrbracket \,\&\, \llbracket A_3\rrbracket \,\&\, \cdots \,\&\, \llbracket A_n\rrbracket) \\
\downarrow !\left(m^2_{\llbracket A_1\rrbracket,\llbracket A_2\rrbracket \& \ldots \llbracket A_n\rrbracket}\right)^{-1} \\
!(!\llbracket A_1\rrbracket \otimes !(\llbracket A_2\rrbracket \,\&\, \llbracket A_3\rrbracket \,\&\, \cdots \,\&\, \llbracket A_n\rrbracket)) \\
\downarrow !\left(id_{!\llbracket A_1\rrbracket} \otimes \left(m^2_{\llbracket A_2\rrbracket,\llbracket A_2\rrbracket \& \cdots \& \llbracket A_n\rrbracket}\right)^{-1}\right) \\
\vdots \\
\downarrow !\left(id_{!\llbracket A_1\rrbracket} \otimes \cdots \otimes \left(m^2_{\llbracket A_{n-1}\rrbracket,\llbracket A_n\rrbracket}\right)^{-1}\right) \\
!(!\llbracket A_1\rrbracket \otimes !\llbracket A_2\rrbracket \otimes !\llbracket A_3\rrbracket \otimes \cdots \otimes !\llbracket A_n\rrbracket) \\
\downarrow !f \\
!\llbracket A\rrbracket
\end{array}
$$

Figure 1.5: Construction of the promotion in a Seely category

Definition 1.1.14. *Let* **Rel** *be the category whose objects are sets and whose morphisms are power sets of the Cartesian product:* $\mathbf{Rel}(X, Y) = \mathcal{P}(X \times Y)$*. For* $R \in \mathbf{Rel}(X, Y)$ *and* $S \in \mathbf{Rel}(Y, Z)$ *the composition* RS *is defined by*

$$RS \triangleq \{(x, z) \in X \times Z \mid \exists y \in Y, (x, y) \in R \text{ and } (y, z) \in S\}$$

and for each object X *of* **Rel** *the identity* id_X *is defined by*

$$id_X \triangleq \{(x, x) \in X \times X\}.$$

The category **Rel** is a symmetric monoidal category with the Cartesian product as its tensor product and a singleton as its unit. It is symmetric monoidal closed, with $X \multimap Y = X \times Y$ and $\star$-autonomous with $\perp$ being a singleton. The category **Rel** also has an additive structure, given by the Cartesian product defined by

$$X_1 \,\&\, X_2 \triangleq \{1\} \times X_1 \cup \{2\} \times X_2$$

which leads to projections π_i and terminal object $\top$ defined by

$$\pi_i \triangleq \{((i, x), x) \mid x \in X_i\} \qquad \top \triangleq \emptyset.$$

For two morphims $R_1 \in \mathbf{Rel}(Y, X_1)$ and $R_2 \in \mathbf{Rel}(Y, X_2)$ we define the pairing as

$$\langle R_1, R_2\rangle \triangleq \{(y, (i, x)) \mid (y, x) \in R_i \text{ for } i = 1, 2\} \in \mathbf{Rel}(Y, X_1 \,\&\, X_2)$$

Finally, we define a comonad over **Rel** to interpret the exponential structure. To do so, we use finite multisets.

Definition 1.1.15. *A* multiset *m over a set X is a function $m : A \to \mathbb{N}$. The* multiplicity *of an element x is the integer $m(x)$. The sum of two multisets m and m' over the same set X is defined as the multiset $m + m' : A \to \mathbb{N}$ where for each $x \in X$, $m + m'(x) = m(x) + m'(x)$. A multiset m is* finite *when it has finite support. It this case we may use a list notation, where the number of occurrences of an element in the list is its multiplicity.*

Now, one defines a comonad $(!, \mathsf{der}, \mathsf{dig})$ as follows:

- for each object X, $!X \triangleq \mathcal{M}_f(X)$, the set of finite multisets of elements of X;
- for a relation $R \in \mathbf{Rel}(X, Y)$,

$$!R \triangleq \{([a_1, \dots, a_n], [b_1, \dots, b_n]) \mid n \in \mathbb{N} \text{ and } \forall i, (a_i, b_i) \in R\}$$

- for $X \in \mathbf{Rel}$, we set

$$\mathsf{der}_X \triangleq \{([a], a) \mid a \in X\} \in \mathbf{Rel}(!X, X);$$

- for $X \in \mathbf{Rel}$, we set

$$\mathsf{dig}_X \triangleq \{(m_1 + \cdots + m_n, [m_1, \dots, m_n]) \mid n \in \mathbb{N},\ m_1, \dots, m_n \in \mathcal{M}_f(X)\} \in \mathbf{Rel}(!X, !!X)$$

  using the sum of multisets.

From this comonad, the Seely isomorphisms are interpreted by the bijection $m^2_{X,Y}$ between $([x_1, \dots x_n], [y_1, \dots, y_m])$ and $[(1, x_1), \dots, (1, x_n), (2, y_1), (2, y_m)]$. and m^0 by the bijection from the singleton 1 to $\{[]\}$.

1.2 Graded linear logic

Some extensions of linear logic have been defined since the original paper by Girard in 1987. One of them named, bounded linear logic (BLL) from Girard *et al.* [GSS91] is a refinement of LL in the sense that the exponential connectives are now indexed by some polynomials. While linear logic was about taking into account the resources consumed by a program during its execution, this new polynomial represents a constraint about the time of execution of the program as proved by Girard *et al.* Following this technique, Ghica and Smith have extended it to a system where the exponential modality is parametrized by a semiring [GS14]. This idea known as grading has been used to represent coeffects [GKO+16, BGMZ14].

1.2.1 Syntax

Graded linear logic is a system that uses semirings to parametrize the exponential modality of linear logic.

Definition 1.2.1. *A* semiring $(\mathcal{S}, +, 0, \times, 1)$ *is given by a set $\mathcal{S}$ with two associative binary operations on $\mathcal{S}$: a sum $+$ which is commutative and has a neutral element $0 \in \mathcal{S}$ and a product $\times$ which is distributive over the sum and has a neutral element $1 \in \mathcal{S}$.*
Such a semiring is said to be commutative *when the product is commutative.*
An ordered semiring *is a semiring endowed with a partial order $\leq$ such that the sum and the product are monotonic.*

In this section, we will sometimes use the term semiring instead of ordered semiring, as every semiring we consider here is ordered. This algebraic structure can then be used to refine the usual exponential rules of linear logic. An important point in our presentation here is the fact that graded linear logic is usually presented as an extension of *intuitionistic* linear logic, since it is used as a typing system. However, we chose to present it in the context of classical linear logic, in order to fit with our previous presentation of linear logic, and since it is the point of view that we will use in the next chapters.

Let $\mathcal{S}$ be a semiring. We define the logic $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, which is parametrized by $\mathcal{S}$, as follows. The grammar is the one of MALL extended with two connectives that are indexed by the elements of $\mathcal{S}$

$$A, B \triangleq X \mid X^{\perp} \mid A \parr B \mid A \otimes B \mid A \,\&\, B \mid A \oplus B \mid !_x A \mid ?_x A$$

where X are atoms and $x \in \mathcal{S}$. The inference rules of this system are the ones of LL except for the exponential ones that are refined. These refined rules are as follows.

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?_0 A}\,\mathsf{w} \qquad \frac{\vdash \Gamma, ?_x A, ?_y A}{\vdash \Gamma, ?_{x+y} A}\,\mathsf{c} \qquad \frac{\vdash \Gamma, A}{\vdash \Gamma, ?_1 A}\,\mathsf{d}$$

$$\frac{\vdash ?_X \Gamma, A}{\vdash ?_{y\times X}\Gamma, !_y A}\,\mathsf{p} \qquad \frac{\vdash \Gamma, ?_x A \qquad x \leq y}{\vdash \Gamma, ?_y A}\,\mathsf{d}_I$$

The capital X in the promotion represents a tuple of elements of $\mathcal{S}$. If $\Gamma = A_1, \ldots, A_n$ and $X = (x_1, \ldots, x_n)$, $?_X\Gamma = ?_{x_1}A_1, \ldots ?_{x_n}A_n$ and $y \times X = (yx_1, \ldots, yx_n)$ Taking the semiring of natural numbers $\mathbb{N}$ with usual sum and product, one can, for example, understand this parametrization as a way to count the minimal number of resources to produce an output. Going back to a two-sided presentation, the weakening

$$\frac{\Gamma \vdash B}{\Gamma, !_0 A \vdash B}\,\mathsf{w}$$

expresses that if we can produce B with Γ, we can also produce it with 0 times A. For the contraction, if we need x occurrences of A and y occurrences of A then we need $x + y$ occurrences of A. The dereliction states that if we produce an output with A, we produce it with one occurrence of A. The promotion is less trivial with this interpretation. Its two-sided version is

$$\frac{!_{x_1}A_1, \ldots, !_{x_n}A_n \vdash B}{!_{yx_1}A_1, \ldots, !_{yx_n}A_n \vdash !_y B}\,\mathsf{p}$$

and means that if one produces B with x_i times the resource A_i for each i, then it is possible to produce y time the output B if having $y \times x_i$ times the ressouce A_i for each i. Finally, comparing this set of rules to the one of exponential rules of linear logic, one rule is added. This is the rule that we call d_I, the indexed dereliction. It uses the order of the semiring, and states that if we can produce B with x times A, we can also produce it with y times A if y is greater than x. Note that the neutrality of some elements of $\mathcal{S}$ reflects the neutrality of some rules: in linear logic the weakening is the neutral of the contraction and the dereliction is the neutral of the promotion. It is algebraically reflected here, since 0 is the neutral of the sum and 1 the neutral of the product.

We have taken the semiring of natural numbers to give simple intuitions about graded linear logic, but, of course, other semirings can be used to index the exponential. If one takes the semiring of polynomials, this system is (almost) equivalent to bounded linear logic. Another possibility is to use the semiring of positive real numbers. In this case, the system corresponds to ideas from differential privacy [GHH$^+$13].

1.2.2 Cut elimination

The logic $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ being quite close to LL in its syntax, it is natural to try to adapt what is done in linear logic to define its cut elimination. This approach works well, but two points have to be solved. For the indexes, some equalities will be required, and one can wonder if the structure of semiring is powerful enough to express these equalities. The second point is about the indexed dereliction, which is not in LL. This question of cut elimination has been studied by Breuvart and Pagani [BP15]. We give their solution to adapt the one from linear logic in Figure 1.6.

Each cut elimination case here has the same form of the one of linear logic. The only differences are the indexations of the exponentials. The difficulty is to decorate these exponentials correctly. For the cut between a promotion and a weakening, we use the fact that $0 \times x = 0$ for each $x \in \mathcal{S}$ and for the one between a promotion and a dereliction, we use that $1 \times x = x$. For the cut between a promotion and a contraction, the left distributivity of the product allows using the same form as the one for linear logic. For the cut between two promotions, it uses the associativity of the product of the semiring.

The last case is the cut between a promotion and an indexed dereliction. We give the reduction in Figure 1.7. This case works thanks to the monotonicity of the order with respect to the product. From the cut elimination theorem of linear logic, one easily deduces a cut elimination for $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$.

Theorem 1.2.2. *For each proof π of a sequent $\vdash \Gamma$ of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, there is a cut-free proof π' of $\vdash \Gamma$ of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ such that $\pi \rightsquigarrow^* \pi'$.*

Note that here, some commutation cases need to be considered to be completely formal.

1.2.3 Model

The categorical axiomatization of graded linear logic is given by Burnel *et al.* [BGMZ14]. They use bimonoidal categories [Lap72], which gives a categorification of the notion of semiring. They then use the notion of strong monads from Melliès [Mel12] to describe the interaction between the category interpreting linear logic and the bimonoidal one. This model as known as a *bounded exponential situation.* Gaboardi *et al.* gave a similar presentation [GKO+16], without bimonoidal categories, by expliciting the diagrams corresponding to the interaction between the algebraic axioms of the semiring and the category interpreting the formulas and proofs. Other equivalent approaches are possible, such as the notion of stratified categories, which are used by Breuvart and Pagani to give an alternative definition of a model of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ [BP15]. We do not present here a detailed categorical semantics with its conditions, but we only express the objects required to interpret $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$. We then give the example of **Rel** graded by a semiring.

Definition 1.2.3. *Let $\mathcal{S}$ be a semiring. A* bounded exponential situation *consists of a symmetric monoidal closed category $(\mathbf{C}, \otimes, 1, \multimap)$ with a bifunctor $!_{\_} : \mathcal{S} \times \mathbf{C} \to \mathbf{C}$ and six natural transformations*

$$
\begin{array}{ll}
m^0_s : 1 \to !_s 1 & m^2_{s,X,Y} : !_s X \otimes !_s Y \to !_s(X \otimes Y) \\
\mathsf{w}_X : !_0 X \to 1 & \mathsf{c}_{s,t,X} : !_{s+t} X \to !_s X \otimes !_t X \\
\mathsf{d}_X : !_1 X \to X & \mathsf{p}_{X,s,t} : !_{st} X \to !_s !_t X
\end{array}
$$

which satisfies some coherence diagrams, similar to those of linear logic but parametrized (see [BGMZ14]).

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdash ?_{y_1}A_1,\dots,?_{y_n}A_n,B\end{array}}{\vdash ?_{0\times y_1}A_1,\dots,?_{0\times y_n}A_n,!_0B}\,\mathsf{p} \qquad \dfrac{\begin{array}{c}\pi_2\\ \vdash \Gamma\end{array}}{\vdash \Gamma,?_0B^{\perp}}\,\mathsf{w}}{\vdash \Gamma,?_0A_1,\dots,?_0A_n}\,cut \quad\rightsquigarrow\quad \dfrac{\begin{array}{c}\dfrac{\dfrac{\begin{array}{c}\pi_2\\ \vdash \Gamma\end{array}}{\vdash \Gamma,?_0A_1}\,\mathsf{w}}{}\,\mathsf{w}\\ \vdots\end{array}}{\vdash \Gamma,?_0A_1,\dots,?_0A_n}\,\mathsf{w}$$

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdash ?_{y_1}A_1,\dots,?_{y_n}A_n,B\end{array}}{\vdash ?_{1\times y_1}A_1,\dots,?_{1\times y_n}A_n,!_1B}\,\mathsf{p} \qquad \dfrac{\begin{array}{c}\pi_2\\ \vdash \Gamma,B^{\perp}\end{array}}{\vdash \Gamma,?_1B^{\perp}}\,\mathsf{d}}{\vdash \Gamma,?_{y_1}A_1,\dots,?_{y_n}A_n}\,cut \quad\rightsquigarrow\quad \dfrac{\begin{array}{c}\pi_1\\ \vdash ?_{y_1}A_1,\dots,?_{y_n}A_n,B\end{array} \qquad \begin{array}{c}\pi_2\\ \vdash \Gamma,B^{\perp}\end{array}}{\vdash ?_{y_1}A_1,\dots,?_{y_n}A_n}\,cut$$

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdash ?_{z_1}A_1,\dots,?_{z_n}A_n,B\end{array}}{\vdash ?_{(x+y)z_1}A_1,\dots,?_{(x+y)z_n}A_n,!_{x+y}B}\,\mathsf{p} \qquad \dfrac{\begin{array}{c}\pi_2\\ \vdash \Gamma,?_xB^{\perp},?_yB^{\perp}\end{array}}{\vdash \Gamma,?_{x+y}B^{\perp}}\,\mathsf{c}}{\vdash \Gamma,?_{xz_1+yz_1}A_1,\dots,?_{xz_1+yz_n}A_n}\,cut \quad\rightsquigarrow$$

$$\dfrac{\begin{array}{c}\dfrac{\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdash ?_{z_1}A_1,\dots,?_{z_n}A_n,B\end{array}}{\vdash ?_{xz_1}A_1,\dots,?_{xz_n}A_n,!_xB}\,\mathsf{p} \qquad \dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdash ?_{z_1}A_1,\dots,?_{z_n}A_n,B\end{array}}{\vdash ?_{yz_1}A_1,\dots,?_{yz_n}A_n,!_yB}\,\mathsf{p} \qquad \begin{array}{c}\pi_2\\ \vdash \Gamma,?_xB^{\perp},?_yB^{\perp}\end{array}}{\vdash \Gamma,?_{yz_1}A_1,\dots,?_{yz_n}A_n,?_xB^{\perp}}\,cut}{\vdash \Gamma,?_{xz_1}A_1,\dots,?_{xz_n}A_n,?_{yz_1}A_1,\dots,?_{yz_n}A_n}\,cut}{}\,\mathsf{c}\\ \vdots\end{array}}{\vdash \Gamma,?_{xz_1+yz_1}A_1,\dots,?_{xz_1+yz_n}A_n}\,\mathsf{c}$$

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdash ?_{z_1}A_1,\dots,?_{z_n}A_n,B^{\perp}\end{array}}{\vdash ?_{xyz_1}A_1,\dots,?_{xyz_n}A_n,!_{xy}B^{\perp}}\,\mathsf{p} \qquad \dfrac{\begin{array}{c}\pi_2\\ \vdash ?_{y_1}B_1,\dots,?_{y_m}B_m,?_yB^{\perp},C\end{array}}{\vdash ?_{xy_1}B_1,\dots,?_{xy_m}B_m,?_{xy}B^{\perp},!_xC}\,\mathsf{p}}{\vdash ?_{xyz_1}A_1,\dots,?_{xyz_n}A_n,?_{xy_1}B_1,\dots,?_{xy_m}B_m,!_xC}\,cut \quad\rightsquigarrow$$

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdash ?_{z_1}A_1,\dots,?_{z_n}A_n,B^{\perp}\end{array}}{\vdash ?_{yz_1}A_1,\dots,?_{yz_n}A_n,!_yB^{\perp}}\,\mathsf{p} \qquad \begin{array}{c}\pi_2\\ \vdash ?_{y_1}B_1,\dots,?_{y_m}B_m,?_yB^{\perp},C\end{array}}{\vdash ?_{yz_1}A_1,\dots,?_{yz_n}A_n,?_{y_1}B_1,\dots,?_{y_m}B_m,C}\,cut}{\vdash ?_{xyz_1}A_1,\dots,?_{xyz_n}A_n,?_{xy_1}B_1,\dots,?_{xy_m}B_m,!_xC}\,\mathsf{p}$$

Figure 1.6: Cut elimination for the exponential rules in graded linear logic

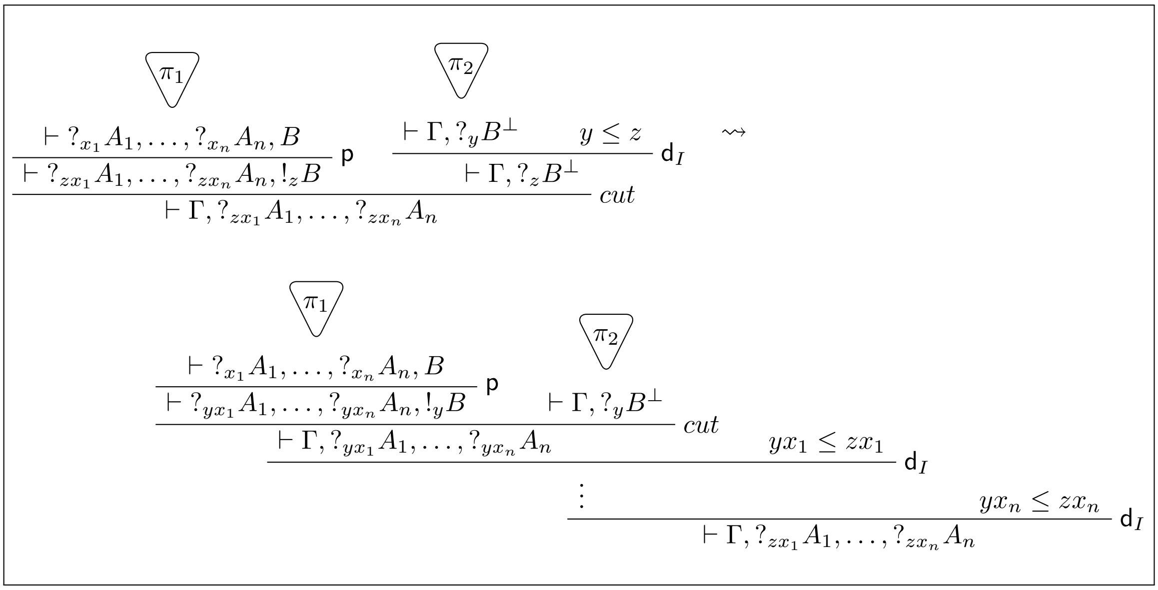


Figure 1.7: Cut elimination for the indexed dereliction in graded linear logic

Remark 1.2.4. To be completely formal, one should define how $\mathcal{S}$ is categorified. Moreover, in the presentation of categorical models of linear logic that we gave, the category where the formulas were interpreted were symmetric monoidal closed, Cartesian closed and $\star$-autonomous. Here it is presented differently due to the fact that $\mathsf{B_{\mathcal{S}}LL}$ is defined as a multiplicative intuitionistic logic. Then, the $\star$-autonomy and the Cartesian closure are not required.

1.2.3.1 A graded version of Rel as a model of $\mathsf{B_{\mathcal{S}}LL}$

We give here the example of a grading, based on the relational model. To do so, we need to extend the notion of multisets which were used for the exponential functor of **Rel**.

Definition 1.2.5. *Let X be a set and $\mathcal{S}$ a semiring. We denote by $\mathcal{S}_f\langle X\rangle$ the set of functions $\mu : X \to \mathcal{S}$ whose support supp$(\mu) = \{x \in X \mid \mu(x) \neq 0_{\mathcal{S}}\}$ is finite. We use $[]$ to denote the function constant at $0_{\mathcal{S}}$ and $[x]$ for the function with value $1_{\mathcal{S}}$ at x and $0_{\mathcal{S}}$ elsewhere, which are elements of $\mathcal{S}_f\langle X\rangle$.*

Proposition 1.2.6. *Let $(\mathcal{X}, \cdot, I)$ be a monoid and $\mathcal{S}$ an ordered semiring. Then $\mathcal{S}_f\langle \mathcal{X}\rangle$ is an ordered semiring with the following structure*

$$0_{\mathcal{S}_f\langle\mathcal{X}\rangle} \triangleq [] \qquad (\mu +_{\mathcal{S}_f\langle\mathcal{X}\rangle} \nu) : x \mapsto \mu(x) +_{\mathcal{S}} \nu(x)$$

$$1_{\mathcal{S}_f\langle\mathcal{X}\rangle} \triangleq [I] \qquad (\mu \times_{\mathcal{S}_f\langle\mathcal{X}\rangle} \nu) : x \mapsto \sum_{x_1 \cdot x_2 = x} \mu(x_1) \times_{\mathcal{S}} \nu(x_2)$$

and the order defined by $\mu \leq_{\mathcal{S}_f\langle\mathcal{X}\rangle} \nu$ iff for all $x \in \mathcal{X}$, $\mu(x) \leq_{\mathcal{S}} \nu(x)$.

From this new semiring, we define a new exponential comonad on **Rel**, through the following exponential functor. For each object X of **Rel**, and each relation $R \in \mathbf{Rel}(X, Y)$,

$$\begin{array}{rcl} !^{\mathcal{S}} X & \triangleq & \mathcal{S}_f\langle X\rangle \\ !^{\mathcal{S}} R & \triangleq & \{(u, v) \in \mathbf{Rel}(!^{\mathcal{S}}X, !^{\mathcal{S}}Y) \mid \exists \mu \in \mathcal{S}_f\langle R\rangle, \quad u(x) = \sum_{y \in Y} \mu(x, y) \text{ and} \\ & & \qquad v(y) = \sum_{x \in X} \mu(x, y)\} \end{array}$$

From this functor, one defines the following natural transformations

$$\begin{array}{rcl} \mathsf{der}_X & \triangleq & \{([x], x) \mid x \in X\} \in \mathbf{Rel}(!^{\mathcal{S}}X, X) \\ \mathsf{dig}_X & \triangleq & \{(m, M) \mid \forall x \in X, m(x) = \sum_{n \in !^{\mathcal{S}}X} n(x) \times M(m)\} \in \mathbf{Rel}(!^{\mathcal{S}}X, !^{\mathcal{S}}!^{\mathcal{S}}X) \end{array}$$

which gives a comonad as proved by Carraro *et al.* [CES10]. On extends this structure, denoted $\mathbf{Rel}^{\mathcal{S}}$, to a Seely category by defining the following natural transformations

$$\begin{array}{rcl} \mathsf{w}_X & \triangleq & \{([], *)\} \in \mathbf{Rel}(!^{\mathcal{S}}X, 1) \\ \mathsf{c}_X & \triangleq & \{(u, (v_1, v_2)) \mid u, v_1, v_2 \in !^{\mathcal{S}}X, u = v_1 + v_2\} \in \mathbf{Rel}(!^{\mathcal{S}}X, !^{\mathcal{S}}X \otimes !^{\mathcal{S}}X) \\ m^0 & \triangleq & \{(*, u) \mid u \in !^{\mathcal{S}}1\} \in \mathbf{Rel}(1, !^{\mathcal{S}}1) \\ m^2_{X,Y} & \triangleq & \{((x \mapsto \sum_{y \in Y} \mu(x, y), y \mapsto \sum_{x \in X} \mu(x, y)), \mu) \mid \mu \in !^{\mathcal{S}}r(X \otimes Y)\} \\ & & \in \mathbf{Rel}(!^{\mathcal{S}}X \otimes !^{\mathcal{S}}Y, !^{\mathcal{S}}(X \otimes Y)). \end{array}$$

Note that, in order to prove that this structure is actually a Seely category, the semiring $\mathcal{S}$ has to be a *multiplicity* semiring (see [CES10, Section 3.1] for the definition). The semiring of natural numbers is the typical example of multiplicity semiring, and one can use it to construct many more.

From the definition of $!^{\mathcal{S}}\_$ and its associated natural transformations, we define a graded comonad to give a graded model of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ based on $\mathbf{Rel}$. To do so, for $s \in \mathcal{S}$ and $X \in \mathbf{Rel}$, we define

$$!_s X \triangleq \{\mu \in !^{\mathcal{S}}X \mid s \geq \sum_{x \in X} \mu(x)\}$$

and the functoriality and the natural transformations are defined through these restrictions. This gives a graded model of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$.

1.3 Differential linear logic

In the early 2000s, Ehrhard studied several vectorial models of linear logic such as finiteness spaces or Köthe spaces [Ehr05, Ehr02]. These frameworks are well suited to study the notion of differentiation. From these considerations, he brought with Regnier this notion of differentiation into the syntax of lambda-calculus to define differential lambda-calculus [ER03] and into linear logic to define differential linear logic (DiLL) [ER06].

In linear logic, the dereliction rule allows to forget the linearity of a proof, going from the linear world to the non-linear one. Semantically, it consists of considering a linear morphism in the Kleisli category of the exponential comonad. Differential linear logic is based on a codereliction rule, which is the syntactical opposite of the dereliction. Its role is the opposite to the one of the dereliction: it allows going from the non-linear world to the linear one. Semantically, it corresponds to taking the best linear approximation of a non-linear map, which is its differential. Two other rules had to be added to the syntax in order to express the cut elimination procedure of differential linear logic. These rules are the coweakening and the cocontraction, which are the syntactical opposites of, respectively, the weakening and the contraction.

1.3.1 Syntax

Originally, the syntax of differential linear logic was given in terms of proof nets, called differential interaction nets. One crucial difference with this kind of proof nets and proof nets of linear logic is that one needs to be able to perform sums of proofs in differential interaction nets. This necessity may be understood from several points of view. Semantically, we will interpret the contraction by the pointwise product of smooth functions and the codereliction by computing the differential at 0 of a function. Expressing the cut elimination of these two rules will then correspond to compute the differential of a product of functions. From the Leibniz rule of differential calculus, we have

$$D(fg) = D(f)g + fD(g)$$

that we will use to express the reduction between a contraction and a codereliction but this works only with sums of proofs. In the syntax, this sum can also be viewed as a form of non-determinism. A contraction cell has two inputs, and to perform the reduction with a codereliction, we must branch it to one of these inputs. We use this non-determinism, and hence this notion of sum, to superpose the two branching possibilities that we have. Note that we chose to illustrate this point with the example of the interaction between the contraction and the codereliction but the sum appears in the interaction between the dereliction and the cocontraction as well. We will detail this later on while giving a precise semantics for DiLL. Recent work by Ehrhard shows how to get rid of this non-determinism, in a framework named coherent differentiation. It is done by controlling precisely the elements that are summed [Ehr23, Ehr24].

The grammar of the formulas of differential linear logic is the same as that of linear logic. For the rules that we still present in a one-sided flavor, it extends those of LL with three new rules

$$\frac{}{\vdash !A}\ \bar{\mathsf{w}} \qquad \frac{\vdash \Gamma, !A \qquad \vdash \Delta, !A}{\vdash \Gamma, \Delta, !A}\ \bar{\mathsf{c}} \qquad \frac{\vdash \Gamma, A}{\vdash \Gamma, !A}\ \bar{\mathsf{d}}$$

the coweakening, the cocontraction and the codereliction. In its first version, differential linear logic was presented as a promotion-free logic: the promotion was removed from the set of rules. We denote by DiLL_0 this logic. In order to take into account the notion of sums of proofs that we mentioned before, we also add two rules that provide an additive structure to the set of proof trees

$$\frac{}{\vdash \Gamma}\ 0 \qquad \frac{\vdash \Gamma \qquad \vdash \Gamma}{\vdash \Gamma}\ +$$

where the rule 0 allows to produce an empty proof (or 0-proof) for each sequent and the rule + to sum two proofs that prove the same sequent.

Remark 1.3.1. Note that the rule 0 allows to prove every sequent. One can then trivially prove `false` is differential linear logic. It can seem quite puzzling, since the consistency property is usually crucial for a logic. However, DiLL is useful for its dynamics: it consists of a logic where the cut elimination procedure underlies the resource analysis of linear logic together with the usual rules of differential calculus. We then focus on the structure of its proofs rather than what can be proved.

1.3.2 Cut elimination

For the cut elimination procedure of DiLL, many new cases appear compared to the procedure of LL. In the exponential rules of linear logic, the only one introducing a formula of the form $!A$ is the promotion. The only possible interaction was then between the promotion and the other structural rules. Each new costructural rules introduce a $!A$ formula. They then interact with every structural rule. We give the cases involving the coweakening in Figure 1.8, the cases with the cocontraction in Figures 1.9 and 1.10 and the ones with codereliction in Figure 1.11. From these cases, one deduces a cut elimination theorem.

Theorem 1.3.2. *For each proof π of a sequent $\vdash \Gamma$ of LL, there is a cut-free proof π' of $\vdash \Gamma$ of LL such that $\pi \rightsquigarrow^* \pi'$.*

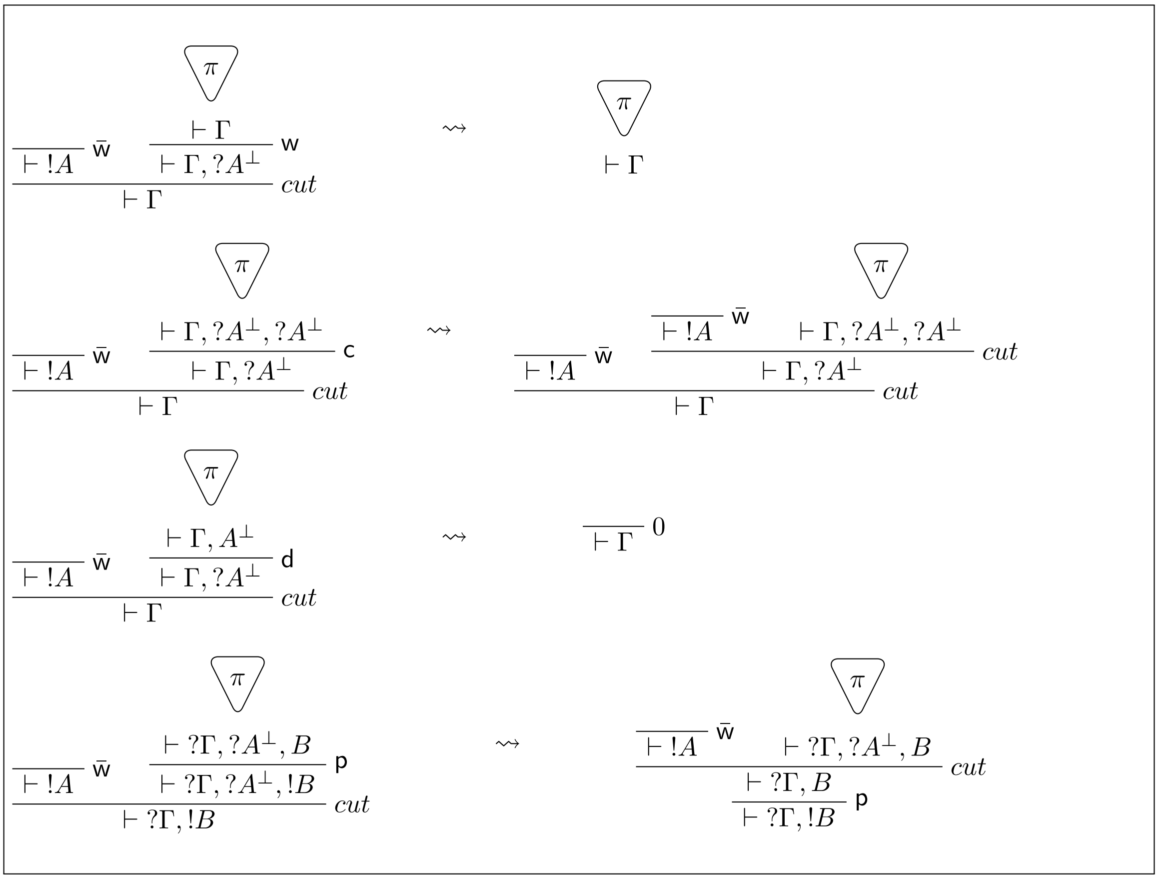


Figure 1.8: Cut elimination cases with $\bar{\mathsf{w}}$ in DiLL

Such a theorem has been proved by Pagani in terms of proof nets [Pag09]. In this context, it has even been proved by Pagani and Tranquilli that the rewriting is strongly normalizing, meaning that there is only one possible sequence of reduction [PT17].

Some semantic intuitions behind the rules and the cut elimination The cut elimination cases from differential linear logic can nicely be described in terms of standard differential calculus. To do so, one should keep one model in mind. It can be **Rel**, or **Köthe** the category of Köthe spaces that we introduce in Section 1.3.4. We chose here to describe this rewriting in terms of smooth maps and distributions, the model **Smooth** that we present in Section 1.3.5. In this model, formulas of MALL are interpreted by finite dimensional real vector spaces, isomorphic to $\mathbb{R}^n$, and exponential formulas by smooth maps and distributions

$$[\![?A]\!] \triangleq \mathcal{C}^\infty([\![A]\!]', \mathbb{R}) \qquad\qquad [\![!A]\!] \triangleq (\mathcal{C}^\infty([\![A]\!], \mathbb{R}))'.$$

The exponential rules are interpreted as follows:

- the weakening introduces the smooth map cst_1 constant at 1;
- the contraction takes two smooth maps and computes their pointwise product;
- the dereliction takes a vector v and computes the smooth map $x \in A' \mapsto x(v)$;

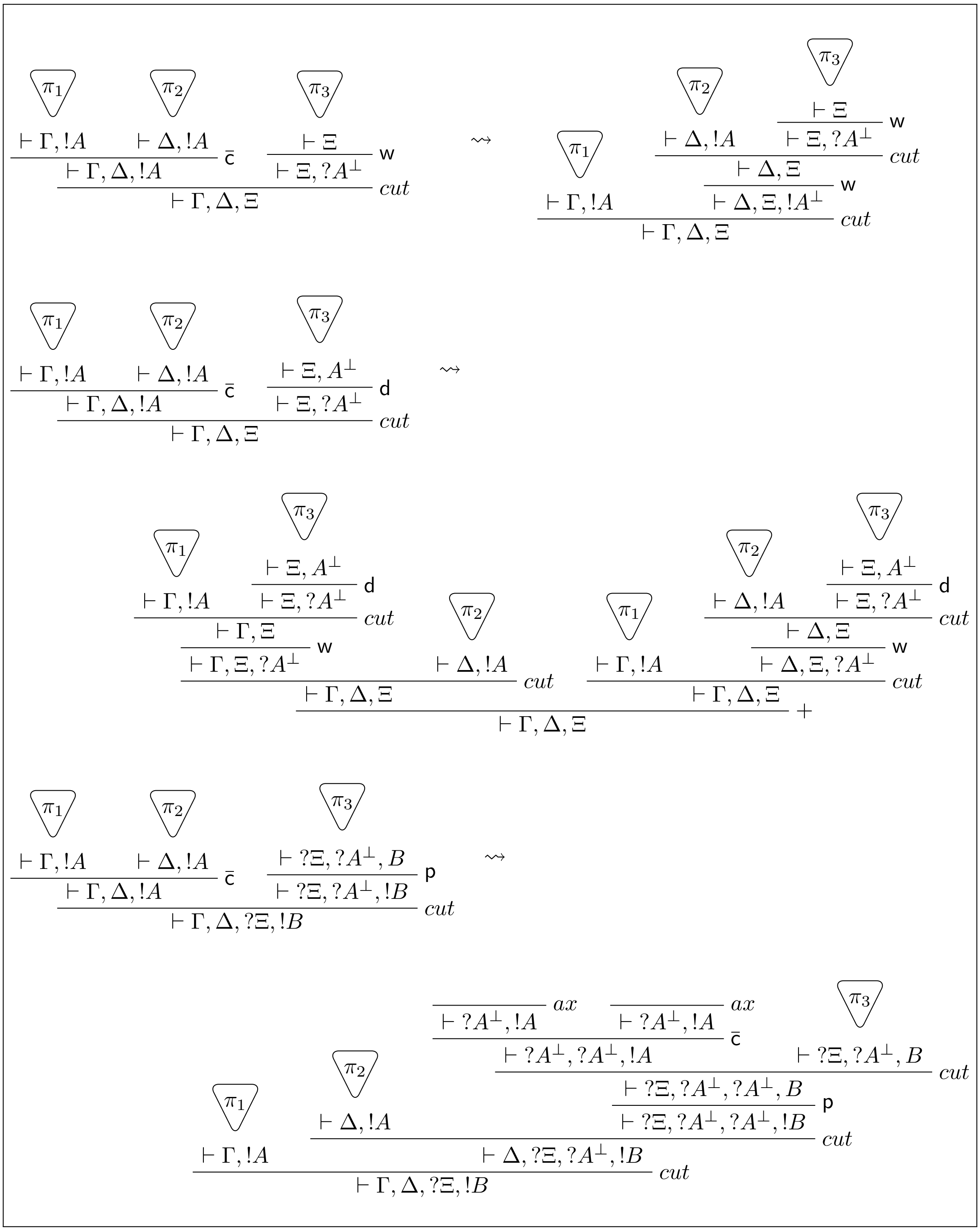


Figure 1.9: Cut elimination cases with c̄ in DiLL

- the promotion takes a vector v and produces the Dirac at v denoted δ_v which is the distribution such that for each smooth function f, $\delta_v(f) = f(v)$;
- the coweakening introduces δ_0, the Dirac at 0;
- the cocontraction takes two distributions and computes their convolution product;

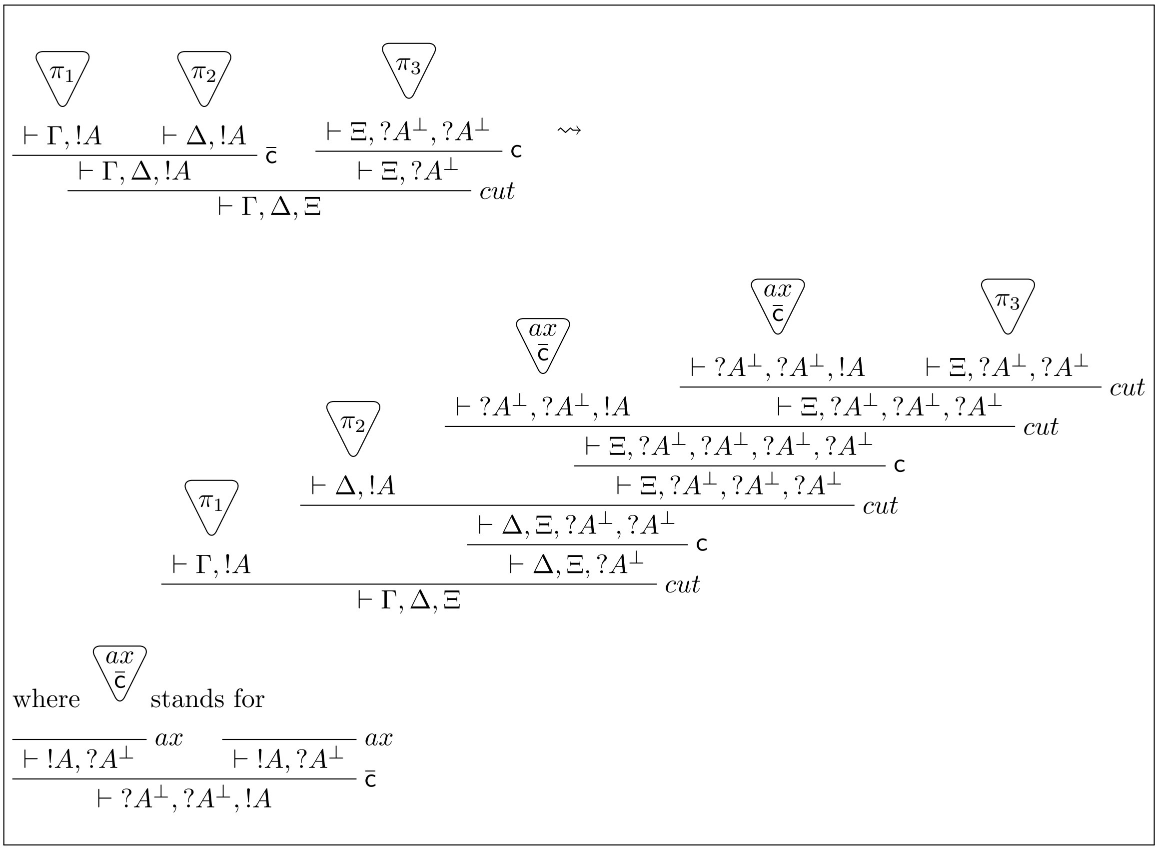


Figure 1.10: Cut elimination case between $\bar{\mathsf{c}}$ and c in DiLL

- the coderelicion takes a vector v, and produces the distribution that takes a smooth map f and computes $D_0(f)(v)$, the differential at 0 of f taken at point v.

Every exponential cut elimination case can be expressed in terms of crucial equalities of this framework. We describe some of them to highlight how DiLL captures differential calculus. To do so, we present the interpretation of the cuts involving a codereliction.

The cut between $\bar{\mathsf{d}}$ and w represents the computation $D_0(cst_1)(v)$ where v is a vector. The rewriting works since $D_0(cst_1)(v) = 0$, which is the case, since the differential of a constant function is 0.

The cut between $\bar{\mathsf{d}}$ and c represents the computation $D_0(f.g)(v)$ where f and g are smooth functions and v a vector. The tree resulting in the cut elimination case in interpreted as follows:

$$\dfrac{\dfrac{\dfrac{}{\vdash \delta_0}\,\bar{\mathsf{w}} \qquad \dfrac{\dfrac{\dfrac{\vdash v}{\vdash D_0(\_)(v)}\,\bar{\mathsf{d}} \qquad \vdash f,g}{\vdash D_0(f)(v),g}\,cut}{}}{\vdash D_0(f)(v).\delta_0(g)}\,cut \qquad \dfrac{\dfrac{\vdash v}{\vdash D_0(\_)(v)}\,\bar{\mathsf{d}} \qquad \dfrac{\dfrac{}{\vdash \delta_0}\,\bar{\mathsf{w}} \qquad \vdash f,g}{\vdash \delta_0(f)}\,cut}{\vdash \delta_0(f).D_0(g)(v)}\,cut}{\vdash D_0(f)(v).g(0) + f(0).D_0(g)(v)}\,+$$

and the equality with the tree before the elimination highlights exactly the product rule (or Leibniz rule) of differential calculus.

The cut between $\bar{\mathsf{d}}$ and d represents the computation $D_0(x \in A^{\perp\perp} \mapsto x(\ell))(v)$ where $\ell \in A^\perp$ is a linear map and $v \in A$ is a vector. Using the isomorphism between A and

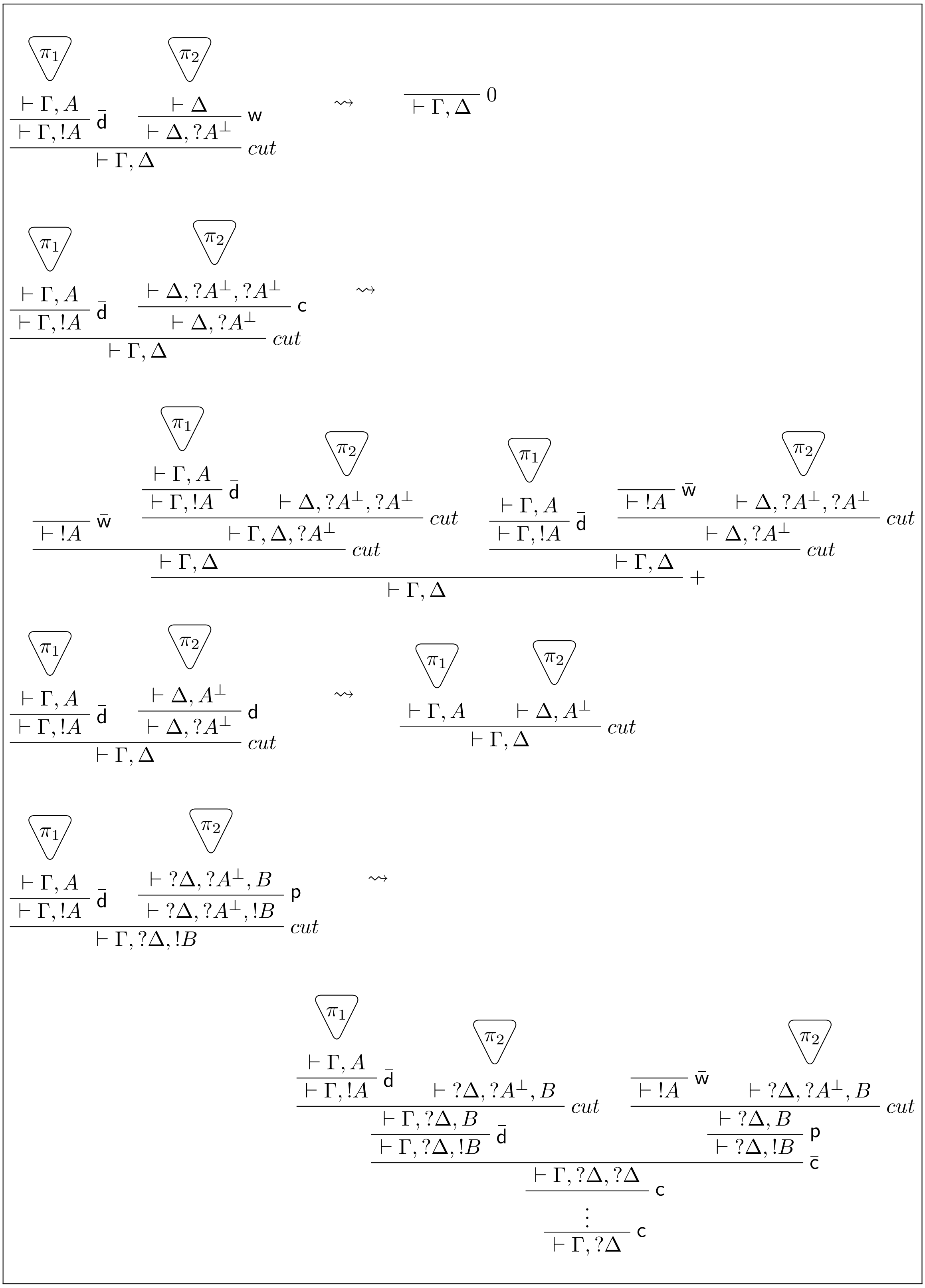


Figure 1.11: Cut elimination cases with $\bar{\mathsf{d}}$ in DiLL

$A^{\perp\perp}$, this computation is equivalent to $D_0(\ell)(v)$. The tree that we get after a rewriting is interpreted by $\ell(v)$ and the equality of the interpretations represents the fact that $D_0(\ell) = \ell$, so that the differential of a linear map is the linear map itself.

The cut between $\bar{\mathsf{d}}$ and p represents the computation of $D_0(\delta_f)(v)$ where v is a vector in A, f is a linear map from $!A$ to B and δ_f is a linear map from $!A$ to $!B$ that takes x in $!A$, and gives $\delta_{f(x)}$ an element of $!B$ which takes $g \in ?B$ and gives $\delta_{f(x)}(g) = g(f(x)) \in \mathbb{R}$. One way to see the interpretation of this cut[2] is by the distribution $g \mapsto D_0(g \circ f)(v)$. After the rewriting, it is interpreted as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdash v}{\vdash D_0(\_)(v)}\,\bar{\mathsf{d}} \quad \vdash f}{\vdash D_0(f)(v)}\,cut}{\vdash D_0(\_)(D_0(f)(v))}\,\bar{\mathsf{d}} \qquad \dfrac{\dfrac{\dfrac{}{\vdash \delta_0}\,\bar{\mathsf{w}} \quad \vdash f}{\vdash \delta_0(f)}\,cut}{\vdash \delta_{\delta_0(f)}}\,\mathsf{p}}{\vdash D_0(\_)(D_0(f)(v)) * \delta_{\delta_0(f)}}\,\bar{\mathsf{c}}$$

Applying this distribution to a smooth function $g \in ?B$, one gets

$$\begin{aligned} D_0(x \mapsto \delta_{f(0)}(y \mapsto g(x+y)))(D_0(f)(v)) &= D_0(x \mapsto g(x+f(0)))(D_0(f)(v)) \\ &= D_{f(0)}(g)(D_0(f)(v)). \end{aligned}$$

This shows that this rewriting corresponds to the chain rule of differential calculus.

The cut elimination cases of the coderelicion rule represent then syntactically the differential calculus rule to compute the differential of constant functions, linear functions, product of functions and composition of functions. While we give intuitions in terms of smooth functions and distributions, all of these interactions can be formalized categorically. This is done in differential categories for example. We mention this at the end of the next section.

1.3.3 The categorical semantics of DiLL

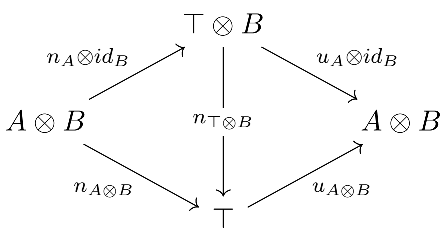
Defining a categorical semantics for differential linear logic consists of extending the semantics of linear logic to be able to interpret the new costructural rules. We will base our definitions on Seely categories, as it is the point of view that we chose to present the categorical semantics of linear logic. Another presentation with exponential structures is given by Ehrhard [Ehr18]. Following work by Fiore [Fio09], we extend Seely categories to interpret the coweakening, the cocontraction and the codereliction. For the coweakening and the cocontraction, the constructions appear naturally extending the product to a biproduct.

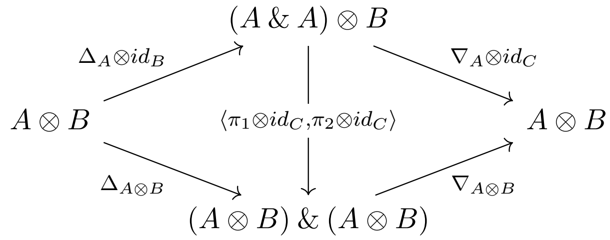
Definition 1.3.3. *Let* **C** *be a symmetric monoidal category. A* biproduct $(\&, \top)$ *of* **C** *is given by four natural transformations*

$$u_A : \top \to A, \qquad n_A : A \to \top, \qquad \nabla_A : A \,\&\, A \to A, \qquad \Delta_A : A \to A \,\&\, A$$

[2] This is not completely right. We present it this way here, as our point is to give intuitions about the syntax but this interpretation lacks some formalism for the promotion rule.

such that (A, u, ∇) *is a commutative* &-*monoid and* (A, n, Δ) *is a commutative* &-*comonoid. Moreover, we say that this biproduct is* compatible with the symmetric monoidal structure *when the following diagrams commute.*





This allows interpreting the coweakening and the cocontraction of differential linear logic. We recall from Section 1.1.4.2 that to interpret the weakening and the contraction, we had to precompose the morphism from the premise with the following morphisms

$$\mathsf{c}_A : !A \xrightarrow{!\Delta_A} !(A \,\&\, A) \xrightarrow{(m^2_{A,A})^{-1}} !A \otimes !A \qquad\qquad \mathsf{w}_A : !A \xrightarrow{!t_A} !\top \xrightarrow{m^0} 1$$

To interpret the coweakening and the cocontraction, we use the biproduct structure to define

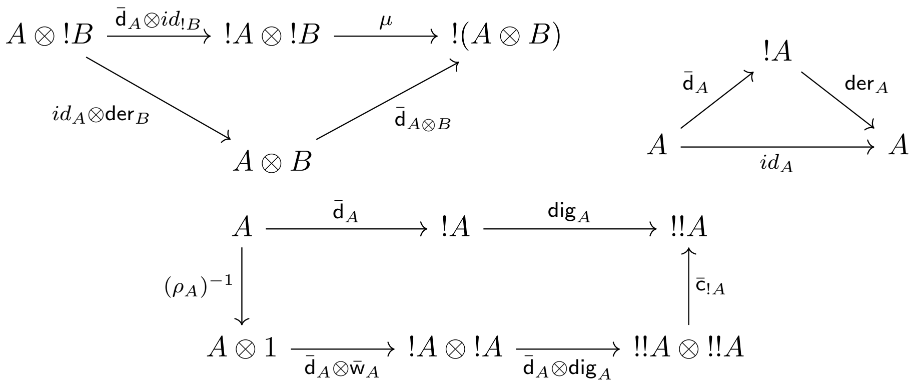
$$\bar{\mathsf{c}}_A : !A \otimes !A \xrightarrow{m^2_{A,A}} !(A \,\&\, A) \xrightarrow{!\nabla_A} !A \qquad\qquad \bar{\mathsf{w}}_A : 1 \xrightarrow{(m^0)^{-1}} !\top \xrightarrow{!u_A} !A$$

From these new morphisms, the interpretation of $\bar{\mathsf{w}}_A$ is exactly the interpretation of the coweakening, and if we take $f : [\![\Gamma]\!] \to !A$ and $g : [\![\Delta]\!] \to !A$ as interpretations of the premises of a cocontraction, the morphism $(f \otimes g); \bar{\mathsf{c}}_A : [\![\Gamma]\!] \otimes [\![\Delta]\!] \to !A$ interprets the rule.

For the codereliction rule, the method is different. One might hope that, since the dereliction of linear logic comes from the unit of the comonad of the Seely category, adding a structure of monad and using its unit should be the right way to interpret the codereliction. This idea has been investigated by Kerjean and Lemay [KL23, KL24]. We will detail it later in Chapter 3. Here, we do not use this idea because it is too strong to describe only differential linear logic. The comonadic structure on $!\_$ underlies both the dereliction and the promotion rule. This is why, to axiomatize the codereliction we do not use a monad but only a natural transformation with the right diagrams. We define $\bar{\mathsf{d}}$ as the following natural transformation

$$\bar{\mathsf{d}}_A : A \to !A$$

which satisfy the following diagrams:



(1.2)

The last challenge in the definition of the categorical axiomatization of differential linear logic is the question of the sums of proofs. As we explain in Section 1.3.1, the structure of the proofs of differential linear logic needs to be endowed with a 0 proof and sums of proofs. Categorically speaking, this corresponds to have an enrichment over the hom-sets of $\mathbf{C}$, meaning that each hom-set is endowed with a commutative monoidal structure $(\mathbf{C}, +, 0)$. This can be asked as a new constraint on $\mathbf{C}$, but things work elegantly here since such an enrichment arises directly from the biproduct structure, as shown by Fiore [Fio09].

Proposition 1.3.4. *If a category* $\mathbf{C}$ *with finite products is endowed with a biproduct structure, then it is enriched over commutative monoids.*

Proof. Let us take $\mathbf{C}$ endowed with a biproduct structure $(\&, \top)$. One defines a monoidal operation $+$ on $\mathbf{C}(A, B)$ together with its neutral element 0 as

$$f + g : A \xrightarrow{\Delta_A} A \,\&\, A \xrightarrow{f \& g} B \,\&\, B \xrightarrow{\nabla_B} B \qquad\qquad 0 : A \xrightarrow{n_A} \top \xrightarrow{u_B} B$$

where $f : A \to B$ and $g : A \to B$ are elements of $\mathbf{C}(A, B)$. □

Definition 1.3.5. *A* Seely model of differential linear logic *is a Seely category equipped with a biproduct compatible with the symmetric monoidal structure and a natural transformation* $\bar{\mathsf{d}}$ *that satisfies the diagrams of equation 1.2.*

Remark 1.3.6. This presentation is one possible axiomatization of differential linear logic. Differential categories [BCS06, BCLS20] can interpret DiLL. The main difference is that the differentiation does not appear with a natural transformation $\bar{\mathsf{d}} : !A \to A$ but from a *deriving transformation* $\mathbf{D} : !A \otimes A \to !A$. This corresponds to give a general operation for the differential, operating non-linearly on the variable along which we differentiate, and linearly on the one we apply the differentiation function. It has the flaw of being further from the syntax of DiLL. It requires more work to interpret DiLL from differential categories. This is why we did not choose to present it here. However, differential categories are closer to intuitions from differential calculus. The commutative diagrams coming with the deriving transformation correspond exactly to the rules of differential calculus that we describe at the end of Section 1.3.2.

Here as well, we have a nice theorem proving that the categorical semantics is invariant over the cut elimination procedure of differential linear logic.

Theorem 1.3.7. *Let* π *and* π' *be two proofs of differential linear logic such that* $\pi \rightsquigarrow \pi'$. *Then* $[\![\pi]\!] = [\![\pi']\!]$.

1.3.4 Köthe spaces as a model of DiLL

We present here one example of a model of differential linear logic where the codereliction corresponds to an actual differentiation. It is possible to extend what we described in Section 1.1.5 to turn **Rel** into a model of DiLL but because of the implicity of **Rel** the differentiation does not really match with the intuitions of approaching a non-linear map by the linear one. The codereliction rule in **Rel** is defined from a deriving transformation, which takes a relation $R \in \mathbf{Rel}(!X, Y)$ and transforms it into

$$DR \triangleq \{((m, x), y) \mid m \in \mathcal{M}_f(X), x \in X, y \in Y, (m + [x], y) \in R\} \in \mathbf{Rel}(!X \times X, Y).$$

Köthe spaces is one of the first models of DiLL, and has been studied by Ehrhard [Ehr02]. It had first been introduced as a model of linear logic, and then used to extend the syntax of LL into DiLL. To define Köthe spaces, we fix X a set at most countable, a field $\mathbb{K}$ (which can be $\mathbb{R}$ or $\mathbb{C}$), and study the space $\mathbb{K}^X$ of sequences of elements of $\mathbb{K}$ indexed by X.

Definition 1.3.8. *If $u, v \in \mathbb{K}^X$ are such that $\sum_{i \in X} |u_i v_i|$ converges, u and v are in* duality. *For $E \subseteq \mathbb{K}^X$, we denote*

$$E^\perp \triangleq \{u \in \mathbb{K}^X \mid \forall v \in E,\ u \text{ and } v \text{ are in duality}\}.$$

For $i \in X$, we define $e_i \in \mathbb{K}^X$ such that $(e_i)_j = 0$ if $i \neq j$ and $(e_i)_i = 1$.

Definition 1.3.9. *A* Köthe space *is a pair $\mathcal{X} = (X, E_X)$ such that X is at most countable, $E_X \subseteq \mathbb{K}^X$ and $E_X^{\perp\perp} = E_X$. Its dual $\mathcal{X}^\perp$ is defined by $X^\perp = X$ and $E_{X^\perp} = E_X^\perp$.*

A crucial remark is that when X is finite, $E_X = \mathbb{K}^X$, since $E_X^\perp = \mathbb{K}^X$ for each $E_X \subseteq \mathbb{K}^X$. For each Köthe space (X, E_X), E_X is a vector space, as a sub-vector space of $\mathbb{K}^X$. It is endowed of the *normal topology* $\varepsilon(E_X)$. This topology is defined from a family of seminorms. We define $(N_x)_{x \in E_X^\perp}$ as

$$N_x(x') = \sum_{i \in X} |x_i x'_i|$$

and the normal topology on E_X is the one induced by this family. We define a correspondence between matrices and linear maps. From a linear map $f : E_X \to E_Y$ we define $M(f) \in \mathbb{K}^{X \times Y}$ where $M(f)_{a,b} = f(e_a)_b$. From a matrix $M \in \mathbb{K}^{X \times Y}$, $f(M) : E_X \to \mathbb{K}^Y$ is defined by $f(M)(x)_b = \sum_{a \in X} M_{a,b} x_a$. Defining $E_{X \multimap Y}$ as the space of $M \in \mathbb{K}^{X \times Y}$ such that

$$\sum_{a \in X, b \in Y} M_{a,b} x_a y'_b$$

converges absolutely for all $x \in E_X$, and $y \in E_{Y^\perp}$, $(X, E_X) \multimap (Y, E_Y)$ is interpreted by $(X \times Y, E_{X \multimap Y})$. The correspondence given above provides an isomorphism between $E_{X \multimap Y}$ and $\mathcal{L}(E_X, E_Y)$. We denote by **Köthe** the category of Köthe spaces with linear continuous maps as morphisms. To define the composition and the identities, we use the correspondence with the matrices and naturally define the composition as the matricial product and the identities thanks to the identity matrix. We now endow **Köthe** with the categorical structure required to make it a model of differential linear logic. The proofs that these constructions fulfill the axioms of models of DiLL are done by Ehrhard [Ehr02].

The multiplicative additive structure on Köthe The monoidal and Cartesian structure in **Köthe** is defined naturally from usual constructions over vector spaces. For two Köthe spaces E_X and E_Y, they tensor product is defined by

$$E_X \otimes E_Y \triangleq (E_X \multimap E_Y^\perp)^\perp = \{(x_a y_b)_{(a,b) \in X \times Y} \mid (x_a)_a \in E_X, (y_b)_b \in E_Y\}^{\perp\perp}.$$

The double duality ensures that the space constructed is, in fact, a Köthe space. This construction gives a monoidal closed structure on **Köthe**. For the Cartesian structure, which interprets the additive subpart of LL, we define is through direct sums. For X and Y at most countable, $X \,\&\, Y$ is the disjoint sum of X and Y. The set $E_X \,\&\, E_Y$ is defined by

$$E_X \,\&\, E_Y \triangleq \{z \in \mathbb{K}^{X \& Y} \mid \pi_1(z) \in E_X \text{ and } \pi_2(z) \in E_Y\}$$

where the projections π_1 represents the subsequence where one keeps the elements whose index is in X (and similarly for π_2 with indexes in Y).

The exponential structure on Köthe To define the exponential structure, we use some ideas from **Rel**. The multiplicative and additive constructions on the set of indices of Köthe spaces are the same as those we present for **Rel**. We then keep this idea here, and the exponential of a Köthe space will be a space of sequences indexed by finite multisets. From a Köthe space (X, E_X), and $x \in E_X$, one defines $x^!$ such that

$$(x^!)_\mu = x^\mu = \prod_{i \in X} x_i^{\mu(i)}$$

where μ is a finite multiset, and $!(X, E_X)$ by the Köthe space

$$!X = \mathcal{M}_f(X) \qquad\qquad E_{!X} = \{x^! \mid x \in E_X\}^{\perp\perp}.$$

From this, one defines the weakening as $\mathsf{w} : E_{!X} \to \mathbb{K}$ by $\mathsf{w}((x)_\mu) = x_{[]}$ where $[]$ is the empty multiset, and $\mathsf{c} : E_{!X} \to E_{!X} \otimes E_{!X}$ by $\mathsf{c}(x^!) = x^! \otimes x^! = (x^!_{\mu_1} x^!_{\mu_2})_{\mu_1, \mu_2}$.

Equivalently, the exponential of a Köthe space can be defined from entire functions.

Definition 1.3.10. *Let E_X and E_Y be two Köthe spaces. A map $h : E_X \to E_Y$ is* entire *when there is a continuous linear function $f : E_{!X} \to E_Y$ such that $h(x) = f(x^!)$ for all $x \in E_X$. In this case, f is called the* power series *defining h.*

This notion of entire function corresponds to functions that are entirely determined by their Taylor expansion, in the sense that they are equal to their Taylor series everywhere. An entire map $h : E_X \to \mathbb{K}$ can then be represented by its power series $M \in E_{!X} \multimap \mathbb{K} = E_{(!X)^\perp}$ and we have

$$h(x) = \sum_{\mu \in \mathcal{M}_f(X)} M_\mu x^\mu.$$

In this case, $M_{\{a\}}$ corresponds to $\frac{\partial f}{\partial x_a}(0)$ and M_μ to $\frac{1}{\mu!} \frac{\partial^\mu f}{\partial x_1^{\mu(a_1)} \dots \partial x_n^{\mu(a_n)}}(0)$.

Finally, the comonadic structure over $!\_$ is defined thanks to the matricial point of view. The dereliction for E_X is defined by the matrix $(\delta_{\mu,[a]})_{\mu \in !X, a \in X}$ which is the matrix with 0 everywhere except when $\mu = [a]$. The digging for E_X by the matrix $(\delta_{\mu, \sum(M)})_{\mu \in !X, M \in !!X}$ where $\sum(M)$ is the sum of multisets as defined in Definition 1.1.15.

The differential structure on Köthe The last step in the definition of **Köthe** as a model of differential linear logic is the biproduct structure to define $\bar{\mathsf{w}}$ and $\bar{\mathsf{c}}$, and the natural transformation $\bar{\mathsf{d}}$. For the coweakening $\bar{\mathsf{w}} : \mathbb{K} \to E_{!X}$, we define $\bar{\mathsf{w}}(x) = (x\delta_{[],\mu})_\mu$ the sequence equal to x when the index is the empty set and 0 otherwise. For the cocontraction $\bar{\mathsf{c}} : E_{!X} \otimes E_{!X} \to E_{!X}$,

$$\bar{\mathsf{c}}((x^!)_{\mu_1} \otimes (y^!)_{\mu_2})_\mu = \sum_{\mu_1 + \mu_2 = \mu} \binom{\mu_1 + \mu_2}{\mu_1} x_{\mu_1} y_{\mu_2} \quad \text{with} \quad \binom{\mu}{\nu} = \prod_{a \in X} \frac{\mu(a)!}{\nu(a)!(\mu(a) - \nu(a))!}$$

the binomial coefficient generalized to multisets. The natural transformation $\bar{\mathsf{d}} : E_X \to E_{!X}$ is expressed as follows:

$$\bar{\mathsf{d}}((x_a))_\mu = \begin{cases} x_a & \text{if } \mu = [a] \text{ for some } a \in X \\ 0 & \text{otherwise} \end{cases}$$

and from this definition, composing $\bar{\mathsf{d}}$ with an entire map gives, in fact, its derivative at 0.

1.3.5 A smooth model with distributions

The models that we have overlooked so far are the categories **Rel** and **Köthe**. These two models have drawbacks. In the relational model, the objects are not topological vector spaces; hence, the notion of differentiation is far from a smooth mathematical perspective. The model with Köthe spaces is made of discrete spaces, in the sense that Köthe spaces are subsets of spaces of sequences. Hence they do not provide notions of continuity, and one cannot perform usual functional analysis with this kind of space. We present here a model with smooth maps and distributions over topological vector spaces. However, this model has some drawbacks as well. It is not a model of classical linear logic: either we define it as an intuitionistic model, or as a model of polarized differential linear logic. Moreover, the model we present here in finitary: it is a model of DiLL without digging because one cannot define the interpretation of formulas of type $!!A$. A possible alternative is to define a model thanks to *bornological* spaces instead of topological ones. Such a model is defined by Blute *et al.* [BET12]. It is neither polarized nor finitary, but it models intuitionistic differential linear logic. We chose to present here a smooth model for polarized finitary DiLL.

1.3.5.1 Distribution theory

First, we denote the space of smooth functions from $\mathbb{R}^n$ to $\mathbb{R}$ by $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ or $\mathcal{E}(\mathbb{R}^n)$, where a *smooth function* is defined as differentiable with a smooth differential. If a smooth function $f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ has compact support[3], f is said to be a *test function* and the set of test functions from $\mathbb{R}^n$ to $\mathbb{R}$ is denoted by $\mathcal{C}_c^\infty(\mathbb{R}^n, \mathbb{R})$ or by $\mathcal{D}(\mathbb{R}^n)$.

Definition 1.3.11. *A* topological vector space *(or* tvs*) is a vector space E endowed with a topology such that the addition and the scalar multiplication of E are continuous for this topology. A tvs is a* locally convex topological vector space *(or* lctvs*) when each point is contained in a convex open set of the topology.*

The topology of the spaces $\mathcal{E}(\mathbb{R}^n)$ and $\mathcal{D}(\mathbb{R}^n)$ are defined through a family of semi-norms. For a precise definition see [Tre67].

In what follows, we call the dual of a lctvs E the algebraic dual of E of linear forms over E restricted to the linear forms, which are continuous with respect to the topology of E. We denote this lctvs E'. Moreover, in the context of topological vector spaces, we say that two lctvs E and F are isomorphic when there is an isomorphism φ of vector spaces between E and F such that φ and φ^{-1} are continuous. Hence, E and F are isomorphic and homeomorphic. Since $\mathcal{D}(\mathbb{R}^n)$ and $\mathcal{E}(\mathbb{R}^n)$ are lctvs (see [Tre67]), they have topological duals. These dual spaces are endowed with the topology of uniform convergence over compact sets. This topology is the one where open sets are generated by the sets $U_{K,r}$, formed with functions such that $\sup_K |f| < r$, where K is compact and $r > 0$.

Definition 1.3.12. *Let n be an integer. The space of* distributions *over $\mathbb{R}^n$ is the topological dual of $\mathcal{D}(\mathbb{R}^n)$ (the space of functions with compact support), denoted by $\mathcal{D}'(\mathbb{R}^n)$. Similarly, the space of distribution with compact support over $\mathbb{R}^n$ is the topological dual of the space of smooth functions $\mathcal{E}(\mathbb{R}^n)$, denoted by $\mathcal{E}'(\mathbb{R}^n)$.*

A distribution should be seen as a generalized function, since for $f \in \mathcal{D}(\mathbb{R}^n)$ and $g \in \mathcal{E}(\mathbb{R}^n)$, one can define two distributions T_f and T_g by

$$\left(T_f : h \in \mathcal{E}(\mathbb{R}^n) \mapsto \int fh\right) \in \mathcal{E}'(\mathbb{R}^n) \qquad \left(T_g : h \in \mathcal{D}(\mathbb{R}^n) \mapsto \int gh\right) \in \mathcal{D}'(\mathbb{R}^n). \quad (1.3)$$

[3] where the compactness is defined with the topology induced by the canonical norm on $\mathbb{R}^n$

Then, this definition rises that distributions (resp. distributions with compact support) generalize smooth functions (resp. smooth functions with compact support).

Example 1.3.13. *A very common example of distribution (with compact support) is the* Dirac distribution at 0*. This distribution, denoted by* δ_0*, is defined by* $\delta_0 : f \mapsto f(0)$*. This example will come back later.*

The right way to consider a monoidal product between functions and distributions is the notion of convolution product. This is a binary operation which can be defined between two smooth functions, two distributions and a function and a distribution with the right hypothesis on the support.

Definition 1.3.14. *Let* n *be an integer and* $f \in \mathcal{D}(\mathbb{R}^n)$*,* $g \in \mathcal{E}(\mathbb{R}^n)$*,* $\phi \in \mathcal{E}'(\mathbb{R}^n)$ *and* $\psi \in \mathcal{E}'(\mathbb{R}^n)$*. The convolution product over smooths functions or distributions is defined as follows :*

- *The convolution between two smooth functions*[4]*, one of them having compact support, is*
$$f * g : x \in \mathbb{R}^n \mapsto \int f(x-t)g(t)\mathrm{d}t \in \mathcal{D}(\mathbb{R}^n).$$
- *The convolution of a distribution with compact support and a smooth function is defined as*
$$\phi * g : x \mapsto \phi(y \mapsto g(x-y)) \in \mathcal{E}(\mathbb{R}^n).$$
- *Finally, for two distributions with compact support, the convolution* $\phi * \psi$ *is the distribution*
$$\phi * \psi : h \in \mathcal{E}(\mathbb{R}^n) \mapsto \phi(x \mapsto \psi(y \mapsto h(x+y))) \in \mathcal{E}'(\mathbb{R}^n).$$

The hypothesis on the support is made to ensure that the integral is well defined.

Some important properties of this convolution product are proved, for example, in [Hor63]. We summarize them in the following proposition.

Proposition 1.3.15. *Let* n *be an integer.*

- *A fundamental property is that the convolution product is associative and commutative : for* $f, g \in \mathcal{D}(\mathbb{R}^n)$ *and* $h \in \mathcal{E}(\mathbb{R}^n)$*,*
$$f * h = h * f \qquad (f * h) * g = f * (h * g)$$
and for $\phi, \psi \in \mathcal{E}'(\mathbb{R}^n)$ *and* $\varphi \in \mathcal{D}'(\mathbb{R}^n)$*,*
$$\phi * \varphi = \varphi * \phi \qquad (\phi * \varphi) * \psi = \phi * (\varphi * \psi).$$
- *Let* $\phi \in \mathcal{D}'(\mathbb{R}^n)$*,* $\psi \in \mathcal{E}'(\mathbb{R}^n)$ *and* $f \in \mathcal{E}(\mathbb{R}^n)$*. From the definition we have*
$$(\phi * \psi) * f = \phi * (\psi * f)$$
since
$$\begin{aligned}(\phi * \psi) * f &= x \mapsto (\phi * \psi)(y \mapsto f(x-y)) \\ &= x \mapsto \phi(z \mapsto \psi(t \mapsto f(x-z-t))) \\ &= x \mapsto \phi(z \mapsto (\psi * f)(x-z)) \\ &= \phi * (\psi * f).\end{aligned}$$

[4]It is possible to define this operation on continuous functions, but we only need to consider it on smooth functions and distributions for our purpose.

The distribution δ_0 from Example 1.3.13 is the neutral element for the convolution product. For a smooth function $f \in \mathcal{E}(\mathbb{R}^n)$,

$$\delta_0 * f(x) = \delta_0(y \mapsto f(x-y)) = f(x-0) = f(x)$$

which leads to $\delta_0 * f = f$. Moreover, since δ_0 has compact support, for each $\psi \in \mathcal{E}'(\mathbb{R}^n)$

$$(\psi * \delta_0)(f) = \psi(x \mapsto \delta_0(y \mapsto f(x+y)) = \psi(x \mapsto f(x-0)) = \psi(f)$$

implying that $\psi * \delta_0 = \psi$. Hence, δ_0 is neutral for smooth functions and for distributions.

Remark 1.3.16. In most contexts, distribution theory is introduced by defining the space of test functions $\mathcal{D}(\mathbb{R}^n)$, and the space of distributions as its topological dual. Here we also introduce the space of smooth functions and its dual, the space of distributions with compact support. The reason is that we seek for a model of DiLL, and hope that distributions can interpret exponential formulas. Since the dereliction rule states that one can consider a linear map as a non-linear one, we cannot use spaces $\mathcal{D}(\mathbb{R}^n)$ to interpret $?A$ formulas because linear maps do not have compact support (in general).

1.3.5.2 Topological prerequisites

Many topological notions have to be carefully defined and proved to give a smooth model of DiLL. We do not detail everything, but we give some basic definitions, which are helpful to partly understand the challenges here. For a more complete presentation, see [Ker18b]. Remember that, in this context, the isomorphisms mentioned are also homeomorphisms.

Definition 1.3.17. *A lctvs E is* reflexive *if the map*

$$\delta : \begin{cases} E & \to E'' \\ x & \mapsto (\delta_x : \ell \in E' \mapsto \ell(x) \in \mathbb{R}) \end{cases}$$

is an isomorphism.

Definition 1.3.18. *An* Euclidean *lctvs is a finite dimensional $\mathbb{R}$-vector space with an inner product, endowed with the topology coming from the usual norm $||\_||_2$. Each Euclidean space E is isomorphic to $\mathbb{R}^n$ where n is the dimension of E.*

The space E is isomorphic to $\mathbb{R}^n$ since finite dimensional vector spaces of the same dimension are always isomorphic and their topologies are equivalent.

Definition 1.3.19. *A* Fréchet space *(or* F-space*) is a complete metrizable lctvs. A space which is dual to a F-space is called a* DF-space*.*

Definition 1.3.20. *Let E be a lctvs and X be a Banach space (a complete normed tvs). A linear continuous map $f : E \to X$ is* nuclear *if there exists a sequence $(a_n)_n$ of E' in the closed ball unit of E', a sequence $(y_n)_n$ in the closed ball unit of X, and a sequence $(\lambda_n)_n$ in ℓ_1 such that*

$$\forall x \in E, f(x) = \sum_n \lambda_n a_n(x) y_n$$

If every linear continuous map from E to any Banach space is nuclear, we say that E is nuclear*.*

To stay at an intuitive level, the idea behind the definition of a nuclear space is that it is a space which *looks like* a normed space. Nuclear spaces, together with F and DF spaces have nice properties for our purposes that we will describe now. First, they are useful to characterize spaces of smooth functions and distributions.

Proposition 1.3.21. *For each $n \in \mathbb{N}$, the space of smooth functions $\mathcal{E}(\mathbb{R}^n)$ is a nuclear F-space, and the space of distributions with compact support $\mathcal{E}(\mathbb{R}^n)'$ is a nuclear DF-space.*

Another crucial property in our quest for a classical smooth model is the reflexivity of these spaces. This is stated in the following theorem.

Theorem 1.3.22. *A lctvs which is both nuclear and a F-space is reflexive. Similarly, a lctvs which is both nuclear and a DF-space is also reflexive.*

In addition, we will need some constructions on these spaces that will allow us to interpret connectives of linear logic as usual constructions on vector spaces.

Proposition 1.3.23 ([Jar81])**.** *The class of nuclear F-spaces is stable by Cartesian products and completed projective tensor products.*

Proposition 1.3.24 ([Gro66])**.** *The class of nuclear DF-spaces is stable by completed projective tensor products.*

Finally, the following theorem is used to prove that the functor $!\_$ that we will define is strongly monoidal.

Theorem 1.3.25 (Kernel theorem [Tre67])**.** *Let $n, m \in \mathbb{N}$. The spaces $\mathcal{E}(\mathbb{R}^{n+m})$ and $\mathcal{E}(\mathbb{R}^n) \otimes \mathcal{E}(\mathbb{R}^m)$ are isomorphic.*

1.3.5.3 Interpreting polarized differential linear logic

We denote by **Eucl** the category of Euclidean spaces with continuous linear maps. Thanks to the usual tensor product, Cartesian product and duality on topological vector spaces, **Eucl** is a $\star$-autonomous symmetric monoidal closed category with finite products. It is then a model of MALL. The $\star$-autonomy property comes from the fact that the duality on Euclidean spaces is reflexive. While this property is easy to get for finite dimensional lctvs, it is much harder for infinite dimensional ones. And it is crucial to get a *classical* model of DiLL. In general, nuclear spaces are not reflexive, but as we saw in Theorem 1.3.22 nuclear F-spaces and nuclear DF-spaces are. We then use this result to interpret exponential formulas. One important point before trying to model differential linear logic by nuclear F and DF spaces is the following remark.

Remark 1.3.26. We have stated that, in general, nuclear spaces are not reflexive. One may hope that working in the category of nuclear F or DF space may be enough to define a smooth model of classical finitary DiLL. However, the tensor product of a F-space with a DF-space is neither a F-space nor a DF-space. Hence, it may not be reflexive. This is the reason why we have to work in a polarized setting here. We separate the exponential interpretations into two subcategories of nuclear spaces, which are stable by tensor and Cartesian product.

Interpreting the exponential formulas We denote by **NuclF** (resp. **NuclDF**) the category of nuclear F-spaces (resp. nuclear DF-spaces) with linear continuous maps as morphisms. In this model, the exponential connectives are interpreted by

$$[\![!A]\!] \triangleq (\mathcal{C}^\infty([\![A]\!], \mathbb{R}))' \qquad\qquad [\![?A]\!] \triangleq \mathcal{C}^\infty([\![A]\!]', \mathbb{R})$$

and by Proposition 1.3.21 we have $[\![!A]\!] \in \mathbf{NuclDF}$ and $[\![?A]\!] \in \mathbf{NuclF}$ as we are in the finitary setting. The formulas A are interpreted by Euclidean spaces which ensure that the definitions of the well definedness of these interpretations. While being in a finitary

$$\begin{array}{rcl} A,B & \triangleq & 0 \mid 1 \mid \top \mid \bot \mid X \mid X^\perp \mid A \parr B \mid A \otimes B \mid A \,\&\, B \mid A \oplus B \\ N,M & \triangleq & A \mid N \parr M \mid N \,\&\, M \mid ?A \\ P,Q & \triangleq & A \mid P \otimes Q \mid P \oplus Q \mid !A \end{array}$$

Figure 1.12: The grammar of the formulas of classical finitary polarized DiLL

polarized setting prevents us from defining the exponential of infinite dimensional spaces, the multiplicative and additive connectives over exponentials need to be interpreted. The formal grammar for the formulas is given in Figure 1.12. The formulas A, B are interpreted in the category **Eucl**, formulas N, M in **NuclF** and formulas P, Q in **NuclDF**. The parr $\parr$ in **NuclF** is defined as the tensor product[5] and the & by the Cartesian product, which is still in **NuclF** from Proposition 1.3.23. For the positive formulas, the interpretations of $\otimes$ and $\oplus$ are made through the negation: for example, $A \otimes B = (A^\perp \parr B^\perp)^\perp$, which works thanks to the reflexivity of these spaces. Note that one can prove that $!\_$ is a strong monoidal functor, using the kernel theorem (Theorem 1.3.25)

Interpreting the proofs To interpret the proofs in this model, we strongly use the fact that the spaces we consider are reflexive. The interpretations are defined as follows:

$$\mathsf{w} : \begin{cases} \mathbb{R} & \to ?A \\ 1 & \mapsto cst_1 \end{cases} \qquad \mathsf{c} : \begin{cases} ?A \otimes ?A & \to ?A \\ f \otimes g & \mapsto f.g \end{cases} \qquad \mathsf{d} : \begin{cases} !A & \to A'' \\ \phi & \mapsto \phi_{|A'} \end{cases}$$

$$\bar{\mathsf{w}} : \begin{cases} \mathbb{R} & \to !A \\ 1 & \mapsto \delta_0 \end{cases} \qquad \bar{\mathsf{c}} : \begin{cases} !A \otimes !A & \to !A \\ \phi \otimes \psi & \mapsto \phi * \psi \end{cases} \qquad \bar{\mathsf{d}} : \begin{cases} A'' & \to !A \\ ev_x & \mapsto (f \mapsto ev_x(D_0(f))) \end{cases}$$

where $\phi_{|A'}$ is the distribution ϕ whose domain is restricted to A' and ev_x the linear map from A' to $\mathbb{R}$ such that $ev_x : \ell \to \ell(x)$ where $x \in A$. In this framework, we see that the weakening is the neutral for the contraction and the coweakening the neutral for the cocontraction, interpreted by the scalar multiplication between smooth maps and the convolution product between distributions. These operations are well suited for our interpretations, as we expressed intuitively in Section 1.3.2. The syntax of the cut elimination between the contraction and the codereliction represents exactly the differentiation of the product of functions. The interpretation of the contraction is used to prove the kernel theorem. For the cocontraction, the convolution is the natural monoidal operation on distributions. The dereliction as defined here is not exactly the rule that forgets linearity. It is expressed dually, but one can see it as a rule from A' to $?A'$, and that would correspond to forget the linearity For the codereliction, we use the fact that the spaces we consider are reflexive. Since the differential at 0 of a function is linear, the rule is well typed.

[5] It is true algebraically, but to be more precise, it is a completed projective tensor product, meaning that the topology makes some canonical maps continuous, and completed in order to get a complete space from this construction.

Chapter 2

Indexed differential linear logic

This chapter adapts results published in [BKM23].

Contents

This chapter studies differentiation in the syntax from two points of view: adding costructural rules to graded linear logic leading to $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$, or indexing differential linear logic leading to IDiLL. These two points of view are, in fact, equivalent. $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ comes from syntactical considerations, while IDiLL from semantical ones.

In Section 2.1, we first present the logic D-DiLL that has been defined by Kerjean [Ker18a] which is on the basis of this work. It extends DiLL by adding a partial differential operator to the syntax. We give the semantics associated with this logic, which is based on smooth maps and distributions.

Then, we define in Section 2.2 an extension of this system called IDiLL. The idea behind this extension is to unify two notions: gradation and differentiation. It consists of *grading* the syntax of D-DiLL, which consists of redefining the exponential rules by adding grades to the exponential connectives. These grades will be elements of a monoid of linear partial differential operators. We then give the semantics of this logic, which uses smooth maps and distributions again. This semantics is based on the interaction between differential operators and functions and distributions.

Finally, we define in Section 2.3 a graded version of D-DiLL, which we call $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$. The syntax of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ is constructed from the following idea: as adding costructural rules to LL gives rise to a logic that fits differentiation, we add costructural rules to graded linear logic. This leads to a system which fits the action of partial differential operators, with their parameters and solutions, when graded by a particular monoid. We show that this syntax is equivalent to IDiLL. We give a cut elimination procedure for $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$, which induces one on IDiLL thanks to the equivalence. We then prove that the semantics of IDiLL with differential operators is coherent with this cut elimination.

Since $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ is constructed from $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ which has a well-known dynamics, it is helpful to study the dynamics of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$. And the fact that IDiLL is close to D-DiLL which has already been semantically studied [Ker18a], helps us to provide a semantics for IDiLL.

Here, we let the promotion rule out of the picture. Considering this rule in our systems corresponds to several challenges, coming with semantical, syntactical and algebraic questions. Semantically, the promotion in linear logic is interpreted by the composition of higher order functions. In our setting, it corresponds to finding the solution of the composition of two differential operators, from the solution of each operator separately. But in the class of differential equations that we will consider, the set of solutions is not stable by composition. Moreover, interpreting partial differential equations requires a canonical basis to interpret the variables along which we partially differentiate. This corresponds to a need for a canonical basis, which clashes with higher order formulas: at higher order, the vector spaces are spaces of functions or distributions whose dimension is infinite and do not have a canonical basis contrary to spaces like $\mathbb{R}^n$. The graded point of view leads to several issues with the promotion as well. Syntactically, a logical system that is both graded and differential gives rises to a lot of possible interactions between the rules, which come with many algebraic properties on the indices. In graded linear logic, having a promotion rule corresponds to having a product operation in the underlying algebraic structure of indexes. The algebraic structure that we consider is composed of partial differential operators. Defining a product operation on this set comes with some restrictions, which is why we will limit ourselves to a monoid of operators here. All of these reasons lead us to consider a promotion free logic, and, moreover, to limit ourselves to a setting where the formulas are finitary. Partial solutions for these questions will be given in Chapters 3 and 4.

2.1 D-DiLL: a differential linear logic indexed by a differential operator

In [Ker18a], Kerjean defined a logic named D-DiLL, which mixes differential linear logic with the action of a partial differential operator D. The goal of this logic is to extend DiLL to the study of differential equations and their solutions. Such is system is developed with the following analogy. In linear logic, the dereliction is interpreted by the rule that *forgets* the linearity of a function. In differential linear logic, the codereliction corresponds to the rule that *linearizes* a function, by differentiating it. This interpretation can be expressed in other words. Considering the operation of differentiation as applying a differential operator, one can see the dereliction as solving the differential equation associated with this operator. Let ℓ be a linear map. The linearity of ℓ implies that its differential is ℓ itself. Since applying the dereliction does not change the linear map, but only its type, another way of understanding this rule is to see it as a resolution of a differential equation.

$$A, B \triangleq 0 \mid 1 \mid \top \mid \bot \mid A \otimes B \mid A \parr B \mid A \oplus B \mid A \& B$$
$$N, M \triangleq ?N \mid ?_D N \mid N \parr M \mid N \& M$$
$$P, Q \triangleq !P \mid !_D P \mid P \otimes Q \mid P \oplus Q$$

(a) The grammar of the formulas of D-DiLL

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?_D A}\ \mathsf{w}_D \qquad \frac{\vdash \Gamma, ?_D A, ?A}{\vdash \Gamma, ?_D A}\ \mathsf{c}_D \qquad \frac{\vdash \Gamma, ?_D A}{\vdash \Gamma, ?A}\ \mathsf{d}_D$$
$$\frac{}{\vdash !_D A}\ \bar{\mathsf{w}}_D \qquad \frac{\vdash \Gamma, !_D A \qquad \vdash \Delta, !A}{\vdash \Gamma, \Delta, !_D A}\ \bar{\mathsf{c}}_D \qquad \frac{\vdash \Gamma, !_D A}{\vdash \Gamma, !A}\ \bar{\mathsf{d}}_D$$

(b) The exponential rules of D-DiLL

Figure 2.1: The syntax of the logic D-DiLL

In fact, the solution of the differential equation of the differentiation with parameter ℓ is ℓ, which is the result of the application of the dereliction to ℓ.

Following this point of view, D-DiLL consists of a system analogous to DiLL. If in DiLL d solves the differential equation associated with the differentiation and $\bar{\mathsf{d}}$ applies the differential operator of differentiation, their counterparts in D-DiLL solve and apply differential equations, but for a partial differential operator.

To do so, the class of differential operators that is considered has to be restricted. We consider linear operators, as they fit particularly well with linear logic, and we ask for constant coefficients on these operators. This condition comes from considerations on solutions of differential equations that we detail in the next section. The detailed syntax of this logic is given in Figure 2.1.

2.1.1 Linear partial differential operators with constant coefficients

Definition 2.1.1. *A* linear partial differential operator (LPDO) *is an operator* $D : \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ *that can be written as*

$$D : f \mapsto \left(x \in \mathbb{R}^n \mapsto \sum_{\alpha \in \mathbb{N}^n} a_\alpha(x) \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}(x) \right) \qquad \text{with } a_\alpha \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}).$$

Such an operator is a linear partial differential operator with constant coefficients (LPDOcc) *when each a_α is a constant function.*

Some well-known examples of differential equations can be represented by LPDOcc. Those are, for instance, the heat equation or the Laplacian equation. To lighten the expressions, we will denote by ∂^α the operator

$$\partial^\alpha : f \mapsto \left(x \mapsto \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}(x) \right)$$

for each $\alpha \in \mathbb{N}^n$.

With this definition of LPDO, one can apply a LPDO to a smooth map, but it can be extended to be applied to a distribution. First, for each LPDO D, we define $\widehat{D}$ as:

$$D = \sum_{\alpha \in \mathbb{N}^n} a_\alpha \partial^\alpha \qquad \widehat{D} = \sum_{\alpha \in \mathbb{N}^n} (-1)^{|\alpha|} a_\alpha \partial^\alpha$$

and for a distribution ϕ,

$$D(\phi) : f \mapsto \phi(\widehat{D}(f)).$$

which is a distribution as well. Note that the operation $\widehat{\_} : D \mapsto \widehat{D}$ is involutive:

$$\widehat{\widehat{D}} = \sum_{\alpha \in \mathbb{N}^n} (-1)^{|\alpha|}(-1)^{|\alpha|} a_\alpha \partial^\alpha = \sum_{\alpha \in \mathbb{N}^n} (-1)^{2|\alpha|} a_\alpha \partial^\alpha = \sum_{\alpha \in \mathbb{N}^n} a_\alpha \partial^\alpha = D$$

with the same notations as above. The idea behind the use of $\widehat{D}$ instead of D comes from the fact that a distribution can be seen as a generalized function. Taking a smooth map with compact support f, one can define $T_f : g \mapsto \int fg$ which is, in fact, a distribution. One would then hope that, for each LPDO D, the definition of the extension of D to distributions would give that $D(T_f) = T_{D(f)}$. But by integration by parts, we have

$$\int_x \partial^\alpha(f)(x) g(x) \mathrm{d}x = (-1)^{|\alpha|} \int_x f(x) \partial^\alpha(g)(x) \mathrm{d}x$$

which gives

$$T_{D(f)}(g) = \int \sum_\alpha a_\alpha \partial^\alpha(f) g = \sum_\alpha a_\alpha \left(\int \partial^\alpha(f) g \right) = \sum_\alpha (-1)^{|\alpha|} a_\alpha \int f \partial^\alpha g = \int f \widehat{D}(g).$$

That shows that $T_{D(f)} = D(T_f)$ thanks to the definition of a LPDO applied to a distribution and the involutivity of $\widehat{\_}$.

The key theorem on LPDOcc comes from Hörmander. It expresses the existence of a highly general notion of solutions of differential equations: fundamental solutions.

Theorem 2.1.2 (Hörmander [Hor63])**.** *Let D be a linear partial differential equation with constant coefficients. There exists a unique distribution $\Phi_D \in (\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}))'$ such that $D(\Phi_D) = \delta_0$. This distribution is called the* fundamental solution *of D.*

Each LPDO represents a differential equation. Having a fundamental solution for an operator expresses a type of independence between finding a solution for this equation and the parameter of this equation. From a fundamental solution, one easily computes the solution for each parameter. Considering a partial differential equation (E) with constant coefficients and a parameter g

$$a_{\alpha_1} \partial^{\alpha_1} f + \cdots + a_{\alpha_n} \partial^{\alpha_n} f = g \tag{E}$$

the equation (E) is associated to the LPDOcc D

$$D = \sum_{i=1}^{n} a_{\alpha_i} \partial^{\alpha_i}$$

and the function $\Phi_D * g$ is a solution of (E), as

$$a_{\alpha_1} \partial^{\alpha_1}(\Phi_D * g) + \cdots + a_{\alpha_n} \partial^{\alpha_n} \Phi_D * g = D(\Phi_D * g) = g$$

thanks to Corollary 2.1.6 which will be introduced and proved later on in this section. We focus now on some properties of LPDOcc that will be useful to give a semantical interpretation of the logics we will introduce.

Lemma 2.1.3. *Let D be a LPDOcc, and ϕ and ψ two distributions. We then have,*

$$D(\phi * \psi) = D(\phi) * \psi = \phi * D(\psi).$$

Proof. For this proof, we first need a result about LPDOcc applied to translated functions. For each x_i, we have

$$\begin{aligned}\frac{\partial}{\partial x_i}(x \mapsto f(x+y))(z) &= \lim_{t \to 0} \frac{f(z+y+tx_i) - f(z+y)}{t} \\ &= \frac{\partial}{\partial x_i}(f)(z+y).\end{aligned}$$

This extends directly to composed derivatives: for each $\alpha \in \mathbb{N}^n$,

$$\partial^\alpha(x \mapsto f(x+y)) = x \mapsto \partial^\alpha(f)(x+y). \tag{2.1}$$

We can now prove the wanted equality. Let $D = \sum_\alpha a_\alpha \partial^\alpha$ be a LPDOcc. We have:

$$\begin{aligned}D(\phi * \psi)(f) &= \phi * \psi(\widehat{D}(f)) \\ &= \phi(x \mapsto \psi(y \mapsto \widehat{D}(f)(x+y))) \\ &= \phi(x \mapsto \psi(y \mapsto \sum_\alpha (-1)^{|\alpha|} a_\alpha \partial^\alpha(f)(x+y))) \\ &= \phi(x \mapsto \psi(\sum_\alpha (-1)^{|\alpha|} a_\alpha (y \mapsto \partial^\alpha(f)(x+y)))) \\ &= \phi(x \mapsto \psi(\sum_\alpha (-1)^{|\alpha|} a_\alpha \partial^\alpha(y \mapsto f(x+y)))) && \text{(by 2.1)} \\ &= \phi(x \mapsto D(\psi)(y \mapsto f(x+y))) \\ &= (\phi * D(\psi))(f)\end{aligned}$$

The other equality is then proved using the commutativity of the convolution

$$D(\phi * \psi) = D(\psi * \phi) = \psi * D(\phi) = D(\phi) * \psi.$$

which concludes the proof of the lemma. □

This lemma will be used again in the next section, as its main interest is to show a proposition about the monoid of LPDOcc that we study in Section 2.2.1. Note that this lemma ensures the unicity of the fundamental solution. Let us take two fundamental solutions Φ_D and Ψ_D. By Lemma 2.1.3 we have

$$\Phi_D = \Phi_D * \delta_0 = \Phi_D * D(\Psi_D) = D(\Phi_D * \Psi_D) = D(\Phi_D) * \Psi_D = \delta_0 * \Psi_D = \Psi_D.$$

In the proof of the lemma, we proved the equation 2.1. This equation is useful to prove another lemma that we will use to give a concrete semantics to D-DiLL.

Lemma 2.1.4. *Let D be a LPDOcc and ϕ a distribution. We then have,*

$$(\Phi_{\widehat{D}} * \phi)(D(f)) = \phi(f).$$

Proof. Let $D = \sum_\alpha a_\alpha \partial^\alpha$ be a LPDOcc. Using equation 2.1 from Lemma 2.1.3, we have

$$(\Phi_{\widehat{D}} * \phi)(D(f)) = \phi(x \mapsto \Phi_{\widehat{D}}(y \mapsto f(x+y)))$$

$$
\begin{aligned}
&= \phi(x \mapsto \Phi_{\widehat{D}}(y \mapsto \sum_\alpha a_\alpha \partial^\alpha(f)(x+y))) \\
&= \phi(x \mapsto \Phi_{\widehat{D}}(\sum_\alpha a_\alpha \partial^\alpha(y \mapsto f(x+y)))) && \text{(by 2.1)} \\
&= \phi(x \mapsto \hat{D}(\Phi_{\widehat{D}})(y \mapsto f(x+y))) \\
&= \phi(x \mapsto \delta_0(y \mapsto f(x+y))) && \text{(definition of } \Phi_{\widehat{D}}) \\
&= \phi(x \mapsto f(x+0)) = \phi(f)
\end{aligned}
$$

which proves the wanted equality. □

We prove similar results on the convolution of a function and a distribution.

Lemma 2.1.5. *Let D be a LPDOcc, ϕ a distribution and f a smooth function. We then have*

$$D(\phi) * f = \phi * D(f) = D(\phi * f).$$

Proof. The proof works with a similar method as for Lemma 2.1.3. We use a result analogous to equation 2.1 but suited for the convolution of a function with a distribution. Let f be a smooth function. For each x_i, we have

$$
\begin{aligned}
\frac{\partial}{\partial x_i}(y \mapsto f(x-y))(z) &= \lim_{t\to 0} \frac{f(x-z-tx_i) - f(x-z)}{t} \\
&= (-1) \lim_{t\to 0} \frac{f(x-z-tx_i) - f(x-z)}{-t} \\
&= (-1) \frac{\partial}{\partial x_i}(f)(x-z).
\end{aligned}
$$

This extends directly to composed derivatives: for each $\alpha \in \mathbb{N}^n$,

$$\partial^\alpha(y \mapsto f(x-y)) = (-1)^{|\alpha|}(y \mapsto \partial^\alpha(f)(x-y)). \tag{2.2}$$

Hence, for a LPDOcc $D = \sum_\alpha a_\alpha \partial^\alpha$, we have

$$
\begin{aligned}
\phi * D(f)(x) &= \phi(y \mapsto D(f)(x-y)) \\
&= \phi(y \mapsto \sum_\alpha a_\alpha \partial^\alpha(f)(x-y)) \\
&= \phi(\sum_\alpha a_\alpha(y \mapsto \partial^\alpha(f)(x-y))) \\
&= \phi(\sum_\alpha (-1)^{|\alpha|} a_\alpha \partial^\alpha(y \mapsto f(x-y))) && \text{(equation 2.2)} \\
&= \phi \circ \widehat{D}(y \mapsto f(x-y)) \\
&= D(\phi) * f(x)
\end{aligned}
$$

which proves the first equality. For the second one, we have

$$
\begin{aligned}
\frac{\partial(\phi * f)}{\partial x_i}(x) &= \lim_{t\to 0} \frac{\phi * f(x+tx_i) - \phi * f(x)}{t} \\
&= \lim_{t\to 0} \frac{\phi(y \mapsto f(x-y+tx_i)) - \phi(y \mapsto f(x-y))}{t} \\
&= \lim_{t\to 0} \phi\left(\frac{y \mapsto (f(x-y+tx_i) - f(x-y))}{t}\right) && \text{(linearity of } \phi)
\end{aligned}
$$

$$
\begin{aligned}
&= \phi\left(y \mapsto \lim_{t\to 0} \frac{f(x-y+tx_i) - f(x-y)}{t}\right) && \text{(continuity of } \phi\text{)}\\
&= \phi\left(y \mapsto \frac{\partial f}{\partial x_i}(x-y)\right)\\
&= \phi * \left(\frac{\partial f}{\partial x_i}\right)(x)
\end{aligned}
$$

which extends easily to

$$\partial^\alpha(\phi * f) = \phi * (\partial^\alpha f)$$

for each $\alpha \in \mathbb{N}^n$. Thanks to the linearity of ϕ, we then have

$$D(\phi * f) = \sum_\alpha a_\alpha(x \mapsto \phi(y \mapsto \partial^\alpha(f)(x-y))) = x \mapsto \phi(y \mapsto D(f)(x-y)) = \phi * D(f)$$

that proves the second equality and concludes the proof of the lemma. □

Applying the previous lemma to the fundamental solution of D gives the following corollary.

Corollary 2.1.6. *For a LPDOcc D and a smooth map f, we have*

$$D(\Phi_D * f) = f = \Phi_D * D(f).$$

2.1.2 Spaces of solutions and parameters as exponential formulas

In order to define a concrete model for D-DiLL, we base ourselves on a concrete model of DiLL. Since their syntax is quite similar, it should be possible to refine models of DiLL to get models of D-DiLL. Moreover, since we want to take into account the action of the LPDO D, we will take a smooth model to fully reflect the behavior of D. Two natural models can be used: the bornological one introduced by Blute *et al.* [BET12] or the topological one presented by Kerjean [Ker18a]. We use the second one as a basis for the concrete semantics of D-DiLL, as distribution theory is well suited to study linear partial differential equations (see Hörmander [Hor63]). We have presented this model in the context of differential linear logic in Section 1.3.5.

Considering DiLL, one interprets exponential formulas in the smooth semantics by sets of smooth functions and distributions

$$[\![!A]\!] \triangleq \mathcal{C}^\infty([\![A]\!], \mathbb{R})' \qquad\qquad [\![?A]\!] \triangleq \mathcal{C}^\infty([\![A]\!]', \mathbb{R})$$

which are linear topological dual to each other. In D-DiLL, some exponential formulas are added to the grammar of the logic: the formulas $!_D A$ and $?_D A$ (see Figure 2.1a). These new formulas come from the analogy that we made at the beginning of this section about solutions and parameters of differential equations. Through this analogy, we want that the dereliction $\mathsf{d} : A \to ?A$ corresponds to solve a differential equation and the codereliction $\bar{\mathsf{d}} : A \to !A$ to apply a differential operator. This analogy cannot be made directly, as there is a mismatch in terms of types. For the dereliction, for example, in differential linear logic, one can only forget the linearity of a linear map, but solving a linear partial differential equation can be done from a non-linear parameter. For the codereliction, applying a linear partial differential operator does not necessarily lead to a linear map. These two new exponential types settle the mismatch. The dereliction is extended to a map $\mathsf{d}_D : ?_D A \to ?A$ and the codereliction to a map $\bar{\mathsf{d}}_D : !_D A \to !A$.

Remark 2.1.7. Note that here there is a choice made on the types of the dereliction and the codereliction. We will make a different choice in the next section when we try to unify differential and graded logics. To be closer to the ideas of graded linear logic, we will have to reverse those types, but keep their semantical interpretation.

The formula $?_D A$ represents the set of functional parameters of the differential equation associated to D and $!_D A$ the set of distributional solutions of this differential equation. Formally, this gives

$$!_D A \triangleq D(\mathcal{C}^\infty(\llbracket A \rrbracket, \mathbb{R}))' \qquad\qquad ?_D A \triangleq D(\mathcal{C}^\infty(\llbracket A \rrbracket', \mathbb{R})).$$

Here, the space $!_D$ represents the space of *distributional solutions* of $\widehat{D}$ in the sense that for a distribution $\phi \in !_D A$,

$$\widehat{D}(\phi) : f \in ?A \mapsto \widehat{D}(\phi)(f) = \phi(\widehat{\widehat{D}}(f)) = \phi(D(f))$$

which is a well-defined distribution in $?A$ as ϕ applies to maps of the form $D(f)$. Note that replacing D by D_0 the differential at 0 in these definitions gives back the interpretation in the context of DiLL.

Remark 2.1.8. One can notice that, as differential equations always have solutions in our case, the space of solutions $\llbracket ?_D A \rrbracket$ is *isomorphic* to the function space $\llbracket ?A \rrbracket$. The isomorphism in question $f \mapsto \Phi_D * f$ is given by the fundamental solution. The same goes for the distribution spaces: $\llbracket !_D A \rrbracket$ is isomorphic to $\llbracket !A \rrbracket$ through $\phi \mapsto \widehat{D}(\phi)$.

About the order of the LPDOcc A specific issue in these semantical definitions has been hidden so far. In Definition 2.1.1, each LPDO D is associated to an integer which represents the dimension of the domain of the maps on which D can be applied. But the definitions of $?_D A$ and $!_D A$ have been made regardless of this dimension. This comes from an abuse of notation on D.

Let $D = \sum_\alpha a_\alpha \partial^\alpha$ be a LPDOcc on $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ and $f \in \mathcal{C}^\infty(\mathbb{R}^m, \mathbb{R})$. If $n \leq m$, we identify each n-tuple $\alpha = (\alpha_1, \dots, \alpha_n)$ to the m-tuple padded by zeros $(\alpha_1, \dots, \alpha_n, 0, \dots, 0)$ when we *apply* D to f. If $n \geq m$ we restrict the n-tuple to its first m elements.

Polarity of the formulas The formulas of D-DiLL are polarized. This is for the same reasons as for the smooth model from Section 1.3.5. This logic is guided by this model, as one wants it to represent well the interactions between a LPDOcc and smooth functions and distributions. To define it as a model of differential linear logic one has to consider a polarized finitary version of DiLL, so the formulas in D-DiLL are polarized and finitary as well. The exponential formulas must be separated between the nuclear F-spaces and the nuclear DF-spaces in order to define the interpretation of MALL connectives over them, and to keep a classical model where the spaces are reflexive. Hence, we have the same polarities here for the formulas of DiLL. Moreover, the new formulas made of $!_{D\_}$ and $?_{D\_}$ are polarized as well. For an Euclidean space E, the space $D(\mathcal{C}^\infty(E, \mathbb{R}))$ is in **NuclF** (as a subspace of $\mathcal{C}^\infty(E, \mathbb{R})$) and by duality $(D(\mathcal{C}^\infty(E, \mathbb{R})))'$ is in **NuclDF**. Hence, the formulas $?_D A$ are considered negative, while the formulas $!_D A$ are considered positive. And the constructions of MALL can be applied while leaving this model classical by keeping the spaces reflexive. Finally, the Seely isomorphisms of $!_\_$ can be extended to $!_{D\_}$.

Theorem 2.1.9. *The mapping* $!_{D\_} : \mathbf{Eucl} \to \mathbf{NuclF}$ *where*

$$!_{D\_} : \begin{cases} A \in \mathbf{Eucl} & \mapsto \quad D(\mathcal{C}^\infty(A, \mathbb{R}))' \\ \ell \in \mathbf{Eucl}(A, B) & \mapsto \quad (\phi \in !_D A \mapsto (f \mapsto \phi(f \circ \ell))) \end{cases}$$

is a functor. Moreover, there are natural isomorphisms

$$m^2_{D,A,B} : !_D A \otimes !B \to !_D (A \,\&\, B).$$

Proof sketch. In this context, $!_D A \otimes !_D B$ and $!_D(A \& B)$ are not isomorphic. It comes from what we explained earlier about the order of the LPDOcc. As when a LPDOcc applies to a space of dimension higher than its order, it is considered padded by zeros, $!_D(A \,\&\, B)$ can be considered as the space where the operator does not apply to the variables coming from B. This is not completely formal, as the case where the dimension is lower than the order of D is not considered. To improve that, the order should be added to the index of the exponential, but that would make this calculus much heavier. To prove the theorem, one combines this idea of padding with zeros, together with the kernel theorem (Theorem 1.3.25). □

2.1.3 The exponential rules of the logic D-DiLL

In D-DiLL, each exponential rule has an interpretation in terms of smooth functions and distributions, which takes into account the action of a LPDOcc. These rules represent the behavior of the new connectives $!_{D\_}$ and $?_{D\_}$, with their interactions with the connectives $!_\_$ and $?_\_$. Their interpretations are the following:

- The weakening w_D is interpreted by:

$$\mathsf{w}_D : \begin{cases} \mathbb{R} & \to D(\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})) \\ 1 & \mapsto D(cst_1) \end{cases}$$

  which follows the interpretation of the weakening of DiLL, but turning it into a function which is, in fact, a parameter of the differential equation associated to D.

- The coweakening $\bar{\mathsf{w}}_D$ is interpreted by:

$$\bar{\mathsf{w}}_D : \begin{cases} \mathbb{R} & \to !_D A \\ 1 & \mapsto \Phi_{\widehat{D}} * \delta_0 = \Phi_{\widehat{D}} \end{cases}$$

  as in differential linear logic, but making it a solution of the differential equation associated to D thanks to the convolution with the fundamental solution.

- The contraction c_D is interpreted by:

$$\mathsf{c}_D : \begin{cases} ?_D A \otimes ?A & \to ?_D A \\ f \otimes g & \mapsto f.g \end{cases}$$

  the pointwise product, as in the smooth model. The well typedness comes from the question of the order of a LPDOcc and uses the natural isomorphism from Theorem 2.1.9.

- The cocontraction $\bar{\mathsf{c}}_D$ is interpreted by:

$$\bar{\mathsf{c}}_D : \begin{cases} !_D A \otimes !A & \to !_D A \\ \phi \otimes \psi & \mapsto \phi * \psi \end{cases}$$

  which is the same as the interpretation of this rule in the smooth model of DiLL. This rule is well typed since

$$\phi * \psi(D(f)) = \phi(x \mapsto \psi(y \mapsto D(f)(x+y)))$$

  and ψ applies to maps of the form $D(f)$.

- The dereliction d_D is interpreted by:

$$\mathsf{d}_D : \begin{cases} !A & \to !_D A \\ \phi & \mapsto \Phi_{\widehat{D}} * \phi \end{cases}$$

which consists, in fact, in solving the differential equation associated to D by using the fundamental solution $\Phi_{\widehat{D}}$. It is well typed, since the distribution $\Phi_{\widehat{D}} * \phi$ applies to functions of the form $D(f)$ with $f \in ?A$ since

$$\Phi_{\widehat{D}} * \phi(D(f)) = \phi(f)$$

by Lemma 2.1.4.

- The codereliction $\bar{\mathsf{d}}_D$ is interpreted by:

$$\bar{\mathsf{d}}_D : \begin{cases} !_D A & \to !A \\ \phi & \mapsto \widehat{D}(\phi) \end{cases}$$

which applies the differential operator D to the distribution. It is well typed as

$$\widehat{D}(\phi)(f) = \phi(\widehat{\widehat{D}}(f)) = \phi(D(f))$$

and ϕ applies to functions of the form $D(f)$ being in $!_D A$.

Note that the dereliction and the codereliction, which represent the interaction between the new spaces $!_D A$ with the usual ones $!A$ are interpreted by the isomorphisms between these spaces, given in Remark 2.1.8.

2.1.4 The cut elimination procedure of D-DiLL and its computational content

The cut elimination procedure of D-DiLL is very similar to the one of DiLL. Syntactically, the (co)weakening and the (co)dereliction give to the connectives $!_D$ and $?_D$ the same structure as ! and ?. Some technical issues appear with the (co)contractions, since the mix indexed and non-indexed connectives. But these issues are solved by applying almost the same rewriting as in DiLL. However, the interest of this procedure arose with the semantics. Considering again the smooth model, the procedure represents, in fact, some crucial equalities about LPDOcc and smooth functions and distributions.

- The cut between w_D and $\bar{\mathsf{w}}_D$ corresponds to

$$\Phi_{\widehat{D}}(D(cst_1)) = \widehat{D}(\Phi_{\widehat{D}})(cst_1) = \delta_0(cst_1) = 1$$

- The cut between c_D and $\bar{\mathsf{w}}_D$ corresponds to

$$\Phi_{\widehat{D}}(D(f).g) = \Phi_{\widehat{D}}(D(f.g)) = \delta_0(f.g) = f(0).g(0) = \Phi_{\widehat{D}}(D(f)).\delta_0(g)$$

where the first equality is based on the Seely isomorphism and the second one uses Lemma 2.1.4.

- The cut between w_D and $\bar{\mathsf{c}}_D$ corresponds to

$$\begin{aligned} \phi * \psi(D(cst_1)) &= \widehat{D}(\phi * \psi)(cst_1) = (\widehat{D}(\phi) * \psi)(cst_1) \\ &= \widehat{D}(\phi)(cst_{\psi(cst_1)}) = \widehat{D}(\phi)(cst_1).\psi(cst_1) = \phi(D(cst_1)).\psi(cst_1) \end{aligned}$$

using again lemmas on distributions and LPDOcc together with the definition of the convolution product.

- The cut between d_D and $\bar{\mathsf{d}}_D$ corresponds to

$$\widehat{D}(\phi)(f) = \phi(D(f))$$

which is the definition of D applied to a distribution.

- The cut between c_D and $\bar{\mathsf{c}}_D$ is the same as the one for the smooth model in DiLL as the interpretation of the contraction is still the pointwise product, and the one of the cocontraction is still the convolution product.

2.2 IDiLL: a semantical gradation of differential linear logic

As we saw in the previous section, LPDOcc are well suited to be incorporated into differential linear logic, in order to extend it to differential operators. As Theorem 2.1.2 ensures the existence of a fundamental solution, one can give a semantics for the dereliction as a resolution of a differential equation. But this class of operators has other interesting properties. In particular, one that connects the convolution of fundamental solutions and the composition of differential operators. For each LPDOcc D_1 and D_2, we have

$$\Phi_{D_1} * \Phi_{D_2} = \Phi_{D_1 \circ D_2}.$$

Since the convolution of distribution is the interpretation of the cocontraction in the semantics of D-DiLL, one may try to use the former equality to generalize D-DiLL into a system where several differential operators are expressed. In such a logic, the cocontraction rule would then produce an exponential formula indexed not only by a single LPDOcc, but by the composition of two LPDOcc.

$$\frac{\vdash \Gamma, !_{D_1} A \quad \vdash \Delta, !_{D_2} A}{\vdash \Gamma, \Delta, !_{D_1 \circ D_2} A}\ \bar{\mathsf{c}}$$

Note that, in the case of the smooth semantics of D-DiLL, we have that $[\![!_{id} A]\!] = [\![!A]\!]$ where id is the identity differential operator. In this sense, this new contraction rule is an actual generalization of the one from D-DiLL. This composition at the level of the indices and the importance of the identity operator hint at a logic with an underlying monoid of differential operators. We will see in this section how to define such a logic, and that bringing some symmetry compared to D-DiLL simplifies both the syntax and the semantics. We describe in Section 2.2.1 the behavior of the monoid of LPDOcc that we use, and we then define the semantical interpretation of the exponential formulas in Section 2.2.2 and the one of the rules in Section 2.2.2 based on the model with smooth functions and distributions.

2.2.1 A monoid of linear partial differential operators with constant coefficients

We first study a natural algebraic structure for LPDOcc.

Proposition 2.2.1. *Let D_1 and D_2 be two LPDOcc, and Φ_{D_1}, Φ_{D_2} their respective fundamental solution. Then we have the following equality*

$$\Phi_{D_1} * \Phi_{D_2} = \Phi_{D_1 \circ D_2}.$$

The proof of this proposition uses the following lemma.

Lemma 2.1.3. *Let D be a LPDOcc, and ϕ and ψ two distributions. We then have,*

$$D(\phi * \psi) = D(\phi) * \psi = \phi * D(\psi).$$

From this lemma that we have proved in the previous section, we can prove the equality between the convolution of fundamental solutions and the fundamental solution of the convolution.

Proof of Proposition 2.2.1. Let D_1 and D_2 be two LPDOcc, Φ_1 and Φ_2 their fundamental solution. We have

$$\begin{aligned}
(D_1 \circ D_2)(\Phi_1 * \Phi_2) &= \Phi_1 * ((D_1 \circ D_2)(\Phi_2)) && \text{(Lemma 2.1.3)}\\
&= \Phi_1 * (D_1(D_2(\Phi_2))) \\
&= \Phi_1 * D_1(\delta_0) && \text{(definition of } \Phi_2)\\
&= D_1(\Phi_1 * \delta_0) && \text{(Lemma 2.1.3)}\\
&= D_1(\Phi_1) \\
&= \delta_0 && \text{(definition of } \Phi_1)
\end{aligned}$$

which proves that $\Phi_1 * \Phi_2$ is the fundamental solution of $D_1 \circ D_2$. □

Thanks to this proposition, we will define a logic where the exponential are indexed by LPDOcc.

Proposition 2.2.2. *Let $\mathcal{D}$ be the set of LPDOcc, and id the identity operator. Then $(\mathcal{D}, \circ, id)$ is a commutative monoid.*

There are two possible ways to prove this proposition. A direct proof uses Schwarz's theorem, which states that partial derivatives are commutative. Another way is to notice an isomorphism between LPDOcc endowed with composition and multivariate polynomials endowed with the polynomial product

$$\psi : \begin{cases} \mathcal{D} & \to \mathbb{R}[X_1, \ldots, X_n, \ldots] \\ \sum_\alpha a_\alpha \partial^\alpha & \mapsto \sum_\alpha a_\alpha X_1^{\alpha_1} \ldots X_n^{\alpha_n}. \end{cases}$$

Proposition 2.2.2 becomes then a consequence of the commutativity of the polynomial product. This isomorphism between differential operators and polynomials will be crucial, in particular in Chapter 3.

2.2.2 The spaces of solutions and parameters

For the interpretation of the exponential formulas, we follow intuitions from D-DiLL. However the generalization of the (co)derelictions leads to a switch of point of view regarding how to interpret $!_D A$ and $?_D A$. Following the idea that $!A$ and $?A$ in D-DiLL corresponds to $!_{id} A$ and $?_{id} A$ and that the generalization of IDiLL consists of replacing the identity operator by a general LPDOcc, one would naively generalize d_D and $\bar{\mathsf{d}}_D$ to

$$\frac{\vdash \Gamma, ?_{D_1 \circ D_2} A}{\vdash \Gamma, ?_{D_1} A}\,\mathsf{d} \qquad\qquad \frac{\vdash \Gamma, !_{D_1 \circ D_2} A}{\vdash \Gamma, !_{D_1} A}\,\bar{\mathsf{d}}$$

in IDiLL. This clashes with intuitions from graded linear logic. Such rules would correspond to forget about the operators that have been applied, while graded linear logic keeps track of the information added in the proof. We then take a dual point of view in order to stick

$$A, B \triangleq 0 \mid 1 \mid \top \mid \bot \mid A \otimes B \mid A \parr B \mid A \oplus B \mid A \,\&\, B$$
$$N, M \triangleq ?_D N \mid N \parr M \mid N \,\&\, M$$
$$P, Q \triangleq !_D P \mid P \otimes Q \mid P \oplus Q$$

(a) The grammar of the formulas of IDiLL

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?_{id}A}\ \mathsf{w} \qquad \frac{\vdash \Gamma, ?_{D_1}A, ?_{D_2}A}{\vdash \Gamma, ?_{D_1 \circ D_2}A}\ \mathsf{c} \qquad \frac{\vdash \Gamma, ?_{D_1}A}{\vdash \Gamma, ?_{D_1 \circ D_2}A}\ \mathsf{d}_I$$

$$\frac{}{\vdash !_{id}A}\ \bar{\mathsf{w}} \qquad \frac{\vdash \Gamma, !_{D_1}A \qquad \vdash \Delta, !_{D_2}A}{\vdash \Gamma, \Delta, !_{D_1 \circ D_2}A}\ \bar{\mathsf{c}} \qquad \frac{\vdash \Gamma, !_{D_1}A}{\vdash \Gamma, !_{D_1 \circ D_2}A}\ \bar{\mathsf{d}}_I$$

(b) The exponential rules of IDiLL

Figure 2.2: The syntax of the logic IDiLL

with the intuitions of the graded point of view. The new dereliction and coderelicion have then the form

$$\frac{\vdash \Gamma, ?_{D_1}A}{\vdash \Gamma, ?_{D_1 \circ D_2}A}\ \mathsf{d}_I \qquad \frac{\vdash \Gamma, !_{D_1}A}{\vdash \Gamma, !_{D_1 \circ D_2}A}\ \bar{\mathsf{d}}_I.$$

Note that these rules are denoted here d_I and $\bar{\mathsf{d}}_I$ as *indexed dereliction* and *indexed codereliction* as they have the same form as the similar rules in $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$.

This change of perspective leads to changes in the semantical interpretation of the spaces $!_D A$ and $?_D A$. We want to keep the idea that d_I solves a differential equation and $\bar{\mathsf{d}}_I$ applies a differential operator. They will then be interpreted as in the smooth model of D-DiLL described in Section 2.1.3. In order to have the following semantics

$$\mathsf{d}_I : \begin{cases} ?_{D_1}A & \to ?_{D_1 \circ D_2}A \\ f & \mapsto \Phi_{D_2} * f \end{cases} \qquad \bar{\mathsf{d}}_I : \begin{cases} !_{D_1}A & \to !_{D_1 \circ D_2}A \\ \phi & \mapsto D_2(\phi) \end{cases}$$

we then define the semantics of the exponential formulas as

$$[\![!_D A]\!] = \left\{ \widehat{D}(\phi) \mid \phi \in \left(\mathcal{C}^\infty([\![A]\!]', \mathbb{R})\right)' \right\} = \widehat{D}([\![!A]\!])$$

and

$$[\![?_D A]\!] = \{ g \in \mathcal{C}^\infty([\![A]\!], \mathbb{R}) \mid \exists f \in \mathcal{C}^\infty([\![A]\!], \mathbb{R}),\ D(g) = f \} = D^{-1}([\![?A]\!]).$$

With these dual definitions, the space $!_D A$ represents then the distributional parameters of the differential equation associated with D and $?_D A$ the space of functional solutions of this equation. In this context, we highlight a point similar to Remark 2.1.8 for D-DiLL.

Remark 2.2.3. One can notice that, as linear partial differential equations with constant coefficients always have solutions, the space of solutions $[\![?_D A]\!]$ is *isomorphic* to the function space $[\![?A]\!]$. The isomorphism is not the indexed dereliction, but its dual $f \mapsto D(f)$. For the space of distributional parameters, we also have a map dual to the indexed codereliction given by $\phi \mapsto \Phi_{\widehat{D}} * \phi$ which is an isomorphism between $[\![!_D A]\!]$ and $[\![!A]\!]$. While our setting might be seen as too simple from the point of view of analysis, it is a first and necessary step before extending IDiLL to more intricate differential equations. If we were to explore the abstract categorical setting for our model, these isomorphisms would be relevant in a *bicategorical* setting.

Notation 2.2.4. We sometime remove the brackets on the semantical interpretation of a formula to lighten the expressions. To avoid confusion, we will use $!A, ?_D B, \dots$ to represent the exponential formulas and $!E, ?_D F, \dots$ for the mathematical spaces interpreting the formulas.

2.2.3 The smooth semantics of IDiLL

We define a concrete model of IDiLL based on smooth maps and distributions, together with their interaction with LPDOcc. This model is, of course, highly similar to the smooth model of DiLL from Section 1.3.5 and to the one with LPDOcc of D-DiLL from Section 2.1.3. There are, however, some differences, due to the gradation and to the fact that we had to twist the interpretation of the exponential formulas.

- The weakening is interpreted by

$$\mathsf{w} : \begin{cases} \mathbb{R} & \to ?_{id}A \\ 1 & \mapsto cst_1 \end{cases}$$

  which is the same interpretation as in the smooth model of DiLL. Since the weakening in IDiLL is the neural element for the composition, it is not indexed by a general LPDOcc as in D-DiLL but by the identity operator. Its interpretation is then also the same as in D-DiLL, with $D = id$.

- The coweakening is interpreted by

$$\bar{\mathsf{w}} : \begin{cases} \mathbb{R} & \to !_{id}A \\ 1 & \mapsto \delta_0 \end{cases}$$

  Similarly to the weakening, this is the same interpretation as in DiLL and as in D-DiLL with $D = id$.

- The contraction is interpreted by

$$\mathsf{c} : \begin{cases} ?_{D_1}A \otimes ?_{D_2}A & \to ?_{D_1 \circ D_2}A \\ f \otimes g & \mapsto \Phi_{D_1 \circ D_2} * (D_1(f).D_2(g)) \end{cases} \tag{2.3}$$

  Here, the interpretation changes from DiLL and D-DiLL. We still use the pointwise product, but we also use the operators D_1 and D_2 and the fundamental solution of $D_1 \circ D_2$. This is due to the fact that, for $f \in ?_{D_1}A$ and $g \in ?_{D_2}A$, we cannot ensure that $f.g$ is in $?_{D_1 \circ D_2}A$. In fact, by definition $D_1(f)$ and $D_2(g)$ are in $?A$, but this does not say anything about $D_1 \circ D_2(f.g)$ since the composition of LPDOcc does not behave well with the product of functions. We then use the fundamental solution to have

$$D_1 \circ D_2 \left(\Phi_{D_1 \circ D_2} * (D_1(f).D_2(g))\right) = D_1(f).D_2(g) \in [\![?A]\!]$$

  by Corollary 2.1.6, which proves that $\mathsf{c}(f \otimes g) \in [\![?_{D_1 \circ D_2}A]\!]$ and the well typedness of the interpretation.

- The cocontraction is interpreted by

$$\bar{\mathsf{c}} : \begin{cases} !_{D_1}A \otimes !_{D_2}A & \to !_{D_1 \circ D_2}A \\ \phi \otimes \psi & \mapsto \phi * \psi \end{cases}$$

which is the same as for DiLL and D-DiLL. It works thanks to Proposition 2.2.1. A distribution ϕ in $!_{D_1}A$ has the form $\Phi_{\widehat{D_1}} * \phi'$ with $\phi' \in ?A$. Similarly, ψ is $!_{D_2}A$ has the form $\Phi_{\widehat{D_2}} * \psi'$ with $\psi' \in !A$. The interpretation of the rule gives then

$$(\Phi_{\widehat{D_1}} * \phi') * (\Phi_{\widehat{D_2}} * \psi') = \Phi_{\widehat{D_1}} * \Phi_{\widehat{D_2}} * (\phi' * \psi') = \Phi_{\widehat{D_1 \circ D_2}} * (\phi' * \psi')$$

which is in $[\![!_{D_1 \circ D_2} A]\!]$ since $\phi' * \psi' \in [\![!A]\!]$.

- The indexed dereliction is interpreted by

$$\mathsf{d}_I : \begin{cases} ?_{D_1}A & \to ?_{D_1 \circ D_2}A \\ f & \mapsto \Phi_{D_2} * f \end{cases}$$

  which solves the differential equation of D_2. It is well typed as $D_1 \circ D_2(\Phi_{D_2} * f) = D_1(f)$ which is in $[\![?A]\!]$ since $f \in [\![?_{D_1}A]\!]$.

- The indexed coderelicion is interpreted by

$$\bar{\mathsf{d}}_I : \begin{cases} !_{D_1}A & \to !_{D_1 \circ D_2}A \\ \phi & \mapsto \widehat{D_2}(\phi) \end{cases}$$

  and is well-typed by definition.

2.3 DB$_{\mathcal{M}}$LL: a syntactical differentiation of first-order graded linear logic

Adapting B$_{\mathcal{S}}$LL to study the dynamics of IDiLL makes sense for several reasons. The logic IDiLL is built from semantical considerations, using the formalism of DiLL together with an analogy between LPDOcc and the operator of differentiation. It generalizes D-DiLL using the fact that the composition of LPDOcc and the codereliction rule are highly related (thanks to Proposition 2.2.1). The necessity of the composition, as the operation on the indexes for the contraction, naturally leads to compare it with the sum of indexes in B$_{\mathcal{S}}$LL. Naively, one could hope to define a semiring of LPDOcc, having the composition of operators as the sum of the semiring. This would be important in the study of the dynamics of IDiLL, as the one of B$_{\mathcal{S}}$LL is well known. However, this appears to be a complicated question that will be detailed in Chapters 3 and 4.

We define the logic DB$_{\mathcal{M}}$LL with its grammar and its rules in Figure 2.3. It is parametrized by a monoid $\mathcal{M}$, which has several properties that will be detailed in the next section. Its grammar is the same as in graded linear logic, and its set of exponential rules consists of six rules. There are the structural rules of B$_{\mathcal{S}}$LL where we have removed the promotion and the dereliction. This is because with a promotion, the underlying algebraic structure has to be strongly restricted and this clashes with a potential semiring of LPDOcc. This is studied in Chapter 4. Removing the promotion implies restricting ourselves to a monoid, removing the product operation. Then the dereliction is also removed, as it would be indexed by the neutral element of the product that we do not consider here. To these three structural rules, we add three costructural ones by dualizing the first rules. The logic DB$_{\mathcal{M}}$LL is the syntactical differential of first order B$_{\mathcal{S}}$LL in this sense. DiLL brings differentiation in the syntax by dualizing the structural rules of LL and the same process is done here.

In Section 2.3.1 we study the notion of monoid that we will need in the definition of DB$_{\mathcal{M}}$LL, which contains a particular algebraic decomposition. We prove that the monoid

$$A, B \triangleq 0 \mid 1 \mid \top \mid \bot \mid A \otimes B \mid A \parr B \mid A \oplus B \mid A \,\&\, B \mid ?_x A \mid !_x A \qquad (x \in \mathcal{M})$$

(a) The grammar of the formulas of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?_0 A}\ \mathsf{w} \qquad \frac{\vdash \Gamma, ?_x A, ?_y A}{\vdash \Gamma, ?_{x+y} A}\ \mathsf{c} \qquad \frac{\vdash \Gamma, ?_x A \quad x \le y}{\vdash \Gamma, ?_y A}\ \mathsf{d}_I$$

$$\frac{}{\vdash !_0 A}\ \bar{\mathsf{w}} \qquad \frac{\vdash \Gamma, !_x A \quad \vdash \Delta, !_y A}{\vdash \Gamma, \Delta, !_{x+y} A}\ \bar{\mathsf{c}} \qquad \frac{\vdash \Gamma, !_x A \quad x \le y}{\vdash \Gamma, !_y A}\ \bar{\mathsf{d}}_I$$

(b) The exponential rules of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$

Figure 2.3: The syntax of the logic $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$

of LPDOcc endowed with the composition satisfies this property. In Section 2.3.2 we detail the cut elimination procedure of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ and prove its convergence. We show that $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ and IDiLL are, in fact, equivalent logics which gives us a cut elimination procedure for IDiLL. We then study the semantical side, defining how to extend the relational model to $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ in Section 2.3.3 and proving that the smooth semantics of IDiLL is coherent with its cut elimination in Section 2.3.4.

2.3.1 Monoidal considerations

In the syntax of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, each exponential rule corresponds to axioms of the definition of the ordered semiring $\mathcal{S}$, as indexes of the exponential connectives. The weakening introduces 0, the neutral element for the sum. The sum itself comes with the contraction. Similarly, the dereliction introduces 1, the neutral of the product, which is in the promotion. Finally, the indexed dereliction uses the order. The axioms of an ordered semiring appear in various ways. For example, the commutativity of the sum is mandatory with the definition of the contraction, while the distributivity of the product over the sum is needed to define a cut elimination procedure.

These considerations on the axioms of a semiring will be detailed in Chapter **??**, as we will need to refine them. However, the algebraic structure that we need here is a bit different. First, as we let the promotion out of the picture in this chapter, the product rule do not need to be considered. Hence, its neutral element as well, since its algebraic content needs to be related to a product. Moreover, as explained in Section 2.2, the computational content of the dereliction and of the codereliction appears through other rules with LPDOs: the indexed dereliction and the indexed coderelicion. This also explains why the neutral of the product is not mandatory, as we do not need the dereliction in this setting.

This leaves us with three rules: the weakening, the contraction and the indexed dereliction, together with their differential counterparts. The algebraic structure that we look for has then three elements in its definition: a sum, with its neutral element, and an order. We require the sum to be commutative, as in $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$. This is mandatory, since we work implicitly with the mix rule, we do not want to focus on the order of the rules in c or in $\bar{\mathsf{c}}$, which are the rules where we use the sum. In the cut elimination of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$, some properties on the sum will be needed. We present them in the next section.

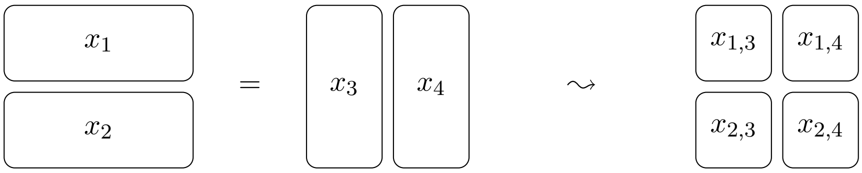
2.3.1.1 On a decomposition over the monoidal operation

Apart from the fact that we do not require to have a product for the indexes, the huge difference between $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ and $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ comes from the interaction between the new rules. In $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, the interactions between the exponential rules are the interactions between the promotion and the other exponential rules [BP15]. In $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$, there is no such interaction, since there is no promotion rule, but the structural rules interact with the costructural ones. For most of these interactions, one can easily adapt what has been done for DiLL, as we will see in Section 2.3.2. But one particular case requires a more careful study: the cut between a contraction and a cocontraction. To perform the elimination of such a cut in our setting, the algebraic structure underlying the indexes will need to have a particular property: the *additive splitting*.

Definition 2.3.1. *A commutative monoid* $(\mathcal{M}, +, 0)$ *is* additive splitting *when for each* $x_1, x_2, x_3, x_4 \in \mathcal{M}$ *such that* $x_1 + x_2 = x_3 + x_4$*, there exist* $x_{1,3}, x_{1,4}, x_{2,3}, x_{2,4} \in \mathcal{M}$ *such that*

$$x_1 = x_{1,3} + x_{1,4} \qquad x_2 = x_{2,3} + x_{2,4} \qquad x_3 = x_{1,3} + x_{2,3} \qquad x_4 = x_{1,4} + x_{2,4}.$$

This property can be graphically represented by the following drawing



where we represent the elements $x_{i,j}$ as the *intersection* (in an informal meaning) of the elements x_i and x_j.

This property appears in a work by Carraro *et al.* [CES10], but at the semantical level. By studying enrichments of the relational model, they need the additive splitting property on the underlying algebraic structure.

Example 2.3.2. *The monoid* $(\mathbb{N}, +, 0)$ *is additive splitting.*

Proof. Let $x_1, x_2, x_3, x_4 \in \mathbb{N}$ such that $x_1 + x_2 = x_3 + x_4$. We define

$$x_{1,3} = \min(x_1, x_3) \qquad x_{1,4} = x_1 - x_{1,3} \qquad x_{2,3} = x_3 - x_{1,3} \qquad x_{2,4} = x_2 - x_{2,3}.$$

The elements $x_{1,3}, x_{1,4}$ and $x_{2,3}$ are in $\mathbb{N}$ by definition. For $x_{2,4}$, since

$$x_{2,4} = x_2 - x_{2,3} = x_2 - x_3 + x_{1,3} = x_2 - x_3 + \min(x_1, x_3)$$

we have to study two cases:

- if $\min(x_1, x_3) = x_1$, then $x_{2,4} = x_2 + x_1 - x_1 = x_2 \in \mathbb{N}$;
- if $\min(x_1, x_3) = x_3$, then $x_{2,4} = x_2 + x_1 - x_3 = x_3 + x_4 - x_3 = x_4 \in \mathbb{N}$

which ensures that these elements are well defined. This construction directly implies that

$$x_1 = x_{1,3} + x_{1,4} \qquad x_2 = x_{2,3} + x_{2,4} \qquad x_3 = x_{1,3} + x_{2,3}.$$

Moreover, we have

$$x_{1,4} + x_{2,4} = (x_1 - x_{1,3}) + (x_2 - x_{2,3}) = x_1 + x_2 - (x_{1,3} + x_{2,3}) = x_3 + x_4 - x_3 = x_4$$

which proves the additive splitting property. □

Remark 2.3.3. One can note that, in general, this decomposition is not unique. For example, for $(\mathbb{N}, +, 0)$ again, the previous proof gives one possible decomposition, but the same proof also works with

$$x_{1,3} = 0 \qquad x_{1,4} = x_1 \qquad x_{2,3} = x_3 \qquad x_{2,4} = x_2 - x_3$$

when $x_2 \geq x_3$. It gives then two distinct decompositions when x_1 and x_3 are non-zero.

In order to study $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ graded by LPDOcc, we first study the monoid of multivariate polynomials and show that it fulfills the additive splitting property.

Definition 2.3.4. *We define inductively the monoid of multivariate polynomials as:*

$$\mathcal{P}_0 = \mathbb{R} \qquad \mathcal{P}_{n+1} = \mathcal{P}_n[X_{n+1}] \qquad \mathcal{P}_\omega = \cup_{n\geq 0}\mathcal{P}_n$$

and the product operation as, for $P, Q \in \mathcal{P}_\omega$ such that

$$P = \sum_{\alpha\in\mathbb{N}^{(\omega)}} a_\alpha X^\alpha \qquad Q = \sum_{\beta\in\mathbb{N}^{(\omega)}} b_\beta X^\beta \qquad P \boxplus Q = \sum_{\alpha,\beta} a_\alpha b_\beta X^{\alpha+\beta}.$$

Theorem 2.3.5. *The monoid $(\mathcal{P}_\omega, \boxplus, 1)$ is additive splitting.*

To prove this theorem, some results about rings of polynomials will be necessary. Even if in this section, the algebraic structures that we consider have only one operation, polynomials are mostly studied as elements of a ring.

Definition 2.3.6. *Let $\mathcal{R}$ be a non-zero commutative ring.*

1. *An element $u \in \mathcal{R}$ is a unit if there is $v \in \mathcal{R}$ such that $uv = 1$.*
2. *An element $x \in \mathcal{R}\backslash\{0\}$ is irreducible if it is not a unit, and not a product of two non-unit elements.*
3. *Two elements $x, y \in \mathcal{R}$ are associates if x divides y and y divides x.*
4. *$\mathcal{R}$ is a factorial ring if it is an integral domain such that for each $x \in \mathcal{R}\backslash\{0\}$ there is a unit $u \in \mathcal{R}$ and $p_1, \dots, p_n \in \mathcal{R}$ irreducible elements such that $x = up_1 \dots p_n$ and for every other decomposition $vq_1 \dots q_m = up_1 \dots p_n$ (with v unit and q_i irreducible for each i) we have $n = m$ and a bijection $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ such that p_i and $q_{\sigma(i)}$ are associated for each i.*

It is a well-known fact that the ring of real multivariate polynomials is factorial (see [Bos09, 2.7.7]). We use the decomposition coming from this property to prove that the monoid $(\mathcal{P}_\omega, \boxplus, 1)$ is additive splitting.

Proof of Theorem 2.3.5. Let $P_1, P_2, P_3, P_4 \in \mathcal{P}_\omega$ such that $P_1 \boxplus P_2 = P_3 \boxplus P_4$. Let n be an integer such that $P_1, P_2, P_3, P_4 \in \mathbb{R}[X_1, \dots, X_n]$, which exists by definition of $\mathcal{P}_\omega$. Using the factoriality of the ring $(\mathbb{R}[X_1, \dots, X_n], 0, +, 1, \times)$, there are decompositions

$$P_i = u_i Q_{i,1} \times \dots \times Q_{i,n_i} \qquad \text{for } 1 \leq i \leq 4$$

such that u_i are units, and $Q_{i,j}$ are irreducible. We then have

$$\begin{aligned} P_1 \boxplus P_2 &= u_1 u_2 Q_{1,1} \dots Q_{1,n_1} Q_{2,1} \dots Q_{2,n_2} \\ &= u_3 u_4 Q_{3,1} \dots Q_{3,n_3} Q_{4,1} \dots Q_{4,n_4} \end{aligned}$$

and since u_1u_2 and u_3u_4 are units, the factoriality implies that there exists a bijection

$$\sigma : \{(1,1),\dots,(1,n_1),(2,1),\dots,(2,n_2)\} \to \{(3,1),\dots,(3,n_3),(4,1),\dots,(4,n_4)\}$$

such that $Q_{i,j}$ and $Q_{\sigma((i,j))}$ are associates. We denote $v_{i,j}$ the unit such that $Q_{\sigma((i,j))} = v_{i,j}Q_{i,j}$. Then, we define

$$\begin{aligned}
I_{1,3} &= \{(1,k) \mid 1 \leq k \leq n_1 \text{ and } \sigma((1,k)) = (3,k')\}\\
I_{1,4} &= \{(1,k) \mid 1 \leq k \leq n_1 \text{ and } \sigma((1,k)) = (4,k')\}\\
I_{2,3} &= \{(2,k) \mid 1 \leq k \leq n_2 \text{ and } \sigma((2,k)) = (3,k')\}\\
I_{2,4} &= \{(2,k) \mid 1 \leq k \leq n_2 \text{ and } \sigma((2,k)) = (4,k')\}.
\end{aligned}$$

We can finally define the four polynomials

$$P_{1,3} = u_1 \prod_{x \in I_{1,3}} Q_x \qquad P_{2,3} = \frac{u_3}{u_1}\left(\prod_{x\in I_{1.3}\cup I_{2.3}} v_x\right)\left(\prod_{x\in I_{2,3}} Q_x\right)$$

$$P_{1,4} = \prod_{x \in I_{1,4}} Q_x \qquad P_{2,4} = \frac{u_2u_1}{u_3}\left(\prod_{x\in I_{1.3}\cup I_{2.3}} v_x\right)^{-1}\left(\prod_{x\in I_{2,4}} Q_x\right)$$

which gives the desired decomposition for P_1, P_2 and P_3 by construction. For P_4, we first have that

$$\begin{aligned}
&u_1u_2Q_{1,1}\dots Q_{1,n_1}Q_{2,1}\dots Q_{2,n_2}\\
&= u_3u_4Q_{3,1}\dots Q_{3,n_3}Q_{4,1}\dots Q_{4,n_4}\\
&= u_3u_4\,(v_{1,1}Q_{1,1})\dots(v_{1,n_1}Q_{1,n_1})\,(v_{2,1}Q_{2,1})\dots(v_{2,n_2}Q_{2,n_2})\\
&= u_3u_4\left(\prod_{x\in I_{1,3}\cup I_{2,3}} v_x\right)\left(\prod_{x\in I_{1,4}\cup I_{2,4}} v_x\right)(Q_{1,1}\dots Q_{1,n_1}Q_{2,1}\dots Q_{2,n_2})
\end{aligned}$$

which implies

$$\prod_{x\in I_{1,4}\cup I_{2,4}} v_x = \frac{u_1u_2}{u_3u_4}\left(\prod_{x\in I_{1,3}\cup I_{2,3}} v_x\right)^{-1}$$

so the decomposition works for P_4, as

$$\begin{aligned}
P_{1,4}\times P_{2,4} &= \frac{u_2u_1}{u_3}\left(\prod_{x\in I_{1,3}\cup I_{2,3}} v_x\right)^{-1}\left(\prod_{x\in I_{1,4}\cup I_{2,4}} Q_x\right)\\
&= u_4\left(\prod_{x\in I_{1,4}\cup I_{2,4}} v_x\right)\left(\prod_{x\in I_{1,4}\cup I_{2,4}} Q_x\right)\\
&= P_4.
\end{aligned}$$

□

Thanks to the isomorphism between multivariate polynomials and LPDOcc introduced for Proposition 2.2.2, we get that LPDOcc with the composition is additive splitting as well.

Corollary 2.3.7. *The monoid of LPDOcc endowed with composition* $(\mathcal{D}, \circ, id)$ *is additive splitting.*

2.3.1.2 Endowing monoids with an order

After studying some properties of the sum of a monoid, we focus on the order. It is used in the rules d_I and $\bar{\mathsf{d}}_I$. The vocabulary of this section is based on [DK09].

Definition 2.3.8. *A commutative monoid* $(\mathcal{M}, +, 0)$ *is*

- ordered *if it is equipped with a partial order* $\leq$ *such that*

$$x \leq y \Rightarrow x + z \leq y + z$$

  for all $z \in \mathcal{M}$*;*

- positively ordered *if for each* $x \in \mathcal{M},\ x \geq 0$*;*
- naturally ordered *if it is equipped by the order* $\leq$ *defined by*

$$x \leq y \Leftrightarrow \exists z \in \mathcal{M},\ x + z = y.$$

Note that a naturally ordered monoid is in particular positively ordered, as $0 + x = x$ for every element x of the monoid.

Remark 2.3.9. In the context of a naturally ordered monoid, the indexed (co)derelictions can be expressed with another form, using the sum instead of the order:

$$\frac{\vdash \Gamma, ?_x A \qquad x \leq y}{\vdash \Gamma, ?_y A}\,\mathsf{d}_I \;=\; \frac{\vdash \Gamma, ?_x A}{\vdash \Gamma, ?_{x+z} A}\,\mathsf{d}_I \qquad \frac{\vdash \Gamma, !_x A \qquad x \leq y}{\vdash \Gamma, !_y A}\,\bar{\mathsf{d}}_I \;=\; \frac{\vdash \Gamma, !_x A}{\vdash \Gamma, !_{x+z} A}\,\bar{\mathsf{d}}_I$$

which can be seen directly from the definition of naturally ordered. This allows us to connect $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ and IDiLL, as this equivalent form is exactly the one that we gave for IDiLL.

2.3.2 Cut elimination

Our main motivation for unifying graded and differential linear logic was to use previous work on $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ to define the dynamics of our logic indexed by differential operators. However, in graded linear logic, the interactions between the exponential rules are only between the promotion and each exponential rule. In $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$, each structural rule and each costructural rule can interact together. While some of these interactions have a behavior similar to the one in DiLL, the indexed (co)derelictions have a completely new behavior. Moreover, since in this framework we consider $\mathcal{M}$ naturally ordered, some cases cannot be naturally described. For example

$$\frac{\dfrac{\dfrac{\pi_1}{\vdash \Gamma, ?_x A}}{\vdash \Gamma, ?_{x+y} A}\,\mathsf{d}_I \qquad \dfrac{\dfrac{\pi_2}{\vdash \Delta, !_z A^{\perp}}}{\vdash \Delta, !_{z+t} A^{\perp}}\,\bar{\mathsf{d}}_I}{\vdash \Gamma, \Delta}\,cut \tag{2.4}$$

with $x + y = z + t$. Noticing the fact that each rule makes the index bigger, if the indexes x and y cannot be related through the natural order, no direct reduction is possible. To be more concrete, let us take $(\mathcal{D}, \circ, id)$ as our monoid, and define $x = \frac{\partial}{\partial x_1} = t$ and $y = \frac{\partial}{\partial x_2} = z$. It verifies $x \circ y = z \circ t$. The fact that we do not have $x \leq z$ nor $z \leq x$ leads to think that there is no direct way to eliminate this cut.

A first approach to tackle this issue would by to *push* these indexed (co)derelictions in the subtrees π_1 and π_2, in order to find another level of the proof to eliminate the cut. But this technique, inspired by subtyping ideas, is not always working. For example, no commutation is possible between a weakening and an indexed dereliction, since an exponential formula has to be introduced before having an index that grows. One may notice that the proof tree 2.4 can equivalently be rewritten as follows.

$$\dfrac{\dfrac{\dfrac{\dfrac{\pi_1}{\vdash \Gamma, ?_x A}}{\vdash \Gamma, ?_x A, ?_0 A}\,\mathsf{w}}{\dfrac{\vdash \Gamma, ?_x A, ?_y A}{\vdash \Gamma, ?_{x+y} A}\,\mathsf{c}}\,\mathsf{d}_I \qquad \dfrac{\dfrac{\pi_2}{\vdash \Delta, !_z A^\perp} \quad \dfrac{\dfrac{}{\vdash !_0 A^\perp}\,\bar{\mathsf{w}}}{\vdash !_t A^\perp}\,\bar{\mathsf{d}}_I}{\vdash \Delta, !_{z+t} A^\perp}\,\bar{\mathsf{c}}}{\vdash \Gamma, \Delta}\,cut \tag{2.5}$$

This cut may be similar to the way cuts between contractions and cocontractions are eliminated in differential linear logic.

The indexed (co)weakenings The fact that indexed (co)derelictions and (co)weakenings cannot commute leads us to introduce two new rules, the indexed weakening w_I and the indexed coweakening $\bar{\mathsf{w}}_I$

$$\dfrac{\vdash \Gamma}{\vdash \Gamma, ?_x A}\,\mathsf{w}_I \qquad\qquad \dfrac{}{\vdash !_x A}\,\bar{\mathsf{w}}_I$$

which generalizes the graded weakening and coweakening. They will be crucial in the definition of the cut elimination procedure of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$. These two rules can easily be expressed in $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ by:

$$\dfrac{\vdash \Gamma}{\vdash \Gamma, ?_x A}\,\mathsf{w}_I \;\triangleq\; \dfrac{\dfrac{\vdash \Gamma}{\vdash \Gamma, ?_0 A}\,\mathsf{w}}{\vdash \Gamma, ?_x A}\,\mathsf{d}_I \qquad\qquad \dfrac{}{\vdash !_x A}\,\bar{\mathsf{w}}_I \;\triangleq\; \dfrac{\dfrac{}{\vdash !_0 A}\,\bar{\mathsf{w}}}{\vdash !_x A}\,\bar{\mathsf{d}}_I$$

which gives the following proposition.

Proposition 2.3.10. *The rules w_I and $\bar{\mathsf{w}}_I$ are admissible in $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$.*

Another remark about these indexed (co)weakenings is that they can replace the indexed (co)derelictions when the monoid $\mathcal{M}$ is naturally ordered, thanks to the following equalities.

$$\dfrac{\vdash \Gamma, ?_x A}{\vdash \Gamma, ?_{x+y} A}\,\mathsf{d}_I \;=\; \dfrac{\dfrac{\vdash \Gamma, ?_x A}{\vdash \Gamma, ?_x A, ?_y A}\,\mathsf{w}_I}{\vdash \Gamma, ?_{x+y} A}\,\mathsf{c} \qquad \dfrac{\vdash \Gamma, !_x A}{\vdash \Gamma, !_{x+y} A}\,\bar{\mathsf{d}}_I \;=\; \dfrac{\vdash \Gamma, !_x A \quad \dfrac{}{\vdash !_y A}\,\bar{\mathsf{w}}_I}{\vdash \Gamma, !_{x+y} A}\,\bar{\mathsf{c}}$$

This implies that, if one defines a logic with the graded rules $\mathsf{w}, \mathsf{c}, \bar{\mathsf{w}}, \bar{\mathsf{c}}$, it has the same expressivity to add d_I and $\bar{\mathsf{d}}_I$ or to add w_I and $\bar{\mathsf{w}}_I$ (when the monoid is naturally ordered). However, the rules d_I and $\bar{\mathsf{d}}_I$ are difficult to study syntactically, since their shapes do not exist in DiLL. Hence the interactions may be very different than those from the cut elimination of DiLL. By expressing them thanks to the indexed (co)weakening, we avoid this issue and the difficulties are only on how to index the exponential.

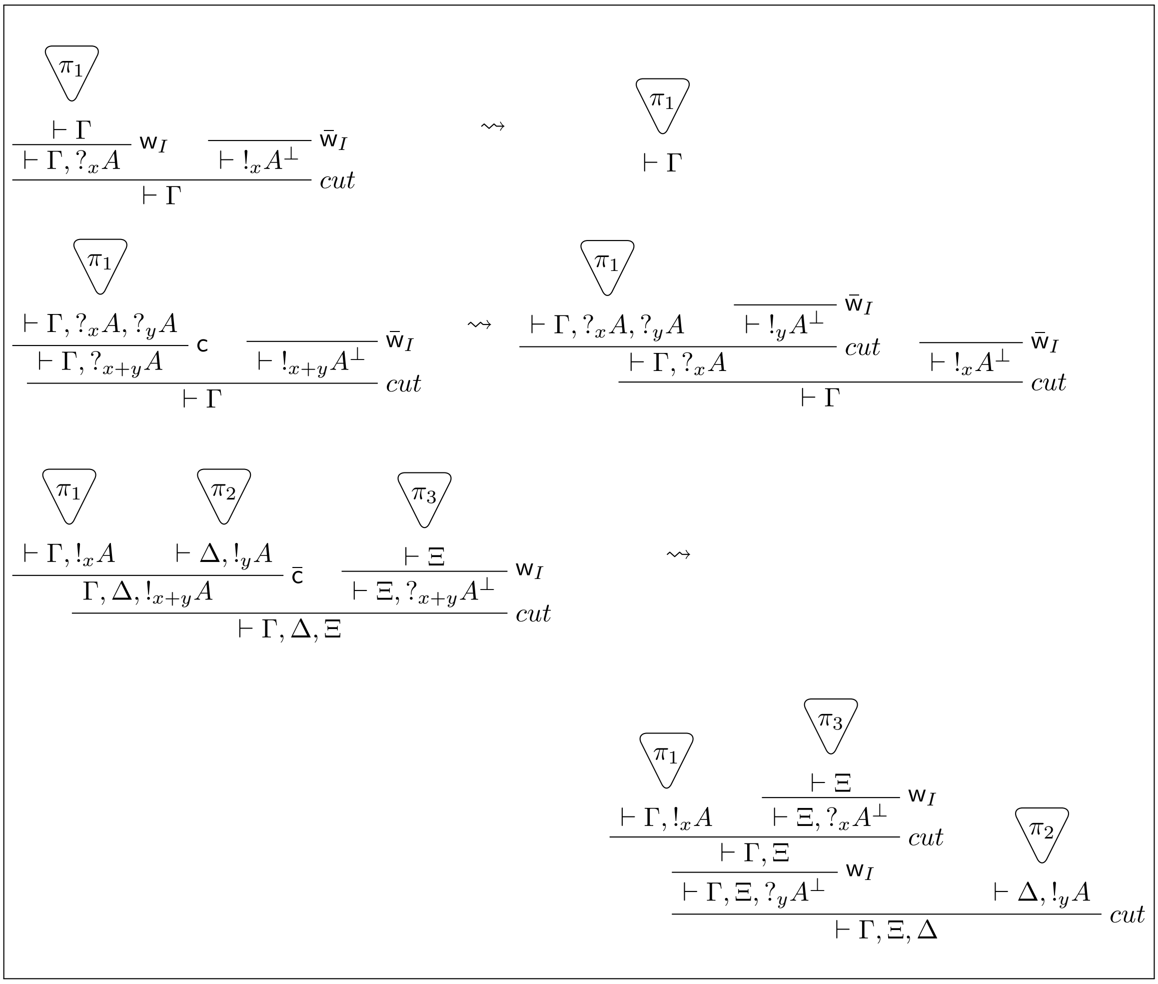


Figure 2.4: Cut elimination for $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$: indexed (co)weakening

The cut elimination cases

Theorem 2.3.11. *If $\mathcal{M}$ is a naturally ordered commutative monoid that is additive splitting, the logic $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ enjoys a cut elimination procedure.*

Remark 2.3.12. The procedure given in [BKM23] combines ideas from subtyping with this intuition where the indexed (co)derelictions are pushed into the tree together with the use of the indexed (co)weakening. Here, we give a much more factorized version of the procedure, thanks to the fact that w_I and $\bar{\mathsf{w}}_I$ can express the action of d_I and $\bar{\mathsf{d}}_I$.

We give the cut elimination cases involving w_I or $\bar{\mathsf{w}}_I$ in Figure 2.4. These cases give rise to trees that have the same shape as those from the cut elimination of DiLL. They are indexed by elements of $\mathcal{M}$. Note that no particular property on $\mathcal{M}$ is required to make these cases work. The last case is between the contraction and the cocontraction. It is given in Figure 2.5. For this one, we use the decomposition given by the additive splitting property. The shape of the reduction is the one of the contraction/cocontraction cut but the difficulty comes from the indexation. Like in the proof tree 2.5, two sums may be equal but coming from non-equal elements. Thanks to the additive splitting, we have a way of decomposing such elements which gives a way to index the exponential for this cut elimination case. Note that the naturally ordered condition in Theorem 2.3.11

Defining π_a and π_b as follows:

$$\pi_a = \dfrac{\dfrac{\dfrac{}{\vdash ?_{x_{2,3}}A^\perp, !_{x_{2,3}}A}\,ax \qquad \dfrac{}{\vdash ?_{x_{2,4}}A^\perp, !_{x_{2,4}}A}\,ax}{\vdash ?_{x_{2,3}}A^\perp, ?_{x_{2,4}}A^\perp, !_{x_2}A}\,\bar{\mathsf{c}} \qquad \dfrac{\pi_1}{\vdash \Gamma, ?_{x_1}A^\perp, ?_{x_2}A^\perp}}{\vdash \Gamma, ?_{x_{2,3}}A^\perp, ?_{x_{2,4}}A^\perp, ?_{x_1}A^\perp}\,cut$$

$$\pi_b = \dfrac{\dfrac{\dfrac{\pi_a}{\vdash \Gamma, ?_{x_{2,3}}A^\perp, ?_{x_{2,4}}A^\perp, ?_{x_1}A^\perp} \qquad \dfrac{\dfrac{}{\vdash ?_{x_{1,3}}A^\perp, !_{x_{1,3}}A}\,ax \qquad \dfrac{}{\vdash ?_{x_{1,4}}A^\perp, !_{x_{1,4}}A}\,ax}{?_{x_{1,3}}A^\perp, ?_{x_{1,4}}A^\perp, !_{x_1}A}\,\bar{\mathsf{c}}}{\vdash \Gamma, ?_{x_{2,3}}A^\perp, ?_{x_{2,4}}A^\perp, ?_{x_{1,3}}A^\perp, ?_{x_{1,4}}A^\perp}\,cut}{\vdash \Gamma, ?_{x_{1,4}}A^\perp, ?_{x_{2,4}}A^\perp, ?_{x_3}A^\perp}\,\mathsf{c}$$

We give the interaction as:

$$\dfrac{\dfrac{\dfrac{\pi_1}{\vdash \Gamma, ?_{x_1}A^\perp, ?_{x_2}A^\perp}}{\vdash \Gamma, ?_{x_1+x_2}A^\perp}\,\mathsf{c} \qquad \dfrac{\dfrac{\pi_2}{\vdash \Delta, !_{x_3}A} \qquad \dfrac{\pi_3}{\vdash \Xi, !_{x_4}A}}{\vdash \Delta, \Xi, !_{x_3+x_4=x_1+x_2}A}\,\bar{\mathsf{c}}}{\vdash \Gamma, \Delta, \Xi}\,cut \quad \rightsquigarrow_{cut}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\pi_b}{\vdash \Gamma, ?_{x_{1,4}}A^\perp, ?_{x_{2,4}}A^\perp, ?_{x_3}A^\perp} \qquad \dfrac{\pi_2}{\vdash \Delta, !_{x_3}A}}{\vdash \Gamma, \Delta, ?_{x_{1,4}}A^\perp, ?_{x_{2,4}}A^\perp}\,cut}{\vdash \Gamma, \Delta, ?_{x_4}A^\perp}\,\mathsf{c} \qquad \dfrac{\pi_3}{\Xi, !_{x_4}A}}{\vdash \Gamma, \Delta, \Xi}\,cut$$

Figure 2.5: Cut elimination for $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$: contraction and cocontraction

does not appear directly in the cut elimination cases. But they are mandatory to express translations on d_I and $\bar{\mathsf{d}}_I$, which is why we do not need to give the cases for these rules.

2.3.2.1 The equivalence between IDiLL and $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$

We first recall Corollary 2.3.7 proved in Section 2.3.1.1.

Corollary 2.3.7. *The monoid of LPDOcc endowed with composition* $(\mathcal{D}, \circ, id)$ *is additive splitting.*

Using this corollary, and endowing this monoid with its natural order, we consider $\mathsf{DB}_{\mathcal{D}}\mathsf{LL}$ which is then graded by LPDOcc. The rules of $\mathsf{DB}_{\mathcal{D}}\mathsf{LL}$ are then the same as those of IDiLL. The difference is in the grammar: IDiLL is polarized, while $\mathsf{DB}_{\mathcal{D}}\mathsf{LL}$ is not. However, the cut elimination cases do not break the polarities, as they do not build new formulas. Taking a proof tree from IDiLL, one can translate its indexed (co)dereliction thanks to indexed (co)weakening, and apply the cases from $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ to eliminate the cuts. Thanks to the admissibility of w_I and $\bar{\mathsf{w}}_I$ in $\mathsf{DB}_{\mathcal{D}}\mathsf{LL}$, the cut free tree obtained after the reduction can be retranslated into a tree of $\mathsf{DB}_{\mathcal{D}}\mathsf{LL}$ that is, in fact, a cut free proof of IDiLL which proves the same sequent.

Theorem 2.3.13. *The logic IDiLL enjoys a cut-elimination procedure.*

2.3.3 The relational model

A natural model to consider for $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ is the relational model. As shown in Chapter 1, **Rel** is a model of LL, of DiLL, and can be extended to a model of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$. To have a model of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$, the interpretation of the formulas and of the structural rules are the same as those of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ from Section 1.2.3.1. What is left to do is to define the interpretations of the rules $\bar{\mathsf{w}}$, $\bar{\mathsf{c}}$ and $\bar{\mathsf{d}}_I$. This is done thanks to the reverse relation

$$\bar{\mathsf{w}}_X \triangleq \{(a,b) \mid (b,a) \in \mathsf{w}_X\} \qquad \bar{\mathsf{c}}_X \triangleq \{(a,b) \mid (b,a) \in \mathsf{c}_X\} \qquad \bar{\mathsf{d}}_X \triangleq \{(a,b) \mid (b,a) \in \mathsf{d}_X\}.$$

Using what is done to extend the relational model to a model of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, this gives a model of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$.

2.3.4 Coherence between the smooth semantics and the cut elimination

After defining the smooth semantics with functions, distributions and LPDOcc of IDiLL, we saw that this logic enjoys a cut elimination procedure thanks to $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$. A crucial point to study is the compatibility between this model and the cut elimination procedure $\rightsquigarrow$. In denotational semantics, one would expect that a model is invariant with respect to the computation. In our case, that would mean that, for each step of rewriting of $\rightsquigarrow$, the interpretation of the proof tree has the same value. To do so, let us consider the interpretation in the smooth model of the indexed (co)weakening. The indexed weakening applies the indexed dereliction to cst_1, the smooth map introduced by the weakening. The indexed coweakening applies the indexed codereliction to δ_0, the distribution introduced by the coweakening. Their semantics is then

$$\mathsf{w}_I : \begin{cases} \mathbb{R} & \to ?_D A \\ 1 & \mapsto \Phi_D * cst_1 \end{cases} \qquad \bar{\mathsf{w}}_I : \begin{cases} \mathbb{R} & \to !_D A \\ 1 & \mapsto \widehat{D}(\delta_0) \end{cases}$$

which are well typed by construction. From this, let us study each exponential commutation case involving w_I or $\bar{\mathsf{w}}_I$.

- For the cut $\mathsf{w}_I/\bar{\mathsf{w}}_I$, the semantical interpretation corresponds to

  $$\widehat{D}(\delta_0)(\Phi_D * cst_1) = \delta_0(D(\Phi_D * cst_1)) = \delta_0(cst_1) = 1$$

  which is the neutral interpretation of a sequent, so this case lets the interpretation invariant.

- For the cut $\mathsf{c}/\bar{\mathsf{w}}_I$, the interpretation before the cut elimination is

  $$\widehat{D_1 \circ D_2}(\delta_0)(\Phi_{D_1 \circ D_2} * (D_1(f).D_2(g))) = \delta_0(D_1(f).D_2(g)) = D_1(f)(0).D_2(g)(0)$$

  where f is in $[\![?_{D_1} A]\!]$ and g is in $[\![?_{D_2} A]\!]$. The interpretation after the cut elimination is

  $$\widehat{D_1}(\delta_0)(f).\widehat{D_2}(\delta_0)(g) = \delta_0(D_1(f)).\delta_0(D_1(f)) = D_1(f)(0).D_2(g)(0)$$

  which shows the invariance.

- For the cut $\bar{\mathsf{c}}/\mathsf{w}_I$, the interpretation before the cut elimination is

$$\begin{aligned}
&(\psi * \phi)(\Phi_{D_1 \circ D_2} * cst_1) \\
&= \psi(x \mapsto \phi(y \mapsto \Phi_{D_1} * (\Phi_{D_2} * cst_1)(x+y))) && \text{(Proposition 2.2.1)} \\
&= \psi(x \mapsto \phi(y \mapsto \Phi_{D_1}(z \mapsto \Phi_{D_2} * cst_1(x+y-z)))) \\
&= \psi(x \mapsto \phi(y \mapsto \Phi_{D_1}(cst_{\Phi_{D_2}(cst_1)}))) \\
&= \psi(x \mapsto \phi(y \mapsto \Phi_{D_1}(\Phi_{D_2}(cst_1).cst_1))) \\
&= \psi(x \mapsto \phi(y \mapsto \Phi_{D_2}(cst_1).\Phi_{D_1}(cst_1))) && \text{(by homogeneity of } \phi) \\
&= \psi(x \mapsto \phi(cst_{\Phi_{D_2}(cst_1).\Phi_{D_1}(cst_1)})) \\
&= \psi(x \mapsto \phi(\Phi_{D_1}(cst_1).cst_{\Phi_{D_2}(cst_1)})) \\
&= \psi(x \mapsto \Phi_{D_1}(cst_1).\phi(cst_{\Phi_{D_2}(cst_1)})) && \text{(by homogeneity of } \phi) \\
&= \psi(cst_{\Phi_{D_1}(cst_1).\phi(cst_{\Phi_{D_2}(cst_1)})}) \\
&= \psi(\phi(cst_{\Phi_{D_2}(cst_1)}).cst_{\Phi_{D_1}(cst_1)}) \\
&= \phi(cst_{\Phi_{D_2}(cst_1)}).\psi(cst_{\Phi_{D_1}(cst_1)}) && \text{(by homogeneity of } \psi)
\end{aligned}$$

where $\psi \in [\![!_{D_1} A]\!]$ and $\phi \in [\![!_{D_2} A]\!]$. The one after the cut elimination is

$$\begin{aligned}
&\psi(\Phi_{D_1} * cst_1).\phi(\Phi_{D_2} * cst_1) \\
&= \psi(x \mapsto \Phi_{D_1}(y \mapsto cst_1(x-y))).\phi(x \mapsto \Phi_{D_2}(y \mapsto cst_1(x-y))) \\
&= \psi(x \mapsto \Phi_{D_1}(cst_1)).\phi(x \mapsto \Phi_{D_2}(cst_1)) \\
&= \psi(cst_{\Phi_{D_1}(cst_1)}).\phi(cst_{\Phi_{D_2}(cst_1)})
\end{aligned}$$

and that proves the invariance for this case.

The last case to consider is the one between a contraction and a cocontraction. This case takes slightly more work. It relies on the density of $\{\delta_x \mid x \in E\}$ in $!E$. First, let us study the dual rule of the contraction.

Lemma 2.3.14. *The dual of the contraction law corresponds to*

$$\mathsf{c}'_{D_1,D_2} : \widehat{D_1 \circ D}_2(\delta_x) \in !_{D_1 \circ D_2} E \mapsto (\widehat{D_1}(\delta_x) \otimes \widehat{D_2}(\delta_x)) \in !_{D_1} E \otimes !_{D_2} E.$$

Proof. Because we are working on finite-dimensional spaces E, an application of the Hahn-Banach theorem gives us that the span of $\{\delta_x \mid x \in E\}$ is dense in $!E$. As such, the interpretation of c' can be restricted to elements of the form $\widehat{D_1 \circ D}_2(\delta_x) \in !_{D_1 \circ D_2} E$. Also remember that for a linear map $\ell : E \multimap F$, its dual $\ell' : F' \multimap E'$ computes as follows

$$\ell' : h \in F' \mapsto (x \in E \mapsto h(\ell(x))) \in E'$$

Indeed, consider $\ell \in (?_{D_1 \circ D_2} E)'$. As all the spaces considered are reflexive, one has:

$$(?_{D_1 \circ D_2} E)' \simeq \{\widehat{D_1 \circ D}_2(\phi) \mid !E\}$$

and as such there is $\phi \in !E$ such that $\ell = \widehat{D_1 \circ D}_2(\phi)$. As such, for any $f \otimes g \in ?_{D_1} E \hat{\otimes} ?_{D_2} E$ one has:

$$\begin{aligned}
(\ell \circ \mathsf{c})(f \otimes g) &= (\phi \circ D_1 \circ D_2)(\Phi_{D_1 \circ D_2} * (D_1(f) \cdot D_2(g))) \\
&= \phi(D_1(f).D_2(g))
\end{aligned}$$

Considering $\phi = \delta_x$, we obtain

$$\begin{aligned}(\ell \circ \mathsf{c})(f \otimes g) &= \delta_x(D_1(f).D_2(g)) \\ &= \delta_x(D_1(f)) \cdot \delta_x(D_2(g)) \\ &= ((\delta_x \circ D_1) \otimes (\delta_x \circ D_2))(f \otimes g).\end{aligned}$$

Hence c' corresponds to $\mathsf{c}'_{D_1,D_2} : !_{D_1 \circ D_2} E \to !_{D_1} E \otimes !_{D_2} E$. □

For this case, suppose that we have $D_1, D_2, D_3, D_4 \in \mathcal{D}$ such that $D_1 \circ D_2 = D_3 \circ D_4$. By the additive splitting property, we have $D_{1,3}, D_{1,4}, D_{2,3}, D_{2,4}$ such that

$$D_1 = D_{1,3} \circ D_{1,4} \qquad D_2 = D_{2,3} \circ D_{2,4} \qquad D_3 = D_{1,3} \circ D_{2,3} \qquad D_4 = D_{1,4} \circ D_{2,4}.$$

The diagrammatic translation of the cut-elimination rule in Figure 2.5 is the following.

$$\begin{array}{ccc} !_{D_1} E \otimes !_{D_2} E & \xrightarrow{\mathsf{c}'_{D_{1,3},D_{1,4}} \otimes \mathsf{c}'_{D_{2,3},D_{2,4}}} & !_{D_{1,3}} E \otimes !_{D_{1,4}} E \otimes !_{D_{2,3}} E \otimes !_{D_{2,4}} E \\ \downarrow{\scriptstyle \bar{\mathsf{c}}_{D_1,D_2}} & & \\ !_{D_1 \circ D_2} E = !_{D_3 \circ D_4} E & & \downarrow \\ \downarrow{\scriptstyle \mathsf{c}'_{D_3,D_4}} & & \\ !_{D_3} E \otimes !_{D_4} E & \xleftarrow[\bar{\mathsf{c}}_{D_{1,3},D_{2,3}} \otimes \bar{\mathsf{c}}_{D_{1,4},D_{2,4}}]{} & !_{D_{1,3}} E \otimes !_{D_{2,3}} E \otimes !_{D_{1,4}} E \otimes !_{D_{2,4}} E \end{array}$$

Remember that the convolution of Dirac operators is the Dirac of the sum of points, and as such, we have:

$$\bar{\mathsf{c}} : (\widehat{D_a}(\delta_x)) \otimes (\widehat{D_b}(\delta_y)) \mapsto (\widehat{D_b \circ D}_a(\delta_{x+y})).$$

We make use of Lemma 2.3.14 to compute easily that the diagram above commutes on elements $(\widehat{D_1}(\delta_x)) \otimes (\widehat{D_2}(\delta_y))$ of $!_{D_1} E \otimes !_{D_2} E$, and as such it commutes on all elements by density and continuity of $\bar{\mathsf{c}}$ and c'. The last step is to prove that the translations of d_I and $\bar{\mathsf{d}}_I$ are invariant as well by the smooth model. In fact, these translations are used to rewrite the proof tree before the procedure to a proof tree of IDiLL. For the indexed dereliction, the translation for a smooth map $f \in [\![?_{D_1} A]\!]$ is interpreted by the contraction with an indexed weakening

$$\Phi_{D_1 \circ D_2} * (D_1(f).D_2(\Phi_{D_2} * cst_1)) = \Phi_{D_1 \circ D_2} * (D_1(f).D_2(cst_1)) = a_0 \Phi_{D_1 \circ D_2} * (D_1(f)) = a_0 \Phi_{D_2} * f$$

which is the interpretation of the indexed dereliction.

Remark 2.3.15. Note that we do not get the exact interpretation of the indexed dereliction. To get it, we need the property that $a_0 = 1$. To do so, we just need to refine the monoid by dividing every LPDOcc by its coefficient a_0, and forbidding the operators having $a_0 = 0$.

For the indexed codereliction, the translation is interpreted by

$$\phi * \widehat{D_2}(\delta_0) = \widehat{D_2}(\phi * \delta_0) = \widehat{D_2}(\phi)$$

for $\phi \in [\![!_{D_1} A]\!]$ and we found the interpretation of the indexed codereliction. We can the conclude with the following theorem.

Theorem 2.3.16. *The interpretation of each morphism* $\mathsf{w}, \bar{\mathsf{w}}, \mathsf{c}, \bar{\mathsf{c}}, \mathsf{d}_I$ *and* $\bar{\mathsf{d}}_I$ *is compatible with the cut elimination procedure.*

Chapter 3

A double indexed differential linear logic

The definitions and results in this chapter are partially given in a preprint.

Contents

This chapter presents an extension of the logic IDiLL to higher order. In the previous chapter, the promotion rule has been removed from the syntax of the logics that we have defined. Adding this rule comes with difficulties at the level of the syntax, of the semantics and of the underlying algebraic structure of the indices. It is however a desirable rule, as it gives the possibility to consider higher order objects and to have a more complete framework. The solution to these difficulties used in this chapter consists of adding a duality operation on the set of indices of the exponential connectives. It corresponds to have two copies of the semiring, which can be used as indices. Compared to polarized linear logic, the exponential connectives are not splitted into $!_{\_}$ which is positive, and $?_{\_}$ which is negative. The connectives $!_{\_+}$ and $?_{\_+}$ are indexed by positive elements of the semiring while $!_{\_-}$ and $?_{\_-}$ by their duals which are negative. The common point with polarized linear logic is that this allows to restrict the usual exponential rules regarding whether the formula uses a positive or a negative index, including the promotion. This idea of a new duality comes from the use of the higher order Laplace transform, which has been used in models of DiLL by Kerjean and Lemay [KL23].

The higher order Laplace transform The Laplace transform consists of a morphism of function spaces, which is a powerful tool to solve differential equations. Its usual form

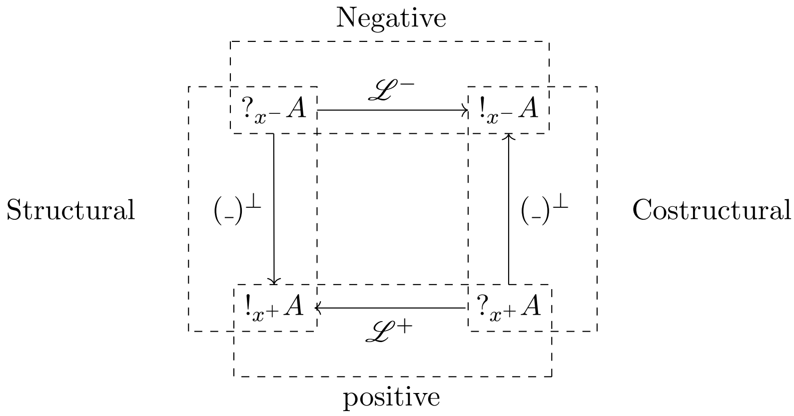


Figure 3.1: Double indexed formulas related by the Laplace transform and the negation

is

$$\mathscr{L} : (f : \mathbb{R}^n \to \mathbb{R}) \overset{\simeq}{\mapsto} \left(x \in \mathbb{R}^n \mapsto \int f(t) e^{-xt} dt \right). \tag{3.1}$$

However, the use of integration does not generalize easily to functions acting on an arbitrary (and potentially infinite dimensional) vector space $f : E \to \mathbb{R}$. The key to the higher-order generalization of Equation 3.1 is to then replace the use of integral with the use of distributions. Indeed, when the domain of Laplace transform is seen as a space of distributions, one can generalize Equation 3.1 as:

$$\mathscr{L} : (\phi \in !A) \overset{\simeq}{\mapsto} \left(\ell \in A^\perp \mapsto \phi(t \in A \mapsto e^{\ell(t)}) \right) \in ?A. \tag{3.2}$$

Then $\mathscr{L}(\phi)$ when $\phi \in !A$ is indeed an element of $?A$, which is a non-linear function acting on A'. Considering the Laplace transform of morphisms of DiLL, Kerjean and Lemay have shown that

$$\mathscr{L}(\bar{\mathsf{w}}) = \mathsf{w} \qquad \mathscr{L}(\bar{\mathsf{c}}) = \mathsf{c} \qquad \mathscr{L}(\bar{\mathsf{d}}) = \mathsf{d}.$$

In other words, the Laplace transform turns the costructural rules into the structural ones. This led us to see Laplace as a transformation orthogonal to the negation: the negation turns smooth maps into distributions, while the Laplace transform acts on the indices instead, transforming a negative one into a positive. The use of two copies of the same semiring allows having this orthogonality in the syntax. We express how those two transformations work together on formulas in Figure 3.1.

Syntactically, this having those four kinds of exponential formulas, with two orthogonal translations, gives the possibility of having a promotion rule. The issue in the syntax with the promotion is that when it has interactions with the costructural rules, and it is indexed, this implies new algebraic axioms on the semiring which does not match well with the ideas of differentiation. But with the four kinds of exponential formulas, these interactions are not possible anymore: there are no cuts between the promotion and the costructural rules, since the formulas produced are not the negation of each other anymore. This allows even to define a *copromotion* rule, based on the categorical codigging defined by Kerjean and Lemay [KL23] which is the morphism dual to the digging. In addition, the Laplace transform fits well with the theory of linear partial differential operators, as it transforms differential equations into polynomial ones.

In Section 3.1, we define the logic DIDiLL, which is based on a relaxed notion of semiring that we call *weak semiring*. It corresponds to the algebraic structure having exactly the algebraic axioms needed to perform the cut elimination. The grammar of DIDiLL uses exponential formulas indexed by positive and negative elements of a weak semiring. Then we give the definition of the Laplace transform and of the negation on these new formulas. We define the exponential rules of DIDiLL, which are the ones of DiLL but graded, where the indices are polarized. There is, in addition, a copromotion rule, which is built from the codigging. From these rules, we express the cut elimination cases between the exponential rules, which are either between structural rules and the promotion or between costructural rules and the copromotion. In Section 3.2, we give a first model of DIDiLL. This model uses LPDOcc and Köthe spaces. We define a weak semiring of LPDOcc, which is an extension of the monoid of LPDOcc from Chapter 2 where we add an operation of multiplication. Then we show how those differential operators can be applied to elements of Köthe spaces. To do so, we have to endow the Köthe spaces with a map that will represent a choice of some elements of a basis of the space. This is needed to interpret the variables along which we will differentiate. Using this, we show how the rules of DIDiLL are interpreted as morphisms between Köthe spaces. Sadly, we have not been able to interpret the copromotion in this setting for non-convergence reasons, but we get that Köthe spaces and LPDOcc form a model of DIDiLL without the copromotion. In Section 3.3, we give a second model of DIDiLL based on smooth maps and distributions. It uses a framework from Ouerdiane *et al.* where the smooth maps are bound by Young functions, which are real convex functions [GHOR00]. Using the Young functions as indices, we are able to give a model of DIDiLL which is smooth and has higher order constructions. To do so, we need to give an alternative version of DIDiLL where the connectives and the rules are polarized. This model also allows to interpret the copromotion rule, following what was done by Kerjean and Lemay on the codigging operation [KL23].

3.1 DIDiLL: a Double Indexed Differential Linear Logic

In this section, we define and study the calculus DIDiLL. This calculus will have exponentials indexed by elements of two copies of an algebraic structure, which is weaker than a semiring. We define the notion of weak semiring.

Definition 3.1.1. *Let* $(\mathcal{S}, +, 0)$ *be a commutative monoid,* 1 *an element of* $\mathcal{S}$*, and* $\times$ *an associative binary operation on* $\mathcal{S}$ *(the product). When the product is right-distributive over the sum, 0 is left-absorbing and* 1 *is a left-neutral for the product, i.e.*

$$(x + y)z = xz + yz \qquad 0 \times x = 0 \qquad 1 \times x = x$$

for each $x, y, z \in \mathcal{S}$*, and* $\leq$ *is a partial order such that the sum is increasing monotone,* $(\mathcal{S}, +, 0, \times, 1, \leq)$ *is a* weak semiring.

Compared to usual semiring, a weak semiring does not require distributivity on both sides, 0 to be right-absorbing, or 1 to be right-neutral. Beware that, in graded linear logic, this restriction is not required, and that it uses the usual definition of a semiring. The additional axioms on the general definition of a semiring are used to interpret the commutation cases in proof nets (see [DG24, Part 1.3] for an example of commutation that corresponds to an algebraic axiom of a semiring, but not required in our weaker definition) which are not necessary here. On the contrary, distributivity on both sides does not hold in the semirings interpreting DIDiLL in Section 3.2 and 3.3. Since we seek for a logic where the indices can be interpreted by differential operators, which do not fulfill every axiom of a semiring, this weaker notion is crucial here. In addition to being an algebraic structure

$$A, B \triangleq 1 \mid 0 \mid \top \mid \bot \mid A \otimes B \mid A \oplus B \mid A \parr B \mid A \& B \qquad \text{(MALL formulas)}$$
$$\mid ?_{x^+}A \mid !_{x^+}A \mid ?_{x^-}A \mid !_{x^-}A \qquad \text{(exponential formulas)}$$

(a) Formulas of DIDiLL

$$(?_{x^+}A)^\perp \triangleq !_{x^-}A^\perp \qquad (!_{x^+}A)^\perp \triangleq ?_{x^-}A^\perp$$
$$(?_{x^-}A)^\perp \triangleq !_{x^+}A^\perp \qquad (!_{x^-}A)^\perp \triangleq ?_{x^+}A^\perp$$
$$\mathscr{L}(?_{x^+}A) \triangleq !_{x^+}A \qquad \mathscr{L}(!_{x^+}A) \triangleq ?_{x^+}A$$
$$\mathscr{L}(?_{x^-}A) \triangleq !_{x^-}A \qquad \mathscr{L}(!_{x^-}A) \triangleq ?_{x^-}A$$

(b) Syntactical transformations of DIDiLL

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?_{0^-}A}\ \mathsf{w} \qquad \frac{}{\vdash !_{0^-}A}\ \bar{\mathsf{w}}$$

$$\frac{\vdash \Gamma, ?_{x^-}A, ?_{y^-}A}{\vdash \Gamma, ?_{x^-+y^-}A}\ \mathsf{c} \qquad \frac{\vdash \Gamma, !_{x^-}A \qquad \vdash \Delta, !_{y^-}A}{\vdash \Gamma, \Delta, !_{x^-+y^-}A}\ \bar{\mathsf{c}}$$

$$\frac{\vdash \Gamma, A}{\vdash \Gamma, ?_{1^-}A}\ \mathsf{d} \qquad \frac{\vdash \Gamma, A}{\vdash \Gamma, !_{1^-}A}\ \bar{\mathsf{d}}$$

$$\frac{\vdash \Gamma, ?_{x^-}A \qquad x \leq y}{\vdash \Gamma, ?_{y^-}A}\ \mathsf{d}_I \qquad \frac{\vdash \Gamma, !_{x^-}A \qquad x \leq y}{\vdash \Gamma, !_{y^-}A}\ \bar{\mathsf{d}}_I$$

$$\frac{\vdash ?_{X^-}\mathcal{A}, B}{\vdash ?_{y^- \times X^-}\mathcal{A}, !_{y^+}B}\ \mathsf{p} \qquad \frac{\vdash !_{x^-}A, B}{\vdash !_{y^- \times x^-}A, ?_{y^+}B}\ \bar{\mathsf{p}}$$

(c) Exponential rules of DIDiLL

Figure 3.2: The logic DIDiLL

which models LPDOcc, this definition corresponds to the minimal set of axioms needed to have a cut elimination procedure.

Let us now detail the sequent calculus of the logic DIDiLL. To the connectors of MALL, we now add 4 families of graded connectors: $?_{x^+}A$, $?_{x^-}A$, $!_{x^+}A$ and $!_{x^-}A$, where x ranges over a weak semiring (see Figure 3.2a). We denote by *polarity* the upper index $+$ or $-$, and simply denote by index the lower index $x \in \mathcal{S}$, where $\mathcal{S}$ is a semiring. We call $?_{\_^-}$ and $!_{\_^-}$ negative exponentials while $?_{\_^+}$ and $!_{\_^+}$ are positive. This division between two classes of exponential connectives is crucial for this logic. It solves the difficulties coming with the promotion rule, and also allows adding a copromotion rule, a rule added to differential linear logic by Kerjean and Lemay [KL23] that we will detail later. While adding these polarities to the connectives seems to be a syntactical operation, it follows intuitions from the so-called Laplace transform, which is a well-known operation in functional analysis. These intuitions will be detailed in Sections 3.2 and 3.3 within the framework of two concrete models. We give in Figure 3.2b the transformations at the level of the formulas induced by this Laplace transform denoted by $\mathscr{L}(\_)$, and by the negation. As usual, we define inductively the involutive negation such that $(?_{x^+}A)^\perp = !_{x^-}A^\perp$ and $(!_{x^+}A)^\perp = ?_{x^-}A^\perp$. The Laplace transform and the negation have a similar behavior, but Laplace does not change the polarity of the index while the negation does.

We now define the rules of DIDiLL as the rules of MALL to which are added the exponential rules in Figure 3.2c. For the promotion rule, $?_X\mathcal{A}$ represents $?_{x_1}A_1, \ldots, ?_{x_n}A_n$, for $X = (x_1, \ldots, x_n)$ and $\mathcal{A} = A_1, \ldots, A_n$. Then, $y \times X = (yx_1, \ldots, yx_n)$. Notice that the structural and costructural rules act on negative exponentials only. Also notice that, to be as general as possible, we present this logic graded by a weak semiring with an order. However, in the concrete models that we present in Section 3.2 and Section 3.3, the order is defined through the sum: $x \leq z$ when there is y such that $x + y = z$. We explain our choices regarding the copromotion rule $\bar{\mathsf{p}}$ in the next section and detail the cut-elimination cases in Section 3.1.2.

3.1.1 Coddiging and copromotion

The costructural rules of differential linear logic bring some symmetry to linear logic. Each costructural rule is the symmetrical to a structural one in linear logic. However, the promotion rule had a special status, since it was the only one without any costructural counterpart. This gap has been solved by Kerjean and Lemay by studying a particular concrete model of differential linear logic [KL23]. This model consists of spaces of functions and distributions whose growth is controlled by some real functions. We present precisely this model in Section 3.3, by turning it into a model of DIDiLL. The semantical interpretation of the promotion rule of linear logic is built using a morphism called the digging, which consists of taking a distribution and turning it into a distribution of distribution, using the Dirac construction. This digging operation dig gives to the functor $!\_$ a comonadic structure, as it is a comultiplication. In their article, Kerjean and Lemay have noticed that, in the model there were considering, one could define a codigging operation $\overline{\mathsf{dig}}$ given by

$$\overline{\mathsf{dig}} : \begin{cases} !!E & \to !E \\ \delta_\phi & \mapsto \sum_n \frac{1}{n!}\phi^{*^n} \end{cases}$$

where ϕ^{*^n} stands for the distribution ϕ convoluted n times with itself. This new operation $\overline{\mathsf{dig}}$ is a multiplication over $!\_$ and together with the codereliction it gives $!\_$ a monadic structure. This gives then a solution to the lack of symmetry at this level in differential linear logic. However, this rule is not well-defined in every model of DiLL. In fact, a model of DiLL has a codigging when every function is equal to its Taylor expansion at 0. This works for the model that we present in Section 3.3 thanks to its strong conditions of convergence, but not for Köthe spaces, for instance. We will detail this in Section 3.2.

After the semantical study of Kerjean and Lemay of this codigging, an important question is the one of the existence of a *copromotion* rule, which would solve the lack of symmetry of DiLL at the level of the syntax. While the promotion was defined syntactically and then semantically interpreted by the digging, here we use the codigging, which is semantical, to define the copromotion. The promotion rule can be decomposed into two rules: the functoriality of the bang ($!_\mathsf{f}$) and the digging (to simplify here, we call digging the digging generalized to a sequent. This has been done properly in Figure 1.5).

$$\frac{\Gamma \vdash A}{!\Gamma \vdash !A}\,!_\mathsf{f} \qquad \frac{!!\Gamma \vdash B}{!\Gamma \vdash B}\,\mathsf{dig} \qquad \mathsf{p} = \cfrac{\cfrac{!\Gamma \vdash A}{!!\Gamma \vdash !A}\,!_\mathsf{f}}{!\Gamma \vdash !A}\,\mathsf{dig}$$

One can then define a copromotion rule that will be dual to the promotion, but one major restriction has to be done. To be able to define a cut-elimination procedure for DIDiLL, we had to restrict for only a one-formula context in the copromotion $\bar{\mathsf{p}}$. This comes from the binary flavor of $\bar{\mathsf{c}}$, compared to the unary one of c. In the cut elimination case between p and c, the promoted part is duplicated and then recombined through contractions. Adapting this idea in the costructural case is possible, but the recombination works only with exactly one formula in the context, because of the binarity. This gives the following rule:

$$\bar{\mathsf{p}} = \dfrac{\dfrac{?A \vdash B}{??A \vdash ?B}\,?\mathsf{f}}{?A \vdash ?B}\,\mathsf{dig} \qquad\qquad \dfrac{\vdash !A, B}{\vdash !A, ?B}\,\bar{\mathsf{p}}$$

which is presented in our graded polarized setting in Figure 3.2c. This rule is built by duality using the functoriality of $?_{\_}$ and the codigging. One important point here is that this version of the copromotion comes from some choices. We might hope for another version in the future, but this is the only one which allowed us to prove cut elimination .

Notice that here, $\bar{\mathsf{p}}$ and p *do not interact through cut elimination*, as they do not involve dual connectors. Grading, and involving the Laplace transform into structural rules, is thus a way to safely include $\bar{\mathsf{p}}$ as a syntactical rule in DiLL presented as a sequent calculus. We have not developed a proof-net version of our calculus: DiLL fits well in this formalism but Graded Linear Logic does not. Note, however, that a tentative to symmetrize promotion was already explored in proof nets by Gimenez [Gim09].

3.1.2 Cut-elimination

In LL and in DiLL, there are only two exponential connectors, the ! and the ?, which are dual to each other. Then, in LL, each structural rule $\mathsf{w}, \mathsf{d}, \mathsf{c}$ interacts with the promotion rule in the cut elimination, but they do not interact together. When one considers DiLL, three costructural rules $\bar{\mathsf{w}}, \bar{\mathsf{d}}$ and $\bar{\mathsf{c}}$ are added. This adds many cases to the cut elimination: the principal formulas of costructural rules are of the form $!A$, and then each costructural rule interacts both with the promotion and with every other structural rule.

In the rules of DIDiLL, the principal formulas of the costructural rules have the form $!_{x^-}A$, while the ones of the structural rules have the form $?_{x^-}A^\perp$. But $!_{y^-}A$ *is not* the dual of $?_{y^-}A^\perp$, since the negation changes the polarity of the index. This implies that the number of possible interactions through the cut rule is very reduced compared to DiLL: $\bar{\mathsf{w}}, \bar{\mathsf{c}}$ and $\bar{\mathsf{d}}$ cannot interact with w, c and d. The only cut elimination cases to consider are the ones involving a promotion rule or a copromotion rule. Moreover, the promotion and the copromotion cannot interact together for similar reasons. The cases have to be refined, compared to what is done in LL or DiLL since we use graded exponentials. These cases with the promotion were presented in the graded but not polarized case by Breuvart and Pagani [BP15]. We present the same cases with our connectives indexed by polarized elements of a weak semiring in Figure 3.3.

The other cut elimination cases are the ones with a copromotion. Those are new, but they should be similar to the ones with a promotion. We do not provide the categorical semantics of DIDiLL here, but it would hopefully consist of the one of DiLL together with dual axioms for the copromotion which would be dual to those for the promotion. Following this remark, the cut elimination cases with a copromotion are strongly similar to those of the promotion. We provide them in Figure 3.4. Some variations should still be noted. First, a natural difference comes from the binary flavor of the costructural rules. This appears in the cut elimination case between a copromotion and a cocontraction which is then similar to the interaction between a promotion and a contraction but has a binary form. The second difference comes from the restriction we have made on the context of

$$\dfrac{\dfrac{\vdash ?_{Y^-}\Gamma', A}{\vdash ?_{(0\times Y)^-}\Gamma', !_{0^+}A}\,\mathsf{p} \qquad \dfrac{\vdash \Gamma}{\vdash \Gamma, ?_{0^-}A^\perp}\,\mathsf{w}}{\vdash \Gamma, ?_{0^-}\Gamma'}\,\mathit{cut} \qquad\rightsquigarrow\qquad \dfrac{\dfrac{\dfrac{\vdash \Gamma}{\vdash \Gamma, ?_{0^-}B_1}\,\mathsf{w}}{\vdots}\,\mathsf{w}}{\vdash \Gamma, ?_{0^-}\Gamma'}\,\mathsf{w}$$

$$\dfrac{\dfrac{\vdash ?_{Z^-}\Gamma', A}{\vdash ?_{(1\times Z)^-}\Gamma', !_{1^+}A}\,\mathsf{p} \qquad \dfrac{\vdash \Gamma, A^\perp}{\vdash \Gamma, ?_{1^-}A^\perp}\,\mathsf{d}}{\vdash \Gamma, ?_{(1\times Z)^-}\Gamma'}\,\mathit{cut} \qquad\rightsquigarrow\qquad \dfrac{\vdash ?_{Z^-}\Gamma', A \qquad \vdash \Gamma, A^\perp}{\vdash \Gamma, ?_{Z^-}\Gamma'}\,\mathit{cut}$$

$$\dfrac{\dfrac{\vdash ?_{Z^-}\Gamma', A}{\vdash ?_{(y\times Z)^-}\Gamma', !_{y^+}A}\,\mathsf{p} \qquad \dfrac{\vdash \Gamma, ?_{x^-}A^\perp \qquad x \leq y}{\vdash \Gamma, ?_{y^-}A^\perp}\,\mathsf{d}_I}{\vdash \Gamma, ?_{(y\times Z)^-}\Gamma'}\,\mathit{cut} \qquad\rightsquigarrow$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdash ?_{Z^-}\Gamma', A}{\vdash ?_{(x\times Z)^-}\Gamma', !_{x^+}A}\,\mathsf{p} \qquad \vdash \Gamma, ?_{x^-}A^\perp}{\vdash \Gamma, ?_{(x\times Z)^-}\Gamma'}\,\mathit{cut} \qquad (x\times z_1) \leq (y \times z_1)}{\ddots}\,\mathsf{d}_I \qquad (x\times z_n) \leq (y\times z_n)}{\vdash \Gamma, ?_{(y\times Z)^-}\Gamma'}\,\mathsf{d}_I$$

$$\dfrac{\dfrac{\vdash ?_{Z^-}\Gamma', A}{\vdash ?_{((x+y)\times Z)^-}\Gamma', !_{(x+y)^+}A}\,\mathsf{p} \qquad \dfrac{\vdash \Gamma, ?_{x^-}A^\perp, ?_{y^-}A^\perp}{\vdash \Gamma, ?_{(x+y)^-}A^\perp}\,\mathsf{c}}{\vdash \Gamma, ?_{((x+y)\times Z)^-}\Gamma'}\,\mathit{cut} \qquad\rightsquigarrow$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdash ?_{Z^-}\Gamma', A}{\vdash ?_{(x\times Z)^-}\Gamma', !_{x^+}A}\,\mathsf{p} \qquad \dfrac{\dfrac{\vdash ?_{Z^-}\Gamma', A}{\vdash ?_{(y\times Z)^-}\Gamma', !_{y^+}A}\,\mathsf{p} \qquad \vdash \Gamma, ?_{x^-}A^\perp, ?_{y^-}A^\perp}{\vdash \Gamma, ?_{(y\times Z)^-}\Gamma', ?_{x^-}A^\perp}\,\mathit{cut}}{\vdash \Gamma, ?_{(x\times Z)^-}\Gamma', ?_{(y\times Z)^-}\Gamma'}\,\mathit{cut}}{\vdots}\,\mathsf{c}}{\vdash \Gamma, ?_{(x\times Z + y\times Z)^-}\Gamma'}\,\mathsf{c}$$

$$\dfrac{\dfrac{\vdash ?_{Y^-}\Gamma', A}{\vdash ?_{(xz_1\times Y)^-}\Gamma', !_{(xz_1)^+}A}\,\mathsf{p} \qquad \dfrac{\vdash ?_{Z^-}\mathcal{M}, ?_{z_1^-}A^\perp, Q}{\vdash ?_{(x\times Z)^-}\mathcal{M}, ?_{(xz_1)^-}A^\perp, !_{x^+}Q}\,\mathsf{p}}{\vdash ?_{(xz_1\times Y)^-}\Gamma', ?_{(x\times Z)^-}\mathcal{M}, !_{x^+}Q}\,\mathit{cut} \qquad\rightsquigarrow$$

$$\dfrac{\dfrac{\dfrac{\vdash ?_{Y^-}\Gamma', A}{\vdash ?_{(z_1\times Y)^-}\Gamma', !_{z_1^+}A}\,\mathsf{p} \qquad \vdash ?_{Z^-}\mathcal{M}, ?_{z_1^-}A^\perp, Q}{\vdash ?_{(z_1\times Y)^-}\Gamma', ?_{Z^-}\mathcal{M}, Q}\,\mathit{cut}}{\vdash ?_{(xz_1\times Y)^-}\Gamma', ?_{(x\times Z)^-}\mathcal{M}, !_{x^+}Q}\,\mathsf{p}$$

Figure 3.3: Cases with the promotion

$$\dfrac{\dfrac{\vdash !_{y^-}B, A}{\vdash !_{(0\times y)^-}B, ?_{0^+}A}\,\bar{\mathsf{p}} \qquad \dfrac{}{\vdash !_{0^-}A^\perp}\,\bar{\mathsf{w}}}{\vdash !_{0^-}B}\,cut \quad \rightsquigarrow \quad \dfrac{}{\vdash !_{0^-}B}\,\bar{\mathsf{w}}$$

$$\dfrac{\dfrac{\vdash !_{y^-}B, A}{\vdash !_{(1\times y)^-}B, ?_{1^+}A}\,\bar{\mathsf{p}} \qquad \dfrac{\vdash \Gamma, A^\perp}{\vdash \Gamma, !_{1^-}A^\perp}\,\bar{\mathsf{d}}}{\vdash \Gamma, !_{y^-}B}\,cut \quad \rightsquigarrow \quad \dfrac{\vdash !_{y^-}B, A \qquad \vdash \Gamma, A^\perp}{\vdash \Gamma, !_{y^-}B}\,cut$$

$$\dfrac{\dfrac{\vdash !_{z^-}B, A}{\vdash !_{(yz)^-}B, ?_{y^+}A}\,\bar{\mathsf{p}} \qquad \dfrac{\vdash \Gamma, !_{x^-}A^\perp \qquad x \leq y}{\vdash \Gamma, !_{y^-}A^\perp}\,\bar{\mathsf{d}}_I}{\vdash \Gamma, !_{(yz)^-}B}\,cut \quad \rightsquigarrow$$

$$\dfrac{\dfrac{\dfrac{\vdash !_{z^-}B, A}{\vdash !_{(xz)^-}B, ?_{x^+}A}\,\bar{\mathsf{p}} \qquad \vdash \Gamma, !_{x^-}A^\perp}{\vdash \Gamma, !_{(xz)^-}B}\,cut \qquad xz \leq yz}{\vdash \Gamma, !_{(yz)^-}B}\,\bar{\mathsf{d}}_I$$

$$\dfrac{\dfrac{\vdash !_{z^-}B, A}{\vdash !_{((x+y)z)^-}B, ?_{(x+y)^+}A}\,\bar{\mathsf{p}} \qquad \dfrac{\vdash \Gamma, !_{x^-}A^\perp \qquad \vdash \Delta, !_{y^-}A^\perp}{\vdash \Gamma, \Delta, !_{(x+y)^-}A^\perp}\,\bar{\mathsf{c}}}{\vdash \Gamma, \Delta, !_{(xz+yz)^-}B}\,cut \quad \rightsquigarrow$$

$$\dfrac{\dfrac{\dfrac{\vdash !_{z^-}B, A}{\vdash !_{(xz)^-}B, ?_{x^+}A}\,\bar{\mathsf{p}} \qquad \vdash \Gamma, !_{x^-}A^\perp}{\vdash \Gamma, !_{(xz)^-}B}\,cut \qquad \dfrac{\dfrac{\vdash !_{z^-}B, A}{\vdash !_{(yz)^-}B, ?_{y^+}A}\,\bar{\mathsf{p}} \qquad \vdash \Delta, !_{y^-}A^\perp}{\vdash \Delta, !_{(yz)^-}B}\,cut}{\vdash \Gamma, \Delta, !_{(xz+yz)^-}B}\,\bar{\mathsf{c}}$$

$$\dfrac{\dfrac{\vdash !_{y^-}B, A}{\vdash !_{(xzy)^-}B, ?_{(xz)^+}A}\,\bar{\mathsf{p}} \qquad \dfrac{\vdash !_{z^-}A^\perp, Q}{\vdash !_{(xz)^-}A^\perp, ?_{x^+}Q}\,\bar{\mathsf{p}}}{\vdash !_{(xz)y^-}B, ?_{x^+}Q}\,cut \quad \rightsquigarrow$$

$$\dfrac{\dfrac{\dfrac{\vdash !_{y^-}B, A}{\vdash !_{(zy)^-}B, ?_{z^+}A}\,\bar{\mathsf{p}} \qquad \vdash !_{z^-}A^\perp, Q}{\vdash !_{(zy)^-}B, Q}\,cut}{\vdash !_{(x(zy))^-}B, ?_{x^+}Q}\,\bar{\mathsf{p}}$$

Figure 3.4: Cut-elimination cases for the copromotion

the copromotion. This context has only one formula, which simplifies the cut elimination cases compared to those of the promotion.

Now, we focus on the cut elimination procedure for DIDiLL.

Conjecture 3.1.2. *The logic DIDiLL enjoys a cut elimination procedure where the rewriting is weakly normalizing.*

This conjecture is based on a translation from the proofs of DIDiLL into the proofs of LL. This translation has the property that each rewriting step in DIDiLL corresponds to a rewriting step in LL. Let us describe this translation.

We define a map $\mathcal{T}$ that transforms formulas and proofs of DIDiLL into formulas and proofs of LL. For atoms and MALL connectors, $\mathcal{T}$ is the identity, and for exponential formulas we define:

$$\mathcal{T}(!_{x^+}A) \triangleq !(\mathcal{T}(A)) \qquad \mathcal{T}(!_{x^-}A) \triangleq ?(\mathcal{T}(A))$$

$$\mathcal{T}(?_{x^+}A) \triangleq !(\mathcal{T}(A)) \qquad \mathcal{T}(?_{x^-}A) \triangleq ?(\mathcal{T}(A)).$$

Then, for the proofs, $\mathcal{T}$ is the identity for each rule, except for the costructural ones. $\mathcal{T}$ transforms $\bar{\mathsf{d}}, \bar{\mathsf{d}}_I, \bar{\mathsf{p}}$ respectively into d, d_I and p. For the coweakening we define

$$\mathcal{T}\left(\dfrac{}{\vdash !_{0^-}A}\;\bar{\mathsf{w}}\right) \triangleq \dfrac{\dfrac{}{\vdash}\; mix_0}{\vdash ?\mathcal{T}(A)}\;\mathsf{w}$$

and for the cocontraction

$$\mathcal{T}\left(\dfrac{\dfrac{\pi_1}{\vdash \Gamma, !_{x^-}A} \qquad \dfrac{\pi_2}{\vdash \Delta, !_{y^-}A}}{\vdash \Gamma, \Delta, !_{x^-+y^-}A}\;\bar{\mathsf{c}}\right) \triangleq \dfrac{\dfrac{\dfrac{\mathcal{T}(\pi_1)}{\vdash \mathcal{T}(\Gamma), ?\mathcal{T}(A)} \qquad \dfrac{\mathcal{T}(\pi_2)}{\vdash \mathcal{T}(\Delta), ?\mathcal{T}(A)}}{\vdash \mathcal{T}(\Gamma), \mathcal{T}(\Delta), ?\mathcal{T}(A), ?\mathcal{T}(A)}\; mix_2}{\vdash \mathcal{T}(\Gamma), \mathcal{T}(\Delta), ?\mathcal{T}(A)}\;\mathsf{c}$$

which concludes the definition of $\mathcal{T}$. This map is such that if $\pi \rightsquigarrow_{cut} \pi'$ in DIDiLL, then $\mathcal{T}(\pi) \rightsquigarrow_{cut} \mathcal{T}(\pi')$ in LL. This property is easy to prove, except for the cases $\bar{\mathsf{p}}/\bar{\mathsf{w}}$ and $\bar{\mathsf{p}}/\bar{\mathsf{c}}$.

- For the cases with the promotion, it is obvious, since $\mathcal{T}$ only forgets the indices. These cases are then the same as those in LL. The only difference is for the cut between a promotion and an indexed dereliction. Since this rule does not exist in LL, this case cannot be compared to one in LL but $\mathcal{T}$ forgets the indices so $\mathcal{T}(\mathsf{d}_I)$ just corresponds to the identity in LL.

- For the copromotion with $\bar{\mathsf{d}}, \bar{\mathsf{d}}_I$ and $\bar{\mathsf{p}}$, $\mathcal{T}$ transforms them into the cases between promotion and $\bar{\mathsf{d}}, \bar{\mathsf{d}}_I$ and p.

- For the case $\bar{\mathsf{p}}/\bar{\mathsf{w}}$, $\mathcal{T}$ transforms it into a special case of the cut p/w. Applying $\mathcal{T}$ to the tree before the cut elimination gives

$$\mathcal{T}(\pi) = \dfrac{\dfrac{\vdash ?B, A}{\vdash ?B, !A}\;\mathsf{p} \qquad \dfrac{\dfrac{}{\vdash}\; mix_0}{\vdash ?A}\;\mathsf{w}}{\vdash ?B}\; cut$$

while after the cut it gives

$$\mathcal{T}(\pi') = \dfrac{\dfrac{}{\vdash}\; mix_0}{\vdash ?B}\;\mathsf{w}$$

and we actually have $\mathcal{T}(\pi) \rightsquigarrow_{cut} \mathcal{T}(\pi')$ in LL.

- For the last case, which is between a copromotion and a cocontraction, we have to use a commutation between mix_2 and cut. For such a cut, in a proof π, $\mathcal{T}(\pi)$ is

$$\dfrac{\dfrac{\vdash ?B, A}{\vdash ?B, !A}\,\mathsf{p} \qquad \dfrac{\dfrac{\vdash \Gamma, ?A^{\perp} \qquad \vdash \Delta, ?A^{\perp}}{\vdash \Gamma, \Delta ?A^{\perp}, ?A^{\perp}}\,mix_2}{\vdash \Gamma, \Delta, ?A^{\perp}}\,\mathsf{c}}{\vdash \Gamma, \Delta, ?B}\,cut$$

which reduces in LL to the proof π_1

$$\dfrac{\dfrac{\dfrac{\vdash ?B, A}{\vdash ?B, !A}\,\mathsf{p} \qquad \dfrac{\dfrac{\vdash ?B, A}{\vdash ?B, !A}\,\mathsf{p} \qquad \dfrac{\vdash \Gamma, ?A^{\perp} \qquad \vdash \Delta, ?A^{\perp}}{\vdash \Gamma, \Delta, ?A^{\perp}, ?A^{\perp}}\,mix_2}{\vdash \Gamma, \Delta, ?B, ?A^{\perp}}\,cut}{\Gamma, \Delta, ?B, ?B}\,cut}{\Gamma, \Delta, ?B}\,\mathsf{c}$$

while $\mathcal{T}(\pi')$, where $\pi \rightsquigarrow_{cut} \pi'$, is

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdash ?B, A}{\vdash ?B, !A}\,\mathsf{p} \qquad \vdash \Gamma, ?A^{\perp}}{\vdash \Gamma, ?B}\,cut \qquad \dfrac{\dfrac{\vdash ?B, A}{\vdash ?B, !A}\,\mathsf{p} \qquad \vdash \Delta, ?A^{\perp}}{\vdash \Delta, ?B}\,cut}{\vdash \Gamma, \Delta, ?B, ?B}\,mix_2}{\vdash \Gamma, \Delta, ?B}\,\mathsf{c}$$

which is equal to π_1 modulo the following commutation rule between mix_2 and cut (see [Pro23, p. 50]) by applying it twice.

However, since the cut elimination of linear logic is not strongly normalizing, it does not imply a normalization property for DIDiLL. One can imagine that the finite chain of reduction on a translated tree cannot be obtained from our translation. We, sadly, have not been able to prove that DIDiLL enjoys a normalizing property, but we think that a precise study of the cases implying that the cut elimination procedure of LL is not strongly normalizing should give the recipe for a proof of Conjecture 3.1.2.

3.2 Grading DIDiLL with LPDOs

After studying the syntax of the logic DIDiLL, we present in this section a first model for this logic. This goal of this model is to define a framework where the exponential are indexed by LPDOcc. This would, in a sense, extend the model of IDiLL which uses smooth functions and distributions from Section 2.2.3. However, it is not possible to use the smooth model to provide a model of DIDiLL indexed by LPDOcc. The reason comes from the definition of linear partial differential operators and from the promotion rule. With a promotion rule, we have to be able to consider higher order objects, such as spaces of smooth functions whose domain is another space of smooth function. For instance, if a formula A is interpreted by $\mathbb{R}^n$, the formula $(!!A)^{\perp}$ will be interpreted by $\mathcal{C}^{\infty}(\mathcal{C}^{\infty}(\mathbb{R}^n, \mathbb{R}), \mathbb{R})$. In IDiLL we have restricted the grammar of the formulas to forbid such formulas, allowing only finitary formulas. But since there is a promotion rule in DIDiLL this restriction is not possible anymore, the construction of the promotion is strongly based on the ability to have non-finitary formulas. This is what clashes with LPDOcc: an element f of $(!!A)^{\perp}$ is a function whose domain is an infinite dimensional vector space. Contrarily to a space like $\mathbb{R}^n$, it does not have a canonical basis, and even building linearly independent families may be hard in such spaces. But this is something crucial to interpret the action of a LPDO on f. For

$$D = \sum_{\alpha} a_\alpha \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

we need to define an interpretation of the x_i as elements of the domain of f to compute $D(f)$. While the x_i are interpreted by elements of the canonical basis of $\mathbb{R}^n$ when it is the domain of f, this is then much harder when f is in the interpretation of a non-finitary formula.

For this reason, the model presented in this section is built using Köthe spaces. Introduced by Ehrhard [Ehr02] as a model of (differential) linear logic, we have presented these spaces in Section 1.3.4. Köthe spaces are spaces of sequences, built on countable sets, that can represent the smooth functions that are analytic. This makes them the perfect candidate for a model of DIDiLL graded by LPDOcc. It is discrete, in the sense that it uses sequences, based on a countable set, and from this countable set we can interpret the variables along with the differential operators can differentiate, and it is powerful enough to represent some smooth functions.

In Section 3.2.1 we extend the monoid of LPDOcc from Section 2.3.1.1 to a weak semiring. Here we build an operation on polynomials which is right distributive over the product of polynomials using a form of composition. Then, we use again the isomorphism between multivariate polynomials to deduce a weak semiring of LPDOcc having the usual composition of differential operators as the sum operation. In Section 3.2.2 we refine the definition of Köthe spaces. In order to use the countable sets of the Köthe spaces, we pair topological vector spaces E with an injection $\varphi : \mathbb{N} \to E$. This provides an interpretation of the partial derivatives $\frac{\partial -}{\partial x_i}$ involved in any differential operator, or more generally for the variable X_i of a polynomial, interpreting x_i or X_i as $\varphi(i)$. In Section 3.2.3 we define the action of LPDOcc and polynomials over Köthe spaces. This gives us a definition for the exponential connectives of DIDiLL in the model of Köthe spaces graded by differential operators. Finally, we express how the exponential rules are interpreted in this setting, using the interpretation of the rules of DiLL and defining the Laplace transform over graded Köthe spaces. Beware that non-indexed dereliction d and codereliction $\bar{\mathsf{d}}$, as well as the copromotion $\bar{\mathsf{p}}$ are sadly *not interpreted* in this model.

3.2.1 A weak semiring of LPDOcc

In Chapter 2 we have defined a monoid $\mathcal{D}$ of LPDOcc, whose monoidal operation was the composition of operators. We have used the fact that this monoid is isomorphic to the one of multivariate polynomials endowed with the polynomial product to prove properties on $\mathcal{D}$. To extend $\mathcal{D}$ to a weak semiring, we use a similar isomorphism, and we start by extending Definition 2.3.4.

Definition 3.2.1. *We define by induction:*

$$\mathcal{P}_0 = \mathbb{R} \qquad \mathcal{P}_{n+1} = \mathcal{P}_n[X_{n+1}] \qquad \mathcal{P}_\omega = \cup_{n \geq 0} \mathcal{P}_n$$

and we give the following operations for each $P, Q \in \mathcal{P}_\omega$ *such that*

$$P = \sum_{\alpha \in \mathbb{N}^{(\omega)}} a_\alpha X^\alpha \qquad Q = \sum_{\beta \in \mathbb{N}^{(\omega)}} b_\beta X^\beta$$

we define $P \boxplus Q = P \times Q$, the usual product of polynomials, and $P \boxtimes Q$ as a composition:

$$P \boxplus Q = \sum_{\alpha,\beta} a_\alpha b_\beta X^{\alpha+\beta} \qquad P \boxtimes Q = \sum_{\alpha} a_\alpha Q^{|\alpha|}.$$

where $Q^{|\alpha|}$ denotes the multiplication of Q by itself $|\alpha| = \sum_i \alpha_i$ times. We denote

$$\widehat{P} = \sum_{\alpha} (-1)^{|\alpha|} a_\alpha X^\alpha.$$

In this definition, the addition $\boxplus$ is still the product of polynomials, and the new operation $\boxtimes$ is a form of composition[1] of multivariate polynomials. We recall that the use of $\widehat{P}$ is necessary to interpret differentiation on distributions. The $(-1)^{|\alpha|}$ coefficients resulting from an integration by part when distributions are considered as generalized functions.

Using the multinomial theorem, we prove the following.

Proposition 3.2.2. *$(\mathcal{P}_\omega, \boxplus, \boxtimes, 1, X_1)$ is a weak semiring.*

Proof. The fact that $(\mathcal{P}_\omega, \boxplus, 1)$ is a commutative monoid comes from the well-known result that the product of polynomials is commutative. The polynomial X_1 is an obvious left neutral for $\boxtimes$ (as would any X_i), and 1 an obvious left absorbing element. The product $\boxtimes$ is associative. Let $P, Q, R \in \mathcal{P}_\omega$, with

$$P = \sum_{i=0}^{n} a_{\alpha_i} X^{\alpha_i} \qquad Q = \sum_{j=0}^{m} b_{\beta_j} X^{\beta_j}.$$

Then, we have

$$\begin{aligned} P \boxtimes (Q \boxtimes R) &= P \boxtimes \left(\sum_{j=0}^{m} b_{\beta_j} R^{|\beta_j|} \right) \\ &= \sum_{i=0}^{n} a_{\alpha_i} \left(\sum_{j=0}^{m} b_{\beta_j} R^{|\beta_j|} \right)^{|\alpha|} \\ &= \sum_{i=0}^{n} a_{\alpha_i} \left(\sum_{k_0+\dots+k_m=|\alpha_i|} \binom{|\alpha_i|}{k_0,\dots,k_m} b_{\beta_0}^{k_0} \dots b_{\beta_m}^{k_m} R^{k_0|\beta_0|+\dots+k_m|\beta_m|} \right) \end{aligned}$$

with the multinomial theorem, which proves one side of the associativity property. For the other side,

$$\begin{aligned} (P \boxtimes Q) \boxtimes R &= \left(\sum_{i=0}^{n} a_{\alpha_i} Q^{|\alpha_i|} \right) \boxtimes R \\ &= \left[\sum_{i=0}^{n} a_{\alpha_i} \left(\sum_{j=0}^{m} b_{\beta_j} X^{\beta_j} \right)^{|\alpha_i|} \right] \boxtimes R \\ &= \left[\sum_{i=0}^{n} a_{\alpha_i} \left(\sum_{k_0+\dots+k_m=|\alpha_i|}^{|\alpha_i|} \binom{|\alpha_i|}{k_0,\dots,k_m} b_{\beta_0}^{k_0} \dots b_{\beta_m}^{k_m} X^{k_0\beta_0+\dots+k_m\beta_m} \right) \right] \boxtimes R \end{aligned}$$

[1] It is a form of composition, as there is no natural definition for the composition of multivariate polynomials. For the form chosen here, $P \boxtimes Q$ consists is replacing every variable in P by Q.

$$= \sum_{i=0}^{n} a_{\alpha_i} \left(\sum_{k_0+\cdots+k_m=|\alpha_i|} \binom{|\alpha_i|}{k_0,\ldots,k_m} b_{\beta_0}^{k_0} \ldots b_{\beta_m}^{k_m} R^{|k_0\beta_0+\cdots+k_m\beta_m|} \right)$$

$$= \sum_{i=0}^{n} a_{\alpha_i} \left(\sum_{k_0+\cdots+k_m=|\alpha_i|} \binom{|\alpha_i|}{k_0,\ldots,k_m} b_{\beta_0}^{k_0} \ldots b_{\beta_m}^{k_m} R^{k_0|\beta_0|+\cdots+k_m|\beta_m|} \right)$$

with the same method, which proves the associativity.

For the right distributivity, let $P, Q, R \in \mathcal{P}_\omega$, with

$$P = \sum_{i=0}^{n} a_{\alpha_i} X^{\alpha_i} \qquad Q = \sum_{j=0}^{m} b_{\beta_j} X^{\beta_j}.$$

$$\begin{aligned}
(P \boxplus Q) \boxtimes R &= \left(\sum_{i=0}^{n} \sum_{j=0}^{m} a_{\alpha_i} X^{\alpha_i} b_{\beta_j} X^{\beta_j} \right) \boxtimes R \\
&= \left(\sum_{i=0}^{n} \sum_{j=0}^{m} a_{\alpha_i} b_{\beta_j} X^{\alpha_i+\beta_j} \right) \boxtimes R \\
&= \sum_{i=0}^{n} \sum_{j=0}^{m} a_{\alpha_i} b_{\beta_j} R^{\alpha_i+\beta_j} \\
&= \sum_{i=0}^{n} \sum_{j=0}^{m} a_{\alpha_i} R^{\alpha_i} b_{\beta_j} R^{\beta_j} \\
&= \left(\sum_{i=0}^{n} a_{\alpha_i} R^{\alpha_i} \right) \boxplus \left(\sum_{j=0}^{m} b_{\beta_j} R^{\beta_j} \right) \\
&= (P \boxtimes R) \boxplus (Q \boxtimes R)
\end{aligned}$$

concluding the proof that we have indeed a weak semiring. □

Note that we strongly use the fact that we only ask for a weak semiring. The axioms that are in the definition of a semiring but not in a weak one are the fact that 0 is right absorbing for the product, 1 is right neutral for the product, and the product is left-distributive. Those axioms are not satisfied here:

- taking $P = 2$, $P \boxtimes 1 = 2 \boxtimes 1 = 2 \neq 1$ (1 being the 0 element, it is not right absorbing);
- taking $P = X_2$, $P \boxtimes X_1 = X_2 \boxtimes X_1 = X_1 \neq X_2$ (X_1 being the 1 element, it is not right neutral);
- taking $P = X_1 + X_2$, $Q = X_3$ and $R = X_4$, we have

  $$P \boxtimes (Q \boxplus R) = (X_1 + X_2) \boxtimes (X_3 X_4) = X_3 X_4 + X_3 X_4 = 2 X_3 X_4$$

  and

  $$(P \boxtimes Q) \boxplus (P \boxtimes R) = ((X_1 + X_2) \boxtimes X_3) \boxplus ((X_1 + X_2) \boxtimes X_4) = 2X_3 \boxplus 2X_4 = 4X_3X_4$$

  which gives a counter-example for the right distributivity.

One may remark that X_1 being, in fact, left neutral for the product. This is also the case for each X_i. Other possibilities could have then been taken for the definition of a weak semiring of multivariate polynomials. Finally, this weak semiring of polynomials gives, in fact, a weak semiring of LPDOcc, with the following definition.

Definition 3.2.3. *Considering $P \in \mathcal{P}_\omega$, we define $P(\partial)$ as:*

$$P = \sum_\alpha a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n} \qquad P(\partial) = \sum_\alpha a_\alpha \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots x_n^{\alpha_n}}.$$

In particular, for $P, Q \in \mathcal{P}_\omega$ one has $(P \boxplus Q)(\partial) = P(\partial) \circ Q(\partial)$ and $1(\partial) = id$.

3.2.2 Köthe spaces

We have already introduced the Köthe spaces in this manuscript. However, we refine here this notion by defining indexed Köthe spaces, adding an injection from the base space of Köthe sequences to the natural numbers. This injection will model a family of free vectors, along which we will differentiate when applying a differential operator. Let us recall the definition of a Köthe space, of linear logic constructions over those spaces and extend those definitions with the new injection that we add. For a set X, we write for $E \subset \mathbb{K}^X$:

$$E^\perp \triangleq \left\{ u \in \mathbb{K}^X \mid \forall v \in E, \sum_{x \in X} |u_x v_x| \text{ converges} \right\}$$

This acts as an orthogonality, meaning that for any E one has $E \subseteq E^{\perp\perp}$ and $E^\perp = E^{\perp\perp\perp}$. As in Section 1.3.4, we define the Köthe spaces using this orthogonality.

Definition 3.2.4. *A* Köthe space *is a pair $\mathcal{X} = (X, E_X)$ such that X is at most countable, $E_X \subseteq \mathbb{K}^X$ and $E_X^{\perp\perp} = E_X$. Its dual $\mathcal{X}^\perp$ is defined by $X^\perp = X$ and $E_{X^\perp} = E_X^\perp$.*

From this, we endow Köthe spaces with a map from X to $\mathbb{N}$ to represent a free family of variables.

Definition 3.2.5. *We consider* indexed Köthe spaces *$E = (X, \varphi, E_X)$ where φ_X is an injection from X to $\mathbb{N}^*$. When i is not in the image of ϕ, then one assigns $\phi^{-1}(i) = 0$. One can think then as $X = (\phi^{-1}(1), \phi^{-1}(2), \dots, \phi^{-1}(n)(, \dots))$, the last $(\dots)$ being optional whether X is finite or infinite. As the use of ϕ is to handle partial derivatives with respect to $\phi^{-1}(i)$ (see Definition 3.2.6), when i is not in the image of ϕ one will consider functions defined on X as constant with respect to the i-th variable.*

In what follows, we may just say Köthe space to mean indexed Köthe space. Consider $E = (X, \phi, E_X)$ and $F = (Y, \psi, F_Y)$ two Köthe spaces. We now recall the constructions making Köthe spaces a model of linear logic and refining them with the use of a formal ordering of their bases.

- The space $E \multimap F$ of linear continuous maps from E_X to F_Y corresponds to the subset of $\mathbb{K}^{X \times Y}$ of all M such that the sum:

  $$\sum_{i,j} M_{i,j} x_i y'_j$$

  is absolutely converging for all $x \in E$ and $y' \in F^\perp$. The cardinal on $X \times Y$ is defined as $\chi : (x, y) \mapsto 2^{\phi(x)-1}(2^{\psi(y)} - 1)$. This is an injection when ϕ and ψ are injections.

- The *tensor product* of two perfect sequence spaces E_X and F_Y is the perfect sequence space $(E \multimap F^\perp)^\perp$. In particular, $E^\perp \simeq E \multimap \mathbb{K}$, where an element $\ell = (\ell_i)_i$ of $E^\perp$ acts on $x = (x_i)_i$ in E as $\ell(x) = \sum_i \ell_i x_i$. The cardinal on $X \times Y$ the same as above.

- The product and coproduct constructions on Köthe spaces, interpreting & and $\oplus$, are defined as the product and coproduct of topological vector spaces and preserve perfect sequence spaces. Their bases are enumerated as above.

- The interpretation of exponentials formulas in Köthe spaces embodies the intuition that non-linear proofs should be represented as analytic functions. Consider a set X and $\mathcal{M}(X)$ the set of all finite multisets of X. If $\mu \in \mathcal{M}(X)$ and $x \in E$, we write:

$$x^\mu = \prod_{x \in X} {x_{\phi(x)}}^{\mu(x)}$$

and $x^!$ the element of $\mathbb{K}^{\mathcal{M}(X)}$ such that $(x^!)_\mu = x^\mu$. Then one defines:

$$!E = (\mathcal{M}(X), \{x^! | x \in E\}^{\perp\perp}).$$

Let $\{p_1, \dots, p_n, \dots\}$ denote the list of the prime numbers, ordered with the usual order on $\mathbb{N}$. Then the cardinal on $\mathcal{M}(X)$ is defined as $\chi : \mu \mapsto \prod_{x \in X} p_{\phi(x)}^{(\mu(x))}$, which gives an injection by unicity of the integer factorization.

The addition of an enumeration of their bases does not change the categorical structure of Köthe spaces, for the simple reason that two different enumerations on the same base are always isomorphic. Let us now detail how to combine the notion of indexed Köthe spaces and LPDOcc.

3.2.3 DIDiLL indexed by polynomials

Consider DIDiLL indexed by the semiring $(\mathcal{P}_\omega, \boxplus, \boxtimes, 1, X_1)$. We interpret MALL formulas of DIDiLL and non-indexed exponentials as the refinement of the interpretation by Ehrhard [Ehr02] recalled in Section 3.2.2. The non-indexed exponentials are not directly needed for interpreting DIDiLL, but they will be used in Definition 3.2.8 to define the interpretation of the indexed exponentials.

Definition 3.2.6. *Consider (X, ϕ, E_X) a Köthe space. Consider $P = \sum_{\alpha \in \mathbb{N}^\omega} a_\alpha X^\alpha \in \mathcal{P}_\omega$, $f : E \Rightarrow \mathbb{K}$ a function and $\phi : !E \equiv (E \Rightarrow \mathbb{K})^\perp$ a distribution.*

- *For $x \in E$, denoting λ_i the coordinate of x along x_i such that $x = \sum_i \lambda_i x_i$, we define*

$$P(x) \triangleq \sum_{\alpha \in \mathbb{N}^\omega} a_\alpha \prod_i \lambda_i^{\alpha_i}$$

- *$P \cdot f$ denote the analytic function $x \mapsto P(x) \cdot f(x)$, where $\_ \cdot \_$ is the scalar multiplication.*

- *$P \cdot \phi$ denotes the distribution $g \mapsto \phi(P \cdot g)$.*

- *$P(\partial) f$ denotes the partial differential operator $P(\partial)$ applied to f, according to the basis X:*

$$P(\partial) f : z \mapsto \sum_\alpha a_\alpha \frac{\partial^{|\alpha|} f}{\partial \phi^{-1}(1)^{\alpha_1} \dots \partial \phi^{-1}(n)^{\alpha_n}}(z).$$

Parameters

$$
\begin{array}{ccc}
?_{P^-}A = \{f \mid \exists g \in ?A, f = \hat{P} \cdot g\} & \xleftarrow{\mathscr{L}^-} & !_{P^-}A = \{\phi \mid \exists \psi \in !A, \phi = P(\partial)(\psi)\} \\
\updownarrow (\_)^\perp & & \updownarrow (\_)^\perp \\
!_{P^+}A = \{\phi \mid \exists \psi \in !A, \psi = \hat{P} \cdot \phi\} & \xrightarrow[\mathscr{L}^+]{} & ?_{P^+}A = \{f \mid \exists g \in ?A, g = P(\partial)(f)\}
\end{array}
$$

Solutions

Figure 3.5: Laplace and duality acting on exponentials graded by polynomials

- *$P(\partial)\phi$ denotes the distribution $\phi \circ \widehat{P}(\partial)$. This is the standard definition of differential operators applied to distributions.*

Note that, if the indexed Köthe space (X, E_X, ϕ) is such that m is not in the image of ϕ and x_m appears in P, following the convention of Definition 3.2.5, functions of X will be considered constant with respect to x_m.

Example 3.2.7. *As an example, let us detail how differential operators act on entire functions $f : E \Rightarrow \mathbb{K}$. Consider (X, E_X) a Köthe space. Consider $P \in \mathcal{P}_\omega$ and write f as $f(x) = \sum_{\mu \in \mathcal{X}} f_\mu x^\mu$. For $\mu \in \mathcal{X}$ let's write μ_i the (possibly null) coefficient of x_i in μ. Then for $a \in E$ $\frac{\partial f}{\partial x_i}(a) = \sum_\mu f_\mu \mu_i \frac{a^\mu}{a_i}$ and as such, with the notations of Definition 3.2.3 we have:*

$$(P(\delta)(f))\,(x) = \sum_\alpha a_\alpha f_{(\mu+\alpha)} \prod_i (\mu_i + \alpha_i) x^\mu.$$

Now that we know how to apply polynomials and partial differential operators to Köthe spaces, let us define the interpretation of DIDiLL formulas in this context.

Definition 3.2.8. *We interpret as follows graded exponentials:*

$$[\![?_{P^-}A]\!] = \hat{P} \cdot ([\![?A]\!]) \qquad [\![!_{P^-}A]\!] = P(\partial)([\![!A]\!])$$

$$[\![!_{P^+}A]\!] = (\hat{P})^{-1} \cdot ([\![!A]\!]) \qquad [\![?_{P^+}A]\!] = P(\partial)^{-1}([\![?A]\!])$$

These definitions and their properties are summarized in Figure 3.5, where the structural rules apply to the connectives on the left and the costructural on the connectives on the right. This also defines four functors acting on the category of Köthe spaces and linear maps:

$$?_{P^-}E = \hat{P} \cdot (?E) \qquad !_{P^-}E = P(\partial)(!E)$$

$$!_{P^+}E = \hat{P}^{-1} \cdot (!E) \qquad ?_{P^+}E = P(\partial)^{-1}(?E)$$

We now turn to the heart of the interpretation of each exponential rule of DIDiLL. As we will see below, the costructural rules $\bar{\mathsf{d}}_I$, $\bar{\mathsf{c}}$ and $\bar{\mathsf{w}}$ are interpreted as in IDiLL [Ker18a], representing the action of differential operators. The structural rules, however, are interpreted more elegantly than in IDiLL, as d_I, c and w now represent operations on polynomials. Furthermore, a new and important addition to IDiLL is the interpretation of the promotion rule, representing the composition of higher-order analytic functions.

The interpretation of the exponential rules of DIDiLL follows the usual smooth interpretation of DiLL rules (see Section 1.3.5), as the latter adapts easily to the bi-graded case.

- The weakening rule w is interpreted by the introduction of the polynomial constant at 1, which is indeed a function of $[\![?_{0^-}N]\!] = \{1 \cdot f | f : N^\perp \Rightarrow \mathbb{K}\}$. The coweakening $\bar{\mathsf{w}}$ is interpreted as the introduction of the Dirac at 0, which is indeed an element of $[\![!_{0^-}N]\!] = \{\phi \circ id \mid \phi : (N \Rightarrow \mathbb{K})^\perp\}$.

- The contraction c is interpreted as usual, as the scalar multiplication of $f_1 \in [\![?_{P^-}N]\!]$ and $f_2 \in [\![?_{Q^-}N]\!]$ leads indeed to a function in $[\![?_{P^- \times Q^-}N]\!]$, since $\hat{P} \cdot g_1 . \hat{Q} \cdot g_2 = \hat{PQ} \cdot (g_1 . g_2)$. The cocontraction $\bar{\mathsf{c}}$ is interpreted as usual by the convolution product, and the convolution $\phi \circ P(\partial) * \psi \circ Q(\partial)$ equals indeed $(\phi * \psi) \circ (P(\partial) \circ Q(\partial)) = \phi * \psi \circ (P \times Q)(\partial)$. Note that, contrary to IDiLL's model (see Equation 2.3), here the interpretation of the contraction follows the usual one (see Section 1.3.5), and grading appears naturally.

- Remember that the order on $\mathcal{P}_\omega$ is interpreted with respect to the additive law: $P \leq Q$ if there is H such that $Q = H \times P$. Indexed dereliction d_I is interpreted by the composition by a polynomial: to a function $f = P \cdot f'$ it maps $H \cdot f$. Indexed codereliction $\bar{\mathsf{d}}_I$ is then interpreted as in IDiLL by the precomposition of a distribution $\phi \circ P(\partial)$ with $H(\partial)$.

Non-indexed (co)dereliction Non-indexed dereliction d and codereliction $\bar{\mathsf{d}}$ are, however, *not* interpreted in this model, mainly because the usual differential at 0 used in the semantics of DiLL is not an element of the indexing semiring. Indeed, since the 1 element of the semiring is X_1, the interpretation of d would map an element $x = \sum \lambda_i x_i$ to a map $x_1 \cdot f$ with $f \in ?E$. Following the usual interpretation in Köthe spaces, the interpretation would be the function $y = \sum \lambda'_i x_i \in E^\perp \mapsto \sum \lambda_i \lambda'_i$. If $\lambda_1 = 0$, this interpretation cannot be factorized as $x_1 \cdot f$. Likewise, interpreting $\bar{\mathsf{d}}$ would mean writing $x \mapsto (f \mapsto D_0(\_)(x))$ as a distribution $\phi \circ D$ with D a LPDOcc. As $f \mapsto D_0(f)$ is clearly not a linear partial differential operator, this tentative fails.

Promotion and digging Graded promotion requires the definition of a new form of composition. Consider informally two functions $f = P \cdot f' : A \Rightarrow B$ and $g = Q \cdot g' : B \Rightarrow C$. The idea is that the composition $g \circ f$ should be defined pairwise on polynomials and functions: $g \tilde{\circ} f = (Q \boxtimes P) \cdot (g' \circ f')$. Note that when $Q = P = 1$, this is the usual composition. When $A = B = C = \mathbb{K}$ and $f' = g' = cst_1$, then this is also the usual composition of univariate polynomials. This composition is associative and has $(1 \cdot \mathsf{id})$ as left and right neutral.

We now define the graded digging:

$$\mu_{Q,P} : \begin{cases} !_{Q^+ \boxtimes P^+} E & \rightarrow \quad !_{Q^+} !_{P^+} E \\ \quad \phi & \mapsto \quad (Q \cdot f' \in ?_{Q^-} ?_{P^-} E^\perp \\ & \quad\ \mapsto f'(P \cdot g' \in ?_{P^-} E^\perp \\ & \quad\ \mapsto \phi(x \in E^\perp \\ & \quad\ \mapsto (Q \boxtimes P)(x) \cdot f'(g'(x))))) \end{cases}$$

Consider a linear map $\ell \in [\![!_{X^+} \mathcal{N}^\perp]\!] \multimap [\![P]\!]$ interpreting a proof of $\vdash ?_{X^-}\mathcal{N}, P$. Then the proof resulting from a graded promotion rule is, as usual, the composition $!_Q \ell \circ \mu_{Q,P}$.

We now study a graded version of the Laplace transform interpreted in Köthe spaces:

$$\mathscr{L}^- : \begin{cases} !_{P^-}E & \to & ?_{P^-}E \\ \phi & \mapsto & (\ell \in E^\perp \mapsto \phi(x \in E \mapsto e^{\ell(x)})) \end{cases} \tag{3.3}$$

$$\mathscr{L}^+ : \begin{cases} !_{P^+}E & \to & ?_{P^+}E \\ \phi & \mapsto & (\ell \in E^\perp \mapsto \phi(x \in E \mapsto e^{\ell(x)})) \end{cases} \tag{3.4}$$

This is a straightforward adaptation of Equation 3.2. Through a careful computation of the Laplace transform on a Köthe space, and using strongly the fact that E and $E^\perp$ have the same basis, one proves the following proposition.

Proposition 3.2.9. *For $\phi \in !_{P^-}E$ we have that $\mathscr{L}^-(\phi)(\ell)$ is well-defined for every $\ell \in E^\perp$, and that $\mathscr{L}^-(\phi) \in ?_{P^-}E$.*

Proof. Consider $P \in \mathcal{P}_\omega$, (X, ϕ, E_X) a Köthe space. To simplify the notation, and up to completion by 0, let's denote $X = (x_i)_{i\in\mathbb{N}}$ where we have then $x_i = x_{\phi^{-1}(x)}$ for all x, i.

Consider $\phi \in !_{P^-}E$, that is there is $\phi \in !E$ such that $\phi = \psi \circ P(\delta)$. Then we have, in particular, $\phi \in !E$. Let us show that $\mathscr{L}^-(\phi)(\ell)$ is well defined, for which we need to show that ϕ can be applied to $\mathscr{L}^-(\phi)(\ell) : x \mapsto e^{\ell(x)}$, meaning that $x \mapsto e^{\ell(x)}$ is in $\{x^! | x \in E\}^\perp$.

The linear function in $\{x^! | x \in E\}^\perp$ corresponding to $\mathscr{L}^-(\phi)(\ell)$ is exactly $(\widetilde{\mathscr{L}^-(\phi)(\ell)}) : x^! \mapsto e^{\ell(x)}$. It is linear in $x^!$, as writing $x = (\lambda_i)_i$ and $\ell = (\ell_i)_i$ one has

$$e^{\ell(x)} = \sum_n \frac{\ell(x)^n}{n!} = \sum_n \frac{1}{n!}\left(\sum_i \ell_i x_i\right)^n = \sum_{\mu\in\mathcal{M}(X)} \frac{1}{|\mu|!}\ell^\mu x^\mu$$

Let us show that $\mathscr{L}^-(\phi)$ is indeed in $?_{P^-}E$.

$$\begin{aligned}\mathscr{L}^-(P(\partial)(\phi)) &= \ell \mapsto \phi \circ \hat{P}(\partial)(x \mapsto e^{\ell(x)}) \\ &= (\ell \in E^\perp) \mapsto \phi(x \in E \mapsto \left(\sum_\alpha (-1)^{|\alpha|} \prod \ell(x_i)^{\alpha_i} \cdot e^{\ell(x)}\right)) \\ &= \ell \mapsto (\sum_\alpha (-1)^{|\alpha|} \prod l(x_i)^{\alpha_i}) \cdot (\mathscr{L}^-(\phi))(\ell)\end{aligned}$$

On the Köthe space $E^\perp$, the basis is also X, and now $(\ell(x_i)_i$ represents the coefficients of ℓ along this basis. Hence, $\sum_\alpha (-1)^{|\alpha|} \prod \ell(x_i)^{\alpha_i} = \hat{P}(\ell)$ and

$$\mathscr{L}^-(P(\partial)(\phi)) = \hat{P} \cdot (\mathscr{L}^-(\phi))$$

Hence $\mathscr{L}^-$ maps an object in $!_{P^-}E$ to an object in $?_{P^-}E$. □

One shows likewise that $\mathscr{L}^+$ is well-defined. This shows that the diagram in Figure 3.5 is well-defined. It commutes thanks to the involutivity of $(\_)^\perp$ in Köthe spaces.

Sadly, looking in particular at the proof of proposition 3.2.9, we see no reason why $\mathscr{L}^-$ or $\mathscr{L}^+$ should be isomorphisms. As such, we cannot obtain the graded codigging as the reverse image by the Laplace transform, as done in Section 3.3. It seems that, as is usual Köthe spaces [KL19, Section 5.1], graded Köthe spaces offer no interpretation for the copromotion.

3.3 A Smooth Higher Order Model of DIDiLL

The previous section provided a higher order model where we were able to interpret linear partial differential operators. However, we are also interested in smooth models. With Köthe spaces, we were studying spaces of sequences, which are then discrete even if they represent some smooth functions. It was crucial in order to have spaces where we can interpret the variables of differential operators. In this section we take another point of view and provide a model made of smooth functions and distributions. We extend to higher order the smooth and polarized model of DiLL made of Fréchet and DF-spaces, which was previously restricted to first order. For this extension, we have to grade the exponentials by Young Functions. These functions will act like bounds, and will allow us to ensure some convergence properties. In this sense, DIDiLL is the right framework to propose a higher order smooth model, as we need the gradation for the higher order constructions. We will still need a restriction of DIDiLL: formally, the logic we will use is a polarized version of DIDiLL, in the fashion of polarized linear logic as we have presented it in Section 1.1.3. This was already the case in differential linear logic; spaces of smooth functions and distributions are a model for a polarized differential linear logic.

Section 3.3.1 defines the spaces $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$, which are spaces of smooth functions bound by a Young function. These spaces were defined by Ouerdiane *et al.* to prove a duality theorem using the Laplace transform [GHOR00]. Kerjean and Lemay have used those spaces as a concrete model to interpret the codigging rule [KL23]. In Section 3.3.2 we define the polarized version of DIDiLL. We show that the spaces $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$ are a higher order model of the polarized DIDiLL in Section 3.3.3, where we use a weak semiring of Young functions.

3.3.1 The spaces $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$

Before defining the spaces $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$, we need some topological definitions. Some of the definitions we give are close to those from Section 1.3.5.

Definition 3.3.1. *A* Hausdorff and locally convex topological vector space *(lcs) is a $\mathbb{K}$-vector space endowed with a Hausdorff topology making scalar multiplication and addition continuous, and such that every point has a basis of convex open neighborhoods.*

The lcs we consider are on $\mathbb{C}$. To get a classical model of differential linear logic, one wants to interpret the formulas by reflexive spaces, which is spaces such that $E \simeq E''$. But those spaces have poor stability properties: they are not stable by tensor product do not generalize well at higher order. This is why we use Fréchet spaces and nuclear spaces as in Section 1.3.5. For a smooth model of DIDiLL we use the same recipe, basing our approach on finding the right elements of **NuclF** and **NuclDF** to interpret the exponential formulas. The key to construct a higher model of DiLL based on **NuclF** and **NuclDF** is, in fact, to consider a good decomposition of these objects. It will allow to construct Fréchet spaces of smooth maps over DF spaces, and as such a higher-order smooth model of DiLL, but graded and polarized.

Proposition 3.3.2. *[Jar81, Chapter 21] The topology on any* **NuclF** *space N can be defined through a countable family of Hilbertian norms $|\_|_p$, $p \in \mathbb{N}$, and if one denotes N_p the Hilbert space resulting in the completion of N with respect to $|\_|_p$, we have that N is the limit of all N_p, while N' is the colimit of all $(N_p)'$:*

$$\bigcap_p N_p = N \qquad \bigcup_p (N_p)' = N'.$$

Now, contrary to the work done in Section 3.2 or in [BKM23], exponentials in the current model will be graded by *Young functions*.

Definition 3.3.3. *A* Young function *is a real continuous function* $\theta : \mathbb{R}_+ \to \mathbb{R}_+$ *which is convex, growing, and such that* $\theta(0) = 0$ *and* $\lim_{+\infty} \frac{\theta(x)}{x} = +\infty$. *We define the* convex conjugate *(or just* conjugate*) of a Young functions* θ *as the function*

$$\theta^* : x \mapsto \sup_{t \geq 0}(tx - \theta(t)).$$

The conjugate of θ can be thought of as an inverse when θ is defined through a measure. Indeed, if $\theta = \int \mu(t)dt$, then $\theta^* = \int \mu^{-1}(t)dt$. It also has the following property.

Proposition 3.3.4. *The convex conjugate of a Young function is a Young function.*

Proof. Let θ be a Young function. θ^* is continuous by continuity of the sup and of θ. For $x \in \mathbb{R}_+$, $\theta^*(x) \geq 0 \times x - \theta(0) = 0$ so $\theta^* : \mathbb{R}_+ \to \mathbb{R}_+$. Now, for $x_1, x_2 \in \mathbb{R}_+$ and $0 \leq r \leq 1$,

$$\begin{aligned}
\theta^*(rx_1 + (1-r)x_2) &= \sup_{t \geq 0}(trx_1 + t(1-r)x_2 - \theta(t)) \\
&= \sup_{t \geq 0}(rtx_1 + (1-r)tx_2 - r\theta(t) - \theta(t) + r\theta(t)) \\
&= \sup_{t \geq 0}(r(tx_1 - \theta(t)) + (1-r)(tx_2 - \theta(t))) \\
&\leq \sup_{t \geq 0}(r(tx_1 - \theta(t))) + \sup_{t \geq 0}((1-r)(tx_2 - \theta(t))) \\
&= r\sup_{t \geq 0}(tx_1 - \theta(t)) + (1-r)\sup_{t \geq 0}(tx_2 - \theta(t)) \\
&= r\theta^*(x_1) + (1-r)\theta^*(x_2)
\end{aligned}$$

which proves the convexity of θ^*. For $x_1 \leq x_2$ in $\mathbb{R}_+$, we have that

$$tx_1 - \theta(t) \leq tx_2 - \theta(t)$$

for each $t \geq 0$ so $\theta^*(x_1) \leq \theta^*(x_2)$. Finally, $\theta^*(0) = \sup_{t \geq 0}(-\theta(t)) = -\theta(0) = 0$ and for each $t_0 > 0, x > 0$,

$$\frac{\theta^*(x)}{x} = \frac{\sup_{t \geq 0}(tx - \theta(t))}{x} \geq \frac{t_0 x - \theta(t_0)}{x} = t_0 - \frac{\theta(t_0)}{x}$$

so

$$\lim_{x \to +\infty} \frac{\theta^*(x)}{x} \geq t_0 - \frac{\theta(t_0)}{x} = t_0$$

for each t_0 which proves that $\lim_{x \to +\infty} \frac{\theta^*(x)}{x} = +\infty$. That concludes the proof that θ^* is a Young function. ☐

Following work from Ouerdiane *et al.* [GHOR00] we can now define the spaces $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$, that are spaces of smooth maps bounded by a Young function.

Definition 3.3.5 ([GHOR00])**.** *Consider a Young function* θ *and a Banach space* B*. Let* $Exp(B, \theta, m)$ *denote the Banach space of holomorphic functions from* B *to* $\mathbb{C}$ *such that there exists* $K \geq 0$ *such that:*

$$|f(z)| \leq Ke^{\theta(m||z||)}.$$

The space $Exp(\theta, m, p)$ *is Banach when endowed with the norm :*

$$||\text{-}||_{\theta,m} : f \mapsto \sup\{|f(z)|e^{-\theta(m||z||)} \mid z \in B\}.$$

This defines two types of functions with exponential growth, depending on whether they take their arguments on a **NuclF** lcs N or on a **NuclDF** lcs N'.

$$\mathscr{F}_\theta(N') \triangleq \bigcap_{m,p} \operatorname{Exp}((N_p)', \theta, m) \qquad \mathscr{G}_\theta(N) \triangleq \bigcup_{m,p} \operatorname{Exp}(N_p, \theta, m).$$

Through an isomorphism with spaces of formal power series, one can show that $\mathscr{F}_\theta(N)$ is a nuclear Fréchet space [GHOR00, Prop 2]. As such, its dual $\mathscr{F}'_\theta(N)$, i.e. the space of distributions acting on $\mathscr{F}_\theta(N)$, is a nuclear DF space. As linear morphisms are bounded, $\mathscr{G}_\theta$: **NuclF** $\to$ **NuclDF** and $\mathscr{F}_\theta$: **NuclDF** $\to$ **NuclF** are indeed functors. We have an isomorphism between those spaces using the Laplace transform and the convex conjugate.

Theorem 3.3.6. *[GHOR00, Thm 1] Suppose that $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ is a Young function. For the conjugate Young function θ^*, we have that the Laplace transform results in an isomorphism:*

$$\mathscr{L}^+ : \begin{cases} \mathscr{F}'_\theta(N') & \simeq & \mathscr{G}_{\theta^*}(N) \\ \phi & \mapsto & \big(\ell \in N' \mapsto \phi(x \in N \mapsto e^{\ell(x)} \in \mathbb{C})\big) \end{cases} \tag{3.5}$$

As usual in lcs, the contraction rule is interpreted by the scalar multiplication of functions. It has the right type at the level of indices.

Proposition 3.3.7. *The scalar multiplication of two functions $f_1 \in \mathscr{F}_{\theta_1}(F')$ and $f_2 \in \mathscr{F}_{\theta_2}(F')$ belongs to $\mathscr{F}_{\theta_1+\theta_2}(F')$.*

Proof. Consider $p, m \in \mathbb{N}$. Let us show that there is $K \in \mathbb{R}$ such that for all $x \in F'$, $|f_1(x)f_2(x)| \leq Ke^{\theta_1(m\|z\|_p)}e^{\theta_2(m\|z\|_p)}$. This is immediate, as there are K_i such that $|f_i(x)| \leq K_i e^{\theta_i(m\|z\|_p)}$ for both indices so $K = K_1K_2$ gives the correct inequality. $\square$

Taking the dual equation 3.5, we get:

$$\mathscr{L}^- : \mathscr{G}'_{\theta^*}(N) \simeq \mathscr{F}_\theta(N') \tag{3.6}$$

Applying it to Proposition 3.3.7, we get the interpretation of the cocontraction $\bar{\mathsf{c}}$:

Proposition 3.3.8. *The convolution of two distributions $\phi_1 \in \mathscr{G}'_{\theta_1}(F)$ and $\phi_2 \in \mathscr{G}'_{\theta_2}(F)$ belongs to $\mathscr{G}'_{(\theta_1^*+\theta_2^*)^*}(F')$.*

In their paper [CEOO02, Lemma 1], Ouerdiane *et al.* also describes a convolution operator of type $\mathscr{F}'_\theta(F') \otimes \mathscr{F}'_\theta(F') \to \mathscr{F}'_\theta(F')$. That is, on $\mathscr{F}'$, the convolution does not change the indices. Applying the Laplace isomorphism (equation 3.5), we get a scalar multiplication of type $\mathscr{G}_\theta(F) \otimes \mathscr{G}_\theta(F) \to \mathscr{G}_\theta(F)$. From this, we get a slogan: *grading with θ acts only Fréchet spaces, leaving DF-spaces unchanged.* As Fréchet spaces interpret negative connectors [BKM23], while DF spaces interpret positive connectors, this justifies the choice of upper indices $+$ $-$ chosen for the indices of the exponentials in DIDiLL. Our slogan then becomes: *grading acts on negative exponential connectors, leaving positive exponential connectors unchanged.*

We also recall what will be the interpretation of promotion, following a proposition proved when studying the codigging [KL23, Prop V.8], while considering the composition of linear functions $f : \mathscr{F}'_{\theta_1}(N_1) \to N_2$ and $g : \mathscr{F}'_{\theta_2}(N_2) \to N_3$

Proposition 3.3.9. *For any Young functions θ_1, θ_2 we have a natural transformation in the category* **NuclDF***:*

$$\mu_P : \mathscr{F}'_{(\theta_2 e^{\theta_1})}(P) \to \left(\mathscr{F}'_{\theta_2}(\mathscr{F}'_{\theta_1}(P))\right).$$

This gives us the multiplication rule for a semiring of Young functions that would index DIDiLL. Work by Ouerdiane *et al.* [CEOO02] [KL23] uses a codigging acting on distributions of type $\mathscr{F}'_\theta$. To interpret DIDiLL (see Figure 3.2c), we, however, need a codigging acting on $\mathscr{G}'_\theta$. From the exact same proof as in the literature [CEOO02, Theorem 1], using the isomorphism $\mathscr{L}^-$, one gets a convolutional exponential interpreting the codigging of type:

$$\bar{\mu}_P : \left(\mathscr{G}'_{\theta_2^*}(\mathscr{G}'_{\theta_1^*}(P))\right) \to \mathscr{G}'_{\left(\theta_2 e^{\theta_1}\right)^*}(P'). \tag{3.7}$$

3.3.2 A polarized DIDiLL

This section introduces DIDiLL_P, a polarized refinement of DIDiLL. This refinement is needed to use the spaces $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$. As expressed in the previous section, the results that we have on $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$ are strongly based on the spaces they act on. We need to know if the spaces N or N' are in **NuclF** or in **NuclDF** to get results on $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$. This is the reason for this polarized refinement. We restrict the spaces where all exponential connectives can be applied to get the correct typing with the mandatory properties.

We define the logic DIDiLL_P in Figure 3.6, with its formulas, which are then polarized in Figure 3.6a, its refined syntactical transformations (the negation and the Laplace transform) in Figure 3.6b and its refined rule with the polarized formulas in Figure 3.6c. We now briefly compare DIDiLL_P with Polarized Linear Logic [Lau02] that we have presented in Section 1.1.3. It refines Linear Logic by specializing the action of the connectors to so-called positive or negative formulas. Its grammar is as follows:

$$P, Q \triangleq 1 \mid 0 \mid P \otimes Q \mid P \oplus Q \mid !N$$
$$N, M \triangleq \top \mid \bot \mid N \parr M \mid N \,\&\, M \mid ?P.$$

The syntactical use of polarization is in proof search: negative connectors are those whose introduction rules are reversible, while the introduction rule of positive connectors are not reversible. But the polarity also has a semantical meaning, notably in game semantics, where two players face each other to define a strategy, interpreting a proof. In a smooth model of DiLL, which is detailed in Section 1.3.5, this notion comes at plays as the involutive duality interpreting the negation acts between two different kinds of vector spaces. Polarization also finds its way back in DIDiLL as we split exponentials, having positive and negative versions of both ! and ?.

Some versions of polarized linear logic feature shift connectors $\uparrow$ and $\downarrow$. These are connectors which change polarity in a covariant way, used for focusing, and that are traditionally included in the grammar of polarized proof systems when the exponential connector does not change the polarity of the formulas [MT10].

$$P, Q \triangleq 1 \mid 0 \mid P \otimes Q \mid P \oplus Q \mid !P \mid \downarrow N$$
$$N, M \triangleq \top \mid \bot \mid N \parr M \mid N \,\&\, M \mid ?N \mid \uparrow P$$

Consider, for example, an exponential connector ! mapping negative formulas to negative formulas. Then to interpret minimal logic into the polarized grammar thanks to the usual call-by-name translation, one needs a shift, as the formula $!N \multimap M \equiv (!N)^\perp \parr M$ is not well constructed if $\parr$ applies only to negative connectors. The arrow is then translated as $N \Rightarrow M = \uparrow(!N)^\perp \parr M$.

In DIDiLL, exponential connectors do not change the polarity of formulas, even if we have four of them. The key is that the role usually played by shift connectors $\uparrow$ and $\downarrow$ is no longer necessary, thanks to the class of unpolarized formulas denoted A, B in Figure 3.2.

$$P, Q := 1 \mid 0 \mid P \otimes Q \mid P \oplus Q \mid ?_{x^+} P \mid !_{x^+} P$$
$$N, M := \top \mid \bot \mid N \parr M \mid N \,\&\, M \mid ?_{x^-} N \mid !_{x^-} N$$
$$A.B := N \mid P \mid A \otimes B \mid A \oplus B \mid A \parr B \mid A \,\&\, B$$

(a) Formulas of DIDiLL_P

$$(?_{x^+} P)^\perp = !_{x^-} P^\perp \qquad (!_{x^+} P)^\perp = ?_{x^-} P^\perp$$
$$(?_{x^-} N)^\perp = !_{x^+} N^\perp \qquad (!_{x^-} N)^\perp = ?_{x^+} N^\perp$$
$$\mathscr{L}(?_{x^+} P) = !_{x^+} P \qquad \mathscr{L}(!_{x^+} P) = ?_{x^+} P$$
$$\mathscr{L}(?_{x^-} N) = !_{x^-} N \qquad \mathscr{L}(!_{x^-} N) = ?_{x^-} N$$

(b) Syntactical transformations of DIDiLL_P

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?_{0^-} N}\ \mathsf{w} \qquad \frac{}{\vdash !_{0^-} N}\ \bar{\mathsf{w}}$$

$$\frac{\vdash \Gamma, ?_{x^-} N, ?_{y^-} N}{\vdash \Gamma, ?_{x^- + y^-} N}\ \mathsf{c} \qquad \frac{\vdash \Gamma, !_{x^-} N \qquad \vdash \Delta, !_{y^-} N}{\vdash \Gamma, \Delta, !_{x^- + y^-} N}\ \bar{\mathsf{c}}$$

$$\frac{\vdash \Gamma, N}{\vdash \Gamma, ?_{1^-} N}\ \mathsf{d} \qquad \frac{\vdash \Gamma, N}{\vdash \Gamma, !_{1^-} N}\ \bar{\mathsf{d}}$$

$$\frac{\vdash \Gamma, ?_{x^-} N \qquad x \leq y}{\vdash \Gamma, ?_{y^-} N}\ \mathsf{d}_I \qquad \frac{\vdash \Gamma, !_{x^-} N \qquad x \leq y}{\vdash \Gamma, !_{y^-} N}\ \bar{\mathsf{d}}_I$$

$$\frac{\vdash ?_{X^-} \mathcal{N}, P}{\vdash ?_{y^- \times X^-} \mathcal{N}, !_{y^+} P}\ \mathsf{p} \qquad \frac{\vdash !_{x^-} N, P}{\vdash !_{y^- \times x^-} N, ?_{y^+} P}\ \bar{\mathsf{p}}$$

(c) Exponential rules of DIDiLL_P

Figure 3.6: The logic DIDiLL_P

$$\begin{array}{ccccc} & [\![?_{\theta^-} N]\!] = \mathscr{F}_\theta([\![N]\!]') & \xleftarrow[\text{Fréchet Nuclear spaces}]{\mathscr{L}^-} & [\![!_{\theta^-} N]\!] = \mathscr{G}'_{\theta^*}([\![N]\!]) & \\ \mathsf{w}, \mathsf{c}, \mathsf{d}, \mu & (\_)^\perp \updownarrow & & (\_)^\perp \updownarrow & \bar{\mathsf{w}}, \bar{\mathsf{c}}, \bar{\mathsf{d}}, \bar{\mu} \\ & [\![!_{\theta^+} P]\!] = \mathscr{F}'_\theta([\![P]\!]) & \xrightarrow[\text{DF Nuclear spaces}]{\mathscr{L}^+} & [\![?_{\theta^+} P]\!] = \mathscr{G}_{\theta^*}([\![P]\!]') & \end{array}$$

Figure 3.7: Laplace and Duality acting on exponential graded by exponential growth

For example, compared with the previous translation, the call-by-name translation of the arrow is simply $N \Rightarrow M = (!_{N^-})^\perp \parr M$, which is well-defined (omitting indices).

Remark 3.3.10. Notice that shifts do not have a natural interpretation in our standard model of Fréchet / DF-spaces, interpreting respectfully negative and positive connectors. There is no natural covariant way to transform a Fréchet space into a DF-space and, conversely, without drastically changing their topology. However, the constraint on the interpretation of MALL connectors is also much softer in topological vector spaces than in polarized linear logic. Indeed, taking the example of the tensor product, while the projective tensor product of DF-spaces is indeed a DF-space, the projective tensor product of a Fréchet space and a DF-space can also be constructed, making it a simple topological vector space. Hence, one can construct a category of formulas A, B on which every MALL operation is defined. Then, compared with the previous translation, the call-by-name translations from minimal logic to DIDiLL_P is simply $N \Rightarrow M = (!_- N)^\perp \parr M$, which is well-defined (omitting indices). This, however, has its limits, as one cannot translate fully LL into DIDiLL_P (consider indeed the impossible translation of $!(!1 \multimap !1)$, as $!1 \multimap !1$ would need to be strictly positive or negative for an exponential to be applicable).

3.3.3 A model for DIDiLL

In this section, we build a denotational model for the polarized version of DIDiLL exposed in the previous section. This model is based on the spaces $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$. It uses the results we have on those spaces, namely the fact that we have isomorphisms between the spaces of functions and of distributions using the Laplace transform. These isomorphisms act on the θ functions, and hints for using the convex conjugate to relate two copies of a semiring of Young functions. We sum up how the interpretation of the exponential formulas are related in this model in Figure 3.7. In order to define this model, we seek a weak semiring of Young functions to intepret the indices, then we give the interpretation of the exponential formulas and of the exponential rules that will follow the usual interpretations of those rules in models of DiLL but taking into account the Young functions.

Definition 3.3.11. *Let us define the set*

$$\Theta \triangleq \{\theta : \mathbb{R}_+ \to \mathbb{R}_+ \mid \theta \text{ is a Young function}\}$$

i.e. θ is continuous, convex, growing and such that $\theta(0) = 0$ and $\lim_{x \to +\infty} \frac{\theta(x)}{x} = +\infty$. We define the addition $\boxplus$ on Θ as the pointwise addition of functions and the multiplication as $\theta_1 \boxtimes \theta_2 \triangleq x \mapsto \theta_1 \circ e^{\theta_2(x)}$

This cannot be turned into a weak semiring, because of the neutral elements. The neutral of the addition $\boxplus$ is the function cst_0, which is constant at 0. It is not a Young function, as

$$\lim_{x\to+\infty} \frac{cst_0(x)}{x} = 0.$$

For the multiplication, its neutral is the logarithm since

$$\theta \boxtimes \ln = x \mapsto \theta \circ e^{\ln(x)} = \theta(x) \qquad \text{and} \qquad \ln \boxtimes \theta = x \mapsto \ln \circ e^{\theta(x)} = \theta(x)$$

which is also not a Young function as it is not convex, not defined at 0 and

$$\lim_{x\to+\infty} \frac{\ln(x)}{x} = 0.$$

However, the operations of sum and product that we have defined fulfill the axioms of a weak semiring.

Proposition 3.3.12. *The operations* $\boxplus$ *and* $\boxtimes$ *over* Θ *fulfill the axioms of associativity and commutativity of a weak semiring.*

Proof. Addition is straightforwardly associative and commutative. The multiplication is associative

$$\theta_1 \boxtimes (\theta_2 \boxtimes \theta_3) = \theta_1 \circ e^{\theta_2 \circ e^{\theta_3(x)}} = (\theta_1 \boxtimes \theta_2) \boxtimes \theta_3$$

thanks to the associativity of the composition. The multiplication is right distributive over the sum as

$$(\theta_1 \boxplus \theta_2) \boxtimes \theta_3 = (\theta_1 + \theta_2) \circ e^{\theta_3} = \theta_1 \circ e^{\theta_3} + \theta_2 \circ e^{\theta_3} = (\theta_1 \boxtimes \theta_3) \boxplus (\theta_2 \boxtimes \theta_3).$$

□

Even we do not have a weak semiring of Young functions, we still can interpret the exponential connectives. The functions cst_0 and ln are not Young functions, but the spaces indexed by those functions still exists. The issue is that the Laplace transform is not an isomorphism anymore when it is indexed by the identities (see [GHOR00, Lemma 3]), hence preventing us of interpreting the (co)weakening or the (co)dereliction. Following the results of Section 3.3.1 we define the intepretation of the exponential connectives with spaces $\mathscr{F}_\theta$ and $\mathscr{G}_\theta$.

Definition 3.3.13. *We interpret the four exponential connectors as follows. The interpretation is summarized in Figure 3.7, where the set of indices is* Θ*, of the functions* cst_0 *and* ln.

- $?_{\theta^+}P$*, where* P *is a positive formula, is a positive formula interpreted by* $\mathscr{G}_{\theta^*}([\![P]\!]')$.
- $!_{\theta^-}N$*, where* N *is a negative formula, is the dual of* $?_{\theta^+}(N^\perp)$ *and a negative formula interpreted by* $\mathscr{G}'_{\theta^*}([\![N]\!])$.
- $?_{\theta^-}N$*, where* N *is a negative formula, is a negative formula interpreted by* $\mathscr{F}_\theta([\![N]\!]')$.
- $!_{\theta^+}P$*, where* P *is a positive formula, is the dual of* $?_{\theta^-}(P^\perp)$ *and a positive formula interpreted by* $\mathscr{F}'_\theta([\![P]\!])$.

Colluding syntax and semantics, we can see how the Laplace transform and the duality act on these connectors

$$\mathscr{L}([\![!_{\theta^+}P]\!]) \simeq [\![?_{\theta^+}P]\!] \qquad \mathscr{L}([\![!_{\theta^-}N]\!]) \simeq [\![?_{\theta^-}N]\!]$$

where θ is a Young function (and not the function cst_0 or ln). The motivations behind the choices of the spaces in Definition 3.3.13 are to get these isomorphisms, which is why we have used the function θ^* for some interpretation. These isomorphisms are then exactly interpreted by the Laplace transforms $\mathscr{L}^+$ and $\mathscr{L}^-$, respectively given in equation 3.5 and equation 3.6. Let us show that the interpretation of the exponential rules in smooth models of DiLL (see Section 1.3.5) adapts well to this graded setting.

- The weakening rule w on a function corresponds to the introduction of the constant function on $[\![P]\!]'$, which is indeed an element of $\mathscr{G}_0([\![P]\!])$.
- The coweakening $\bar{\mathsf{w}}$ is interpreted as the introduction of δ_0, which acts linearly and continuously on every $\mathscr{F}_0(P)$.
- Contraction c and cocontraction $\bar{\mathsf{c}}$ are interpreted by scalar multiplication and convolution, as detailed in Proposition 3.3.8 and Proposition 3.3.7.
- The indexed dereliction d_I has an immediate interpretation and no computational content, as for $\theta' \leq \theta$ in the semi-ring, that is $\theta \leq \theta'$ pointwise, we have $(\theta')^* \leq \theta^*$ pointwise, and thus $[\![?_{\theta'^+} P]\!] = \mathscr{G}_{(\theta')^*}([\![P]\!]) \subseteq [\![?_{\theta^+} P]\!] = \mathscr{G}_{\theta^*}([\![P]\!])$. The indexed codereliction $\bar{\mathsf{d}}_E$ works likewise by inclusion.
- For the promotion, consider the interpretation $\ell \in \mathcal{L}(\mathscr{F}'_{\theta_1}[\![N]\!]', [\![Q]\!])$ of the sequent $\vdash ?_{x^-} N, Q$. Then by the functorality of $\mathscr{F}'_{\theta_2}$, and precomposing by $\mu_{[\![N]\!]'} : \mathscr{F}'_{(\theta_2 e^{\theta_1})}([\![N]\!]') \to \left(\mathscr{F}'_{\theta_2}(\mathscr{F}'_{\theta_1}([\![N]\!]'))\right)$ described in Proposition 3.3.9, one get a linear map in $\mathcal{L}(\mathscr{F}'_{(\theta_2 e^{\theta_1})}([\![N]\!]'), \mathscr{F}'_{\theta_2}([\![N]\!]'))$ interpreting the sequent $\vdash ?_{y\times x^-} N, !_{y^+} Q$.
- The interpretation of the copromotion rules work likewise, thanks to the natural transformation $\bar{\mu}$ described in equation 3.7.
- Consider a Nuclear Fréchet space $N = \lim N_p$, then $\mathsf{d} : N \to ?_{1^-} N$ maps $\ell \in (N')'$ to the same map in $\mathscr{G}_1(N')$, and there is indeed a K such that ℓ is bounded by $Ke^{\ln(||\_||_p)} = K||\_||_p$ on every N_p.
- Likewise, the map $\bar{\mathsf{d}} : N \to !_{1^-} N$ mapping a vector v to $f \mapsto D_0(f)(v)$ is well typed, as $D_0(\_)(v)$ acts continuously on $\mathscr{F}'_{\ln}(N)$.

When the rules introducing identities to the indices ((co)weakening and non indexed (co)dereliction) are removed, the cut-elimination rules of DIDiLL are preserved, as they follow the usual interpretation of calculus in smooth models of DiLL.

Chapter 4

Graded differential linear logic

Contents

This chapter takes an orthogonal point of view on what we presented in Chapter 3. To extend our indexed differential linear logic to higher order, we have refined the indices with some kind of polarization inspired by semantic considerations using the Laplace transform. Here, we try to achieve the same goal but using syntactical conditions. Concretely, we add a promotion rule to IDiLL, which enables us to construct objects at higher order and see what other conditions our formal system has to respect to get a *working* logic. Building such a system comes with new cases in the cut elimination procedure. Those cases induce several algebraic conditions on the indices. The cut elimination corresponds to an equality between formulas, which comes with an equality on the indices. Studying those cases, the algebraic structure of the indices has to fulfill several new axioms such as the discretness of the semiring, or a particular decomposition over the elements. In Section 4.1, we define the logic $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$. It is based on the notion of *differential semiring*, which represents the algebraic conditions needed on the indices. We then give the cut elimination cases of $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$ in Section 4.2. We do not prove any theorem on a potential cut elimination procedure, but we shortly explain which kind of property should be true for such a procedure. Having graded rules may change a lot the dynamics compared to non-graded differential linear logic: the shape of some cut elimination cases has to be redefined. Especially while DiLL strongly uses the notion of sums of proofs, in $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$ the sums of proofs are not possible. In Section 4.3 we explain the link between the notion of differential semiring and the cut elimination cases. We then give some examples and counter-examples of differential semirings.

4.1 The syntax of Graded Differential Linear Logic

We define here graded differential linear logic, which is an extension of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ to higher order. Being able to perform higher-order constructions means that we add the promotion rule to this logic. Adding a promotion rule alone in a graded setting does not really make sense. At the level of the indices, having a graded promotion means that we perform a multiplication (see Section 1.2). Hence, it should come with the rule introducing the unit element of this multiplication, which is the dereliction rule. Since we are in a differential framework, we also add a graded codereliction rule. We have then a logic where the exponential connectives are indexed by elements not just of a monoid but also of a semiring. The grammar of the formulas of our graded differential linear logic ($\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$) is the same as the one of $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, and the exponential rules are given in Figure 4.1

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?_0 A}\ \mathsf{w} \qquad \frac{\vdash \Gamma, ?_x A, ?_y A}{\vdash \Gamma, ?_{x+y} A}\ \mathsf{c} \qquad \frac{\vdash \Gamma, A}{\vdash \Gamma, ?_1 A}\ \mathsf{d}$$

$$\frac{}{\vdash !_0 A}\ \bar{\mathsf{w}} \qquad \frac{\vdash \Gamma, !_x A \qquad \vdash \Delta, !_y A}{\vdash \Gamma, \Delta, !_{x+y} A}\ \bar{\mathsf{c}} \qquad \frac{\vdash \Gamma, A}{\vdash \Gamma, !_1 A}\ \bar{\mathsf{d}}$$

$$\frac{\vdash \Gamma, ?_x A \qquad x \leq y}{\vdash \Gamma, ?_y A}\ \mathsf{d}_I \qquad \frac{\vdash \Gamma, !_x A \qquad x \leq y}{\vdash \Gamma, !_y A}\ \bar{\mathsf{d}}_I \qquad \frac{\vdash ?_X \Gamma, A}{\vdash ?_{y \times X} \Gamma, !_y A}\ \mathsf{p}$$

Figure 4.1: Exponential rules of $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$

Remark 4.1.1. The extension to the higher order that we define here uses the promotion rule. It may also be extended with a copromotion, as we did for DIDiLL in Chapter 3 but this is not something we have studied so far. Relaxing polarization implies to already consider many cases, some of them being very different with their version in differential linear logic. Moreover, the copromotion has not been studied a lot yet. The codigging has recently been introduced by Kerjean and Lemay [KL23] as a categorical operation, and we have defined the copromotion in the previous chapter but making some choices on its form. We hope that the system presented in this chapter can be extended to include a copromotion rule.

These exponential rules are based on a set of indices, having a particular algebraic structure. This structure is usually a semiring for graded logics, but as we saw with our notion of weak semiring (see Definition 3.1.1) this may need to be refined. For DIDiLL, we have introduced weak semirings for two reasons: it was the set of algebraic axioms corresponding exactly to those needed to perform the cut elimination cases, and it was weak enough to have the set of LPDOcc fulfill those axioms. Graded linear logic uses semirings as it has the axioms needed for cut elimination and also some other that correspond to commutation rules useful in proof nets. For $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$ we have to define a new, relaxed notion of semiring because of the new interactions coming from the cut elimination cases. We chose to give a minimal definition: our relaxed definition has axioms only based on the cut elimination cases. It may allow us to grade our exponential connectives with more interesting semirings, and as for DIDiLL we prefer this possibility compared to having a system closed to what is done in proof nets. Let us now give our relaxed definition of semiring, called differential semiring.

Definition 4.1.2. *A* differential semiring $\mathcal{S}$ *is the data of two operations* $+$ *and* $\times$*,two elements* $0, 1 \in \mathcal{S}$ *and a partial order* $\leq$ *such that*

1. $(\mathcal{S}, +, 0)$ *is a commutative monoid, which means that* $+$ *is associative and commutative and 0 is its neutral elements: for each* $x \in \mathcal{S}$, $x + 0 = x = 0 + x$*;*
2. $\mathcal{S}$ *is additive splitting, meaning that for each* $x_1, x_2, x_3, x_4 \in \mathcal{S}$ *such that* $x_1 + x_2 = x_3 + x_4$*, there exist* $x_{1,3}, x_{1,4}, x_{2,3}, x_{2,4} \in \mathcal{S}$ *such that*
$$x_1 = x_{1,3} + x_{1,4} \qquad x_2 = x_{2,3} + x_{2,4} \qquad x_3 = x_{1,3} + x_{2,3} \qquad x_4 = x_{1,4} + x_{2,4};$$
3. *0 is left absorbing for the product, that is for each* $x \in \mathcal{S}$, $0 \times x = 0$*;*
4. *1 is a left neutral for the product, meaning that for each* $x \in \mathcal{S}$, $1 \times x = x$*;*
5. *the product is distributive over the sum, which means that for each* $x, y, z \in \mathcal{S}$,
$$(x + y)z = xz + yz \qquad x(y + z) = xy + xz;$$
6. *the product is associative, meaning that for each* $x, y, z \in \mathcal{S}$,
$$(xy)z = x(yz);$$
7. $\mathcal{S}$ *is naturally ordered, which means that the order is defined by*
$$x \leq y \qquad \Leftrightarrow \qquad \exists z \in \mathcal{S},\ x + z = y;$$
8. $\mathcal{S}$ *is discrete, that is if* $x, y \in \mathcal{S}$ *such that* $x + y = 1$ *then* $x = 0$ *or* $y = 0$*;*
9. $\mathcal{S}$ *is integral domain, meaning that is* $x, y \in \mathcal{S}$ *such that* $xy = 0$ *then* $x = 0$ *or* $y = 0$*;*
10. $\mathcal{S}$ *is multiplicative splitting, which implies that for* $x, y, r, z \in \mathcal{S}$ *such that* $x + y = rz$*, there exist elements* $r_1, \ldots, r_k$ *and* $z_1, \ldots, z_l$ *in* $\mathcal{S}$ *and a set* $U \subseteq \{1, \ldots, k\} \times \{1, \ldots, l\}$ *such that*
$$r = \sum_{i=1}^{k} r_i \qquad z = \sum_{j=1}^{l} z_j \qquad x = \sum_{(i,j) \in U} r_i z_j \qquad y = \sum_{(i,j) \notin U} r_i z_j;$$
11. $\mathcal{S}$ *satisfies a non-unity property, which states that if* $xy = 1$ *then* $x = 1 = y$*.*

Compared to the usual notion of semiring, some conditions are weakened. For example, we only ask for distributivity on one side, as well as for the neutrality or the absorbance. But some conditions are also stronger than what is required in a usual semiring. The condition of discretness or the multiplicative splitting for example are new axioms.

4.2 The cut elimination cases of graded differential linear logic

In $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$, there are a lot possible cut elimination cases. Some of them have already been studied in previous chapters. Since $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$ is an extension of the logic $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ defined in Chapter 2, some techniques used for the cut elimination of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$ will be reused here. In particular, we introduce again the rules w_I and $\bar{\mathsf{w}}_I$ that were introduced in Section 2.3.2. These rules are defined by

$$\frac{\vdash \Gamma}{\vdash \Gamma, ?_x A}\,\mathsf{w}_I \quad \triangleq \quad \dfrac{\dfrac{\vdash \Gamma}{\vdash \Gamma, ?_0 A}\,\mathsf{w}}{\vdash \Gamma, ?_x A}\,\mathsf{d}_I \qquad\qquad \frac{}{\vdash !_x A}\,\bar{\mathsf{w}}_I \quad \triangleq \quad \dfrac{\dfrac{}{\vdash !_0 A}\,\bar{\mathsf{w}}}{\vdash !_x A}\,\bar{\mathsf{d}}_I$$

We recall that these rules are equivalent to d_I and $\bar{\mathsf{d}}_I$ when the semiring order is defined through the sum. It is the case here and it makes the cut elimination cases easier to define. Let us consider now which of these cases are new and which ones have been studied before.

1. The interactions between the structural rules and the promotion have been given by Breuvart and Pagani (see [BP15] or Figure 1.6).

2. The interactions between the structural monoidal rules ($\mathsf{w}, \mathsf{w}_I, \mathsf{c}$) and the costructural monoidal rules ($\bar{\mathsf{w}}, \bar{\mathsf{w}}_I, \bar{\mathsf{c}}$) are given in our definition of the cut elimination procedure of IDiLL in Chapter 2 (see Figures 2.4 and 2.5).

This lets us four new kinds of cases to consider:

3. the interactions between the dereliction and the costructural rules;

4. the interactions between the codereliction and the structural rules;

5. the interaction between the promotion and the indexed weakening;

6. the interactions between the promotion and the costructural rules.

4.2.1 The interactions between the dereliction and the costructural rules

These interactions have been studied in DiLL in their non-graded version. We extend them to the graded framework, giving a rewriting of the same shape and decorating it at the level of the indices, introducing algebraic axioms.

- The cut between d and $\bar{\mathsf{w}}$ cannot exist: the dereliction introduces formulas $?_1 A$ and the coweakening formulas $!_0 A$ which cannot be dual to each other[1]

- The cut between d and $\bar{\mathsf{w}}_I$ can exist, as w_I introduces a formula $!_x A$ where x can be any element of the semiring, including 1. For this case, the reduction is the same as the one of DiLL, which uses the 0 (or empty) proof

$$\dfrac{\dfrac{\vdash \Gamma, A^\perp}{\vdash \Gamma, ?_1 A^\perp}\,\mathsf{d} \qquad \dfrac{}{\vdash !_1 A}\,\bar{\mathsf{w}}_I}{\vdash \Gamma}\,\mathit{cut} \qquad \rightsquigarrow \qquad \dfrac{}{\vdash \Gamma}\,0$$

- The cut between d and $\bar{\mathsf{c}}$ is not similar to the one in DiLL. In differential linear logic, the dereliction introduces a formula $?A^\perp$ from $A^\perp$ and the cocontraction contracts two occurrences of $!A$ into one. To eliminate this cut in DiLL, one cuts $?A^\perp$ with the occurrences of $!A$, but since there are two occurrences, it gives two parts of the proof tree that are combined using the sum of proofs (see Figure 1.9). Here, the graded flavor brings a major difference: the sum of proofs is not needed anymore. A dereliction introduces $?_1 A^\perp$ while a cocontraction sums the indices of the exponential formulas. This means that this cut corresponds to the situation

[1] They can, in fact, be dual to each other, but this would mean that $0 = 1$ in the semiring we consider which implies that every elements of the semiring are equal, and then the logic collapses to differential linear logic.

$$\dfrac{\dfrac{\vdash \Gamma, A^{\perp}}{\vdash \Gamma, ?_1 A^{\perp}}\,\mathsf{d} \qquad \dfrac{\vdash \Delta, !_x A \qquad \vdash \Xi, !_y A}{\vdash \Delta, \Xi, !_{x+y=1} A}\,\bar{\mathsf{c}}}{\vdash \Gamma, \Delta, \Xi}\,\mathit{cut}$$

where either $x \neq 1$ or $y \neq 1$. In this situation, $?_1 A^{\perp}$ cannot be cut with both $!_x A$ and $!_y A$ and this is why the sum of proofs is not needed anymore. However, to be able to eliminate this cut, we need a new axiom on the semiring. We need to assume that $x + y = 1$ implies that $x = 1$ and $y = 0$ or $x = 0$ and $y = 1$. We say that the semiring is *discrete*. Using this axiom, the reduced proof is

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdash \Gamma, A^{\perp}}{\vdash \Gamma, ?_1 A^{\perp}}\,\mathsf{d} \qquad \vdash \Delta, !_{x=1} A}{\vdash \Gamma, \Delta}\,\mathit{cut}}{\vdash \Gamma, \Delta, ?_0 A^{\perp}}\,\mathsf{w} \qquad \vdash \Xi, !_{y=0} A}{\vdash \Gamma, \Delta, \Xi}\,\mathit{cut}$$

or

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdash \Gamma, A^{\perp}}{\vdash \Gamma, ?_1 A^{\perp}}\,\mathsf{d} \qquad \vdash \Xi, !_{y=1} A}{\vdash \Gamma, \Xi}\,\mathit{cut}}{\vdash \Gamma, \Xi, ?_0 A^{\perp}}\,\mathsf{w} \qquad \vdash \Delta, !_{x=0} A}{\vdash \Gamma, \Delta, \Xi}\,\mathit{cut}$$

depending on the value of x and y.

- The cut between d and $\bar{\mathsf{d}}$ is the same as its non-graded version

$$\dfrac{\dfrac{\vdash \Gamma, A}{\vdash \Gamma, ?_1 A}\,\mathsf{d} \qquad \dfrac{\vdash \Delta, A^{\perp}}{\vdash \Delta, !_1 A^{\perp}}\,\bar{\mathsf{d}}}{\vdash \Gamma, \Delta}\,\mathit{cut} \quad \rightsquigarrow \quad \dfrac{\vdash \Gamma, A \qquad \vdash \Delta, A^{\perp}}{\vdash \Gamma, \Delta}\,\mathit{cut}$$

and does not need any algebraic axiom.

4.2.2 The interactions between the codereliction and the structural rules

They are similar to the cases between the codereliction and the structural rules. They use the same axiom, the axiom of discreteness, to perform the cut elimination case between the codereliction and the contraction.

- The cut between $\bar{\mathsf{d}}$ and w cannot exist for the same reasons as the one between d and $\bar{\mathsf{w}}$.

- The cut between $\bar{\mathsf{d}}$ and w_I is here again reduced using the empty proof, similar to the cut $\bar{\mathsf{d}}/\mathsf{w}_I$ in DiLL

$$\dfrac{\dfrac{\vdash \Gamma, A}{\vdash \Gamma, !_1 A}\,\bar{\mathsf{d}} \qquad \dfrac{\vdash \Delta}{\vdash \Delta, ?_1 A^{\perp}}\,\mathsf{w}_I}{\vdash \Gamma, \Delta}\,\mathit{cut} \quad \rightsquigarrow \quad \dfrac{}{\vdash \Gamma, \Delta}\,0$$

- The cut between $\bar{\mathsf{d}}$ and c changes from DiLL, as it removes the use of the sum of proofs for similar reasons to the one between d and $\bar{\mathsf{c}}$. We use the axiom of discreteness once more, as a way to branch the formula produced by the codereliction into one of the premises of the contraction. Before the rewriting, this case has the following shape:

$$\dfrac{\dfrac{\vdash \Gamma, A}{\vdash \Gamma, !_1 A}\,\bar{\mathsf{d}} \qquad \dfrac{\vdash \Delta, ?_x A^\perp, ?_y A^\perp}{\vdash \Delta, ?_{x+y=1} A^\perp}\,\mathsf{c}}{\vdash \Gamma, \Delta}\,cut \ .$$

After the rewriting it can then have two forms, depending on the value of the index of the formulas in this premise, which are

$$\dfrac{\dfrac{}{\vdash !_0 A}\,\bar{\mathsf{w}} \qquad \dfrac{\dfrac{\vdash \Gamma, A}{\vdash \Gamma, !_1 A}\,\bar{\mathsf{d}} \qquad \vdash \Delta, ?_{x=0} A^\perp, ?_{y=1} A^\perp}{\vdash \Gamma, \Delta, ?_0 A^\perp}\,cut}{\vdash \Gamma, \Delta}\,cut$$

when $x = 0$ and $y = 1$ and

$$\dfrac{\dfrac{\vdash \Gamma, A}{\vdash \Gamma, !_1 A}\,\bar{\mathsf{d}} \qquad \dfrac{\dfrac{}{\vdash !_0 A}\,\bar{\mathsf{w}} \qquad \vdash \Delta, ?_{x=1} A^\perp, ?_{y=0} A^\perp}{\vdash \Delta, ?_1 A^\perp}\,cut}{\vdash \Gamma, \Delta}\,cut$$

when $x = 0$ and $y = 1$, which are the only two cases thanks to the discreteness.

4.2.3 The interaction between the promotion and the indexed weakening

It is not exactly a new case. The indexed weakening being the application of a weakening and an indexed dereliction, this case is

$$\dfrac{\dfrac{\vdash ?_{x_1} A_1, \dots, ?_{x_n} A_n, B}{\vdash ?_{yx_1} A_1, \dots, ?_{yx_n} A_n, !_y B}\,\mathsf{p} \qquad \dfrac{\vdash \Gamma}{\vdash \Gamma, ?_y B^\perp}\,\mathsf{w}_I}{\vdash \Gamma, ?_{yx_1} A_1, \dots, ?_{yx_n} A_n}\,cut$$

which corresponds to

$$\dfrac{\dfrac{\vdash ?_{x_1} A_1, \dots, ?_{x_n} A_n, B}{\vdash ?_{yx_1} A_1, \dots, ?_{yx_n} A_n, !_y B}\,\mathsf{p} \qquad \dfrac{\dfrac{\vdash \Gamma}{\vdash \Gamma, ?_0 B^\perp}\,\mathsf{w}}{\vdash \Gamma, ?_y B^\perp}\,\mathsf{d}_I}{\vdash \Gamma, ?_{yx_1} A_1, \dots, ?_{yx_n} A_n}\,cut$$

and the interaction between p and d_I is already studied in the cut elimination procedure of graded linear logic. Applying the rewriting defined in this procedure, we get

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdash ?_{x_1} A_1, \dots, ?_{x_n} A_n, B}{\vdash ?_0 A_1, \dots, ?_0 A_n, !_0 B}\,\mathsf{p} \qquad \dfrac{\vdash \Gamma}{\vdash \Gamma, ?_0 B^\perp}\,\mathsf{w}}{\Gamma, ?_0 A_1, \dots, ?_0 A_n}\,cut}{\Gamma, ?_{yx_1} A_1, \dots, ?_0 A_n}\,\mathsf{d}_I}{\vdots}\,\mathsf{d}_I}{\vdash \Gamma, ?_{yx_1} A_1, \dots, ?_{yx_n} A_n}\,\mathsf{d}_I$$

Since we work in a setting using indexed weakenings instead of indexed derelictions, we use the equivalence between those two rules, which is

$$\dfrac{\vdash \Gamma, ?_x A}{\vdash \Gamma, ?_{x+y} A}\,\mathsf{d}_I \quad = \quad \dfrac{\dfrac{\vdash \Gamma, ?_x A}{\vdash \Gamma, ?_x A, ?_y A}\,\mathsf{w}_I}{\vdash \Gamma, ?_{x+y} A}\,\mathsf{c}$$

to rephrase the tree after the rewriting step as

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdash ?_{x_1} A_1, \ldots, ?_{x_n} A_n, B}{\vdash ?_0 A_1, \ldots, ?_0 A_n, !_0 B}\,\mathsf{p} \qquad \dfrac{\vdash \Gamma}{\vdash \Gamma, ?_0 B^\perp}\,\mathsf{w}}{\Gamma, ?_0 A_1, \ldots, ?_0 A_n}\,cut}{\Gamma, ?_0 A_1, ?_{yx_1} A_1, \ldots, ?_0 A_n}\,\mathsf{w}_I}{\vdash \Gamma ?_{yx_1} A_1, \ldots, ?_0 A_n}\,\mathsf{c}}{\vdots}\,\mathsf{w}_I}{\vdash \Gamma, ?_{yx_1} A_1, \ldots, ?_{yx_n} A_n}\,\mathsf{c}$$

4.2.4 The interactions between the promotion and the costructural rules

Those cases are new in some sense. In the non-graded setting, they have been studied for the cut elimination procedure of DiLL. However, some of these cases are quite different when the exponentials are graded. We have partly presented it in the long version of our paper defining IDiLL [BKM26].

- The cut between p and $\bar{\mathsf{w}}$ uses the integral domain property, which states that when x and y are elements of the semiring such that $xy = 0$ then either $x = 0$ or $y = 0$. This cut case is

  $$\dfrac{\dfrac{}{\vdash !_0 A}\,\bar{\mathsf{w}} \qquad \dfrac{\vdash ?_x A^\perp, ?_{x_1} B_1, \ldots, ?_{x_n} B_n, C}{\vdash ?_{yx} A^\perp, ?_{yx_1} B_1, \ldots, ?_{yx_n} B_n, !_y C}\,\mathsf{p}}{\vdash ?_{yx_1} B_1, \ldots, ?_{yx_n} B_n, !_y C}\,cut$$

  with $yx = 0$. If $y = 0$ it is reduced to

  $$\dfrac{\dfrac{\dfrac{\dfrac{}{\vdash !_0 C}\,\bar{\mathsf{w}}}{\vdash ?_0 B_1, !_0 C}\,\mathsf{w}}{\vdots}\,\mathsf{w}}{\vdash ?_0 B_1, \ldots, ?_0 B_n, !_0 C}\,\mathsf{w}$$

  which is correct since for each i, $yx_i = 0$ using that $y = 0$. This rewriting corresponds to a simple decoration of the same case in DiLL. When $x \neq 0$, the integral domain property implies that $y = 0$ and the reduction is given by

  $$\dfrac{\dfrac{\dfrac{}{\vdash !_0 A}\,\bar{\mathsf{w}} \qquad \vdash ?_0 A^\perp, ?_{x_1} B_1, \ldots, ?_{x_n} B_n, C}{\vdash ?_{x_1} B_1, \ldots, ?_{x_n} B_n, C}\,cut}{\vdash ?_{yx_1} B_1, \ldots, ?_{yx_n} B_n, !_y C}\,\mathsf{p}$$

- The cut between p and w_I is new. Moreover, since the interaction between p and d_I has not been studied either, we cannot use it. This cut before the rewriting is

$$\dfrac{\dfrac{}{\vdash !_{yx}A}\,\bar{\mathsf{w}}_I \qquad \dfrac{\vdash ?_x A^\perp, ?_{x_1}B_1,\dots,?_{x_n}B_n, C}{\vdash ?_{yx}A^\perp, ?_{yx_1}B_1,\dots,?_{yx_n}B_n, !_y C}\,\mathsf{p}}{\vdash ?_{yx_1}B_1,\dots,?_{yx_n}B_n,!_yC}\,cut$$

To deal with this cut, we rewrite the proof to get a case that we know: the cut elimination case between p and $\bar{\mathsf{w}}$. To do so, we will need to use carefully indexed (co)weakenings, but we will use them after the cut. This gives the following rewritten tree.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\vdash !_0A}\,\bar{\mathsf{w}} \qquad \dfrac{\vdash ?_x A^\perp, ?_{x_1}B_1,\dots,?_{x_n}B_n, C}{\vdash ?_0A^\perp, ?_0B_1,\dots,?_0B_n,!_0C}\,\mathsf{p}}{\dfrac{\vdash ?_0B_1,\dots,?_0B_n,!_0C}{\vdots}\,\mathsf{w}_I}\,cut}{\vdash ?_{yx_1}B_1,\dots,?_{yx_n}B_n,?_0B_1,\dots,?_0B_n,!_0C}\,\mathsf{w}_I}{\vdots}\,\mathsf{c}}{\vdash ?_{yx_1}B_1,\dots,?_{yx_n}B_n,!_0C}\,\mathsf{c} \qquad \dfrac{}{\vdash !_yC}\,\bar{\mathsf{w}}_I}{\vdash ?_{yx_1}B_1,\dots,?_{yx_n}B_n,!_yC}\,\bar{\mathsf{c}}$$

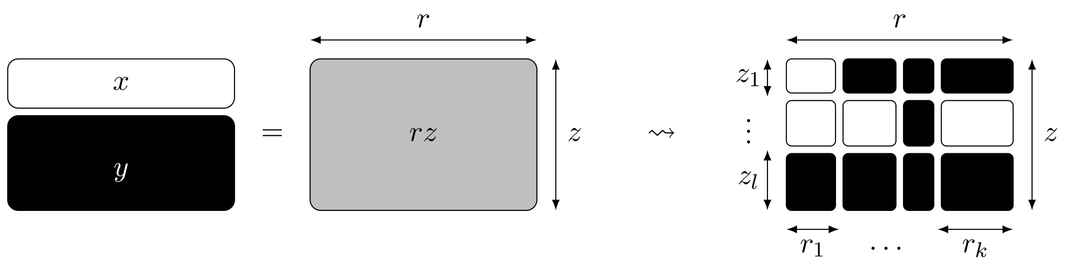
- The cut between p and $\bar{\mathsf{c}}$ has been studied in DiLL when the rules are non-graded. But being graded completely changes the shape of the rewriting. This interaction is

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdash \Gamma, !_xA\end{array} \qquad \begin{array}{c}\pi_2\\ \vdash \Delta, !_yA\end{array}}{\vdash \Gamma,\Delta,!_{x+y}A}\,\bar{\mathsf{c}} \qquad \dfrac{\begin{array}{c}\pi_3\\ \vdash ?_zA^\perp, ?_{t_1}B_1,\dots,?_{t_n}B_n, C\end{array}}{\vdash ?_{rz}A^\perp, ?_{rt_1}B_1,\dots,?_{rt_n}B_n,!_rC}\,\mathsf{p}}{\vdash \Gamma,\Delta,?_{rt_1}B_1,\dots,?_{rt_n}B_n,!_rC}\,cut \tag{4.1}$$

with $x+y = rz$. The difficulty comes from the fact that we have to relate the sum and the product of the semiring. To do so, we use the multiplicative splitting property of the semiring. It implies[2] the existence of elements $r_1,\dots,r_k$ and $z_1,\dots,z_l$ in the semiring and of a set $U \subseteq \{1,\dots,k\}\times\{1,\dots,l\}$ such that

$$r = \sum_{i=1}^{k} r_i \qquad z = \sum_{j=1}^{l} z_j \qquad\qquad x = \sum_{(i,j)\in U} r_iz_j \qquad y = \sum_{(i,j)\notin U} r_iz_j.$$

To give more intuition about this decomposition, we represent it graphically by



[2] Note that several versions of this property exist. One of them simplifies the decomposition by using only six elements of the semiring (see [CES10] or [Mun24]). Here we take a more technical version of this splitting but which is weaker. Our version is strong enough to deal with this cut elimination case and may be weak enough to be satisfied with more semirings.

where we decompose the value of r horizontally and the one of z vertically. Thanks to these decompositions, the products between the r_i and the z_j have to be assigned to either x or y (using black or white in our graphics) and by summing those elements, we recall the value of x and of y. To describe the cut elimination between p and $\bar{\mathsf{c}}$ we will use the rule $\bar{\mathsf{c}}^\perp$, dual of the cocontraction, defined by

$$\dfrac{\dfrac{\pi}{\vdash \Gamma, ?_{x+y}A}}{\vdash \Gamma, ?_xA, ?_yA}\,\bar{\mathsf{c}}^\perp \quad\triangleq\quad \dfrac{\dfrac{\pi}{\vdash \Gamma, ?_{x+y}A} \qquad \dfrac{\dfrac{}{\vdash !_xA^\perp, ?_xA}\,ax \qquad \dfrac{}{\vdash !_yA^\perp, ?_yA}\,ax}{\vdash !_{x+y}A^\perp, ?_xA, ?_yA}\,\bar{\mathsf{c}}}{\vdash \Gamma, ?_xA, ?_yA}\,cut$$

Now, from the proof tree π_3 in equation 4.1 we define, for each $1 \leq i \leq k$ the proof tree $\pi_{3,i}$ by

$$\pi_{3,i} \quad\triangleq\quad \dfrac{\dfrac{\dfrac{\pi_3}{\vdash ?_{\sum_{j=1}^{l} z_j}A^\perp, ?_{t_1}B, \ldots, ?_{t_n}B_n, C}}{\vdash ?_{\sum_{\{j|(i,j)\in U\}} z_j}A^\perp, ?_{\sum_{\{j|(i,j)\notin U\}} z_j}A^\perp, ?_{t_1}B_1, \ldots, ?_{t_n}B_n, C}\,\bar{\mathsf{c}}^\perp}{\vdash ?_{\sum_{\{j|(i,j)\in U\}} r_iz_j}A^\perp, ?_{\sum_{\{j|(i,j)\notin U\}} r_iz_j}A^\perp, ?_{r_it_1}B_1, \ldots, ?_{r_it_n}B_n, !_{r_i}C}\,\mathsf{p}$$

where we use $\bar{\mathsf{c}}^\perp$ to split the sum $\sum_{j=1}^{l} z_j = z$ into two sums: the elements z_j such that r_iz_j is in the decomposition of x and the others (which are then in the decomposition of y thanks to the multiplicative splitting). We need one last intermediate step before giving the rewriting. That is, defining a new tree π_3' that will recombine the elements of x and of y thanks to the trees $\pi_{3,i}$ and (co)contractions.

$$\pi_3' \quad\triangleq\quad \dfrac{\dfrac{\vdots}{\dfrac{\dfrac{\dfrac{\dfrac{\pi_{3,1}}{\vdash \Gamma_1, !_{r_1}C} \qquad \dfrac{\pi_{3,2}}{\vdash \Gamma_2, !_{r_2}C}}{\vdash \Gamma_1, \Gamma_2, !_{r_1+r_2}C}\,\bar{\mathsf{c}} \\ \ddots \qquad \dfrac{\pi_{3,k}}{\vdash \Gamma_k, !_{r_k}C}}{\vdash \Gamma_1, \ldots, \Gamma_k, !_rC}\,\bar{\mathsf{c}}}{}\,\mathsf{c}}}{\vdash ?_xA^\perp, ?_yA^\perp, ?_{rt_1}B_1, \ldots, ?_{rt_n}B_n, !_rC}\,\mathsf{c}$$

where we have used the notation Γ_i to represent the context of the conclusion of the trees $\pi_{3,i}$. Using, at first, the cocontractions allows us to concatenate all of these contexts, and to obtain a formula $!_{r_1+\cdots+r_k}C$ which is $!_rC$ thanks to the decomposition of r. For the contexts, we use contractions to recombine every formula. Each Γ_i has in its context the formula $?_{\sum_{\{j|(i,j\in U)\}r_iz_j}}A^\perp$ and contracting all of these formulas gives $?_xA^\perp$ since

$$\left(\sum_{(i,1)\in U} r_1z_j\right) + \cdots + \left(\sum_{(i,1)\in U} r_kz_j\right) = \sum_{(i,j)\in U} r_iz_j = x$$

and similarly contracting the formulas $?_{\sum_{\{j|(i,j\notin U)\}r_iz_j}}A^\perp$ produces $?_yA^\perp$. In addition, each Γ_i contains formulas $?_{r_it_m}B_m$ for $1 \leq m \leq n$, and contracting them produces

$?_{rt_m}B_m$ using

$$r_1t_m + \cdots + r_kt_m = (r_1 + \cdots + r_k)t_m = rt_m$$

which is allowed by the right distributivity of the semiring. This shows that our proposed form for the conclusion of π_3' is correct. Finally, we are able to give the form of the rewriting for this cut case, using π_3':

$$\dfrac{\overset{\pi_1}{\vdash \Gamma, !_xA} \qquad \dfrac{\overset{\pi_2}{\vdash \Delta, !_yA} \qquad \overset{\pi_3'}{\vdash ?_xA^\perp, ?_yA^\perp, ?_{rt_1}B_1, \ldots, ?_{rt_n}B_n, !_rC}}{\vdash ?_xA^\perp, ?_{rt_1}B_1, \ldots, ?_{rt_n}B_n, !_rC}\,\mathit{cut}}{\vdash ?_{rt_1}B_1, \ldots, ?_{rt_n}B_n, !_rC}\,\mathit{cut}$$

Note that this form of the rewriting is similar to the one in DiLL, but we had to use many additional contractions and cocontractions to recombine the indices correctly.

- For the cut between p and $\bar{\mathsf{d}}$, the shape will be the same as for DiLL and we only need to decorate in correctly. But for this decoration, we will have to use a new algebraic axiom. The tree before the rewriting is

$$\dfrac{\dfrac{\overset{\pi_1}{\vdash ?_{x_1}C_1, \ldots, ?_{x_n}C_n, ?_xB, A}}{\vdash ?_{yx_1}C_1, \ldots, ?_{yx_n}C_n, ?_{yx}B, !_yA}\,\mathsf{p} \qquad \dfrac{\overset{\pi_2}{\vdash \Gamma, B^\perp}}{\vdash \Gamma, !_1B^\perp}\,\bar{\mathsf{d}}}{\vdash \Gamma, ?_{yx_1}C_1, \ldots, ?_{yx_n}C_n, !_yA}\,\mathit{cut}$$

where $xy = 1$. We are then able to use the non-unity property to assume that $x = 1 = y$. Using this, we can use the rewriting of DiLL and decorate it.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{\pi_1}{\vdash ?_{x_1}C_1, \ldots, ?_{x_n}C_n, ?_xB, A}}{\vdash ?_0C_1, \ldots, ?_0C_n, ?_0B, !_0A}\,\mathsf{p} \qquad \dfrac{}{\vdash !_0B^\perp}\,\bar{\mathsf{w}}}{\vdash ?_0C_1, \ldots, ?_0C_n, !_0A}\,\mathit{cut} \qquad \dfrac{\dfrac{\overset{\pi_2}{\vdash \Gamma, B^\perp}}{\vdash \Gamma, !_1B^\perp}\,\bar{\mathsf{d}} \qquad \dfrac{\overset{\pi_1}{\vdash ?_{x_1}C_1, \ldots, ?_{x_n}C_n, ?_1B, A}}{\vdash ?_{x_1}C_1, \ldots, ?_{x_n}C_n, ?_1B, !_1A}\,\bar{\mathsf{d}}}{\vdash \Gamma, ?_{x_1}C_1, \ldots, ?_{x_n}C_n, !_1A}\,\mathit{cut}}{\vdash \Gamma, ?_0C_1, ?_{x_1}C_1, \ldots, ?_0C_n, ?_{x_n}C_n, !_1A}\,\bar{\mathsf{c}}}{\vdots}\,\mathsf{c}}{\vdash \Gamma, ?_{x_1}C_1, \ldots, ?_{x_n}C_n, !_1A}\,\mathsf{c}$$

This decoration is straight forward: the use of the coweakening forces us to index by 0 in the promotion. The rest works by simply applying the rules.

This concludes the definition of the new elimination cases that need to be considered in $\mathsf{G}_{\mathcal{S}}$DILL.

4.2.5 Discussion on a cut elimination procedure

We sadly do not give any proof that the cut elimination cases that we gave led to some rewriting strategies that converge. However, we hope that adapting work by Pagani and Tranquilli may lead to a weakly normalizing rewriting relation [Pag09,PT17]. Those papers study the cut elimination procedure of DiLL using proof nets. Pagani proved that the cut elimination procedure is weakly normalizing [Pag09] and Pagani and Tranquilli extend this result to show that the rewriting is strongly normalizing (quotienting the rewriting by some cases) [PT17].

The proof of strong normalization is based on several techniques. The proof sketch used by Pagani and Tranquilli, that may hopefully be applied for $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$, starts from the fact that that DiLL weakly normalizing, then they define a system equivalent to DiLL and prove that in this system the weak normalization implies the strong one and they deduce that the rewriting procedure of DiLL is strongly normalizing. The equivalent system replaces the rules of dereliction and codereliction with the rules of annihilation and creation. This is something that can be done in graded sequent calculus as well by defining those new rules as

$$\dfrac{\vdash \Gamma, A \qquad \vdash \Delta, !_x A}{\vdash \Gamma, \Delta, !_{x+1} A}\ \partial \ \triangleq\ \dfrac{\dfrac{\vdash \Gamma, A}{\vdash \Gamma, !_1 A}\ \bar{\mathsf{d}} \qquad \vdash \Delta, !_x A}{\vdash \Gamma, \Delta, !_{x+1} A}\ \bar{\mathsf{c}} \qquad \dfrac{\vdash \Gamma, A, ?_x A}{\vdash \Gamma, ?_{x+1} A}\ \varrho \ \triangleq\ \dfrac{\dfrac{\vdash \Gamma, A, ?_x A}{\vdash \Gamma, ?_1 A, ?_x A}\ \mathsf{d}}{\vdash \Gamma, ?_{x+1} A}\ \mathsf{c}$$

and there are in fact equivalent to have a dereliction and a codereliction using

$$\dfrac{\vdash \Gamma, A}{\vdash \Gamma, !_1 A}\ \bar{\mathsf{d}} \ \triangleq\ \dfrac{\vdash \Gamma, A \qquad \dfrac{}{\vdash !_0 A}\ \bar{\mathsf{w}}}{\vdash \Gamma, !_1 A}\ \partial \qquad \dfrac{\vdash \Gamma, A}{\vdash \Gamma, !_1 A}\ \mathsf{d} \ \triangleq\ \dfrac{\dfrac{\vdash \Gamma, A}{\vdash \Gamma, A, ?_0 A}\ \mathsf{w}}{\vdash \Gamma, ?_1 A}\ \varrho$$

The quantities used by Pagani and Tranquilli on proof nets should be definable on sequents, but the main difficulty comes from the new cases and, in particular, the cut elimination case between the promotion and the cocontraction.

4.3 About differential semirings

The notion of differential semiring that we give in Definition 4.1.2 is new. It comes from some equalities that appear at the level of the indices during the cut elimination procedure as defined in Section 4.2

4.3.1 From cut elimination to algebraic restrictions

We have seen that to perform some cut elimination cases, we had to use several of the axioms of the definition of a differential semiring. Actually, these axioms are exactly what is needed to define every cut elimination case of $\mathsf{G}_{\mathcal{S}}\mathsf{DILL}$. Let us see where each algebraic condition from Definition 4.1.2 is used.

1. We need $+$ to be a commutative operation to be able to define the contraction and the cocontraction. Otherwise, the order of the formulas in a sequent could change the index but in the logic we consider we do not care about this order. Furthermore, we must have 0 as the neutral element of the sum to define w_I and $\bar{\mathsf{w}}_I$: since we operate in a naturally ordered semiring, $0 + x = x$ implies that $0 \leq x$ for each $x \in \mathcal{S}$, and this is essential for defining w_I and $\bar{\mathsf{w}}_I$. Note that we could have asked for 0 to be neutral on only one side, but the commutativity of the sum implies that it would then be neutral in general.

2. The additive splitting property is required for the cut between a contraction and a cocontraction. This is explained in Chapter 2 when we define the cut elimination procedure of $\mathsf{DB}_{\mathcal{M}}\mathsf{LL}$.

3. Having 0 left absorbing for the product is required for the cut p/w, which was already present in $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, and for the cut p/w_I, which is new.

4. The element 1 needs to be left neutral for the product to perform the cut between the promotion and the dereliction. Here again, this was needed for $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ as well.

5. The distributivity of the product over the sum comes from several cases. The right distributivity was already mandatory for $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, because of the case between the promotion and the contraction. The left distributivity is used in a new case defined here: the interaction between the promotion and the cocontraction.

6. The associativity is used for the cut between two promotions, and is needed for $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$ as well.

7. The naturally ordered condition is helpful for the case between a promotion and an indexed dereliction. This case requires the order to be right monotonous: we need that if $x \leq y$ then $xz \leq yz$. This is implied by the fact that the order is natural: if $x + x' = y$ then $x + x' + z = y + z$. The naturalness condition is also used for the interaction p/w_I, as it implies that each element of the semiring is greater than 0.

8. The discreteness condition is used in the new cases $\bar{\mathsf{d}}/\mathsf{c}$ and $\mathsf{d}/\bar{\mathsf{c}}$. As we explained, those cases change when they are graded, and we use the discreteness to have results on x and y when $x + y = 1$. Note that this condition states that if $x + y = 1$ then $x = 0$ or $y = 0$. This directly implies that the other element is equal to 1, which is what we use to perform the cut elimination.

9. The integral domain property is used for the case between promotion and coweakening.

10. We need the multiplicative splitting property for the cut between a promotion and a cocontraction.

11. Finally, we ask for a non-unity property. This is used in the case with p and $\bar{\mathsf{d}}$ to be able to characterize what happens when $xy = 1$.

4.3.2 Examples

Now that we have described the origin of the notion of differential semiring, let us give some examples. Note that this new notion is neither weaker nor stronger than the usual notion of semiring. For example, in a semiring, we ask for 0 to be absorbing and 1 to be neutral for the product, but in differential semiring, this condition is weakened by asking only for left absorbance and left neutrality. On the other hand, some axioms in the definition of a differential semiring are strong conditions not required for usual semirings such as the discreteness or the multiplicative splitting property.

Some semirings are often used as a set of indices for $\mathsf{B}_{\mathcal{S}}\mathsf{LL}$, as they can model interesting situations. Breuvart and Pagani present some of them, which are the booleans, the natural numbers, the polynomials with natural number coefficients or the positive real numbers [BP15].

Example 4.3.1. *The boolean semiring* $(\{0,1\}, 0, \vee, 1, \wedge)$ *is not a differential semiring.*

Proof. One axiom that is not satisfied is the one of discreteness. In the boolean semiring, $1 + 1 = 1$ which is a counter-example of this axiom. □

Example 4.3.2. *The natural number semiring* $(\mathbb{N}, 0, +, 1, \times)$ *is a differential semiring.*

Proof. Each axiom from Definition 4.1.2 is obviously satisfied by the natural semiring, except for the splitting axioms. We have proved in Example 2.3.2 that natural numbers satisfy the additive splitting property. For the multiplicative splitting, suppose that we have $x, y, z, r \in \mathbb{N}$ such that $x + y = rz$. Then, defining $k = r$, $l = z$ and $r_i = 1 = z_j$ for each $1 \leq i \leq k$ and $1 \leq j \leq l$ we have

$$r = \sum_{i=1}^{k} r_i \qquad\qquad z = \sum_{j=1}^{l} z_j$$

and defining U as the subsets of the x first elements of $\{1, \ldots, k\} \times \{1, \ldots, l\}$ (with the lexicographic order) we have

$$x = \sum_{(i,j)\in U} r_i z_j \qquad\qquad y = \sum_{(i,j)\notin U} r_i z_j$$

since each $r_i z_j = 1$ and U has x elements and $\{1, \ldots, k\} \times \{1, \ldots, l\} \backslash U$ has $rz - x = y$ elements. This proves that $\mathbb{N}$ fulfills the multiplicative splitting property and is then a differential semiring. □

Example 4.3.3. *The semiring of polynomials with natural number coefficients* $(\mathbb{N}[X_i]_{i\in\mathbb{N}}, 0, +, 1, \times)$ *is a differential semiring.*

Proof. It is well known that most of the axioms are satisfied, since $\mathbb{N}[X_i]_{i\in\mathbb{N}}$ is a semiring. Hence, those that we need to prove are the two splitting axioms, the discreteness, the integral domain and the non-unity property.

- Let $P_1, P_2, P_3, P_4 \in \mathbb{N}[X_i]_{i\in\mathbb{N}}$ such that $P_1 + P_2 = P_3 + P_4$. Writing $P_i = \sum_{\alpha\in\mathbb{N}^\omega} a_{i,\alpha} X^\alpha$ we have
$$a_{1,\alpha} + a_{2,\alpha} = a_{3,\alpha} + a_{4,\alpha}$$
for each $\alpha \in \mathbb{N}^\omega$. Since these $a_{i,\alpha}$ are integers, we use the additive splitting property over the integers to deduce that there are elements $a_{1,3,\alpha}, a_{1,4,\alpha}, a_{2,3,\alpha}, a_{2,4,\alpha} \in \mathbb{N}$ such that
$$\begin{aligned} a_{1,\alpha} &= a_{1,3,\alpha} + a_{1,4,\alpha} & \qquad a_{2,\alpha} &= a_{2,3,\alpha} + a_{2,4,\alpha} \\ a_{3,\alpha} &= a_{1,3,\alpha} + a_{2,3,\alpha} & \qquad a_{4,\alpha} &= a_{1,4,\alpha} + a_{2,4,\alpha} \end{aligned}$$
for each α. Defining $P_{i,j} = \sum_\alpha a_{i,j,\alpha} X^\alpha$ for $i = 1$ or 2 and $j = 3$ or 4, we get the desired decomposition.

- For the multiplicative splitting, let us suppose that there are four polynomials, such that $PQ = R + S$. We write
$$P = \sum_{i=0}^{n} p_i X^{\alpha_i} \qquad Q = \sum_{j=0}^{m} q_j X^{\beta_j} \quad \text{and we have} \quad PQ = \sum_{\substack{1\leq i\leq n \\ 1\leq j\leq m}} p_i q_j X^{\alpha_i+\beta_j} = R + S$$
We use then a technique similar to our decomposition for natural numbers. Writing $r_{i,j}$ and $s_{i,j}$ the coefficients of $X^{\alpha_i+\beta_j}$ in respectively R and S, we know that $p_i q_j = r_{i,j} + s_{i,j}$. The splitting for P and Q is defined as follows:
$$P = \underbrace{X^{\alpha_1} + \cdots + X^{\alpha_1}}_{p_1 \text{ times}} + \cdots + \underbrace{X^{\alpha_n} + \cdots + X^{\alpha_n}}_{p_n \text{ times}}$$

for P and

$$Q = \underbrace{X^{\beta_1} + \cdots + X^{\beta_1}}_{q_1 \text{ times}} + \cdots + \underbrace{X^{\beta_m} + \cdots + X^{\beta_m}}_{q_m \text{ times}}$$

for Q. Using this decomposition, the product PQ is equal to

$$PQ = \underbrace{X^{\alpha_1+\beta_1} + \cdots + X^{\alpha_1+\beta_1}}_{p_1 q_1 \text{ times}} + \cdots + \underbrace{X^{\alpha_n+\beta_m} + \cdots + X^{\alpha_n+\beta_m}}_{p_n q_m \text{ times}}$$

and the set U is defined by putting the first $r_{i,j}$ elements of each coefficient in the set, and the other not in the set. This produces a correct splitting, using that $p_i q_j = r_{i,j} + s_{i,j}$.

- For the discreteness, if P and Q are two polynomials with natural number coefficients, the degree of $P + Q$ is greater or equal than the of P and than the one of Q. So if $P + Q = 0$ then both P and Q are in $\mathbb{N}$, and we use that the semiring of natural numbers is discrete to conclude.

- For the integral domain property and the non-unity, we use a similar result on the sum: if PQ has degree zero, either $P = 0$ or $Q = 0$ or both are natural numbers, which allow us to conclude these two properties.

□

Example 4.3.4. *The semiring of positive real numbers* $(\mathbb{R}^+, 0, +, 1, \times)$ *is not a differential semiring.*

Proof. This semiring does not fulfill the discreteness property, as for instance, $2 \times \frac{1}{2} = 1$. □

Conclusion

In this thesis, we aimed to study proof systems representing the computation and the resolution of partial differential equations. To do so, we have presented several ways to unify differential linear logic [ER06] and graded linear logic [GS14]. This unification is crucial, as several models of differential linear logic have been studied over the years. Those models give multiple ways of doing functional analysis, and have various properties that are useful for the study of differential equations. But the syntax of differential linear logic is too restricted to represent partial differential equations. This is why we have looked for a unification with graded linear logic. This setting has been defined to endow linear logic with some parameters. Depending on the choice of these parameters, some quantitative properties can be represented in graded linear logic. Our starting point was to try building a framework, based on differential linear logic, but having the possibility of adding parameters to it, hoping for a system where we could interpret those parameters by partial differential equations.

For the partial differential equations, we have limited ourselves to linear partial differential equations with constant coefficients (LPDOcc). The linearity condition fits well with (differential) linear logic, and asking for constant coefficients simplifies the study of the solutions. Each LPDOcc has a *fundamental solution*, which gives a direct expression of the solutions of the equation for every possible parameter. This strong property is crucial in this work, as it is on the basis of the links we made between the differentiation and the gradation. We use an analogy stating that the dereliction of linear logic could be seen as an operation solving a differential equation, while the codereliction of differential linear logic applies a differential operator. Extending this idea to LPDOcc causes a mismatch on the types, but this mismatch is solved using the subtyping rule of graded linear logic. Using this rule to represent the resolution of a differential equation gives us a first hint at how to relate graded and differential linear logic. But to do so, we need to be able to solve differential equations easily, so we restricted ourselves to LPDOcc.

We have used two mathematical frameworks to represent the action of the differential operators. The first one uses smooth functions and their dual which is the space of distributions. This setting is a model of differential linear logic when we consider polarized connectives [Ker18a]. Smooth functions and distributions are well suited for a study of linear partial differential operators. Hormander conducted quite a complete review of LPDOcc acting on distributions [Hor63]. In addition, the notion of fundamental solutions is defined using distributions. However, we had to consider another framework as a model. We used the notion of Köthe spaces, which is a model of linear logic studied by Ehrhard [Ehr02], that has led Ehrhard and Regnier to the definition of differential linear logic. This model is more discrete than smooth functions and distributions, since it is based on spaces of sequences. But those sequences can be identified to analytic functions, on which we can compute partial differentiation.

Contributions

The concrete contributions of this thesis were to present several proof systems, enjoying both graded connectives and costructural rules to interpret differential operators. For those systems, we have adapted the model with smooth functions and distribution to incorporate the action of LPDOcc, and the restrictions coming from the syntax of the proof systems.

Unified graded and differential linear logic without higher order. In Chapter 2 we present our first attempt of a proof system unifying graded and differential linear logic. We show that adding indices to differential linear logic or adding the rules used to differentiate to graded linear logic results in fact in the same logic, called IDiLL. We define a cut elimination procedure on this logic, based on a refined notion of monoid where the indices belong. Then, we give a model which consists of smooth functions and distributions, which are respectively spaces of solutions and of parameters of differential equations. In this model, the exponentials are indexed by a monoid of LPDOcc. But in this model, the formulas cannot be considered at higher order, which prevents us from considering a promotion rule. We have then removed this rule, following the first version of differential linear logic. The next goal that we have focused on was the extension of this logic to a promotion rule, and the extension of this model to higher order.

An extension of IDiLL through the Laplace transform. In Chapter 3 we refine the syntax of IDiLL, and, in particular, the grammar of its exponential rules. Rather than having an algebraic structure containing the possible indices, we consider two copies of the same algebraic structure. This comes with an additional duality on the exponential formulas. In addition to the usual negation, we add an orthogonal duality, which is the Laplace transform. This new transformation comes from semantical considerations. In a model studied by Kerjean and Lemay, the Laplace transform is, in fact, an isomorphism that turns the structural rules into the costructural ones [KL23]. This new logic with two copies of the indices is called DIDiLL. It has a promotion rule which allows considering higher order objects and also we also define a copromotion rule, inspired by the codigging morphism studied by Kerjean and Lemay [KL23]. We define the cut elimination cases of this logic and conjecture a normalization property. We give two models of DIDiLL. The first one uses Köthe spaces and LPDOcc. We had to restrain ourselves to Köthe spaces to take into account the variables of differentiation of the operators. The second one is a refinement of smooth functions and distributions where the growth of the functions is bounded by exponentials of Young functions.

A higher order unified logic indexed by a discrete semiring. The last proof system that we consider is defined in Chapter 4. This is another extension of IDiLL to higher order. The difference with DIDiLL is that, rather than taking ideas from the semantics to refine the formulas, we study what is happening in the syntax when one adds a promotion rule to IDiLL. This study is based on the cut elimination cases occurring with the promotion. While in graded linear logic, those cases have the same dynamics as in linear logic, it is not the case anymore with the costructural rules. For instance, the sums of proofs needed in differential linear logic disappear in a graded framework. In addition, this study led to the definition of a new algebraic notion, which we called differential semiring. This structure corresponds to the structure needed on the indices to perform the cut elimination cases. One crucial property of differential semirings is that they are discrete. It prevents LPDOcc from being a differential semiring. In a sense this justifies the logic DlDiLL to be the right framework for a graded differential linear logic where we can interpret the action of partial differential operators.

What's next?

By building bridges between differential linear logic and graded linear logic, this thesis leads to many interesting new questions to study.

- One missing point of this manuscript is the study of the categorical semantics of the systems that we have consider. While the categorical semantics of differential linear logic has been studied a lot [BCLS20], as well as the one of graded linear logic [BGMZ14, BP15, GKO+16], one would like to unify those semantics, leading to a definition for our logics. Lemay and Vienney have made some advances in this direction [LV23]. They, however, do not consider the indexed (co)derelictions, which are crucial for us in the models that we consider. Kerjean and Lemay have also defined a categorical framework for the Laplace transform [KL24] which should be part of the axiomatization of the logic DlDiLL.

- A point that we would like to improve is the differential operators that we use. We have limited ourselves to linear partial differential operators with constant coefficients. They represent interesting differential equations, such as the heat or the Laplacian equations. But being able to generalize to more general operators would be a nice result. In order to do so, one would need to study carefully the resolutions of such operators, as the existence of fundamental solutions would not be ensured anymore. This may also need to incorporate a notion of fixpoint to approach some solutions. Some work on connecting differentiation and fixpoints has been made by Galal and Lemay [GPL24].

- It would also be interesting to study frameworks closer to the computation. The logics that we have defined give nice hints for the types and their inhabitants of a system solving partial differential equations. But to take into account the computations more precisely, one may look for a λ-calculus, which would be typed by graded differential linear logic. Here again, the use of fixpoint would probably be crucial in order to approach functions and their derivatives.

Bibliography

[And90] Jean-Marc Andreoli. *Proposition pour une synthese des paradigmes de la programmation logique et de la programmation par objets.* PhD thesis, Paris 6, 1990.

[BCLS20] R. F. Blute, J. R. B. Cockett, J.-S. P. Lemay, and R. A. G. Seely. Differential categories revisited. *Applied Categorical Structures*, 2020. doi:10.1007/s10485-019-09572-y.

[BCS06] Rick Blute, Robin Cockett, and Robert Seely. Differential categories. *Mathematical Structures in Computer Science*, 16(6), 2006.

[BET12] Rick Blute, Thomas Ehrhard, and Christne Tasson. A convenient differential category. *Les cahiers de topologie et de géométrie différentielle catégorique*, 2012.

[BGMZ14] Aloïs Brunel, Marco Gaboardi, Damiano Mazza, and Steve Zdancewic. A core quantitative coeffect calculus. In *Programming Languages and Systems*. Springer Berlin Heidelberg, 2014.

[BKM23] Flavien Breuvart, Marie Kerjean, and Simon Mirwasser. Unifying Graded Linear Logic and Differential Operators. In *8th International Conference on Formal Structures for Computation and Deduction (FSCD 2023)*, Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.FSCD.2023.21.

[BKM26] Flavien Breuvart, Marie Kerjean, and Simon Mirwasser. Unifying graded linear logic and differential operators. *Logical Methods in Computer Science*, Volume 22, Issue 1, Jan 2026. URL: https://lmcs.episciences.org/13064, doi:10.46298/lmcs-22(1:7)2026.

[Bos09] Siegfried Bosch. *Algebra.* Springer-Lehrbuch. Springer, 2009.

[BP15] Flavien Breuvart and Michele Pagani. Modelling Coeffects in the Relational Semantics of Linear Logic. In *Computer Science Logic (CSL)*, Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl, 2015.

[CEOO02] M. B. Chrouda, M. El Oued, and H. Ouerdiane. Convolution calculus and applications to stochastic differential equations. *Soochow Journal of Mathematics*, 28(4), 2002.

[CES10] Alberto Carraro, Thomas Ehrhard, and Antonino Salibra. Exponentials with infinite multiplicities. In Anuj Dawar and Helmut Veith, editors, *Computer*

Science Logic, 24th International Workshop, CSL 2010, 19th Annual Conference of the EACSL, Brno, Czech Republic, August 23-27, 2010. Proceedings, volume 6247 of *Lecture Notes in Computer Science*, pages 170–184. Springer, 2010. doi:10.1007/978-3-642-15205-4_16.

[Chu32] Alonzo Church. A set of postulates for the foundation of logic. *Annals of Mathematics*, 33(2):346–366, 1932. URL: http://www.jstor.org/stable/1968337.

[Cur34] H. B. Curry. Functionality in combinatory logic. *Proceedings of the National Academy of Sciences of the United States of America*, 20(11):584–590, 1934. URL: http://www.jstor.org/stable/86796.

[DE11] Vincent Danos and Thomas Ehrhard. Probabilistic coherence spaces as a model of higher-order probabilistic computation. *Inf. Comput.*, 209(6):966–991, June 2011. doi:10.1016/j.ic.2011.02.001.

[DG24] Rémi Di Guardia. *Identity of Proofs and Formulas using Proof-Nets in Multiplicative-Additive Linear Logic*. PhD thesis, 2024. Thèse de doctorat dirigée par Laurent, Olivier Informatique Lyon, École normale supérieure 2024.

[DK09] Manfred Droste and Werner Kuich. *Semirings and Formal Power Series*, pages 3–28. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. doi:10.1007/978-3-642-01492-5_1.

[Ehr02] Thomas Ehrhard. On Köthe Sequence Spaces and Linear Logic. *Mathematical Structures in Computer Science*, 12(5), 2002.

[Ehr05] Thomas Ehrhard. Finiteness spaces. *Mathematical Structures in Computer Science*, 15(4), 2005.

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

[Ehr23] Thomas Ehrhard. Coherent differentiation. *Mathematical Structures in Computer Science*, 33(4–5):259–310, 2023. doi:10.1017/S0960129523000129.

[Ehr24] Thomas Ehrhard. From differential linear logic to coherent differentiation, 2024. URL: https://arxiv.org/abs/2401.14834, arXiv:2401.14834.

[ER03] Thomas Ehrhard and Laurent Regnier. The differential lambda-calculus. *Theoretical Computer Science*, 309, 2003. doi:10.1016/S0304-3975(03)00392-X.

[ER06] Thomas Ehrhard and Laurent Regnier. Differential interaction nets. *Theoretical Computer Science*, 364(2), 2006.

[Fio09] Marcelo P. Fiore. Differential structure in models of multiplicative biadditive intuitionistic linear logic. In *TLCA 2007,Proceedings*, volume 4583 of *Lecture Notes in Computer Science*. Springer, 2009.

[Flo67] Robert W. Floyd. *Assigning Meanings to Programs*, volume 19, pages 91–32. American Mathematical Society, 1967.

[GHH+13] Marco Gaboardi, Andreas Haeberlen, Justin Hsu, Arjun Narayan, and Benjamin C. Pierce. Linear dependent types for differential privacy. *SIGPLAN Not.*, 48(1):357–370, January 2013. `doi:10.1145/2480359.2429113`.

[GHOR00] R. Gannoun, R. Hachaichi, H. Ouerdiane, and A. Rezgui. Un théorème de dualité entre espaces de fonctions holomorphes à croissance exponentielle. *Journal of Functional Analysis*, 171(1), 2000.

[Gim09] Stéphane Gimenez. *Programmer, calculer et raisonner avec les réseaux de la Logique Linéaire*. Theses, Université Paris-Diderot, 2009.

[Gir87] Jean-Yves Girard. Linear logic. *Theoretical Computer Science*, 50(1), 1987.

[Gir88a] Jean-Yves Girard. Normal functors, power series and λ-calculus. *Annals of Pure and Applied Logic*, 1988.

[Gir88b] Jean-Yves Girard. Normal functors, power series and λ-calculus. *Ann. Pure Appl. Logic*, 37(2):129–177, 1988. `doi:10.1016/0168-0072(88)90025-5`.

[Gir91] Jean-Yves Girard. A new constructive logic: classic logic. *Mathematical Structures in Computer Science*, 1(3):255–296, 1991. `doi:10.1017/S0960129500001328`.

[Gir98] Jean-Yves Girard. Light linear logic. *Information and Computation*, 143(2):175–204, 1998. URL: `https://www.sciencedirect.com/science/article/pii/S0890540198927006`, `doi:https://doi.org/10.1006/inco.1998.2700`.

[GKO+16] Marco Gaboardi, Shin-ya Katsumata, Dominic Orchard, Flavien Breuvart, and Tarmo Uustalu. Combining effects and coeffects via grading. In *Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming*, International Conference on Functional Programming, ICFP. Association for Computing Machinery, 2016.

[GPL24] Zeinab Galal and Jean-Simon Pacaud Lemay. Combining fixpoint and differentiation theory. In *Proceedings of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '24, New York, NY, USA, 2024. Association for Computing Machinery. `doi:10.1145/3661814.3662108`.

[Gri90] Timothy G. Griffin. A formulae-as-types notion of control. In *In Conference Record of the Seventeenth Annual ACM Symposium on Principles of Programming Languages*. ACM Press, 1990.

[Gro66] Alexandre Grothendieck. Produits tensoriels topologiques et espaces nucléaires. *Memoirs of the AMS*, 16, 1966. Publisher: American Mathematical Society.

[GS14] Dan Ghica and Alex I. Smith. Bounded linear types in a resource semiring. In *Programming Languages and Systems,*, European Symposium on Programming, (ESOP). Springer Berlin Heidelberg, 2014.

[GSS91] Jean-Yves Girard, Andre Scedrov, and Philip Scott. Bounded linear logic. *Theoretical Computer Science*, 9, 08 1991.

[Hor63] Lars Hormander. *Linear partial differential operators*. Springer Berlin, 1963.

[How69] William A. Howard. The formulae-as-types notion of construction. 1969. URL: https://api.semanticscholar.org/CorpusID:118720122.

[Jar81] Hans Jarchow. Locally convex spaces. B. G. Teubner Stuttgart, 1981. Mathematical Textbooks.

[Ker18a] Marie Kerjean. A logical account for linear partial differential equations. In *Logic in Computer Science (LICS), Proceedings*. Association for Computing Machinery, 2018.

[Ker18b] Marie Kerjean. *Reflexive spaces of smooth functions : a logical account of linear partial differential equations*. PhD thesis, 2018. Thèse de doctorat dirigée par Ehrhard, Thomas Informatique. Informatique fondamentale Sorbonne Paris Cité 2018.

[KL19] Marie Kerjean and Jean-Simon Pacaud Lemay. Higher-order distributions for differential linear logic. In *Foundations of Software Science and Computation Structures FOSSACS 2019 Proceedings*, Lecture Notes in Computer Science. Springer, 2019.

[KL23] Marie Kerjean and Jean-Simon Pacaud Lemay. Taylor expansion as a monad in models of dill. In *38th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2023, Boston, MA, USA, June 26-29, 2023*, 2023. doi:10.1109/LICS56636.2023.10175753.

[KL24] Marie Kerjean and Jean-Simon Pacaud Lemay. Laplace distributors and laplace transformations for differential categories. In Jakob Rehof, editor, *FSCD 2024 Tallinn, Estonia*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024. doi:10.4230/LIPICS.FSCD.2024.9.

[Laf04] Yves Lafont. Soft linear logic and polynomial time. *Theoretical Computer Science*, 318(1):163–180, 2004. Implicit Computational Complexity. URL: https://www.sciencedirect.com/science/article/pii/S0304397503005231, doi:https://doi.org/10.1016/j.tcs.2003.10.018.

[Lai16] J. Laird. Fixed points in quantitative semantics. In *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '16, page 347–356, New York, NY, USA, 2016. Association for Computing Machinery. doi:10.1145/2933575.2934569.

[Lam92] François Lamarche. Quantitative domains and infinitary algebras. *Theoretical Computer Science*, 94(1):37–62, 1992. URL: https://www.sciencedirect.com/science/article/pii/0304397592903238, doi:https://doi.org/10.1016/0304-3975(92)90323-8.

[Lap72] Miguel L. Laplaza. Coherence for distributivity. In G. M. Kelly, M. Laplaza, G. Lewis, and Saunders Mac Lane, editors, *Coherence in Categories*, pages 29–65, Berlin, Heidelberg, 1972. Springer Berlin Heidelberg.

[Lau02] O. Laurent. *Etude de la polarisation en logique*. Thèse de Doctorat, Université Aix-Marseille II, March 2002.

[Lau04] Olivier Laurent. Polarized games. *Annals of Pure and Applied Logic*, 130(1):79–123, 2004. Papers presented at the 2002 IEEE Symposium on Logic in Computer Science (LICS). URL: `https://www.sciencedirect.com/science/article/pii/S0168007204000697`, `doi:https://doi.org/10.1016/j.apal.2004.04.006`.

[Lau08] Olivier Laurent. Théorie de la démonstration. University lecture, 2008.

[LMMP13] Jim Laird, Giulio Manzonetto, Guy McCusker, and Michele Pagani. Weighted relational models of typed lambda-calculi. In *2013 28th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 301–310, 2013. `doi:10.1109/LICS.2013.36`.

[LV23] Jean-Simon Pacaud Lemay and Jean-Baptiste Vienney. Graded differential categories and graded differential linear logic, 2023. preprint.

[Mel09] Paul-André Melliès. Categorical semantics of linear logic. Société Mathématique de France, 2009.

[Mel12] Paul-André Melliès. Parametric monads and enriched adjunctions, 2012. Technical report. URL: `https://www.irif.fr/~mellies/tensorial-logic/8-parametric-monads-and-enriched-adjunctions.pdf`.

[MT10] Paul-André Melliès and Nicolas Tabareau. Resource modalities in tensor logic. *Annals of Pure and Applied Logic*, 161(5):632–653, 2010. The Third workshop on Games for Logic and Programming Languages (GaLoP). URL: `https://www.sciencedirect.com/science/article/pii/S0168007209001602`, `doi:https://doi.org/10.1016/j.apal.2009.07.018`.

[Mun24] Nicolas Munnich. *Operational and categorical models of PCF : addressing machines and distributing semirings*. PhD thesis, 2024. Thèse de doctorat dirigée par Manzonetto, Giulio et Breuvart, Flavien Informatique Paris 13 2024. URL: `http://www.theses.fr/2024PA131015`.

[Pag09] Michele Pagani. The cut-elimination theorem for differential nets with promotion. In *International Conference on Typed Lambda Calculus and Applications*, 2009.

[Pro23] The LL Handbook Project. Handbook of linear logic. Draft, online handbook, 2023. URL: `https://ll-handbook.frama.io/ll-handbook/ll-handbook-public.pdf`.

[PSV14] Michele Pagani, Peter Selinger, and Benoît Valiron. Applying quantitative semantics to higher-order quantum computing. In *Proceedings of the 41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, POPL '14, page 647–658, New York, NY, USA, 2014. Association for Computing Machinery. `doi:10.1145/2535838.2535879`.

[PT17] MICHELE PAGANI and PAOLO TRANQUILLI. The conservation theorem for differential nets. *Mathematical Structures in Computer Science*, 27(6):939–992, 2017. `doi:10.1017/S0960129515000456`.

[Ros14] J. Rosser. Curry haskell b. and feys robert. combinatory logic. volume i. with two sections by william craig. studies in logic and the foundations of mathematics. north-holland publishing company, amsterdam 1958, xvi + 417 pp. *The Journal of Symbolic Logic*, 32:267–268, 08 2014. doi:10.1017/S0022481200114203.

[Sco77] Dana Scott. Outline of a mathematical theory of computation. *Kibernetich-eskij Sbornik. Novaya Seriya*, 14, 01 1977.

[Sco82] Dana Scott. Domains for denotational semantics. pages 577–613, 01 1982. doi:10.1007/BFb0012801.

[Tre67] F. Treves. *Topological Vector Spaces, Distributions and Kernels*. Elsevier Science, 1967.