

# Elements of Quantitative Rewriting

FRANCESCO GAVAZZO, University of Pisa, Italy

CECILIA DI FLORIO, University of Bologna, Italy

We introduce a general theory of quantitative and metric rewriting systems, namely systems with a rewriting relation enriched over quantales modelling abstract quantities. We develop theories of abstract and term-based systems, refining cornerstone results of rewriting theory (such as Newman's Lemma, Church-Rosser Theorem, and critical pair-like lemmas) to a metric and quantitative setting. To avoid distance trivialisation and lack of confluence issues, we introduce non-expansive, linear term rewriting systems, and then generalise the latter to the novel class of graded term rewriting systems. These systems make quantitative rewriting modal and context-sensitive, this way endowing rewriting with coeffectful behaviours.

CCS Concepts: • **Theory of computation** → **Equational logic and rewriting**; **Program semantics**.

Additional Key Words and Phrases: quantitative rewriting, metric rewriting, modal graded rewriting, quantitative equational theory, quantal relations, quantitative calculus of relations

## ACM Reference Format:

Francesco Gavazzo and Cecilia Di Florio. 2023. Elements of Quantitative Rewriting. *Proc. ACM Program. Lang.* 7, POPL, Article 63 (January 2023), 32 pages. <https://doi.org/10.1145/3571256>

## 1 INTRODUCTION

Modern mathematics begins with *symbolic manipulation*. The central role of signs and symbols *per se* is one of the main achievement of the Medieval culture [Meier-Oeser 2011] leading, among others, to the development of elementary or *symbolic algebra*. Starting from the latter, the syntactic manipulation of symbols more or less independently of their meaning — i.e. to what symbols stand for — has become an essential part of mathematical reasoning, not to say of reasoning *in general*. Today, symbolic manipulation is not just a pillar of mathematics, but it is at the very hearth of *computation*. Indeed, the symbolic manipulations of elementary algebra carry a computational content and, vice versa, computational processes can be fully described symbolically.

Rewriting theory [Bezem et al. 2003; Newman 1942] is the discipline that studies (the computational content of) symbolic manipulation in general. As such, rewriting has its origin both in symbolic algebra as the study of the algorithmic properties of equational reasoning, and in computability and programming language theory, where rewriting systems have been used to define symbolic models of computation — such as the  $\lambda$ -calculus [Barendregt 1984] and combinatory logic [Curry and Feys 1958a; Hindley and Seldin 2008] — as well as the (operational) semantics and implementation of programming languages [Asperti and Guerrini 1998; Jones 1987]. In both cases, rewriting is motivated by the need to define *operational* notions of equality revealing the computational content of equational deductions. Remarkably, operationality is ultimately achieved by making equality asymmetric, so that the aforementioned computational content can be fully uncovered by orienting equations. Nowadays, these oriented equations (and the evolution thereof)

---

Authors' addresses: Francesco Gavazzo, francesco.gavazzo@unipi.it, University of Pisa, Italy; Cecilia Di Florio, cecilia.diflorio2@unibo.it, University of Bologna, Italy.

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2023 Copyright held by the owner/author(s).

2475-1421/2023/1-ART63

<https://doi.org/10.1145/3571256>

are known as *rewriting* — or *reduction* — relations. All of that highlights a crucial trait of rewriting theory, namely its deep connection with equational reasoning. In fact, rewriting does not actually focus on arbitrary symbolic transformations, but with *equality-preserving* ones: a rewriting relation *refines* equality by making the latter operational, and it is thus contained in it.

Recent advances in theoretical computer science, however, have questioned the central role played by equality in semantics, arguing for more quantitative and approximated forms of equivalence. For instance, equality is a too strong notion for reasoning about probabilistic computations, where even small perturbations break the equivalence between probabilistic processes. To overcome this problem, researchers have thus refined equality to *distances* between probabilistic processes, this way replacing equivalences with *metrics*. Similarly, metric-based and approximated equivalences have been used to reason about privacy and security of systems [de Amorim et al. 2017; Reed and Pierce 2010], not to mention reasoning about intensional aspects of computation, such as resource consumption [Dal Lago and Gavazzo 2021b, 2022b].

Prompted by that, several theories of semantic equality have been refined giving rise to *quantitative theories of semantic differences*, prime examples being general theories of program [Arnold and Nivat 1980; Crubillé and Dal Lago 2015, 2017; Dal Lago and Gavazzo 2020, 2021a, 2022a,b; Dal Lago et al. 2019; de Amorim et al. 2017; de Bakker and Zucker 1982; Escardo 1999; Gavazzo 2018, 2019; Reed and Pierce 2010] and system distances [Baldan et al. 2014, 2015; Du et al. 2016; Ferns et al. 2004, 2005; Gebler et al. 2016] and the theory of quantitative algebras and quantitative equational reasoning [Bacci et al. 2018, 2021; Mardare et al. 2016, 2017, 2018, 2021; Mio et al. 2021]. The latter, in particular, aim to provide a common foundation for general quantitative reasoning by refining traditional, set-based algebraic structures to metric-like ones and by replacing traditional equations with *quantitative equations* bounding the difference (or distance) between the equated elements. Accordingly, classic equations  $t = s$  are replaced by expressions of the form  $t \stackrel{\varepsilon}{=} s$ , with the informal reading that  $t$  and  $s$  are at most  $\varepsilon$  apart, or that they are equal up to an error  $\varepsilon$ . Thus, quantitative algebraic theories are not theories about *equality* between objects, but about *distances* between them, and can thus be seen as the quintessence of quantitative and metric reasoning.

But what about the *computational content* of quantitative equational reasoning? What is an *operational* notion of quantitative equality or distance allowing us to effectively compute distances by means of quantitative equations? And, more generally, what is the theory of *quantitative symbolic manipulation*, where symbolic transformations can break semantic equivalence? The development of such a theory, which is the main topic of this work, is of paramount importance not only to make quantitative equational deduction effective, but also to develop a general quantitative theory of programming language semantics. In this paper, we introduce the theory of *quantitative and metric rewriting systems* as a first step towards a general theory of the computational content of quantitative symbolic manipulations. Such a theory is rich and largely extends traditional rewriting in a nontrivial way. The goal of this paper is to lay the foundation of quantitative rewriting systems, this way opening the door to a larger research program. In particular, the main contribution of the present work is threefold.

1. We introduce *quantitative abstract rewriting systems* —  $\Omega$ -ARSs — a *quantitative relational* foundation for the whole development of quantitative and metric rewriting. We develop the (meta)theory of  $\Omega$ -ARSs introducing, among others, quantitative refinements of the well-known notions of confluence and termination. Additionally, we refine cornerstone results of abstract rewriting, such as Newman [Newman 1942] and Hindley-Rosen [Hindley 1964; Rosen 1970] Lemma, to a quantitative and metric setting.
2. We then introduce *quantitative non-expansive term rewriting systems* —  $\Omega$ -TRSs, for short — namely term-based rewriting systems where rewriting distances are non-expansively propagated

throughout term constructors of the language. Non-expansive systems provide a powerful tool for the quantitative analysis of computational effects and programming languages, and offer a symbolic perspective on optimisation problems, such as those relating to edit distances. Forcing non-expansiveness of term constructors, however, may lead to undesired phenomena, such as distance trivialisation [Crubillé and Dal Lago 2017; Gavazzo 2019] and the failure of major confluence theorems. To avoid all of that we require non-expansive systems to be linear. In particular, we shall prove general quantitative critical pair-like lemmas [Huet 1980] ensuring confluence of large families of linear non-expansive systems.

3. Finally, we go beyond linearity and non-expansiveness by introducing *graded quantitative term rewriting systems* —  $(\Omega, \Phi)$ -TRS, for short. In such systems, rewriting is not only quantitative but also *modal* and *context-sensitive*, meaning that contexts are not required to non-expansively propagate distances — as in non-expansive systems — but they directly act on them, this way behaving as generalised Lipschitz continuous functions. We will extend the confluence results proved for (linear) non-expansive systems to  $(\Omega, \Phi)$ -TRS, as well as prove an additional confluence result for orthogonal systems. As an example, we shall obtain a quantitative confluence theorem for graded combinatory logic [Abramsky 2002; Abramsky et al. 2002; Atkey 2018] (as well as for its effectful extension), a foundational calculus for coefficientful programming languages [Bernardy et al. 2018; Gaboardi et al. 2016; Ghica and Smith 2014; Orchard et al. 2019; Petricek et al. 2014].

In addition to the aforementioned results, a further contribution of the paper is the methodology employed. All our theory, in fact, is developed following the abstract relational theory of distances initiated by Lawvere [1973], whereby we work with relations taking values in arbitrary quantales [Hofmann et al. 2014; Rosenthal 1990]. Consequently, since abstract metric and modal reasoning are essentially equivalent [Abel and Bernardy 2020; Dal Lago and Gavazzo 2021b, 2022b], our theory can be seen both as a general theory of metric and quantitative rewriting systems and as a theory of modal and substructural rewriting, this way suggesting possible connections with modal and coefficientful systems [Abel and Bernardy 2020; Gaboardi et al. 2016; Ghica and Smith 2014; Orchard et al. 2019; Petricek et al. 2014]. An extended version [Gavazzo and Di Florio 2022] of this paper with detailed proofs and further (minor) results is available.

## 2 QUANTITATIVE AND METRIC REWRITING, AN INVITATION

Before diving into the technical development of quantitative rewriting systems, we gently introduce the reader to the subject by informally looking at a nontrivial running example: the *quantitative Church-Rosser theorem* for systems of quantitative combinatory logic.

### 2.1 From Curry to Lawvere, via Church-Rosser

To motivate the shift from traditional to quantitative rewriting, in this introductory section we consider the problem of extending the well-known Church-Rosser theorem [Church and Rosser 1936] (CR theorem, for short) to quantitative extensions of Curry’s system of combinators [Curry and Feys 1958a,b; Hindley and Seldin 2008], which we generically denote by  $\mathcal{K}$ . Terms of  $\mathcal{K}$ , known as combinators, are built as first order terms over the signature  $\Sigma_{\mathcal{K}}$  consisting of a binary operation for application (as usual, we write application simply as juxtaposition and assume it associates to the right) and the four basic combinators B, C, K, W. System  $\mathcal{K}$  is then defined as the *equational theory* generated by the following axioms:

$$Bxyz \approx x(yz) \quad Cxyz \approx xzy \quad Kxy \approx x. \quad Wxy \approx xyy.$$

The power of  $\mathcal{K}$ , however, is fully exploited only when endowing it with the rewriting relation  $\rightarrow$  obtained by orienting  $\approx$  left-to-right. Accordingly,  $\rightarrow$  provides an operational account of  $\approx$  (and of its associated theory of equational deduction) as well as a formal model of program execution. In

fact, since  $\approx$  coincides with the convertibility relation<sup>1</sup>  $\leftrightarrow^*$  induced by  $\rightarrow$ , so that two combinators  $t$  and  $s$  are equivalent if and only if there is a bidirectional  $\rightarrow$ -path from  $t$  to  $s$ , we can rely on  $\leftrightarrow^*$ , and thus on  $\rightarrow$ , to prove operational properties of  $\approx$ . The celebrated Church-Rosser theorem [Church and Rosser 1936] states, in the case of system  $\mathcal{K}$ , that two combinators are equal if and only if they have a common reduct. Symbolically,  $\leftrightarrow^* = \rightarrow^*; \leftarrow^*$ . Such a theorem, which is equivalent to confluence of  $\rightarrow$ , has several highly nontrivial consequences: it entails uniqueness of normal forms, and thus consistency and semi-decidability of  $\approx$ . Moreover, it can be used to prove soundness of implementations of programming languages based on  $\mathcal{K}$  as well as soundness of powerful optimisation techniques, such as parallelisation [Asperti and Guerrini 1998; Jones 1987].

*Quantitative Combinatory Systems.* Even if system  $\mathcal{K}$  models the basic dynamic of program execution, its minimal syntax is not very well-suited for the analysis of finer programming language features. For that reason, researchers have designed and studied several systems of combinators extending  $\mathcal{K}$ , each focusing on some specific features of interest. To make the point concrete, here we extend  $\mathcal{K}$  with primitive data for numerical computation.<sup>2</sup> For the sake of the argument, here we just consider combinators  $Z$  (zero),  $S$  (successors), and  $A$  (addition) for arithmetic, together with the reductions  $A x Z \rightarrow x$  and  $A x (S y) \rightarrow S (A x y)$  (and their associated equational counterparts). The system thus obtained – which we still denote as  $\mathcal{K}$  – being inherently equational, it allows us to study *equivalence* between combinators, and thus between numerical computations. When dealing with the latter, however, it is oftentimes more interesting to compare programs producing different results and to measure *differences* between them (think about the case of an inefficient but precise numerical computation and of a more efficient although only approximately correct one). It is now well-known that these and similar analyses can be obtained by replacing equality with metric-like functions, to which we generically refer to as *distances*. Remarkably, quantitative and metric reasoning can still be performed in an equational fashion following the methodology of quantitative equational theories Mardare et al. [2016, 2017]. Accordingly, we move from traditional equations  $t \approx s$  to *quantitative equations*, that is *ternary relations*  $\varepsilon \Vdash t \approx s$  (or  $t \stackrel{\varepsilon}{\approx} s$ ) relating pairs of objects  $t, s$  with non-negative real numbers  $\varepsilon$ ,<sup>3</sup> the informal reading of a quantitative equation  $\varepsilon \Vdash t \approx s$  being that  $t$  and  $s$  are at most  $\varepsilon$ -apart.<sup>4</sup>

We can thus endow system  $\mathcal{K}$  with a quantitative equational theory simply by replacing its defining equations with zero-weighted equations – i.e.  $0 \Vdash B x y z \approx x (y z)$ ,  $0 \Vdash C x y z \approx x z y$ ,  $0 \Vdash K x y \approx x$ ,  $0 \Vdash W x y \approx x y y$  – plus equations for numerical combinators – such as  $0 \Vdash A x Z \approx x$  and  $0 \Vdash A x (S y) \approx S (A x y)$  – and adding suitable distance-producing equations. For this example, we consider the simple equation  $1 \Vdash S x \approx x$  defining the quantitative equational theory of the Euclidean distance (even if extremely simple, we will see that developing a well-behaved meta-theory for this system is nontrivial). Indeed, the quantitative equational theory  $\approx$  defines a distance

<sup>1</sup>We use standard notational conventions for binary relations [Bezem et al. 2003].

<sup>2</sup>Further extensions we shall study in this paper include primitives for computational effects and for the intensional analysis of programs. As an example of the former, we will consider, among others, a binary operation  $+_\epsilon$  for ( $\epsilon$ -biased) probabilistic choice together with equations and reductions orienting the axioms of a barycentric algebra [Stone 1949]; as an example of the latter, we will consider the ticking operation  $\surd$  performing cost and improvement analysis [Sands 1998].

<sup>3</sup>Even if quantitative equational systems are one of the main motivations behind the introduction of quantitative rewriting systems, we will not deal with them in this paper (some sections are dedicated to that topic in the extended version of this work [Gavazzo and Di Florio 2022]). Nonetheless, we remark that even if equational in nature, quantitative equational deduction is considerably different from its traditional counterpart, as its deduction rules are inherently quantitative. Transitivity, for instance, describes the triangular inequality axiom of metric spaces: from  $\varepsilon \Vdash t \approx s$  and  $\delta \Vdash s \approx u$  we infer  $\varepsilon + \delta \Vdash t \approx u$ ; whereas reflexivity gives the identity of indiscernibles axiom ( $0 \Vdash t \approx t$ ).

<sup>4</sup>Other possible readings come from the world of metric spaces (*the distance between  $t$  and  $s$  is at most  $\varepsilon$* ), resource analysis (*given resource  $\varepsilon$ , the terms  $t$  and  $s$  can be proved equal*), and fuzzy and graded logic(s) ( *$t$  is equal to  $s$  with degree  $\varepsilon$* ).

$E$  on combinators as  $E(t, s) \triangleq \inf\{\varepsilon \mid \varepsilon \Vdash t = s\}$ , and the defining rules of quantitative equational deduction ensure  $E$  to be a(n equationally defined) pseudometric.

*The Computational Content of a Distance.* Quantitative equational theories model distances equationally and provide a quantitative refinement of traditional equality. This puts ourselves in a situation similar to the one outlined at the very beginning of this section, the only(!) difference being that now equalities have been replaced by pseudometrics. In analogy to the case of traditional equational reasoning and rewriting, it is natural to ask whether (equationally defined) distances also come with quantitative notions of rewriting exploiting their operational content and, if so, whether we have powerful results, such as a quantitative counterpart of the CR theorem, giving interesting computational and algorithmic properties of equational distances.

Quantitative rewriting relations are quantitative relations (as introduced in the next paragraph) describing step-by-step distance-producing *dynamics*, i.e. those step-by-step transformations producing differences. In the case of system  $\mathcal{K}$ , quantitative rewriting relations can be defined taking advantages of its term-structure (a concept discussed in the next section). In fact, by orienting the ternary equations of previous paragraph, we obtain a ternary relation  $\mapsto_K$  relating pairs of combinators with non-negative extended real numbers.<sup>5</sup> Such a ternary relation is then extended to the rewriting relation  $\rightarrow_K$  that propagates distances produced by  $\mapsto_K$  inside arbitrary combinators. This point will be crucial in the next section, but for the moment it is not relevant, and thus we postpone its analysis. What matters is that by orienting quantitative equations we obtain a ternary rewriting relation  $\rightarrow_K$ , exactly in the same way of traditional rewriting. Moreover,  $\rightarrow_K$  induces a quantitative relation (a distance)  $K$  associating to combinators their rewriting distance:  $K(t, s) \triangleq \inf\{\varepsilon \mid \varepsilon \Vdash t \rightarrow_K s\}$ . The rewriting distance  $K$  thus acts (at least at a definitional level) as the operational counterpart of the equational pseudometric  $E$  previously introduced; and indeed it is our first example of a *quantitative abstract rewriting system*.

Summing up, we have obtained quantitative refinements of the equality  $\approx$  and rewriting  $\rightarrow$  relations on combinators defined at the beginning of this section, namely the equational pseudometric  $E$  and the rewriting distance  $K$ , respectively. As in the traditional case the Church-Rosser theorem provides a (fruitful) link between  $\approx$  and  $\rightarrow$ , we now ask whether a similar link exists between  $E$  and  $K$ . That is, whether we have a quantitative CR theorem for system  $\mathcal{K}$ . To answer this question, we first need to understand what a CR theorem states in a quantitative setting, as well as its operational meaning. The first main contribution of this paper, namely a theory of quantitative *abstract* rewriting system precisely gives that information. Next, we have to prove that indeed the quantitative CR theorem holds for  $\mathcal{K}$ . This is highly nontrivial and requires to develop suitable proof techniques both at the level of quantitative abstract rewriting systems (such as quantitative refinements of Hindley-Rosen and Newman's lemmas) and at the level of term-based quantitative rewriting systems, which is the second main contribution of this paper and the topic we shall discuss in the next section. Let us begin with the first point.

*The Art and Craft of Quantitative Relations.* To quantitatively refine CR and understand its operational meaning, we take a top-down approach whereby we first derive the formal statement of quantitative CR in a purely algebraic way, and then analyse its operational meaning. We do so moving from two main observations:

1. First, we notice that traditional (abstract) rewriting systems are relational notions [Doornbos et al. 1997] whose theory can be developed in the framework of abstract relational calculi [Bird and de Moor 1997; Freyd and Scedrov 1990; Schmidt 2011; Tarski 1941]. For instance, the

<sup>5</sup>Actually, our theory of quantitative rewriting is developed in terms of relations enriched in suitable quantales [Hofmann et al. 2014], non-negative real numbers giving just one example of such quantales (see section 3).



Church-Rosser property is described by the identity<sup>6</sup>  $(R \cup R^-)^* = R^*; R^{*-}$ , and a simple relational calculation shows that such an equality is equivalent to  $R^{*-}; R^* \subseteq R^*; R^{*-}$ , which precisely encodes confluence of  $R$ .

2. Secondly, as pointed out by Lawvere [1973], we observe that quantitative relations (or distances) are governed by an algebra close<sup>7</sup> to the one of ordinary relations, so that a large part of the aforementioned calculi of relations can be refined to calculi of quantitative relations. In fact, by viewing binary relations as maps  $R : A \times B \rightarrow \{\perp, \top\}$ , we see that a quantitative relation simply refines the Boolean structure  $(\{\perp, \top\}, \leq, \wedge)$  by replacing it with  $([0, \infty], \geq, +)$  (notice the opposite natural order), so that we can use this similarity to generalise many relational constructions and their properties to a quantitative setting. For instance, by refining the existential quantifier  $\exists$  as the infimum  $\inf$  and the Boolean meet  $\wedge$  as addition  $+$ , we can define the composition between quantitative relations  $R, S$  by  $(R; S)(a, c) \triangleq \inf_b R(a, b) + S(b, c)$ . Consequently, we will say that a quantitative relation  $R$  is transitive if  $R; R \geq R$ , i.e. if  $\inf_b R(a, b) + R(b, c) \geq R(a, c)$ , which is nothing but the usual triangle inequality law. In the same way, we can refine the notions of reflexivity and symmetry to quantitative relations — altogether obtaining exactly the defining axioms of a pseudometric — as well as the notions of reflexive, transitive (and symmetric) closure of a quantitative relation.

Putting these facts together, we shall define quantitative abstract rewriting systems as sets  $A$  endowed with quantitative (endo)relations, i.e. maps  $R : A \times A \rightarrow [0, \infty]$ . By observation (2) above, then, we have an algebra of quantitative relations that we can rely on to study quantitative abstract rewriting systems; and by observation (1), we can use such an algebra to decline several rewriting notions in a quantitative setting. For instance, the *convertibility distance*  $R^\equiv$  induced by a quantitative rewriting relation  $R$  is defined as the reflexive, symmetric, and transitive closure of  $R$ , so that  $R^\equiv$  is indeed a pseudometric, much in the same way as convertibility is an equivalence in traditional rewriting. More interestingly, instantiating the CR equality of point (1) in the quantitative setting — which gives  $R^\equiv = R^*; R^{*-}$  — we obtain the desired notions of quantitative Church-Rosser. Similarly, quantitative confluence is given by  $R^{*-}; R^* \geq R^*; R^{*-}$ , and the abstract calculus of quantitative relations shows that quantitative confluence and CR are indeed equivalent notions.

Now that we have a quantitative CR property, we can read it back to understand its operational meaning. Suppose we want to compute the distance  $E(t, s) = K^\equiv(t, s)$  between two combinators. Since such a distance is obtained by means of bidirectional quantitative rewriting paths, it is desirable to be able to compute, or at least approximate, it as the sum of the rewriting distances into their common reducts:  $R^\equiv(t, s) = \inf_u R^*(t, u) + R^*(s, u)$ , which is nothing but the pointwise reading of quantitative CR property defined above. Consequently, the quantitative CR theorem entails that we can approximate convertibility distances as sums of directional rewriting distances (and obtain an exact characterisation by taking their infimum).<sup>8</sup>

*Back to Curry.* Having the machinery of quantitative abstract rewriting systems at our disposal, we come back to our running example: the quantitative CR theorem for system  $\mathcal{K}$ . By the general theory of quantitative abstract rewriting systems we shall develop in section 4, we know that quantitative CR and quantitative confluence are equivalent, so that it is enough to prove quantitative

<sup>6</sup>Here, we denote by  $R^-$  the converse (or dual) of a binary relation, and by  $R^*$  its reflexive and transitive closure. See section 4 for precise definitions.

<sup>7</sup>Not identical, though.

<sup>8</sup>In the case of term-based systems, convertibility distances are obtained as limits of bidirectional reduction path of the form  $t \xrightarrow{\varepsilon_1} \cdot \xleftarrow{\varepsilon_2} \cdot \xrightarrow{\varepsilon_3} \cdot \cdot \cdot \xleftarrow{\varepsilon_{n-1}} \cdot \xrightarrow{\varepsilon_n} s$ , so that the convertibility distance between  $t$  and  $s$  given by this path is  $\sum_{i=1}^n \varepsilon_i$ . For these systems, we shall prove an even stronger CR result whereby for each such a rewriting path there exists term  $u$  such that  $t \xrightarrow{\delta_1} \cdot \cdot \cdot \xrightarrow{\delta_m} u \xleftarrow{\eta_p} \cdot \cdot \cdot \xleftarrow{\eta_1} s$  and  $\sum_{i=1}^n \varepsilon_i \geq \sum_{j=1}^m \delta_j + \sum_{k=1}^p \eta_k$ .

confluence of the rewriting distance  $K$ . One way to do so is by brute force, meaning that we attempt a confluence proof for the whole system  $\mathcal{K}$  viewed monolithically. Obviously, this is not a far-sighted strategy, especially in light of the further extensions we will consider in this work (e.g. probabilistic nondeterminism). A better way to go is to proceed *modularly*, noticing that we can view  $\mathcal{K}$  as the composition of two systems: the original system of combinators plus a system of arithmetic. It is then desirable to prove confluence of these systems separately and to rely on some compositionality result stating that, under suitable conditions, confluence is preserved by joining systems together. Such a result is nothing but the quantitative refinement of the well-known Hindley-Rosen Lemma, which we shall prove in [section 4](#).

The quantitative Hindley-Rosen Lemma indeed improves modularity of confluence proofs, but we still have to prove confluence of each subsystem of  $\mathcal{K}$ ; and the theory of abstract systems is of little help for that (although it still offers useful results: for instance, our quantitative refinement of Newman's lemma allows us to infer quantitative confluence from quantitative weak confluence for terminating systems). At this point the crucial observation is that all systems considered so far are examples of term-based systems, meaning that their rewriting distances are obtained starting with a ground distance on terms which is then propagated through term constructs. To achieve full modularity of confluence proofs, consequently, it is thus desirable to have *local* confluence criteria, whereby suitable confluence-like properties of ground rewriting distances entail confluence for full rewriting. This directly leads us to the second main contribution of this work, namely the theory of quantitative *term-based* rewriting systems.

## 2.2 Quantitative Term Rewriting Systems: A Short Phenomenology

In this paper, we shall study both quantitative *abstract* and *term* rewriting systems. The latter are systems measuring distances between *first-order* terms with the constraint that rewriting distances must be compatible with term constructs. In the theory of traditional term-rewriting systems there is just one type of such constraints: rewriting relations must be compatible with term constructs, this way ensuring convertibility distances to be congruences. In a quantitative setting, however, there are many types of compatibility constraints, and thus many types of quantitative term rewriting systems. Here, we recap the two main ones we will deal with.

*Non-Expansive Systems.* The quantitative counterpart of compatibility (and thus of congruence) is *non-expansiveness*. Non-expansive term rewriting systems are quantitative systems in which reducing terms inside contexts *non-expansively* propagates distances. Given a signature  $\Sigma$  (i.e. a collection of operation symbols with their arities), a non-expansive term rewriting system  $\mathcal{R}$  is defined by a ternary ground rewriting relation  $\mapsto_R$  on  $\Sigma$ -terms and non-negative extended real numbers, which is then *non-expansively* extended to the (ternary) rewriting relation  $\rightarrow_R$  thus:<sup>9</sup>

$$\frac{\varepsilon \Vdash t \mapsto_R s}{\varepsilon \Vdash C[t^\sigma] \rightarrow_R C[s^\sigma]}$$

Therefore, in a non-expansive system contexts cannot increase distances produced by reducing their subterms. Thinking about the context  $C$  as a function on terms, then  $C$  is non-expansive with respect to the rewriting distance  $R$  induced by  $\rightarrow_R$ . Formally,  $R(t, s) \geq R(C[t^\sigma], C[s^\sigma])$ . Looking back at the previous section, we see that system  $\mathcal{K}$  is an example of a non-expansive system.

*Distance Trivialisation.* Forcing non-expansiveness of contexts ultimately means assuming function symbols in  $\Sigma$  to behave as non-expansive functions at a semantic level (i.e. at the level of convertibility distances and quantitative equational theories). Contrary to traditional term-rewriting

<sup>9</sup>Given a term  $t$  and a substitution  $\sigma$  (i.e. a map from variables to terms), we write  $t^\sigma$  for the application of the substitution  $\sigma$  to  $t$ . We also write  $C[t]$  for the term obtained by plugging-in the term  $t$  into the context  $C$ .

systems, where forcing compatibility gives a well-behaved theory of first-order rewriting, in a quantitative setting the non-expansive assumption is not innocent and, in absence of further syntactic constraints, it makes the theory of non-expansive systems inconsistent. The syntactic constraints we are referring to ultimately amounts to *linearity* of rewriting rules, meaning that each variable in the defining rules of  $\mapsto_R$  appears (at most) once on each side of the rule.

We will see in [subsection 5.2](#) that the absence of linearity makes modularity as advocated at the end of the previous section simply not available (invalidating, for instance, cornerstone results such as orthogonality and critical pairs theorems). For the moment, we just recall that linearity assumptions in the context of metric reasoning are well-known in programming language semantics [[Crubillé and Dal Lago 2014, 2017](#); [Gavazzo 2018](#)], where it has been shown that, in absence of linearity, non-expansiveness leads to the so-called distance trivialisation [[Crubillé and Dal Lago 2017](#)]. Rephrased in our language, this means that the convertibility distance collapses on an equivalence relation, and thus quantitative reasoning trivialises.

To see that, it is enough to look at our running example. First, we notice that the rewriting rule  $0 \Vdash Wx y \rightarrow x y y$  is non-linear. Then, we observe that: (i) the convertibility distance  $K^\equiv$  is a combinatory version of the Euclidean distance; (ii) by non-expansiveness, the convertibility distance  $\varepsilon$  between  $Z$  and  $SZ$  must be bigger or equal than the one between  $WAZ$  and  $WA(SZ)$ . But the latter distance is the distance between  $Z$  and  $SSZ$ , which is  $1 + \varepsilon$ , since  $1 \Vdash Sx \rightarrow x$ . We thus conclude the desired thesis.

*Failure of Confluence.* The very same example witnessing distance trivialisation, additionally, shows the failure of quantitative confluence for system  $\mathcal{K}$ . In fact, we have the peak

$$Z \xleftarrow{0} AZZ \xleftarrow{0} WAZ \xleftarrow{1} WA(SZ) \xrightarrow{0} A(SZ)(SZ)$$

and since  $Z$  cannot be reduced any further, quantitative confluence should give a reduction from  $A(SZ)(SZ)$  to  $Z$  with distance at most 1. However, it does not take much to realise that all such reductions produce distance 2, hence making quantitative confluence failing.

Having observed that non-linearity leads to distance trivialisation, as well as to the failure of confluence and its term-based proof techniques ([subsection 5.2](#)), it is thus natural to ask if non-linearity is indeed the source of all problems. We shall answer this question in the affirmative by showing that under the linearity assumption many important confluence results, such as orthogonality and critical pairs theorems, hold for non-expansive systems. Consequently, we can conclude that non-expansive systems are an adequate formalism to provide quantitative analyses of rewriting in presence of linearity (although they do not scale to non-linearity).

*Towards Church-Rosser.* Even if linearity is a restrictive condition, the class of linear non-expansive systems covers many interesting examples. For instance, several quantitative computational effects can be given as quantitative linear systems, as well as linear programming languages. Coming back to our running example, even if there is no hope to prove confluence, and thus CR, of system  $\mathcal{K}$ , we can still consider its linear subsystem obtained by dropping the non-linear combinator  $W$ . We shall prove the resulting system to be indeed confluent (and thus CR), and we will do so in a highly modular fashion relying on the aforementioned quantitative refinements of the Hindley-Rosen Lemma and Newman's Lemma as well as of critical-pair and orthogonality theorems.

*Graded Systems.* Notwithstanding the interesting examples of linear non-expansive systems, asking linearity of rewriting rules may be a too strong condition, our running example (system  $\mathcal{K}$ ) being a prime witness of that. To go beyond linearity and non-expansiveness, in [section 6](#) we introduce the theory of *graded systems*. Graded systems constitute the largest class of term-based quantitative rewriting systems we consider. Contrary to non-expansive systems, in a graded system



the distance generated by a reduction  $\varepsilon \Vdash t \rightarrow s$  can be amplified when performed inside a context, although in a controlled way. Semantically, such a controlled way generalises the well-known notion of a Lipschitz constant [Searcoid 2006]. This allows us to model non-linear systems avoiding, at the same time, distance trivialisation and other undesired behaviours. In a graded system, in fact, the quantitative rewriting relation  $\rightarrow_R$  is defined by the rule

$$\frac{\varepsilon \Vdash t \mapsto_R s}{\partial_C(\varepsilon) \Vdash C[t^\sigma] \rightarrow_R C[s^\sigma]}$$

where  $\partial_C : [0, \infty] \rightarrow [0, \infty]$  describes how much  $C$  amplifies rewriting distances. Consequently, the non-expansive inequality  $R(t, s) \geq R(C[t^\sigma], C[s^\sigma])$ , characterising non-expansive systems is replaced by the more liberal one  $\partial_C(R(t, s)) \geq R(C[t^\sigma], C[s^\sigma])$ .

The map  $\partial_C$  — called the *sensitivity* or *degree* of  $C$  — acts as a generalised Lipschitz constant; and indeed, multiplication by a constant is a typical example of a context degree. To compute the sensitivity of a context, in graded systems we built terms relying on *modal signatures* [Dagnino and Pasquali 2022], whereby  $n$ -ary function symbols  $f$  come equipped with *modal arities*  $(\phi_1, \dots, \phi_n)$  specifying that  $f$  has sensitivity  $\phi_i$  on its  $i$ th argument. Since terms are built using variables and function symbols, modal arities allow us to compute the sensitivity of any context, and thus to define  $\rightarrow_R$  as previously described. All of that makes graded term rewriting systems inherently *modal* and *coeffectful* [Ghica and Smith 2014; Orchard et al. 2019; Petricek et al. 2014] and, at the same time, makes them a powerful tool to define the operational semantics of modal and coeffectful languages [Abel and Bernardy 2020; Choudhury et al. 2021]. For example, we will prove general quantitative critical-pair and orthogonality theorems for graded systems that gives us powerful proof techniques to prove confluence (and thus CR) of graded systems. We will then apply such techniques (the orthogonality theorem, in particular) to infer quantitative confluence and CR of a system of graded combinators [Abramsky 2002; Abramsky et al. 2002; Atkey 2018], which is nothing but the graded counterpart of system  $\mathcal{K}$ . More specifically, we shall extend the linearisation of system  $\mathcal{K}$  with a graded operation symbol  $!_n$  whose sensitivity is given by the multiplication by  $n$  function. Consequently, whenever  $\varepsilon \Vdash t \mapsto_R s$ , we also have  $n\varepsilon \Vdash !_n t \mapsto_R !_n s$ . This will allow us to reintroduce non-linear combinators, but in a controlled way: for instance, the duplication combinator  $W$  will now be given as a graded family of combinators  $W_{n,m}$  together with the quantitative rewriting rule  $0 \Vdash W_{n,m} x \mapsto_{n+m} y \mapsto x \mapsto_{!_n} y \mapsto_{!_m} y$ . This way, we avoid distance trivialisation and obtain quantitative confluence even in presence of non-linearity. Thinking about the counterexample to confluence seen in previous paragraph, we see that the source of divergence, namely the reduction  $1 \Vdash WA(SZ) \rightarrow WAZ$ , is refined as  $2 \Vdash W_{1,1} A !_2(SZ) \rightarrow W_{1,1} A !_2 Z$ , so that the rewriting distance obtained by reducing  $SZ$  to  $Z$  (i.e.  $1 \Vdash SZ \rightarrow Z$ ) is duplicated, this way ensuring quantitative confluence.<sup>10</sup>

### 2.3 Closing the Circle

The theory of graded systems closes the circle opened by our running example at the beginning of this section. We based this introductory narrative on the problem of proving a quantitative Church-Rosser theorem for (quantitative extensions) of Curry's system of combinators. After having introduced quantitative abstract rewriting systems as a foundational theory inside which we can give precise definitions of, and prove general results about, quantitative rewriting (e.g. definitions of quantitative Church-Rosser, confluence, and termination; proofs of quantitative Newman's and Hindley-Rosen lemmas), we have approached possible proofs of quantitative confluence/CR for system  $\mathcal{K}$ , focusing on modularity. First, we have noticed that  $\mathcal{K}$  can be seen as composed of several

<sup>10</sup>Notice that, to be formal, we also have to specify how combinators and  $!_n$  interact. For instance, the graded system of section 6 has a combinator  $D$  (dereliction) described by the reduction  $0 \Vdash D !_1 x \mapsto x$ .

subsystems (so that we can rely on the quantitative Hindely-Rosen Lemma to prove confluence of each subsystem in isolation); and then that confluence of all such subsystems could be obtained *locally* by taking advantage of their term structure.

Doing so requires the development of a theory of quantitative term rewriting systems which, however, turns out to be considerably more involved than its traditional counterpart. In fact, we noticed that quantitative extensions of Curry's systems being non-expansive, they fail to satisfy quantitative confluence, due to the presence of non-linear rewriting rules. One way to overcome the issue is to linearise them by dropping all non-linear rules. This way, we indeed obtain the desired quantitative CR theorem, but for linear (sub)systems only. Another, more satisfactory, solution is to refine system  $\mathcal{K}$  to a graded system of combinators. This way, we can account for non-linear combinators and, at the same time, rely on the quantitative rewriting theory of graded systems that allow us to obtain the desired quantitative confluence and Church-Rosser theorem in a completely modular fashion.

*What's Next?* Now that the reader has familiarised herself with quantitative rewriting systems, we move to the technical development of their theory, proceeding in three steps. We will first define quantitative *abstract* rewriting systems, which provide a relational foundation of quantitative rewriting. We shall then introduce *non-expansive* quantitative term rewriting systems and, finally, move to *graded* systems. In doing so, we will review the examples seen in this section in a formal and precise way.

### 3 PRELIMINARIES: QUANTITATIVE RELATIONAL CALCULUS, À LA LAWVERE

Before introducing quantitative abstract rewriting systems, we recall some necessary mathematical preliminaries.

*Quantales.* Traditional abstract rewriting systems can be naturally defined and studied *relationally*. To define a theory of quantitative rewriting, it thus seems natural to rely on *quantitative relational calculi*. Here, we follow the analysis of generalised metric spaces as enriched categories by Lawvere [1973] and work with relations taking values (i.e. enriched) in a quantale [Hofmann et al. 2014; Rosenthal 1990].

**Definition 1.** A (unital) quantale  $\Omega = (\Omega, \leq, \otimes, k)$  consists of a monoid  $(\Omega, k, \otimes)$  and a sup-semilattice  $(\Omega, \leq)$  satisfying the following distributivity laws:

$$\delta \otimes \bigvee_{i \in I} \varepsilon_i = \bigvee_{i \in I} (\delta \otimes \varepsilon_i) \qquad \left( \bigvee_{i \in I} \varepsilon_i \right) \otimes \delta = \bigvee_{i \in I} (\varepsilon_i \otimes \delta).$$

The element  $k$  is called unit of the quantale, whereas  $\otimes$  is called its tensor (or multiplication).

It is easy to see that  $\otimes$  is monotone in both arguments. We denote the top and bottom element of a quantale by  $\top$ ,  $\perp$ , respectively. Quantales having unit  $k$  coinciding with the top element are called *integral* quantales. Moreover, we say that a quantale is commutative if its underlying monoid is, and that it is non-trivial if  $k \neq \perp$ . Integral quantales are particularly well-behaved: for instance, in an integral quantale we have  $\varepsilon \otimes \perp = \perp$ , for any  $\varepsilon \in \Omega$ .<sup>11</sup> If the opposite direction holds, i.e. whenever  $\varepsilon \otimes \delta = \perp$ , either  $\varepsilon = \perp$  or  $\delta = \perp$  hold, we say that the quantale is *cointegral*. From now on, we assume quantales to be commutative, (co)integral, and non-trivial. We refer to such quantales as *Lawvereian*. Among such conditions, cointegrality is definitely the strongest one. Such a condition is needed in subsection 4.2 only (where it is used to prove the equivalence between quantitative termination and induction). In all the remaining parts of the paper, cointegrality is not needed, so that the reader can freely drop such an assumption. Finally, we say that a quantale is

<sup>11</sup>For  $\varepsilon \otimes \perp \leq \top \otimes \perp = k \otimes \perp = \perp$ .

*idempotent* if  $\varepsilon \otimes \varepsilon = \varepsilon$ . Notice that any quantale  $(\Omega, \leq, k, \otimes)$  induces an idempotent quantale as  $(\Omega, \leq, \top, \wedge)$  and that in any integral idempotent quantale  $\wedge$  and  $\otimes$  coincide.

- Example 1.** 1. The *boolean quantale*  $\mathbb{B} = (2, \leq, \wedge, \top)$ , where  $2 = \{\top, \perp\}$  and  $\perp \leq \top$ , is an idempotent Lawverian quantale.
2. The *Lawvere quantale*  $\mathbb{L} = ([0, \infty], \geq, +, 0)$  consisting of the extended real half-line ordered by the “greater or equal” relation  $\geq$  and with extended addition as tensor product is a Lawverian quantale. Notice that we use the opposite natural ordering, so that, e.g., 0 is the top element of  $\mathbb{L}$ .
3. The *Strong Lawvere quantale*  $\mathbb{L}^{\max} = ([0, \infty], \geq, \max, 0)$  obtained by replacing addition with maximum in the Lawvere quantale is an idempotent Lawverian quantale.
4. The unit interval  $\mathbb{I} = ([0, 1], \leq, *)$  endowed with a left continuous *triangular norm* [Hájek 1998] (*t-norm* for short)  $*$  is an integral quantale. If, additionally,  $x * y = 0$  implies  $x = 0$  or  $y = 0$ , then we obtain a Lawverian quantale. Examples of such *t-norms* are the *product t-norm* ( $x *_p y \triangleq x \cdot y$ ) and the *Gödel t-norm* ( $x *_g y \triangleq \min\{x, y\}$ ). An example of a *t-norm* breaking cointegrality is the Łukasiewicz *t-norm*  $x *_l y \triangleq \max\{0, x + y - 1\}$ . Due to their applications in Fuzzy reasoning, we refer to  $\mathbb{I} = ([0, 1], \leq, *, 1)$  as the Fuzzy quantale(s).
5. The set of *monotone modal predicates*  $2^W$  on a preordered monoid (of possible worlds) with top element  $(W, \leq, +, 0, \top)$ , endowed with the tensor product  $(p \otimes q)(w)$  defined by  $\exists u, v. w \geq u + v \wedge p(u) \wedge q(v)$  is a Lawverian quantale. Such a quantale is used to study modal and coeffectful properties of programs [Dal Lago and Gavazzo 2021b, 2022b]. Replacing 2 with a (Lawverian) fuzzy quantale  $([0, 1], \leq, *, 1)$ , we obtain the (Lawverian) quantale of Fuzzy modal predicates.

Since any quantale is, in particular, a complete lattice and tensor product is monotone in both arguments, the latter has both left and right adjoints (which coincide in our case, as we assume  $\otimes$  to be commutative), which we denote by  $\multimap$ . Explicitly, we define  $\varepsilon \multimap \delta \triangleq \bigvee \{\eta \mid \varepsilon \otimes \eta \leq \delta\}$ . For instance, in the Boolean quantale  $\multimap$  is ordinary implication, whereas in the Lawvere quantale it is truncated subtraction.

*Quantale-valued Relations.* We now move to quantale-valued relations, our main tool to model quantitative rewriting. As quantales model abstract quantities, quantale-valued relations provide abstract notions of distances.

**Definition 2.** Given a quantale  $\Omega = (\Omega, \leq, k, \otimes)$ , a  $\Omega$ -relation  $R : A \multimap B$  between sets  $A$  and  $B$  is a function  $R : A \times B \rightarrow \Omega$ . For any set  $A$ , we define the identity (or diagonal)  $\Omega$ -relation  $\Delta_A : A \multimap A$  mapping diagonal elements  $(a, a)$  to  $k$ , and all other elements to  $\perp$ . Moreover, the composition  $R; S : A \multimap C$  of  $\Omega$ -relations  $R : A \multimap B$  and  $S : B \multimap C$  is defined by the so-called matrix multiplication formula [Hofmann et al. 2014]:

$$(R; S)(a, c) \triangleq \bigvee_{b \in B} R(a, b) \otimes S(b, c).$$

In general, we think about a  $\Omega$ -relation as giving the distance or the degree of relatedness of two elements [Flagg 1992; Hofmann et al. 2014]. For instance, when the quantale is Boolean, elements are either related or not, whereas for Fuzzy quantales  $\Omega$ -relations coincide with Fuzzy relations [Belohlávek 2002], and thus they give the degree to which elements are related, as well as proximity and similarity relations. When we move to the Lawvere quantale (and quantales alike),  $\Omega$ -relations give general notions of distances [Lawvere 1973], and thus act as a foundation for metric reasoning [Bonsangue et al. 1998]. Finally, considering the quantale of (fuzzy) modal predicates, we obtain (fuzzy) modal and coeffectful relations [Dal Lago and Gavazzo 2022b; Routley and Meyer 1972a,b, 1973; Urquhart 1972], whereby (the degree of) relatedness of elements is given with respect a possible world (such as the available resources).

Since  $\Omega$ -relation composition is associative and has  $\Delta$  as unit element, for any quantale  $\Omega$  we have a category, denoted by  $\Omega\text{-Rel}$ , with sets as objects and  $\Omega$ -relations as arrows. Moreover, the complete lattice structure of  $\Omega$  lifts to  $\Omega$ -relations pointwise, so that we can say that a  $\Omega$ -relation  $R : A \rightarrow B$  is *reflexive* if  $\Delta \leq R$ ; *transitive* if  $R; R \leq R$ ; and *symmetric* if  $R^- \leq R$ , where the transpose of  $R : A \rightarrow B$  is the  $\Omega$ -relation  $R^- : B \rightarrow A$  defined by  $R^-(b, a) \triangleq R(a, b)$ . When read pointwise, reflexivity, transitivity, and symmetry give the following inequalities, respectively:

$$k \leq R(a, a) \quad R(a, b) \otimes R(b, c) \leq R(a, c) \quad R(a, b) \leq R(b, a).$$

Altogether, we obtain the notion of a preorder (i.e. reflexive and transitive) and equivalence (i.e. reflexive, transitive, and symmetric)  $\Omega$ -relation.

**Notation 1.** Fixing a quantale  $\Theta$ , we oftentimes refer to  $\Omega$ -relation on  $\Theta$  as  $\Theta$ -relations. Thus, for example,  $\mathbb{L}$ -relations are just  $\Omega$ -relations on the Lawvere quantale  $\mathbb{L}$ .

- Example 2.** 1. On the Boolean quantale,  $\mathbb{B}$ -relations are ordinary (binary) relations, and preorder and equivalence  $\mathbb{B}$ -relations coincide with traditional preorders and equivalences.
2. On the Lawvere quantale,  $\mathbb{L}$ -relations are distances. Instantiating transitivity on  $\mathbb{L}$ , we obtain the usual *triangle inequality* formula:  $\inf_b R(a, b) + R(b, c) \geq R(a, c)$ . Similarly, reflexivity gives the identity of indiscernibles inequality:  $0 \geq R(a, a)$ . Altogether, we see that preorder  $\mathbb{L}$ -relations coincide with *generalised metrics* [Bonsangue et al. 1998; Lawvere 1973] and equivalence  $\mathbb{L}$ -relations with *pseudometrics* [Steen and Seebach 1995].
3. On the Strong Lawvere quantale  $\mathbb{L}^{\max}$ , transitivity gives the *strong triangle inequality* formula:  $\inf_b \max(R(a, b), R(b, c)) \geq R(a, c)$ . Equivalence  $\mathbb{L}^{\max}$ -relations coincide with *ultra-pseudometrics*.
4. On the unit interval (fuzzy) quantale(s),  $\mathbb{I}$ -relations coincide with fuzzy relations [Belohlávek 2002]. Equivalence  $\mathbb{I}$ -relations are often called similarity or proximity relations.

We notice that the “algebra” of  $\Omega$ -relations is close to the one of ordinary relations,<sup>12</sup> so that we can refine a large part of calculi of relations [Schmidt 2011] to a quantale-based setting. Since many notions of traditional rewriting can be given in purely relational terms, we can take advantage of that and rephrase them in terms of  $\Omega$ -relations.

To do so, it is useful to exploit fixed point characterisations of relational constructions, as well as their adjunction properties [Backhouse 2000; Bird and de Moor 1997]. Recall that  $\Omega\text{-Rel}(A, B)$  carries a complete lattice structure, so that any monotone map  $F : \Omega\text{-Rel}(A, B) \rightarrow \Omega\text{-Rel}(A, B)$  has least and greatest fixed points, denoted by  $\mu X.F(X)$  and  $\nu X.F(X)$ , respectively. We denote by  $\perp$  the  $\Omega$ -relation  $\mu X.X$  assigning distance  $\perp$  to all elements, and by  $\top$  the  $\Omega$ -relation  $\nu X.X$ , i.e. the indiscrete  $\Omega$ -relation assigning distance  $k$  to all elements. Moreover, given a  $\Omega$ -relation  $R : A \rightarrow A$ , we define the following  $\Omega$ -relations: (i) the *reflexive closure* of  $R$  as  $R^+ \triangleq R \vee \Delta$ ; (ii) the *transitive and reflexive closure* of  $R$  as  $R^* \triangleq \mu X.\Delta \vee R; X$ ; (iii) the *equivalence closure* of  $R$  as  $R^\equiv \triangleq (R \vee R^-)^*$ . As usual, we have  $R^* = \bigvee_{n \geq 0} R^n$ , where  $R^0 \triangleq \Delta$  and  $R^{n+1} \triangleq R; R^n$ .

Finally, since  $\Omega\text{-Rel}(A, A)$  is not only a complete lattice, but a quantale,  $\Omega$ -relation composition has both left and right adjoints, often referred to as left and right division [Bird and de Moor 1997]:

$$R; S \leq P \iff S \leq R \setminus P \quad R; S \leq P \iff R \leq P / S.$$

Notice that  $R; (R \setminus S) \leq S$  and  $(R / S); S \leq R$ .

**Ternary Relations.** Even if we model quantitative rewriting relations as  $\Omega$ -relations, in section 2 we have defined rewriting systems by means of suitable ternary relations, from which we have then extracted a  $\Omega$ -relations. This process, known as *strata extension* [Hofmann et al. 2014], is an instance of a more general correspondence [Dal Lago and Gavazzo 2021b] between abstract

<sup>12</sup>As the category of traditional relations, the category  $\Omega\text{-Rel}$  is a *quantaloid* [Hofmann et al. 2014; Stubbe 2014].

distances and suitable ternary relations akin to substructural Kripke relations [Routley and Meyer 1972a,b, 1973; Urquhart 1972]. Since we will extensively switch between  $\Omega$ -relations and ternary relations, we recall the notion of a  $\Omega$ -ternary relation.

**Definition 3.** Given a quantale  $\Omega$ , a  $\Omega$ -ternary relation over  $A \times B$  is a ternary relation  $R \subseteq A \times \Omega \times B$  antitone in its second argument (meaning that  $R(a, \varepsilon, b)$  implies  $R(a, \delta, b)$ , for any  $\delta \leq \varepsilon$ ).

Any ternary  $\Omega$ -relation  $R$  induces a  $\Omega$ -relation  $R^\bullet(a, b) \triangleq \bigvee_{R(a, \varepsilon, b)} \varepsilon$ . Vice versa, any  $\Omega$ -relation  $R$  induces a  $\Omega$ -ternary relation  $R^\circ$  defined by  $R^\circ(a, \varepsilon, b) \iff \varepsilon \leq R(a, b)$ . Moreover, these two processes are each other inverses, meaning that  $R^{\circ\bullet} = R$  and  $R^{\bullet\circ} = R$ , so that we can freely switch between  $\Omega$ -ternary relations and  $\Omega$ -relations.

**Notation 2.** Oftentimes, we will use  $\Omega$ -ternary relations to define rewriting systems. In those cases, we will use notations of the form  $\rightarrow_R$  and write  $\varepsilon \Vdash a \rightarrow_R b$  in place of  $\rightarrow_R(a, \varepsilon, b)$ . Moreover, we shall denote by  $R$  the  $\Omega$ -relation associated to  $\rightarrow_R$ . That is,  $R(a, b) \triangleq \bigvee_{\varepsilon \Vdash a \rightarrow_R b} \varepsilon$ .

## 4 QUANTITATIVE ABSTRACT REWRITING SYSTEMS

Now that the reader has the necessary mathematical background, we introduce *quantitative abstract rewriting systems* and their theory. Throughout this and later sections, let  $\Omega = (\Omega, \leq, k, \otimes)$  be a fixed (Lawverian) quantale.

**Definition 4.** A *Quantitative Abstract Rewriting Systems* ( $\Omega$ -ARS, for short) is a pair  $(A, R : A \rightarrow A)$ .

**Definition 4** is extremely simple: as a traditional abstract rewriting system is a set of objects together with a binary (rewriting) relation on it, a  $\Omega$ -ARS is a set of objects together with a binary (rewriting)  $\Omega$ -relation on it. Given elements  $a, b \in A$ , we say that  $a$  rewrites into (or reduces to)  $b$  if  $R(a, b) \neq \perp$ : in that case, we say that the (rewriting) distance between  $a$  and  $b$  is  $R(a, b)$ . Further possible informal reading (possibly depending on the quantale considered) refer to  $R(a, b)$  as the *degree* of the reduction, the *cost* of the reduction, or as the *resource* required for the reduction.<sup>13</sup> Rewriting paths are obtained by iterating  $R$ . In particular, we say that a finite sequence  $(a_0, \dots, a_n)$  is a  $R$ -reduction sequence if  $\bigotimes_i R(a_i, a_{i+1}) \neq \perp$  and that an infinite sequence  $(a_0, \dots, a_n, \dots)$  is a  $R$ -reduction sequence if  $R(a_0, a_1) \otimes \dots \otimes R(a_{n-1}, a_n) \neq \perp$ , for any  $n \geq 0$ . Notice that since  $\Omega$  is Lawverian, for any reduction sequence  $(a_0, \dots, a_n)$ , we have  $R(a_i, a_{i+1}) \neq \perp$ , for any  $i$ .

Finally, notice that given a  $\Omega$ -ARS  $(A, R : A \rightarrow A)$ , the convertibility  $\Omega$ -relation  $R^\equiv$  generated by  $R$  is a  $\Omega$ -equivalence and thus endows  $A$  with a metric-like structure.

### 4.1 Confluence

Given a set  $A$  of objects together with an equivalence  $\equiv$  on it, traditional rewriting systems exploit the computational content of  $\equiv$ . Accordingly, one considers a rewriting relation  $\rightarrow \subseteq A \times A$  on  $A$  such that  $\rightarrow$ -convertibility coincides with  $\equiv$ . At this point, properties of  $\rightarrow$  are proved so to ensure  $\equiv$  to be computationally well-behaved. Among those, the so-called Church-Rosser property states that whenever  $a \equiv b$ , there exists an object  $c$  such that both  $a$  and  $b$  can be reduced to  $c$  in a finite number of steps. Formally,  $\equiv$  coincides with  $\rightarrow^*$ ;  $^* \leftarrow$  (where  $\leftarrow$  stands for  $\rightarrow^-$ ), so that to study  $\equiv$  it is enough to study directional rewriting.

In a quantitative setting, the relation  $\equiv$  is replaced by a  $\Omega$ -equivalence  $E$ , and the rewriting relation  $\rightarrow$  is replaced by a  $\Omega$ -rewriting relation  $R$  such that  $R^\equiv = E$ . One then looks at properties of  $R$  ensuring  $E$  to be computationally well-behaved. In this section, we explore some of these properties, viz. (*quantitative*) *confluence*, *Church-Rosser*, and *termination*. We begin with confluence,

<sup>13</sup>This is the case, in particular, for quantales of modal predicates, where rewriting is ultimately performed in a possible world describing intensional aspects of the rewriting process, such as the available resource.



which states that two reductions originating from the same element can be joined into a common element, as in the classical case, but with the additional property that the merging reduction is achieved without increasing distances.

**Definition 5.** Let  $R, S : A \rightarrow A$  be  $\Omega$ -relations. We say that  $R$  *commutes* with  $S$  if  $R^-; S \leq S; R^-$ , and that  $R$  satisfies the *diamond property* if  $R$  commutes with itself, i.e.  $R^-; R \leq R; R^-$ .

Let us comment on Definition 5 by analysing the diamond property. On the Boolean quantale, we recover the usual diamond property as defined for traditional rewriting systems. Indeed, in pointwise notation, we obtain  $\exists c. R(c, a) \wedge R(c, b) \rightarrow \exists d. R(a, d) \wedge R(b, d)$ , which is equivalent to  $\forall c. (R(c, a) \wedge R(c, b) \rightarrow \exists d. R(a, d) \wedge R(b, d))$ , i.e. the usual diamond property.

More interesting is the case of the Lawvere quantale, which we use as vehicle to move to the general case. Pointwise, the diamond property reads as  $\inf_c R(c, a) + R(c, b) \geq \inf_d R(a, d) + R(b, d)$ . The left-hand-side of the inequality gives the minimal *peak distance* between  $a$  and  $b$ , that is the shortest connection between  $a$  and  $b$  obtained through a pick  $c$  reducing to both  $a$  and  $b$ . The right-hand-side, instead, gives the minimal *valley distance* between  $a$  and  $b$ . Let us say that  $c$  is a peak over  $a$  and  $b$  if  $R(c, a) + R(c, b) \neq \infty$ , so that none of  $R(c, a)$  and  $R(c, b)$  is  $\infty$ . The diamond property ensures that whenever we have a peak  $c$  over  $a$  and  $b$ , then we also have a collection of valleys under  $a$  and  $b$ , i.e. elements  $d$  such that  $R(a, d) + R(b, d) \neq \infty$ , such that the infimum of such valleys is smaller or equal than  $R(c, a) + R(c, b)$ . Notice, however, that there is no guarantee that there is an actual valley  $d$  such that  $R(c, a) + R(c, b) \geq R(a, d) + R(b, d)$ .

**Example 3.** Consider the Lawvere quantale and the  $\mathbb{L}$ -ARS over the set  $A \triangleq \mathbb{R}^+ \cup \{a, b_1, b_2\}$  (notice that  $0 \notin \mathbb{R}^+$ ) with  $R(a, b_1) \triangleq R(a, b_2) \triangleq 0$ ,  $R(b_1, \varepsilon) \triangleq R(b_2, \varepsilon) \triangleq \frac{\varepsilon}{2}$ , for each  $\varepsilon \in \mathbb{R}^+$ , and  $R(x, y) \triangleq \infty$  otherwise. Then, there is no  $c$  such that  $R(b_1, c) + R(b_2, c) = 0$ , although  $\inf_{\varepsilon} R(b_1, \varepsilon) + R(b_2, \varepsilon) = \inf_{\varepsilon > 0} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = 0$ .

In the general setting of an arbitrary quantale  $\Omega$ , we see that the diamond property has the following pointwise reading:

$$\bigvee_c R(c, a) \otimes R(c, b) \leq \bigvee_d R(a, d) \otimes R(b, d).$$

The abstract formulation suggests further non-distance-based readings of the diamond property (and properties alike); and among those, noticeable ones are obtained in terms of *enriched properties* and *degrees of reductions*. Accordingly, we read  $\bigvee_c R(c, a) \otimes R(c, b)$  as the *divergence degree* of  $a$  and  $b$ , and  $\bigvee_d R(a, d) \otimes R(b, d)$  as the *convergence degree* of  $a$  and  $b$ . The diamond property then states that the divergence degree between any two elements is always smaller or equal than their convergence degree; that is, the system tends more to converge than to diverge. Instantiating  $\Omega$  with the Boolean element (and thus recovering the traditional diamond property and properties alike), we stipulate degrees of convergence and divergence to be absolute values. On the other hand, taking the unit interval quantale, we let convergence and divergence be *fuzzy* notions.

**Remark 1.** We have seen that both confluence and the diamond property involve *enriched properties*, i.e. non-Boolean properties taking values in a quantale. Nonetheless, both confluence and the diamond property are *Boolean* properties of  $\Omega$ -relations, as they are essentially of the form  $R \leq S$ . It is natural to push the quantitative perspective one step further and consider an *enriched* version of, e.g., commutation. For instance, we see that requiring  $R$  to commute with  $S$ , i.e.  $R^-; S \leq S; R^-$ , means requiring  $\Delta \leq R^- \setminus (S; R^-) / S$  and thus we may forget about  $\Delta$  and think about the  $\Omega$ -relation  $R^- \setminus (S; R^-) / S$  as assigning to elements the degree of commutation of  $R$  and  $S$  on them, i.e. their

divergence-convergence distance. Pointwise, we thus obtain:

$$(R^- \setminus (S; R^-) / S)(a, b) = \bigvee_c R(c, a) \otimes S(c, b) \multimap \bigvee_d S(a, d) \otimes R(b, d).$$

Notice that the latter is an element of  $\Omega$  rather than a (Boolean) truth value, and thus it indicates *how much*  $R$  commutes with  $S$ . For instance, on the Lawvere quantale,  $(R^- \setminus (R; R^-) / R)(a, b)$  gives the difference between the divergence and convergence distance on  $a$  and  $b$ , and thus a measure of *how much*  $R$  has the diamond property on  $a$  and  $b$ . We leave the exploration of this further form of enrichment for future investigation.

As for the traditional case, we are interested in rewriting paths rather than in single rewriting steps. Given a  $\Omega$ -ARS  $(A, R)$ , we say that: (i)  $R$  is *confluent* if  $R^*$  has the diamond property; (ii)  $R$  is *locally confluent* if  $R^-; R \leq R^*; R^{*-}$ ; (iii) and that  $R$  is *Church-Rosser* (CR, for short) if  $R^\equiv = R^*; R^{*-}$ .

Our quantitative Church-Rosser theorem states that if a rewriting  $\Omega$ -relation is confluent, then we can characterise the convertibility distance  $R^\equiv$  in terms convergent sequences of rewriting steps.

**Proposition 1** (Quantitative Church-Rosser). *Let  $(A, R)$  be a  $\Omega$ -ARS. Then  $R$  is confluent if and only if it is CR.*

PROOF SKETCH. If  $R$  is CR, then it is confluent. We prove the vice versa by least fixed point induction on  $R^{*-}$  (notice that  $R^{*-} = R^{*-}$ ).  $\square$

**Notation 3.** Given a  $\Omega$ -ARS  $\mathcal{R} = (A, R)$  and a property  $\varphi$  on  $\Omega$ -relations, such as being confluent, we say that  $\mathcal{R}$  has property  $\varphi$  if  $R$  has  $\varphi$ . Thus, for instance, we say that  $\mathcal{R}$  is confluent if  $R$  is.

Thanks to [Proposition 1](#), we see that confluence is a crucial property in quantitative and metric reasoning. Proving confluence of quantitative systems, however, can be cumbersome: indeed, quantitative systems are often built *compositionally* by joining systems together. Hindley-Rosen Lemma [[Hindley 1964](#); [Rosen 1970](#)] is arguably one of the most well-known techniques to prove confluence of composed systems modularly. We generalise such a result to  $\Omega$ -ARSs.

**Proposition 2** (Quantitative Hindley-Rosen). *If  $R^*$  commutes with  $S^*$  and  $R, S$  are confluent, then  $R \vee S$  is confluent.*

PROOF SKETCH. Since  $(R^* \vee S^*)^* = (R \vee S)^*$ , it is enough to show that  $R \vee S$  is confluent, which amounts to prove  $(R \vee S)^{-*} \leq ((R \vee S)^*; (R \vee S)^{-}) / (R \vee S)^*$ . We do that by (least) fixed point induction noticing that, for all  $\Omega$ -relations  $P, Q, T$  if  $P; Q \leq Q; P$  and  $P; T \leq T; P$ , then  $P; (Q \vee T)^* \leq (Q \vee T)^*; P$ .  $\square$

## 4.2 Locality and Termination

By [Proposition 1](#), we know that nice operational properties of a  $\Omega$ -equivalence  $E$  can be obtained by characterising  $E$  as the convertibility  $\Omega$ -relation of a *confluent* (rewriting)  $\Omega$ -relation  $R$ . Even if [Proposition 2](#) gives a technique to prove confluence of composed systems compositionally, proving confluence of “atomic” systems may still be difficult. In fact, by its very definition, confluence is a *global* property of a system, in the sense that it refers to rewriting sequences, rather than to single rewriting steps. Newman’s Lemma [[Newman 1942](#)] is a celebrated result in the theory of abstract rewriting stating that if a system is *terminating*, then confluence follows from *local confluence*. We now refine Newman’s Lemma to  $\Omega$ -ARSs. To do so, we first define the notion of a terminating  $\Omega$ -relation and prove that terminating  $\Omega$ -relations satisfy a suitable induction principle. We then use the latter to extend Newman’s Lemma to  $\Omega$ -relations.

Due to space constraints, here we depart from the algebraic style followed so far and give a *pointwise* analysis of quantitative termination together with a *pointwise* proof of (the quantitative

refinement of) Newman's Lemma. Nonetheless, we remark that all these results can be given algebraically by extending the relational theory of induction and abstract rewriting by [Doornbos et al. \[1997\]](#) to a quantitative,  $\Omega$ -enriched setting. Such an extension, which can be found in the extended version of this paper [[Gavazzo and Di Florio 2022](#)], is nontrivial and thus it is incompatible with space constraints.

As a first step towards a quantitative refinement of Newman's Lemma, we extend the notion of termination to  $\Omega$ -relations. To understand what is the right notion of termination in an enriched setting, we take an operational approach and stipulate that a (rewriting)  $\Omega$ -relation is terminating if it supports an induction principle. But what could the latter possibly be? Let us recall [[Doornbos et al. 1997](#)] that, given a binary relation  $R \subseteq A \times A$ , a property  $p$  on  $A$  is *R-inductive* if it satisfies the law  $(\forall x. xRy \rightarrow p(x)) \rightarrow p(y)$ , for any  $y \in A$ . We then say that the relation  $R$  *admits induction* if for any  $R$ -inductive property  $p$ ,  $p(a)$  holds for any  $a \in A$ . Consequently, if  $R$  admits induction and we want to prove that each element of  $A$  has a given property  $p$ , it is enough to prove  $p$  to be  $R$ -inductive. This is nothing but the familiar formulation of well-founded induction. We now generalise this idea to  $\Omega$ -relations and  $\Omega$ -properties.

**Definition 6.** Let  $(A, R)$  be a  $\Omega$ -ARS. A  $\Omega$ -property  $p : A \rightarrow \Omega$  is *R-inductive* if, for any  $b \in A$ , we have  $\bigwedge_a R(a, b) \multimap p(a) \leq p(b)$ . We say that  $R$  *admits induction* if for any  $a \in A$ ,  $p(a) = k$ , for any  $R$ -inductive predicate  $p$ .

We now define terminating  $\Omega$ -relations in such a way that terminating  $\Omega$ -relations coincide with those admitting induction.

**Definition 7.** Let  $(A, R)$  be a  $\Omega$ -ARS. We say that: (i)  $a \in A$  is a *normal form* if  $\bigvee_b R(a, b) = \perp$ ; (ii) a reduction sequence  $(a_0, \dots, a_n)$  *terminates* if  $a_n$  is a normal form; (iii)  $R$  is *terminating* if for each  $a \in A$ , there is no infinite reduction sequence starting from  $a$ .

**Proposition 3.** Let  $(A, R)$  be a  $\Omega$ -ARS. Then  $R^-$  admits induction if and only if  $R$  is terminating.

**PROOF SKETCH.** Suppose  $R^-$  admits induction. We prove that  $R$  is terminating by defining  $p(a) = k$  if there is no infinite reduction sequence starting from  $a$ , and  $p(a) = \perp$ , otherwise, and then showing that  $p$  is inductive. Vice versa, suppose  $R$  is terminating and let  $p$  be  $R^-$ -inductive. We prove  $p(a) = k$ , for any  $a \in A$  by contradiction showing that if there exists  $a \in A$  such that  $p(a) < k$ , then there also exists  $b \in A$  such that  $R(a, b) \neq \perp$  and  $p(b) < k$ .  $\square$

We now have all the ingredients to quantitatively refining Newman's Lemma.

**THEOREM 1 (QUANTITATIVE NEWMAN'S LEMMA).** Let  $(A, R)$  be a terminating  $\Omega$ -ARS. Then,  $R$  is confluent if and only if it is locally confluent.

**PROOF SKETCH.** Obviously, if  $R$  is confluent, then it is locally confluent too. To prove the converse, we define the  $\Omega$ -property  $p$  by stipulating  $p(a) = k$  if for all  $b_1, b_2 \in A$ ,  $R^*(a, b_1) \otimes R^*(a, b_2) \leq \bigvee_b R(b_1, b) \otimes R(b_2, b)$ . We define  $p(a) = \perp$  otherwise. Therefore,  $p$  is a Boolean property, in the sense that for any  $a \in A$ ,  $p(a)$  is either  $k$  or  $\perp$ . Moreover, we see that  $p(a) = k$  if and only if  $R$  is confluent on  $a$ . Since  $R$  is terminating,  $R^-$  admits induction, and thus we prove the thesis by showing that  $p$  is inductive.  $\square$

**Remark 2.** Notice that both [Proposition 3](#) and [Theorem 1](#) rely on cointegrality of  $\Omega$  (these are the only results requiring the cointegrality condition). Thus, for instance, they do not hold for the Fuzzy quantale of the Łukasiewicz  $t$ -norm (see, e.g., [[Belohlávek et al. 2010](#)]).

## 5 QUANTITATIVE TERM REWRITING: NON-EXPANSIVE SYSTEMS

Having defined  $\Omega$ -ARSSs, we now move to term-based systems, beginning with non-expansive quantitative term rewriting systems. Through this section, let  $\Omega = (\Omega, \leq, \otimes, k)$  be a fixed quantale.

**Notation 4.** Before going any further, we shortly recall some (standard) notions we will use in the rest of the paper.

*Terms* For a signature  $\Sigma$  and a countable set of variables  $X$ , we write  $\Sigma(X)$  for the collection of  $(\Sigma)$ -terms over  $X$ . We use small Latin letters to range over terms.

*Positions* A position  $p$  is a finite string of positive integers. We denote by  $\lambda$  the empty string and by  $pq$  the concatenation of positions  $p, q$ ; we write  $p \leq q$  if  $p$  is a prefix of  $q$ , i.e, if there is  $r$  such that  $q = pr$ . We write  $p \parallel q$  if  $p \not\leq q$  and  $q \not\leq p$ . Finally, we denote by  $t|_p$  the subterm of  $t$  at position  $p$ . If  $t|_p = s$ , we will also write  $t[s]_p$ .

*Context* A context is a term over the signature  $\Sigma \cup \{\square\}$ . We write  $C[\cdot]$  for a context containing a single occurrence of  $\square$  and use the notation  $C[t]$  to denote the term obtained by replacing the (single) occurrence of  $\square$  with  $t$  in  $C[\cdot]$ .

*Substitution* We denote substitutions by  $\sigma, \tau, \dots$  and write  $t^\sigma$  for  $\sigma(t)$ . Given substitutions  $\sigma, \tau$ , we write  $\sigma \leq \tau$  if there exists  $\rho$  such that  $\tau = \sigma\rho$ , where  $(\sigma\rho)(t) \triangleq \sigma(\rho(t))$ . Given terms  $t$  and  $s$ , if  $t^\sigma = s^\sigma$ , then  $\sigma$  is a unifier of  $t$  and  $s$ , while  $t$  and  $s$  are said to be *unifiable*. Finally, recall that the *most general unifier (mgu)* of two unifiable terms is their minimal unifier with respect to  $\leq$ .

*Linearity* We say that a term  $t$  is *linear* if it has no multiple occurrences of the same variable. We say that a mathematical expression involving terms is linear if all terms appearing in it are linear.

We are now ready to define non-expansive quantitative term rewriting systems, which we simply refer to as  $\Omega$ -term rewriting systems.

**Definition 8.** A  $\Omega$ -term rewriting system ( $\Omega$ -TRS, for short) is a pair  $\mathcal{R} = (\Sigma, \mapsto_{\mathcal{R}})$  consisting of a signature  $\Sigma$  and a  $\Omega$ -ternary relation  $\mapsto_{\mathcal{R}}$  over  $\Sigma$ -terms.<sup>14</sup> The (rewriting)  $\Omega$ -ternary relation  $\rightarrow_{\mathcal{R}}$  generated by  $\mapsto_{\mathcal{R}}$  is defined by the following rules:

$$\frac{\varepsilon \Vdash a \mapsto_{\mathcal{R}} b}{\varepsilon \Vdash C[a^\sigma] \rightarrow_{\mathcal{R}} C[b^\sigma]} \quad \frac{\varepsilon \Vdash t \rightarrow_{\mathcal{R}} s \quad \delta \leq \varepsilon}{\delta \Vdash t \rightarrow_{\mathcal{R}} s}$$

**Remark 3.** 1. In quantitative algebra, it is customary to consider ternary relations on a *base* of the quantale, rather than on the quantale itself [Dahlqvist and Neves 2022]. Although this choice makes definitions computationally lighter, it is irrelevant for our results. Consequently, we will continue working with full  $\Omega$ -ternary relations (nonetheless, the reader can safely pretend such relations to be over a base of  $\Omega$ ).

2. **Definition 8** stipulates that  $\rightarrow_{\mathcal{R}}$  must be closed under the structural rule of weakening, which states that we can always loose information and (under)approximate rewriting distances. To achieve (denotational) completeness, quantitative equational theories consider two additional structural rules, viz. closure under finite joins and the so-called Archimedean rule. From a rewriting perspective, however, such rules may not be as natural as they appear in an equational setting making, for instance, the defining rules of  $\rightarrow_{\mathcal{R}}$  *infinitary*. Nonetheless, we remark that such rules can be freely added to the definition of  $\rightarrow_{\mathcal{R}}$ , leaving our results essentially unchanged.

<sup>14</sup>To be precise, we require that if  $t \mapsto_{\mathcal{R}} s$ , then  $t$  is not a variable and all variables appearing in  $s$  also appear in  $t$ . This condition is necessary to avoid trivial lack of confluence issues (think about the Boolean rewriting rule  $f(x) \mapsto g(y)$ ). All systems we will consider indeed meet these requirements, except the system of Barycentric algebras we are going to introduce. That system, however, does not constitute a problem, since the only rule violating the aforementioned syntactic conditions is invertible, and hence it cannot lead to lack of confluence.

As usual, when  $R$  is clear from the context, we write  $\rightarrow$  in place of  $\rightarrow_R$  and use the notation  $t \xrightarrow{\varepsilon} s$  in place of  $\varepsilon \Vdash t \rightarrow s$ . We refer to triples  $a \xrightarrow{\varepsilon} b$  as reduction *rules*, and call  $a$  and  $b$  the *redex* and *contractum*, respectively.

Let us now comment on [Definition 8](#). The first defining rule of the relation  $\rightarrow_R$  is the very essence of non-expansive systems: it states that rewriting can be performed inside any context and on any instance of reductions in  $\mapsto_R$  *without* increasing rewriting distances. This rule reflects the (semantic) assumption that operation symbols in  $\Sigma$  behave as *non-expansive* functions: accordingly, contexts do not amplify rewriting distances. As we shall see, to make such an assumption consistent (both equationally and operationally) we have to require non-expansive systems to be linear. Finally, we notice that any  $\Omega$ -TRS  $(\Sigma, \mapsto_R)$  induces a  $\Omega$ -ARS with  $\Sigma$ -terms as object and with rewriting  $\Omega$ -relation  $R(t, s) \triangleq \bigvee \{\varepsilon \mid \varepsilon \Vdash t \rightarrow_R s\}$ . Consequently, all definitions and results seen so far extend to  $\Omega$ -TRSs. For that reason, we oftentimes say that a  $\Omega$ -TRS  $(\Sigma, \mapsto_R)$  has a given property when we actually mean that its associated  $\Omega$ -ARS has it.

**Example 4.** Let us now see some examples of  $\Omega$ -TRSs, which are summarised in [Table 1](#).

1. As a first example, we consider a simple system of natural number with addition. System  $\mathcal{N} = (\Sigma_{\mathcal{N}}, \mapsto_{\mathcal{N}})$  has signature  $\Sigma_{\mathcal{N}} \triangleq \{Z, S, A\}$  — containing a constant  $Z$  for zero, a unary function symbol  $S$  for the successor function, and a binary function symbol  $A$  for addition — and the  $\mathbb{L}$ -relation  $\mapsto_{\mathcal{N}}$  defined thus:

$$A(x, Z) \xrightarrow{0}_{\mathcal{N}} x \quad A(x, S(y)) \xrightarrow{0}_{\mathcal{N}} S(A(x, y)) \quad S(x) \xrightarrow{1}_{\mathcal{N}} x$$

The relation  $\mapsto_{\mathcal{N}}$  stipulates that actual distances are produced by deleting successor functions. The convertibility  $\mathbb{L}$ -relation  $N^{\equiv}$  coincides with the Euclidean distance on natural numbers.

2. Our second example is a quantitative string rewriting system<sup>15</sup> for the computation of edit distances on DNA molecules. Let us consider the  $\mathbb{L}$ -TRS  $\mathcal{M} = (\Sigma_{\mathcal{M}}, \mapsto_{\mathcal{M}})$  defined thus: the signature  $\Sigma_{\mathcal{M}} \triangleq \{A, C, G, T, \text{nil}\}$  contains *unary* function symbols  $A, C, G, T$  (DNA-bases) and a constant  $\text{nil}$  acting as the empty string (this is a standard way to model string systems as term ones), whereas the  $\mathbb{L}$ -relation  $\mapsto_{\mathcal{M}}$  is defined by the following rules, where  $b, c \in \{A, C, G, T\}$ , and  $b \neq c$  in the third rule.

$$x \xrightarrow{1}_{\mathcal{M}} b(x) \quad b(x) \xrightarrow{1}_{\mathcal{M}} x \quad b(x) \xrightarrow{1}_{\mathcal{M}} c(x).$$

The rewriting distance  $M$  operationally describes the Levenshtein distance [\[Gusfield 1997\]](#) between DNA sequences, the latter coinciding with convertibility distance  $M^{\equiv}$  induced by  $\mapsto_{\mathcal{M}}$ . Further edit distances on DNA molecules can be easily obtained modifying system  $\mathcal{M}$ . For instance, dropping the first two rules, we measure the number of mutations between DNA sequences, thus obtaining the *Hamming distance* between molecules [\[Gusfield 1997\]](#). And if we modify the rule for the Hamming distance allowing mutations between purines ( $A, G$ ) and pyrimidines ( $C, T$ ) only, then we obtain the so-called Eigen–McCaskill–Schuster distance [\[Deza and Deza 2009\]](#), (these mutations are used to study virus and cancer proliferation).

3. We now see how  $\Omega$ -TRSs provide an operational analysis of quantitative computational effects. We begin with quantitative nondeterminism. System  $\mathcal{L} = (\Sigma_{\mathcal{L}}, \mapsto_{\mathcal{L}})$  of quantitative semilattices is the  $\mathbb{L}^{\max}$ -TRS given by the signature  $\Sigma_{\mathcal{L}}$  containing a single binary operation  $\cup$  for nondeterministic choice and  $\mathbb{L}^{\max}$ -relation  $\mapsto_{\mathcal{L}}$  defined thus:

$$x \xrightarrow{0}_{\mathcal{L}} x \cup x \quad (x \cup y) \cup z \xrightarrow{0}_{\mathcal{L}} x \cup (y \cup z) \quad (x \cup y) \xrightarrow{0}_{\mathcal{L}} (y \cup x).$$

<sup>15</sup>It is well-known [\[Bezem et al. 2003\]](#) that string rewriting systems [\[Book and Otto 1993; Thue 1914\]](#) can be modelled as TRSs. This result holds *mutatis mutandis* in a quantitative setting.



Table 1. Main Examples of  $\Omega$ -TRSs

System	Objects/Name	Distance Induced
$\mathcal{N} = (\Sigma_{\mathcal{N}}, \mapsto_{\mathcal{N}})$	Natural Numbers	Euclidean Distance
$\mathcal{B} = (\Sigma_{\mathcal{B}}, \mapsto_{\mathcal{B}})$	Multi-distributions/Barycentric algebras	Total Variation distance
$\mathcal{K}_{\mathcal{N}} = (\Sigma_{\mathcal{K}_{\mathcal{N}}}, \mapsto_{\mathcal{K}_{\mathcal{N}}})$	Affine combinators	Weak Reduction
$\mathcal{T} = (\Sigma_{\mathcal{T}}, \mapsto_{\mathcal{T}})$	Ticking	Cost distance
$\mathcal{M} = (\Sigma_{\mathcal{M}}, \mapsto_{\mathcal{M}})$	DNA molecules	Edit distances
$\mathcal{L} = (\Sigma_{\mathcal{L}}, \mapsto_{\mathcal{L}})$	Quantitative (semi)lattices	Hausdorff distance

The definition of  $\mapsto_L$  is not interesting in itself. What is interesting is the choice of the quantale used for distances, namely the *strong* Lawvere quantale  $\mathbb{L}^{\max}$ . This impacts on the definition of  $\mapsto_L$ , which now gives a form of non-expansiveness of  $\cup$  reflecting the *idempotent* structure of  $\mathbb{L}^{\max}$ . In particular, if  $\varepsilon \Vdash x \rightarrow x'$  and  $\delta \Vdash y \rightarrow y'$ , then  $\max(\varepsilon, \delta) \Vdash x \cup y \rightarrow^* x' \cup y'$ . The convertibility distance  $L^{\equiv}$  gives the so-called theory of quantitative semilattices [Mardare et al. 2016] and axiomatises the Hausdorff distance between sets [Munkres 2000]. Since  $\mathbb{L}^{\max}$  is idempotent, quantitative reasoning on it becomes similar to traditional Boolean reasoning, and linearity is not necessary to avoid distance trivialisation and to ensure confluence properties of systems.

4. The  $\mathbb{L}$ -TRS  $\mathcal{B} = (\Sigma_{\mathcal{B}}, \mapsto_{\mathcal{B}})$  of Barycentric algebras is given by a signature  $\Sigma_{\mathcal{B}}$  containing a family of binary probabilistic choice operations  $+_{\varepsilon}$  indexed by rational numbers  $\varepsilon \in \mathbb{Q} \cap [0, 1]$  and by the  $\mathbb{L}$ -relation  $\mapsto_{\mathcal{B}}$  defined thus, where  $\varepsilon, \kappa \in (0, 1)$  and  $\varepsilon \leq \varepsilon$  in the last rule.

$$x +_1 y \mapsto_{\mathcal{B}}^0 x \quad x +_{\varepsilon} y \mapsto_{\mathcal{B}}^0 y +_{1-\varepsilon} x \quad (x +_{\varepsilon} y) +_{\kappa} z \mapsto_{\mathcal{B}}^0 x +_{\varepsilon\kappa} (y +_{\frac{\varepsilon-\varepsilon\kappa}{1-\varepsilon\kappa}} z) \quad x +_{\varepsilon} y \mapsto_{\mathcal{B}}^{\varepsilon} z +_{\varepsilon} y.$$

Notice that system  $\mathcal{B}$  does not have the idempotency rule  $x +_{\varepsilon} x \mapsto_{\mathcal{B}}^0 x$ , meaning that we are actually modelling *multi-distributions* [Avanzini et al. 2020] rather than distributions: this guarantees linearity of  $\mathcal{B}$  and agrees with the definition of a probabilistic rewriting system by Avanzini et al. [2020]. The convertibility distance  $B^{\equiv}$  coincides with the *total variation distance* [Villani 2008] between multi-distributions.

5. The  $\mathbb{L}$ -TRS  $\mathcal{T} = (\Sigma_{\mathcal{T}}, \mapsto_{\mathcal{T}})$  of ticking is defined by the signature  $\Sigma_{\mathcal{T}}$  of  $\mathbb{N}$ -indexed unary operation symbols  $n.(\cdot)$  and the following  $\mathbb{L}$ -relation, where  $\varepsilon \geq |n - m|$  in the last rule.

$$0.x \mapsto_{\mathcal{T}}^0 x \quad n.(m.x) \mapsto_{\mathcal{T}}^0 (n+m).x \quad n.x \mapsto_{\mathcal{T}}^{\varepsilon} m.x.$$

System  $\mathcal{T}$  is an example of quantitative output [Bacci et al. 2020] and we can read  $n.t$  as *count  $n$  unit of cost, then continue as  $t$* . Oftentimes, one writes terms of the form  $1.t$  as  $\checkmark t$  and decorates programs with  $\checkmark$  annotations to count computation steps. In that case, we actually obtain a simplified system  $\mathcal{T}_{\checkmark}$  whose signature contains the unary function symbol  $\checkmark$  only and whose (unique) rewriting rule is  $\checkmark x \mapsto^1 x$ . System  $\mathcal{T}_{\checkmark}$  can be then used to perform improvement-based and monadic cost analysis [Dal Lago and Gavazzo 2019; Sands 1998].

### 5.1 New Systems from Old Ones

In the previous example, we have introduced several  $\Omega$ -TRSs, namely systems  $\mathcal{N}$ ,  $\mathcal{M}$ ,  $\mathcal{B}$ ,  $\mathcal{T}$ , and  $\mathcal{L}$ . We now show how to combine such systems together by means of their sum [Bezem et al. 2003].

**Definition 9.** Given  $\Omega$ -TRSs  $\mathcal{R} = (\Sigma_{\mathcal{R}}, \mapsto_{\mathcal{R}})$  and  $\mathcal{S} = (\Sigma_{\mathcal{S}}, \mapsto_{\mathcal{S}})$  with disjoint signatures, we define their *sum* as the  $\Omega$ -TRS  $\mathcal{R} + \mathcal{S} = (\Sigma_{\mathcal{R}+\mathcal{S}}, \mapsto_{\mathcal{R}+\mathcal{S}})$  defined thus:  $\Sigma_{\mathcal{R}+\mathcal{S}} \triangleq \Sigma_{\mathcal{R}} \cup \Sigma_{\mathcal{S}}$  and  $\mapsto_{\mathcal{R}+\mathcal{S}} \triangleq \mapsto_{\mathcal{R}} \cup \mapsto_{\mathcal{S}}$ .

We immediately notice that the  $\Omega$ -relation  $RS$  associated to  $\mathcal{R} + \mathcal{S}$  coincides with  $R \vee S$ , so that we can rely on [Proposition 2](#) to prove confluence of  $\mathcal{R} + \mathcal{S}$  modularly.

**Example 5.** System  $\mathcal{K} = (\Sigma_{\mathcal{K}}, \mapsto_K)$  of affine combinators [[Hindley 2008](#)] is defined by the signature  $\Sigma_{\mathcal{K}}$  containing the three constants (basic combinators)  $B$ ,  $C$ , and  $K$  and a single binary operation symbol  $\cdot$  for application (we assume application to associate to the left omitting unnecessary parentheses), together with the  $\mathbb{L}$ -relation  $\mapsto_K$  defined thus:

$$B \cdot x \cdot y \cdot z \xrightarrow{0}_K x \cdot (y \cdot z) \quad C \cdot x \cdot y \cdot z \xrightarrow{0}_K x \cdot z \cdot y \quad K \cdot x \cdot y \xrightarrow{0}_K x$$

To make system  $\mathcal{K}$  truly quantitative, we combine it with one or more of the  $\Omega$ -TRSs previously introduced. For instance, system  $\mathcal{K} + \mathcal{B}$  gives a quantitative system of *probabilistic* affine combinators. Similarly, extensions of  $\mathcal{K}$  with ticking are obtained as  $\mathcal{K} + \mathcal{T}$  and  $\mathcal{K} + \mathcal{T}_{\checkmark}$ , whereas  $\mathcal{K} + \mathcal{L}$  gives nondeterministic affine combinators.

## 5.2 Confluence, Part I

At this point, we have introduced several examples of  $\Omega$ -TRSs and it is now the time to investigate their properties. In light of the connection (that, due to space constraints, we cannot formally spell out in the present paper) between quantitative rewriting and quantitative equational theories, here we focus on proving (quantitative) confluence (and thus CR) of  $\Omega$ -TRSs. From (quantitative) Newman's Lemma ([Theorem 1](#)), we know that to prove confluence of a terminating  $\Omega$ -TRS we only need to verify its *local* confluence. Proving local confluence of a  $\Omega$ -TRS, however, can be difficult, as reductions may happen inside arbitrary contexts and on arbitrary instances of reduction rules. It is thus natural to ask whether we can prove local confluence *locally*, i.e. by looking at *ground* rewriting only.

In this section, we show that local confluence of a *linear*  $\Omega$ -TRS follows directly from local confluence of its *critical pairs* [[Bezem et al. 2003](#); [Huet 1980](#)]. Linearity, as we have already discussed, is a crucial property in quantitative and metric reasoning: forcing non-expansiveness on non-linear systems often let distance trivialise [[Crubillé and Dal Lago 2014, 2017](#); [Gavazzo 2018, 2019](#)], this way collapsing quantitative equational deduction to traditional, Boolean reasoning. On rewriting systems, non-linearity leads to further undesired consequences, as we shall see.

Let us now recall the notion of a critical pair and refine the well-known critical pair lemma [[Huet 1980](#)] to a quantitative setting.

**Definition 10.** Let  $(\Sigma, \mapsto_R)$  be a  $\Omega$ -TRS and  $a_1 \xrightarrow{\varepsilon} b_1, a_2 \xrightarrow{\delta} b_2$  be renamings of rewrite rules without common variables. We say that  $a_1 \xrightarrow{\varepsilon} b_1$  and  $a_2 \xrightarrow{\delta} b_2$  overlap at position  $p$  if: (i)  $p$  is a function symbol position of  $a_2$ ; (ii)  $a_1$  and  $a_2|_p$  are unifiable; (iii) if  $p = \lambda$ , then the two rules are not variants (i.e. they cannot be obtained one from the other by variables renaming). If  $a_1 \xrightarrow{\varepsilon} b_1, a_2 \xrightarrow{\delta} b_2$  overlap at position  $p$  and  $\sigma$  is the mgu of  $a_2|_p$  and  $a_1$ , then we have the reductions  $a_2^\sigma \xrightarrow{\delta} b_2^\sigma$  and  $a_2^\sigma \xrightarrow{\varepsilon} a_2^\sigma[b_1^\sigma]_p$ . We call the triple  $(a_1 \xrightarrow{\varepsilon} b_1, p, a_2 \xrightarrow{\delta} b_2)$  a *critical overlap* and the pair  $(a_2^\sigma[b_1^\sigma]_p, b_2^\sigma)$  a *critical pair*. We denote by  $CP(\mathcal{R})$  the collection of its critical pairs of  $\mathcal{R}$ .

**Example 6.** Consider system  $\mathcal{B}$  of Barycentric algebras and the (no-common-variable renamings of) commutativity and associativity rules  $x' +_{\epsilon_1} y' \xrightarrow{0}_\delta y' +_{1-\epsilon_1} x'$  and  $(x' +_{\epsilon_1} y') +_{\epsilon_2} z \xrightarrow{0}_\delta x' +_{\epsilon_1 \epsilon_2} (y' +_{\frac{\epsilon_1 - \epsilon_1 \epsilon_2}{1 - \epsilon_1 \epsilon_2}} z)$ . Then, we see that the substitution  $\sigma$  mapping the variable  $x'$  to  $x$  and  $y'$  to  $y$  is the mgu of  $((x +_{\epsilon_1} y) +_{\epsilon_2} z)|_1$  and  $x' +_{\epsilon} y'$ . Consequently, the following triple and pair form a *critical*

overlap and a critical pair, respectively:

$$(x' +_{\epsilon_1} y' \xrightarrow{\delta}_\delta y' +_{1-\epsilon} x', 1, x +_{\epsilon_1 \epsilon_2} (y +_{\frac{\epsilon_1 - \epsilon_1 \epsilon_2}{1 - \epsilon_1 \epsilon_2}} z)) \quad ((y +_{1-\epsilon_1} x) +_{\epsilon_2} z, x +_{\epsilon_1 \epsilon_2} (y +_{\frac{\epsilon_2 - \epsilon_1 \epsilon_2}{1 - \epsilon_1 \epsilon_2}} z))$$

We are now ready to prove the quantitative refinement of the so-called Critical Pair Lemma [Huet 1980]. In its traditional version, such a lemma states that to prove local confluence of a rewriting relation it is enough to prove its local confluence on critical pairs only. When we move to quantitative rewriting, the Critical Pair Lemma needs a further assumption, namely *linearity*.

**Example 7.** Consider the signature  $\Sigma \triangleq \{f, e, i\}$  with  $f$  a binary function symbol and  $e, i$  constants. Let  $\mapsto_R$  be the  $\mathbb{L}$ -relation defined by  $f(x, x) \xrightarrow{0} x$  and  $e \xrightarrow{1} i$ . Notice that the system has no critical pair, and thus all its critical pairs are trivially locally confluent. Consider the peak  $\pi$  given by  $e \xleftarrow{0} f(e, e) \xrightarrow{1} f(i, e)$ , so that  $R(f(e, e), e) = 0$ ,  $R(f(e, e), f(i, e)) = 1$ . To close  $\pi$ , we need to reduce  $e$  twice:  $e \xrightarrow{1} i \xleftarrow{0} f(i, i) \xleftarrow{1} f(i, e)$ . This gives  $R^*(e, i) = 1$  and  $R^*(f(i, e), f(i, i)) = 1$ , this way breaking (local) confluence, and thus showing the necessity of the linearity assumption in Lemma 1.

**Lemma 1 (Critical Pair).** *Let  $\mathcal{R} = (\Sigma, \mapsto_R)$  be a linear  $\Omega$ -TRS. If  $R$  is locally confluent on all critical pairs of  $\mathcal{R}$ , then it is locally confluent. Moreover, if  $\Omega$  is idempotent, then linearity is not needed.*

**PROOF SKETCH.** We have to show  $R^-; R \leq R^*$ ;  $R^{*-}$  given that  $(R^-; R)(t, s) \leq (R^*; R^{*-})(t, s)$ , for any  $(t, s) \in CP(\mathcal{R})$ . Pointwise, we need to prove  $\bigvee_t R(t, s_1) \otimes R(t, s_2) \leq \bigvee_u R^*(s_1, u) \otimes R^*(s_2, u)$ , for arbitrary terms  $s_1$  and  $s_2$ . Since  $R(t, s) = \bigvee \{\varepsilon \mid \varepsilon \Vdash t \rightarrow s\}$  and the join distributes over the tensor, it is enough to show that for any local peak  $s_1 \xleftarrow{\varepsilon_1} t \xrightarrow{\varepsilon_2} s_2$ , we have  $\varepsilon_1 \otimes \varepsilon_2 \leq \bigvee_u R^*(s_1, u) \otimes R^*(s_2, u)$ . We proceed by structural induction on  $\varepsilon_1 \Vdash t \rightarrow s_1$  and  $\varepsilon_2 \Vdash t \rightarrow s_2$ . The case for weakening is straightforward. The interesting case is for  $\varepsilon_1 \Vdash t \rightarrow s_1$  obtained by applying a rule  $\varepsilon_1 \Vdash a_1 \mapsto b_1$  at position  $p_1$ , and  $\varepsilon_2 \Vdash t \rightarrow s_2$  obtained by applying a rule  $\varepsilon_2 \Vdash a_2 \mapsto b_2$  at position  $p_2$ . We proceed by cases, depending on the relationship between  $p_1$  and  $p_2$ . We show one case as an example. Suppose that  $p_1 \leq p_2$ , so that there is  $q$  such that  $p_2 = p_1 q$ , and  $\varepsilon \Vdash a_1 \mapsto b_1$  and  $\delta \Vdash a_2 \mapsto b_2$  overlap at position  $q$ . Then,  $a_1^\sigma|_q = (t|_{p_1})|_q = t|_{p_2} = a_2^\sigma$ . That is,  $\sigma$  is a unifier of  $a_1|_q$  and  $a_2$ . Let  $\tau$  be their most general unifier, so that  $\sigma = \tau\rho$ , for some substitution  $\rho$ . Then  $(a_1^\tau[b_2^\tau]_q, b_1^\tau)$  is a critical pair. By hypothesis, we know that  $(R^-; R)(a_1^\tau[b_2^\tau]_q, b_1^\tau) \leq \bigvee_u R^*(a_1^\tau[b_2^\tau]_q, u) \otimes R^*(b_1^\tau, u)$ , and thus  $\varepsilon \otimes \delta \leq \bigvee_u R^*(a_1^\tau[b_2^\tau]_q, u) \otimes R^*(b_1^\tau, u)$ . To prove the thesis, it is sufficient to prove

$$\bigvee_u R^*(a_1^\tau[b_2^\tau]_q, u) \otimes R^*(b_1^\tau, u) \leq \bigvee_v R^*(s_1, v) \otimes R^*(s_2, v).$$

As usual, it is enough to show that for any  $u$  such that  $\eta \Vdash a_1^\tau[b_2^\tau]_q \rightarrow u$  and  $\iota \Vdash b_1^\tau \rightarrow u$  we have  $\eta \otimes \iota \leq \bigvee_v R^*(s_1, v) \otimes R^*(s_2, v)$ . Fixed  $u$  as above, we have  $s_1 = t[b_1^{\tau\rho}]_{p_1}$  and  $s_2 = t[a_1^{\tau\rho}[b_2^{\tau\rho}]_q]_{p_1}$  and thus  $s_2 = t[(a_1^\tau[b_2^\tau]_q)^\rho]_{p_1} \xrightarrow{\eta} t[u^\rho]_{p_1} \xleftarrow{\iota} t[b_1^{\tau\rho}]_{p_1} = s_1$ .  $\square$

**THEOREM 2.** *Any linear and terminating  $\Omega$ -TRS locally confluent on its critical pairs is confluent.*

**PROOF.** It directly follows from Theorem 1 and Lemma 1.  $\square$

**Example 8.** 1. System  $\mathcal{N}$  of natural numbers is terminating and locally confluent on all its critical pairs. By Theorem 2, we conclude that  $\mathcal{N}$  is confluent.  
2. System  $\mathcal{M}$  (and variations thereof) is not terminating and thus we cannot rely on Theorem 2 to prove their confluence. One easy way to overcome the problem is to rephrase  $\mathcal{M}$  as terminating system. For instance, we can stipulate that molecules can only be deleted and that substitution of molecules is directed, in the sense that, e.g., A can become C, G, and T, but not vice-versa (similarly, C can become G and T, and G can only become a T). This way, we indeed obtain a terminating

- system locally confluent on its critical pairs, and thus a confluent system. Although this approach works fine as long as we are interested in reachability problems (and alike), some care is needed when dealing with optimal distances (i.e. the task of finding reductions  $\varepsilon \Vdash t \rightarrow_R s$  such that  $R(t, s) = \varepsilon$ ), as forcing termination may lead to increasing minimal distances between molecules.
3. System  $\mathcal{T}_\checkmark$  is terminating and locally confluent, and thus confluent. System  $\mathcal{T}$  is not terminating, due to the rule  $n.x \xrightarrow{\varepsilon} m.x$ , with  $\varepsilon \geq |n - m|$ . We can easily fix that by imposing  $m < n$ , this way obtaining a terminating and locally confluent — and thus confluent — system.

**Theorem 2** constitutes a powerful tool to prove confluence of  $\Omega$ -TRSs. In a quantitative setting, however, termination might be a too strong condition, and interesting  $\Omega$ -TRSs may not satisfy it. As an example, consider system  $\mathcal{B}$  of Barycentric algebras. To solve this issue, we modify **Lemma 1** replacing local confluence with a stronger condition, namely *strong confluence*.

**Definition 11.** We say that a  $\Omega$ -relation  $R : A \rightarrow A$  is *strongly confluent* if  $R^-; R \leq R^-; R^{*-}$ . If, additionally,  $R$  satisfies  $R^-; R \leq R^*; R^{*-}$ , then we say that it is *strongly closed*.

A lexicographic induction shows that if  $R$  is strongly confluent, then it is confluent. Moreover, as it happens with local confluence, if a  $\Omega$ -TRS is linear, then strong confluence is implied by strong confluence on critical pairs.

**THEOREM 3.** *If a linear  $\Omega$ -TRS is strongly closed on all its critical pairs, then it is confluent. Moreover, if  $\Omega$  is idempotent, then linearity is not needed.*

**PROOF SKETCH.** Adapting the proof of **Lemma 1** one proves that if  $\mathcal{R} = (\Sigma, \mapsto_R)$  is linear, then  $R$  is strongly closed if and only if  $R$  is strongly closed on all the critical pairs of  $\mathcal{R}$ .  $\square$

We can now rely on **Theorem 3** to prove confluence of  $\Omega$ -TRSs.

**Example 9.** Inspecting the critical pairs of system  $\mathcal{B}$ , we see that  $B$  is strongly closed, and thus  $\mathcal{B}$  is confluent. In a similar fashion, we see that system  $\mathcal{L}$  is confluent too.

We now have proved confluence of all systems of **Table 1** (confluence of  $\mathcal{K}$  follows since  $K$  collapses to the traditional reduction of affine combinators, which is confluent [Hindley 2008]). But what about their combinations? Is system  $\mathcal{K} + \mathcal{B}$  confluent? And system  $\mathcal{K} + \mathcal{B} + \mathcal{N}$ ? Given  $\Omega$ -TRSs  $\mathcal{R} = (\Sigma_{\mathcal{R}}, \mapsto_R)$ ,  $\mathcal{S} = (\Sigma_{\mathcal{S}}, \mapsto_S)$ , we have observed that the sum  $\Omega$ -relation  $RS$  coincides with  $R \vee S$ , so that we find ourselves in the condition to rely on the quantitative Hindley-Rosen Lemma to prove confluence of  $\mathcal{R} + \mathcal{S}$ . Accordingly, to infer confluence of  $R \vee S$  we need to have confluence of  $R$  and  $S$  as well as commutation of  $R$  with  $S$ . Whereas the former usually is our starting hypothesis, the latter requires a specific analysis. For the cases we are interested in, such an analysis is smooth, as systems  $\mathcal{R}$  and  $\mathcal{S}$  are essentially independent, viz.  $\mathcal{R} + \mathcal{S}$  does not create new critical pairs.

**Lemma 2.** *Given two linear  $\Omega$ -TRS  $\mathcal{R}, \mathcal{S}$  as above, if the collection of critical pairs obtained by overlapping of a rule of  $\mathcal{R}$  and a rule of  $\mathcal{S}$  is empty, then  $R$  strongly commutes with  $S$ , and thus  $R$  commutes with  $S$ . Moreover, if  $\Omega$  is idempotent, linearity is not necessary.*

**PROOF SKETCH.** The proof that  $R$  strongly commutes with  $S$  is a simplified instance of the proof of **Lemma 1** and **Theorem 3**. From that, commutation of  $R$  with  $S$  follows.  $\square$

**Example 10.** Since systems  $\mathcal{K}$  and  $\mathcal{B}$  have no common critical pair, **Lemma 2** gives confluence of  $\mathcal{K} + \mathcal{B}$ . Similarly, systems  $\mathcal{K} + \mathcal{N}$ ,  $\mathcal{K} + \mathcal{T}$ , and  $\mathcal{K} + \mathcal{L}$  are confluent.

## 6 BEYOND NON-EXPANSIVENESS: GRADED SYSTEMS

In the previous section, we have outlined a theory of non-expansive quantitative term rewriting systems showing, in particular, how *linearity* plays a key role to ensure their correct behaviours. The results achieved show that non-expansive systems provide a powerful formalism for the quantitative analysis of linear systems that, however, cannot cope with non-linearity. In this section, we introduce a new class of quantitative term rewriting systems — that we dub *graded* rewriting systems — that allows us to model non-linear systems avoiding, at the same time, distance trivialisation and lack of confluence issues.

Non-expansive systems, in fact, treat term constructors as non-expansive functions (with respect to rewriting distances): given a  $\Omega$ -TRS  $\mathcal{R} = (\Sigma_{\mathcal{R}}, \mapsto_{\mathcal{R}})$ , we have  $R(t, s) \leq R(C[t], C[s])$  (and more generally,  $\bigotimes_i R^*(t_i, s_i) \leq R^*(f(t_1, \dots, t_n), f(s_1, \dots, s_n))$ ). Non-expansive maps, however, are not the only maps one is interested in metric reasoning. Another more liberal class of functions is the one of *Lipschitz continuous* functions [Searcoid 2006]. Moving from an original observation by Lawvere [1973], generalisations of Lipschitz continuous maps to  $\Omega$ -relations have been given in terms of change of base functors [Gavazzo 2018, 2019], and corelaters [Dal Lago and Gavazzo 2022b]. In a nutshell, we allow contexts  $C$  to amplify distances, but in a controlled way. Such a way is given by a (family of suitable) function(s)  $\partial_C : \Omega \rightarrow \Omega$ , so that we replace non-expansiveness with the inequality  $\partial_C(R(t, s)) \leq R(C[t], C[s])$ , which allows contexts to amplify rewriting distances, but in a controlled way. In fact, the map  $\partial_C$  — called the *sensitivity* or *degree* of  $C$  — gives the law describing how  $C$  amplifies distances. Accordingly, we think about sensitivity as generalising Lipschitz constants; and indeed, multiplication by a constant is a typical example of a sensitivity on the Lawvere quantale.

To ensure such a general form of Lipschitz continuity of contexts (and, more generally, of term constructors), graded term rewriting systems rely on two distinguished features. First, the rewriting  $\Omega$ -ternary relation  $\rightarrow_R$  is defined relying on the rule

$$\frac{\varepsilon \Vdash t \mapsto_R s}{\partial_C(\varepsilon) \Vdash C[t^\sigma] \rightarrow_R C[s^\sigma]}$$

where  $\partial_C$  is the sensitivity of  $C$ . Second, terms in graded systems are built relying on *modal signatures* [Dagnino and Pasquali 2022], whereby  $n$ -ary function symbols  $f$  come with *modal arities*  $(\phi_1, \dots, \phi_n)$  specifying that  $f$  has sensitivity  $\phi_i$  on its  $i$ th argument. Since terms are built using variables and function symbols, modal arities (and their algebra) allow us to compute the sensitivity of any context, and thus to define  $\rightarrow_R$  as previously described. All of that makes graded term rewriting systems *modal* and *coeffectful* [Orchard et al. 2019; Petricek et al. 2014] and, at the same time, makes them a powerful tool to model the operational semantics of modal and coeffectful languages, as we shall see.

### 6.1 Modal and Graded Rewriting: $(\Omega, \Phi)$ -Systems

We now introduce graded systems formally. To do so, we first recall the notion of a quantale homomorphism which we will use to define context sensitivity.

**Definition 12.** Given quantales  $\Omega = (\Omega, \leq, \otimes, k)$ ,  $\Theta = (\Theta, \sqsubseteq, \boxtimes, j)$  a quantale homomorphism is a monotone map  $h : \Omega \rightarrow \Theta$  such that:  $j = h(k)$ ,  $h(\varepsilon) \boxtimes h(\delta) = h(\varepsilon \otimes \delta)$ ; and  $h(\bigvee_i \varepsilon_i) = \bigsqcup_i h(\varepsilon_i)$ .

From now on, we shall work with quantale homomorphisms on the same quantale  $\Omega$ . We denote such maps by  $\phi, \psi, \dots$  and refer to them as *change of base (endo)functors* (CBEs, for short).

**Example 11.** 1. The main example of CBEs we consider is multiplication by a constant on the Lawvere quantale (and variations thereof). Given  $\kappa \in \mathbb{R}_{\geq 0}$ , we regard  $\kappa$  as mapping  $\varepsilon \in [0, \infty]$  to  $\kappa\varepsilon \in [0, \infty]$ . Notice that we do not allow multiplication by infinity.



2. The map  $\psi \circ \varphi$  is a CBE, where  $\psi : 2 \rightarrow \Omega$  is defined by  $\psi(\top) \triangleq k$ ,  $\psi(\perp) \triangleq \perp$  and  $\varphi : \Omega \rightarrow 2$  is its right adjoint defined by  $\varphi(\varepsilon) \triangleq \top$  if  $\varepsilon = k$  and  $\varphi(\varepsilon) \triangleq \perp$ , otherwise.
3. Other examples of CBEs, especially on quantales of modal predicates, can be found in the literature on relational reasoning on coeffects [Dal Lago and Gavazzo 2022b].

CBEs are closed under composition and the identity function  $1 : \Omega \rightarrow \Omega$  is a CBE. Moreover, we extend the order  $\leq$  and the multiplication  $\otimes$  of  $\Omega$  to CBEs pointwise. Finally, we denote by  $k^\star$  the constant  $k$  CBE. These operations are precisely what we need to compute degree of variables in terms starting from the modal arity of operation symbols. Indeed, the reader familiar with coeffects may have recognised that such operations endow CBEs with a grade (or resource) algebra, i.e. a pre-ordered semiring [Gaboridi et al. 2016; Orchard et al. 2019]. Additionally, any CBE  $\phi$  induces an action  $[\phi]$  on  $\Omega$ -relations defined by  $[\phi]R(a, b) \triangleq \phi(R(a, b))$ . The map  $[\phi]$  is an example of a corelator [Dal Lago and Gavazzo 2022b].

We now introduce a new class of rewriting systems, which we dub  $(\Omega, \Phi)$ -systems. Let us fix a quantale  $\Omega$  and a structure  $\Phi = (\Phi, \leq, \circ, 1, \otimes, k^\star)$ , where  $\Phi$  is a set of CBEs containing the identity and constant  $k$ -functions, and closed under function composition and tensor.

- Definition 13.**
1. The *modal arity* of an  $n$ -ary function symbol  $f$  is a tuple  $(\phi_1, \dots, \phi_n)$  with  $\phi_i \in \Phi$ . Given a function symbol  $f$  with modal arity  $(\phi_1, \dots, \phi_n)$  (notation  $f : (\phi_1, \dots, \phi_n)$ ), we say that  $f$  has sensitivity (or modal grade)  $\phi_i$  on its  $i$ th argument.
  2. A  $\Phi$ -graded signature is a set  $\Sigma$  containing function symbols with their modal arity. Given a  $\Phi$ -graded signature  $\Sigma$  and a set  $X$  of variables, the collection of  $\Sigma(X)$  is defined as usual.
  3. Given a term  $t$  and a position  $p$  for a variable in  $t$ , we define the grade  $\partial_p(t)$  of  $p$  in  $t$  as follows, where  $f : (\phi_1, \dots, \phi_n) \in \Sigma$ :

$$\partial_\lambda(t) \triangleq 1 \qquad \partial_{ip}(f(t_1, \dots, t_n)) \triangleq \phi_i \circ \partial_p(t_i).$$

Given a term  $t$  and a variable  $x$ , we can compute the grade  $\partial_x(t)$  of  $x$  in  $t$  by ‘summing’ the grades of all position  $p$  such that  $t|_p = x$ . Formally,  $\partial_x(t) \triangleq \bigotimes \{\partial_p(t) \mid t|_p = x\}$ , where  $\bigotimes \emptyset \triangleq k^\star$ . We write  $\partial_C$  in place  $\partial_p(C)$ , where  $p$  is the position of the hole in  $C$ .

**Definition 14.** A  $(\Omega, \Phi)$ -term rewriting system  $((\Omega, \Phi)$ -TRS, for short) is a pair  $\mathcal{R} = (\Sigma, \mapsto_R)$  consisting of a  $\Phi$ -graded signature  $\Sigma$  and a  $\Omega$ -ternary relation. The (rewriting)  $\Omega$ -ternary relation  $\rightarrow_R$  generated by  $\mapsto_R$  is defined thus:

$$\frac{\varepsilon \Vdash a \mapsto_R b}{\partial_C(\varepsilon) \Vdash C[a^\sigma] \rightarrow_R C[b^\sigma]} \quad \frac{\varepsilon \Vdash t \rightarrow_R s \quad \delta \leq \varepsilon}{\delta \Vdash t \rightarrow_R s}$$

We say that  $\mathcal{R}$  is *balanced* if for any rule  $\varepsilon \Vdash a \mapsto b$  we have  $\partial_x(a) = \partial_x(b)$ , for any variable  $x$ . From now on, we assume all  $(\Omega, \Phi)$ -TRSs to be balanced.

Notice that Definition 14 allows contexts (and function symbols) to amplify distances, but seems to leave substitution non-expansive. This is not the case (as we shall clearly see when defining multi-reductions), but it seems so because in the definition of  $\rightarrow_R$  we apply *the same* substitution on terms. Intuitively, this reflects the fact that passing identical (i.e. at a null distance) arguments to a Lipschitz continuous function produces identical results, and thus there is no distance amplification. Notice also that an  $\Omega$ -TRS is a  $(\Omega, \Phi)$ -TRS with  $\Phi = \{k^\star\}$  and that, in this case, the balanced condition forces rules not to duplicate variables. Finally, any  $(\Omega, \Phi)$ -TRS  $(\Sigma, \mapsto_R)$  induces a  $\Omega$ -ARS whose objects are  $\Sigma$ -terms and whose rewriting  $\Omega$ -relation is defined by  $R(t, s) \triangleq \bigvee \{\varepsilon \mid \varepsilon \Vdash t \rightarrow_R s\}$ .

**Example 12.** 1. The main example of a  $(\Omega, \Phi)$ -TRS we consider is system  $\mathcal{W}$  *graded combinatory logic* [Atkey 2018; Dagnino and Pasquali 2022], a generalisation of Abramsky’s bounded combinatory logic [Abramsky 2002; Abramsky et al. 2002]. Graded combinators have been introduced as

a way to extend linear combinatory logic with non-linear combinators (such as  $W$ ) in a controlled way. To do so, one refines system  $\mathcal{K}$  by introducing graded exponential modalities  $!_n$  allowing to break linearity up to usage  $n$ . Formally, system  $\mathcal{W}$  is defined by the modal signature — notice that application remains non-expansive —  $\Sigma_{\mathcal{W}} \triangleq \{B, C, K, W_{n,m}, D, \delta_{n,m}, F_n, !_n : n, \cdot : (1, 1) \mid n, m \in \mathbb{N}\}$ , and by the  $\mathbb{L}$ -relation  $\mapsto_W$  defined by extending  $\mapsto_K$  with the rules (for readability, we write  $t$   $s$  in place of  $t \cdot s$ ):

$$D !_1 x \mapsto_W x \quad \delta_{n,m} !_n m x \mapsto_W !_n !_m x \quad F_n !_n x !_n y \mapsto_W !_n (x y) \quad W_{n,m} x !_n m y \mapsto_W x !_n y !_m y$$

2. We can grade the signature of system  $\mathcal{B}$  by giving to any function symbol  $+_\epsilon$  modal signature  $(\epsilon, 1 - \epsilon)$ , where by  $\epsilon$  (resp.  $1 - \epsilon$ ) we mean multiplication by  $\epsilon$  (resp.  $1 - \epsilon$ ). [Mardare et al. \[2016\]](#) have shown that the system obtained from such a signature together with the (equational) rules of idempotency, commutativity, and associativity provides a (quantitative) equational axiomatisation of the (finitary) Wasserstein-Kantorovich distance [[Villani 2008](#)].

## 6.2 Confluence, Part II

Let us now extend the theory of  $\Omega$ -TRSs to  $(\Omega, \Phi)$ -TRSs. The notion of an overlap and of a critical pair straightforwardly extend to  $(\Omega, \Phi)$ -TRSs. Notice, however, that if  $a_1 \xrightarrow{\varepsilon_1} b_1$  and  $a_2 \xrightarrow{\varepsilon_2} b_2$  overlap at position  $p$ , so that there is a substitution  $\sigma$  such that  $a_2^\sigma = C[a_1^\sigma]$  (with  $C = a_2^\sigma[-]_p$ ), then the critical peak is  $b_2^\sigma \xleftarrow{\varepsilon_2} a_1^\sigma \xrightarrow{\partial_C(\varepsilon_2)} C[b_2^\sigma]$ . We thus have all the ingredients to extend [Lemma 1](#) to  $(\Omega, \Phi)$ -TRSs. Compared to its  $\Omega$ -TRS counterpart, however, the critical pair lemma for  $(\Omega, \Phi)$ -TRSs presents a major difference: we can relax the linearity assumption and require rewriting rules to be *left-linear* only.

**Lemma 3** (Critical Pair, Graded). *Let  $\mathcal{R} = (\Sigma, \mapsto_{\mathcal{R}})$  be a left-linear (balanced)  $(\Omega, \Phi)$ -TRS. If  $\mathcal{R}$  is locally confluent on all critical pairs of  $\mathcal{R}$ , then it is locally confluent.*

**PROOF SKETCH.** The proof proceeds as for [Lemma 1](#), the main difference being the case of nested, non-critical redexes. We analyse this case in isolation (this straightforwardly generalise to the case of redexes inside larger terms by structural properties of CBEs). Suppose to have rules  $a_1 \xrightarrow{\varepsilon_1} b_1$ ,  $a_2 \xrightarrow{\varepsilon_2} b_2$ . Without loss of generality, we consider the case in which the second reduction happens inside (an instance) of the first one. So there is a variable  $x$  in  $a_1$  and a substitution instance such that  $x^\sigma$  contains  $a_2^\sigma$ . Since  $\mathcal{R}$  is left linear, we know that there is just one occurrence — say it is at position  $p$  — of  $x$  in  $a_1$ , so that we have  $a_1[x]_p$ . Say also that the relevant occurrence of  $a_2^\sigma$  in  $x^\sigma$  is at position  $q$ , so that  $x^\sigma[a_2^\sigma]_q$  and, consequently,  $a_1^\sigma[x^\sigma[a_2^\sigma]_q]_p$  and  $a_1^\sigma[a_2^\sigma]_{pq}$ . The rule  $a_1 \xrightarrow{\varepsilon_1} b_1$  may, in general, duplicate the single occurrence of  $x$  in  $a_1$ . Say we have  $b_1[x]_{p_1, \dots, p_n}$ , meaning that  $b_1$  has  $n$  occurrences of  $x$ , each at position  $p_i$ . Therefore, reducing  $a_1^\sigma$  gives  $b_1^\sigma[x^\sigma]_{p_1, \dots, p_n}$ , and thus  $b_1^\sigma[a_2^\sigma]_{p_1 q, \dots, p_n q}$ . We can now reduce each of the  $n$  occurrences of  $a_2^\sigma$  in  $b_1^\sigma$ . The distance obtained for each reduction is  $\partial_{p_i q}(b_2^\sigma)(\varepsilon_2)$ . Local confluence now follows from  $\varepsilon_1 \otimes \partial_{pq}(a_1^\sigma)(\varepsilon_2) \leq \bigotimes_i \partial_{p_i q}(b_1^\sigma)(\varepsilon_2) \otimes \varepsilon_1$ , which holds since left-linearity and balanceness entail  $\partial_p(a_1) = \partial_x(a_1) = \partial_x(b_1) = \bigotimes_i \partial_{p_i}(b_1)$ ,  $\partial_{pq}(a_1^\sigma) = \partial_p(a_1) \circ \partial_q(x^\sigma)$  and  $\bigotimes_i \partial_{p_i q}(b_1^\sigma) = \bigotimes_i \partial_{p_i}(b_1) \circ \partial_q(x^\sigma)$ .  $\square$

From [Theorem 1](#) and [Lemma 3](#) we then obtain the following result.

**THEOREM 4.** *Any left-linear and terminating (balanced)  $(\Omega, \Phi)$ -TRS locally confluent on its critical pairs is confluent.*

*Orthogonality.* Even if useful on many  $(\Omega, \Phi)$ -TRSs, [Theorem 4](#) can only be used to infer *local* confluence of non-terminating systems. We now generalise the well-known result [[Rosen 1970](#)]

$$\begin{array}{c}
\frac{}{k \Vdash x \multimap_R x} \quad \frac{\varepsilon \Vdash t \multimap_R s \quad \delta \leq \varepsilon}{\delta \Vdash t \multimap_R s} \\
\\
\frac{(\forall i) \varepsilon_i \Vdash t_i \multimap_R s_i \quad f : (\phi_1, \dots, \phi_n) \in \Sigma_{\mathcal{R}}}{\bigotimes_i \phi_i(\varepsilon_i) \Vdash f(t_1, \dots, t_n) \multimap_R f(s_1, \dots, s_n)} \quad \frac{(\forall i) \delta_i \Vdash v_i \multimap_R w_i \quad \varepsilon \Vdash a \mapsto_R b}{\varepsilon \otimes \bigotimes_i \partial_{x_i}(a)(\delta_i) \Vdash a[\bar{v}/\bar{x}] \multimap_R b[\bar{w}/\bar{x}]}
\end{array}$$

Fig. 1. Multi-step reduction  $\multimap_R$ 

that orthogonality implies confluence to a quantitative and graded setting. Using such a result, we will obtain (quantitative) confluence of system  $\mathcal{W}$  for free.

**Definition 15.** A  $(\Omega, \Phi)$ -TRS is *orthogonal* if it is left-linear and has no critical pair.

Our prime example of an orthogonal  $(\Omega, \Phi)$ -TRS is system  $\mathcal{W}$  of graded combinators. To prove confluence of orthogonal systems we employ Tait and Martin-Löf technique [Aczel 1978; Barendregt 1984], properly instantiated to our rewriting setting. In what follows, we often employ the vector notation  $\bar{\varphi}$  for finite sequences  $\varphi_1, \dots, \varphi_n$  of symbols.

**Definition 16.** Given a  $(\Omega, \Phi)$ -TRS  $\mathcal{R} = (\Sigma, \mapsto_R)$ , we inductively define the multi-step reduction  $\multimap$  by the rules in Figure 1. We define the  $\Omega$ -relation  $\dot{R}$  by  $\dot{R}(t, s) \triangleq \bigvee \{ \varepsilon \mid \varepsilon \Vdash t \multimap_R s \}$ .

We immediately notice that since  $\multimap$  allows us to reduce several redexes in a term simultaneously, it gives the following substitution lemma (whose proof follows the pattern of graded and quantitative substitution lemmas [Dal Lago and Gavazzo 2022b; Gavazzo 2018, 2019]).

**Lemma 4** (Substitution Lemma). *We have the following substitution inequality*

$$\dot{R}(t, s) \otimes \bigotimes_i [\partial_{x_i}(t)] \dot{R}(v_i, w_i) \leq \dot{R}(t[\bar{v}/\bar{x}], s[\bar{w}/\bar{x}]).$$

Given a  $(\Omega, \Phi)$ -TRS  $\mathcal{R} = (\Sigma, \rightarrow_R)$ , we are going to prove confluence of  $R$  by actually proving a stronger result, namely confluence of  $\rightarrow_R$ . To achieve such a result, we shall prove that  $\multimap_R$  has the diamond property. Since  $\rightarrow_R \subseteq \multimap_R \subseteq \rightarrow_R^*$  (and thus  $R \leq \dot{R} \leq R^*$ ), confluence of  $\rightarrow_R$  follows.

**Proposition 4.** *Let  $\mathcal{R} = (\Sigma, \mapsto_R)$  be an orthogonal  $(\Omega, \Phi)$ -TRS. Then, the relation  $\multimap_R$  has the diamond property. That is, if  $s_1 \xleftarrow{\varepsilon_1} t \xrightarrow{\varepsilon_2} s_2$ , there there exists a term  $s$  such that  $s_1 \xrightarrow{\delta_1} s \xleftarrow{\delta_2} s_2$  and  $\varepsilon_1 \otimes \varepsilon_2 \leq \delta_1 \otimes \delta_2$ . Consequently,  $\dot{R}$  has the diamond property.*

**PROOF SKETCH.** The proof is by induction on  $t$  with a case analysis on the defining clauses of  $\multimap$  using the aforementioned substitution lemma and noticing that if we have a (necessarily unique) rule  $a \xrightarrow{\varepsilon} b$  and we reduce a term of the form  $a[\bar{v}/\bar{x}]$ , then either we reduce the (instance) of redex  $a$ , or the term obtained is itself an instance of the redex  $a$ , i.e. it is of the form  $a[\bar{w}/\bar{x}]$ , for some terms  $\bar{w}$ .

□

**THEOREM 5.** *Let  $\mathcal{R} = (\Sigma, \mapsto_R)$  be an orthogonal  $(\Omega, \Phi)$ -TRS. Then,  $R$  is confluent.*

**PROOF SKETCH.** Let  $S \triangleq R^-$ . We prove  $S^*; R^* \leq R^*; S^*$ . By adjunction, it is sufficient to prove  $S^* \leq (R^*; S^*)/R^*$ , which follows by fixed point induction using the general inequality  $P \leq \dot{P} \leq P^*$ . □

Notice that Theorem 5 trivially entails that orthogonal non-expansive systems are confluent. We conclude this section by observing that system  $\mathcal{W}$  is orthogonal, and thus confluent.

**THEOREM 6.** *System  $\mathcal{W}$  of graded combinatory logic is confluent.*

To the best of the authors' knowledge, this is the first confluence result for a system of graded combinators endowed with a quantitative and modal operational (reduction) semantics, and thus it is a first step towards a foundational study of operational properties of coeffectful calculi.

## 7 CONCLUSION, RELATED, AND FUTURE WORK

In this paper, we have started the development of a systematic theory of metric and quantitative rewriting systems. The abstract nature of the notion of distance employed makes our framework robust and allows for several conceptual interpretations of our rewriting systems. The latter, in fact, can be thought not only as metric and quantitative systems, but also as substructural (e.g. fuzzy or monoidal) and modal or coeffectful systems, this way suggesting possible applications of our theory to the development of quantitative and modal operational semantics of coeffectful programming languages.

We have focused on fundamental definitions and confluence properties of abstract and term-based systems. Developing a general theory of quantitative rewriting systems is an ambitious project that cannot be exhausted in a single paper. Among the many possible extensions of the theory presented in this paper, we mention the development of a theory of reduction strategies, the design of completion algorithms for quantitative term rewriting systems (both linear and graded), and the study of inductive and termination properties of quantitative systems. The latter, in particular, seem to suggest that new rewriting properties can be discovered by pushing the quantitative enrichment one step forward, this way making the notions of termination, induction, confluence, etc quantitative themselves. In fact,  $\Omega$ -relations are the enrichment of Boolean relations in a quantale. Since a quantale is itself defined relying on a binary (preorder) relation  $\leq$ , we may push the enrichment one step further, and enrich the preorder of the quantale itself in a (possibly different) quantale. This way, we naturally obtain the aforementioned quantitative notions of induction, termination, etc. Of course, we can iterate the enrichment one step further, i.e. enrich the preorder of the second quantale in a third quantale, and so on and so forth. The result of this process is a relational infinite enrichment that may be used as a base for the development of a general theory of enriched rewriting systems (cf. higher and  $\infty$ -categories [Riehl and Verity 2022]). Connected to that, we notice that by relaxing constraints of quantales, we can give new interpretations to quantitative rewriting systems. For instance, allowing for non-commutative quantales, we can consider the quantale  $2^M$ , with  $M$  non-commutative monoid of rewriting paths and path equalities. We leave the study of such quantales for further investigation.

Finally, an interesting topic for future research is the study of *metric word problems* [Gavazzo and Di Florio 2022], i.e. quantitative refinements of the well-known word problem for equational systems [Bezem et al. 2003]. Contrary to traditional word problems (which ultimately ask for decidability of equality), metric word problems can take several forms. Here are the main ones: (i) The *reachability problem* asks whether the convertibility distance between two elements is not  $\perp$ . Confluence of a rewriting system entails semi-decidability of the reachability problem, whereas confluence and termination gives decidability. (ii) We can strengthen the reachability problem by replacing  $\perp$  with a fixed element of the quantale. Confluence and termination are in general not enough to solve this stronger version of the reachability problem, since looking at the rewriting paths leading to the common normal form (if any) of two objects can give too coarse (over)approximations of their convertibility distance only [Gavazzo and Di Florio 2022] (notice that, assuming the system to be confluent, one can try to obtain better approximations by enlarging the state space and looking at arbitrary common reducts: that, however, does not give decidability either, as there may be infinitely many such reducts. (iii) The *shortest path* is a problem

specific to term-based systems and ask whether the convertibility distance between two terms, which is defined as the join of the distances of (ternary) rewriting paths, is attained by an actual rewriting path. (iv) Connected to the shortest path problem is the *optimal strategy* problem which asks, assuming shortest paths to exist, for a quantitative rewriting strategy to compute such paths.

*Related Work.* To the best of the authors' knowledge, this is the first systematic analysis of quantitative and metric rewriting systems. This, of course, does not mean that isolated forms of quantitative rewriting have not been proposed in the literature. For instance, specific forms of weighted reductions have been employed in the study of cost analysis of rewriting systems [Moser and Schneckenreither 2020; Naaf et al. 2017]. Measured abstract rewriting systems, i.e. *abstract* rewriting systems with a reduction relation enriched in a monoid, have been introduced by van Oostrom and Toyama [2016] to study normalisation properties by random descent. In that context, a quantitative notion of confluence is introduced which, however, differs from ours in the way it compares distances between objects. In fact, given a peak  $b_1 \xleftarrow{\varepsilon_1} a \xrightarrow{\varepsilon_2} b_2$  and a valley  $b_1 \xrightarrow{\delta_1} b \xleftarrow{\delta_2} b_2$ , it is required  $\varepsilon_1 \otimes \delta_1 \leq \varepsilon_2 \otimes \delta_2$  rather than  $\varepsilon_1 \otimes \varepsilon_2 \leq \delta_1 \otimes \delta_2$ . Even if this requirement has a natural reading when it comes to study normalisation properties of rewriting, it seems ineffective when applied to the study of distances and metrics; additionally, it does not fit the algebra of quantitative relations. Moreover, we also remark that measured rewriting systems have been studied in the context of *abstract* rewriting only, whereas our theory of quantitative rewriting covers both abstract and (non-expansive and graded) term-based systems.

At the time of writing, the authors have discovered that *abstract* fuzzy rewriting systems have been studied by Belohlávek et al. [2009, 2010] relying on the theory of fuzzy relations [Belohlávek 2002]. Even if the aforementioned theory of fuzzy rewriting systems does *not* cover term-based systems (neither non-expansive nor graded), the development of fuzzy abstract rewriting systems is in line with our section 4. In particular, Belohlávek et al. [2009, 2010] define fuzzy notions of confluence and prove a quantitative Newman's lemma similar to (the pointwise version of) ours.

Contrary to the case of rewriting systems, general theories of quantitative equational reasoning have been developed following (at least) two lines of work: (quantitative) Fuzzy equational and algebraic theories [Belohlávek and Vychodil 2005] and the more recent quantitative algebras and equational theories by Mardare et al. [2016, 2017] [Bacci et al. 2018, 2021; Mardare et al. 2018, 2021; Mio et al. 2021]. With the exception of the recent work by Dagnino and Pasquali [2022] and Mio et al. [2022], such theories are usually not graded and, to the best of the authors' knowledge, are not capable of describing non-linear systems.

The situation changes if one looks at operational semantics for graded type systems and programming languages, where instrumented *resource-sensitive* operational semantics have been studied. In particular, we mention the weighted abstract machine models by Brunel et al. [2014] and Abel and Bernardy [2020], as well as the recent quantitative heap model by Choudhury et al. [2021] (see also the work by Marshall et al. [2022]). All these models instrument operational semantics in a quantitative fashion relying on additional data structures, such as stacks or heaps, properly decorated with resource annotations. This way, reductions carry explicit information on both resource consumption and production. Many interesting program behaviours (such as noninterference [Abadi et al. 1999] and metric preservation [Reed and Pierce 2010]) can be proved in terms of resource sensitive type safety results which, ultimately, state that programs do not run out of resources during evaluation. Notice that this property reminds of our balanced condition for graded systems. Understanding the exact relationship between the aforementioned operational models and quantitative rewriting systems is an interesting topic for future research.



## REFERENCES

- Martín Abadi, Anindya Banerjee, Nevin Heintze, and Jon G. Riecke. 1999. A Core Calculus of Dependency. In *POPL '99, Proceedings of the 26th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, San Antonio, TX, USA, January 20-22, 1999*. 147–160.
- Andreas Abel and Jean-Philippe Bernardy. 2020. A unified view of modalities in type systems. *Proc. ACM Program. Lang.* 4, ICFP (2020), 90:1–90:28.
- Samson Abramsky. 2002. Predicative Copying and Polynomial Time. *Clifford Lectures, Tulane* (2002).
- Samson Abramsky, Esfandiar Haghighi, and Philip Scott. 2002. Geometry of interaction and linear combinatory algebras. *Mathematical Structures in Computer Science* 12, 5 (2002), 625–665.
- Peter Aczel. 1978. A general Church-Rosser theorem. *Draft, Manchester* (1978).
- A. Arnold and M. Nivat. 1980. Metric Interpretations of Infinite Trees and Semantics of non Deterministic Recursive Programs. *Theor. Comput. Sci.* 11 (1980), 181–205.
- Andrea Asperti and Stefano Guerrini. 1998. *The optimal implementation of functional programming languages*. Cambridge tracts in theoretical computer science, Vol. 45. Cambridge University Press.
- Robert Atkey. 2018. Syntax and Semantics of Quantitative Type Theory. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018*. 56–65.
- Martin Avanzini, Ugo Dal Lago, and Akihisa Yamada. 2020. On probabilistic term rewriting. *Sci. Comput. Program.* 185 (2020).
- Giorgio Bacci, Radu Mardare, Prakash Panangaden, and Gordon Plotkin. 2020. *Foundations of Probabilistic Programming*. Cambridge University Press, Chapter Quantitative Equational Reasoning, 333–360. <https://doi.org/10.1017/9781108770750.011>
- Giorgio Bacci, Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. 2018. An Algebraic Theory of Markov Processes. In *Proc. of LICS 2018*, Anuj Dawar and Erich Grädel (Eds.). ACM, 679–688. <https://doi.org/10.1145/3209108.3209177>
- Giorgio Bacci, Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. 2021. Tensor of Quantitative Equational Theories. In *9th Conference on Algebra and Coalgebra in Computer Science, CALCO 2021, August 31 to September 3, 2021, Salzburg, Austria (LIPIcs, Vol. 211)*, Fabio Gadducci and Alexandra Silva (Eds.). Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 7:1–7:17. <https://doi.org/10.4230/LIPIcs.CALCO.2021.7>
- Roland Carl Backhouse. 2000. Galois Connections and Fixed Point Calculus. In *Algebraic and Coalgebraic Methods in the Mathematics of Program Construction, International Summer School and Workshop, Oxford, UK, April 10-14, 2000, Revised Lectures (Lecture Notes in Computer Science, Vol. 2297)*, Roland Carl Backhouse, Roy L. Crole, and Jeremy Gibbons (Eds.). Springer, 89–148. [https://doi.org/10.1007/3-540-47797-7\\_4](https://doi.org/10.1007/3-540-47797-7_4)
- Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. 2014. Behavioral Metrics via Functor Lifting. In *Proc. of FSTTCS*. 403–415.
- Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. 2015. Towards Trace Metrics via Functor Lifting. In *Proc. of CALCO 2015*. 35–49.
- H.P. Barendregt. 1984. *The lambda calculus: its syntax and semantics*. North-Holland.
- R. Belohlávek. 2002. *Fuzzy Relational Systems: Foundations and Principles*. Springer US.
- Radim Belohlávek, Tomas Kuhr, and Vilém Vychodil. 2009. Confluence and related properties of fuzzy relations. In *FUZZ-IEEE 2009, IEEE International Conference on Fuzzy Systems, Jeju Island, Korea, 20-24 August 2009, Proceedings*. IEEE, 569–574. <https://doi.org/10.1109/FUZZY.2009.5277248>
- Radim Belohlávek, Tomas Kuhr, and Vilém Vychodil. 2010. Confluence and termination of fuzzy relations. *Inf. Sci.* 180, 17 (2010), 3288–3303. <https://doi.org/10.1016/j.ins.2010.04.007>
- Radim Belohlávek and Vilém Vychodil. 2005. *Fuzzy Equational Logic*. Studies in Fuzziness and Soft Computing, Vol. 186. Springer. <https://doi.org/10.1007/b105121>
- Jean-Philippe Bernardy, Mathieu Boespflug, Ryan R. Newton, Simon Peyton Jones, and Arnaud Spiwack. 2018. Linear Haskell: practical linearity in a higher-order polymorphic language. *PACMPL* 2, POPL (2018), 5:1–5:29.
- M. Bezem, J.W. Klop, E. Barendsen, R. de Vrijer, and Terese. 2003. *Term Rewriting Systems*. Cambridge University Press.
- Richard S. Bird and Oege de Moor. 1997. *Algebra of programming*. Prentice Hall.
- M.M. Bonsangue, F. Van Breugel, and J.J.M.M. Rutten. 1998. Generalized Metric Spaces: Completion, Topology, and Powerdomains via the Yoneda Embedding. *Theor. Comput. Sci.* 193, 1-2 (1998), 1–51.
- R.V. Book and F. Otto. 1993. *String-rewriting Systems*. Springer-Verlag.
- Alois Brunel, Marco Gaboardi, Damiano Mazza, and Steve Zdancewicz. 2014. A Core Quantitative Coeffect Calculus. In *Proc. of ESOP 2014*. 351–370.
- Pritam Choudhury, Harley Eades III, Richard A. Eisenberg, and Stephanie Weirich. 2021. A Graded Dependent Type System with a Usage-Aware Semantics. *Proc. ACM Program. Lang.* 5, POPL, Article 50 (jan 2021), 32 pages. <https://doi.org/10.1145/3434331>
- Alonzo Church and J. B. Rosser. 1936. Some Properties of Conversion. *Trans. Amer. Math. Soc.* 39, 3 (1936), 472–482.

- R. Crubillé and U. Dal Lago. 2014. On Probabilistic Applicative Bisimulation and Call-by-Value lambda-Calculi. In *Proc. of ESOP 2014*. 209–228.
- R. Crubillé and U. Dal Lago. 2015. Metric Reasoning about lambda-Terms: The Affine Case. In *Proc. of LICS 2015*. 633–644.
- R. Crubillé and U. Dal Lago. 2017. Metric Reasoning About lambda-Terms: The General Case. In *Proc. of ESOP 2017*. 341–367.
- H.B. Curry and R. Feys. 1958a. *Combinatory Logic*. Number v. 1 in *Combinatory Logic*. North-Holland Publishing Company.
- H.B. Curry and R. Feys. 1958b. *Combinatory Logic*. Number v. 1 in *Combinatory Logic*. North-Holland Publishing Company.
- Francesco Dagnino and Fabio Pasquali. 2022. Logical Foundations of Quantitative Equality. In *Proceedings of the 37th ACM/IEEE Symposium on Logic in Computer Science, LICS 2022*. ACM. <https://doi.org/10.1145/3531130.3533337> to appear.
- Fredrik Dahlqvist and Renato Neves. 2022. An Internal Language for Categories Enriched over Generalised Metric Spaces. In *Proc. of CSL 2022 (LIPICs, Vol. 216)*, Florin Manea and Alex Simpson (Eds.). Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 16:1–16:18. <https://doi.org/10.4230/LIPICs.CSL.2022.16>
- Ugo Dal Lago and Francesco Gavazzo. 2019. Effectful Normal Form Bisimulation. In *Proc. of ESOP 2019*. 263–292.
- Ugo Dal Lago and Francesco Gavazzo. 2020. Differential Logical Relations Part II: Increments and Derivatives. In *Proc. of ICTCS 2020*. 101–114.
- Ugo Dal Lago and Francesco Gavazzo. 2021a. Differential logical relations, part II increments and derivatives. *Theoretical Computer Science* (2021). <https://doi.org/10.1016/j.tcs.2021.09.027>
- Ugo Dal Lago and Francesco Gavazzo. 2021b. Modal Reasoning = Metric Reasoning, via Lawvere. *CoRR* abs/2103.03871 (2021). [arXiv:2103.03871](https://arxiv.org/abs/2103.03871) <https://arxiv.org/abs/2103.03871>
- Ugo Dal Lago and Francesco Gavazzo. 2022a. Effectful Program Distancing. *Proc. ACM Program. Lang.* 6, POPL (2022). <https://doi.org/10.1145/3498680>
- Ugo Dal Lago and Francesco Gavazzo. 2022b. A relational theory of effects and coeffects. *Proc. ACM Program. Lang.* 6, POPL (2022), 1–28. <https://doi.org/10.1145/3498692>
- Ugo Dal Lago, Francesco Gavazzo, and Akira Yoshimizu. 2019. Differential Logical Relations, Part I: The Simply-Typed Case. In *Proc. of ICALP 2019*. 111:1–111:14. <https://doi.org/10.4230/LIPICs.ICALP.2019.111>
- A.A. de Amorim, M. Gaboardi, J. Hsu, S. Katsumata, and I. Cherigui. 2017. A semantic account of metric preservation. In *Proc. of POPL 2017*. 545–556.
- J.W. de Bakker and J.I. Zucker. 1982. Denotational Semantics of Concurrency. In *STOC*. 153–158.
- M.M. Deza and E. Deza. 2009. *Encyclopedia of Distances*. Springer Berlin Heidelberg.
- Henk Doornbos, Roland Carl Backhouse, and Jaap van der Woude. 1997. A Calculational Approach to Mathematical Induction. *Theor. Comput. Sci.* 179, 1-2 (1997), 103–135. [https://doi.org/10.1016/S0304-3975\(96\)00154-5](https://doi.org/10.1016/S0304-3975(96)00154-5)
- W. Du, Y. Deng, and D. Gebler. 2016. Behavioural Pseudometrics for Nondeterministic Probabilistic Systems. In *Proc. of SETTA 2016*. 67–84.
- M.H. Escardo. 1999. A metric model of PCF. In *Workshop on Realizability Semantics and Applications*.
- N. Ferns, P. Panangaden, and D. Precup. 2004. Metrics for Finite Markov Decision Processes. In *Prof. of UAI 2004*. 162–169.
- N. Ferns, P. Panangaden, and D. Precup. 2005. Metrics for Markov Decision Processes with Infinite State Spaces. In *Prof. of UAI 2005*. 201–208.
- Robert C Flagg. 1992. Completeness in continuity spaces. In *AMS Conference proceedings*, Vol. 13. 183–199.
- Peter J. Freyd and Andre Scedrov. 1990. *Categories, allegories*. North-Holland mathematical library, Vol. 39. North-Holland.
- Marco Gaboardi, Shin-ya Katsumata, Dominic A. Orchard, Flavien Breuvar, and Tarmo Uustalu. 2016. Combining effects and coeffects via grading. In *Proc. of ICFP 2016*. 476–489.
- Francesco Gavazzo. 2018. Quantitative Behavioural Reasoning for Higher-order Effectful Programs: Applicative Distances. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018*. 452–461.
- Francesco Gavazzo. 2019. *Coinductive Equivalences and Metrics for Higher-order Languages with Algebraic Effects*. Ph.D. Dissertation. University of Bologna, Italy. <http://amsdottorato.unibo.it/9075/>
- Francesco Gavazzo and Cecilia Di Florio. 2022. Quantitative and Metric Rewriting: Abstract, Non-Expansive, and Graded Systems. *CoRR* abs/2206.13610 (2022). <https://doi.org/10.48550/arXiv.2206.13610>
- D. Gebler, Larsen. K.G., and S. Tini. 2016. Compositional bisimulation metric reasoning with Probabilistic Process Calculi. *Logical Methods in Computer Science* 12, 4 (2016).
- Dan R. Ghica and Alex I. Smith. 2014. Bounded Linear Types in a Resource Semiring. In *Programming Languages and Systems - 23rd European Symposium on Programming, ESOP 2014, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2014, Grenoble, France, April 5-13, 2014, Proceedings*. 331–350.
- D. Gusfield. 1997. *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press.
- P. Hájek. 1998. *Metamathematics of Fuzzy Logic*. Springer Netherlands.
- J.R. Hindley. 1964. *The Church-Rosser Property and a Result in Combinatory Logic*. Ph.D. Dissertation. University of Newcastle-upon-Tyne.

- J.R. Hindley. 2008. *Basic Simple Type Theory*. Cambridge University Press.
- J.R. Hindley and J.P. Seldin. 2008. *Lambda-Calculus and Combinators: An Introduction*. Cambridge University Press.
- D. Hofmann, G.J. Seal, and W. Tholen (Eds.). 2014. *Monoidal Topology: A Categorical Approach to Order, Metric, and Topology*. Number 153 in Encyclopedia of Mathematics and its Applications. Cambridge University Press.
- G  rard P. Huet. 1980. Confluent Reductions: Abstract Properties and Applications to Term Rewriting Systems: Abstract Properties and Applications to Term Rewriting Systems. *J. ACM* 27, 4 (1980), 797–821. <https://doi.org/10.1145/322217.322230>
- Simon L. Peyton Jones. 1987. *The Implementation of Functional Programming Languages*. Prentice-Hall.
- F.W. Lawvere. 1973. Metric spaces, generalized logic, and closed categories. *Rend. Sem. Mat. Fis. Milano* 43 (1973), 135–166.
- Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. 2016. Quantitative Algebraic Reasoning. In *Proc. of LICS 2016*, Martin Grohe, Eric Koskinen, and Natarajan Shankar (Eds.). ACM, 700–709. <https://doi.org/10.1145/2933575.2934518>
- Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. 2017. On the axiomatizability of quantitative algebras. In *Proc. of LICS 2017*. IEEE Computer Society, 1–12. <https://doi.org/10.1109/LICS.2017.8005102>
- Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. 2018. Free complete Wasserstein algebras. *Log. Methods Comput. Sci.* 14, 3 (2018). [https://doi.org/10.23638/LMCS-14\(3:19\)2018](https://doi.org/10.23638/LMCS-14(3:19)2018)
- Radu Mardare, Prakash Panangaden, and Gordon D. Plotkin. 2021. Fixed-Points for Quantitative Equational Logics. In *Proc. of LICS 2021*. IEEE, 1–13. <https://doi.org/10.1109/LICS52264.2021.9470662>
- Daniel Marshall, Michael Vollmer, and Dominic Orchard. 2022. Linearity and Uniqueness: An Entente Cordiale. In *Programming Languages and Systems*, Ilya Sergey (Ed.). Springer International Publishing, Cham, 346–375.
- Stephan Meier-Oeser. 2011. Medieval Semiotics. In *The Stanford Encyclopedia of Philosophy* (Summer 2011 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.
- Matteo Mio, Ralph Sarkis, and Valeria Vignudelli. 2021. Combining Nondeterminism, Probability, and Termination: Equational and Metric Reasoning. In *Proc. of LICS 2021*. IEEE, 1–14. <https://doi.org/10.1109/LICS52264.2021.9470717>
- Matteo Mio, Ralph Sarkis, and Valeria Vignudelli. 2022. Beyond Nonexpansive Operations in Quantitative Algebraic Reasoning. *CoRR* abs/2201.09087 (2022). <https://arxiv.org/abs/2201.09087>
- Georg Moser and Manuel Schneckenreither. 2020. Automated amortised resource analysis for term rewrite systems. *Sci. Comput. Program.* 185 (2020). <https://doi.org/10.1016/j.scico.2019.102306>
- J.R. Munkres. 2000. *Topology*. Prentice Hall, Incorporated.
- Matthias Naaf, Florian Frohn, Marc Brockschmidt, Carsten Fuhs, and J  rgen Giesl. 2017. Complexity Analysis for Term Rewriting by Integer Transition Systems. In *Proc. of FroCoS 2017 (Lecture Notes in Computer Science, Vol. 10483)*, Clare Dixon and Marcelo Finger (Eds.). Springer, 132–150. [https://doi.org/10.1007/978-3-319-66167-4\\_8](https://doi.org/10.1007/978-3-319-66167-4_8)
- M. H. A. Newman. 1942. On Theories with a Combinatorial Definition of "Equivalence". *Annals of Mathematics* 43, 2 (1942), 223–243.
- Dominic Orchard, Vilem-Benjamin Liepelt, and Harley Eades III. 2019. Quantitative Program Reasoning with Graded Modal Types. *Proc. ACM Program. Lang.* 3, ICFP, Article 110 (2019), 110:1–110:30 pages.
- Tomas Petricek, Dominic A. Orchard, and Alan Mycroft. 2014. Coeffects: a calculus of context-dependent computation. In *Proc. of ICFP 2014*. 123–135.
- J. Reed and B.C. Pierce. 2010. Distance makes the types grow stronger: a calculus for differential privacy. In *Proc. of ICFP 2010*. 157–168.
- Emily Riehl and Dominic Verity. 2022. *Elements of  $\infty$ -Category Theory*. Cambridge University Press. <https://doi.org/10.1017/9781108936880>
- Barry K. Rosen. 1970. Tree-Manipulating Systems and Church-Rosser Theorems. In *Proceedings of the 2nd Annual ACM Symposium on Theory of Computing, May 4-6, 1970, Northampton, Massachusetts, USA*. ACM, 117–127. <https://doi.org/10.1145/800161.805157>
- K.I. Rosenthal. 1990. *Quantaes and their applications*. Longman Scientific & Technical.
- Richard Routley and Robert K. Meyer. 1972a. The semantics of entailment - II. *J. Philos. Log.* 1, 1 (1972), 53–73.
- Richard Routley and Robert K. Meyer. 1972b. The semantics of entailment - III. *J. Philos. Log.* 1, 2 (1972), 192–208.
- Richard Routley and Robert K. Meyer. 1973. The Semantics of Entailment. In *Truth, Syntax and Modality*, Hugues Leblanc (Ed.). Studies in Logic and the Foundations of Mathematics, Vol. 68. Elsevier, 199 – 243.
- D. Sands. 1998. Improvement Theory and Its Applications. In *Higher Order Operational Techniques in Semantics*, A. D. Gordon and A. M. Pitts (Eds.). Cambridge University Press, 275–306.
- Gunther Schmidt. 2011. *Relational Mathematics*. Encyclopedia of Mathematics and its Applications, Vol. 132. Cambridge University Press.
- M  iche  l    Searc  id. 2006. *Metric Spaces*. Springer London.
- L.A. Steen and J.A. Seebach. 1995. *Counterexamples in Topology*. Dover Publications.
- M. H. Stone. 1949. Postulates for the barycentric calculus. *Ann. Mat. Pura Appl. (4)* 29, 1 (1949), 25–30. <https://doi.org/10.1007/BF02413910>

- Isar Stubbe. 2014. An introduction to quantaloid-enriched categories. *Fuzzy Sets Syst.* 256 (2014), 95–116. <https://doi.org/10.1016/j.fss.2013.08.009>
- Alfred Tarski. 1941. On the Calculus of Relations. *J. Symb. Log.* 6, 3 (1941), 73–89. <https://doi.org/10.2307/2268577>
- A. Thue. 1914. *Probleme über Veränderungen von Zeichenreihen nach gegebenen Regeln*. J. Dybwad.
- Alasdair Urquhart. 1972. Semantics for Relevant Logics. *J. Symb. Log.* 37, 1 (1972), 159–169.
- Vincent van Oostrom and Yoshihito Toyama. 2016. Normalisation by Random Descent. In *Proc. of FSCD 2016 (LIPIcs, Vol. 52)*, Delia Kesner and Brigitte Pientka (Eds.). Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 32:1–32:18. <https://doi.org/10.4230/LIPIcs.FSCD.2016.32>
- C. Villani. 2008. *Optimal Transport: Old and New*. Springer Berlin Heidelberg.