Tracelet Hopf algebras and decomposition spaces

Nicolas Behr
Université de Paris, CNRS, IRIF
nicolas.behr@irif.fr

Joachim Kock
Universitat Autònoma de Barcelona &
Centre de Recerca Matemàtica
kock@mat.uab.cat

Tracelets are the intrinsic carriers of causal information in categorical rewriting systems. In this work, we assemble tracelets into a symmetric monoidal decomposition space, inducing a cocommutative Hopf algebra of tracelets. This Hopf algebra captures important combinatorial and algebraic aspects of rewriting theory, and is motivated by applications of its representation theory to stochastic rewriting systems such as chemical reaction networks.

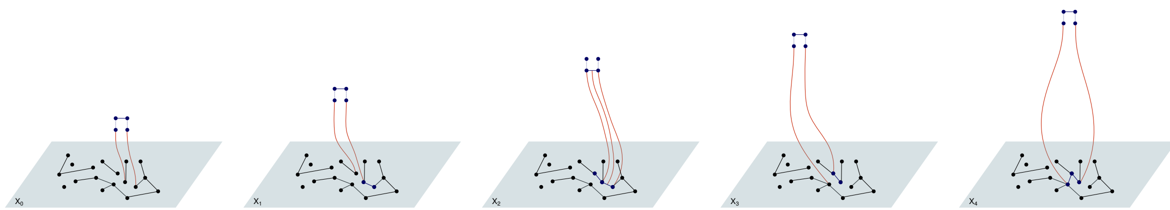
1 Introduction

Double-Pushout (DPO) [13] and more generally compositional categorical rewriting systems [1, 7] provide a versatile and mathematically sound framework for modeling complex transition systems, with a paradigmatic example the modeling of reaction systems in biochemistry [10] and in organic chemistry [6]. The specification of an individual rewriting operation (*direct derivation*) in essence amounts to providing a rewrite rule, i.e., a span of monomorphisms, that acts as a sort of template for the operation, together with a *match*, which permits to specify the location within a host object where the local replacement operation is to be performed. In practical applications, it is often the case that the rewriting rules themselves involve only comparatively small graph-like objects. In contrast, the host objects to which the rewrites are applied could easily be several orders of magnitude larger, so that an enormous number of matches may be possible for a given rule and a given host object.

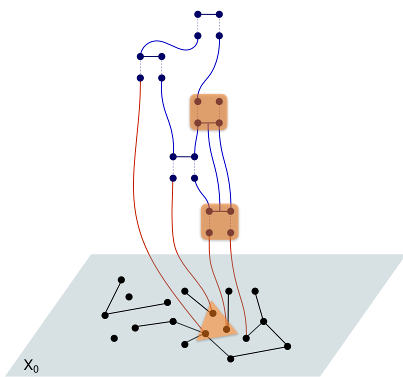
A natural and powerful approach to overcome this fundamental problem consists in focusing on the combinatorial, statistical and structural properties of *interactions* of rewriting rules within derivation traces, and to aim for a classification of traces in terms of "interaction patterns". Unlike in compositional diagrammatic calculi such as in particular the theory of string diagrams, the key obstacle for such a type of analysis in rewriting theories resides in the fact that two given rules may in general interact in a multitude of ways, i.e., there does not exist a notion of deterministic rule composition. Instead, as first demonstrated in [4] and further developed in [1, 9, 6, 7], it is necessary to define a notion of non-deterministic rule composition via a form of recursive application of the concurrency theorem, which then indeed permits to define tractable methods to reason statically about classes of rule compositions. Taking inspiration from the notion of *pathways* in chemical reaction systems, this approach was then further refined in [2] to the notion of *tracelets*, which in essence act as the carriers of causal information in derivation sequences.

The main objective of the present paper consists in establishing a principled mathematical approach to formalize the *combinatorics* of tracelets. Generalizing results of [5] on rewriting systems over directed multi-graphs to the categorical rewriting theory setting, we will demonstrate that it is indeed the notion of *combinatorial Hopf algebras* that naturally captures the rich structure of tracelets. Apart from a rewriting-theoretic construction of the Hopf algebras (Section 5), we report on our original discovery that this at first sight seemingly ad hoc construction is in fact interpretable in terms of the theory of *decomposition spaces* (Sections 3 and 4). Our motivation for this approach has been the analogy between

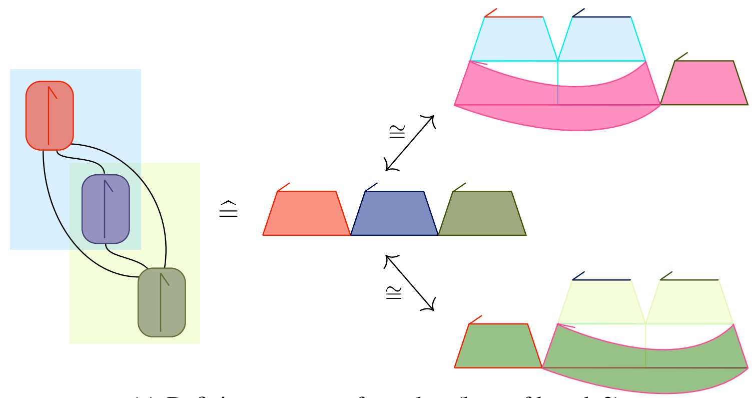
Submitted to:
ACT 2021



(a) A direct derivation sequence of length 5 (with edge creation/deletion rules, and where“wires” indicate matches).



(b) Tracelet and shift equivalence example.



(c) Defining property of tracelets (here of length 3).

Figure 1: An illustration of graph rewriting sequences (top) and of the tracelet picture (bottom).

the inductive definition of tracelets and the decomposition-space axioms in homotopy combinatorics. The benefit of this ‘detour’ is to situate the tracelet Hopf algebra in a general framework covering most combinatorial Hopf algebras, thereby exhibiting the constructions and proofs as general ideas.

Decomposition spaces were introduced in combinatorics by Gálvez, Kock and Tonks [16], [17] as a far-reaching homotopical generalization of posets for the purpose of incidence algebras and Möbius inversion, and independently by Dyckerhoff and Kapranov [12] in homological algebra and representation theory, motivated mainly by the Waldhausen S-construction and Hall algebras. Classically, since the work of Rota [22] in the 1960s, incidence algebras were defined from posets through the process of decomposing intervals. The construction can fruitfully be formulated in terms of the nerve of a poset. However, a great many combinatorial co- and Hopf algebras are not of incidence type, meaning that they are not the incidence coalgebra of a poset. The basic observation of [16] is that simplicial objects more general than nerves of posets admit the incidence coalgebra construction, allowing most of the basic features of the theory to carry over — and that most combinatorial Hopf algebras do arise as incidence Hopf algebras of (monoidal) decomposition spaces.

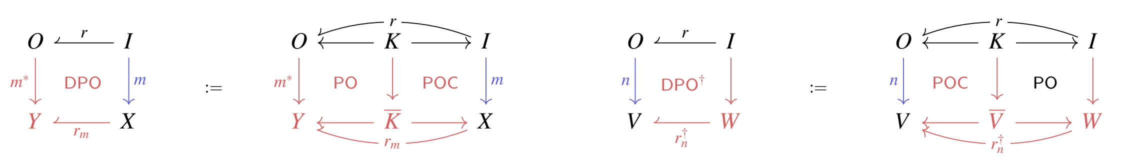
2 Double-Pushout rewriting and tracelet theory

Consider a rewriting system of undirected multi-graphs such as the one depicted in Figure 1(a), with some elementary rules that link and unlink vertices with edges. Starting from some initial graph X_0, each rewriting step (*direct derivation*) consists in choosing a rewriting rule together with an occurrence (*match*) of the input motif (drawn at the bottom of the rule diagrams) within the graph that is rewritten. It is intuitively clear that sequences of rewrites (*derivation traces*) have a rich intrinsic structure, since the graph-like nature of the input and output interfaces of the rules as well as the internal structure of the rules entail that the nature of *interactions* between sequential rewriting steps is not easily encapsulated in any simple form of composition structure. To wit, consider the diagram in Figure 1(b), which provides a sort of movie-script depiction of the five-step derivation trace depicted in the top part of the figure. For clarity, red wires are used to indicate inputs to rules that are present in the original configuration X_0, while blue wires indicate inputs to rules that have originated from outputs of preceding rules.

The main purpose of the theory of *tracelets* [2] then consists in rendering mathematically precise the meaning of this intuitive picture. In particular, according to the *Tracelet Characterization Theorem* [2, Thm. 2], each derivation trace of length n is uniquely characterized by a *tracelet* of length n (cf. the sub-diagram consisting of the five rules and all blue wires in Figure 1(b)) and a *match* of the tracelet into the initial configuration X_0 (depicted as red "wires" in Figure 1(b)). Crucially, the compositional structure of tracelets offers a form of statics causal analysis via algebraic relations such as commutator relations. This type of analysis takes advantage of algebraic relations such as *shift equivalence*, which in the example of Figure 1(b) amounts to the observation that the rules in the boxes highlighted in orange may be freely moved "along the wires" so as to exchange their order (i.e., without changing the overall effect of the rewriting sequence). Finally, as sketched via the highlighted triangle pattern that is produced via the rewriting sequence in the example, tracelets permit to statically reason about the combinatorics of pattern-counting problems in an efficient manner (cf. [3] for a prototype of such an analysis in the setting of counting patterns in planar rooted binary trees).

As depicted in Figure 1(c), a tracelet of length 3 exhibits already quite a non-trivial compositional structure, in that as sketched the internal structure of partial overlaps of rule inputs and outputs in such a tracelet is not inherently of a purely sequential nature; to wit, the diagram encodes a special kind of trace of length 3, with the defining property that it may be equivalently (up to isomorphisms) be obtained via nested composition operations. It is in this particular sense that tracelets offer a *minimal* causal presentation of the structure of rewriting sequences, since via the equivalences to nested pairwise composition operations, they permit to efficiently express n-step sequences as just a special type of derivation sequences.

2.1. Categorical rewriting theory. For simplicity, we will focus here on the variant of tracelet theory for so-called *double-pushout (DPO) rewriting*, and for rewriting rules without application conditions (referring to [2] for the general[1] theory). Throughout this paper, let $\mathbf{C}$ denote an *adhesive category* [20] that is *finitary* (in the sense of [11], i.e., with only finitely many subobjects for each object up to isomorphisms), and that possesses a strict initial object $\varnothing \in \mathsf{obj}(\mathbf{C})$. *Rewriting rules* are defined as spans of monomorphisms $r = (O \leftarrow o - K - i \rightarrow I)$, also denoted for brevity as $r = (O \leftharpoonup I)$. In the tradition of rewriting theory, we refer to such rules as *linear rules* (with "linear" referring to the nature of the span as a span of monos), and denote the class of all such rules as $\mathsf{Lin}(\mathbf{C})$. For every $r \in \mathsf{Lin}(\mathbf{C})$ and object $X \in \mathsf{obj}(\mathbf{C})$, let $\mathcal{M}_r(X)$ denote the *set of (DPO-admissible) matches of r into X*, where $m \in \mathcal{M}_r(X)$ iff m is a monomorphism and the *pushout complement* marked POC in the diagram below left exists:



(1)

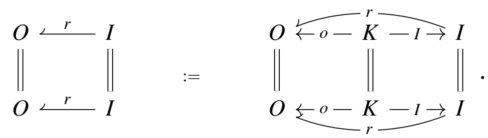
Note that in an adhesive category, pushouts along monomorphisms (here marked PO) are guaranteed to exist; in contrast, pushout complements (here marked POC) may fail to exist, since not every composable

[1] Available generalizations include the type of rewriting (with *Sesqui-Pushout* semantics an alternative option), the choice of base-categories (with $\mathcal{M}$-*adhesive categories* [14] a more general option) as well as the inclusion of constraints and application conditions into the compositional rewriting semantics (cf. also [7]).

pair of arrow can be completed into a pushout square. Moreover, r_m as well as $Y = r_m(X)$ are evidently only defined up to universal isomorphisms. It is customary to refer to the data of the aforementioned diagram as a *direct derivation*. For later convenience, taking advantage of the symmetry of the definition, we mark by $\mathsf{DPO}^\dagger$ diagrams that arise as DPO-type direct derivations in the "opposite direction", i.e., "against" the direction of rules (here, from left to right; cf. the diagram above right).

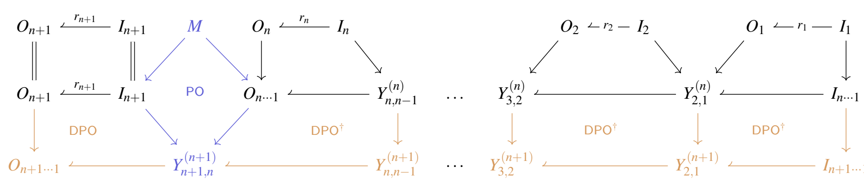
2.2. Tracelets. The class of *tracelets* $\mathcal{T}$ for DPO-type type rewriting over $\mathbf{C}$ is defined recursively:

- *Tracelets of length* 1: for every rule $r = (O \leftharpoonup I) \in \mathsf{Lin}(\mathbf{C})$, define $T(r) \in \mathcal{T}_1$ as a diagram



(2)

- *Tracelets of length* $n+1$*:* denoting by $\mathcal{T}_n$ (for $n > 1$) the class of trecelets of length n, the class $\mathcal{T}_{n+1}$ is defined to consist of diagrams as below, where $r_k = (O_k \leftharpoonup I_k) \in \mathsf{Lin}(\mathbf{C})$ are linear rules (for $k = 1, \dots, n+1$), where the top right part of the diagram encodes a tracelet $T \in \mathcal{T}_n$ of length n, where $\mu = (I_{n+1} \leftarrow M \rightarrow O_{n\cdots 1})$ is a span of monos, with the cospan $I_{n+1} \rightarrow Y^{(n+1)}_{n+1,n} \leftarrow O_{n\cdots 1}$ its pushout, and such that the direct derivations marked DPO and $\mathsf{DPO}^\dagger$ exist:



(3)

Then this data defines a tracelet $T(r_{n+1}) \,{}^{\mu}\!\angle\, T$ of length $n+1$ (the *tracelet composition* of $T(r_{n+1})$ with T along μ) uniquely up to universal isomorphisms (see comments below) as

$T(r_{n+1}) \,{}^{\mu}\!\angle\, T :=$

(4)

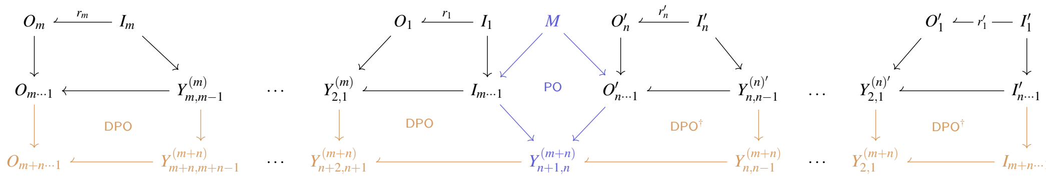
For later convenience, we will sometimes speak of the commutative square below rule r_i in a tracelet as the *i-th plaquette*. Let $\mathcal{T} := \cup_{n\geq 1}\mathcal{T}_n$ denote the class of all (finite length) tracelets. For later convenience, we also introduce the notations $\mathsf{in}(T) := I_{n\cdots 1}$ ("input interface" of $T \in \mathcal{T}_n$), $\mathsf{out}(T) := O_{n\cdots 1}$ ("output interface" of $T \in \mathcal{T}_n$), $\mathsf{MT}_{T(r_{n+1})}(T)$ (for "matches", i.e., admissible partial overlaps μ of the length 1 tracelet $T(r_{n+1})$ with the tracelet $T \in \mathcal{T}_n$) and $[[T]]$ for the so-called *evaluation* of the tracelet T, which for $T \in \mathcal{T}_n$ is defined with notations as in the top right of (4) as

$$[[T]] := (O_{n\cdots 1} \leftharpoonup I_{n\cdots 1}) = (O_{n\cdots 1} \leftharpoonup Y^{(n)}_{n,n-1}) \circ \cdots \circ (Y^{(n)}_{2,1} \leftharpoonup I_{n\cdots 1})\,. \tag{5}$$

Here, $\circ$ denotes the operation of *span composition* (considered up to span-isomorphisms).

Up to this point, one might say that DPO-type tracelets are some form of data structure that encodes a certain form of *sequential compositions* of rewriting rules. This point of view is augmented via the following definition, which finally reveals tracelets as a particular notion of compositional diagrams.

2.3. Tracelet composition. Let $T \in \mathcal{T}_m$ and $T' \in \mathcal{T}_n$ be two tracelets of length m and n, respectively (for $m, n > 0$). Let $\mu := (I_{m\cdots 1} \leftarrow M \to O'_{n\cdots 1})$ be a partial overlap (of the "input interface" $I_{m\cdots 1}$ of T with the "output interface" $O'_{n\cdots 1}$ of T') whose pushout $I_{m\cdots 1} \to Y^{(m+n)}_{n+1,n} \leftarrow O'_{n\cdots 1}$ satisfies that in the diagram below, all direct derivations marked DPO and DPO†, respectively, exist:



(6)

In this case, we write $\mu \in \mathsf{MT}_T(T')$ to say that μ is a(n DPO-admissible) *match* of T into T', and we denote the *composition* of T with T' along μ as

$$T \,{}^{\mu}\!\angle\, T' :=$$


(7)

The definition of tracelets and their composition might appear somewhat ad hoc at first sight, yet it is very natural if viewed in diagrammatic form. To this end, consider the example of a tracelet of length 3 such as in Figure 1(c). The "wires" in the schematic diagram that link individual length 1 tracelets encode the partial overlaps; as indicated, the tracelet of length 3 may be realized recursively via either determining the partial overlap of the first and the second sub-tracelet, composing, and then determining the resulting overlap of the composite tracelet of length 2 with the third tracelet of length 1, or (equivalently as it will turn out) computing the composition of the third and second tracelets of length 1, and of that composite with the first tracelet of length 1. This so-called *associativity property* of tracelet compositions is at the heart of the algebraic properties of tracelets. It will be further illustrated when we now pass to discuss tracelets in the framework of decomposition spaces.

3 The decomposition space of rewrite rules

In this section we describe a decomposition space $\mathbf{X}_\bullet$ of rewrite rules (for a fixed rewrite system in a fixed adhesive category $\mathbf{C}$ as above), whose incidence algebra is the rule algebra.

3.1. Decomposition spaces. A decomposition space [12], [16] is a simplicial groupoid $X_\bullet : \mathbb{\Delta}^{\mathrm{op}} \to \boldsymbol{Grpd}$ satisfying a certain exactness property designed precisely to allow the incidence coalgebra construction, classically defined for posets. (The nerve of a poset or a category is an example of a decomposition space. Where categories encode composition, decomposition spaces owe their name to encoding decomposition.) Many situations where compositionality is hard to achieve can be dealt with instead with decompositions, as is often the case in combinatorics, where combinatorial structures can be split into smaller ones without the ability to compose [18], [15]. Often non-deterministic composition structures can be turned around and constitute instead a decomposition.

There are different ways to formulate the decomposition-space axioms. One (simplified) version states that for any endpoint-preserving monotone map $\alpha : [2] \twoheadrightarrow [n]$, defining a decomposition of any

n-simplex into an n_1-simplex and an n_2-simplex, the natural square

$$\begin{array}{ccc} X_n & \longrightarrow & X_{n_1} \times X_{n_2} \\ \big\downarrow & \lrcorner & \big\downarrow \text{long edges} \\ X_2 & \underset{\text{short edges}}{\longrightarrow} & X_1 \times X_1 \end{array} \tag{8}$$

is a pullback. It says that an n-simplex can be reconstructed from the two smaller simplices of the decomposition together with the information of a gluing of the the long edges of the two simplices onto the short edges of a base 2-simplex.

This condition is considerably weaker than the Segal condition (which characterizes categories, hence composition rather than just decomposition), which says that a single-vertex overlap between the two smaller simplices is enough to glue. In the decomposition-space case, the base 2-simplex is required as a kind of context for the gluing.

3.2. Groupoids of tracelets. An *isomorphism* between two tracelets of length n is by definition a family of object-wise isomorphisms between the involved objects in $\mathbf{C}$ making all squares commute. We denote by $\mathbf{X}_n$ the groupoid of all tracelets of length n. (In particular, the only tracelet of length 0 is the empty one (which evaluates to the trivial rule), so $\mathbf{X}_0 = \{*\}$.)

Theorem 3.3. *The groupoids $\mathbf{X}_n$ assemble into a simplicial groupoid $\mathbf{X}_\bullet : \mathbb{\Delta}^{\mathrm{op}} \to$ **Grpd** (whose face and degeneracy maps we proceed to describe below).*

To describe the simplicial structure, we give the standard presentation in terms of face and degeneracy maps, so that the data is a diagram in the category of groupoids of shape

$$\{*\} \; \underset{\leftarrow d_0}{\overset{\leftarrow d_1}{\xrightarrow{s_0}}} \; \mathbf{X}_1 \; \begin{smallmatrix} \leftarrow d_2 \\ \xrightarrow{s_1} \\ \leftarrow d_1 \\ \xrightarrow{s_0} \\ \leftarrow d_0 \end{smallmatrix} \; \mathbf{X}_2 \; \begin{smallmatrix} \leftarrow d_3 \\ \xrightarrow{s_2} \\ \leftarrow d_2 \\ \xrightarrow{s_1} \\ \leftarrow d_1 \\ \xrightarrow{s_0} \\ \leftarrow d_0 \end{smallmatrix} \; \mathbf{X}_3 \qquad \cdots \tag{9}$$

3.4. Description of the face maps. The *top face map* $d_n : \mathbf{X}_n \to \mathbf{X}_{n-1}$ (resp. the *bottom face map* $d_0 : \mathbf{X}_n \to \mathbf{X}_{n-1}$) is defined via (1) performing *tracelet surgery* (cf. Corollary A.7) to exhibit the tracelet as a composition of the first (the last) rule with an $(n-1)$-tracelet, followed by (2) extracting the $(n-1)$-length tracelet. This is illustrated in Figure 1(c) for the case of $n = 3$, with d_0 (d_3) defined to return the tracelet shaded in light blue (in light yellow).

Recall from Section 2.2 that a *plaquette in position i* of a given tracelet is defined as the i-th direct derivation, i.e., the commutative subdiagram of the tracelet involving the i-th rule (read from the right). Then the *inner face maps* $d_i : \mathbf{X}_n \to \mathbf{X}_{n-1}$ (for $0 < i < n$) replace the two plaquettes p_i and p_{i+1} in the chain with a single new plaquette p' with the same starting point as p_i and the same endpoint as p_{i+1}, by applying the "synthesis" part of the concurrency theorem (cf. Theorem A.2) to convert the sub-sequence of plaquettes p_{i+1} after p_i into a one-step direct derivation along the composite rule $p'_{i+1,i}$. The result of this operation is guaranteed to be a tracelet of length $n-1$. Referring once again Figure 1(c) for an illustration of the case $n = 3$, d_1 (d_2) are defined to return the tracelet shaded in green (in pink).

The *degeneracy maps* $s_i : \mathbf{X}_n \to \mathbf{X}_{n+1}$ (for $0 \le i \le n$) insert a copy of the trivial rule $\emptyset \leftarrow \emptyset \to \emptyset$ in the tracelet at position i.

In the form stated, $\mathbf{X}_\bullet$ is only a pseudo-simplicial groupoid. This means that the simplicial identities only hold up to (specified) isomorphism, and that there are coherence issues to deal with. The reason for this pseudo-ness is that composition of rules and tracelets, as involved in the face maps, is only well defined up to isomorphism, relying as it does on pushouts and pullbacks. To actually get well-defined face maps, it is necessary to make choices of these universal constructions, and these choices screw up the strict simplicial identities. (A well-known example of this phenomenon is how composition of spans by means of pullbacks defines a bicategory, not an ordinary category.)

This pseudo-ness is not at all a problem for the sake of decomposition-space theory, designed to be up to homotopy, and it does not affect the incidence algebra we construct from this decomposition space (which in any case is spanned by iso-classes of rewrite rules). Nevertheless it is very fruitful to provide also a strict model of $\mathbf{X}_\bullet$. The standard technique for constructing this (which goes back to insight from algebraic topology from the 1970s (notably Quillen,[2] Waldhausen, and Segal)) is to beef up the groupoid of n-simplices to something equivalent that contains all the (redundant) data involved in the face maps.

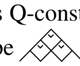
Specifically, a 2-simplex should not just be a 2-tracelet, but rather a 2-tracelet *together* with a choice of composite rule. In this way the middle face map $d_1 : \mathbf{X}_2 \to \mathbf{X}_1$ does not have to compute any composite by means of choices; it can simply return the choice already built in. The fact that these choices are unique up to universal isomorphisms say precisely that this bigger groupoid is equivalent to the original, and hence that the homotopy properties of the bigger simplicial groupoids are the same. In Figure 2(b) we see such a fully specified 2-simplex. The two short edges (01 and 12) are the two rules in a 2-tracelet, and the squares marked PO and POC are the plaquettes constituting altogether the 2-tracelet. The pullback square (blue, marked as PB) is not part of the data of the tracelet, but it is included in the fully specified notion of 2-simplex.

In degree 3 we arrive at the first point where there is an interesting simplicial identity to establish, namely commutativity of the square

$$\begin{array}{ccc} \mathbf{X}_3 & \xrightarrow{d_1} & \mathbf{X}_2 \\ {\scriptstyle d_2}\downarrow & & \downarrow{\scriptstyle d_1} \\ \mathbf{X}_2 & \xrightarrow[d_1]{} & \mathbf{X}_1, \end{array} \tag{10}$$

which in essence states that a sequential composition of three rules may be recovered equivalently from two steps of pairwise rule compositions in either of the nesting orders.

For the groupoids of bare tracelets, this simplicial identity cannot be strict, due to the choices of pushouts and pullbacks involved in composition of rules and tracelets. That the equation holds up to natural isomorphism is a nontrivial statement which involves the concurrency theorem (in the particular form called associativity theorem [9, 4, 5]). We explain how the same theorem implies the strict equation for the fully specified 3-simplices. This exhibits the beautiful geometry inherent in the associativity theorem. As always, the idea is that a fully specified 3-simplex should contain all information about all choices. In particular (in order for the four face maps to be forgetful) it should contain four 2-simplices of the form of Figure 2(b). A full picture of such a subdivided tetrahedron is given in Figure 2(e). One can chase through how this is built up from composition of tracelets, over specified overlaps: Consider the diagram depicted in Figure 2(c), which is formed by (1) a 2-simplex encoding a composition of two rules r_{21} and r_{10} into some rule r_{20}, and (2) another 2-simplex of which one "short edge" is the rule r_{20}, and which contains another rule r_{32} and the data of the composition of r_{32} with r_{20} into some rule r_{30}. Upon closer inspection, it is possible (via a number of somewhat intricate steps) to construct

[2] Historical remark: Bénabou (1963) had described a bicategory spans. Quillen used the techniques of big redundant n-simplices to exhibit the same structure as a *strict* simplicial groupoid, now called Quillen's Q-construction. Instead of having simply chains of n composable spans in degree n, he defined it to be diagrams of shape , that is composable spans, *together* with all the relevant pullbacks. Similar constructions were given in related situation by Waldhausen and Segal, and today the technique is standard in algebraic topology.

from this data the interior and the other two faces of a tetrahedron. To this end, one first invokes the "analysis" part of the DPO-type concurrency theorem (cf. Theorem A.2 in Appendix A) in order to obtain from the sub-diagram that encodes the one-step direct derivation of the object I_{03} along the composite rule $O_{20} \leftarrow K_{20} \rightarrow K_{20}$ the data of a sequence of two direct derivations along the "constituent" rules $O_{21} \leftarrow K_{21} \rightarrow I_{21}$ after $O_{10} \leftarrow K_{10} \rightarrow I_{10}$. This construction in particular delivers an object Z located in the interior of the tetrahedron. Over several further steps (involving pushout and pullback operations), it is then possible to fill the remaining two faces of the 3-simplex with the structure of two sequential rule compositions, ultimately resulting in the diagram of Figure 2(e). The fact that all these constructions are given by universal properties (pushouts and pullbacks, together with the axioms of adhesive categories), ensures that the groupoid of such fully specified 3-simplices is equivalent to the groupoid of bare 3-tracelets. The face maps are now obvious (or even tautological) and all the simplicial identities are clearly strict for this reason: they merely return data already contained in (the beefed-up version of) $\mathbf{X}_3$.

The higher simplices are increasingly cumbersome to describe, due to our limited vision of geometry in dimension higher than 3, but the principle is easy to follow: just include all information about all possible composites, and the overall geometric shape is always a geometric n-simplex whose edges are rules, whose 2-dimensional faces are as in Figure 2(b) and whose 3-dimensional faces are as in Figure 2(e).

The fact that in each dimension the bare tracelets contain information necessary and sufficient to reconstruct the full specified simplex is an expression of the the central result of [2] that it is indeed tracelets that provide the minimal carriers of causal information in sequential rule compositions.

We proceed to establish that $\mathbf{X}_\bullet$ is a decomposition space. Since this is a homotopy invariant property, we may work with the simple version of groupoids of n-tracelets. Before the check, let us just note that $\mathbf{X}_\bullet$ is not a Segal space (a category), because of the non-deterministic nature of composition. Specifically, a 2-simplex cannot be reconstructed from knowing its two short edges.

Theorem 3.5. $\mathbf{X}_\bullet$ *is a decomposition space. This means that for all* $0 < i < n$ *the two squares*

$$\begin{array}{ccc} \mathbf{X}_{n+1} & \xrightarrow{d_{n+1}} & \mathbf{X}_n \\ {\scriptstyle d_i}\downarrow & & \downarrow{\scriptstyle d_i} \\ \mathbf{X}_n & \xrightarrow[d_n]{} & \mathbf{X}_{n-1} \end{array} \qquad\qquad \begin{array}{ccc} \mathbf{X}_{n+1} & \xrightarrow{d_0} & \mathbf{X}_n \\ {\scriptstyle d_{i+1}}\downarrow & & \downarrow{\scriptstyle d_i} \\ \mathbf{X}_n & \xrightarrow[d_0]{} & \mathbf{X}_{n-1} \end{array} \tag{11}$$

are (homotopy) pullbacks.

To check this, it is enough to show that the fibers of the maps pictured vertically are equivalent. We shall see that indeed all fibers of inner face maps are canonically identified with the fiber of $d_1 : \mathbf{X}_2 \rightarrow \mathbf{X}_1$).

3.6. Fiber calculations. Consider $d_1 : \mathbf{X}_2 \rightarrow \mathbf{X}_1$ which sends a pair of composable rules with minimal gluing (r_2, w, r_1) to the composite rule r'. The fiber over $r' \in \mathbf{X}_1$ is thus the groupoid of all (r_2, w, r_1) that compose to r'. We denote this groupoid $(\mathbf{X}_2)_{r'}$. Notice that the objects of $\mathbf{X}_2$ are composable pairs of plaquettes with the property that the intermediate point between the two plaquettes is a minimal gluing (of the output of rule r_1 with the input of rule r_2; said otherwise, the middle cospan in the two-step direct derivation sequence is a pushout of its own pullback).

Lemma 3.7. *The (homotopy) fiber of* $d_i : \mathbf{X}_n \rightarrow \mathbf{X}_{n-1}$ *(for* $0 < i < n$*) over a tracelet which in position* i *has a plaquette with rule* r' *is equivalent to the groupoid* $(\mathbf{X}_2)_{r'}$*. In particular, it does not depend on the whole plaquette* p' *under* r'*, and it does not depend on the context in any way. (*Proof: *cf. Appendix B.1)*

One can now unpack the general construction of incidence algebras of decomposition spaces (cf. Appendix C.4) to establish:

Proposition 3.8. *The incidence algebra is the rule algebra of [9].*

This algebra is not our main focus in this work. Rather do we regard the decomposition space $\mathbf{X}_\bullet$ as a stepping stone towards more interesting decomposition spaces and Hopf algebras, notably the tracelet Hopf algebra.

4 Decomposition spaces of tracelets

So far we have defined the decomposition space $\mathbf{X}_\bullet$ of rules, whose incidence algebra is the rule algebra of [9]. We now proceed towards Hopf algebras spanned by tracelets.

The Hopf algebra of tracelets should be spanned by iso-classes of tracelets. Furthermore, we now impose one more restriction on tracelets, namely that they should be *non-degenerate* as simplices of $\mathbf{X}_\bullet$. This means that the rules involved are not allowed to be the trivial rule.[3] The non-degenerate simplices of $\mathbf{X}_\bullet$ do not form a simplicial object, since inner faces of non-degenerate simplices are not always non-degenerate, but the outer face maps survive (as a consequence of the decomposition-space axioms), so as to define a presheaf

$$\vec{\mathbf{X}}_\bullet : \mathbb{\Delta}_{\text{inert}}^{\text{op}} \to \boldsymbol{Grpd}.$$

Left Kan extension along the inclusion functor $j : \mathbb{\Delta}_{\text{inert}} \to \mathbb{\Delta}$ defines a new simplicial groupoid:

$$\mathbf{Y}_\bullet := j_!\vec{\mathbf{X}}_\bullet \ : \ \mathbb{\Delta}^{\text{op}} \to \boldsymbol{Grpd}.$$

This is a general construction that makes sense for any (complete) decomposition space, and by a result of Hackney and Kock [19] it always produces a decomposition space again. Furthermore, one can expand explicitly what its simplices are:

$$\mathbf{Y}_k = \sum_{\alpha:[k]\twoheadrightarrow[n]} \vec{\mathbf{X}}_n.$$

(The sum is over active maps, cf. Appendix C.2.) In particular

$$\mathbf{Y}_0 = \mathbf{X}_0 \qquad \text{and} \qquad \mathbf{Y}_1 = \sum_{n\in\mathbb{N}} \vec{\mathbf{X}}_n.$$

So the new 1-simplices are the non-degenerate tracelets of any length. The higher simplices are 'subdivided tracelets'. To see this, recall that the decomposition space axioms can be written (cf. [16, Prop. 6.9]) as saying that for any active map $\alpha : [k] \twoheadrightarrow [n]$ the canonical square

$$\begin{array}{ccc} \mathbf{X}_n & \longrightarrow & \mathbf{X}_{n_1} \times \cdots \times \mathbf{X}_{n_k} \\ \downarrow & \lrcorner & \downarrow \\ \mathbf{X}_k & \longrightarrow & \mathbf{X}_1 \times \cdots \times \mathbf{X}_1 \end{array} \tag{12}$$

is a pullback. Here the vertical maps are active and the horizontal maps are combinations of inert maps. What the condition says is that it is possible to glue together k simplices (of different dimensions n_i) if just one has available a 'mould' to glue them together in, namely a k-simplex whose k principal edges match the long edges of the k simplices. (This is also the essence of the very definition of tracelet.)

[3] By imposing this condition, we account directly for an equivalence relation imposed in [2] called 'equivalence up to trivial tracelets' (cf. Definition 5.2).

An example of such a composition is depicted in Figure 2(c), in which a length 2 tracelet (depicted as the 2-simplex 012) is composed along the short edge 02 of the 2-simplex 023 with a tracelet of length 1 (here depicted as the edge 23). Figure 2(d) then depicts the method for computing the resulting tracelet of length 3, which itself is depicted in Figure 2(f).

Since non-degeneracy in a decomposition space can be measured on principal edges (cf. [17]), we also have the pullback

$$\begin{array}{ccc} \vec{\mathbf{X}}_n & \xrightarrow{\quad} & \vec{\mathbf{X}}_{n_1} \times \cdots \times \vec{\mathbf{X}}_{n_k} \\ \big\downarrow & \lrcorner & \big\downarrow \\ \mathbf{X}_k & \xrightarrow{\quad} & \mathbf{X}_1 \times \cdots \times \mathbf{X}_1 . \end{array} \tag{13}$$

We see that a k-simplex in $\mathbf{Y}_\bullet$ is the data of a tracelet τ of length k (not necessarily non-degenerate) together with a non-degenerate tracelet σ_i glued onto each of the principal edges of this base tracelet along their evaluation. (That is, the rule given by evaluating the tracelet σ_i must match the rule corresponding to the ith principal edge of τ.)

The corresponding algebra, given by the standard incidence algebra construction (cf. Appendix C.4), is spanned by isomorphism classes of non-degenerate tracelets, and the product of two tracelets is given by summing over all possible tracelet composites.

We now proceed to extend this structure into a Hopf algebra. This is not straightforward, because it is easy to see that the decomposition space $\mathbf{Y}_\bullet$ is not monoidal under sum. A monoidal structure exists in degree 1, by declaring the product of two tracelets to be the composite along trivial overlap:

$$T \odot T' := T \overset{\emptyset}{\angle} T'.$$

But this definition is not compatible with higher simplices.

Our task is now to explain how this is fixed in a canonical way. The solution amounts to imposing the so-called *shift equivalence* relation on tracelets, an equivalence relation already important in the theory. In the graphical interpretation it is about saying that for tracelets that are not connected, it should make no difference in which order they are applied. After passing to this equivalence relation, the monoidal structure $\odot$ will be well defined in all simplicial degrees. This final symmetric monoidal decomposition space of tracelets up to shift-equivalence will be denoted $\mathbf{Z}_\bullet$. We shall go deeper into the notion of shift equivalence in Section 5 (and interested reader are referred to [8, 1, 7] for the full background information and details). Here we just state the following proposition, which gives an alternative approach to shift equivalence.

A *splitting vertex* of a tracelet is an inner vertex for which the corresponding rule overlap is trivial. This property is invariant under precomposition with active maps. (That is, if $\sigma' = g(\sigma)$ for g an active map not eliminating vertex v, then v is splitting for σ' if and only if it is splitting for σ.) Second, there are expected transitive properties in connection with 'stages' in the sense of higher-order simplices of $\mathbf{Y}_\bullet$. Note that this transitive property does *not* imply that irreducibility is compatible with inert maps (outer face maps).

A non-degenerate tracelet $T \in \mathbf{Y}_1 = \sum_n \vec{\mathbf{X}}_n$ is *primitive* if it does not admit any splitting. (A higher-dimensional simplex $\sigma \in \mathbf{Y}_k$ (that is a subdivided tracelet) is primitive if its long edge is primitive in $\mathbf{Y}_1 = \sum_n \vec{\mathbf{X}}_n$ (that is, its underlying tracelet is primitive).)

Lemma 4.1. *Every maximal splitting of a given simplex has, up to isomorphism and permutation, the same primitive pieces.*

Proposition 4.2. *Tracelets are shift-equivalent in the restricted sense of trivial overlaps if and only if they have the same factorization into primitives.*

Proposition 4.3. *Shift-equivalence is compatible with the simplicial structure. This defines a simplicial groupoid $\mathbf{Z}_\bullet$ with $\mathbf{Z}_k = \mathbf{Y}_k/\sim$. This simplicial groupoid is a (locally finite) decomposition space.*

Note that if $\widetilde{\mathbf{Y}}_n$ denotes the groupoid of shift-equivalence classes, then we have

$$\mathbf{Z}_k = \sum_{[k]\twoheadrightarrow[n]} \mathbf{Y}_k \times_{\mathbf{X}_1^k} \left(\widetilde{\mathbf{Y}}_{n_1} \times \cdots \times \widetilde{\mathbf{Y}}_{n_k}\right)$$

This makes sense: in the fiber product, the maps from the factors $\widetilde{\mathbf{Y}}_{n_i}$ return the long edge, which is invariant under shift-equivalence.

Theorem 4.4. *There is a level-wise equivalence of groupoids*

$$\mathbf{Z}_k \simeq \mathsf{S}(\mathbf{Y}_k^{\mathrm{irr}}),$$

assembling into an equivalence of simplicial groupoids. In particular, $\mathbf{Z}_\bullet$ is symmetric monoidal under $\odot$.

Note that the primitive tracelets themselves do not form a simplicial groupoid. Only the active maps are well defined. The outer face maps applied to an primitive tracelet is not necessarily primitive. (Said in another way, a splitting vertex of a face might not be splitting for the whole simplex.) But after we apply S (the free-symmetric-monoidal-category monad), which is just a fancy way of saying 'monomials of' or 'families of', it does work.[4]

The upshot is now that the standard incidence algebra construction (cf. Appendix C.4) yields a Hopf algebra H of tracelets up to shift equivalence. This is the Hopf algebra we are really interested in, and towards which the previous ones were preliminary constructions. By Poincaré–Birkhoff–Witt, H is the enveloping algebra of the Lie algebra of primitive tracelets. In the next section we spell out the structure maps of this Hopf algebra in details.

5 The Hopf algebra of tracelets

Independently of the constructions of decomposition spaces presented in the preceding sections, it is possible to construct the tracelet Hopf algebra via an extension of the rule diagram Hopf algebra construction of [5] (which was based upon relational calculus) into the category-theoretical setting provided by tracelet theory. The readers more familiar with rewriting theory might appreciate however that the constructions presented in this section, while perhaps somewhat ad hoc at first sight, have a clear interpretation from decomposition space theory. Throughout this section, we fix a field $\mathbb{K}$ that will typically be chosen as either $\mathbb{R}$ or $\mathbb{C}$ (or, possibly, $\mathbb{Q}$). An essential prerequisite for our Hopf algebra construction is given by the following notions of equivalence relations.

5.1. Shift equivalence (cf. [2]). Let $\equiv_S$ denote the equivalence relation on $\mathcal{T}$ defined as the reflexive symmetric transitive closure of the relation on pairwise composition operations on tracelets: let $T = T_B \,{}^{\mu}\!\angle\, T_A$ (for some admissible match $\mu = (I_B \leftarrow M \rightarrow O_A)$), and denote by $[[T_B]] = (O_B \leftarrow K_B \rightarrow I_B)$ and $[[T_A]] = (O_A \leftarrow K_A \rightarrow I_A)$ the evaluations of T_B and T_A, respectively. Suppose $[[T_B]]$ and $[[T_A]]$ are *sequentially independent* in the composition along μ, which entails that M is isomorphic to both the pullbacks of the cospans $K_B \rightarrow I_B \leftarrow M$ and $M \rightarrow O_A \leftarrow K_A$, respectively. In this situation we define the composite tracelet $\overline{T} = T_A \,{}^{\overline{\mu}}\!\angle\, T_B$ (for $\overline{\mu} = I_A \leftarrow M \rightarrow O_B$) to be *shift equivalent* to the tracelet $T = T_B \,{}^{\mu}\!\angle\, T_A$.

[4] Note how this is analogous to the decomposition space of forests [18]: every forest is of course a disjoint union of trees, but the top face of a tree is not in general a tree, only a forest.

5.2. Normal form equivalence (cf. [2]). Let $\equiv_A$ denote an equivalence relation on $\mathfrak{T}$ (so-called *abstraction equivalence*) whereby $T \equiv_A T'$ iff T and T' are tracelets of the same length, and if moreover there exists an isomorphism $T \xrightarrow{\cong} T'$ (induced from isomorphisms on objects so that the resulting diagram commutes). Let $\equiv_T$ be defined as the reflexive symmetric transitive closure of a relation whereby for any $T \in \mathfrak{T}$, we let $T \equiv_T T \uplus T_\varnothing \equiv_T T_\varnothing \uplus T$ (with $T_\varnothing := T(\varnothing \leftarrow \varnothing \rightarrow \varnothing) \in \mathfrak{T}_1$, and where $\uplus := {}^{\mu_\varnothing}\angle$ denotes tracelet composition along trivial overlap). Then we define the *tracelet normal form* equivalence relation as $\equiv_N := rst(\equiv_A \cup \equiv_T \cup \equiv_S)$, i.e., as the reflexive transitive closure of the union of the aforementioned three relations.

Definition 5.3 (Primitive tracelets)**.** Denote by $\mathfrak{Prim}(\mathfrak{T}_N)$ the set of *primitive tracelets*, defined as

$$\mathfrak{Prim}(\mathfrak{T}_N) := \{[T]_{\equiv_N} | T \neq T_\varnothing \wedge \nexists T_A, T_B \neq T_\varnothing : T \equiv_N T_A \uplus T_B\}\,. \tag{14}$$

Primitive tracelets play a central role in our construction, since they are in a certain sense the smallest "indecomposable" building blocks of tracelets with respect to (de-)composition (just as primitive *rule diagrams* in [5]).

Proposition 5.4 (Tracelet normal form)**.** *Every tracelet $T \in \mathfrak{T}$ is $\equiv_N$-equivalent to a* tracelet normal form *in the sense that $T_\varnothing \equiv_N T_\varnothing$, and*[5] $\forall T \neq T_\varnothing : \quad T \equiv_N \biguplus_{i \in I} T_i$*, where $T_i \in \mathfrak{Prim}(\mathfrak{T}_N)$ for all $i \in I$, and with I a (finite) index set. (*Proof: *cf. Appendix B.2)*

Definition 5.5 (Tracelet $\mathbb{K}$-vector space $\hat{\mathfrak{T}}$)**.** Let $\hat{\mathfrak{T}}$ be the $\mathbb{K}$-vector space spanned by a basis indexed by $\equiv_N$-equivalence classes, in the sense that there exists an isomorphism $\delta : \mathfrak{T}_N \xrightarrow{\cong} basis(\hat{\mathfrak{T}})$ from the set[6] of $\equiv_N$-equivalence classes of tracelets $\mathfrak{T}_N := \mathfrak{T}/_{\equiv_N}$ to the set of basis vectors of $basis(\hat{\mathfrak{T}})$. We will use the notation $\hat{T} := \delta(T)$ for the basis vector associated to some class $T \in \mathfrak{T}_N$. We denote by $Prim(\hat{\mathfrak{T}}) \subset \hat{\mathfrak{T}}$ the sub-vector space of $\hat{\mathfrak{T}}$ spanned by basis vectors indexed by primitive tracelets.

Definition 5.6 (Tracelet algebra product and unit)**.** Let $\otimes \equiv \otimes_{\mathbb{K}}$ be the tensor product operation on the $\mathbb{K}$-vector space $\hat{\mathfrak{T}}$. Then the *multiplication map* μ and the *unit map* $\eta : \mathbb{K} \rightarrow \hat{\mathfrak{T}}$ are defined via their action on basis vectors of $\hat{\mathfrak{T}}$ as follows:

$$\mu : \hat{\mathfrak{T}} \otimes \hat{\mathfrak{T}} \rightarrow \hat{\mathfrak{T}} : \hat{T} \otimes \hat{T}' \mapsto \hat{T} \diamond \hat{T}'\,, \qquad \hat{T} \diamond \hat{T}' := \sum_{\mu \in \mathsf{MT}_T(T')} \delta\left(\left[T\, {}^{\mu}\angle\, T'\right]_{\equiv_N}\right) \tag{15}$$

$$\eta : \mathbb{K} \rightarrow \hat{\mathfrak{T}} : k \mapsto k \cdot \hat{T}_\varnothing\,. \tag{16}$$

Both definitions are suitably extended by (bi-)linearity to generic (pairs of) elements of $\hat{\mathfrak{T}}$.

Proposition 5.7. *The morphisms μ and η give rise to an* associative, unital $\mathbb{K}$-algebra $(\hat{\mathfrak{T}}, \mu, \eta)$*, which we refer to as* tracelet algebra*. (*Proof: *cf. Appendix B.3)*

Definition 5.8 (Tracelet coproduct and counit)**.** Fixing the *notational convention* $\uplus_{i \in \emptyset} T_i := T_\varnothing$ for later convenience, let $T \equiv_N \uplus_{i \in I} T_i$ be the tracelet normal form for a given tracelet $T \in \mathfrak{T}$ (where $T_i \in \mathfrak{Prim}(\mathfrak{T}_N)$ for all $i \in I$ if $T \neq T_\varnothing$). Then the *tracelet coproduct* Δ and *tracelet counit* ε are defined via their action on basis vectors $\hat{T} = \delta(T)$ of $\hat{\mathfrak{T}}$ as

$$\Delta : \hat{\mathfrak{T}} \rightarrow \hat{\mathfrak{T}} \otimes \hat{\mathfrak{T}} : \hat{T} \mapsto \Delta(\hat{T}) := \sum_{X \subset I} \delta\left([\uplus_{x \in X} T_x]_{\equiv_N}\right) \otimes \delta\left(\left[\uplus_{y \in I \setminus X} T_x\right]_{\equiv_N}\right) \tag{17}$$

and $\varepsilon : \hat{\mathfrak{T}} \rightarrow \mathbb{K} : \hat{T} \mapsto coeff_{\hat{T}_\varnothing}(\hat{T})$. Both definitions are extended by linearity to generic elements of $\hat{\mathfrak{T}}$.

[5] We chose to make the case distinction explicit in order to emphasize that the normal form of a non-trivial tracelet $T \neq T_\varnothing$ does itself not contain trivial sub-tracelets, such that manifestly $T_i \in \mathfrak{Prim}(\mathfrak{T}_N)$ in $T \equiv_N \uplus_{i \in I} T_i$. This is clearly the case, since invoking $\equiv_T$ on $\uplus_{i \in I} T_i$ would in effect remove any trivial constituent $T_i = T_\varnothing$.

[6] Here, we tacitly assume that the $\equiv_N$-equivalence classes indeed form a proper *set*, which is in all known applications the case since abstraction equivalence $\equiv_A$ is part of the definition of $\equiv_N$. For example, it is well known that isomorphism classes of finite directed multigraphs indeed form a set.

Proposition 5.9. *The data* $(\hat{\mathfrak{T}}, \Delta, \varepsilon)$ *defines a* coassociative, cocommutative and counital coalgebra*.*

Proof. Since the construction of Δ and ε is the standard construction for a deconcatenation coalgebra (cf. e.g. [21]), the proof is omitted here for brevity. □

The algebra and coalgebra structures on $\hat{\mathfrak{T}}$ are compatible in the following sense:

Theorem 5.10 (Bialgebra structure)**.** *The data* $(\hat{\mathfrak{T}}, \mu, \eta, \Delta, \varepsilon)$ *defines a* bialgebra*. (*Proof: *cf. Appendix B.4)*

By virtue of the definition of the tracelet normal form, it is evident that both composition and decomposition of tracelets is compatible with a filtration structure given by number of "connected components" in the following sense:

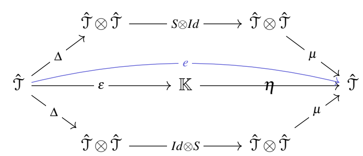
Theorem 5.11. *The tracelet bialgebra* $(\hat{\mathfrak{T}}, \mu, \eta, \Delta, \varepsilon)$ *is* connected and filtered*, with connected component* $\hat{\mathfrak{T}}^{(0)} := span_{\mathbb{K}}\{\hat{T}_{\varnothing}\}$*, and with the higher components of the filtration given by the subspaces*

$$\forall n > 0: \quad \hat{\mathfrak{T}}^{(n)} := span_{\mathbb{K}} \left\{ \hat{T}_1 \uplus \ldots \uplus \hat{T}_n \,\middle|\, \hat{T}_1, \ldots, \hat{T}_n \in Prim(\hat{\mathfrak{T}}) \right\}, \tag{18}$$

where in a slight abuse of notations $\hat{T}_1 \uplus \ldots \uplus \hat{T}_n := \delta(T_1 \uplus \ldots \uplus T_n)$ *(*Proof: *cf. Appendix B.6)*

Utilizing results from the general theory of Hopf algebras (cf. e.g. [21] for an excellent review), we finally obtain one of the central results of the present paper:

Theorem 5.12 (Compare [5], Sec. 3.4 and Thm. 3.2)**.** *The tracelet bialgebra* $(\hat{\mathfrak{T}}, \mu, \eta, \Delta, \varepsilon)$ *admits the structure of a* Hopf algebra*, where the* antipode *S, which is to say the endomorphism of* $\hat{\mathfrak{T}}$ *that makes the diagram below commute,*



(19)

is given by $S := Id^{\star^{-1}}$*. The latter denotes the inverse of the identity morphism* $Id : \hat{\mathfrak{T}} \to \hat{\mathfrak{T}}$ *under the* convolution product $\star$ *of linear endomorphisms on* $\hat{\mathfrak{T}}$*. More concretely, letting* $e := \eta \circ \varepsilon$ *denote the unit for the convolution product* $\star$*,*

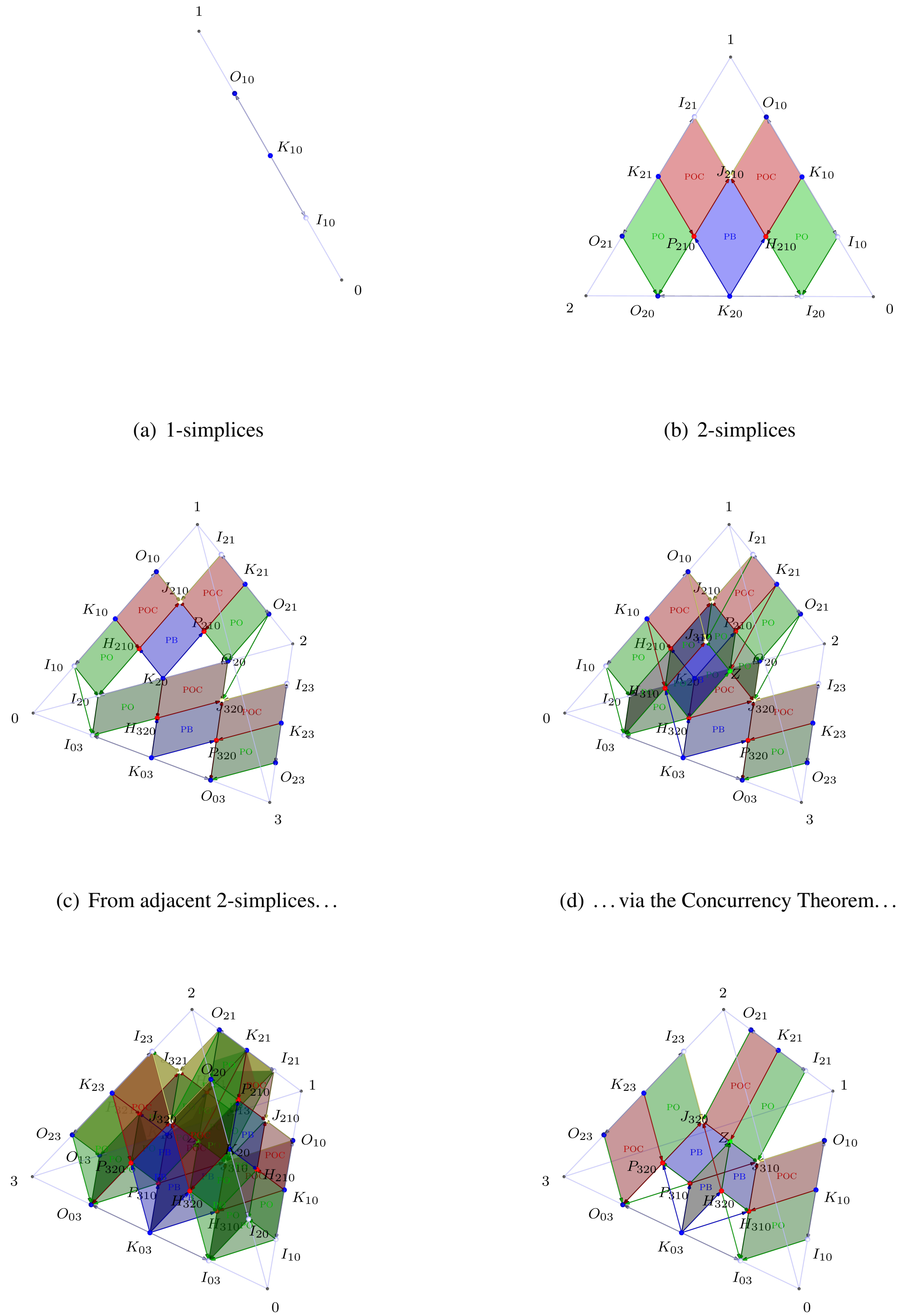
$$S(\hat{T}) = Id^{\star^{-1}}(\hat{T}) = (e - (e - Id))^{\star^{-1}} = e(\hat{T}) + \sum_{k \geq 1} (e - Id)^{\star k}(\hat{T}). \tag{20}$$

It might be instructive to compute the action of the coproduct and of the antipode on tracelets of low filtration degree. To this end, in the equations below let $\hat{T}_i \in Prim(\hat{\mathfrak{T}})$ denote primitive elements of $\hat{\mathfrak{T}}$.

$$\begin{aligned}
\Delta(\hat{T}_{\varnothing}) &= \hat{T}_{\varnothing} \otimes \hat{T}_{\varnothing} & S(\hat{T}_{\varnothing}) &= \hat{T}_{\varnothing} \\
\Delta(\hat{T}_1) &= \hat{T}_{\varnothing} \otimes \hat{T}_1 + \hat{T}_1 \otimes \hat{T}_{\varnothing} & S(\hat{T}_1) &= -\hat{T}_1 \\
\Delta(\hat{T} = \hat{T}_1 \uplus \hat{T}_2) &= \hat{T}_{\varnothing} \otimes \hat{T} + \hat{T} \otimes \hat{T}_{\varnothing} + \hat{T}_1 \otimes \hat{T}_2 + \hat{T}_2 \otimes \hat{T}_1 & S(\hat{T}_1 \uplus \hat{T}_2) &= \hat{T}_1 \diamond \hat{T}_2 + \hat{T}_2 \diamond \hat{T}_1 - \hat{T}_1 \uplus \hat{T}_2
\end{aligned} \tag{21}$$

Finally, yet again taking inspiration from [5], one may demonstrate that the tracelet Hopf algebra is isomorphic to a Hopf algebra that is well-known in the setting of the Heisenberg-Weyl diagram Hopf algebra and the Poincaré-Birkhoff-Witt theorem for "normal-ordering" of elements of the Hopf algebra:

Theorem 5.13. *Let* $\mathcal{L}_{\mathfrak{T}} := (Prim(\hat{\mathfrak{T}}), [.,.]_{\diamond})$ *denote the* tracelet Lie algebra*, where* $[\hat{T}, \hat{T}']_{\diamond} := \hat{T} \diamond \hat{T}' - \hat{T}' \diamond \hat{T}$ *is the* commutator operation *(w.r.t.* $\diamond$*). Then the tracelet Hopf algebra is isomorphic (in the sense of Hopf algebra isomorphisms) to the* universal enveloping algebra *of* $\mathcal{L}_{\mathfrak{T}}$*.*



(a) 1-simplices

(b) 2-simplices

(c) From adjacent 2-simplices...

(d) ...via the Concurrency Theorem...

(e) ...to 3-simplices.

(f) Tracelets of length 3 and evaluation

Figure 2: Elements of Tracelet Decomposition Space theory (*Note:* in order to allow for a more in-detail inspection, the figures are hyperlinked to on-line interactive 3D views of the respective diagrams).

References

[1] Nicolas Behr (2019): *Sesqui-Pushout Rewriting: Concurrency, Associativity and Rule Algebra Framework*. In: Proceedings of *GCM 2019*, *EPTCS* 309, pp. 23–52, doi:10.4204/eptcs.309.2.

[2] Nicolas Behr (2020): *Tracelets and Tracelet Analysis Of Compositional Rewriting Systems*. In: Proceedings of *ACT 2019*, *EPTCS* 323, pp. 44–71, doi:10.4204/EPTCS.323.4.

[3] Nicolas Behr (2021): *On Stochastic Rewriting and Combinatorics via Rule-Algebraic Methods*. In: *Proceedings of TERMGRAPH 2020*, 334, pp. 11–28, doi:10.4204/eptcs.334.2.

[4] Nicolas Behr, Vincent Danos & Ilias Garnier (2016): *Stochastic mechanics of graph rewriting*. In: Proceedings of *LICS '16*, doi:10.1145/2933575.2934537.

[5] Nicolas Behr, Vincent Danos, Ilias Garnier & Tobias Heindel (2016): *The algebras of graph rewriting*. arXiv *preprint 1612.06240*.

[6] Nicolas Behr & Jean Krivine (2020): *Rewriting theory for the life sciences: A unifying framework for CTMC semantics*. In: *Graph Transformation (ICGT 2020)*, *TCS* 12150, doi:10.1007/978-3-030-51372-6.

[7] Nicolas Behr & Jean Krivine (2021): *Compositionality of Rewriting Rules with Conditions*. Compositionality 3, doi:10.32408/compositionality-3-2.

[8] Nicolas Behr & Pawel Sobocinski (2018): *Rule Algebras for Adhesive Categories*. In: Proceedings of *(CSL 2018)*, *LIPIcs* 119, pp. 11:1–11:21, doi:10.4230/LIPIcs.CSL.2018.11.

[9] Nicolas Behr & Pawel Sobocinski (2020): *Rule Algebras for Adhesive Categories (extended journal version)*. *LMCS* Volume 16, Issue 3. Available at `https://lmcs.episciences.org/6615`.

[10] Pierre Boutillier et al. (2018): *The Kappa platform for rule-based modeling*. Bioinformatics 34(13), pp. i583–i592, doi:10.1093/bioinformatics/bty272.

[11] Benjamin Braatz, Hartmut Ehrig, Karsten Gabriel & Ulrike Golas (2014): *Finitary $\mathcal{M}$ -adhesive categories*. Mathematical Structures in Computer Science 24(4), pp. 240403–240443, doi:10.1017/S0960129512000321.

[12] Tobias Dyckerhoff & Mikhail Kapranov (2019): *Higher Segal spaces*. Lecture Notes in Mathematics 2244, Springer-Verlag.

[13] H. Ehrig, K. Ehrig, U. Prange & G. Taentzer (2006): *Fundamentals of Algebraic Graph Transformation*. Monographs in Theoretical Computer Science. An EATCS Series, doi:10.1007/3-540-31188-2.

[14] Hartmut Ehrig et al. (2014): *$\mathcal{M}$-adhesive transformation systems with nested application conditions. Part 1: parallelism, concurrency and amalgamation*. *MSCS* 24(04), doi:10.1017/s0960129512000357.

[15] Imma Gálvez-Carrillo, Joachim Kock & Andrew Tonks (2016): *Decomposition spaces in combinatorics*. Preprint, arXiv:1612.09225.

[16] Imma Gálvez-Carrillo, Joachim Kock & Andrew Tonks (2018): *Decomposition spaces, incidence algebras and Möbius inversion I: Basic theory*. Adv. Math. 331, pp. 952–1015, doi:10.1016/j.aim.2018.03.016.

[17] Imma Gálvez-Carrillo, Joachim Kock & Andrew Tonks (2018): *Decomposition spaces, incidence algebras and Möbius inversion II: Completeness, length filtration, and finiteness*. Adv. Math. 333, pp. 1242–1292, doi:10.1016/j.aim.2018.03.017.

[18] Imma Gálvez-Carrillo, Joachim Kock & Andrew Tonks (2020): *Decomposition spaces and restriction species*. Int. Math. Res. Notices 2020(21), pp. 7558–7616, doi:10.1093/imrn/rny089.

[19] Philip Hackney & Joachim Kock (2021): *Free decomposition spaces*. In preparation.

[20] Stephen Lack & Paweł Sobociński (2004): *Adhesive Categories*. In: Proceedings of *FoSSaCS 2004*, *LNCS* 2987, pp. 273–288, doi:10.1007/978-3-540-24727-2_20.

[21] Dominique Manchon (2008): *Hopf algebras in renormalisation*. Handbook of algebra 5, pp. 365–427.

[22] Gian-Carlo Rota (1964): *On the foundations of combinatorial theory. I. Theory of Möbius functions*. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 2, pp. 340–368 (1964). Available at `https://www.maths.ed.ac.uk/~v1ranick/papers/rota1.pdf`.

A Background material

A.1. Double-Pushout (DPO) rewriting theory. An important ingredient in analyzing rewriting theories based upon DPO semantics over adhesive categories is the following theorem, which is a central result of categorical rewriting theory [13], and which we state here in the variant used in [2]. Note that while the following statement is formulated for *rules*, which in the present context represent the special case of length 1 tracelet, applying the theorem repeatedly leads precisely to the notion of tracelets of arbitrary length (compare Section 2.2). Referring to [8] for the full details of the categorical rule-based calculus, suffice it here to quote the theorem, and inviting the readers to inspect the relevant commutative diagrams in (24) and (25) below:

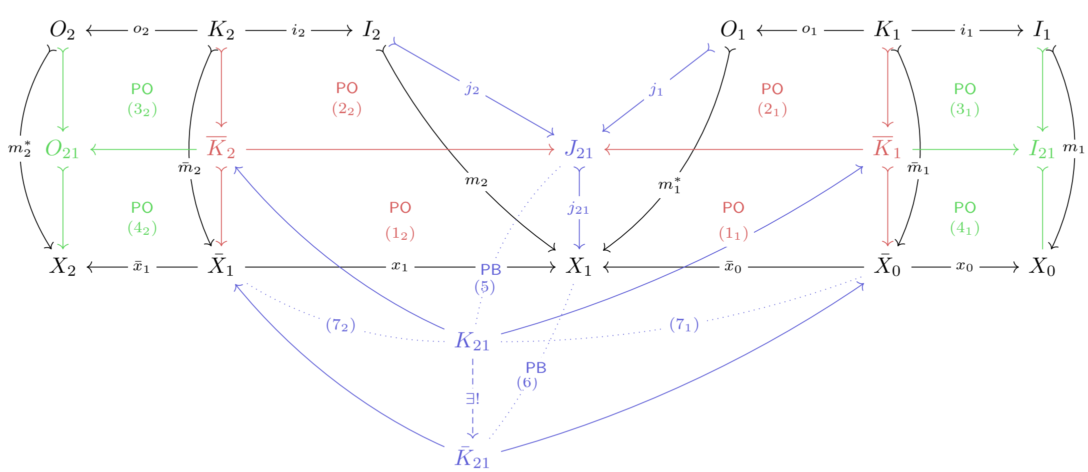
Theorem A.2 (Concurrency theorem [8]). *There exists a bijection $\varphi : A \xrightarrow{\cong} B$ on pairs of DPO-admissible matches between the sets A and B,*

$$\begin{aligned} & A=\{(m_2,m_1)\mid m_1\in\mathcal{M}_{R_1}(X_0)\,,;m_2\in\mathcal{M}_{R_2}(X_1)\}\\ \cong\quad & B=\{(\mu_{21},m_{21})\mid \mu_{21}\in\mathsf{M}_{R_2}(R_1)\,,\ m_{21}\in\mathcal{M}_{R_{21}}(X_0)\}\,, \end{aligned} \tag{22}$$

where $X_1 = R_{1_{m_1}}(X_0)$ and $R_{21} = R_2{}^{\mu_{21}}\!\triangleleft R_1$ such that for each corresponding pair $(m_2,m_1) \in A$ and $(\mu_{21},m_{21}) \in B$, it holds that

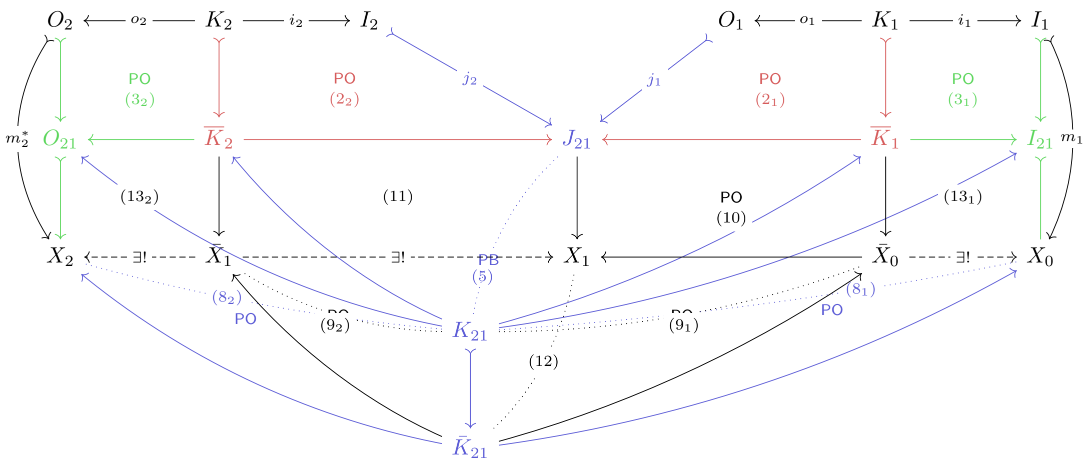
$$R_{21_{m_{21}}}(X_0) \cong R_{2_{m_2}}(R_{1_{m_1}}(X_0))\,. \tag{23}$$

The importance of this theorem with regard to the rule decomposition spaces constructed in the present paper resides in the fact that from any two-step sequence of direct derivations, one can construct (uniquely up to universal isomorphisms) a one-step direct derivation along a composite rule, which amounts to the “synthesis” part of the concurrency theorem:



(24)

Conversely, given a direct derivation along a composite rule, one may construct (again uniquely up to isomorphisms) a two-step sequence of direct derivations along the constituent rules, which amounts to the “analysis” part of the theorem:



(25)

The concurrency theorem may be used to derive the following technical result, which is useful for interpreting the notion of tracelet composition as well as the internal structure of three-simplices as depicted in Figure 2(d):

Corollary A.3 (Cf. [2], Cor. 2)**.** *Let $r_{n\cdots 1} = (O_{n\cdots 1} \leftharpoonup I_{n\cdots 1})$ be a span of monomorphisms, and let $(Y^{(n)}_{j+1,j} \leftharpoonup Y^{(n)}_{j,j-1})$ be n spans of monomorphisms (for $0 \leq j \leq n$) with $Y^{(n)}_{n+1,n} = O_{n\cdots 1}$, $Y^{(n)}_{1,0} = I_{n\cdots 1}$, and such that*

$$(O_{n\cdots 1} \leftharpoonup I_{n\cdots 1}) = (O_{n\cdots 1} \leftharpoonup Y^{(n)}_{n,n-1}) \circ \ldots \circ (Y^{(n)}_{2,1} \leftharpoonup I_{n\cdots 1}) \,.$$

Then for each object X_0 and for each DPO-admissible match $(I_{n\cdots 1} \hookrightarrow X_0) \in \mathcal{M}_{R_{n\cdots 1}}(X_0)$ (for $R_{n\cdots 1} = (r_{n\cdots 1})$, the direct derivation of X_0 by $R_{n\cdots 1}$ to X_0 along this match

$$\begin{array}{ccc} O_{n\cdots 1} & \leftharpoonup & I_{n\cdots 1} \\ \downarrow & \text{DPO} & \downarrow \\ X_n & \leftharpoonup & X_0 \end{array} \tag{26a}$$

uniquely (up to isomorphisms) encodes an n-step DPO-type derivation sequence of the following form, and vice versa:

$$\begin{array}{ccccccc} O_{n\cdots 1} & \leftharpoonup & Y^{(n)}_{n,n-1} & \cdots & Y^{(n)}_{2,1} & \leftharpoonup & I_{n\cdots 1} \\ \downarrow & \text{DPO} & \downarrow & & \downarrow & \text{DPO} & \downarrow \\ X_n & \leftharpoonup & X_{n-1} & \cdots & X_1 & \leftharpoonup & X_0 \end{array} \tag{26b}$$

A.4. Partial overlaps and minimal gluings. An interesting property of adhesive categories concerns *gluings* of objects:

Lemma A.5. *For every cospan of monomorphisms $A \xrightarrow{a} X \xleftarrow{b} B$, there exist a span $A \xleftarrow{\alpha} I \xrightarrow{\beta} B$ of monomorphisms, a cospan $A \xrightarrow{a'} X' \leftarrow B$ of monomorphisms and a monomorphism $X' \xrightarrow{x} X$ such that*

- $[\alpha, \beta]$ *is a* pullback *of* $]a, b[$
- $]a', b'[$ *is a* pushout *of* $[\alpha, \beta]$*, and in particular a jointly epimorphic cospan*
- $a = x \circ a'$ *and* $b = x \circ b'$

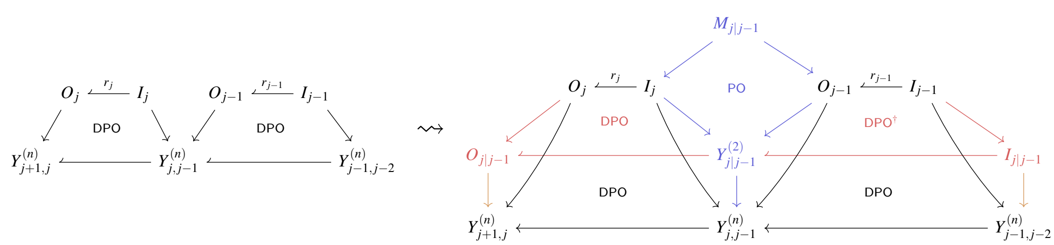
We refer to the span $[\alpha,\beta]$ *as a* partial overlap *and to the cospan* $]a',b'[$ *as a* minimal gluing *of* A *and* B. *Partial overlaps and minimal gluings are unique up to isomorphisms in* X'.

Proof. The statement follows from the property of so-called *effective binary unions* [20], i.e., in an adhesive category binary unions of subobjects are computed as pushouts of their intersection (which is itself computed via pullback). Uniqueness up to isomorphisms then follows from the universal property of pushouts. □

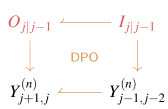
Minimal gluings of objects are a key concept for constructing tracelets, as they model partial interactions of rules in sequences of rewriting steps.

A.6. Tracelet surgery. For several of the constructions presented in this paper (such as in particular for the definition of face maps in tracelet decomposition spaces), it is necessary to consider the following type of operation that permits to extract information from tracelets. The basis for this type of reasoning is once again the concurrency theorem (cf. Theorem A.2).

Corollary A.7 (Tracelet surgery; [2], Cor. 1)**.** *Let* $T \in \mathcal{T}_n$ *be a tracelet of length* n*, so that it consists of* n *sequential direct derivations* $t_n,\dots,t_1$ *(written in the following as* $T \equiv t_n|\dots|t_1$*). Then for any consecutive direct derivations* $t_j|t_{j-1}$ *in* T*, one may uniquely (up to isomorphisms) construct a diagram* $t_{(j|j-1)}$ *and a tracelet* $T_{(j|j-1)}$ *of length* 2 *as follows (where* $\rightsquigarrow$ *amounts to an application of the "analysis" part of the concurrency theorem):*



(27a)

$t_{(j|j-1)} :=$  $,\quad T_{(j|j-1)} := T(r_j)\,{}^{\mu}\!\angle\, T(r_{j-1})$ (27b)

Here, $\mu := (I_j \hookleftarrow M \hookrightarrow O_{j-1})$ *is the span of monomorphisms obtained by taking the pullback of the cospan* $(I_j \hookrightarrow Y^{(n)}_{j,j-1} \hookleftarrow O_{j-1})$*, and this span is (by virtue of the concurrency theorem) an admissible match. By associativity of the tracelet composition, this extends to consecutive sequences* $t_j|\dots|t_{j-k}$ *of adjacent direct derivations in* T *inducing diagrams* $t_{(j|\dots|j-k)}$ *and tracelets of length* 1*, denoted* $T_{(j|\dots|j-k)}$*, where for* $k=0$*,* $t_{(j)} = t_j$ *and* $T_{(j)} = T(r_j)$*.*

B Proofs

B.1. Proof of Lemma 3.7. We already noted that d_i operates entirely within the pair of consecutive plaquettes, and that it does not modify the outer feet. The fiber thus consists of pairs of composable plaquettes (p_{i+1},p_i) with some middle foot y, but with fixed outer feet (the same as the feet of p'). By definition of d_i, we should first construct w as the pushout of the pullback and then compose the two rules along w, and this is required to be r'. That is precisely the description of the fiber $(\mathbf{X}_2)_{r'}$. It remains to see that the rest of the data is uniquely reconstructible. But this follows via the "analysis" part of the concurrency theorem, whereby if r' is realized as a composition of rule r_{i+1} with rule r_i along some admissible match into rule r', the direct derivation along r' may be equivalently expressed as a two-step sequence of direct derivations along r_i followed by r_{i+1} (cf. Theorem A.2 of Appendix A.1).

B.2. Proof of Proposition 5.4. For the case $T = T_{\varnothing}$, since $\varnothing$ is by assumption a strict initial object, there exists precisely one inhabitant in the isomorphism class of $T_{\varnothing}$ (i.e. $T_{\varnothing}$ itself). Moreover, as $T_{\varnothing}$ is a tracelet of length 1, neither $\equiv_S$ nor $\equiv_T$ may be invoked non-trivially, thus proving the claim in this case. For the case $T \neq T_{\varnothing}$, invoking $\equiv_N$ repeatedly in order to compute the $\equiv_N$-equivalence class of T in effect removes any occurrence of sub-tracelets containing the trivial rule; since by assumption $T \neq T_{\varnothing}$ and since T is of finite length, the process returns an $\equiv_N$-irreducible class of minimal representatives $T' \equiv T$ with $T' \neq T_{\varnothing}$. Such a T' could then itself be obtainable by repeated operations $\uplus$ on shorter tracelets (all non-trivial), which by invoking $\equiv_A$ and $\equiv_S$ yield manifestly equivalent "permutations" of $\uplus$ composants, thus proving the claim.

B.3. Proof of Proposition 5.7. Due to linearity, it suffices to verify the claim on basis vectors of $\hat{\mathfrak{T}}$. Associativity of μ follows from the associativity of tracelet composition (cf. [2, Thm. 1]):

$$\mu \circ (\mu \otimes id)[\hat{T} \otimes \hat{T}' \otimes \hat{T}''] = (\hat{T} \diamond \hat{T}') \diamond \hat{T}'' = \hat{T} \diamond (\hat{T}' \diamond \hat{T}'') = \mu \circ (id \otimes \mu)[\hat{T} \otimes \hat{T}' \otimes \hat{T}''] .$$

Unitality is established via

$$\mu \circ (\eta \otimes id)[\hat{T}] = \mu[\hat{T}_{\varnothing} \otimes \hat{T}] = \hat{T} = \mu[\hat{T} \otimes \hat{T}_{\varnothing}] = \mu \circ (id \otimes \eta)[\hat{T}] .$$

More precisely, we have invoked the natural isomorphism $1 \otimes \hat{T} = \hat{T} = \hat{T} \otimes 1$ in several steps.

B.4. Proof of Theorem 5.10. It is necessary to provide proofs of all four axioms of bialgebras in order to verify the claim. Due to linearity, one may again focus on calculations in terms of basis vectors only.

Axiom I:

$$\varepsilon \circ \eta[1_{\mathbb{K}}] = \varepsilon[\hat{T}_{\varnothing}] = 1_{\mathbb{K}} = id[1_{\mathbb{K}}] . \tag{28}$$

Axiom II:

$$\Delta \circ \eta[1_{\mathbb{K}}] = \Delta[\hat{T}_{\varnothing}] = \hat{T}_{\varnothing} \otimes \hat{T}_{\varnothing} = (\eta \otimes \eta)[1_{\mathbb{K}}] . \tag{29}$$

Axiom III: The verification of the following property

$$\varepsilon \circ \mu[\hat{T} \otimes \hat{T}'] = \varepsilon[\hat{T} \diamond \hat{T}'] \stackrel{!}{=} (\varepsilon \otimes \varepsilon)[\hat{T} \otimes \hat{T}'] \tag{30}$$

requires a little additional work. We have to prove that $\hat{T} \diamond \hat{T}'$ contains the summand $\hat{T}_{\varnothing}$ (with coefficient 1) if and only if $\hat{T} = \hat{T}' = \hat{T}_{\varnothing}$. The proof follows from the fact that $\hat{T}_{\varnothing}$ is the *unit element* for $\diamond$ – for all $\hat{T} \neq \hat{T}_{\varnothing}$,

$$\hat{T} \diamond \hat{T}_{\varnothing} = \hat{T}_{\varnothing} \diamond \hat{T} = \hat{T} \neq \hat{T}_{\varnothing} , \tag{31}$$

which proves the claim for all cases but for the case $\hat{T} \neq \hat{T}_{\varnothing}$ and $\hat{T}' \neq \hat{T}_{\varnothing}$. However, it is straightforward to verify from the precise definition of tracelets and of the tracelet composition operations that it is not possible to obtain a tracelet $\equiv_N$-equivalent to $T_{\varnothing}$ by composition of tracelets not themselves equal to $T_{\varnothing}$, which completes the proof of the Axiom III property.

Axiom IV: The most difficult to prove property is the compatibility of product and coproduct, which must be verified on arbitrary basis vectors $\hat{T}, \hat{T}' \in \hat{\mathfrak{T}}$:

$$\Delta \circ \mu[\hat{T} \otimes \hat{T}'] \stackrel{!}{=} (\mu \otimes \mu) \circ (id \otimes \tau \otimes id) \circ (\Delta \otimes \Delta)[\hat{T} \otimes \hat{T}'] . \tag{32}$$

Here, $\tau : \hat{\mathfrak{T}} \otimes \hat{\mathfrak{T}} \to \hat{\mathfrak{T}} \otimes \hat{\mathfrak{T}}$ is the *interchange morphism*, defined to act as $\tau[\hat{T} \otimes \hat{T}'] := \hat{T}' \otimes \hat{T}$. While the property is straightforward to prove for the case where either both or one of $\hat{T}$ or $\hat{T}'$ is equal to $\hat{T}_{\varnothing}$, the general case requires an intricate case study. Let us first introduce some auxiliary notations.

Definition B.5. Given two spans $s_i := (A_i \xleftarrow{a_i} X_i \xrightarrow{b_i} B_i)$ $(i=1,2)$, denote by

$$s_i \uplus s_j := (A_1 + A_2 \xleftarrow{[a_1,a_2]} X_1 + X_2 \xrightarrow{[b_1,b_2]} B_1 + B_2)$$

their disjoint union (i.e., the pushout of $s_1 \leftarrow s_\varnothing \rightarrow s_2$ over the span $s_\varnothing := (\varnothing \leftarrow \varnothing \rightarrow \varnothing)$ in the category $span(\mathbf{C})$ of spans over $\mathbf{C}$). Given a span $s = \sum_k A'_k \leftarrow \sum_\ell X_\ell \rightarrow \sum_m B_m$ (where all objects of the span are finite disjoint unions of objects), a *minimal presentation* of s is defined as a disjoint union of spans of the form

$$s = \biguplus_{\gamma} s_\alpha\,, \quad s_\alpha = (A_\gamma \leftarrow X_\gamma \rightarrow B_\gamma) \tag{33}$$

which satisfies that $s_\alpha \neq s_\varnothing$ for all indices α. We furthermore define the *format* of s, denoted $\Phi(s)$, as

$$\Phi(s) := \uplus_\gamma \{(|A_\gamma|, |B_\gamma|)\}\,, \tag{34}$$

where $|A_\gamma|$ and $|B_\gamma|$ denote the number of connected components of A_γ and B_γ, respectively. Finally, we let $\mathcal{F}_{m,n}$ denote the set of all formats with the sum of first entries equal to m and the sum of second entries to n.

Note that with the above definitions, the format of a span of the form

$$s = (\uplus_\alpha A_\alpha) \leftarrow \varnothing \rightarrow (\uplus_\beta B_\beta) \tag{35}$$

evaluates to

$$\Phi(s) = \left(\biguplus_{\alpha} \{(1,0)\}\right) \uplus \left(\biguplus_{\beta} \{(0,1)\}\right)\,. \tag{36}$$

Let $\hat{T}_X := \biguplus_{x \in X} \hat{T}_x$ and $\hat{T}'_Y := \biguplus_{y \in Y} \hat{T}'_y$ (for some finite index sets X and Y, and so that $\hat{T}_x \neq \hat{T}_\varnothing$ and $\hat{T}'_y \neq \hat{T}_\varnothing$ for all indices x and y) denote two tracelets in normal form. Equipped with the notion of formats of spans (which in our applications encode partial overlaps), we may refine the formula for the composition of the two tracelets as follows:

$$\begin{aligned}
\hat{T}_X \diamond \hat{T}'_Y &= \sum_{\mu \in \mathsf{MT}_{T_X}(T'_Y)} \delta\left(T_X \,{}^{\mu}\!\angle\, T'_Y\right) \\
&= \sum_{F \in \mathcal{F}_{|in(T_X)|,|out(T'_Y)|}} \sum_{\substack{\mu \in \mathsf{MT}_{T_X}(T'_Y) \\ \Phi(\mu)=F}} \delta\left(T_X \,{}^{\mu}\!\angle\, T'_Y\right) \equiv \sum_{F \in \mathcal{F}_{|in(T_X)|,|out(T'_Y)|}} \hat{T}_X \circledast_F \hat{T}'_Y
\end{aligned} \tag{37}$$

Here, we made use of the fact that $\mathsf{MT}_{T_X}(T'_Y) = \mathcal{M}_{[[T_X]]}([[T'_Y]])$ (i.e., matches of tracelets are determined by matches of their evaluations), as well as of the fact that each admissible match μ is in fact a span of the form $in(T_X) \leftarrow M \rightarrow out(T'_Y)$, and thus a span of the format $\Phi(\mu) \in \mathcal{F}_{|in(T_X)|,|out(T_Y)|}$.

To proceed, note that for a given class of contributions $\hat{T}_X \circledast_F \hat{T}'_Y$, the number of connected components is given by $n_X - n_F^X + n_Y - n_F^Y + |F|$ (where n_X and n_Y denote the number of connected components of T_X and T_Y, respectively, where n_F^X and n_F^Y denote the sums of first/second entries in the elements of F, and where $|F|$ denotes the total number of pairs of integers F). Moreover, if we let $\pi \in \Pi_F$ denote a *partition* of F (into a part π and a part $F \setminus \pi$), one may verify that for a given admissible match μ with

$\Phi(\mu) = F$, if we let $\mu|_\pi$ and $\mu|_{F\setminus\pi}$ denote the restrictions of μ to the sub-formats π and $F\setminus\pi$, respectively, then $\mu|_\pi$ and $\mu|_{F\setminus\pi}$ are again admissible matches (of their respective domains and codomains). We thus obtain the following intermediate result:

$$\begin{aligned}\Delta(\hat{T}_X \diamond \hat{T}'_Y) &= \Delta\left(\sum_{F\in\mathfrak{F}_{|in(T_X)|,|out(T'_Y)|}} \hat{T}_X \circledast_F \hat{T}'_Y\right)\\ &= \sum_{F\in\mathfrak{F}_{|in(T_X)|,|out(T'_Y)|}} \sum_{\pi\in\Pi_F}\sum_{V\subseteq X}\sum_{W\subseteq Y} (\hat{T}_V \circledast_\pi \hat{T}'_W)\otimes(\hat{T}_{X\setminus V}\circledast_{F\setminus\pi}\hat{T}'_{Y\setminus W})\end{aligned} \tag{38}$$

Note that in the last step, we took advantage of the definition of $.\circledast_F.$, in that the factors $\hat{T}_V \circledast_\pi \hat{T}'_W$ cover indeed all possibilities for connected components arising from composing some connected components $\hat{T}_V$ of $\hat{T}_X$ with some connected components $\hat{T}_W$ of $\hat{T}_Y$ in the format π (and analogously for the other tensor factors with compositions of the format $F\setminus\pi$). It is now straightforward to verify that since we are summing over *all* possible formats F, and since $.\circledast_F.$ by definition only retains the non-zero contributions (i.e., those that come from admissible matches μ of format F), it follows that we are in fact evaluating all possible formats $F'\mathfrak{F}_{|in(V)|,|out(W)|}$ for the left and $F''\in\mathfrak{F}_{|in(X\setminus V)|,|out(Y\setminus W)|}$ for the right tensor factors, respectively, which finally leads to the proof of axiom IV:

$$\begin{aligned}\Delta(\hat{T}_X \diamond \hat{T}'_Y) &= \sum_{V\subseteq X}\sum_{W\subseteq Y}\sum_{F'\in\mathfrak{F}_{|in(T_V)|,|out(T'_W)|}}\sum_{F''\in\mathfrak{F}_{|in(T_{X\setminus V})|,|out(T'_{Y\setminus W})|}} (\hat{T}_V\circledast_{F'}\hat{T}'_W)\otimes(\hat{T}_{X\setminus V}\circledast_{F''}\hat{T}'_{Y\setminus W})\\ &= \sum_{V\subseteq X}\sum_{W\subseteq Y}(\hat{T}_V\diamond\hat{T}'_W)\otimes(\hat{T}_{X\setminus V}\diamond\hat{T}'_{Y\setminus W}) = (\mu\otimes\mu)\circ(Id\otimes\tau\otimes Id)\left(\Delta(\hat{T}_X)\otimes\Delta(\hat{T}'_Y)\right).\end{aligned} \tag{39}$$

B.6. Proof of Theorem 5.11. By definition, the coproduct Δ satisfies that for all $n\geq 0$ and $\hat{T}\in\mathfrak{T}^{(n)}$,

$$\Delta(\hat{T}) \in \sum_{m=0}^{n}\hat{\mathfrak{T}}^{(m)}\otimes\hat{\mathfrak{T}}^{(n-m)}. \tag{40}$$

Note in particular that $\hat{T}_\varnothing$ is indeed the only basis element for which the coproduct is an element of $\hat{\mathfrak{T}}^{(0)}\otimes\hat{\mathfrak{T}}^{(0)}$. Finally, by definition of tracelet superposition and of tracelet composition, it is clear that for all $m,n\geq 0$ and basis elements $\hat{T}\in\mathfrak{T}^{(m)}$ and $\hat{T}'\in\mathfrak{T}^{(n)}$,

$$T\diamond T' \in \sum_{r=0}^{m+n}\mathfrak{T}^{(r)}, \tag{41}$$

since the trivial overlap $\mu_\varnothing$ is always an admissible match (hence realizing $\hat{T}\uplus\hat{T}'\in\mathfrak{T}^{(m+n)}$, i.e., the contribution to $T\diamond T'$ with the highest possible filtration degree), and since non-trivial overlaps will reduce the overall number of connected components in general.

C Glossary

C.1. Simplicial groupoids. Let $\mathbb{\Delta}$ denote the category whose objects are the nonempty finite orders

$$[n] := \{0\leq 1\leq\cdots\leq n\}$$

and whose arrow are the monotone maps. A simplicial groupoid is a functor $X_\bullet : \mathbb{\Delta}^{\mathrm{op}}\to\boldsymbol{Grpd}$. The value of $X_\bullet$ on $[n]$ is denoted X_n. Exploiting the generators-and-relations description of $\mathbb{\Delta}$, one can describe a simplicial set as a diagram

$$X_0 \; \substack{\xleftarrow{\;d_1\;} \\ \xrightarrow{\;s_0\;} \\ \xleftarrow{\;d_0\;}} \; X_1 \; \substack{\xleftarrow{\;d_2\;} \\ \xrightarrow{\;s_1\;} \\ \xleftarrow{\;d_1\;} \\ \xrightarrow{\;s_0\;} \\ \xleftarrow{\;d_0\;}} \; X_2 \; \substack{\xleftarrow{\;d_3\;} \\ \xrightarrow{\;s_2\;} \\ \xleftarrow{\;d_2\;} \\ \xrightarrow{\;s_1\;} \\ \xleftarrow{\;d_1\;} \\ \xrightarrow{\;s_0\;} \\ \xleftarrow{\;d_0\;}} \; X_3 \qquad \cdots \tag{42}$$

subject to the *simplicial identities*: $d_i s_i = d_{i+1} s_i = 1$ and

$$d_i d_j = d_{j-1} d_i, \quad d_{j+1} s_i = s_i d_j, \quad d_i s_j = s_{j-1} d_i, \quad s_j s_i = s_i s_{j-1}, \qquad (i < j).$$

The indexing convention is that d_i deletes the ith vertex and s_i repeats the ith vertex.

C.2. Active and inert maps. Inside the simplex category $\mathbb{\Delta}$, we have the subcategory $\mathbb{\Delta}_{\text{inert}}$ of *inert* maps: these are the maps that preserve distance, meaning $\phi : [m] \to [n]$ such that $\phi(i+1) = \phi(i) + 1$. While $\mathbb{\Delta}$ is generated by face maps and degeneracy maps, $\mathbb{\Delta}_{\text{inert}}$ is generated only by the outer face maps. We also have the subcategory of *active* maps, which are precisely the endpoint-preserving maps in $\mathbb{\Delta}$. These two classes of maps form a factorization system on $\mathbb{\Delta}$: every map factors uniquely as an active map followed by an inert map.

For a simplicial groupoid $X_\bullet : \mathbb{\Delta}^{\text{op}} \to \boldsymbol{Grpd}$, we use the same terminology for simplicial operators induced by active and inert maps in $\mathbb{\Delta}$. For example, the inert face maps are precisely the outer face maps $d_0 : X_n \to X_{n-1}$ and $d_n : X_n \to X_{n-1}$ (for all $n > 0$), and the active maps are the inner face maps $d_i : X_n \to X_{n-1}$ for all $0 < i < n$. (All degeneracy maps are active.)

C.3. Decomposition spaces. It is a fact that active and inert maps in $\mathbb{\Delta}$ admit pushouts along each other and that the resulting new maps are again active and inert. A simplicial simplicial groupoid $X_\bullet$: $\mathbb{\Delta}^{\text{op}} \to \boldsymbol{Grpd}$ is called a *decomposition space* [16] when it sends active-inert pushouts in $\mathbb{\Delta}$ to (homotopy) pullbacks in $\boldsymbol{Grpd}$. It turns out it is enough to check the following squares to be (homotopy) pullbacks (for all $0 < i < n$):

$$\begin{array}{ccc} \mathbf{X}_{n+1} & \xrightarrow{d_{n+1}} & \mathbf{X}_n \\ {\scriptstyle d_i}\downarrow & & \downarrow{\scriptstyle d_i} \\ \mathbf{X}_n & \xrightarrow[d_n]{} & \mathbf{X}_{n-1} \end{array} \qquad\qquad \begin{array}{ccc} \mathbf{X}_{n+1} & \xrightarrow{d_0} & \mathbf{X}_n \\ {\scriptstyle d_{i+1}}\downarrow & & \downarrow{\scriptstyle d_i} \\ \mathbf{X}_n & \xrightarrow[d_0]{} & \mathbf{X}_{n-1} \end{array} \tag{43}$$

Particularly important are the first two instances of this:

$$\begin{array}{ccc} \mathbf{X}_3 & \xrightarrow{d_3} & \mathbf{X}_2 \\ {\scriptstyle d_1}\downarrow & & \downarrow{\scriptstyle d_1} \\ \mathbf{X}_2 & \xrightarrow[d_2]{} & \mathbf{X}_1 \end{array} \qquad\qquad \begin{array}{ccc} \mathbf{X}_3 & \xrightarrow{d_0} & \mathbf{X}_2 \\ {\scriptstyle d_2}\downarrow & & \downarrow{\scriptstyle d_1} \\ \mathbf{X}_2 & \xrightarrow[d_0]{} & \mathbf{X}_1 \end{array} \tag{44}$$

These conditions can be read as saying that a 3-simplex can be reconstructed by gluing two 2-simplices along a 1-simplex: the long edge of one along a short edge of the other.

An equivalent way of stating the decomposition-space axioms is that for every active map $\alpha : [k] \twoheadrightarrow [n]$, the square

$$\begin{array}{ccc} \mathbf{X}_n & \longrightarrow & \mathbf{X}_{n_1} \times \cdots \times \mathbf{X}_{n_k} \\ \downarrow & \lrcorner & \downarrow \\ \mathbf{X}_k & \longrightarrow & \mathbf{X}_1 \times \cdots \times \mathbf{X}_1 \end{array} \tag{45}$$

is a (homotopy) pullback (cf. [16, Prop. 6.9]). This time the condition says that an n-simplex can be reconstructed by gluing k simplices (of dimensions $n_1, \ldots, n_k$) onto the principal edges of a base k-simplex, in close analogy with the tracelet axioms.

C.4. Incidence coalgebras. The motivating property of decomposition space is that they allow the incidence-coalgebra construction (and with it many algebraic constructions from combinatorics, such as Möbius inversion). The incidence coalgebra of X has as underlying vector space $\mathbb{Q}_{\pi_0 X_1}$ (where $\pi_0 X_1$ is the set of iso-classes of 1-simplices), and the comultiplication law is defined as

$$\Delta(f) = \sum_{\sigma \in X_2 | d_1(\sigma) = f} d_2(\sigma) \otimes d_0(\sigma)$$

The decomposition-space axiom is designed to ensure that this comultiplication law is coassociative. (The 3-simplices enter the proof of coassociativity, and the higher simplices are useful for various purposes, as in the present case where tracelets are higher simplices.) If $X_\bullet$ is *monoidal* (for this notion, see [16]) then the incidence coalgebra becomes a bialgebra.

In the present paper, we are interested in the incidence algebra instead of the incidence coalgebra. It can generally be derived from the coalgebra as a convolution algebra. Even if the sum defining the incidence coalgebra is infinite, the corresponding sum in the convolution product becomes finite if restricted to finitely-supported functions [17].

C.5. About groupoids and homotopy pullbacks. One technicality that complicates the framework of decomposition spaces and combinatorial Hopf algebra is the necessity to work with groupoids instead of sets. Even if ultimately the incidence bialgebra will be generated by iso-classes, it does not work with decomposition 'sets' of iso-classes of objects (such as tracelets). The problem is that taking iso-classes too early kills important symmetries which must be respected in order for the constructions to work. This is explained in detail in [15]. Once this is understood it is actually a big simplification to work with groupoids, because with sets of iso-classes there are many pitfalls in connection with symmetries. With the groupoid formalism, all these issues are taken care of automatically.

In conclusion, one should always work with the naturally defined groupoids of objects, without trying to take iso-classes of pick representatives. A large amount of conditions take the form of pullbacks conditions. In the setting of groupoids, it is crucial to work with homotopy pullbacks instead of ordinary pullbacks, as the latter are not invariant under homotopy equivalence. An ordinary pullback of sets

$$\begin{array}{ccc} X \times_S Y & \longrightarrow & Y \\ \downarrow & & \downarrow q \\ X & \xrightarrow[p]{} & S \end{array} \tag{46}$$

is given by $X \times_S Y = \{(x,y) \mid px = qy\}$. For the homotopy pullback of groupoids, an explicit isomorphism is included as data:

$$X \times_S Y = \{(x,y,\sigma) \mid x \in X, y \in Y, \sigma : px \simeq qy\}$$

(These are the objects; the arrows are pairs of arrows compatible with the specified isos.) This can complicate certain calculations, but very often it is actually universal properties and standard manipulations that are important, and at that level of abstraction, homotopy pullbacks work very much like ordinary pullbacks do for sets.

In spite of their complicated appearance, homotopy pullbacks are often much closer to actual mathematical practice. For example, when we say 'glue together two 2-simplices along the long edge of one and the short edge of the other', it is unrealistic to assume that the long edge of one is literally equal to the short edge of the other. What actually happens is that they are *isomorphic*, and that a specific identification is employed (and must be referenced) in the constructions.

(We stress that homotopy pullbacks are used for groupoids, whereas inside ordinary categories, such as our fixed adhesive category $\mathcal{C}$, the notion of pullback is the ordinary 1-categorical one.)