Math. Struct. in Comp. Science (2008), *vol.* 18, *pp.* 1133–1167. © 2008 Cambridge University Press
doi:10.1017/S0960129508007160 First published online 22 October 2008 Printed in the United Kingdom

Using groups for investigating rewrite systems

PATRICK DEHORNOY[†] and VINCENT VAN OOSTROM[‡]

[†]*Laboratoire de Mathématiques Nicolas Oresme, Université de Caen, 14032 Caen, France*
Email: dehornoy@math.unicaen.fr
Website: www.math.unicaen.fr/ dehornoy
[‡]*Universiteit Utrecht, Heidelberglaan 6, 3584CS Utrecht, Netherlands*
Email: Vincent.vanOostrom@phil.uu.nl
Website: www.phil.uu.nl/ oostrom

Received 15 September 2006; revised 24 April 2008

We describe several technical tools that prove to be efficient for investigating the rewrite systems associated with an equational specification. These tools consist of introducing a monoid of partial maps, listing the monoid relations corresponding to the local confluence diagrams of the rewrite system, introducing the group presented by these relations, and, finally, replacing the initial rewrite system with an internal process entirely sitting in this group. When the approach can be completed, one typically obtains a practical method for constructing algebras satisfying prescribed equations and for solving the associated word problem. The above techniques have been developed by the first author in a context of general algebra. The goal of this paper is to bring them to the attention of the rewrite system community. We hope that these techniques may be useful for more general rewrite systems.

1. Introduction

Let $\mathscr{E}$ be an equational specification (typically associativity, commutativity, distributivity, *etc.*) involving a certain signature $\mathscr{F}$ of function symbols, and let $\mathscr{V}$ be an infinite set of variables. By choosing for each equation $l = r$ of $\mathscr{E}$ one preferred orientation, we obtain a rewrite rule and, therefore, a rewrite system $\mathscr{R}_{\mathscr{E}}$ on the family of all $\mathscr{F}$-terms constructed on $\mathscr{V}$. The aim of this paper is to present a general method for investigating the rewrite systems $\mathscr{R}_{\mathscr{E}}$ by introducing an associated monoid and, in good cases, a group. This approach has proved to be useful for various specifications $\mathscr{E}$, typically allowing one to solve the word problem or to construct models of $\mathscr{E}$ by means of the geometry monoid of $\mathscr{R}_{\mathscr{E}}$. The method may also be relevant for more general rewrite systems (Baader and Nipkow 1998; Dershowitz and Jouannaud 1994).

The approach comprises three steps. The first consists of associating with the rewrite system $\mathscr{R}_{\mathscr{E}}$ a certain inverse monoid $\mathscr{G}eom_{\mathscr{E}}$ of partial maps by taking into account where and in which direction the rules are applied; this monoid is called the *two-sided geometry monoid* for $\mathscr{E}$ as it captures several geometrical phenomena specific to $\mathscr{E}$ – and involves the two possible orientations of the equations under consideration.

The second step consists of replacing the inverse monoid $\mathscr{G}eom_{\mathscr{E}}$ with a group $\mathtt{Geom}_{\mathscr{E}}$ that resembles it: when the equations of $\mathscr{E}$ are simple enough, which typically means linear, the group $\mathtt{Geom}_{\mathscr{E}}$ can be defined simply to be the universal group of $\mathscr{G}eom_{\mathscr{E}}$; in more complicated cases, a convenient group can be obtained by investigating the local

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

confluence diagrams holding in $\mathcal{G}eom_{\mathcal{E}}$ and defining $\mathtt{Geom}_{\mathcal{E}}$ to be the group presented by the relations corresponding to these diagrams.

The third step, and in terms of technical development the main one, is an internalisation process that associates with every term a so-called blueprint of that term that lies in $\mathtt{Geom}_{\mathcal{E}}$ and replaces the external action of $\mathcal{G}eom_{\mathcal{E}}$ on terms by an internal multiplication in $\mathtt{Geom}_{\mathcal{E}}$.

When all three steps can be completed, various questions about the equational specification $\mathcal{E}$ can be successfully addressed, typically by constructing a model of $\mathcal{E}$, or by solving the universal word problem of $\mathcal{E}$.

In this paper we shall present the approach in a general setting and mention some of the existing examples. However, in order to make this paper more than a survey, we shall put the emphasis on a new example, for which the method had not been considered before, namely the case of the augmented left self-distributivity laws, or $\mathcal{ALD}$-laws, which are defined to be the following three laws:

$$\begin{cases} x * (y * z) = (x * y) * (x * z) \\ x * (y \circ z) = (x * y) \circ (x * z) \\ x * (y * z) = (x \circ y) * z. \end{cases} \qquad (\mathcal{ALD})$$

These laws and their models, called *$\mathcal{ALD}$-algebras* in this paper, have appeared recently in several frameworks (Dehornoy 2005; Dehornoy 2006; Dehornoy 2007), but they have never been addressed from the point of view of the geometry monoid. We shall see that several technical questions remain open in this seemingly difficult case. Nevertheless, the method is sufficient to lead naturally to the construction of a (highly non-trivial) example of a model, and to a solution of the word problem. Technically, the main step is the construction of what is called a blueprint for the $\mathcal{ALD}$-laws.

The leading principle underlying the approach is to use the geometry monoid to *guess* some formulas, and then to check the latter by a direct verification. This guessing process is guided by two simple principles: in order to guess relations in the geometry monoid $\mathcal{G}eom_{\mathcal{E}}$, we concentrate on the confluence diagrams of $\mathcal{R}_{\mathcal{E}}$ and translate them into equalities in $\mathcal{G}eom_{\mathcal{E}}$; in order to guess what to use as blueprint of a term, we associate with every term a distinguished reduction path connecting the term to some fixed base-term. The point is that this approach, which might seem loose, actually works in some definitely non-trivial cases: thus what legitimates the scheme is not some general *a priori* argument, but rather its *a posteriori* success.

The paper is organised as follows. The two-sided geometry monoid $\mathcal{G}eom_{\mathcal{E}}$ is introduced in Section 2. The process for going from this monoid to a group $\mathtt{Geom}_{\mathcal{E}}$ is described in Section 3. The principle of internalising terms in the geometry monoid/group is explained in Section 4. Next, we show in Section 5 how the general study developed in Sections 3 and 4 can be used, in good cases, to investigate an equational specification, with a special emphasis on the example of $\mathcal{ALD}$.

2. The geometry monoid $\mathcal{G}eom_{\mathcal{E}}$

The first step in our study consists of analysing a rewrite system by means of a monoid of partial maps. The general idea is as follows. If $\mathcal{R}$ is a term rewrite system, then, in the

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

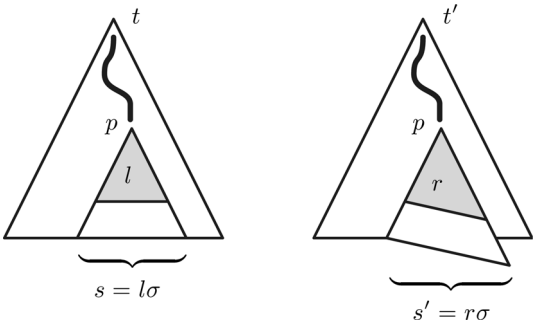


Fig. 1. Applying the law $l = r$ to a term t, here viewed as a rooted tree – we replace some subterm s of t that is a substitution instance of l with the corresponding substitution instance of r; or *vice versa.*

most general situation, several rules may be applied to a given term t. On the other hand, not every rule need be applicable to t, so the action of $\mathscr{R}$ on terms cannot be described by maps in a natural way. However, by fixing the rule and the position, we obtain uniqueness, and can describe $\mathscr{R}$ in terms of partial maps.

2.1. *Applying an algebraic law*

To fix terminology, we will recapitulate from Baader and Nipkow (1998), Dershowitz and Jouannaud (1994) and Terese (2003) the basic notions of rewriting needed for this paper. Our perspective here is that we will consider rewriting to be the application of a map, and that we study rewrite systems *via* the associated family of application maps.

For $\mathscr{F}$ a signature consisting of function symbols, and $\mathscr{V}$ a non-empty set of variables, we use $\mathscr{T}(\mathscr{F}, \mathscr{V})$ to denote the family of all terms built using function symbols from $\mathscr{F}$ and variables from $\mathscr{V}$. In practice, we shall assume that $\mathscr{F}$ and $\mathscr{V}$ are clear from the context, and simply write $\mathscr{T}$ for $\mathscr{T}(\mathscr{F}, \mathscr{V})$. An equation is a pair of terms (l, r), which we will usually denote as $l = r$.

In this context, the fundamental operation of *applying* an equation, typically applying an algebraic law such as associativity, to a term corresponds to a rewrite system: applying $l = r$ to some term t means replacing some subterm s of t which happens to be a substitution instance of l with the corresponding substitution instance s' of r, or *vice versa.* This means that there exists a position p and a substitution σ, that is, a mapping from $\mathscr{V}$ into $\mathscr{T}(\mathscr{F})$, such that the pth subterm of t is $l\sigma$ and t' is obtained from t by replacing the pth subterm with $r\sigma$; or *vice versa*, with the roles of l and r exchanged (Figure 1). In other words, we apply one of the two rules $l \to r$, $r \to l$.

Definition 2.1. For each equational specification $\mathscr{E}$ involving the signature $\mathscr{F}$, we use $\mathscr{R}_{\mathscr{E}}^{\leftrightarrow}$ to denote the rewrite system on $\mathscr{T}$ consisting of all rules $l \to r$ and $r \to l$ for $l = r$ an equation of $\mathscr{E}$; we use $\mathscr{R}_{\mathscr{E}}$ to denote the rewrite system consisting of the rules $l \to r$ alone.

Thus, trivially, we have the following proposition.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

Proposition 2.2. For each equational specification $\mathscr{E}$, and for all terms t, t' in $\mathscr{T}$, the relation $t =_{\mathscr{E}} t'$ holds if and only if we have $t =_{\mathscr{R}^{\leftrightarrow}_{\mathscr{E}}} t'$.

2.2. *Geometry monoid: the principle*

The rewrite systems $\mathscr{R}^{\leftrightarrow}_{\mathscr{E}}$ are not functional: there may be many ways to apply one of its rules to a given term. However, provided all equations are regular, these systems can be viewed as the union of a family of partial functions, each of which corresponds to choosing an equation, an orientation, and a position.

For the rest of this paper it is convenient to fix a way to access the subterms of a term by means of positions. As is usual, we shall see terms as rooted oriented trees (*cf.* Figure 1), where nodes are labelled using function symbols and variables. Then a subterm of a term t is naturally specified by the node where its root lies, which is itself determined by the path that connects the root of the tree to that node. If all function symbols are binary, which will be the case in the examples considered below, we can, for instance, use finite sequences of 0's and 1's to describe such a path, using 0 for 'forking to the left' and 1 for 'forking to the right'. We use ε for the empty position, that is, the position of the root. If t is a term and p is a position, we use $t|_p$ to denote the subterm of t at position p. Note that $t|_p$ exists only for p short enough: $t|_\varepsilon$ always exists and equals t, but, for instance, $t|_0$ and $t|_1$, which are the left and right subterms, respectively, of t, exist only if t is not a variable or a constant.

Definition 2.3.

(i) Assume that E is an equation. For each position p, we use $\rightarrow_{E,p}$ (respectively, $\leftarrow_{E,p}$) to denote the (partial) map on terms corresponding to applying E at position p in left–right (respectively, right–left) direction.

(ii) For $\mathscr{E}$ an equational certification, we define the *two-sided geometry monoid of* $\mathscr{E}$, denoted $\mathscr{G}eom_{\mathscr{E}}$, to be the monoid generated by all maps $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$ where E ranges over $\mathscr{E}$ and p ranges over all positions, equipped with the composition of maps viewed as relations, and with the identity mapping of $\mathscr{T}$.

We shall always think of the maps $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$ as acting on the right, thus writing $t \cdot g$ rather than $g(t)$ for the image of the term t under the map g. To be coherent, we use the reversed composition, denoted '$;$', in the geometry monoid. Thus, $g\,;g'$ means 'g then g''.

Remark 2.4. When a term s is a substitution instance of the term l, the substitution σ such that $s = l\sigma$ is not unique, since the value of $x\sigma$ is uniquely determined only for those variables x that occur in l. Hence, for $E = (l, r)$, the map $\rightarrow_{E,p}$ is functional only if all variables occurring in r occur in l. The equations that satisfy this condition for both $\rightarrow_{E,\varepsilon}$ and $\leftarrow_{E,\varepsilon}$ are those for which the same variables occur in l and in r. Such equations are usually called *regular*. Although this is not necessary, we shall always restrict to regular equations in the rest of this paper.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

By construction, every element in a geometry monoid $\mathcal{G}eom_{\mathcal{E}}$ is a finite product of maps $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$ with E in $\mathcal{E}$. It will be convenient in the rest of this paper to fix the following notation.

Notation 2.5. Assume that $\mathcal{E}$ is an equational specification. For each equation E in $\mathcal{E}$ and each position p, we introduce two letters $\mathtt{E}_p$ and $\mathtt{E}_p^{-1}$, and use $W_{\mathcal{E}}$ to denote the free monoid consisting of all words on the letters $\mathtt{E}_p$ and $\mathtt{E}_p^{-1}$. Then we use *eval* to denote the canonical evaluation morphism

$$eval : W_{\mathcal{E}} \rightarrow \mathcal{G}eom_{\mathcal{E}}$$

that maps $\mathtt{E}_p$ to $\rightarrow_{E,p}$ and $\mathtt{E}_p^{-1}$ to $\leftarrow_{E,p}$ for all E and p.

Thus, $W_{\mathcal{E}}$ consists of all formal products of letters $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$, while $\mathcal{G}eom_{\mathcal{E}}$ consists of actual maps on terms. We shall see soon that, in general (as may be expected), the evaluation mapping is far from injective, that is, the geometry monoid $\mathcal{G}eom_{\mathcal{E}}$ is far from free.

2.3. *Geometry monoid: the example of* $\mathcal{ALD}$

To illustrate our approach, we consider the specification ($\mathcal{ALD}$), standing for 'Augmented Left self-Distributivity', consisting of the following three equations:

$$x * (y * z) = (x * y) * (x * z) \tag{D}$$

$$x * (y \circ z) = (x * y) \circ (x * z) \tag{D'}$$

$$x * (y * z) = (x \circ y) * z. \tag{A}$$

The specific interest in this choice is that, unlike the case for the specification ($\mathcal{LD}$) consisting of D alone, the above mixed equations, which in the rest of this paper will be collectively referred to as ($\mathcal{ALD}$), have not yet been investigated from the viewpoint of the geometry monoid and, therefore, the results we present below are new.

Here the signature consists of two binary function symbols $*, \circ$. The equation D is the left self-distributivity law, which was extensively studied in Dehornoy (2000). The additional equations D' and A express the fact that $*$ is left distributive with respect to $\circ$ and that $\circ$ behaves like a sort of composition relative to $*$. Many examples of $\mathcal{LD}$-algebras, that is, structures satisfying D, happen to be equipped with a second operation that satisfies the mixed equations D' and A. This is, in particular, the case for every group equipped with the $\mathcal{LD}$-operation $x * y := xyx^{-1}$: in this case, defining the second operation to be the multiplication $x \circ y := xy$ yields an $\mathcal{ALD}$-algebra – and more in fact, namely an $\mathcal{LD}$-monoid in the sense of Dehornoy (2000, Chapter XI).

By definition, the geometry monoid $\mathcal{G}eom_{\mathcal{ALD}}$ is generated by three families of maps, corresponding to the three equations, namely $\rightarrow_{D,p}$, $\rightarrow_{D',p}$, $\rightarrow_{A,p}$, and their inverses. These maps correspond to the left-to-right orientation of the equations as written above – this choice is not important, as, in the two-sided geometry monoid, both orientations are considered. In this specific case we observe that the actions of the maps $\rightarrow_{D,p}$ and $\rightarrow_{D',p}$ are similar in that both consist of distributing the left subterm to the two halves of the

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

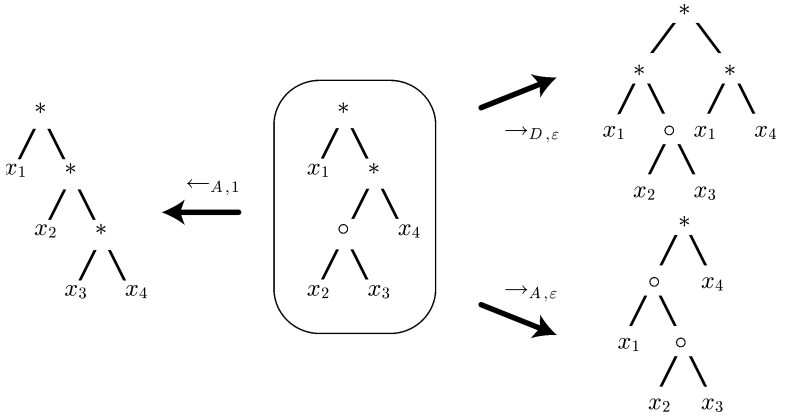


Fig. 2. Case of $\mathscr{ALD}$: two left-to-right maps apply to the term $x_1 * ((x_2 \circ x_3) * x_4)$, namely $\rightarrow_{D,\varepsilon}$ and $\rightarrow_{A,\varepsilon}$, while only one right-to-left map applies to it, namely $\leftarrow_{A,1}$, that is, the copy of $\leftarrow_{A,\varepsilon}$ acting at position 1.

right subterm. However, their domains are disjoint, as $\rightarrow_{D,p}$ applies only when the symbol at $p1$ is $*$, while $\rightarrow_{D',p}$ applies only when the symbol at $p1$ is $\circ$. In order to obtain a smaller family of generators, rather than considering two maps separately, it will be convenient to introduce their union, which is still functional, and denote it by $\rightarrow_{D,p}$. So, we have

$$\rightarrow_{D,\varepsilon} : t_1 * (t_2 \,\square\, t_3) \rightarrow (t_1 * t_2) \,\square\, (t_1 * t_3), \tag{1}$$

that is, $\rightarrow_{D,\varepsilon}$ maps every term of the form $t_1 * (t_2 \,\square\, t_3)$ to the corresponding term $(t_1 * t_2) \,\square\, (t_1 * t_3)$, where $\square$ stands for either $*$ or $\circ$. On the other hand, we have

$$\rightarrow_{A,\varepsilon} : t_1 * (t_2 * t_3) \rightarrow (t_1 \circ t_2) * t_3, \tag{2}$$

that is, $\rightarrow_{A,\varepsilon}$, also denoted $\rightarrow_{A,\varepsilon}$, maps every term of the form $t_1 * (t_2 * t_3)$ to $(t_1 \circ t_2) * t_3$. By definition, the geometry monoid $\mathscr{Geom}_{\mathscr{ALD}}$ is the monoid generated by all partial maps $\rightarrow_{D,p}$, $\leftarrow_{D,p}$, $\rightarrow_{A,p}$, and $\leftarrow_{A,p}$ with p ranging over $\{0,1\}^*$. As shown in Figure 2, a term generally belongs to the domain of several such maps.

Remark 2.6. For a signature $\mathscr{F}$ comprising more than one function symbol, as is the case with $\mathscr{ALD}$, just taking into account the position where a equation is applied does not exhaust all information. Indeed, when a map $\rightarrow_{E,p}$ is applied to a term t, the complete list of the function symbols that occur in t above p may be important, typically in terms of the relations possibly connecting various maps. In order to include such data, we can use addresses instead of positions to index the maps of the geometry monoid, an *address* being defined to be a finite sequence of even length consisting of alternating function symbols and forking directions: for instance, $(*,1)$ and $(*,1,\circ,0)$ are addresses that refine the positions 1 and 10 respectively. In this framework, the map $\mathtt{A}_1^-$ mentioned in Figure 2 would become $\mathtt{A}_{(*,1)}^-$, that is, $\mathtt{A}_1^-$ applied to a term with $*$ at the root. Observe that a map $\rightarrow_{E,p}$ as defined in Definition 2.3 is, as a set of pairs of terms, just the disjoint union of all maps $\rightarrow_{E,\alpha}$ with α an address that projects on the position p when the function symbols are forgotten. Although these refinements may be necessary in some cases, they

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

are not required in the examples considered in this paper, in particular, they are not required for $\mathscr{ALD}$, and there will be no need to split the maps $\rightarrow_{E,p}$ into more elementary components.

2.4. *Connection between $\mathscr{Geom}_{\mathscr{E}}$ and equality in $\mathscr{E}$*

We saw in Proposition 2.2 that equality in $\mathscr{E}$ is directly connected with the equivalence relation associated with the rewrite system $\mathscr{R}^{\leftrightarrow}_{\mathscr{E}}$. As a consequence, it is also connected with the two-sided geometry monoid $\mathscr{Geom}_{\mathscr{E}}$, which is a geometric way of expressing the so-called logicality, that is, the coincidence between derivability in equational logic and the convertibility relation of the corresponding rewrite system (Bertling and Ganzinger 1989).

Proposition 2.7. Let $\mathscr{E}$ be an equational specification. Then two terms t, t' are equal in $\mathscr{E}$ if and only if some element of the two-sided geometry monoid $\mathscr{Geom}_{\mathscr{E}}$ maps t to t'.

Proof. Let us write $t \equiv t'$ if some map of $\mathscr{Geom}_{\mathscr{E}}$ maps t to t'. First, $t \equiv t'$ implies $t =_{\mathscr{E}} t'$. Indeed, by construction, both $t' = t\cdot \rightarrow_{E,p}$ and $t' = t\cdot \leftarrow_{E,p}$ imply $t =_{\mathscr{E}} t'$. As $=_{\mathscr{E}}$ is transitive, the same is true when several maps $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$ are composed.

Conversely, by Proposition 2.2, $=_{\mathscr{E}}$ coincides with $=_{\mathscr{R}^{\leftrightarrow}_{\mathscr{E}}}$, so, in order to prove that $t =_{\mathscr{E}} t'$ implies $t \equiv t'$, it is enough to prove that $\equiv$ is a congruence on terms, and that it contains all instances of the laws of $\mathscr{E}$. Now, $\equiv$ is an equivalence relation because $\mathscr{Geom}_{\mathscr{E}}$ is closed under composition and contains the identity mapping, and it is a congruence, that is, it is compatible with all operations of $\mathscr{F}$. Indeed, assume that $\rightarrow_{E,p}$ maps t to t'. Let s be any term and $\Box$ be any operation in $\mathscr{F}$. Then, assuming that all operations of $\mathscr{F}$ are binary and $0, 1$ positions are used, the map $\rightarrow_{E,0p}$ maps $s \Box t$ to $s \Box t'$, while $\rightarrow_{E,1p}$ maps $t \Box s$ to $t' \Box s$. Finally, assume that (t, t') is an instance of some equation E of $\mathscr{E}$. This means that there exists a substitution σ such that, assuming that E is $l = r$, we have $t = l\sigma$ and $t' = r\sigma$. Now, by definition, $\rightarrow_{E,\varepsilon}$ maps t to t', so $t \equiv t'$ holds. □

It is not our aim in this paper to develop a general theory of the geometry monoid. However, we will mention two results: one about the structure of $\mathscr{Geom}_{\mathscr{E}}$; the other about its dependence on $\mathscr{E}$.

First, note that, by construction, each map in the geometry monoid $\mathscr{Geom}_{\mathscr{E}}$ is close to admitting an inverse. Indeed, the map $\leftarrow_{E,p}$ is, as a set of pairs, the inverse of the map $\rightarrow_{E,p}$, that is, $\rightarrow_{E,p}$ maps t to t' if and only if $\leftarrow_{E,p}$ maps t' to t. This is not enough to make $\mathscr{Geom}_{\mathscr{E}}$ into a group, as the composition $\rightarrow_{E,p}\ ;\ \leftarrow_{E,p}$ is the identity map of its domain only, and *not* the identity map of $\mathscr{T}(\mathscr{F})$ in general, but we do have the following proposition.

Proposition 2.8. Assume that $\mathscr{E}$ is a family of regular equations. Then $\mathscr{Geom}_{\mathscr{E}}$ is an inverse monoid.

Proof. Assume that g is a non-empty map in $\mathscr{Geom}_{\mathscr{E}}$. Then, by construction, there exists a word w in $W_{\mathscr{E}}$ such that g equals $eval(w)$. Let w^{-1} be the formal inverse of w, that is, the word obtained from w by exchanging $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$ everywhere in w and reversing the order of the letters, and let $g' := eval(w^{-1})$. An immediate induction shows that a

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

pair of terms (t,t') belongs to g if and only if the pair (t',t) belongs to g', and we deduce

$$g\,;g'\,;g=g' \quad \text{and} \quad g'\,;g\,;g'=g,$$

so every element in $\mathtt{Geom}_{\mathscr{E}}$ possesses an inverse, as required for an inverse monoid: in particular, this applies to $g=g'=\varnothing$. □

As for the second result, the *equational variety* associated with an equational specification $\mathscr{E}$ involving the signature $\mathscr{F}$ is, by definition, the collection of all $\mathscr{F}$-algebras that are models of $\mathscr{E}$. Different specifications may give rise to the same variety: for instance, when adjoined to the commutativity law $xy=yx$, the associativity law $x(yz)=(xy)z$ and the law $x(yz)=z(yx)$ define the same variety.

Proposition 2.9 (Dehornoy 1993). Up to isomorphism, the monoid $\mathscr{G}eom_{\mathscr{E}}$ only depends on the equational variety defined by $\mathscr{E}$: if $\mathscr{E}$ and $\mathscr{E}'$ define the same variety, the monoids $\mathscr{G}eom_{\mathscr{E}}$ and $\mathscr{G}eom_{\mathscr{E}'}$ are isomorphic.

Sketch of proof. It suffices to show that if we add to a specification $\mathscr{E}$ a new equation E that is a consequence of the equations of $\mathscr{E}$, then the geometry monoid is not changed. Assume $E=(l,r)$. By Proposition 2.7, the hypothesis that E is a consequence of $\mathscr{E}$ implies that some map g in $\mathscr{G}eom_{\mathscr{E}}$ maps l to r. The geometry monoid $\mathscr{G}eom_{\mathscr{E}\cup\{E\}}$ is obtained from $\mathscr{G}eom_{\mathscr{E}}$ by adding a new generator that is a product of the canonical generators of $\mathscr{G}eom_{\mathscr{E}}$, and, therefore, it is isomorphic to $\mathscr{G}eom_{\mathscr{E}}$. □

3. Replacing the geometry monoid with a group

A first step that often proves useful in investigating the geometry monoids is to replace the monoid $\mathscr{G}eom_{\mathscr{E}}$ with a *group* $\mathtt{Geom}_{\mathscr{E}}$ that resembles it, in that they share some defining relations. The benefit of this lies in the possibility of freely computing with inverses and avoiding the problems connected with the fact that the composition of two non-empty maps may be the empty map. However, there is no universal recipe for going from $\mathscr{G}eom_{\mathscr{E}}$ to a (non-trivial) group, and we shall complete the approach in specific cases only.

3.1. *The linear case*

For each inverse monoid M, there exists a biggest quotient of M that is a group, namely the *universal group* $U(M)$ of M obtained by collapsing all idempotents to 1 (Paterson 1998). When the geometry monoid $\mathscr{G}eom_{\mathscr{E}}$ does not contain the empty map, the idempotent elements are those maps that are identity on their domain, and the associated group then keeps enough information to be non-trivial. We shall first mention one case when this favourable situation occurs, namely the case of *linear* equations. This case can be seen as a first step for more difficult cases to come later.

Definition 3.1. A term t is said to be *linear* if no variable occurs twice in t. An equation $l=r$ is said to be *linear* if the terms l and r are linear and, moreover, the equation is regular. An equational specification $\mathscr{E}$ is said to be *regular* (respectively, *linear*) if all equations in $\mathscr{E}$ are regular (respectively, linear).

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

For instance, associativity and commutativity are linear equations, but self-distributivity is not, as the variable x is repeated twice in the right-hand term of $x(yz) = (xy)(xz)$.

In order to describe the case of linear equations, we will start with a general result about the elements of a geometry monoid.

Definition 3.2. Assume that g is a map on $\mathscr{T}$. We say that a pair of terms (l, r) is a *seed* for g if g, as a set of pairs, consists of all instances of (l, r), that is, it consists of all pairs $(l\sigma, r\sigma)$ with σ a $\mathscr{T}$-valued substitution.

If E is the equation $l = r$, then $\rightarrow_{E,\varepsilon}$ consists of all instances of (l, r), so (l, r) is a seed for $\rightarrow_{E,\varepsilon}$. More generally, we have the following lemma.

Lemma 3.3 (Dehornoy 2000). Assume that $\mathscr{E}$ is a regular equational specification involving a single binary function symbol. Then each non-empty map g in $\mathscr{G}eom_{\mathscr{E}}$ admits a seed, that is, it is an instance of some pair of $\mathscr{R}^{\leftrightarrow}_{\mathscr{E}}$-convertible terms.

Sketch of proof. First, (x, x) is a seed for the identity mapping of $\mathscr{T}$. Then, every map $\rightarrow_{E,p}$ admits a seed, because, if (l, r) is a seed for $\rightarrow_{E,\varepsilon}$ and x is a variable not occurring in l and r, then, assuming that $*$ is the only function symbol, $(x * l, x * r)$ is a seed for $\rightarrow_{E,1}$, $(l * x, r * x)$ is a seed for $\rightarrow_{E,0}$, and an easy induction gives the result for every $\rightarrow_{E,p}$.

Then, by construction, every map in $\mathscr{G}eom_{\mathscr{E}}$ is a finite composition of maps $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$ with E in $\mathscr{E}$, and, by definition, the latter admit seeds. Hence, the point is to show that the composition of two maps admitting a seed still admits a seed provided it is non-empty. So we assume that g_1 and g_2 consist of all instances of (l_1, r_1) and (l_2, r_2), respectively. Then, either there exists no unifier for r_1 and l_2, and in this case $g_1 ; g_2$ is empty, or there exists a unifier, and then (Robinson 1963) there exists a most general unifier (MGU) of the terms r_1 and l_2, that is, there exist two substitutions σ, τ satisfying $r_1\sigma = l_2\tau$ and such that every common instance of r_1 and l_2 is an instance of $r_1\sigma$. In this case, it is easy to check that $(l_1\sigma, r_2\tau)$ is a seed for $g_1 ; g_2$. □

Remark 3.4. The restriction on the signature in Lemma 3.3 can be dropped at the expense of splitting the maps $\rightarrow_{E,p}$ according to addresses as explained in Remark 2.6, or, alternatively, by considering a more general notion of terms and instances in which substitution is possible not only for variables but also for some specific function symbols considered as variables of higher type. Once again, we shall not go into details here, as such notions will not be used in the rest of this paper.

The following result appears as Lemma 2.3 in Dehornoy (2005).

Lemma 3.5. Assume that $\mathscr{E}$ is a linear equational specification involving a single binary function symbol. Then the seed of every map in $\mathscr{G}eom_{\mathscr{E}}$ is a pair of linear terms. Moreover, $\mathscr{G}eom_{\mathscr{E}}$ does not contain the empty map.

Proof. The MGU of two linear terms always exists and it is a linear term: indeed, the only reason that may cause a unifier not to exist is a variable clash, that is, a so-called occur-check, and this cannot happen with linear terms. So the result for the seed follows

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

from an induction. The fact that the required unifier always exists guarantees that the empty map cannot appear. □

Recall that a term t is called *linear* if no variable appears more than once in t.

Definition 3.6. For $\mathscr{E}$ an equational specification, and g, g' in $\mathscr{G}eom_{\mathscr{E}}$, we say that $g \sim g'$ (respectively, $g \approx g'$) holds if we have $t \cdot g = t \cdot g'$ for at least one linear term t (respectively, for each term t for which both $t \cdot g$ and $t \cdot g'$ are defined.

When the geometry monoid $\mathscr{G}eom_{\mathscr{E}}$ contains the empty map, the relations $\sim$ and $\approx$ need not be transitive. However, this cannot happen in the linear case.

Proposition 3.7 (Dehornoy 2005). Assume that $\mathscr{E}$ is a linear equational specification involving a single binary function symbol. Then the relations $\sim$ and $\approx$ coincide, and they are congruences on the monoid $\mathscr{G}eom_{\mathscr{E}}$. Furthermore, the quotient-monoid $\mathscr{G}eom_{\mathscr{E}}/\approx$ is a group, and it is the universal group of $\mathscr{G}eom_{\mathscr{E}}$.

Proof. Because $\mathscr{E}$ contains only linear equations, each map in $\mathscr{G}eom_{\mathscr{E}}$ is non-empty, and it admits a seed consisting of linear terms. In particular, for all g, g' in $\mathscr{G}eom_{\mathscr{E}}$, the map $g^{-1}\,;g'$ is non-empty, and its seed consists of linear terms. Hence there exists a linear term t such that both $t \cdot g$ and $t \cdot g'$ are defined. Hence $g \approx g'$ implies $g \sim g'$.

Conversely, assume $g \sim g'$, say $t \cdot g = t \cdot g'$ with t a linear term. By hypothesis, g has a seed (l, r), and g' has a seed (l', r'). Let l_* be the MGU of l and l'. By construction, there exist substitutions σ, σ', τ satisfying $l_* = l\sigma = l'\sigma'$ and $t = l_*\tau$, hence $t = l\sigma\tau = l'\sigma'\tau$. Then, we find $t \cdot g = r\sigma\tau = t \cdot g' = r'\sigma'\tau$. As t is linear, we deduce $r\sigma = r'\sigma'$, hence $l_* \cdot g = l_* \cdot g'$. Now let t_* be any term such that both $t_* \cdot g$ and $t_* \cdot g'$ are defined. By hypothesis, t_* is an instance of both l and l', hence we must have $t_* = l_*\sigma_*$ for some substitution σ_*, and we find $t_* \cdot g = (l_* \cdot g)\sigma_* = (l_* \cdot g')\sigma_* = t_* \cdot g'$. So $g \approx g'$ holds.

The remaining verifications are now straightforward. □

Under the hypotheses of Proposition 3.7, we use $U(\mathscr{G}eom_{\mathscr{E}})$ to denote the quotient-group $\mathscr{G}eom_{\mathscr{E}}/\approx$, that is, the universal group of the (two-sided) geometry monoid $\mathscr{G}eom_{\mathscr{E}}$. So in this case we have a scheme of the form

$$W_{\mathscr{E}} \xrightarrow{\text{onto}} \mathscr{G}eom_{\mathscr{E}} \xrightarrow{\text{onto}} U(\mathscr{G}eom_{\mathscr{E}}). \tag{3}$$

Moreover, the partial action of $\mathscr{G}eom_{\mathscr{E}}$ on terms induces a well-defined action of $\mathtt{Geom}_{\mathscr{E}}$ since, by definition, all maps in a $\approx$-class agree on the terms that lie in their domain. Furthermore, no information is lost when we replace $\mathscr{G}eom_{\mathscr{E}}$ with $U(\mathscr{G}eom_{\mathscr{E}})$ in the precise sense that the counterpart of Proposition 2.7 is true: two terms t, t' are equal in $\mathscr{E}$ if and only if we have $t' = t \cdot g$ for some g in $U(\mathscr{G}eom_{\mathscr{E}})$. So in this case we can replace the monoid $\mathscr{G}eom_{\mathscr{E}}$ with the group $U(\mathscr{G}eom_{\mathscr{E}})$ in all further uses.

Example 3.8. Let $\mathscr{A}$ denote the associativity law $x(yz) = (xy)z$. Then the corresponding group $U(\mathscr{G}eom_{\mathtt{A}})$ is a well-known group, namely R. Thompson's group F (Thompson 1980). In the case of associativity together with commutativity, the corresponding group is R. Thompson's group V – *cf.* Dehornoy (2005).

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

Remark 3.9. Even in the smooth case of linear equations, the action of the group $U(\mathscr{G}eom_{\mathscr{E}})$ on terms is a partial action: for $t \cdot g$ to be defined, t must be large enough. However, we could obtain an everywhere defined action by considering infinite trees.

3.2. *Local confluence relations: the principle*

Whenever the empty map occurs in the geometry monoid, the previous approach fails badly.

Lemma 3.10. Assume that $\mathscr{G}eom_{\mathscr{E}}$ contains the empty map. Then the universal group $U(\mathscr{G}eom_{\mathscr{E}})$ is trivial, that is, it reduces to $\{1\}$.

Proof. Assume that π is a homomorphism of $\mathscr{G}eom_{\mathscr{E}}$ to a group. Then, for each g in $\mathscr{G}eom_{\mathscr{E}}$, we have $\varnothing ; g = \varnothing$, hence $\pi(\varnothing) \cdot \pi(g) = \pi(\varnothing)$, and thus $\pi(g) = 1$. □

Example 3.11. The case when there is no unifier frequently occurs. For instance, in the case of the self-distributivity law ($\mathscr{LD}$), every term t belonging to the range of $\rightarrow_{D,\varepsilon}$ satisfies $t|_{00} = t|_{10}$; it then follows that every term t in the range of $\rightarrow_{D,\varepsilon} ; \rightarrow_{D,1}$ satisfies $t|_{00} = t|_{100}$, and therefore $t|_{00} \neq t|_{10}$. Hence no term in the image of $\rightarrow_{D,\varepsilon} ; \rightarrow_{D,1}$ may belong to the domain of $\leftarrow_{D,\varepsilon}$: in other words, the composed map $\rightarrow_{D,\varepsilon} ; \rightarrow_{D,1} ; \leftarrow_{D,\varepsilon}$, as a set of ordered pairs, is the empty set, and, in particular, its domain of definition is empty. In other words, we have in $\mathscr{G}eom_{\mathscr{LD}}$, as well as in $\mathscr{G}eom_{\mathscr{ALD}}$, the relation

$$eval(\mathtt{D}_\varepsilon \mathtt{D}_1 \mathtt{D}_\varepsilon^{-1}) = \varnothing.$$

This implies that any word w in $W_{\mathscr{LD}}$ containing $\mathtt{D}_\varepsilon \mathtt{D}_1 \mathtt{D}_\varepsilon^{-1}$ as a factor evaluates in $\mathscr{G}eom_{\mathscr{LD}}$ to the empty map. In particular, this is the case for the symmetric word

$$\mathtt{D}_\varepsilon \mathtt{D}_1 \mathtt{D}_\varepsilon^{-1} \mathtt{D}_\varepsilon \mathtt{D}_1^{-1} \mathtt{D}_\varepsilon^{-1},$$

though the latter appears as freely reducing to the empty word and might therefore be expected to represent a map that is close to the identity.

The above situation almost always occurs when non-linear equations are involved. As the example of $(xy)(xz) = (xz)(xy)$ shows, it is not true that the presence of at least one non-linear equation in $\mathscr{E}$ forces the empty map to belong to $\mathscr{G}eom_{\mathscr{E}}$: what matters here are the various geometric ways in which a variable is moved. For instance, it is sufficient for implementing the argument used in Example 3.11 that $\mathscr{E}$ contains an equation $l = r$ such that, for some variable x, the sets $\{p_1, \ldots, p_i\}$ and $\{q_1, \ldots, q_j\}$ showing where x occurs in l and r, respectively, are distinct, at least one of them is not a singleton and the length of all positions q_m (as sequences of 0's and 1's) are different from the length of all positions p_ℓ.

In such cases, the universal group $U(\mathscr{G}eom_{\mathscr{E}})$ is of no use, and we have to look for another method. One could try to modify the construction of the geometry monoid so as to artificially discard the empty map, but we doubt that anything interesting would result: indeed, any group that could be subsequently be attached is likely to be trivial – see Dehornoy (2000) for a more thorough discussion. Instead, we shall now develop a completely different method for associating with $\mathscr{G}eom_{\mathscr{E}}$ a group that keeps the meaningful

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

information of $\mathcal{G}eom_{\mathcal{E}}$, namely finding a presentation of $\mathcal{G}eom_{\mathcal{E}}$ and introducing the group that admits this presentation – whatever its connection with $\mathcal{G}eom_{\mathcal{E}}$ is.

Actually, finding a presentation of $\mathcal{G}eom_{\mathcal{E}}$ is in general beyond reach, at least by a direct approach. So, once again, we use an indirect approach consisting of isolating *some* relations satisfied in $\mathcal{G}eom_{\mathcal{E}}$ using a uniform scheme, but not trying to prove that these relations make a presentation: the latter may come at the very end when further constructions have been performed.

So, at this point, the problem is to find relations connecting the various maps $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$ of $\mathcal{G}eom_{\mathcal{E}}$. In the rest of this paper we shall concentrate on *positive* relations, that is, on relations that involve the maps $\rightarrow_{E,p}$ but not their inverses.

Definition 3.12. Given an equational specification $\mathcal{E}$, the *one-sided geometry monoid* $\mathcal{G}eom^{+}_{\mathcal{E}}$ of $\mathcal{E}$ is defined to be the submonoid of $\mathcal{G}eom_{\mathcal{E}}$ generated by all maps $\rightarrow_{E,p}$ with E in $\mathcal{E}$ and p in $\{0,1\}^*$.

The connection between the rewrite system $\mathcal{R}_{\mathcal{E}}$ and the one-sided geometry monoid $\mathcal{G}eom^{+}_{\mathcal{E}}$ is similar to the connection between the rewrite system $\mathcal{R}^{\rightarrow}_{\mathcal{E}}$ and the two-sided geometry monoid $\mathcal{G}eom_{\mathcal{E}}$.

Proposition 3.13. Let $\mathcal{E}$ be an equational specification. Then t rewrites in t' with $\mathcal{R}_{\mathcal{E}}$ if and only if some element of $\mathcal{G}eom^{+}_{\mathcal{E}}$ maps t to t'.

The verification is the same as for Proposition 2.7.

We shall be looking for relations in $\mathcal{G}eom^{+}_{\mathcal{E}}$, that is, for relations connecting the maps $\rightarrow_{E,p}$ in $\mathcal{G}eom^{+}_{\mathcal{E}}$. The principle is to investigate the relations that may arise when two maps are applied to one and the same term t, which amounts to investigating the local confluence of the rewrite system $\mathcal{R}_{\mathcal{E}}$. This means that, for all equations E, E' in $\mathcal{E}$ and all positions p, q, we look for relations of the generic form

$$\rightarrow_{E,p} \ ; \ldots = \rightarrow_{E',q} \ ; \ldots \tag{4}$$

In algebraic terms, this amounts to looking for common right multiples in $\mathcal{G}eom^{+}_{\mathcal{E}}$.

Definition 3.14. For g a (partial) map on $\mathcal{T}(\mathcal{F})$ and p a position, we use $sh_p(g)$ to denote the *p-shift* of g, which is defined to be the partial map on $\mathcal{T}(\mathcal{F})$ that consists of applying g to the pth subterm of its argument: $t \cdot sh_p(g)$ is defined if and only if $t|_p$ exists and $t|_p \cdot g$ is defined, and, in this case, $t \cdot sh_p(g)$ is obtained from t by replacing the pth subterm by its image under g.

So, for instance, we have $\rightarrow_{E,p} = sh_p(\rightarrow_{E,\varepsilon})$ for each equation E and each position p, and, more generally, $\rightarrow_{E,pq} = sh_p(\rightarrow_{E,q})$ for all p, q.

What follows is a refinement of Huet's Critical Pair Lemma (see Lemma 2.7.15 in Terese (2003)) in which we explicitly take positions into account. Two general schemes will be used to identify local confluence relations in the monoid $\mathcal{G}eom^{+}_{\mathcal{E}}$. The first corresponds to the disjoint/parallel case.

Lemma 3.15. Assume that p and q are parallel positions, that is, there exists o such that $o0$ is a prefix of p and $o1$ is a prefix of q, or *vice versa*. Then, for all partial maps g, h, we

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

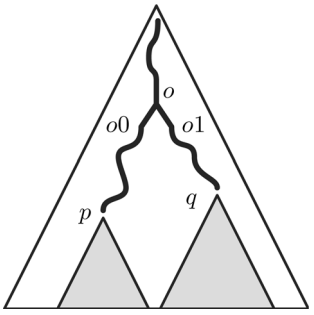


Fig. 3. Parallel case: the pth and qth subterms are disjoint, so maps acting on them commute.

have

$$sh_p(g)\,;sh_q(h) = sh_q(h)\,;sh_p(g). \tag{5}$$

Proof. (Figure 3) The maps $sh_p(g)$ and $sh_q(h)$ act on disjoint subtrees, and therefore they commute. □

In particular, we deduce the following proposition.

Proposition 3.16. Assume that p and q are parallel positions, that is, we have $o0 \leqslant p$ and $o1 \leqslant q$, or conversely, for some o. Then for all equations E, E', we have

$$\rightarrow_{E,p}\ ;\ \rightarrow_{E',q} = \rightarrow_{E',q}\ ;\ \rightarrow_{E,p}\ . \tag{6}$$

The second general scheme corresponds to nested redexes.

Lemma 3.17. Assume that E is the oriented equation (l, r) and that some variable x occurs at positions $p_1, \ldots, p_i$ in l and at $q_1, \ldots, q_j$ in r. Then, for each (partial) map g acting on terms, we have

$$sh_{p_1}(g)\,;\ldots;sh_{p_i}(g)\,;\ \rightarrow_{E,\varepsilon}\ =\ \rightarrow_{E,\varepsilon}\ ;sh_{q_1}(g)\,;\ldots;sh_{q_j}(g). \tag{7}$$

More generally, for each position o, we have

$$sh_{op_1}(g)\,;\ldots;sh_{op_i}(g)\,;\ \rightarrow_{E,o}\ =\ \rightarrow_{E,o}\ ;sh_{oq_1}(g)\,;\ldots;sh_{oq_j}(g). \tag{8}$$

Proof. (Figure 4) Assume that $\rightarrow_{E,\varepsilon}$ maps t to t'. This means that there exists a substitution σ such that t is $l\sigma$ and t' is $r\sigma$. Let σ_1 be the substitution defined by $y\sigma_1 = y\sigma$ for $y \neq x$, and $x\sigma_1 = x\sigma g$. Let $t_1 := l\sigma_1$ and $t'_1 := r\sigma_1$. Then, by construction, $\rightarrow_{E,\varepsilon}$ maps t_1 to t'_1. Now, t_1 is obtained from t by replacing the subterms at positions $p_1, \ldots, p_i$ with their image under g, so we have

$$t_1 = t \cdot (sh_{p_1}(g)\,;\ldots;sh_{p_i}(g)).$$

Similarly, t'_1 is obtained from t' by replacing the subterms at positions $q_1, \ldots, q_j$ with their image under g, so we have

$$t'_1 = t' \cdot (sh_{q_1}(g)\,;\ldots;sh_{q_j}(g)),$$

and (7) follows.

Relation (8) is deduced by applying sh_o to both terms of (7). □

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

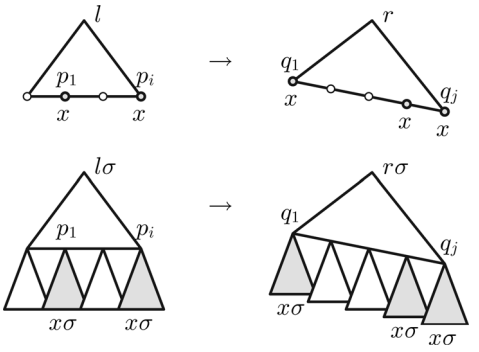


Fig. 4. Nested case: if the variable x occurs at $p_1,\dots,p_i$ in l and at $q_1,\dots,q_j$ in r and nowhere else, then for each term $l\sigma$, the subterm $x\sigma$ occurs at $p_1,\dots,p_i$ in $l\sigma$, and t $q_1,\dots,q_j$ in $r\sigma$, and applying g in each $x\sigma$ before or after applying $l \to r$ leads to the same result.

By applying Lemma 3.17 to the case when the map g has the form $\to_{E',p'}$, we obtain the following refinement of Terese (2003, Lemma 2.7.15(b)).

Proposition 3.18. Assume that E is the equation (l,r) and that some variable x occurs at positions $p_1,\dots p_i$ in l and at $q_1,\dots,q_j$ in r. Then for all positions o,o' and each equation E', we have

$$\to_{E,o} \;;\; \to_{E',oq_1o'} \;;\dots;\; \to_{E',oq_jo'} = \to_{E',op_1o'} \;;\dots;\; \to_{E',op_io'} \;;\; \to_{E,o}\,. \tag{9}$$

In general, the two previous general schemes do not exhaust all possible local confluence relations in $\mathscr{G}eom_{\mathscr{E}}$. Typically, no relation is obtained for pairs of the form (o,op) where p is so short that it gives an overlap. In this case, which corresponds to the overlap case (c) of Terese (2003, Lemma 2.7.15), no general scheme may exist, and one has to look at the specific rewrite system under consideration in order to find possible local confluence relations.

3.3. *Local confluence relations: the example of $\mathscr{ALD}$*

We illustrate the previous scheme in the case of $\mathscr{ALD}$. The first family of local confluence relations comprises the trivial commutation relations corresponding to the parallel cases, here all relations

$$\to_{E,p} \;;\; \to_{E',q} = \to_{E',q} \;;\; \to_{E,p} \quad \text{for } E,E' = D \text{ or } A, \text{ and } p,q \text{ parallel.} \tag{10}$$

The second family of local confluence relations corresponds to nested redexes. As for the laws D and D', that is, for the map $\to_{D,\varepsilon}$, three variables are involved in the rule

$$x * (y \,\square\, z) \to (x * y) \,\square\, (x * z). \tag{11}$$

The variable x occurs at position 0 on the left-hand side of (11), while it occurs at 00 and 10 on the right-hand side. So relation (9) is here

$$\to_{D,p} \;;\; \to_{E,p0q} = \to_{E,p00q} \;;\; \to_{E,p10q} \;;\; \to_{D,p} \tag{12}$$

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

for $E = D$ or A. Similarly, the variable y occurs at 10 on the left-hand side of (11), and at 01 on the right-hand side, while the variable z occurs at 11 on both sides. Thus Relations (9) become

$$\rightarrow_{D,p} \ ; \rightarrow_{E,p10q} = \rightarrow_{E,p01q} \ ; \rightarrow_{D,p} \tag{13}$$

$$\rightarrow_{D,p} \ ; \rightarrow_{E,p11q} = \rightarrow_{E,p11q} \ ; \rightarrow_{D,p} \quad \text{for } E = D \text{ or } A. \tag{14}$$

The treatment is analogous for the equation A, that is, for the map $\rightarrow_{A,\varepsilon}$. Three variables occur in the rule

$$x * (y * z) \rightarrow (x \circ y) * z. \tag{15}$$

The variable x occurs at 0 on the left-hand side of (15), and at 00 on the right-hand side; y occurs at 10 and at 01, respectively; finally, z occurs at 11 and at 1. The corresponding relations (9) are

$$\rightarrow_{A,p} \ ; \rightarrow_{E,p0q} = \rightarrow_{E,p00q} \ ; \rightarrow_{A,p} \tag{16}$$

$$\rightarrow_{A,p} \ ; \rightarrow_{E,p10q} = \rightarrow_{E,p01q} \ ; \rightarrow_{A,p} \tag{17}$$

$$\rightarrow_{A,p} \ ; \rightarrow_{E,p11q} = \rightarrow_{E,p1q} \ ; \rightarrow_{A,p}, \tag{18}$$

again for $E = D$ or A.

With the previous approach, we found one relation of type (4) for all pairs p, q, with the exception of the pairs $\{\rightarrow_{A,p}, \rightarrow_{D,p}\}$, as well as all pairs $\{\rightarrow_{E,p}, \rightarrow_{E',p1}\}$. According to the current option, the next question is whether we can find in $\mathscr{Geom}_{\vec{\mathscr{ALD}}}$ relations for the critical pairs of the overlap cases

$$\rightarrow_{E,p} \ ; \ldots = \rightarrow_{E',p} \ ; \ldots \quad \text{and} \quad \rightarrow_{E,p} \ ; \ldots = \rightarrow_{E',p1} \ ; \ldots$$

when $(\rightarrow_{E,\varepsilon}, \rightarrow_{E',\varepsilon})$ ranges over the various combinations of $\rightarrow_{D,\varepsilon}$ and $\rightarrow_{A,\varepsilon}$. As shown in Figure 5, the following relations are satisfied:

$$\rightarrow_{D,p} \ ; \rightarrow_{D,p1} \ ; \rightarrow_{D,p} = \rightarrow_{D,p1} \ ; \rightarrow_{D,p} \ ; \rightarrow_{D,p1} \ ; \rightarrow_{D,p0} \tag{19}$$

$$\rightarrow_{D,p} \ ; \rightarrow_{D,p1} \ ; \rightarrow_{A,p} = \rightarrow_{A,p1} \ ; \rightarrow_{D,p} \ ; \rightarrow_{D,p0} \tag{20}$$

$$\rightarrow_{A,p} \ ; \rightarrow_{D,p} = \rightarrow_{D,p1} \ ; \rightarrow_{D,p} \ ; \rightarrow_{A,p1} \ ; \rightarrow_{A,p0} . \tag{21}$$

Remark 3.19. Not all possible cases are covered: there is *no* relation $\rightarrow_{A,p} \ldots = \rightarrow_{D,p} \ldots$, or $\rightarrow_{A,p} \ldots = \rightarrow_{A,p1} \ldots$ in the list above. There may exist no such relation involving positive maps $\rightarrow_{A,\ldots}$ and $\rightarrow_{D,\ldots}$ only: for instance, there is no way to rewrite the terms $(x \circ y) * z$ and $(x * y) \square (x * z)$ into a common third term. So using a critical pair completion is hopeless here.

On the other hand, note that for all positions p, q and all combinations of $\rightarrow_{E,\varepsilon}, \rightarrow_{E',\varepsilon}$, we have found at most one relation $\rightarrow_{E,p} \ ; \ldots = \rightarrow_{E',q} \ ; \ldots$. We do not claim that we found found all of the confluence relations holding in the geometry monoid $\mathscr{Geom}_{\mathscr{E}}$: we have only applied a heuristic approach that has led us to some relations – in good cases, we can subsequently prove that the list is actually complete, but there is no *a priori* evidence.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

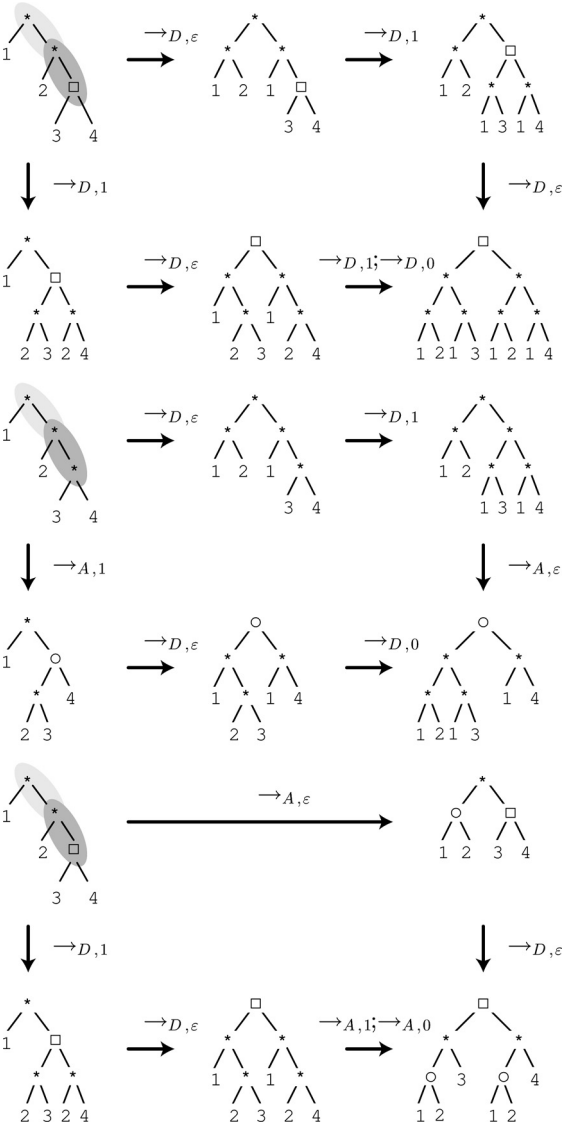


Fig. 5. Overlapping case: here □ stands for both ∗ or ∘, the numbers stand for the indices of the variables and the overlapping patterns are highlighted in grey.

3.4. *The geometry group*

According to the principle stated above, we now introduce the group for which the confluence relations of $\mathscr{G}eom_{\mathscr{E}}$ make a presentation. According to Remark 3.19, there is some fuzziness in the definition, as there is in general no proof that the list of confluence relations is complete.

Definition 3.20. For $\mathscr{E}$ a family of oriented equations, the *geometry group* $\mathtt{Geom}_{\mathscr{E}}$ of $\mathscr{E}$ is defined to be the group generated by elements $\mathtt{E}_p$ with E in $\mathscr{E}$ and p a position, subject

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

to the confluence relations connecting the corresponding maps in the monoid $\mathcal{G}eom^{+}_{\mathscr{E}}$, as found using the approach of Section 3.2. We use `eval` to denote the canonical evaluation morphism of $W_{\mathscr{E}}$ onto $\mathtt{Geom}_{\mathscr{E}}$.

Example 3.21. In the case of the $\mathscr{ALD}$-laws, the group $\mathtt{Geom}_{\mathscr{ALD}}$ is, by definition, a group generated by two infinite series of generators $\mathtt{D}_p$ and $\mathtt{A}_p$ indexed by finite sequences of 0's and 1's, subject to the relations (10), (12), (13), (14), (19), (20) and (21).

Example 3.22. In the case of the (linear) associativity law $\mathscr{A}$, one easily checks that, for each pair of positions p, q, there exists a local confluence relation of the form $\rightarrow_{A,p}\ ;\ldots=\rightarrow_{A,q}\ ;\ldots$ in the positive monoid $\mathcal{G}eom^{+}_{\mathscr{A}}$ – see also Melliès (2002). Indeed, besides the quasi-commutativity relations

$$\rightarrow_{A,p}\ ;\ \rightarrow_{A,p0q}=\rightarrow_{A,p00q}\ ;\ \rightarrow_{A,p},\qquad \rightarrow_{A,p}\ ;\ \rightarrow_{A,p10q}=\rightarrow_{A,p01q}\ ;\ \rightarrow_{A,p},$$
$$\rightarrow_{A,p}\ ;\ \rightarrow_{A,p11q}=\rightarrow_{A,p1q}\ ;\ \rightarrow_{A,p}$$

that come from nested redexes, the only missing pairs are the pairs $p, p1$, and the well-known Mac Lane–Stasheff pentagon relation (Mac Lane 1963)

$$\rightarrow_{A,p}\ ;\ \rightarrow_{A,p}=\rightarrow_{A,p1}\ ;\ \rightarrow_{A,p}\ ;\ \rightarrow_{A,p1}$$

completes the picture. Then the group $\mathtt{Geom}_{\mathtt{A}}$ presented by the above relations turns out to coincide with the group $U(\mathcal{G}eom_{\mathtt{A}})$ of Proposition 3.7, that is, Thompson's group F – see Dehornoy (2005): so, in this case, the method based on presentation by confluence relations subsumes that of Section 3.1.

Question 3.23. Does the group $\mathtt{Geom}_{\mathscr{E}}$ coincide with the group $U(\mathcal{G}eom_{\mathscr{E}})$ for each linear equational specification $\mathscr{E}$?

The connection between the monoid $\mathcal{G}eom_{\mathscr{E}}$ and the group $\mathtt{Geom}_{\mathscr{E}}$ is not obvious in general. Indeed, instead of the sequence of (3), the scheme is now

$$\begin{array}{ccc} W_{\mathscr{E}} & \xrightarrow[eval]{\text{onto}} & \mathcal{G}eom_{\mathscr{E}} \\ {\scriptstyle \text{onto}}\big\downarrow{\scriptstyle \mathtt{eval}} & & \\ \mathtt{Geom}_{\mathscr{E}} & & \end{array} \tag{22}$$

and it is not clear whether any factorisation connects $\mathcal{G}eom_{\mathscr{E}}$ and $\mathtt{Geom}_{\mathscr{E}}$. In order to prove that there exists a morphism from $\mathcal{G}eom_{\mathscr{E}}$ to $\mathtt{Geom}_{\mathscr{E}}$, we ought to know that the selected confluence relations exhaust all relations holding in $\mathcal{G}eom_{\mathscr{E}}$, which cannot be the case if the empty map occurs. And, in order to prove that there exists a morphism from $\mathtt{Geom}_{\mathscr{E}}$ to $\mathcal{G}eom_{\mathscr{E}}$, we ought to be able to control the pairs $\rightarrow_{E,p}\ ;\ \leftarrow_{E,p}$ in $\mathcal{G}eom_{\mathscr{E}}$. More precisely, assume that w, w' are words in $W_{\mathscr{E}}$ that satisfy $\mathtt{eval}(w) = \mathtt{eval}(w')$, that is, that represent the same element of $\mathtt{Geom}_{\mathscr{E}}$. This means that there exists a finite sequence of words $w = w_0, w_1, \ldots, w_n = w'$ such that each w_i is obtained from the previous one either by applying one of the local confluence relations, or by deleting some subfactor $\mathtt{E}_p\mathtt{E}_p^{-1}$ or $\mathtt{E}_p^{-1}\mathtt{E}_p$, or by inserting such a subfactor. In the first two cases, the associated maps are equal, but, in the third case, the associated maps need not be equal, and,

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

in particular, the empty map may appear. So we cannot deduce from the fact that w, w' represent the same element in $\mathtt{Geom}_{\mathscr{E}}$ the fact that they represent the same element in $\mathscr{Geom}_{\mathscr{E}}$. However, the important point is that, even if we cannot directly compare $\mathscr{Geom}_{\mathscr{E}}$ and $\mathtt{Geom}_{\mathscr{E}}$, we can use the properties of $\mathscr{Geom}_{\mathscr{E}}$ as a sort of oracle for guessing properties of $\mathtt{Geom}_{\mathscr{E}}$ that can then possibly be proved using syntactic arguments based on the explicit presentation – according to a scheme that was used successfully in the investigation of the self-distributivity law in Dehornoy (2000).

4. Internalisation of terms

At this point we have associated a monoid $\mathscr{Geom}_{\mathscr{E}}$, and in some cases, which have not been characterised precisely, a group $\mathtt{Geom}_{\mathscr{E}}$, with each family of equations $\mathscr{E}$. By construction, there is a partial action of the monoid $\mathscr{Geom}_{\mathscr{E}}$ on terms. The next step is to try to carry terms inside the monoid $\mathscr{Geom}_{\mathscr{E}}$ and, simultaneously, inside the group $\mathtt{Geom}_{\mathscr{E}}$, so as to replace the external action of the monoid or the group on terms by an internal operation, typically a multiplication. Given this vague description, many solutions could be thought of: for instance, mapping every term t to the partial mapping that is the identity on t and every instance of t. We do not know how to use such 'trivial' solutions; the solutions we shall present are much more specific and they depend heavily on the equational specification under consideration. We will explain how to complete the construction in good cases, and, in particular, in the case of the $\mathscr{ALD}$-laws.

4.1. *Blueprint of a term: the principle*

A priori, terms and maps of the geometry monoid live in disjoint worlds: the only connection is that maps act on terms. The principle we shall apply in the following, which turns out to be efficient in good cases, is to build inside the geometry monoid $\mathscr{Geom}_{\mathscr{E}}$, and simultaneously inside the geometry group $\mathtt{Geom}_{\mathscr{E}}$ that mimics its algebraic properties, a copy of each term, so that the action of maps on terms translates into a simple operation inside $\mathscr{Geom}_{\mathscr{E}}$ and $\mathtt{Geom}_{\mathscr{E}}$, typically a multiplication. This copy of the term t in $\mathscr{Geom}_{\mathscr{E}}$ (respectively, in $\mathtt{Geom}_{\mathscr{E}}$) will be called the $\mathscr{Geom}_{\mathscr{E}}$*-blueprint* (respectively, the $\mathtt{Geom}_{\mathscr{E}}$-blueprint) of t, to suggest that it mimics the construction of t *via* some *ad hoc* conversion.

Convention 4.1. In the following, we shall develop parallel constructions in the geometry monoid $\mathscr{Geom}_{\mathscr{E}}$ and in its abstract counterpart the group $\mathtt{Geom}_{\mathscr{E}}$. To make the distinction easier, we use italic fonts for notions involving $\mathscr{Geom}_{\mathscr{E}}$, and the corresponding typescript fonts for the $\mathtt{Geom}_{\mathscr{E}}$-counterparts, as we have done with the names $\mathscr{Geom}_{\mathscr{E}}$ and $\mathtt{Geom}_{\mathscr{E}}$ themselves.

Recall that, for $\mathscr{F}$ a list of function symbols, $\mathscr{T}(\mathscr{F})$ denotes the family of all terms constructed using the operations of $\mathscr{F}$ and variables from an infinite list $\mathscr{V}$. In the following, we assume that x is a fixed element of $\mathscr{V}$, and use $\mathscr{T}_x$ for the family of all terms constructed using the operations of $\mathscr{F}$ and the single variable x.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

The general principle is as follows. Assume that $\mathscr{E}$ is an equational specification involving the signature $\mathscr{F}$, and we have an injective map ('blueprint')

$$\mathscr{B}p : \mathscr{T}_X \to \mathscr{G}eom_{\mathscr{E}}, \tag{23}$$

that is, a representation of terms inside the geometry monoid $\mathscr{G}eom_{\mathscr{E}}$. If t, t' are equal in $\mathscr{E}$, that is, by Proposition 2.7, if some element g of $\mathscr{G}eom_{\mathscr{E}}$ maps t to t', we can expect that there exists a simple relation between the copies $\mathscr{B}p(t)$ and $\mathscr{B}p(t')$. We shall be interested in the (certainly very special) case when there exists an endomorphism $\mathscr{B}p_*$ of $\mathscr{G}eom_{\mathscr{E}}$ such that the relation takes the form

$$\mathscr{B}p(t \cdot g) \sim \mathscr{B}p(t)\,;\mathscr{B}p_*(g) \tag{24}$$

that is, when the action of $\mathscr{G}eom_{\mathscr{E}}$ on terms becomes a right multiplication at the level of the copies – optimally, we might hope for an equality in (24), but the problem of the empty map makes a true equality impossible in most cases: this is precisely why we shall subsequently resort to the group $\mathtt{Geom}_{\mathscr{E}}$. Recall that $W_{\mathscr{E}}$ denotes the family of all words in the letters $\mathtt{E}_p^{\pm 1}$ for E in $\mathscr{E}$ and p an position. By construction, every map in $\mathscr{G}eom_{\mathscr{E}}$ is a finite product of elementary maps $\mathtt{E}_p^{\pm 1}$, so it can be expressed as $eval(w)$ where w is a word in $W_{\mathscr{E}}$, and (24) can be restated as

$$\mathscr{B}p(t \cdot eval(w)) \sim \mathscr{B}p(t)\,;\mathscr{B}p_*(eval(w)), \tag{25}$$

for t in $\mathscr{T}_X$ and w in $W_{\mathscr{E}}$ such that $t \cdot eval(w)$ is defined.

Now, according to the general principle of Section 3, we wish to replace the monoid $\mathscr{G}eom_{\mathscr{E}}$ with the group $\mathtt{Geom}_{\mathscr{E}}$. If the construction of the mapping $\mathscr{B}p$ of $\mathscr{T}_X$ to $\mathscr{G}eom_{\mathscr{E}}$ and of the endomorphism $\mathscr{B}p$ of $\mathscr{G}eom_{\mathscr{E}}$ is explicit enough, we can mimic them in the group $\mathtt{Geom}_{\mathscr{E}}$, thus defining a similar (hopefully injective) map

$$\mathtt{Bp} : \mathscr{T}_X \to \mathtt{Geom}_{\mathscr{E}}, \tag{26}$$

and an endomorphism $\mathtt{Bp}_*$ of $\mathtt{Geom}_{\mathscr{E}}$, and then using $\mathtt{Bp}(t)$ as a copy of t inside $\mathtt{Geom}_{\mathscr{E}}$. If $\mathtt{Geom}_{\mathscr{E}}$ resembles $\mathscr{G}eom_{\mathscr{E}}$ closely enough, we may hope that the counterpart of (25) holds in $\mathtt{Geom}_{\mathscr{E}}$ so that we can carry the action of $\mathscr{G}eom_{\mathscr{E}}$ on terms inside the group $\mathtt{Geom}_{\mathscr{E}}$. We are thus led to the following notion.

Definition 4.2. Assume that $\mathscr{E}$ is an equational specification involving the signature $\mathscr{F}$, and that $\mathtt{Bp}_*$ is an endomorphism of $\mathtt{Geom}_{\mathscr{E}}$. A mapping $\mathtt{Bp} : \mathscr{T}_X \to \mathtt{Geom}_{\mathscr{E}}$ is said to be a $\mathtt{Bp}_*$-*blueprint* if the relation

$$\mathtt{Bp}(t \cdot eval(w)) = \mathtt{Bp}(t) \cdot \mathtt{Bp}_*(\mathtt{eval}(w)) \tag{27}$$

holds in $\mathtt{Geom}_{\mathscr{E}}$ for every term t and every word w in $W_{\mathscr{E}}$ such that $t \cdot eval(w)$ is defined.

Thus, a $\mathtt{Bp}_*$-blueprint transforms the operation of applying the laws of $\mathscr{E}$ into a $\mathtt{Bp}_*$-twisted multiplication in the group $\mathtt{Geom}_{\mathscr{E}}$. The interest of a $\mathtt{Bp}_*$-blueprint will be discussed in Section 5 below. The general idea is that it allows us to carry the problems about equality in $\mathscr{E}$ into the presented group $\mathtt{Geom}_{\mathscr{E}}$, and therefore may lead to solutions when the latter is under control. The general scheme is as follows. Assume that the map $\mathtt{Bp}$

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

is injective. Then carrying the relation $=_{\mathscr{E}}$ to $\mathtt{Geom}_{\mathscr{E}}$ using $\mathtt{Bp}$ yields a new equivalence relation inside $\mathtt{Geom}_{\mathscr{E}}$, and our goal is to investigate the latter equivalence relation directly.

Remark 4.3. A constant mapping provides a blueprint, though certainly a trivial and uninteresting one. In the following, we are mainly interested in blueprints that are injective or close to, but we shall see that in certain cases, such as $\mathscr{ALD}$, even a blueprint that is not proved to be injective may be useful. That is why we do not require injectivity in Definition 4.2.

4.2. *Blueprint of a term: the example of* $\mathscr{ALD}$

In order to carry out the approach sketched above, that is, to construct a blueprint, in the case of the $\mathscr{ALD}$-laws, the first step is to represent terms in $\mathscr{Geom}_{\mathscr{E}}$, that is, to select for each term t a certain map $\mathscr{Bp}(t)$ in the geometry monoid $\mathscr{Geom}_{\mathscr{E}}$ that in some sense characterises t. A general idea is to choose a map that *constructs* t starting from some fixed absolute base-term. This procedure depends heavily on the specific equations under investigation, which are the $\mathscr{ALD}$-laws here. We use $\mathscr{T}^{*,\circ}_{x}$ for the family of all well-formed terms involving two binary maps $*, \circ$ and the single variable x.

The solution we develop relies on some specific property of the $\mathscr{ALD}$-laws, namely the existence of an *absorption* phenomenon. To decribe this phenomenon, let us define *right combs* to be those terms of $\mathscr{T}^{*,\circ}_{x}$ specificed inductively by

$$x^{[n]} := \begin{cases} x & \text{for } n = 1 \\ x * x^{[n-1]} & \text{for } n \geqslant 2. \end{cases} \tag{28}$$

The associated trees are 'all to the right' degenerate trees, which are also called spines in lambda calculus. The following result expresses the fact that, in the presence of the laws of $\mathscr{ALD}$, every term t of $\mathscr{T}^{*,\circ}_{x}$ is absorbed by all sufficiently large right combs. We begin by associating with every term a natural number. We use $\mathbb{N}$ for the set of non-negative integers.

Lemma 4.4. For p, q in $\mathbb{N}$, define $p * q = q$, $p \circ q = p + q$. Then $(\mathbb{N}, *, \circ)$ is an $\mathscr{ALD}$-algebra.

The verification is straightforward.

Definition 4.5. For each term t, we use $\varphi(t)$ to denote the evaluation of t in $(\mathbb{N}, *, \circ)$ when x is given the value 1.

Lemma 4.6 (Absorption lemma). For each term t in $\mathscr{T}^{*,\circ}_{x}$, the equality

$$x^{[n]} =_{\mathscr{ALD}} t * x^{[n-\varphi(t)]} \tag{29}$$

holds for n large enough.

Proof. The property is true for $t = x$ and $n \geqslant 2$ (in which case the equivalence is an equality), so, in order to establish it for every term in $\mathscr{T}^{*,\circ}_{x}$, it is enough to show that if (29) holds for t_1 and t_2, then it holds for $t_1 * t_2$ and for $t_1 \circ t_2$ as well. So we assume

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

$x^{[n]} =_{\mathscr{ALD}} t_1 * x^{[n-\varphi(t_1)]}$ for $n \geqslant m_1$ and $x^{[n]} =_{\mathscr{ALD}} t_2 * x^{[n-\varphi(t_2)]}$ for $n \geqslant m_2$. We obtain for $n \geqslant \max(m_1 + \varphi(t_2), m_2 + \varphi(t_1))$

$$\begin{aligned} x^{[n]} &=_{\mathscr{ALD}} t_1 * x^{[n-\varphi(t_1)]} && \text{(by hypothesis)} \\ &=_{\mathscr{ALD}} t_1 * (t_2 * x^{[n-\varphi(t_1)-\varphi(t_2)]}) && \text{(by hypothesis)} \\ &=_{\mathscr{ALD}} (t_1 * t_2) * (t_1 * x^{[n-\varphi(t_1)-\varphi(t_2)]}) && \text{(by } D\text{)} \\ &=_{\mathscr{ALD}} (t_1 * t_2) * x^{[n-\varphi(t_2)]} && \text{(by hypothesis)} \\ &= t * x^{[n-\varphi(t)]}. && \text{(by definition of } \varphi\text{)} \end{aligned}$$

Similarly, we have for $n \geqslant \max(m_1, m_2 + \varphi(t_1))$

$$\begin{aligned} x^{[n]} &=_{\mathscr{ALD}} t_1 * x^{[n-\varphi(t_1)]} && \text{(by hypothesis)} \\ &=_{\mathscr{ALD}} t_1 * (t_2 * x^{[n-\varphi(t_1)-\varphi(t_2)]}) && \text{(by hypothesis)} \\ &=_{\mathscr{ALD}} (t_1 \circ t_2) * x^{[n-\varphi(t_1)-\varphi(t_2)]} && \text{(by } A\text{)} \\ &= t * x^{[n-\varphi(t)]} && \text{(by definition of } \varphi\text{)} \end{aligned}$$

so the induction is complete. □

By Proposition 2.7, the equivalence of (29) must be witnessed for by some map of the geometry monoid $\mathscr{Geom}_{\mathscr{ALD}}$: for each term t in $\mathscr{T}^{*,\circ}_{x}$ and every integer n, there must exist a map $\mathscr{Bp}(t)$ in $\mathscr{Geom}_{\mathscr{ALD}}$, hence a composition of maps $\rightarrow_{D,p}$, $\leftarrow_{D,p}$, $\rightarrow_{A,p}$, and $\leftarrow_{A,p}$, that maps $x^{[n]}$ to $t * x^{[n-\varphi(t)]}$. Actually, the inductive proof of Lemma 4.6 not only proves the existence of such a witness, but also gives an inductive construction for it.

Lemma 4.7. For t in $\mathscr{T}^{*,\circ}_{x}$, define $\mathscr{Bp}(t)$ in $\mathscr{Geom}_{\mathscr{ALD}}$ inductively by

$$\mathscr{Bp}(t) = \begin{cases} \mathrm{id} & \text{for } t = x \\ \mathscr{Bp}(t_1)\,;sh_1(\mathscr{Bp}(t_2))\,;\rightarrow_{D,\varepsilon}\,;sh_1(\mathscr{Bp}(t_1))^{-1} & \text{for } t = t_1 * t_2 \\ \mathscr{Bp}(t_1)\,;sh_1(\mathscr{Bp}(t_2))\,;\rightarrow_{A,\varepsilon} & \text{for } t = t_1 \circ t_2. \end{cases} \tag{30}$$

Then, for every term t in $\mathscr{T}^{*,\circ}_{x}$ and every n large enough, we have

$$x^{[n]} \cdot \mathscr{Bp}(t) = t * x^{[n-\varphi(t)]}. \tag{31}$$

Proof. The formulas of (43) are just a translation of the successive equivalence steps in the proof of Lemma 4.6, and the result is then a straightforward verification. □

In other words, $\mathscr{Bp}(t)$, which is a map on $\mathscr{T}^{*,\circ}_{\mathscr{V}}$, hence, in particular, on $\mathscr{T}^{*,\circ}_{x}$, maps $x^{[n]}$ to $t * x^{[n-\varphi(t)]}$ for n large enough. Note that (31) guarantees that the mapping $\mathscr{Bp}$ is injective, since the term t can be recovered from the map $\mathscr{Bp}(t)$. So it is coherent to use the map $\mathscr{Bp}(t)$ as a counterpart of the term t inside the geometry monoid $\mathscr{Geom}_{\mathscr{ALD}}$. According to the general scheme of Section 4.1, we now analyse the counterpart of the action of $\mathscr{Geom}_{\mathscr{E}}$ on terms.

Lemma 4.8. For each term t in $\mathscr{T}^{*,\circ}_{x}$ and each map g in $\mathscr{Geom}_{\mathscr{ALD}}$ such that $t \cdot g$ is defined, we have

$$\mathscr{Bp}(t \cdot g) \sim \mathscr{Bp}(t)\,;sh_0(g). \tag{32}$$

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

Proof. Let $t' = t \cdot g$. Then, for each term t_1, the map $sh_0(g)$ maps the term $t * t_1$ to $t' * t_1$, and, similarly, $t \circ t_1$ to $(t \cdot g) \circ t_1$. So, in particular, $sh_0(g)$ maps $t * x^{[n]}$ to $t' * x^{[n]}$ for each n. Now, by Lemma 4.7 (and for n large enough), the map $\mathscr{Bp}(t)$ maps $x^{[n]}$ to $t * x^{[n-\varphi(t)]}$, while $\mathscr{Bp}(t')$ maps $x^{[n]}$ to $t' * x^{[n-\varphi(t')]}$, hence to $t' * x^{[n-\varphi(t)]}$, as, by Lemma 4.4, the function φ takes equal values on $\mathscr{ALD}$-equivalent terms. This means that both $\mathscr{Bp}(t)\,;sh_0(g)$ and $\mathscr{Bp}(t')$ map $x^{[n]}$ to $t' * x^{[n-\varphi(t)]}$. Hence, in the monoid $\mathscr{Geom}_{\mathscr{ALD}}$, the two maps $\mathscr{Bp}(t)\,;sh_0(g)$ and $\mathscr{Bp}(t')$ agree on at least one term, namely t, which, by definition, means that (32) holds. □

We have thus obtained in the case of $\mathscr{ALD}$ a relation of the form (24), the endomorphism $\mathscr{Bp}_*$ of $\mathscr{Geom}_{\mathscr{ALD}}$ involved here being the shift endomorphism sh_0. Expressed in the shape of (25), relation (32) reads

$$\mathscr{Bp}(t \cdot eval(w)) \sim \mathscr{Bp}(t)\,;sh_0(eval(w)) \tag{33}$$

for t a term in $\mathscr{T}^{*,\circ}_{x}$ and w a word in $W_{\mathscr{ALD}}$ such that $t \cdot eval(w)$ is defined.

Following the scheme of Section 4.1, we now mimic the construction of $\mathscr{Bp}$ inside the group $\mathtt{Geom}_{\mathscr{ALD}}$. Using $\mathtt{sh}_p$ to denote the endomorphism of $\mathtt{Geom}_{\mathscr{ALD}}$ that maps $\mathtt{D}_p$ to $\mathtt{D}_{1p}$ and $\mathtt{A}_p$ to $\mathtt{A}_{1p}$ for each position p, this amounts to using the following definition.

Definition 4.9. We associate inductively with every term t in $\mathscr{T}^{*,\circ}_{x}$ an element $\mathtt{Bp}(t)$ of $\mathtt{Geom}_{\mathscr{ALD}}$ by

$$\mathtt{Bp}(t) = \begin{cases} 1 & \text{for } t = x \\ \mathtt{Bp}(t_1) \cdot \mathtt{sh}_1(\mathtt{Bp}(t_2)) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\mathtt{Bp}(t_1))^{-1} & \text{for } t = t_1 * t_2 \\ \mathtt{Bp}(t_1) \cdot \mathtt{sh}_1(\mathtt{Bp}(t_2)) \cdot \mathtt{A}_\varepsilon & \text{for } t = t_1 \circ t_2. \end{cases} \tag{34}$$

Example 4.10. Let t be $x * ((x \circ x) * x)$. Starting from $\mathtt{Bp}(x) = 1$, we first find

$$\mathtt{Bp}(x \circ x) = 1 \cdot \mathtt{sh}_1(1) \cdot \mathtt{A}_\varepsilon = \mathtt{A}_\varepsilon,$$

then

$$\mathtt{Bp}((x \circ x) * x) = \mathtt{A}_\varepsilon \cdot \mathtt{sh}_1(1) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\mathtt{A}_\varepsilon)^{-1} = \mathtt{A}_\varepsilon \mathtt{D}_\varepsilon \mathtt{A}_1^{-1},$$

and, finally,

$$\mathtt{Bp}(t) = 1 \cdot \mathtt{sh}_1(\mathtt{A}_\varepsilon \mathtt{D}_\varepsilon \mathtt{A}_1^{-1}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(1)^{-1} = \mathtt{A}_1 \mathtt{D}_1 \mathtt{A}_{11}^{-1} \mathtt{D}_\varepsilon.$$

If our intuition is correct, that is, if the local confluence relations that define the group $\mathtt{Geom}_{\mathscr{ALD}}$ capture enough of the geometry of the $\mathscr{ALD}$-laws, relation (33) should follow from these relations, and, thus, it should induce an equality in the group $\mathtt{Geom}_{\mathscr{ALD}}$, that is, we *should* have the relation

$$\mathtt{Bp}(t \cdot eval(w)) = \mathtt{Bp}(t) \cdot \mathtt{sh}_0(\mathtt{eval}(w)), \tag{35}$$

that is, with our former definition, the mapping $\mathtt{Bp}$ should be an $\mathtt{sh}_0$-blueprint . This is indeed the case, and this is the key point for our current analysis of the ALD-laws. The nice feature is that if the result is true (and it is), its proof must consist of simple verifications, namely checking that some explicit equalities follow from the relations defining $\mathtt{Geom}_{\mathscr{ALD}}$.

Lemma 4.11. The mapping $\mathtt{Bp}$ is an $\mathtt{sh}_0$-blueprint for the $\mathscr{ALD}$-laws.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

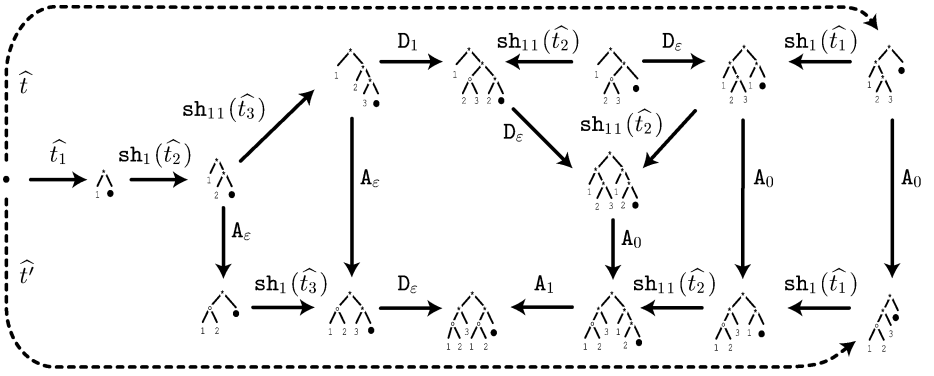


Fig. 6. Proof of Lemma 4.11: deformation of $\widehat{t} \cdot \mathtt{A}_0$ into $\widehat{t'}$ when $t \to_{A,\varepsilon} t'$ holds, that is, for $t = t_1 * (t_2 * t_3)$ and $t' = (t_1 \circ t_2) * t_3$. In the diagram, i stands for t_i and $\bullet$ represents a right comb of the required size.

Proof. In order to get shorter expressions, we will write $\widehat{t}$ for $\mathtt{Bp}(t)$ in the following.

For an induction on the length of the word w, it is enough to prove the result when w consists of a single letter, that is, when it is one of $\mathtt{A}_p^{\pm}$, $\mathtt{D}_p^{\pm}$. Moreover, the cases of $\mathtt{A}_p^{-1}$ and $\mathtt{D}_p^{-1}$ follow immediately from the cases of $\mathtt{A}_p$ and $\mathtt{D}_p$, respectively. So, we need to establish the equalities

$$\mathtt{Bp}(t \cdot \mathtt{A}_p) = \mathtt{Bp}(t) \cdot \mathtt{A}_{0p}$$
$$\mathtt{Bp}(t \cdot \mathtt{D}_p) = \mathtt{Bp}(t) \cdot \mathtt{D}_{0p}$$

whenever the terms involved are defined, that is, using the above notational convention,

$$\begin{aligned} \widehat{t'} &= \widehat{t} \cdot \mathtt{A}_{0p} \quad \text{for } t' = t \cdot \mathtt{A}_p \\ \widehat{t'} &= \widehat{t} \cdot \mathtt{D}_{0p} \quad \text{for } t' = t \cdot \mathtt{D}_p. \end{aligned} \tag{36}$$

We prove (36) using induction on the length of the position p. Let us begin with $p = \varepsilon$ and the case of $\mathtt{A}_p$, that is, of $\mathtt{A}_\varepsilon$. Saying that $t \cdot \mathtt{A}_\varepsilon$ (that is, $t \cdot \mathtt{A}_\varepsilon$) is defined and equal to t' means that t can be decomposed as $t = t_1 * (t_2 * t_3)$, and, then, we have $t' = (t_1 \circ t_2) * t_3$. Using the commutation and quasi-commutation relations of $\mathtt{Geom}_{\mathscr{ALD}}$ (see Figure 6) we find

$$\begin{aligned} \widehat{t'} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{A}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_3}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{A}_1^{-1} \cdot \mathtt{sh}_{11}(\widehat{t_2})^{-1} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \\ &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{A}_\varepsilon \cdot \mathtt{D}_\varepsilon \cdot \mathtt{A}_1^{-1} \cdot \mathtt{sh}_{11}(\widehat{t_2})^{-1} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \\ \widehat{t} \cdot \mathtt{A}_{0p} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{D}_1 \cdot \mathtt{sh}_{11}(\widehat{t_2})^{-1} \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \cdot \mathtt{A}_0 \\ &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{D}_1 \cdot \mathtt{D}_\varepsilon \cdot \mathtt{A}_0 \cdot \mathtt{sh}_{11}(\widehat{t_2})^{-1} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1}, \end{aligned}$$

and the equality follows from (21), which gives $\mathtt{A}_\varepsilon \cdot \mathtt{D}_\varepsilon \cdot \mathtt{A}_1^{-1} = \mathtt{D}_1 \cdot \mathtt{D}_\varepsilon \cdot \mathtt{A}_0$.

We consider now the case of $\mathtt{D}_\varepsilon$. The hypothesis that $t \cdot \mathtt{D}_\varepsilon$ is defined and equal to t' means that t can be decomposed as $t_1 * (t_2 \,\square\, t_3)$, and, then, t' is $(t_1 * t_2) \,\square\, (t_1 * t_3)$. Assume first that $\square = *$ (this is the most complicated case). Using the commutation and quasi-commutation

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

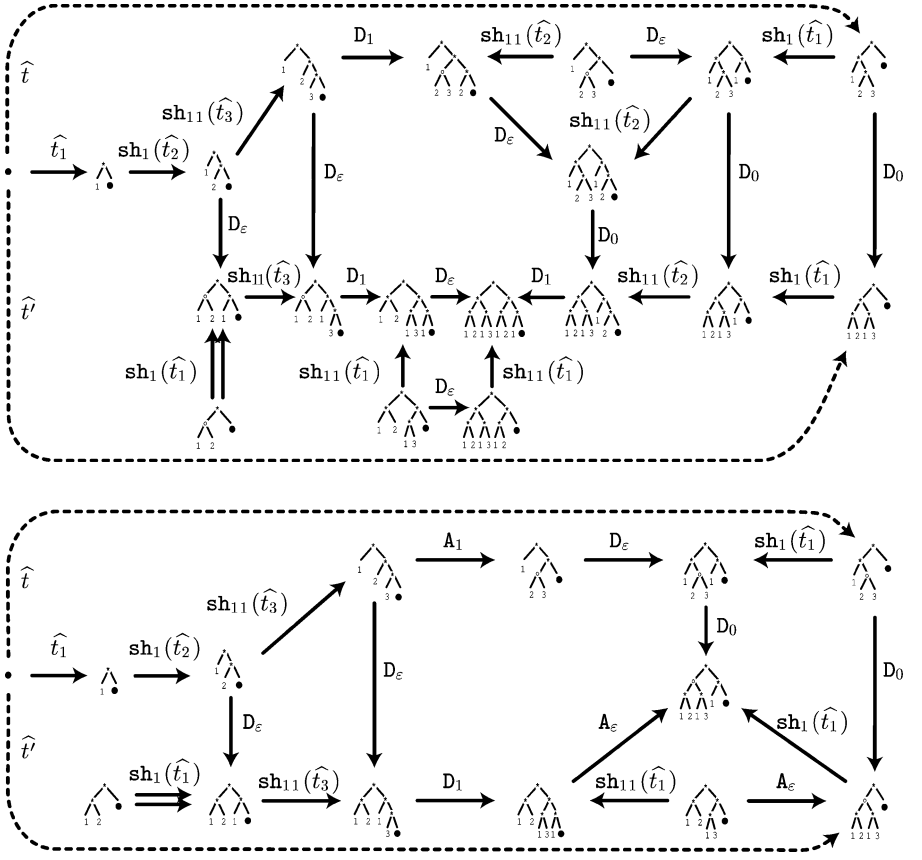


Fig. 7. Proof of Lemma 4.11: deformation of $\widehat{t} \cdot \mathtt{D}_0$ into $\widehat{t'}$ when $t \to_{D,\varepsilon} t'$ holds, that is, for $t = t_1 * (t_2 \square t_3)$ and $t' = (t_1 * t_2) \square (t_1 * t_3)$. There are two cases, depending on whether $\square$ is $*$ (top) or $\circ$ (bottom).

relations of $\mathtt{Geom}_{\mathscr{ALD}}$ (see Figure 7 (top)) we find

$$\begin{aligned}
\widehat{t'} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \cdot \mathtt{sh}_1(\widehat{t_1}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{D}_1 \cdot \mathtt{sh}_{11}(\widehat{t_1})^{-1} \\
&\qquad \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_{11}(\widehat{t_1}) \cdot \mathtt{D}_1^{-1} \cdot \mathtt{sh}_{11}(\widehat{t_2})^{-1} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \\
&= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{D}_1 \cdot \mathtt{D}_\varepsilon \cdot \mathtt{D}_1^{-1} \cdot \mathtt{sh}_{11}(\widehat{t_2})^{-1} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \\
\widehat{t} \cdot \mathtt{D}_{0p} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{D}_1 \cdot \mathtt{sh}_{11}(\widehat{t_2})^{-1} \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \cdot \mathtt{D}_0 \\
&= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{D}_1 \cdot \mathtt{D}_\varepsilon \cdot \mathtt{D}_0 \cdot \mathtt{sh}_{11}(\widehat{t_2})^{-1} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1},
\end{aligned}$$

and the equality follows from the (19) relation $\mathtt{D}_1 \cdot \mathtt{D}_\varepsilon \cdot \mathtt{D}_1 \cdot \mathtt{D}_0 = \mathtt{D}_\varepsilon \cdot \mathtt{D}_1 \cdot \mathtt{D}_\varepsilon$. Assume now that $\square = \circ$. We get (see Figure 7 (bottom))

$$\begin{aligned}
\widehat{t'} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \cdot \mathtt{sh}_1(\widehat{t_1}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{D}_1 \cdot \mathtt{sh}_{11}(\widehat{t_1})^{-1} \cdot \mathtt{A}_\varepsilon \\
&= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{D}_1 \cdot \mathtt{A}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \\
\widehat{t} \cdot \mathtt{D}_{0p} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{A}_1 \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \cdot \mathtt{D}_0 \\
&= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{sh}_{11}(\widehat{t_3}) \cdot \mathtt{A}_1 \cdot \mathtt{D}_\varepsilon \cdot \mathtt{D}_0 \cdot \mathtt{sh}_1(\widehat{t_1})^{-1},
\end{aligned}$$

and the equality follows from the (20) relation $\mathtt{A}_1 \cdot \mathtt{D}_\varepsilon \cdot \mathtt{D}_0 = \mathtt{D}_\varepsilon \cdot \mathtt{D}_1 \cdot \mathtt{A}_\varepsilon$.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

The case $p = \varepsilon$ is now complete. From now on, we shall treat the cases of $\mathtt{A}_p$ and $\mathtt{D}_p$ simultaneously, using $\mathtt{E}_p$ as a common notation. Assume first that the position p is $0q$ for some q. The hypothesis that $t \cdot \mathtt{E}_p$ is defined implies that t can be decomposed as $t_1 \,\square\, t_2$, and, then, t' is $t'_1 \,\square\, t_2$, with $t'_1 = t_1 \cdot \mathtt{E}_q$. Assume first that $\square = *$. The induction hypothesis implies $\widehat{t'_1} = \widehat{t_1} \cdot \mathtt{E}_{0q}$. We find

$$\begin{aligned}\widehat{t'} &= \widehat{t_1} \cdot \mathtt{E}_{0q} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{E}_{10q}^{-1} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \\ &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{E}_{0q} \cdot \mathtt{D}_\varepsilon \cdot \mathtt{E}_{10q}^{-1} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \\ \widehat{t} \cdot \mathtt{E}_{0p} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \cdot \mathtt{E}_{00q} \\ &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{E}_{00q} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1},\end{aligned}$$

and the equality follows from the (12) relation $\mathtt{E}_{0q} \cdot \mathtt{D}_\varepsilon = \mathtt{D}_\varepsilon \cdot \mathtt{E}_{00q} \cdot \mathtt{E}_{10q}$. Similarly, for $\square = \circ$, we find

$$\begin{aligned}\widehat{t'} &= \widehat{t_1} \cdot \mathtt{E}_{0q} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{A}_\varepsilon = \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{E}_{0q} \cdot \mathtt{A}_\varepsilon \\ \widehat{t} \cdot \mathtt{E}_{0p} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{A}_\varepsilon \cdot \mathtt{E}_{00q},\end{aligned}$$

and the equality follows from the (16) relation $\mathtt{E}_{0q} \cdot \mathtt{A}_\varepsilon = \mathtt{A}_\varepsilon \cdot \mathtt{E}_{00q}$.

The argument is similar when the position p is $1q$ for some q. The hypothesis that $t \cdot \mathtt{E}_p$ is defined implies that t can de decomposed as $t_1 \,\square\, t_2$, and, then, t' is $t_1 \,\square\, t'_2$, with $t'_2 = t_2 \cdot \mathtt{E}_q$. Assume first that $\square = *$. The induction hypothesis implies $\widehat{t'_1} = \widehat{t_1} \cdot \mathtt{E}_{0q}$. We find now that

$$\begin{aligned}\widehat{t} \cdot \mathtt{E}_{0p} = \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \cdot \mathtt{E}_{01q} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{E}_{01q} \cdot \mathtt{sh}_1(\widehat{t_1})^{-1} \\ \widehat{t'} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{E}_{10q} \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(\widehat{t_1})^{-1},\end{aligned}$$

and the equality follows from the (13) relation $\mathtt{E}_{10q} \cdot \mathtt{D}_\varepsilon = \mathtt{D}_\varepsilon \cdot \mathtt{E}_{01q}$. Finally, for $\square = \circ$, we have

$$\begin{aligned}\widehat{t} \cdot \mathtt{E}_{0p} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{A}_\varepsilon \cdot \mathtt{E}_{01q} \\ \widehat{t'} &= \widehat{t_1} \cdot \mathtt{sh}_1(\widehat{t_2}) \cdot \mathtt{E}_{01q} \cdot \mathtt{A}_\varepsilon,\end{aligned}$$

and the equality follows from the (17) relation $\mathtt{E}_{10q} \cdot \mathtt{A}_\varepsilon = \mathtt{A}_\varepsilon \cdot \mathtt{E}_{01q}$. The induction is now complete. □

We have thus completed the construction of a blueprint for the $\mathscr{ALD}$-laws. It may be observed that this construction induces the construction of a similar $\mathtt{sh}_0$-blueprint for the $\mathscr{LD}$-law considered alone: indeed, in the case of terms not containing the map $\circ$, the blueprint does not involve any generator $\mathtt{A}_p$, and it can be checked that the only relations needed to check the blueprint condition are present in the group $\mathtt{Geom}_{\mathscr{LD}}$.

Remark 4.12. Here the blueprints have been defined for terms in one variable only. Developing a similar approach for terms involving several variables is possible, at the expense of extending the geometry monoid $\mathscr{Geom}_\mathscr{E}$ by introducing additional maps whose action is to shift the indices of the variables so as to still generate all terms starting from right combs. See Dehornoy (2000) for details in the specific case of the $\mathscr{LD}$-law.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

5. Using $\mathtt{Geom}_{\mathscr{E}}$ to study $\mathscr{E}$

We claim that the (two-sided) geometry group $\mathtt{Geom}_{\mathscr{E}}$ is an interesting object that provides useful information about the equational specification $\mathscr{E}$. The way this vague statement can be made precise depends on the equational specification under consideration. If the latter involves simple equations, typically the associativity and/or commutativity laws, constructing an $\mathscr{E}$-algebra or solving the word problem for $\mathscr{E}$ is not a challenge, and, in particular, appealing to $\mathtt{Geom}_{\mathscr{E}}$ is not necessary. However, the group $\mathtt{Geom}_{\mathscr{E}}$ may be of interest in itself, as the enormous literature devoted to Thompson's groups F and V shows (Thompson 1980; McKenzie and Thompson 1973; Cannon *et al.* 1996). In the case of complicated laws, typically the self-distributivity law or its variants, constructing $\mathscr{E}$-algebras and solving the word problem of $\mathscr{E}$ may be difficult (and often even open) questions, and then the geometry group may be useful. The basic scheme consists of exploiting the blueprint construction, when it exists, to obtain the structure of an $\mathscr{E}$-algebra on some quotient of the group $\mathtt{Geom}_{\mathscr{E}}$.

5.1. *Construction of $\mathscr{E}$-algebras: the principle*

The first, and more direct, application of the previous approach is the construction of a model for an equational specification. The general principle is as follows.

Proposition 5.1. Assume that $\mathscr{F}$ is a signature consisting of binary operations, that $\mathscr{E}$ is a regular equational specification on $\mathscr{F}$, that $\mathtt{Bp}$ is a $\mathtt{Bp}_*$-blueprint for $\mathscr{E}$ and that H is a subgroup of $\mathtt{Geom}_{\mathscr{E}}$ that includes the image of $\mathtt{Bp}_*$. For t a term in $\mathscr{T}_x$, let $\overline{\mathtt{Bp}}(t)$ be the left coset of $\mathtt{Bp}(t)$ modulo H. Then the map $\overline{\mathtt{Bp}}$ is constant on each $=_{\mathscr{E}}$-class, that is, terms that are equal in $\mathscr{E}$ have the same image.

Proof. Assume that t and t' are terms in $\mathscr{T}_x^{\mathscr{F}}$ that are equal in $\mathscr{E}$. By Proposition 2.7, there must exist a map in $\mathscr{G}eom_{\mathscr{E}}$ that maps t to t', and, therefore, there must exist a word w in $W_{\mathscr{E}}$ such that $t \cdot eval(w)$ is defined and equal to t'. Then, by (27), we have $\mathtt{Bp}(t') = \mathtt{Bp}(t) \cdot \mathtt{Bp}_*(\mathtt{eval}(w))$ in $\mathtt{Geom}_{\mathscr{E}}$, so the hypothesis gives

$$(\mathtt{Bp}(t))^{-1} \cdot (\mathtt{Bp}(t')) \in H,$$

that is, $\mathtt{Bp}(t)$ and $\mathtt{Bp}(t')$ lie in the same H-coset. □

Corollary 5.2. Under the same hypotheses, let M be the image of the mapping $\overline{\mathtt{Bp}}$. For each function symbol $\square$ in $\mathscr{F}$, define an operation in M by

$$\mathtt{Bp}(t)H \square \mathtt{Bp}(t')H := \mathtt{Bp}(t \square t')H.$$

Then M equipped with these operations is an $\mathscr{E}$-algebra, that is, an $\mathscr{F}$-structure satisfying all equations of $\mathscr{E}$.

Note that, in general, nothing guarantees that the system obtained in this way is non-trivial: it might happen that the quotient-structure $H\backslash\mathtt{Geom}_{\mathscr{E}}$ reduces to a one-element algebra. This does not happen in the cases we shall consider: on the contrary, the algebraic systems obtained will turn out to be free, that is, as far as possible from trivial.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

5.2. *Construction of algebraic systems: the example of $\mathscr{ALD}$*

We will now return to our leading example, namely the $\mathscr{ALD}$-laws. In Section 4 we constructed an $\mathtt{sh}_0$-blueprint $\mathtt{Bp}$ for $\mathscr{ALD}$, so Proposition 5.1 and Corollary 5.2 apply directly for each subgroup H of $\mathtt{Geom}_{\mathscr{ALD}}$ including the image of $\mathtt{sh}_0$, typically $H := \mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})$. In this way, we obtain an $\mathscr{ALD}$-algebra whose domain is some subset of the coset set $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})\backslash\mathtt{Geom}_{\mathscr{ALD}}$.

Actually, we can adapt the results to make them simpler and more useful. Indeed, instead of restricting to the image of the mapping $\overline{\mathtt{Bp}}$, we can extend the construction to the whole group $\mathtt{Geom}_{\mathscr{ALD}}$. To do this, the idea is obvious: we look at the inductive definition of the blueprint, and define $*$ and $\circ$ to be the operations used to construct $\mathtt{Bp}(t * t')$ and $\mathtt{Bp}(t \circ t')$ from $\mathtt{Bp}(t)$ and $\mathtt{Bp}(t')$, that is, we choose the operations on $\mathtt{Geom}_{\mathscr{ALD}}$ that make $\mathtt{Bp}$ a homomorphism from the free algebra $\mathscr{T}_X^{*,\circ}$ to $\mathtt{Geom}_{\mathscr{ALD}}$.

Lemma 5.3. On the group $\mathtt{Geom}_{\mathscr{ALD}}$, define two new binary operations $*, \circ$ by

$$\begin{aligned} x * y &:= x \cdot \mathtt{sh}_1(y) \cdot \mathtt{D}_\varepsilon \cdot \mathtt{sh}_1(x)^{-1} \\ x \circ y &:= x \cdot \mathtt{sh}_1(y) \cdot \mathtt{A}_\varepsilon. \end{aligned} \tag{37}$$

Then, for all x, y, z in $\mathtt{Geom}_{\mathscr{ALD}}$ and $\square$ in $\{*, \circ\}$, we have

$$(x * y) \square (x * z) = (x * (y \square z)) \cdot \mathtt{D}_0 \tag{38}$$

$$(x \circ y) * z = (x * (y * z)) \cdot \mathtt{A}_0 \tag{39}$$

$$(x \cdot \mathtt{sh}_0(z)) \square y = x \square y \cdot \mathtt{sh}_{00}(z) \tag{40}$$

$$x \square (y \cdot \mathtt{sh}_0(z)) = x \square y \cdot \mathtt{sh}_{01}(z). \tag{41}$$

Proof. The verifications are the same as those already given in the proof of Lemma 4.11. The only difference is that in Section 4 we only considered elements of $\mathtt{Geom}_{\mathscr{ALD}}$ that are blueprints of terms, while, now, we consider arbitrary elements of $\mathtt{Geom}_{\mathscr{ALD}}$. Now, inspecting the proof of Section 4 shows that the specific form of the elements was never used, and, therefore, the whole computation remains valid; in particular, the diagrams of Figures 6 and 7 remain commutative when the elements $\widehat{t_1}, \widehat{t_2}, \widehat{t_3}$ are replaced by arbitrary elements of the group $\mathtt{Geom}_{\mathscr{ALD}}$ – but, in that case, the nodes of the diagram no longer correspond to terms. □

Relations (38) and (39) control the obstruction for $(\mathtt{Geom}_{\mathscr{ALD}}, *, \circ)$ to be an $\mathscr{ALD}$-algebra, and show that the latter belongs to the subgroup $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})$. Relations (40) and (41) show that the operations on $\mathtt{Geom}_{\mathscr{ALD}}$ induce well-defined operations on the coset set $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})\backslash\mathtt{Geom}_{\mathscr{ALD}}$. Thus we may state the following proposition.

Proposition 5.4. Let M be the coset set $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})\backslash\mathtt{Geom}_{\mathscr{ALD}}$. Then M equipped with the operations induced by those of Lemma 5.3 is an $\mathscr{ALD}$-algebra.

The subgroup $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})$ is not normal, and, therefore, the associated coset set is not a group. When we replace $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})$ with a larger subgroup H of $\mathtt{Geom}_{\mathscr{ALD}}$, typically the normal subgroup generated by $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})$, we can still apply Proposition 5.1, but it is not *a priori* certain that the operations $*$ and $\circ$ induce well-defined operations on the whole of $H\backslash\mathtt{Geom}_{\mathscr{ALD}}$. However, this does happen in good cases, as here with $\mathscr{ALD}$.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

We are thus led to investigate the normal subngroups of the group $\mathtt{Geom}_{\mathscr{ALD}}$ that include the image of the shift endomorphism $\mathtt{sh}_0$. This is easy.

Lemma 5.5. Every normal subgroup of $\mathtt{Geom}_{\mathscr{ALD}}$ that includes the image of $\mathtt{sh}_0$ contains all generators $\mathtt{D}_p$ and $\mathtt{A}_p$ such that p contains at least one 0.

Proof. For $\mathtt{E} = \mathtt{D}$ or $\mathtt{A}$, the commutation relation (17) gives $\mathtt{E}_{10} = \mathtt{A}_\varepsilon \cdot \mathtt{E}_{01} \cdot \mathtt{A}_\varepsilon^{-1}$. Hence, inductively,

$$\mathtt{E}_{1^i 0 p} = \mathtt{A}_{1^{i-1}} \cdot \ldots \cdot \mathtt{A}_1 \cdot \mathtt{A}_\varepsilon \cdot \mathtt{E}_{01^i p} \cdot \mathtt{A}_\varepsilon^{-1} \cdot \mathtt{A}_1^{-1} \cdot \ldots \cdot \mathtt{A}_{1^{i-1}}^{-1},$$

which shows that any normal subgroup containing all $\mathtt{E}_{0q}$ must contain all $\mathtt{E}_p$ such that p contains at least one 0. □

Thus, collapsing all generators $\mathtt{D}_p$ and $\mathtt{A}_p$ such that p begins with 0 in $\mathtt{Geom}_{\mathscr{ALD}}$ requires the collapse of all generators $\mathtt{E}_p$ such that p contains at least one 0, in which case the quotient-group is generated by the images of the remaining generators, namely the generators $\mathtt{D}_{1^{i-1}}$ and $\mathtt{A}_{1^{i-1}}$ with $i \geqslant 1$. Considering what remains from the defining relations of $\mathtt{Geom}_{\mathscr{ALD}}$ leads directly to the following quotient of $\mathtt{Geom}_{\mathscr{ALD}}$. We use σ_i and a_i as simplified notation for $\mathtt{D}_{1^{i-1}}$ and $\mathtt{A}_{1^{i-1}}$.

Definition 5.6. We let $B_\bullet$ be the group generated by two infinite sequences $\sigma_1, \sigma_2, \ldots$ and $a_1, a_2, \ldots$ of generators subject to the relations

$$\begin{cases} \sigma_i\sigma_j = \sigma_j\sigma_i, \quad \sigma_i a_j = a_j\sigma_i, \quad a_i a_{j-1} = a_j a_i, \quad a_i\sigma_{j-1} = \sigma_j a_i, \\ \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}, \quad \sigma_{i+1}\sigma_i a_{i+1} = a_i\sigma_i, \quad \sigma_i\sigma_{i+1}a_i = a_{i+1}\sigma_i \end{cases} \tag{42}$$

for $i \geqslant 1$ and $j \geqslant i+2$.

We do not claim that $B_\bullet$ is exactly the quotient of $\mathtt{Geom}_{\mathscr{ALD}}$ by the normal subgroup generated by $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})$: collapsing $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})$ might result in additional relations. By construction, mapping $\mathtt{D}_{1^{i-1}}$ to σ_i and $\mathtt{A}_{1^{i-1}}$ to a_i defines a surjective homomorphism π of $\mathtt{Geom}_{\mathscr{ALD}}$ onto $B_\bullet$ whose kernel includes $\mathtt{sh}_0(\mathtt{Geom}_{\mathscr{ALD}})$, and therefore the normal subgroup N it generates. However, it might be that π collapses more than N. Proving that this does not happen would require a more complete study of the group $\mathtt{Geom}_{\mathscr{ALD}}$, which is not our current aim. Here, the only thing we wish to observe is that, by mimicking once again the construction of the blueprint, we can construct an $\mathscr{ALD}$-structure on the group $B_\bullet$.

Proposition 5.7 (Dehornoy 2006; 2007). Let $\mathtt{sh}$ denote the endomorphism of the group $B_\bullet$ that maps σ_i to σ_{i+1} and a_i to a_{i+1} for every positive integer i. Define binary operations $*, \circ$ on $B_\bullet$ by

$$\begin{aligned} x * y &:= x \cdot \mathtt{sh}(y) \cdot \sigma_1 \cdot \mathtt{sh}(x)^{-1} \\ x \circ y &:= x \cdot \mathtt{sh}(y) \cdot a_1. \end{aligned} \tag{43}$$

Then $(B_\bullet, *, \circ)$ is an $\mathscr{ALD}$-algebra. Moreover, this $\mathscr{ALD}$-algebra is torsion-free, that is, every element of $B_\bullet$ generates a free $\mathscr{ALD}$-subsystem.

The group $B_\bullet$ was introduced by M. Brin in Brin (2004; 2006) and, independently, by the author in Dehornoy (2005), and its elements have been interpreted in Dehornoy (2006)

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

as parenthesised braids, which are a refinement of standard Artin braids in which one takes into account the distances between the strands. The above results show that, in the context of the $\mathscr{ALD}$-laws, the connection between $\mathtt{Geom}_{\mathscr{ALD}}$ and $B_\bullet$ is similar to the connection between the geometry group of self-distributivity and the braid group B_∞, as investigated in Dehornoy (2000). The benefit of the geometry monoid approach is to make the operations (43) natural, at least once the blueprint based on Lemma 4.7 has been chosen, and to explain why they had to appear in this form and in this group.

5.3. *Presentation of $\mathscr{G}eom_{\mathscr{E}}$*

A different use for a blueprint is to allow a closer comparison between the monoid $\mathscr{G}eom_{\mathscr{E}}$ and the group $\mathtt{Geom}_{\mathscr{E}}$ in the non-linear case. We observed that $\mathtt{Geom}_{\mathscr{E}}$ is constructed by means of a list of confluence relations holding in $\mathscr{G}eom_{\mathscr{E}}$, but, in general, there is no reason why this list should be complete, that is, provide a presentation of $\mathscr{G}eom_{\mathscr{E}}$. In fact, when the empty map belongs to $\mathscr{G}eom_{\mathscr{E}}$, it cannot be the case that the confluence relations exhaust all relations of $\mathscr{G}eom_{\mathscr{E}}$ as no such relation involves $\varnothing$. However, the real problem is to control the relations of $\mathscr{G}eom_{\mathscr{E}}$ that do not involve the empty map, which amounts to describing the relation $\sim$ of Definition 3.6 on $\mathscr{G}eom_{\mathscr{E}}$. Then, we have the following solution.

Proposition 5.8. Assume that $\mathtt{Bp}_*$ is injective and $\mathtt{Bp}$ is a $\mathtt{Bp}_*$-blueprint for the equations $\mathscr{E}$. Then the confluence relations used in the definition of $\mathtt{Geom}_{\mathscr{E}}$ generate all non-trivial relations in $\mathscr{G}eom_{\mathscr{E}}$ in the following sense: for all words w, w' in $W_{\mathscr{E}}$, the relation $eval(w) \sim eval(w')$ holds only if w and w' represent the same element of the group $\mathtt{Geom}_{\mathscr{E}}$.

Proof. Assume that w, w' are words in $W_{\mathscr{E}}$ such that the associated maps $eval(w)$ and $eval(w')$ in $\mathscr{G}eom_{\mathscr{E}}$ are connected by $\sim$, that is, there exists at least one term t on which they agree. Let t' be the common image of t under these maps. The hypothesis that $\mathtt{Bp}$ is a $\mathtt{Bp}_*$-blueprint gives

$$\mathtt{Bp}(t) \cdot \mathtt{Bp}_*(\mathtt{eval}(w)) = \mathtt{Bp}(t') = \mathtt{Bp}(t) \cdot \mathtt{Bp}_*(\mathtt{eval}(w'))$$

in $\mathtt{Geom}_{\mathscr{E}}$, and thus

$$\mathtt{Bp}_*(\mathtt{eval}(w)) = \mathtt{Bp}_*(\mathtt{eval}(w')).$$

If $\mathtt{Bp}_*$ is injective, we deduce $\mathtt{eval}(w) = \mathtt{eval}(w')$, that is, the words w and w' represent the same element of the group $\mathtt{Geom}_{\mathscr{E}}$. □

Example 5.9. We conjecture that the endomorphism $\mathtt{sh}_0$ of the group $\mathtt{Geom}_{\mathscr{ALD}}$ is injective. This would imply that the confluence relations listed in Section 3.3 generate, in the sense described above, all relations holding in the geometry monoid $\mathscr{G}eom_{\mathscr{ALD}}$.

When we restrict from $\mathscr{ALD}$ to $\mathscr{LD}$ alone, the argument is similar, and, in that case, the injectivity of $\mathtt{sh}_0$ on $\mathtt{Geom}_{\mathscr{LD}}$ is known (Dehornoy 2000). What makes the case of $\mathscr{ALD}$ more difficult is that in the latter case some confluence relations are missing, and the group $\mathtt{Geom}_{\mathscr{ALD}}$ is not a group of fractions for the positive monoid $\mathtt{Geom}^{\rightarrow}_{\mathscr{ALD}}$, while $\mathtt{Geom}_{\mathscr{LD}}$ is a group of fractions for $\mathtt{Geom}^{\rightarrow}_{\mathscr{LD}}$– see Section 6 for a discussion of what happens when another orientation of the $\mathscr{ALD}$-laws is chosen. Both in the case of $\mathscr{LD}$

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

and $\mathscr{ALD}$, proving the injectivity of $\mathtt{sh}_0$ on the positive monoid is easy, but extending the result to the group is not obvious as, in this case, the latter is not a group of fractions.

5.4. *Solving the word problem*

Still another possibility is to use the blueprint for solving the word problem of $\mathscr{E}$, that is, to construct an algorithm that decides whether two terms t, t' are equivalent modulo the equations of $\mathscr{E}$.

Proposition 5.10. Assume that $\mathtt{Bp}$ is a Turing computable $\mathtt{Bp}_*$-blueprint for the equations $\mathscr{E}$, that P, Q are disjoint recursively enumerable subsets of $\mathtt{Geom}_{\mathscr{E}}$ such that P includes the image of $\mathtt{Bp}_*$ and that there is a binary relation $\diamond$ on $\mathscr{T}_x$ such that for all terms t and t', at least one of $t =_{\mathscr{E}} t'$ or $t \diamond t'$ holds and $t \diamond t'$ implies $\mathtt{Bp}(t)^{-1}\mathtt{Bp}(t') \in Q$. Then the word problem for the equations $\mathscr{E}$ is solvable.

Proof. Let $\widehat{P}, \widehat{Q}$ be disjoint recursive sets that include P and Q, respectively. First assume that $t =_{\mathscr{E}} t'$ holds. Under the hypotheses, this implies

$$\mathtt{Bp}(t)^{-1}\mathtt{Bp}(t') \in \mathrm{Im}(\mathtt{Bp}_*) \subseteq P \subseteq \widehat{P}.$$

Conversely, assume that $t =_{\mathscr{E}} t'$ fails. Then $t \diamond t'$ necessarily holds, which implies

$$\mathtt{Bp}(t)^{-1}\mathtt{Bp}(t') \in Q \subseteq \widehat{Q},$$

so $t =_{\mathscr{E}} t'$ is equivalent to $\mathtt{Bp}(t)^{-1}\mathtt{Bp}(t') \in \widehat{P}$. As the function $\mathtt{Bp}$ is assumed to be Turing computable and the set P is assumed to be Turing decidable, the latter condition is Turing decidable. □

Example 5.11. Let us consider the case of the $\mathscr{LD}$-law. We know that there exists an $\mathtt{sh}_0$-blueprint $\mathtt{Bp}$ for $\mathscr{LD}$. Let $t \diamond t'$ be the symmetric closure of the relation 't is $\mathscr{LD}$-equivalent to some iterated left subterm of some term $\mathscr{LD}$-equivalent to t', where an iterated left subterm is defined to be any subterm corresponding to a position of the form 0^k. Let P be the subset of $\mathtt{Geom}_{\mathscr{LD}}$ consisting of those elements that can be expressed without using either $\mathtt{D}_\varepsilon$ or $\mathtt{D}_\varepsilon^{-1}$, and let Q be the subset of $\mathtt{Geom}_{\mathscr{LD}}$ consisting of those elements that can be expressed using exactly one of $\mathtt{D}_\varepsilon$ or $\mathtt{D}_\varepsilon^{-1}$. It is easy to show that every element of $\mathtt{Geom}_{\mathscr{LD}}$ in the image of $\mathtt{sh}_0$ belongs to P, while $t \diamond t'$ implies $\mathtt{Bp}(t)^{-1}\mathtt{Bp}(t') \in Q$. Moreover, one can show that P and Q are disjoint and recursive, so Proposition 5.10 implies that the word problem for $\mathscr{LD}$ is solvable. This was the first solution for a long standing open question – other solutions are known now (Dehornoy 2000).

A similar scheme was used in Dehornoy (2002) to solve the word problem for the central duplication law $x(yz) = (xy)(yz)$.

As for the word problem for the $\mathscr{ALD}$-laws, for which a direct solution involving the group $B_\bullet$ of Definition 5.6 is known (Dehornoy 2007), the scheme may also work, but, as in Example 5.9, some pieces are missing as the group $\mathtt{Geom}_{\mathscr{ALD}}$ fails to be a group of fractions for the positive monoid $\mathtt{Geom}^{\rightarrow}_{\mathscr{ALD}}$, making the verification of certain technical conditions problematic, typically, the fact that the expected sets P and Q are disjoint. We shall come back to this question in the following section.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

6. Reversing orientation

In this short section we discuss the effect of changing the orientation of equations in the current approach. Although in theory nothing much changes, we shall see when considering our leading example, namely the case of the $\mathscr{ALD}$-laws, that changing the orientation may lead to different developments, which are more or less convenient according to the specific problem we are interested in.

6.1. *The general case*

Introducing the two-sided geometry monoid $\mathscr{Geom}_{\mathscr{E}}$ requires us to fix an orientation for each equation of the equational specification $\mathscr{E}$ under consideration. At a theoretical level, changing the orientation of an equation E does not change much, as it just amounts to replacing each map $\rightarrow_{E,p}$ with its inverse $\leftarrow_{E,p}$, resulting in an isomorphic inverse monoid.

However, in practice things change when we investigate local confluence relations in order to, if we can, introduce the group $\mathtt{Geom}_{\mathscr{E}}$. Indeed, the heuristic principle applied in Section 3 consists of looking for *positive* maps (no map $\leftarrow_{E,p}$), and, therefore, exchanging $\rightarrow_{E,p}$ and $\leftarrow_{E,p}$ may result in different relations, and thus to different developments.

6.2. *The case of $\mathscr{ALD}$*

Concentrating on the case of $\mathscr{ALD}$ will enable us to make the discussion more precise. If we stick to the principle of grouping the equations D and D', we can think of reversing the maps $\rightarrow_{D,p}$, and/or reversing the maps $\rightarrow_{A,p}$.

As for the maps $\rightarrow_{D,p}$, choosing $\leftarrow_{D,p}$, that is, orienting the distributivity law in the contracting direction rather than in the expanding direction may appear natural, as, in particular, it leads to a rewrite system that is terminating, which is not the case for the system obtained in Section 3. Although this option may seem to be natural, it turns out definitely not to be convenient. Indeed, the associated one-sided rewrite systems are not globally confluent: if t, t' are $\mathscr{ALD}$-equivalent terms, there is in general no way to contract them into a common third term, and, as in the analogous case of the self-distributivity law $\mathscr{LD}$ alone, we know of no application based on privileging the contracting direction, that is, on using the maps $\leftarrow_{D,p}$. Critical pair completion of the rules associated with $\mathscr{LD}$ in the contracting direction yields an infinite system.

Question 6.1. Can this infinite rewrite system be described finitely?

The same question can be raised for the central duplication law of Example 5.11; in that case, the completion seems to generate rules that are monotonically increasing in length, which, if true, would give another proof of the decidability of the word problem.

The situation is different for the maps $\rightarrow_{A,p}$: orienting the equation A from $(x \circ y) * z$ to $x * (y * z)$ rather than from $x * (y * z)$ to $(x \circ y) * z$ leads to new and interesting results, which we shall describe now.

Notation 6.2. For each position p, we write $\rightarrow_{\bar{A},p}$ for $\leftarrow_{A,p}$.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

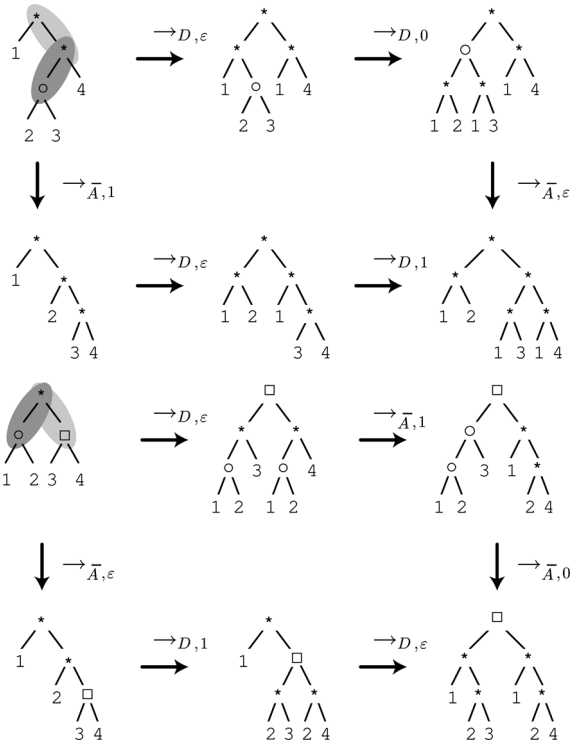


Fig. 8. Local confluence for the maps $\rightarrow_{D,p}$ and $\rightarrow_{\bar{A},q}$: the overlapping case. Recall that $\Box$ stands for both $*$ or $\circ$; the numbers stand for the indices of the variables.

According to the scheme of Section 3, we investigate the local confluence relations that connect the various maps $\rightarrow_{D,p}$ and $\rightarrow_{\bar{A},q}$. The analyses of the parallel and nested cases are analogous to those given for $\rightarrow_{D,p}$ and $\rightarrow_{A,q}$, and we will skip them in order to concentrate on the overlapping case.

As shown in Figure 8, in addition to the relation (19)

$$\rightarrow_{D,p} \;;\; \rightarrow_{D,p1} \;;\; \rightarrow_{D,p} \;=\; \rightarrow_{D,p1} \;;\; \rightarrow_{D,p} \;;\; \rightarrow_{D,p1} \;;\; \rightarrow_{D,p0},$$

which only involves maps $\rightarrow_{D,p}$ and remains valid, the following two mixed relations are satisfied:

$$\rightarrow_{D,p} \;;\; \rightarrow_{D,p0} \;;\; \rightarrow_{\bar{A},p} \;=\; \rightarrow_{\bar{A},p1} \;;\; \rightarrow_{D,p} \;;\; \rightarrow_{D,p1} \tag{44}$$

$$\rightarrow_{D,p} \;;\; \rightarrow_{\bar{A},p1} \;;\; \rightarrow_{\bar{A},p0} \;=\; \rightarrow_{\bar{A},p} \;;\; \rightarrow_{D,p1} \;;\; \rightarrow_{D,p}\,. \tag{45}$$

The only cases where there is no local confluence relation are those of $\rightarrow_{\bar{A},p}$ and $\rightarrow_{\bar{A},p0}$: in these cases no relation may exist because no term may simultaneously belong to the domain of $\rightarrow_{\bar{A},p}$, which requires that the operation symbol at position $p0$ is $\circ$, and to the domain of $\rightarrow_{\bar{A},p0}$, which requires that the operation symbol at $p0$ is $*$.

The relations (44) and (45) are equivalent to (20) and (21), respectively, and, therefore, applying the approach of Section 3 leads us to introduce the same group $\mathtt{Geom}_{\mathscr{ALD}}$.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

However, things are different when we think in terms of oriented rewrite systems: for instance, we shall now describe how to use the approach to solve the word problem of $\mathscr{ALD}$.

Proposition 6.3.

(i) Let $\mathscr{R}$ be the rewrite system associated with the maps $\rightarrow_{\bar{A},p}$ and $\rightarrow_{D',p}$. Then $\mathscr{R}$ is complete, and $\mathscr{R}$-normal forms are of the form $t\sigma$ with t a $\circ$-term and σ a substitution with values in $*$-terms.

(ii) Two $\mathscr{R}$-normal forms $t\sigma, t'\sigma'$ are $\mathscr{ALD}$-equivalent if and only if we have $t = t'$ and $x\sigma =_{\mathscr{LD}} x\sigma'$ for each variable x occurring in t.

Proof.

(i) First, $\mathscr{R}$ is locally confluent. For the parallel and nested cases, this follows from the general analysis. For overlaps, we observed above that the case of $\rightarrow_{\bar{A},p}$ *versus* $\rightarrow_{\bar{A},p0}$ cannot occur due to a symbol conflict; a similar argument excludes the case of $\rightarrow_{D',p}$ *versus* $\rightarrow_{D',p1}$. Finally, the case of $\rightarrow_{\bar{A},p}$ *versus* $\rightarrow_{D',p}$ is covered by relation (45).
Next, $\mathscr{R}$ is terminating as the number of symbols $\circ$ strictly decreases when $\rightarrow_{\bar{A},p}$ maps are applied, and it is preserved, but the sum of the lengths of the positions where $\circ$ occurs strictly decreases when $\rightarrow_{D',p}$ is applied.
Finally, a term that is not of the form $t\sigma$ with t an $\circ$-term and σ a substitution with values in $*$-terms must contain a symbol $\circ$ below a symbol $*$. Let p be the position of the $*$ under consideration. If $\circ$ occurs at $p0$, the term lies in the domain of $\rightarrow_{\bar{A},p}$, and if it occurs at $p1$, the term lies in the domain of $\rightarrow_{D',p}$. In both cases, it cannot be an $\mathscr{R}$-normal form.

(ii) A direct inspection shows that the rules of $\mathscr{R}$ commute with the remaining rules of $\mathscr{G}eom_{\mathscr{ALD}}$. It follows that if t^o denotes the unique $\mathscr{R}$-normal form into which t rewrites, then $t \rightarrow_{\mathscr{ALD}} t'$ implies $t^o =_{\mathscr{LD}} t'^o$. □

As the word problem for $\mathscr{LD}$ is solvable (Dehornoy 2000), it follows that the word problem for $\mathscr{ALD}$ is solvable also.

So we see in this example that orienting the equation A from right to left leads to a solution of the word problem for $\mathscr{ALD}$, which was not the case with the other orientation. On the other hand, we saw in Section 4 that orienting A from left to right leads to a blueprint and to the construction of a torsion-free $\mathscr{ALD}$-algebra, which would be possible, but much less natural when using the other orientation. The conclusion is that both orientations have their own interest, and makes it difficult to predict *a priori* which orientation should be chosen.

Remark 6.4. In terms of the geometry monoid $\mathscr{G}eom_{\mathscr{ALD}}$, Proposition 6.3 (i) says that every non-empty element admits a decomposition of the form $g\,;h\,;g'^{-1}$, where g and g' belong to the submonoid generated by the maps $\rightarrow_{\bar{A},p}$ and $\rightarrow_{D,p}$, and h belongs to the submonoid generated by the maps $\rightarrow_{D,p}$ and $\leftarrow_{D,p}$. A similar result holds in $\mathtt{Geom}_{\mathscr{ALD}}$, but, as the empty map is an element of $\mathscr{G}eom_{\mathscr{ALD}}$, the decomposition is guaranteed only for those elements of $\mathtt{Geom}_{\mathscr{ALD}}$ that act on at least one term. Typically, the element $\bar{\mathtt{A}}_\varepsilon^{-1}\bar{\mathtt{A}}_0$, which corresponds to an empty map, is not eligible for the decomposition.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

7. Summary

We have shown how introducing by means of local confluence relations an abstract group that is supposed to mimic the properties of the geometry monoid, and hence of the initial rewrite system, and then internalising terms in that group so as to transform the initial external action into an internal multiplication may allow us to solve non-trivial questions about a given equational specification.

In the case of non-linear equations, that is, when some variable is repeated at least twice, due to the problem of the empty map, the geometry monoid is an intrinsically inconvenient object, and our approach of replacing it with a group is the only known one. Of course, one might consider other relations besides local confluence relations. The latter proved to be suitable for the examples considered here, but different schemes may prove relevant for other identities: for instance, when commutativity is involved, the maps are involutive, and the associated relations $\rightarrow_{E,p}\ ;\ \rightarrow_{E,p} = \mathrm{id}$ are not confluence relations. In some cases one can keep the principle of introducing the structure $\mathtt{Geom}_{\mathscr{E}}$ presented by the confluence relations in $\mathscr{G}eom_{\mathscr{E}}$, but taking $\mathtt{Geom}_{\mathscr{E}}$ to be a monoid rather than a group: typically, this has to be done in the case of the idempotency law $x = xx$, as, in this case, confluence relations of the type $\rightarrow_{E,\varepsilon}\ ;\ \rightarrow_{E,\varepsilon} = \rightarrow_{E,\varepsilon}\ ;\ \rightarrow_{E,0}\ ;\ \rightarrow_{E,1}$ are satisfied, prevenfting $\mathscr{G}eom_{\mathscr{E}}$ from admitting left cancellation (Jedlicka 2008). So we see that the methods described here require some flexibility in their application.

However, it should be possible to adapt all three main steps, namely introducing a monoid of partial maps, replacing it with a group using a presentation, and internalising the rewrite system by representing the objects on which the initial action is defined by copies inside the group, to more general frameworks. An example is given in Dehornoy (2005) where equations are replaced with a more complicated action on terms ('twisted commutativity'); we think that other rewrite systems, possibly of a completely different type, could be investigated using such tools. We are thinking, in particular, of combinatory logic (Smullyan 1995), for instance, of the word problem for the combinator $\boldsymbol{S}$ defined by $\boldsymbol{S}xyz = (xy)xz$, which is open, or the combinator $\boldsymbol{L}$ defined by $\boldsymbol{L}xy = x(yy)$, which has been solved – see, for example, Sprenger and Wymann-Böni (1993).

References

Bertling, H. and Ganzinger (1989) Completion-time optimization of rewrite-time goal solving. *Springer-Verlag Lecture Notes in Computer Science* **355** 45–58.

Baader, F. and Nipkow, T. (1998) *Term Rewriting and All That*, Cambridge University Press.

Book, R. (1987) Thue systems as rewriting systems. *J. Symb. Comput.* **3** 39–68.

Brin, M. (2004) The algebra of strand splitting: I. A braided version of Thompson's group V. arXiv math.GR/040642.

Brin, M. (2006) The algebra of strand splitting: II. A presentation for the braid group on one strand. *Intern. J. of Algebra and Computation* **16** 203–219.

Cannon, J. W., Floyd, W. J. and Parry, W. R. (1996) Introductory notes on Richard Thompson's groups. *Ens. Math.* **42** 215–257.

Clifford, A. H. and Preston, G. B. (1961) The Algebraic Theory of Semigroups, volume 1. *AMS Surveys* **7**.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press

Dehornoy, P. (1993) Structural monoids associated to equational varieties. *Proc. Amer. Math. Soc.* **117** (2) 293–304.
Dehornoy, P. (1996) The structure group for the associativity identity. *J. Pure Appl. Algebra* **111** 59–82.
Dehornoy, P. (2000) Braids and Self-Distributivity. *Progress in Math.* **192**, Birkhäuser.
Dehornoy, P. (2002) Study of an identity. *Alg. Universalis* **48** 223–248.
Dehornoy, P. (2003) Complete positive group presentations. *J. of Algebra* **268** 156–197.
Dehornoy, P. (2005) Geometric presentations of Thompson's groups. *J. Pure Appl. Algebra* **203** 1–44.
Dehornoy, P. (2006) The group of parenthesized braids. *Advances in Math.* **205** 354–409.
Dehornoy, P. (2007) Free augmented LD-systems. *J. Algebra and Appl.* **6-1** 173–187.
Dershowitz, N. and Jouannaud, J.P. (1994) Rewrite Systems. In: van Leeuwen, J. (ed.) *Handbook of Theor. Comp. Sc.* volume B, Elsevier.
Huet, G. (1980) Confluent reductions: Abstract properties and applications to term rewriting systems. *Journal of the ACM* **27** 797–821.
Jedlicka, P. (2008) Geometry monoid of left distributivity and left idempotency. *J. Algebra and Discrete Math.* (to appear).
Larue, D. (1994) *Left-distributive and left-distributive idempotent algebras*, Ph.D. Thesis, University of Colorado, Boulder.
Mac Lane, S. (1963) Natural associativity and commutativity. *Rice Univ. Studies* **49** 28–46.
McKenzie, R. and Thompson, R. (1973) An elementary construction of unsolvable word problems in group theory. In: Boone *et al.* (eds.) Word Problems. *Studies in Logic* **71**, North Holland.
Melliès, P.A. (2002) Axiomatic rewriting theory VI: residual theory revisited. Proceedings of RTA'02. *Springer-Verlag Lecture Notes in Computer Science* **2378** 23–50.
Paterson, A. (1998) Groupoids, inverse semigroups, and their operator algebras. *Progress in Math.* **170**, Birkhäuser.
Robinson, J. (1963) A machine-oriented logic based on the resolution principle. *J. of the ACM* **10** 163–174.
Smullyan, R. (1995) *To mock a mocking bird, and other logic puzzles including combinatory logic*, Knopf, New York.
Sprenger, M. and Wymann-Böni, M. (1993) How to decide the lark. *Theor. Comp. Sc.* **110** 419–432.
Terese (2003) *Term Rewriting Systems*, Cambridge Tracts in Theoretical Computer Science **55**, Cambridge University Press.
Thompson, R.J. (1980) Embeddings into finitely generated simple groups which preserve the word problem. In: Adjan, S.I. *et al.* (eds.) Word problems II: The Oxford book. *Studies in Logic* **95**, North Holland 401–440.

https://doi.org/10.1017/S0960129508007160 Published online by Cambridge University Press