A new finiteness condition for monoids presented by complete rewriting systems (after Craig C. Squier)

Yves LAFONT

CNRS
Laboratoire de Mathématiques Discrètes*
Email: lafont@lmd.univ-mrs.fr

January 11, 1994

Abstract

Recently, Craig Squier introduced the notion of *finite derivation type* to show that some finitely presentable monoid has no presentation by means of a finite complete rewriting system. A similar result was already obtained by the same author using homology, but the new method is more direct and more powerful. Here, we present Squier's argument with a bit of categorical machinery, making proofs shorter and easier. In addition we prove that, if a monoid has finite derivation type, then its third homology group is of finite type.

An invariant for a structure is something which can be calculated in many ways, but only depends on the structure itself. Typical examples are the dimension of a vector space or the genus of a surface. Squier's finiteness condition for monoids is of this kind: It can be defined in terms of a finite presentation, but does not depend on the choice of this presentation.

To begin with a simpler case, consider the following theorem, which is not hard to prove: If M is a finitely presentable monoid, Σ a finite alphabet and φ a surjective morphism from the free monoid Σ^* to M, then the congruence on Σ^* induced by φ is generated by some finite set $\mathcal{R} \subset \Sigma^* \times \Sigma^*$. In other words, the existence of a finite presentation for M does not depend on the choice of the set of generators, provided it is finite.

The invariance of Squier's finiteness condition is of the same nature, but it is a *2-dimensional word problem* in the sense of [Bu]. Therefore, we have to introduce a little more algebraic material (sections 1–2) before we get to the heart of the matter (sections 3–6). With this geometrical viewpoint, the connection with homology becomes quite natural (section 7).

I am grateful to Volker Diekert for pointing out the preprint of Squier, and to Friedrich Otto for the bibliographical references. Theorem 3 has been proved independently by Robert Cremanns and Friedrich Otto [CrOt].

*Address: 163 avenue de Luminy, case 930, 13288 MARSEILLE CEDEX 9, FRANCE.

1 Strict monoidal categories

A *strict monoidal category* is a category C equipped with an associative bifunctor $x, y \mapsto xy$ and a unit object 1. This means that we have a structure of monoid on the set of objects of C. Moreover, for each pair of arrows $x \xrightarrow{f} x'$ and $y \xrightarrow{g} y'$, there is an arrow $xy \xrightarrow{fg} x'y'$. This multiplication on arrows is associative with unit $1 \xrightarrow{id_1} 1$, and satisfy the following extra properties:

(1) $(f' \circ f)(g' \circ g) = f'g' \circ fg$ for any $x \xrightarrow{f} x' \xrightarrow{f'} x''$ and $y \xrightarrow{g} y' \xrightarrow{g'} y''$,

(2) $id_x id_y = id_{xy}$.

If x is an object and $y \xrightarrow{f} y'$ is an arrow, we shall write xf for $id_x f$ and fx for $f id_x$. With this convention, property (1) can be replaced by the following ones:

(1') $x(f' \circ f)y = xf'y \circ xfy$ for any x, y and $z \xrightarrow{f} z' \xrightarrow{f'} z''$,

(1") $fg = x'g \circ fy = fy' \circ xg$ for any $x \xrightarrow{f} x'$ and $y \xrightarrow{g} y'$.

In particular, the multiplication $f, g \mapsto fg$ is completely determined by the operations $x, f \mapsto xf$ and $f, x \mapsto fx$.

A *strict monoidal groupoid* is a strict monoidal category C such that the category C is a groupoid. In other words, every $x \xrightarrow{f} y$ has an *inverse* $y \xrightarrow{f^{-1}} x$ such that $f^{-1} \circ f = id_x$ and $f \circ f^{-1} = id_y$.

From now on, by *monoidal category*, we shall mean *strict monoidal category*, and similarly for monoidal groupoids.

Special kinds of monoidal groupoids arise in linear algebra. First, any abelian group U can be seen as a monoidal groupoid as follows:

- The set of objects is the trivial monoid $\{1\}$.
- An arrow $1 \xrightarrow{f} 1$ is given by an element of U.
- If $1 \xrightarrow{f} 1$ is given by $u \in U$ and $1 \xrightarrow{g} 1$ is given by $v \in U$, then both $1 \xrightarrow{g \circ f} 1$ and $1 \xrightarrow{fg} 1$ are given by $u + v$.

Note that here, property (1") is a direct consequence of the commutativity of addition in U. More generally, let M be a monoid, and assume that U is a $\mathbf{Z}M$-bimodule. This means that we have a left linear action $x, u \mapsto x \cdot u$ of M on U and a right linear one $u, x \mapsto u \cdot x$ such that $(x \cdot u) \cdot y = x \cdot (u \cdot y)$ for any $x, y \in M$ and $u \in U$. Then, U can be seen as a monoidal groupoid as follows:

- The set of objects is the monoid M.
- An arrow $x \xrightarrow{f} y$ is given by an element of U if $x = y$. Otherwise, there is no such arrow.
- If $x \xrightarrow{f} x$ is given by $u \in U$ and $x \xrightarrow{g} x$ is given by $v \in U$, then $x \xrightarrow{g \circ f} x$ is given by $u + v$.

- If $x \xrightarrow{f} x$ is given by $u \in U$ and $y \xrightarrow{g} y$ is given by $v \in U$, then $xy \xrightarrow{fg} xy$ is given by $u \cdot y + x \cdot v$.

Now, if U is just a left (respectively right) $\mathbf{Z}M$-module, it can also be seen as a $\mathbf{Z}M$-bimodule with $u \cdot x = u$ (respectively $x \cdot u = u$) for all $x \in M$ and $u \in U$. In that case $xy \xrightarrow{fg} xy$ is given by $u + x \cdot v$ (respectively $u \cdot y + v$).

A *2-congruence* on a monoidal category C is an equivalence relation $f \sim g$ on pairs of parallel arrows $x \overset{f,g}{\Longrightarrow} x'$ in C, which is compatible with composition and multiplication, namely:

- $k \circ f \circ h \sim k \circ g \circ h$ for any $x \xrightarrow{h} y \overset{f,g}{\Longrightarrow} y' \xrightarrow{k} z$ such that $f \sim g$,
- $xfy \sim xgy$ for any x, y and $z \overset{f,g}{\Longrightarrow} z'$ such that $f \sim g$.

Note that for compatibility with multiplication, it is enough to consider the operation $x, f, y \mapsto xfy$. The two basic examples of 2-congruences are:

- $f \sim g$ if $f = g$ (the smallest 2-congruence),
- $f \sim g$ for any $x \overset{f,g}{\Longrightarrow} x'$ (the largest one, or *full 2-congruence*).

In the case of a monoidal groupoid, we get for free the fact that $f^{-1} \sim g^{-1}$ if $f \sim g$, and furthermore, $\sim$ is completely determined by the set of arrows $x \xrightarrow{f} x$ such that $f \sim id_x$. In the case of a (left or right) $\mathbf{Z}M$-module, this set is a submodule, and conversely, any submodule corresponds to a 2-congruence.

If $\mathcal{P}$ is a set of pairs of parallel arrows in C, the smallest 2-congruence $\equiv_{\mathcal{P}}$ containing $\mathcal{P}$ is called the *2-congruence generated by* $\mathcal{P}$. In case $\mathcal{P}$ is finite, we say that the 2-congruence $\equiv_{\mathcal{P}}$ is *finitely generated*.

If C and C' are monoidal categories, a *2-morphism* $\Phi : C \to C'$ is a functor preserving the multiplicative structure. In the case of monoidal groupoids, we get for free the preservation of inverses. Finally the inverse image of a 2-congruence by such a 2-morphism is a 2-congruence. As a consequence we get:

Lemma 1 *Let $\mathcal{P}$ be a set of pairs of parallel arrows in C and $\sim$ a 2-congruence on C' such that $\Phi(f) \sim \Phi(g)$ for each $(f, g) \in \mathcal{P}$. Then $\Phi(f) \sim \Phi(g)$ for any $x \overset{f,g}{\Longrightarrow} y$ in C such that $f \equiv_{\mathcal{P}} g$.*

2 Presentations and categories of derivations

A *presentation* of a monoid M consists of an alphabet Σ and a binary relation $\mathcal{R}$ on the free monoid Σ^* such that M is isomorphic to $\Sigma^*/\equiv_{\mathcal{R}}$ where $\equiv_{\mathcal{R}}$ is the congruence generated by $\mathcal{R}$. If x is a word in Σ^*, we write $\overline{x}$ for the corresponding element in M. In particular, the unit of M is $\overline{1}$ where 1 is the empty word.

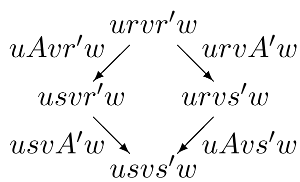
As a first step, we construct the *category of derivations* for this presentation as follows:

- An object is a word in Σ^*.
- An *atomic derivation* $r \xrightarrow{A} s$ is given by a pair $(r, s) \in \mathcal{R}$.
- An *elementary derivation* $x \xrightarrow{E} y$ is given by two words $u, v \in \Sigma^*$ and an atomic derivation $r \xrightarrow{A} s$ such that $x = urv$ and $y = usv$. If $u = v = 1$, we identify E with the atomic derivation A.
- A *derivation* $x \xrightarrow{F} y$ is given by a sequence

$$x = x_0 \xrightarrow{E_1} x_1 \xrightarrow{E_2} \cdots \xrightarrow{E_n} x_n = y$$

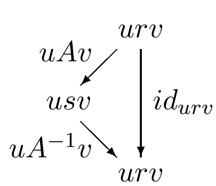
of elementary derivations. If $n = 1$, we identify F with the elementary derivation E_1. If $n = 0$, we get the *identity derivation* id_x.

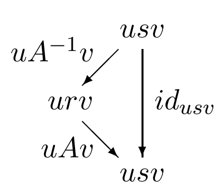
Composition of derivations is defined in the obvious way. Also, if x, y are words and $z \xrightarrow{F} z'$ is a derivation, the derivation $xzy \xrightarrow{xFy} xz'y$ is defined in the obvious way, so that property (1') is satisfied, but not (1"). In order to get a monoidal category, me must consider the set of derivations modulo the equivalence relation generated by *permutation of disjoint derivations.* This means that, if u, v, w are words and $r \xrightarrow{A} s$ and $r' \xrightarrow{A'} s'$ are atomic derivations, we identify the derivations $usvA'w \circ uAvr'w$ and $uAvs'w \circ urvA'w$:



More generally, if $x \xrightarrow{F} urvr'w$ and $usvs'w \xrightarrow{G} y$ are arbitrary derivations, we identify the derivations $G \circ usvA'w \circ uAvr'w \circ F$ and $G \circ uAvs'w \circ urvA'w \circ F$. It is easy to see that, if $x \xrightarrow{F} x'$ and $y \xrightarrow{G} y'$ are arbitrary derivations, then $x'G \circ Fy$ and $Fy' \circ xG$ are equivalent by permutation of disjoint derivations. Therefore, the multiplication of equivalence classes of derivations is well defined by means of (1"), and we get the *free monoidal category* $\mathbf{M}(\Sigma, \mathcal{R})$ generated by the presentation $\Sigma, \mathcal{R}$. It is indeed characterized by a universal property whose precise formulation is left to the reader.

The *free monoidal groupoid* $\mathbf{G}(\Sigma, \mathcal{R})$ is constructed in the same way, introducing a *positive atomic derivation* $A : r \to s$ and a *negative* one $A^{-1} : s \to r$ for each $(r, s) \in \mathcal{R}$. Here, we must consider the set of derivations modulo the equivalence relation generated by permutation of disjoint derivations and *cancellation.* This means that, if u, v are words and $r \xrightarrow{A} s$ is a positive atomic derivation, we identify $uA^{-1}v \circ uAv$ with id_{urv} and $uAv \circ uA^{-1}v$ with id_{usv}:





Again, this applies inside a derivation: for example, if $x \xrightarrow{F} urv$ and $usv \xrightarrow{G} y$ are arbitrary derivations, we identify $G \circ uA^{-1}v \circ uAv \circ F$ with $G \circ F$.

Following Squier, we shall handle derivations rather than equivalence classes of derivations. This is just a matter of rhetoric. For example, a 2-congruence on $\mathbf{M}(\Sigma, \mathcal{R})$ can be seen as an equivalence relation $\sim$ on the set of derivations satisfying the following properties:

- $K \circ F \circ H \sim K \circ G \circ H$ for any $x \xrightarrow{H} y \overset{F,G}{\Longrightarrow} y' \xrightarrow{K} z$ such that $F \sim G$,
- $xFy \sim xGy$ for any x, y and $z \overset{F,G}{\Longrightarrow} z'$ such that $F \sim G$,
- $suA' \circ Aur' \sim Aus' \circ ruA'$ for any u and for any atomic derivations $r \xrightarrow{A} s$ and $r' \xrightarrow{A'} s'$.

In the case of $\mathbf{G}(\Sigma, \mathcal{R})$, there is an extra condition:

- $A^{-1} \circ A \sim id_r$ and $A \circ A^{-1} \sim id_s$ for any positive atomic derivation $r \xrightarrow{A} s$.

If $\Sigma', \mathcal{R}'$ is a presentation of another monoid M', it is clear that any 2-morphism $\Phi : \mathbf{G}(\Sigma, \mathcal{R}) \to \mathbf{G}(\Sigma', \mathcal{R}')$ induces a unique morphism $\varphi : M \to M'$ such that $\varphi(\overline{x}) = \overline{\Phi(x)}$ for all $x \in \Sigma^*$.

Conversely, any morphism $\varphi : M \to M'$ is induced by such a 2-morphism. Indeed, for each $a \in \Sigma$, we can choose a word $x_a \in \Sigma'^*$ such that $\overline{x_a} = \varphi(\overline{a})$. This map extends to a morphism $\xi : \Sigma^* \to \Sigma'^*$ such that $\varphi(\overline{x}) = \overline{\xi(x)}$ for all $x \in \Sigma^*$. Now, for each positive atomic derivation $r \xrightarrow{A} s$ in $\mathbf{G}(\Sigma, \mathcal{R})$, we have $\overline{\xi(r)} = \varphi(\overline{r}) = \varphi(\overline{s}) = \overline{\xi(s)}$ and we can choose a derivation $\xi(r) \xrightarrow{F_A} \xi(s)$. This map extends to a 2-morphism $\Phi : \mathbf{G}(\Sigma, \mathcal{R}) \to \mathbf{G}(\Sigma', \mathcal{R}')$ such that $\Phi(x) = \xi(x)$ for all $x \in \Sigma^*$.

Of course, there are many arbitrary choices in the construction of Φ. If $\Psi : \mathbf{G}(\Sigma, \mathcal{R}) \to \mathbf{G}(\Sigma', \mathcal{R}')$ is another 2-morphism inducing the same morphism $\varphi : M \to M'$, then $\overline{\Phi(x)} = \overline{\Psi(x)}$ for all $x \in \Sigma^*$. In particular, we can choose a derivation $\Phi(a) \xrightarrow{H_a} \Psi(a)$ for each $a \in \Sigma$. In this way, we define a derivation $\Phi(x) \xrightarrow{H_x} \Psi(x)$ for all $x \in \Sigma^*$ such that $H_{xy} = H_x H_y$ for all $x, y \in \Sigma^*$. This H does not define a natural transformation between Φ and Ψ, but we have the following lemma, which is proved by a straightforward induction:

Lemma 2 *Let $\sim$ be a 2-congruence such that $H_s \circ \Phi(A) \sim \Psi(A) \circ H_r$ for each positive atomic derivation $r \xrightarrow{A} s$ in $\mathbf{G}(\Sigma, \mathcal{R})$. Then $H_y \circ \Phi(F) \sim \Psi(F) \circ H_x$ for any derivation $x \xrightarrow{F} y$ in $\mathbf{G}(\Sigma, \mathcal{R})$.*

$$\begin{array}{ccc} \Phi(r) & \xrightarrow{\Phi(A)} & \Phi(s) \\ {\scriptstyle H_r}\downarrow & & \downarrow{\scriptstyle H_s} \\ \Psi(r) & \xrightarrow[\Psi(A)]{} & \Psi(s) \end{array} \qquad\qquad \begin{array}{ccc} \Phi(x) & \xrightarrow{\Phi(F)} & \Phi(y) \\ {\scriptstyle H_x}\downarrow & & \downarrow{\scriptstyle H_y} \\ \Psi(x) & \xrightarrow[\Psi(F)]{} & \Psi(y) \end{array}$$

3 The finiteness condition

We say that a finite presentation $\Sigma, \mathcal{R}$ is of *finite derivation type* if the full 2-congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ is finitely generated.

Theorem 1 [Sq2]

Let $\Sigma, \mathcal{R}$ and $\Sigma', \mathcal{R}'$ be finite presentations of isomorphic monoids M and M'. Then $\Sigma, \mathcal{R}$ is of finite derivation type if and only if $\Sigma', \mathcal{R}'$ is of finite derivation type.

Proof: Assume for example that the full 2-congruence on $\mathbf{G}(\Sigma', \mathcal{R}')$ is generated by a finite set $\mathcal{P}$ of pairs of parallel derivations. By hypothesis, we have two morphisms $\varphi : M \to M'$ and $\varphi' : M' \to M$ such that $\varphi' \circ \varphi$ is the identity morphism on M and $\varphi \circ \varphi'$ is the identity morphism on M'. The first one is induced by a 2-morphism $\Phi : \mathbf{G}(\Sigma, \mathcal{R}) \to \mathbf{G}(\Sigma', \mathcal{R}')$ and the second one by a 2-morphism $\Phi' : \mathbf{G}(\Sigma', \mathcal{R}') \to \mathbf{G}(\Sigma, \mathcal{R})$. In particular, the identity morphism on M is induced by the identity 2-morphism on $\mathbf{G}(\Sigma, \mathcal{R})$ and also by $\Phi' \circ \Phi$. The above construction yields a derivation $x \xrightarrow{H_x} \Phi'(\Phi(x))$ for each $x \in \Sigma^*$.

Let $\sim$ be the smallest 2-congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ such that $\Phi'(F) \sim \Phi'(G)$ for each $(F, G) \in \mathcal{P}$ and $H_s \circ A \sim \Phi'(\Phi(A)) \circ H_r$ for each positive atomic derivation $r \xrightarrow{A} s$ in $\mathbf{G}(\Sigma, \mathcal{R})$. This 2-congruence is finitely generated. More precisely, if $\mathcal{P}$ has p elements and $\mathcal{R}$ has q elements, then $\sim$ is generated by at most $p + q$ elements.

Let $x \overset{F,G}{\Longrightarrow} y$ be any pair of parallel derivations in $\mathbf{G}(\Sigma, \mathcal{R})$. Since $\equiv_{\mathcal{P}}$ is the full 2-congruence on $\mathbf{G}(\Sigma', \mathcal{R}')$, we have $\Phi(F) \equiv_{\mathcal{P}} \Phi(G)$. Hence, we get $\Phi'(\Phi(F)) \sim \Phi'(\Phi(G))$ by lemma 1, and

$$F \sim H_y^{-1} \circ \Phi'(\Phi(F)) \circ H_x \sim H_y^{-1} \circ \Phi'(\Phi(G)) \circ H_x \sim G$$

by lemma 2.

$$\begin{array}{ccc} x & \overset{F}{\underset{G}{\rightrightarrows}} & y \\ {\scriptstyle H_x}\big\downarrow & & \big\downarrow{\scriptstyle H_y} \\ \Phi'(\Phi(x)) & \overset{\Phi'(\Phi(F))}{\underset{\Phi'(\Phi(G))}{\rightrightarrows}} & \Phi'(\Phi(y)) \end{array}$$

This means that $\sim$ is the full 2-congruence on $\mathbf{G}(\Sigma, \mathcal{R})$. **Q.e.d.**

If a finite presentation of a monoid M is of finite derivation type, we say that the monoid M *has finite derivation type*. We have just proved that this property does not depend on this presentation, provided it is finite. In fact, we have only used the fact that $\mathcal{R}$ is finite, not Σ.

Let $\Sigma, \mathcal{R}$ be an arbitrary presentation of a monoid M and $\sim$ a 2-congruence on $\mathbf{G}(\Sigma, \mathcal{R})$. We say that $\mathcal{R}$ is *$\sim$-finite* if there is a finite subset $\mathcal{R}_0$ of $\mathcal{R}$ such that, for any derivation $x \xrightarrow{F} y$ in $\mathbf{G}(\Sigma, \mathcal{R})$, there is a derivation $x \xrightarrow{\widetilde{F}} y$ in $\mathbf{G}(\Sigma, \mathcal{R}_0)$ such that $F \sim \widetilde{F}$. It is clearly enough to check this when F is atomic.

Lemma 3 *If M has finite derivation type and $\mathcal{R}$ is $\sim$-finite, then there is a finite set $\mathcal{P}_0$ of pairs of parallel derivations in $\mathbf{G}(\Sigma, \mathcal{R})$ such that the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ is generated by $\sim$ and $\mathcal{P}_0$.*

Proof: Obviously $\Sigma, \mathcal{R}_0$ is a presentation of M, and since $\mathcal{R}_0$ is finite, there is a finite set $\mathcal{P}_O$ of pairs of parallel derivations in $\mathbf{G}(\Sigma, \mathcal{R}_0)$ generating the full congruence on $\mathbf{G}(\Sigma, \mathcal{R}_0)$. It is clear that the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ is generated by $\sim$ and $\mathcal{P}_0$. **Q.e.d.**

4 Derivations in a complete presentation

Let $\Sigma, \mathcal{R}$ be a *noetherian* presentation of a monoid M. This means that there is no infinite sequence $x_0 \xrightarrow{E_1} x_1 \xrightarrow{E_2} \cdots \xrightarrow{E_n} x_n \xrightarrow{E_{n+1}} \cdots$ of elementary derivations. Then, for any $x \in \Sigma^*$, there is a derivation $x \xrightarrow{F} y$ where y is *reduced*, which means that no elementary derivation starts from y. This y is called a *normal form* of x.

Let $\sim$ be a 2-congruence on $\mathbf{M}(\Sigma, \mathcal{R})$. A *peak* is an unordered pair of elementary derivations $x \xrightarrow{E} y$ and $x \xrightarrow{E'} y'$ starting from the same word x. Such a peak is called *confluent* if there is a word z and two derivations $y \xrightarrow{F} z$ and $y' \xrightarrow{F'} z$.

$$\begin{array}{ccccc} & & x & & \\ & {}^{E}\swarrow & & \searrow^{E'} & \\ y & & & & y' \\ & {}_{F}\searrow & & \swarrow_{F'} & \\ & & z & & \end{array}$$

It is called $\sim$*-confluent* if it is confluent and F, F' can be chosen in such a way that $F \circ E \sim F' \circ E'$. It is called *critical* if $E \neq E'$ and if it is of the form

$$\begin{array}{ccc} & rv = u'r' & \\ {}^{Av}\swarrow & & \searrow^{u'A'} \\ sv & & u's' \end{array} \qquad \text{or} \qquad \begin{array}{ccc} & urv = r' & \\ {}^{uAv}\swarrow & & \searrow^{A'} \\ usv & & s' \end{array}$$

where, in the first case, u' is a strict prefix of r, or equivalently, v is a strict suffix of r'. Note that, in case $\mathcal{R}$ is finite, there are only finitely many critical peaks.

Lemma 4 *If all critical peaks are $\sim$-confluent, then all peaks are $\sim$-confluent.*

Proof: A peak is of the form

$$\begin{array}{ccc} & x & \\ {}^{E}\swarrow & & \searrow^{E} \\ y & & y \end{array} \qquad \text{or} \qquad \begin{array}{ccc} & urvr'w & \\ {}^{uAvr'w}\swarrow & & \searrow^{urvA'w} \\ usvr'w & & urvs'w \end{array} \qquad \text{or} \qquad \begin{array}{ccc} & uxv & \\ {}^{uEv}\swarrow & & \searrow^{uE'v} \\ uyv & & uy'v \end{array}$$

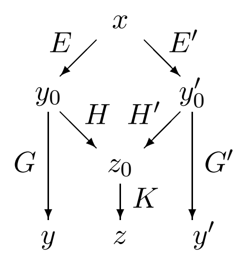
where, in the third case, E, E' form a critical peak. In the first case, take $z = y$ and $F = F' = id_y$. In the second case, take $z = usvs'w$, $F = usvA'w$ and $F' = uAvs'w$. In the third case, apply the hypothesis of the lemma. **Q.e.d.**

Now we assume that the hypothesis of lemma 4 is satisfied.

Lemma 5 *If $x \xrightarrow{F} y$ and $x \xrightarrow{F'} y'$ are any derivations such that y and y' are reduced, then $y = y'$ and $F \sim F'$.*

In other words, the normal form of a word is unique, and the derivation leading to this normal form is unique modulo $\sim$.

Proof: By noetherian induction on x. If x is reduced then $y = x = y'$ and $F = id_x = F'$. Otherwise, $F = G \circ E$ and $F' = G' \circ E'$ where $x \xrightarrow{E} y_0$ and $x' \xrightarrow{E'} y'_0$ are elementary derivations. By lemma 4, we get z_0 and two derivations $y_0 \xrightarrow{H} z_0$ and $y'_0 \xrightarrow{H'} z_0$ such that $H \circ E \sim H' \circ E'$. Moreover, there is a derivation $z_0 \xrightarrow{K} z$ where z is reduced. We can apply the induction hypothesis to y_0 and y'_0.



We get $y = z = y'$ and $F = G \circ E \sim K \circ H \circ E \sim K \circ H' \circ E' \sim G' \circ E' = F'$. **Q.e.d.**

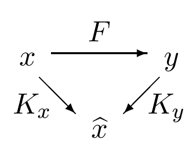
A noetherian presentation is called *complete* (or *canonical*) if all critical peaks are confluent. This implies the confluence of all peaks and the uniqueness of normal forms. In particular, if this presentation is finite, the word problem is decidable. For more details, we refer to [LaPr]. The more general case of term rewriting systems is explained in [Hu, Le].

Theorem 2 [Sq2]

If M has a finite complete presentation, then M has finite derivation type.

Proof: Let $\sim$ be the 2-congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ generated by the confluence diagrams of critical peaks. Since there are only finitely many such critical peaks, $\sim$ is finitely generated. Furthermore, the restriction of $\sim$ to $\mathbf{M}(\Sigma, \mathcal{R})$ is also a 2-congruence, and we can apply the previous lemma.

For each word x, we choose a derivation $x \xrightarrow{K_x} \widehat{x}$ in $\mathbf{M}(\Sigma, \mathcal{R})$ where $\widehat{x}$ is the unique normal form of x. By lemma 5, we have $K_y \circ F \sim K_x$ for any derivation $x \xrightarrow{F} y$ in $\mathbf{M}(\Sigma, \mathcal{R})$.



By induction, this property extends to all derivations in $\mathbf{G}(\Sigma, \mathcal{R})$, and if $x \overset{F,G}{\Longrightarrow} y$ is a pair of parallel derivations in $\mathbf{G}(\Sigma, \mathcal{R})$, we have $F \sim K_y^{-1} \circ K_x \sim G$. This means that $\sim$ is the full 2-congruence on $\mathbf{G}(\Sigma, \mathcal{R})$. **Q.e.d.**

In the case of an infinite complete presentation, the same argument shows that the full 2-congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ is generated by the confluence diagrams of critical peaks.

5 First counterexample

Theorem 2 can be used to show that a monoid has no finite complete presentation, by checking that it does not have finite derivation type.

The following counterexample comes from [LaPr]. We consider the monoid M presented by means of $\Sigma = \{a, b, c, d, d'\}$ and

$$\mathcal{R}_0 = \{(ab, a), (da, ac), (d'a, ac)\}.$$

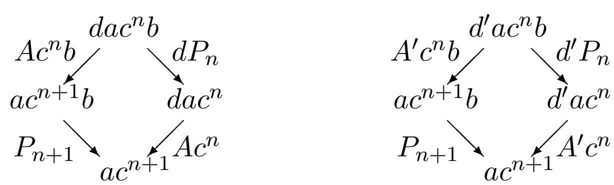
There is an infinite complete presentation of M by means of Σ and

$$\mathcal{R} = \{(ac^n b, ac^n); n \in \mathbf{N}\} \cup \{(da, ac), (d'a, ac)\}.$$

In particular, the word problem for M is clearly decidable. The atomic derivations for this presentation are:

$$ac^n b \xrightarrow{P_n} ac^n, \qquad da \xrightarrow{A} ac, \qquad d'a \xrightarrow{A'} ac.$$

There are two infinite families of confluent critical peaks:



We know that the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ wo families of confluence diagrams. On the other hand, it is clear that $\mathcal{R}$ is $\sim$-finite, where $\sim$ is the 2-congruence generated by the first family of confluence diagrams.

Now, we consider the right $\mathbf{Z}M$-module $\mathbf{Z}M$ as a monoidal groupoid. For simplicity, we identify each arrow of the monoidal groupoid $\mathbf{Z}M$ with the corresponding element of the right $\mathbf{Z}M$-module $\mathbf{Z}M$, and we define a 2-morphism $\Phi : \mathbf{G}(\Sigma, \mathcal{R}) \to \mathbf{Z}M$ as follows:

- $\Phi(x) = \overline{x}$ for all $x \in \Sigma^*$,
- $\Phi(P_n) = 0$, $\Phi(A) = 0$ and $\Phi(A') = \overline{1}$.

This definition makes sense since $\overline{x} = \overline{y}$ for any derivation $x \xrightarrow{F} y$ in $\mathbf{G}(\Sigma, \mathcal{R})$. To evaluate Φ on arbitrary derivations, the following formulae can be applied:

$$\Phi(G \circ F) = \Phi(F) + \Phi(G), \qquad \Phi(F^{-1}) = -\Phi(F), \qquad \Phi(xFy) = \Phi(F)\overline{y}.$$

The asymmetry in the last formula comes from the fact that $\mathbf{Z}M$ is considered as a right $\mathbf{Z}M$-module. For example, we get:

- $\Phi(P_{n+1} \circ Ac^n b) - \Phi(Ac^n \circ dP_n) = 0$,
- $\Phi(P_{n+1} \circ A'c^n b) - \Phi(A'c^n \circ d'P_n) = \overline{c^n b} - \overline{c^n}$.

In particular, $\Phi(G) - \Phi(F) = 0$ for any $x \stackrel{F,G}{\Longrightarrow} y$ such that $F \sim G$.

If M has finite derivation type, then by lemma 3, there is a finite set $\mathcal{P}_0$ of pairs of parallel derivations in $\mathbf{G}(\Sigma, \mathcal{R})$ such that the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ is generated by $\sim$ and $\mathcal{P}_0$. Clearly, this implies that the right module V generated by the infinite family $(\overline{c^n b} - \overline{c^n})_{n \in \mathbf{N}}$ is of finite type. We shall see that this is impossible.

We say that a word $x \in \Sigma^*$ is *prefix reduced* if xy is reduced for any reduced word $y \in \Sigma^*$. A prefix reduced word is reduced, but the converse is not true. For example a is reduced, but not prefix reduced, because ab is not reduced. On the other hand, both $c^n b$ and c^n are prefix reduced. Note that if x is prefix reduced, then $\overline{x}$ is left simplifiable in M.

First we show that the family $(\overline{c^n b})_{n \in \mathbf{N}}$ is free in the right $\mathbf{Z}M$-module $\mathbf{Z}M$. Assume indeed that $\sum_{n \in \mathbf{N}} \overline{c^n b} \cdot \lambda_n = 0$ where the λ_n are elements of the ring $\mathbf{Z}M$ and all but finitely many of them are zero. Since the $c^n b$ are prefix reduced, and none is a prefix of another, it is easy to see that $\lambda_n = 0$ for all $n \in \mathbf{N}$.

If $n \in \mathbf{N}$, we define a $\mathbf{Z}$-linear map $\Pi_n : \mathbf{Z}M \to \mathbf{Z}M$ by $\Pi_n(\overline{x}) = \overline{x}$ if x is a reduced word of length n and $\Pi_n(\overline{x}) = 0$ if x is a reduced word of length $m \neq n$. If $\lambda \in \mathbf{Z}M$ and $\lambda \neq 0$, we write $|\lambda|$ for the largest n such that $\Pi_n(\lambda) \neq 0$.

Now we can show that the family $(\overline{c^n b} - \overline{c^n})_{n \in \mathbf{N}}$ is free in the right $\mathbf{Z}M$-module $\mathbf{Z}M$. Assume indeed that $\sum_{n \in \mathbf{N}} (\overline{c^n b} - \overline{c^n}) \cdot \lambda_n = 0$. If not all λ_n are zero, let $N = \max\{n + 1 + |\lambda_n|; \lambda_n \neq 0\}$. Since the $c^n b$ and the c^n are prefix reduced, we get

$$\Pi_N\Big(\sum_{n \in \mathbf{N}} (\overline{c^n b} - \overline{c^n}) \cdot \lambda_n\Big) = \sum_{n \in \mathbf{N}} \overline{c^n b} \cdot \Pi_{N-n-1}(\lambda_n) = 0.$$

Since $(\overline{c^n b})_{n \in \mathbf{N}}$ is free, we get $\Pi_{N-n-1}(\lambda_n) = 0$ for all $n \in \mathbf{N}$, which is in contradiction with the definition of N.

The right module V cannot be of finite type, and therefore the finitely presentable monoid M does not have finite derivation type. In particular, it has no finite complete presentation. This was already shown in [LaPr], using the homological characterization of [Sq1].

6 Second counterexample

The following counterexample comes from [Sq1] and is treated in [Sq2]. We consider the monoid M presented by means of $\Sigma = \{a, b, c, d, e\}$ and

$$\mathcal{R}_0 = \{(ab, 1), (da, acd), (db, bd), (dc, cd), (de, 1)\}.$$

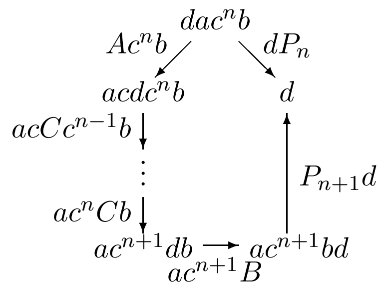
There is an infinite complete presentation of M by means of Σ and

$$\mathcal{R} = \{(ac^n b, 1); n \in \mathbf{N}\} \cup \{(da, acd), (db, bd), (dc, cd), (de, 1)\}.$$

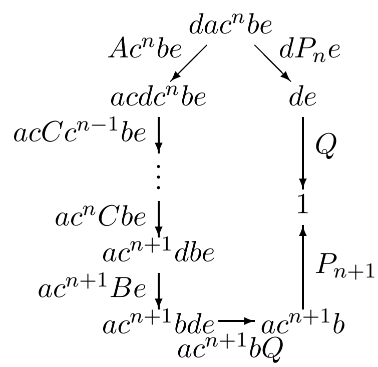
In particular, the word problem for M is clearly decidable. The atomic derivations for this presentation are:

$$ac^n b \xrightarrow{P_n} 1, \qquad da \xrightarrow{A} acd, \qquad db \xrightarrow{B} bd, \qquad dc \xrightarrow{C} cd, \qquad de \xrightarrow{Q} 1.$$

There is an infinite family of confluent critical peaks:



We know that the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ is generated by this family of confluence diagrams. On the other hand, it is clear that $\mathcal{R}$ is $\sim$-finite, where $\sim$ is the 2-congruence generated by the following family of diagrams:



As in the previous case, we consider the right $\mathbf{Z}M$-module $\mathbf{Z}M$ as a monoidal groupoid and we define a 2-morphism $\Phi : \mathbf{G}(\Sigma, \mathcal{R}) \to \mathbf{Z}M$ as follows:

- $\Phi(x) = \overline{x}$ for all $x \in \Sigma^*$,
- $\Phi(P_n) = 0$, $\Phi(A) = \overline{1}$, $\Phi(B) = 0$, $\Phi(C) = 0$ and $\Phi(Q) = 0$.

From this definition, we get:

- $\Phi(P_{n+1}d \circ ac^{n+1}B \circ ac^nCb \circ \cdots \circ acCc^{n-1}b \circ Ac^nb) - \Phi(dP_n) = c^nb$,
- $\Phi(P_{n+1} \circ ac^{n+1}bQ \circ ac^{n+1}Be \circ ac^nCbe \circ \cdots \circ acCc^{n-1}be \circ Ac^nbe) - \Phi(Q \circ dP_ne) = c^nbe$.

If M has finite derivation type, there is a finite set $\mathcal{P}_0$ of pairs of parallel derivations in $\mathbf{G}(\Sigma, \mathcal{R})$ such that the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$ is generated by $\sim$ and $\mathcal{P}_0$. Clearly, this implies that the right module V generated by the infinite families $(\overline{c^nbe})_{n\in\mathbf{N}}$ and $(\overline{c^nb})_{n\in\mathbf{N}}$ is also generated by $(\overline{c^nbe})_{n\in\mathbf{N}}$ and a finite subfamily of $(\overline{c^nb})_{n\in\mathbf{N}}$. This is impossible because both c^nb and c^nbe are prefix reduced, no c^nb is a prefix of another one, and no c^nbe is a prefix of a c^mb, so that no c^nb is a superfluous generator for V.

Thus we have another example of a finitely presentable monoid which has no finite complete presentation. The point here is that the homological argument fails because the homology groups of M are all of finite type. This follows easily from [Sq1] (up to dimension 3) and [Gr, Ko] (in all dimensions). In other words, this example shows the superiority of the new condition over the homological one.

7 Finiteness condition and homology

The following fact was conjectured in [Sq2]:

Theorem 3 *If a finitely presentable monoid M has finite derivation type, then there is a partial resolution*

$$C_3 \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} \mathbf{Z} \to 0$$

where the C_i are finite dimensional free left $\mathbf{Z}M$-modules. In particular, the third homology group $H_3(M)$ is of finite type.

Here we consider $\mathbf{Z}$ as a left $\mathbf{Z}M$-module where $\overline{x} \cdot n = n$ for any $\overline{x} \in M$ and $n \in \mathbf{Z}$.

Proof: Here is the geometrical intuition: From a finite presentation of M, we get a 3-dimensional cellular complex with a single 0-cell, one 1-cell for each generator, one 2-cell for each relation and one 3-cell for each pair of parallel derivations belonging to some finite set generating the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$. By making M act freely on all those cells, we shall build the 3-skeleton of a contractible space whose chain complex is a free resolution of $\mathbf{Z}$. For more details, we refer to [LaPr].

Let $\Sigma, \mathcal{R}$ be a finite presentation of M, and $\mathcal{P}$ a finite set of pairs of parallel derivations generating the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$. Let $C_0 = \mathbf{Z}M$, and let C_1 (respectively C_2 and C_3) be the free left $\mathbf{Z}M$-module generated by the finite set Σ (respectively $\mathcal{R}$ and $\mathcal{P}$).

If $a \in \Sigma$, we write $[a]$ for the corresponding generator in C_1. This notation is extended to all words by means of the following formulae:

$$[1] = 0, \qquad [xy] = [x] + \overline{x} \cdot [y].$$

Similarly, if $r \xrightarrow{A} s$ is a positive atomic derivation in $\mathbf{G}(\Sigma, \mathcal{R})$, we write $[A]$ for the corresponding generator in C_2, and this notation is extended to all derivations by means of the following formulae:

$$[xFy] = \overline{x} \cdot [F], \qquad [id_x] = 0, \qquad [G \circ F] = [F] + [G], \qquad [F^{-1}] = -[F].$$

Finally, if $x \overset{F,G}{\Longrightarrow} y$ is in $\mathcal{P}$, we write $[F, G]$ for the corresponding generator in C_3. The $\mathbf{Z}M$-linear boundary maps are defined as follows:

- $\varepsilon(\overline{1}) = 1$,
- $\partial_1[a] = \overline{a} - \overline{1}$ for each $a \in \Sigma$,
- $\partial_2[A] = [s] - [r]$ for each positive atomic derivation $r \xrightarrow{A} s$,
- $\partial_3[F, G] = [G] - [F]$ for each $x \overset{F,G}{\Longrightarrow} y$ in $\mathcal{P}$.

As consequences, we get the following formulae:

- $\varepsilon(\overline{x}) = 1$,
- $\partial_1[x] = \overline{x} - \overline{1}$ for any $x \in \Sigma^*$,

- $\partial_2[F] = [y] - [x]$ for any derivation $x \xrightarrow{F} y$ in $\mathbf{G}(\Sigma, \mathcal{R})$.

In particular, $\varepsilon\partial_1 = 0$, $\partial_1\partial_2 = 0$ and $\partial_2\partial_3 = 0$. To show that this sequence is exact, we shall construct a contracting homotopy, *i.e.* a sequence of $\mathbf{Z}$-linear maps

$$C_3 \xleftarrow{\sigma_3} C_2 \xleftarrow{\sigma_2} C_1 \xleftarrow{\sigma_1} C_0 \xleftarrow{\eta} \mathbf{Z}$$

such that $\varepsilon\eta = id_{\mathbf{Z}}$, $\partial_1\sigma_1 + \eta\varepsilon = id_{C_0}$, $\partial_2\sigma_2 + \sigma_1\partial_1 = id_{C_1}$ and $\partial_3\sigma_3 + \sigma_2\partial_2 = id_{C_2}$.

First we define the $\mathbf{Z}$-linear map $\mathbf{Z} \xrightarrow{\eta} C_0$ by $\eta(1) = \overline{1}$, so that $\varepsilon\eta = id_{\mathbf{Z}}$. Then, for each $\overline{x} \in M$, we choose a word $\widehat{x}$ in the congruence class of x and we define the $\mathbf{Z}$-linear map $C_0 \xrightarrow{\sigma_1} C_1$ by $\sigma_1(\overline{x}) = [\widehat{x}]$. In particular,

$$(\partial_1\sigma_1 + \eta\varepsilon)(\overline{x}) = \overline{x} - \overline{1} + \overline{1} = \overline{x},$$

so that $\partial_1\sigma_1 + \eta\varepsilon = id_{C_0}$.

Similarly, for each $\overline{x} \in M$ and for each $a \in \Sigma$, we choose a derivation $\widehat{xa} \xrightarrow{\Lambda(x,a)} \widehat{x}a$ and we define the $\mathbf{Z}$-linear map $C_1 \xrightarrow{\sigma_2} C_2$ by $\sigma_2(\overline{x} \cdot [a]) = [\Lambda(x,a)]$. In particular,

$$(\partial_2\sigma_2 + \sigma_1\partial_1)(\overline{x} \cdot [a]) = [\widehat{x}a] - [\widehat{xa}] + [\widehat{xa}] - [\widehat{x}] = \overline{x} \cdot [a],$$

so that $\partial_2\sigma_2 + \sigma_1\partial_1 = id_{C_1}$. We also define a derivation $\widehat{x} \xrightarrow{\Lambda(x)} x$ for each $x \in \Sigma^*$ by means of the following formulae:

$$\Lambda(1) = id_1, \qquad \Lambda(xa) = \Lambda(x)a \circ \Lambda(x,a),$$

so that $\sigma_2[x] = [\Lambda(x)]$ for all $x \in \Sigma^*$.

Finally, using the fact that $\mathcal{P}$ generates the full congruence on $\mathbf{G}(\Sigma, \mathcal{R})$, it is easy to see that for each pair $x \overset{F,G}{\Longrightarrow} y$ of parallel derivations in $\mathbf{G}(\Sigma, \mathcal{R})$, there is an element $\langle F, G\rangle \in C_3$ such that $\partial_3\langle F, G\rangle = [G] - [F]$. We define the $\mathbf{Z}$-linear map $C_2 \xrightarrow{\sigma_3} C_3$ by $\sigma_3(\overline{x} \cdot [A]) = \langle \Lambda(\widehat{x}s), \widehat{x}A \circ \Lambda(\widehat{x}r)\rangle$. In particular,

$$(\partial_3\sigma_3 + \sigma_2\partial_2)(\overline{x} \cdot [A]) = [\widehat{x}A \circ \Lambda(\widehat{x}r)] - [\Lambda(\widehat{x}s)] + [\Lambda(\widehat{x}s)] - [\Lambda(\widehat{x}r)] = \overline{x} \cdot [A],$$

so that $\partial_3\sigma_3 + \sigma_2\partial_2 = id_{C_2}$. Hence, our sequence is indeed a partial free resolution of $\mathbf{Z}$. **Q.e.d.**

In particular, we get an alternative proof of the following theorem:

Theorem 4 [Sq1]

If M has a finite complete presentation, then $H_3(M)$ is of finite type.

The second counterexample in section 6 shows that a finitely presentable monoid M such that $H_3(M)$ is of finite type does not necessarily have finite derivation type. In particular the converse of theorem 4 does not hold, even if M is a finitely presentable monoid with a decidable word problem.

8 Questions

At the end of his paper, Squier asks several questions, in particular:

- If a finitely presented monoid M has finite derivation type, does M have a finite complete presentation?
- If a finitely presented monoid M has finite derivation type, does M have a solvable word problem?

We believe that both get a negative answer. He also suggests that his notions may be relevant for context-free languages, since context-free grammars are a special kind of rewriting systems.

Finally, he argues that, unlike those of [Sq1], his new results can be adapted to more general algebraic systems, allowing many operations of arbitrary arities. We agree with this: For example, Mac Lane coherence theorem for monoidal categories (see [Ma]) can be proved by means of a complete term rewriting system, using the same argument as for theorem 2. However, we claim that the viewpoint of [Bu] is even more general and allows to maintain a geometrical interpretation. In this way, Mac Lane coherence theorem can be seen as a 3-dimensional word problem involving a 4-dimensional space.

Now, we suggest another kind of generalization. For a monoid M, we have already the following hierarchy of conditions:

- M is finitely generated (dimension 1);
- M is finitely presentable (dimension 2);
- M has finite derivation type (dimension 3).

There should be a finiteness condition in dimension n for each n. This condition would be satisfied if M has a finite complete presentation and would imply that $H_n(M)$ is of finite type. In particular, this would give an alternative proof of the generalization of theorem 4 to all dimensions (see [Gr, Ko]).

References

[Bu] **A. Burroni**. *Higher Dimensional Word Problem.* Theoretical Computer Science **115**, 43–62, 1993.

[CrOt] **R. Cremanns & F. Otto**. *Finite Derivation Type Implies The Homological Finiteness Condition FP3.* Mathematische Schriften Kassel, preprint 12/93.

[Gr] **J. R. J. Groves**. *Rewriting systems and homology of groups*, in L. G. Kovács (ed.), *Groups–Canberra 1989.* Lecture Notes in Mathematics **1456**, 114–141, 1990.

[Hu] **G. Huet**. *Confluent reductions: abstract properties and applications to term rewriting systems.* Journal of the ACM **27**, 797–821, 1980.

[Ko] **Y. Kobayashi**. *Complete rewriting systems and homology of monoid algebras*. Journal of Pure and Applied Algebra **65**, 263–275, 1990.

[LaPr] **Y. Lafont & A. Prouté**. *Church-Rosser property and homology of monoids*. Mathematical Structures in Computer Science **1**, 297–326, 1991.

[Le] **P. Le Chenadec**. *Canonical Forms in Finitely Presented Algebras*. Pitman, London, John Wiley & Sons, New York, Toronto, 1986.

[Ma] **S. Mac Lane**. *Categories for the Working Mathematician*. Graduate Texts in Mathematics **5**, Springer-Verlag, 1971.

[Sq1] **C. C. Squier**. *Word problems and a homological finiteness condition for monoids*. Journal of Pure and Applied Algebra **49**, 201–217, 1987.

[Sq2] **C. C. Squier**. *A Finiteness Condition for Rewriting Systems*. Revision by F. Otto & Y. Kobayashi, to appear in Theoretical Computer Science.

[SqOt] **C. C. Squier & F. Otto**. *The word problem for finitely presented monoids and finite canonical rewriting systems*, in J. P. Jouannaud (ed.), *Rewriting Techniques and Applications*. Lecture Notes in Computer Science **256**, 74–82, 1987.